Lie algebras, their representation theory and $GL_n$
Minor Thesis

Greta Panova
June 20, 2008

Contents
1 Introduction 1
2 Lie Algebras 2
  2.1 Definitions and main theorems .............................................. 2
  2.2 Representations of Lie algebras ............................................ 7
  2.3 Representations of $\mathfrak{sl}(2, F)$ ....................................... 9
3 Lie Groups and their Lie algebras 11
4 Representations of semisimple Lie algebras 15
  4.1 Root space decomposition .................................................. 15
  4.2 Roots and weights; the Weyl group ....................................... 19
5 The special linear Lie algebra $\mathfrak{sl}(n, \mathbb{C})$ and the general linear group $GL_n(\mathbb{C})$ 22
  5.1 Structure of $\mathfrak{sl}(n, \mathbb{C})$ ......................................... 22
  5.2 Representations of $\mathfrak{sl}(n, \mathbb{C})$ ................................... 24
  5.3 Weyl’s construction, tensor products and some combinatorics ........ 25
  5.4 Representations of $GL_n(\mathbb{C}^n)$ ...................................... 27
  5.5 Some explicit construction, generators of $S_\lambda V$ ................. 28
6 Conclusion 30
A Schur functors, Weyl’s construction and representations of $GL_n(\mathbb{C})$ 30

1 Introduction

The goal of this minor thesis is to develop the necessary theory of Lie algebras, Lie groups and their representation theory and explicitly determine the structure and representations of $\mathfrak{sl}_n(\mathbb{C})$ and $GL_n(\mathbb{C})$. My interest in the representations of $GL(V)$ come from their strong connection to combinatorics as developed in Chapter 7 and its appendix in [3]. Even though the irreducible holomorphic representations of $GL(V)$ can be derived without passing through the general theory of Lie algebras, as Schur did in his dissertation, any attempt to generalize the result or to obtain analogous results for the unitary or symplectic groups will pass through the theory of Lie algebras.

Here we will develop the basic theory of Lie algebras and their representations, focusing on semisimple Lie algebras. We will develop most of the necessary theory to show facts like complete irreducibility of representations of semisimple Lie algebras, develop the theory necessary to decompose the lie algebras into root spaces and use these root spaces to decompose representations into weight spaces and list the irreducible
representations via weights. We will establish connections between Lie groups and Lie algebras, which will,
for example, enable us to derive the irreducible representations of $GL(V)$ through the ones for $\mathfrak{gl}(V)$.

In our development of the basic theory of Lie algebras we will follow mostly [2], while studying Lie groups,
roots and weights, $\mathfrak{sl}(n,\mathbb{C})$ we will follow [1]. We will encounter some combinatorial facts which will be taken
for granted and whose proofs are found in [3].

2 Lie Algebras

In this section we will follow [2]. We will develop the basic theory of Lie algebras and later we’ll establish
how they arise from Lie groups and essentially motivate their existence.

2.1 Definitions and main theorems

We will, of course, start with the definition of a Lie algebra.

Definition 1. A Lie algebra is a vector space over a field $F$ endowed with a bracket operation $L \times L \to L$
denoted $(x, y) \to [xy]$, which satisfies the following axioms:

(L1) $[xy]$ is bilinear,

(L2) $[xx] = 0$ for all $x \in L$,

(L3) $[xy]$ satisfies the Jacobi identity $[x[yz]] + [y[zx]] + [z[xy]] = 0$ for all $x, y, z \in L$.

A homomorphism (isomorphism) of Lie algebras will be a vector space homomorphism (resp. isomor-
phism) that respects the bracket operation, a subalgebra would a subspace closed under the bracket operation.

An important example of Lie algebra is the general linear algebra $\mathfrak{gl}(V)$, which coincides as a vector
space with $\text{End} V$ (or $M_n - $ space of $n \times n$ matrices) and has a bracket operation defined as $[XY] = XY - YX$
(it is a straightforward check to verify the axioms). Next we can define the special linear algebra $\mathfrak{sl}(V)$ as
the subspace of $\text{End} V$ of trace 0 endomorphisms and the same bracket operation, it is clearly a subalgebra
of $\mathfrak{gl}(V)$. Similarly we can define other subalgebras of $\mathfrak{gl}(V)$ by imposing conditions on the matrices, e.g.
uppertringular, strictly uppertringular, diagonal.

Another important example of a $\mathfrak{gl}(\mathfrak{U})$ subalgebra, where $\mathfrak{U}$ is a vector space endowed with a bilinear
operation (denoted as multiplication), is the algebra of derivations $\text{Der} \mathfrak{U}$. It is first of all the vector space
of linear maps $\delta : \mathfrak{U} \to \mathfrak{U}$, such that $\delta(XY) = X\delta(Y) + \delta(X)Y$. In order to verify that this is a subalgebra
we need to check that $[\delta\delta'] \in \text{Der} \mathfrak{U}$. For the sake of exercise we will do that now. We need to check that
$[\delta\delta'] = \delta\delta' - \delta'\delta$ satisfies the derivation condition (it is clear it’s a linear map), so for $X, Y \in \mathfrak{U}$ we apply
the derivation rule over and over $^1$ Now if we let $\mathfrak{L} = L$ with the bilinear operation defined by the bracket
of $L$, we can talk about $\text{Der} L$. Some of these derivations arise as endomorphisms $\text{ad} x : L \to L$ given by
$\text{ad} x(y) = [xy]$. It is a derivation, since by the Jacobi identity $\text{ad} x(yz) = [x[yz]] = [y[zx]] + [[xy]z] = [y \text{ad} x(z)] + [\text{ad} y(z)]$. We can now make a

Definition 2. The map $L \to \text{Der} L$ given by $x \to \text{ad} x$, where $\text{ad} x(y) = [xy]$ for all $y \in L$, is called the
adjoint representation of $L$.

$^1$ $[\delta\delta'](XY) = \delta(\delta'(XY)) - \delta'(\delta(XY)) = \delta(\delta'(X)Y + X\delta'(Y)) - \delta'(\delta(X)Y + X\delta(Y)) = \delta(\delta'(X)Y) + \delta(X\delta'(Y)) - \delta'(\delta(X)Y) = \delta\delta'(X)Y + \delta'(X)\delta(Y) + \delta(X)\delta'(Y) + X\delta(\delta'(Y)) - \delta'(\delta(X)Y) - \delta'(X)\delta'(Y) - \delta'(X)\delta(Y) - \delta'(X\delta'(Y)) = X(\delta\delta'(Y) - \delta'(X\delta(Y)) + (\delta\delta'(X) - \delta'(X)\delta(Y)) = X(\delta\delta'(Y) + [\delta\delta'](XY))$
L as $N_L(K) = \{x \in L[|xK| \subset K]\}$ and the centralizer of a subset $X$ of $L$ as $C_L(X) = \{x \in L[|xX| = 0]\}$, both of them turn out to be subalgebras of $L$ by the Jacobi identity.

We need to define what we mean by a representation of a Lie algebra since this is the ultimate goal of this paper. A representation of $L$ is a homomorphism $\phi : L \to \mathfrak{gl}(V)$. The most important example of which is the already mentioned adjoint representation $ad : L \to \mathfrak{gl}(L)$. Since linearity is clear, all we need to check is that it preserves the bracket, which again follows from the Jacobi identity.

We proceed now to some very important concepts in the theory of Lie algebras. We call a Lie algebra $L$ solvable if the sequence of ideals $L^{(0)} = L, L^{(1)} = [LL], \ldots, L^{(n)} = [L^{(n-1)}L^{(n-1)}]$ eventually equals $0$. Note for example that simple algebras are not solvable, as $L/I$ as noted previously, but is the already mentioned adjoint representation $ad$. Theorem 1 in particular this is true if $\mathfrak{gl}(n,F)$ is a nilpotent matrix, $X^n = 0$, then for any $Y \in \mathfrak{gl}(V)$ we have $(adX)^k(Y) = \sum_i a_i X^i Y X^{k-i}$ by induction on $k$ for some $a_i \in F$ and so we have for $k \geq 2n$ that either $i$ or $k-i$ is $\geq n$, so $X^i = 0$ or $X^{k-i} = 0$ and ultimately $(adX)^k(Y) = 0$, so $adX$ is nilpotent. If $x$ is semisimple, then there is a basis of $V$, $v_1, \ldots, v_n$, relative to which $x = diag(a_1, \ldots, a_n)$. Then for the basis elements $e_{ij}$ of $\mathfrak{gl}(V)$ (matrices with only nonzero entry equal to 1 at row $i$, column $j$) we have $ad(x(e_{ij})) = xe_{ij} - e_{ij}x = a_i e_{ij} - a_j e_{ij}$ so $ad x$ is a diagonal matrix with $a_i - a_j$ on the diagonal. Since Jordan decomposition is unique and since the properties of nilpotency and semisimplicity are preserved we have that $ad x = ad x_1 + ad x_2$ is the Jordan decomposition of $ad x$.

We will consider now what it means for a Lie algebra $L$ to be nilpotent and semisimple.

**Definition 3.** A Lie algebra $L$ is called semisimple if its radical is $0$, $\text{Rad} L = 0$.

We now define another very important notion to play a role later. As we know from linear algebra every matrix $X$ can be written uniquely as $X = X_s + X_n$, where $X_n$ is nilpotent, $X_s$ is diagonalizable and $X_s$ and $X_n$ commute. We would like to do something similar with Lie algebras, i.e. find their nilpotent and semisimple elements. If $X \in \mathfrak{gl}(V)$ is a nilpotent matrix, $X^n = 0$, then for any $Y \in \mathfrak{gl}(V)$ we have $(adX)^k(Y) = \sum_i a_i X^i Y X^{k-i}$ by induction on $k$ for some $a_i \in F$ and so we have for $k \geq 2n$ that either $i$ or $k-i$ is $\geq n$, so $X^i = 0$ or $X^{k-i} = 0$ and ultimately $(adX)^k(Y) = 0$, so $adX$ is nilpotent. If $x$ is semisimple, then there is a basis of $V$, $v_1, \ldots, v_n$, relative to which $x = diag(a_1, \ldots, a_n)$. Then for the basis elements $e_{ij}$ of $\mathfrak{gl}(V)$ (matrices with only nonzero entry equal to 1 at row $i$, column $j$) we have $ad(x(e_{ij})) = xe_{ij} - e_{ij}x = a_i e_{ij} - a_j e_{ij}$ so $ad x$ is a diagonal matrix with $a_i - a_j$ on the diagonal. Since Jordan decomposition is unique and since the properties of nilpotency and semisimplicity are preserved we have that $ad x = ad x_1 + ad x_2$ is the Jordan decomposition of $ad x$.

We will consider now what it means for a Lie algebra $L$ to be nilpotent and semisimple.

**Definition 4.** A Lie algebra $L$ is called nilpotent if the series $L^0 = L, L^1 = [LL], \ldots, L^i = [LL^{i-1}], \ldots$ eventually becomes $0$.

For example any nilpotent algebra is also solvable since $L^{(i)} \subset L^i$. If $L^n = 0$, this means that for any $x_0, \ldots, x_n \in L$ we have $[x_n][x_{n-1}] \ldots [x_1x_0] = 0$ or in operator notation $ad x_n ad x_{n-1} \ldots ad x_1(x_0) = 0$, in particular this is true if $x_n = \cdots = x_1 = x$, so $(ad x)^n = 0$ for all $x \in L$. It turns out that the converse is also true, we have

**Theorem 1** (Engel’s theorem). If all elements of $L$ are ad-nilpotent (i.e. $(ad x)^k = 0$) then $L$ is nilpotent.

\footnote{Linearity is clear, we need only check that they are closed under the bracket operation. If $x, y \in N_L(K)$ then $0 = [[xy]k] + [[yk]x] + [[xk]y]$ and since $[yk] \in K, [xk] = -[xk] \in K$ we have $[x[yk]] = -[[yk]x] \in K$ and $[y[kz]] \in K$, so $[[xy]k] \in K$. The case for $C_L$ follows from the second two terms being 0.}

\footnote{For $x, y, z \in L$, we have $[ad x ad y(z) = ad(x(ad(y(z))) - ad y(ad(x(z))) = [x[yz]] - [y[xz]] = [[xy]z] = ad(xy)(z)$.}
For example then by this theorem then $n(n, F)$ is a nilpotent Lie algebra. We will prove Engel’s theorem, because it will illustrate some important techniques.

The proof requires a lemma, important on its own, we will prove it first.

**Lemma 1.** If $L$ is a subalgebra of $\mathfrak{gl}(V)$ ($\dim V < \infty$) consisting of nilpotent endomorphisms, then $L.v = 0$ for some nonzero $v \in V$.

**Proof of lemma.** In order to prove this lemma we will proceed by induction on $\dim L$ (this is the first common trick in Lie algebras). The goal is to find a codimension 1 subalgebra of $L$ and use the induction hypothesis on it. The dimension 0 and 1 cases are trivial. Now let $K$ be a maximal proper subalgebra of $L$, it acts on $L$ via the adjoint operator and hence on $L/K$. Since the elements of $K$ are nilpotent, so are their ad-s by the preceding paragraph. We can use the induction hypothesis on $K \subset \mathfrak{gl}(L/K)$, as clearly $\dim K < \dim L$, so there is a vector $x + K \in L/K$, such that $x \not\in K$ and such that $[yx] = 0$ for all $y \in K$, in particular $x \in N_L(K) \setminus K$. Since $K \subset N_L(K)$, $N_L(K)$ is a subalgebra of $L$ as we showed and $K$ is maximal proper subalgebra, we should have $N_L(K) = L$ as a strictly larger subalgebra. But that means that for all $x \in L$ \([xK] \subset K\), i.e. $K$ is an ideal. If $K$ is an ideal, then the vector space spanned by $K$ and $z \in L$ is a subalgebra for any $z \in L$ (since $[zK] \subset K \subset \text{span}(K, z)$). So in particular if $\dim L/K > 1$, then the space spanned by $K$ and $z \in L \setminus K$ is a proper subalgebra violating the maximality of $K$. Therefore $\dim L/K = 1$ and so $L = \text{span}(K, z)$. Now we can apply the induction hypothesis to $K$, so there is a nonzero vector space $W = \{w | Kw = 0\}$. The element $z$ is nilpotent, so is then its restriction to $W$ $(z^k = 0$ then $(z|w)^k = 0)$, so it has an eigenvector in $W$ corresponding to the 0 eigenvalue, $zv = 0$, then $Lv = \text{span}\{Kv, zv\} = 0$ and we are done.

This lemma in particular implies that we can find a basis of $V$, for which the matrices of $L$ are strictly upper triangular.

**Proof of Engel’s theorem.** The condition that the elements of $L$ are ad-nilpotent, suggests that we look at the algebra $\text{ad}L$ first, it acts on the vector space $\mathfrak{gl}(L)$ and its elements are nilpotent, so we can use the lemma and find an $x \in L$, such that $[Lx] = 0$. The existence of this $x$ enables us to find a proper subalgebra of $L$ of smaller dimension and use again induction on $\dim L$. In particular, since $x \in Z(L)$ (centralizer of $L$) then $Z(L)$ is a nonempty subalgebra of $L$ as we showed earlier. It is also an ideal of $L$ \(^4\), so we can form the quotient algebra $L/Z(L)$, which will then have a smaller dimension than $L$, and will still consist of ad-nilpotent elements. By induction it is nilpotent, so there is, such that $(L/Z(L))^n = 0$, i.e. $L^n \subset Z(L)$, but then $L^{n+1} = [LL] \subset [LZ(L)] = 0$ and so $L$ is nilpotent.

Engel’s theorem essentially implies existence of a common eigenvector for 0, we would like to extend this result for any eigenvalue. Here we should assume that the underlying field $F$ is algebraically closed and of characteristic 0.

**Theorem 2** (Lie’s theorem.). If $L$ is a solvable subalgebra of $\mathfrak{gl}(V)$, $V$ of finite dimension, then $L$ stabilizes a flag in $V$, i.e. there is a basis for $V$ with respect to which the matrices in $L$ are upper triangular.

**Proof.** We can proceed by induction on the $\dim V$, we just need to show that there is a common eigenvector $v$ for all elements in $L$ and then apply the induction to $V/v$. In order to prove the existence of such vector, we again need to find an ideal of codimension one. Since $L$ is solvable we know that $[LL] \neq L$, otherwise $L^{(n)} = L$, it is clearly an ideal and any subspace $K'$ of the quotient $L/[LL]$ is an ideal of $L/[LL]$, since $[Kx] \subset [LL]$ goes to 0. So we take a codimension one subspace of $L/[LL]$, its inverse image by the quotient map is then a codimension one ideal in $L$, call it $K$, $L = K + Fz$. As $K^{(n)} \subset L^{(n)}$, $K$ is solvable, so we apply the induction hypothesis and find an eigenvector $v$ for $K$, i.e. for any $x \in K$, there is a $\lambda(x) \in F$, such that $xv = \lambda(x)v$. Here is the place to note that the map $\lambda : K \rightarrow F$ is actually linear \(^5\), it will play important roles further in our discussion. Now instead of fixing $v$, we can fix $\lambda$ and consider the possibly larger subspace $W = \{w \in V | xw = \lambda(x)w \text{ for all } x \in K\}$.

\(^4\)If $y \in Z(L)$ and $z \in L$, then for any $u \in L$ $[u|yz] = -[z|yu] - [y|uz] = 0$

\(^5\)Since $(ax + by)v = a(xv) + b(yv) = a\lambda(x)v + b\lambda(y)v$. 

4
Now the goal is to show that \( z(W) \subset W \), so that we can find an eigenvector for \( z|_W \), which will then be an eigenvector for \( L \).

We have to introduce now another important trick, that is the **fundamental equality** \( xz.w = zx.w + [x,z]w \). Now if \( x \in K \), then we would have \( x(z.w) = \lambda(x)(z.w) + \lambda([z,x])w \) and if \( \lambda([z,x]) = 0 \), then we would have \( w \in W \), which is what we need. Consider the spaces \( W_i = \{ w, zw, z^2w, \ldots, z^iw \} \), we have that \( zW_i \subset W_{i+1} \). Employing our fundamental trick again and using the fact that \( K \) is an ideal we have that for any \( x \in K \), \( x(zw) = z(xzw) + [x,z](z^iw) \), so by induction if \( x \in W_{i-1} \) then \( xW_i \subset W_i \) and \( K W_i \subset W_i \), for \( i \). Moreover we get from the same induction that \( xz^i.w - \lambda(x)z^iw = [x,z]z^{i-1}w \). Since we have \( [x,z]w \in W_{i-1} \) then \( xz^i.w \). Moreover, we get from the same induction that \( xz^i.w - \lambda(x)z^iw = [x,z]z^{i-1}w \). Since \( \lambda(\ldots) = 0 \), then \( xz^i.w \).

As a corollary of this, if we consider the adjoint representation of \( L \) in \( \mathfrak{gl}(L) \) we get that \( L \) fixes a flag of subspaces \( L_0 \subset L_1 \subset \cdots \subset L_n = L \), which are necessarily ideals. If we choose a basis for this flag, then with respect to it \( \text{ad} L \) consists of upper triangular matrices and hence \( \text{ad} L, \text{ad} L LL \) are strictly upper triangular and hence nilpotent, i.e. \( \text{ad} x \) is nilpotent for \( x \in [L,L] \) and so by Engel’s theorem \([L,L]\) is nilpotent.

As we know from linear algebra every matrix \( X \) can be written uniquely as \( X \equiv X_s + X_n \), where \( X_n \) is nilpotent, \( X_s \) is diagonalizable and \( X_s \) and \( X_n \) commute. We would like to do something similar with Lie algebras, i.e. find their nilpotent and semisimple elements. By a theorem from linear algebra, we know that for any matrix, its Jordan decomposition into the sum of commuting semisimple (diagonalizable) and nilpotent parts, can be achieved concretely by giving this parts as polynomials in our original matrix, namely given \( X \) there exist polynomials \( p \) and \( q \) with 0 constant terms, such that \( X = p(X) \) is the semisimple part of \( X \) and \( X = q(X) \) - the nilpotent. This fact enables us to give an easy checkable criterion for solvability.

**Theorem 3** (Cartan’s criterion). Let \( L \) be a subalgebra of \( \mathfrak{gl}(V) \), \( V \)-finite dimensional. If \( Tr(xy) = 0 \) for all \( x \in [L,L] \) and \( y \in L \), then \( L \) is solvable.

**Proof.** If we show that \([L,L]\) is nilpotent, then \([L,L](\pi) \subset [L,L](\pi-1)\) would be solvable, so by Engel’s theorem we need to show that the elements of \([L,L]\) are nilpotent, so we need to figure out a trace criterion for nilpotency.

Suppose first that \( A \subset B \subset \mathfrak{gl}(V) \) are subspaces and \( M = \{ x \in \mathfrak{gl}(V) | [x,B] \supset A \} \). Suppose that there is an \( x \) such that \( Tr(xy) = 0 \) for all \( y \in M \), we will show now \( x \) is nilpotent.

We can find a basis, relative to which the semisimple part of \( x \), \( x_s \), is diagonal, so fix this basis and write \( x = diag(a_1, \ldots, a_n) + x_n \). We want to show that \( a_{i+1} \), consider the space \( E = \mathbb{Q}\{a_1, \ldots, a_n\} \), we will show that \( E = 0 \) by showing that \( E^* = \{ f : E \to \mathbb{Q} | f \text{ linear} \} = 0 \).

For any \( f \in E^* \), let \( y = diag(f(a_1), \ldots, f(a_n)) \), then \( e(x) \) is the corresponding basis of \( \mathfrak{gl}(V) \) of matrices with only nonzero entry, equal to 1, at \( (i,j) \), then \( ad_{x_s} e_{ij} = x_s e_{ij} - e_{ij} x_s = (a_i - a_j) e_{ij} \) and \( ad_{y} e_{ij} = (f(a_i) - f(a_j)) e_{ij} \). By Lagrange interpolation we can find a polynomial \( r(T) \) in \( F[T] \) with 0 constant term, such that \( r(a_i - a_j) = f(a_i) - f(a_j) \) for all \( i, j \). We then have that \( r(ad_{x_s}) e_{ij} = r(a_i - a_j) e_{ij} = (f(a_i) - f(a_j)) e_{ij} \) and \( ad_{y} e_{ij} \) for all \( i, j \), and since \( e_{ij} \) are a basis for \( \mathfrak{gl}(V) \), we must have that \( ad y = r(ad_{x_s}) \).

Thus \( x_s \) is the semisimple part of \( x \) as we remarked earlier, it will be a polynomial in \( ad x \) without constant term and so \( ad y = r(ad_{x_s}) \) will be a polynomial in \( ad x \) without constant term. Since by hypothesis \( ad x(B) = [x,B] \subset A \), we will have that \( ad y \) as a polynomial in \( ad x \) will also map \( B \) into \( A \), so \( y \in M \). Hence we have that \( Tr(xy) = 0 \), but \( Tr(xy) = \sum_i a_i f(a_i) = 0 \). We can apply \( f \) to this linear combination.

---

6One can do this explicitly: if \( \prod (T-a_i)^{m_i} \) is the characteristic polynomial of \( X \), then by the Chinese remainder theorem there is a polynomial \( p(T) \equiv a_i (mod (T-a_i)^{m_i}) \) and \( p(T) \equiv 0 (mod T) \) and so on each eigensapce \( V_i \) for \( X \) we have \( p(X) - a_i I = 0 \), so \( p(X) \) acts diagonally and clearly fixes the eigenspaces as a polynomial in \( X \), so it is the semisimple part, \( X_n = X - q(X) \) would be the nilpotent.

7So in particular ad \( x_s \) is semisimple and since ad \( x_n \) will still be nilpotent as we showed earlier, ad \( x_s \) is the semisimple part of ad \( x \).
of $a_i$s and use the fact that $f(a_i) \in \mathbb{Q}$ to get that $\sum_i f(a_i)^2 = 0$. Since the $f(a_i)$s are rationals, they must then be 0, so $f = 0$ and $E = 0$, so $a_0 = 0$ and so $x_0 = 0$ and so $x = x_n$ is nilpotent.

Having shown that, we can apply it to $A = [LL], B = L$, so $M = \{x \in \mathfrak{gl}(V) | x[L] \subset [LL]\}$. We then have $L \subset M$, but the hypothesis of the theorem is only that $Tr(xy) = 0$ for $y \in L$, not $y \in M$. We will show now that $Tr(xy) = 0$ for $x \in [LL], y \in L$ implies $Tr(xy) = 0$ for $x \in [LL]$ and $y \in M$ so using the above paragraph we will have $x$ nilpotent. Let $y \in M$ and let $x = [a, b] \in [LL]$ for $a, b \in L$. Then $Tr([a, b]y) = Tr(aby - bab) = Tr(aby) - Tr(bay) = Tr(bay) - Tr(ba) = 0$ since $Tr(AB) = Tr(BA)$ for any two matrices. But then $b \in [LL]$ and $[y, a] \in L$ by definition of $M$, so by the hypothesis of the theorem $Tr(xy) = Tr(b[y, a]) = 0$ and we can apply the preceding paragraph to get $x$ nilpotent.

As we already mentioned, a semisimple algebra is the one for which $\text{Rad} L = 0$, but this is a difficult condition to check. We need other criteria and for that we will introduce the Killing form.

**Definition 5.** Let $L$ be a Lie algebra. If $x, y \in L$, we define $\kappa(x, y) = Tr(ad x, ad y)$. Then $\kappa$ is a symmetric bilinear form on $L$, it is called the Killing form.

We can now state the criterion for semisimplicity as

**Theorem 4.** A Lie algebra $L$ is semisimple if and only if its Killing form is nondegenerate.

**Proof.** The proof naturally involves Cartan’s criterion we just showed. For any Lie algebra then, its adjoint representation is a subalgebra of $\mathfrak{gl}(L)$ and if $Tr(ad x ad y) = 0$ $\text{ad} L$ will be solvable and since $\text{ad} L \sim L/Z(L)$ and $Z(L)$ is solvable we have $L$ is solvable. So if we let $S = \{x \in L | Tr(x, y) = 0$ for all $y \in L\}$ we have $S$ is solvable and so $S \subset \text{Rad} L = 0$, so $\kappa$ is nondegenerate.

If now $\kappa$ is nondegenerate we have $S = 0$. Since $\text{Rad} L$ is the union of all abelian ideals, we can show that for every abelian ideal $I, I \subset S$. If $x \in I$ and $y \in L$, then for $z \in L$ we have $(ad x ad y)(z) = [x[y]] \in I$ and then $(ad y ([x[yz]])) \in I$ and so $(ad x ad y)^2(z) \in [I^2] = 0$. So $\text{ad} x ad y$ is nilpotent and hence has trace 0 for all $y \in L$ which makes $x \in S$, so $\text{Rad} L \subset S = 0$ and $L$ is semisimple.

We say that a Lie algebra is a direct sum of ideals if there are ideals $I_1, \ldots, I_k$, such that $L = I_1 + \cdots + I_k$ as a vector space. Semisimple algebras have the nice property to be such direct sums, in other words

**Theorem 5.** Let $L$ be semisimple. Then there exist simple ideals (as Lie algebras) $L_1, \ldots, L_k$ of $L$, such that $L = L_1 \oplus \cdots \oplus L_k$ and this decomposition is unique in the sense that any simple ideal is one of the $L_i$s.

**Proof.** The Killing form has a nice property of associativity, i.e. $\kappa([xy], z) = \kappa(x, [yz])$ and so if we take the orthogonal subspace to an ideal $I$, i.e. $I^\perp = \{x \in L | \kappa(x, y) = 0$ for all $y \in I\}$, then $I^\perp$ will also be an ideal( for $z \in L, x \in I^\perp, y \in I$ we have $\kappa([xz], y) = \kappa(x, [yz]) = 0$ since $[yz] \in I$). Moreover $I \cap I^\perp$ is 0 since it is solvable by Cartan’s criterion and $L$ is semisimple. We must then have $L = I \oplus I^\perp$.

Now let $L_1$ be a minimal nonzero ideal of $L$, we have $L = L_1 \oplus L_1^\perp$. If $I \subset L_1$ is an ideal of $L_1$, then $[L_1] \subset L_1$ is an ideal of $L_1$ by the following reason. Let $x \in I$ and any element of $L$ can be expressed as $y + z$, where $y \in L_1$ and $z \in L_1^\perp$, then $[x(y + z)] = [xy] + [xz]$. We already have $[xy] \in I$, we will now use the fact that $L$ semisimple implies that the Killing form is nondegenerate and show that $[xz] = 0$ for any $x \in L_1$ and $z \in L_1^\perp$. Let $u \in L$, then by associativity $\kappa([xz], u) = \kappa(z, [xu]) = 0$ since $[xu] \in L_1$. As this holds for all $u \in L$ we need to have $[xz] = 0$ for $x$ to be nondegenerate. So $[x(y + z)] = [zy] \in I$ and so $I$ is an ideal of $L$. Then by the choice of minimality of $L_1$ it follows that $I = 0$ or $L_1$, so $L_1$ has no nontrivial ideals and so is simple (or abelian, which is not the case as $\text{Rad} L = 0$). Similarly any ideal of $L_1^\perp$ is an ideal of $L$ and so $L_1^\perp$ cannot have any solvable ideals, so it is semisimple. Then by we reduce to the case of $L_1^\perp$ and by induction it is a direct sum.

In order to show uniqueness suppose $I$ is a simple ideal of $L$, then $[IL]$ is an ideal of $I$ and since $L$ is semisimple it cannot be 0, so $[IL] = I$. But then $L = L_1 \oplus \cdots \oplus L_k$, so $I = [IL] = [IL_1] \oplus \cdots \oplus [IL_k]$ (note that $[IL_i] \subset L_i \cap L_j = 0$ and therefore all but one summand must be 0, $I = [IL_i] \subset L_i$ and so $I$ is an ideal of $L_i$, so it must be $I = L_i$ as $L_i$ is simple.

\[\square\]
Since every simple Lie algebra $L$ is isomorphic to the image of its adjoint representation (the kernel being the center, i.e. 0), then it is isomorphic to a subalgebra of $\mathfrak{gl}(L)$. If $L$ is semisimple it is the direct sum of simple $L_1 \oplus \ldots \oplus L_k$, so $L$ is isomorphic to a subalgebra of $\mathfrak{gl}(L_1 \oplus \ldots \oplus L_k)$, i.e. we have the

**Corollary 1.** Every semisimple Lie algebra is isomorphic to a subalgebra of $\mathfrak{gl}(V)$ for some $V$.

### 2.2 Representations of Lie algebras

Here we will develop some of the basic Representation Theory of semisimple Lie algebras. We start of with a

**Definition 6.** A $\mathbb{F}$-vector space $V$ endowed with an operation $L \times V \to V$ ( $(x,v) \to x.v$) is called an $L$-module if

1. $(ax + by).v = a(x.v) + b(y.v)$,
2. $x.(av + bw) = a(x.v) + b(x.w)$,
3. $[xy]v = x.(y.v) - y.(x.v)$,

for all $x, y \in L$, $a, b, \in F$ and $v, w \in V$.

Homomorphisms are defined as usual, an $L$-module is irreducible if it has no nontrivial $L$-submodules and is completely reducible if it is a direct sum of irreducible $L$-submodules. Given an $L$-module $V$, its dual $V^*$ is also an $L$-module as follows: if $f \in V^*$, $x \in L$ and $v \in V$, then $(x.f)(v) = -f(x.v)$. The reason for the minus sign is to satisfy the third axiom in the definition, since $(\mathord{[xy].f})(v) - f((xy).v) = -f(x(y.v) + f(y.(x.v)) - (x.f)(y.v) - (y.f)(x.v) = -(x.(y.f))(v) + (x.(y.f))(v) = ((xy - yx).f)(v)$. Another peculiarity comes with the action of $L$ on a tensor product $V \otimes W$, where we define $x(v \otimes w) = x.v \otimes w + v \otimes x.w$, one checks again that it respects the third axiom.

No discussion of representation theory could be complete without mentioning

**Lemma 2** (Schur’s lemma). If $\phi : L \to \mathfrak{gl}(V)$ is irreducible, then the only endomorphisms of $V$ commuting with all $\phi(x)$ are the scalars.

We will now exhibit a commuting endomorphism and apply Schur’s lemma. We can extend our discussion about the Killing form by defining a symmetric bilinear form on a semisimple $L$ for any faithful representation $\phi : L \to \mathfrak{gl}(V)$ as $\beta(x,y) = Tr(\phi(x)\phi(y))$, which follows the properties of the Killing form, namely associativity and nondegeneracy. Now if $(x_1, \ldots, x_n)$ is a basis for $L$, let $(y_1, \ldots, y_n)$ be its dual basis with respect to $\beta$, i.e. $\beta(x_i, y_j) = \delta_{ij}$, define the Casimir element of $\phi$ as $c_\phi(\beta) = \sum_{i=1}^n \phi(x_i)\phi(y_i)$. Since $\phi(x_i), \phi(y_i) \in \text{End}(V)$, the so is $c_\phi$, our goal now is to show that it commutes with every $\phi(x)$.

We need to use the associativity of $\beta$. If $z \in L$, we should consider its bracket rather than product with the basis since the bracket is preserved under $\phi$. We can write $[zx_j] = \sum_j a_{ij}x_j$ and $[zy_j] = \sum_j b_{ij}y_j$ for some coefficients $a_{ij}, b_{ij}$ and use the orthogonality with respect to $\beta$, then $\beta([zx_j], y_j) = \beta([-x_i], y_j) = -\beta(x_i, [zy_j]) = -\sum b_{ij}\beta(x_i, y_j) = -b_{ij}$, and $\beta([zx_j], y_j) = \sum a_{ij}\beta(x_i, y_j) = a_{ij}$, so $a_{ij} = -b_{ij}$. Now we can consider

$$[\phi(z),c_\phi(\beta)] = \sum_j [\phi(z), \phi(x_j)\phi(y_j)] = \sum_j \phi(z)\phi(x_j)\phi(y_j) - \phi(x_j)\phi(y_j)\phi(z) =$$

$$\sum_j \phi(z)\phi(x_j)\phi(y_j) - \phi(x_j)\phi(z)\phi(y_j) + \phi(x_j)\phi(z)\phi(y_j) - \phi(x_j)\phi(y_j)\phi(z) =$$

$$\sum_j [\phi(z)\phi(x_j)]\phi(y_j) + \phi(x_j)[\phi(z)\phi(y_j)] = \sum_j \phi([zx_j])\phi(y_j) + \phi(x_j)\phi([zy_j]) =$$

$$\sum_j \sum_k a_{jk}\phi(x_k)\phi(y_j) - \sum_j \sum_l a_{lj}\phi(x_j)\phi(y_l) = 0.$$
So $c_\phi(\beta)$ commutes with any $\phi(z)$ and if $\phi$ is irreducible then by Schur’s lemma we should have $c_\phi(\beta)$ be scalar multiplication, in particular it is $1/(\dim V)\text{Tr}(c_\phi) = 1/\dim V \sum_y \text{Tr}(\phi(x_i)\phi(y_j)) = 1/\dim V \sum_{i=1}^L \beta(x_i, y_j) = \dim L/\dim V$ and so it does not depend on the bases we choose.

Casimir’s element is important in the sense that it shows there is an endomorphism commuting with the action of $L$ and with nonzero trace (equal to $\dim L$); it will be used in the proof of our main theorem as a projection to an irreducible module.

We now proceed to prove the main theorem of this section, namely

**Theorem 6** (Theorem (Weyl)). Let $L$ be semisimple and $\phi : L \to \mathfrak{gl}(V)$ be a finite dimensional representation of $L$. Then $\phi$ is completely reducible.

**Proof.** First of all, since $L$ is semisimple we can write it as a direct sum of simple ideals $L = I_1 \oplus \ldots I_k$ and then $[LL] = [I_1 L] \oplus \ldots [I_k L]$. Since $I_j$ is simple we have $I_j \supset [I_j L] \supset [I_j, I_j] = I_j$, so $[LL] = I_1 \oplus \ldots I_k = L$.

As a consequence of this fact we have that $L$ acts trivially on any one-dimensional space since there $\phi(xy) = \phi(yx) = \phi(x)\phi(y)$ and so $\phi([x, y]) = [\phi(x), \phi(y)] = 0$.

This nice fact suggests that we look first at the case where $W$ is a codimension $1$ $L$–submodule of $V$.

Since as we observed $L$ acts trivially on $V/W$, we have that $\phi(L)(V) \subset W$. Since $c_\phi$ is an $L$–module endomorphism and it is sum of products of elements in $\phi(L)$, then $c_\phi(V) \subset \phi(L)(V) \subset W$. Now we have just exhibited a projection operator as the averaging operators in the representation theory of finite groups.

The goal is to show that $V = W \oplus \ker c_\phi$. Consider $\ker c_\phi \cap W$ - an $L$–submodule of $W$ so it must be either $0$ or $W$. If it is $0$ we are done as then $V = W \oplus \ker c_\phi$. If it is $W$, then we have $c_\phi^2(V) \subset c_\phi(W) = 0$ so $c_\phi$ is nilpotent and in particular $\text{Tr}(c_\phi) = 0$, but we already showed that this trace is $\dim L$ and since $\text{char} F = 0$ we reach a contradiction.

We required that $W$ be irreducible, but we can reduce it to this case easily by induction on $\dim W$. Again, $\dim V/W = 1$. Since again $L(V/W) = 0$ we can write an exact sequence of $L$–module homomorphisms $0 \to W \to V \to F \to 0$. If $W'$ is a proper module of $W$, we can quotient by it and get $0 \to W/W' \to V/W' \to F \to 0$ and since $\dim W/W' < \dim W$ by induction this sequence splits, i.e. $V/W' = W/W' \oplus \tilde{W}/W'$, where $\dim \tilde{W}/W' = 1$. So now looking at $\tilde{W}$ and $W'$ we have again an exact sequence $0 \to W' \to \tilde{W} \to F \to 0$ and again by induction we have $\tilde{W} = W' \oplus W''$. Finally $V/W' = W/W' \oplus \tilde{W}/W' = W/W' \oplus W''/W'$. Since $W \cap \tilde{W} \subset W' \subset W$ we have $W \cap W'' = 0$, so $V = W \oplus W''$.

Now consider the general case where $W$ is a submodule of $V$. The space of linear maps $\text{Hom}(V, W)$ can be viewed as an $L$–module since $\text{Hom}(V, W) = V^* \otimes W$ by $(f \otimes w)v = f(v)w$ for $v \in V$ and $w \in W$. The action of $L$ on $\text{Hom}(V, W)$ can be described using the tensor product rule, namely if $x \in L$, then $(x(f \otimes w))(v) = (x.f \otimes w)(v) + (f \otimes x.w)(v) = x.f(v)w + f(v)x.w = -f(x.v)w + x.f(v).w$ and if $g(v) = (f \otimes w)(v)$ is the actual homomorphism, then $x(g)(v) = -g(x.v) + x.g(v)$. Now let $\mathcal{V}$ be the subspace of $\text{Hom}(V, W)$ of maps whose restriction to $W$ is scalar multiplication. Then this space is also an $L$–submodule of $\text{Hom}(V, W)$: if $f \in \text{Hom}(V, W)$ and $f|W = a 1|W$, then $x.f(v) = -f(x.v) + x.f(w) = -ax.v + x.aw = 0$, in particular $x.f \in \mathcal{V}$. In fact if $\mathcal{W} \subset \mathcal{V}$ is the subspace consisting of $f|W = 0$, then it is again an $L$–submodule and $L(\mathcal{V}) \subset \mathcal{W}$. Moreover $\dim \mathcal{V}/\mathcal{W} = 1$ since $\mathcal{V}/\mathcal{W}$ is determined by that scalar $a$. So we end up with a familiar situation, namely $\mathcal{V}$ has a codimension one submodule $\mathcal{W}$ and by the special case we get $\mathcal{V} = \mathcal{W} \oplus \{f\}$, where $f \in \mathcal{V}$ spans the complementary to $\mathcal{W}$ one-dimensional module, we can assume that $f|W = 1_W$. As we showed earlier for the action of $L$ on $\mathcal{V}$ we have $0 = (x.f)(v) = -f(x.v) + x(f(v))$, i.e. $x.f = f.x$ on $\mathcal{V}$, so $f$ is an $L$–homomorphism and since clearly $W \cap \ker f = 0$ and $f$ is a projection of $V$ into $W$ we have $V = W \oplus \ker f$ as $L$–submodules.

As a result using this theorem we can state another important theorem which will later help us analyze the action of $L$ on a vector space.

**Theorem 7.** Let $L \subset \mathfrak{gl}(V)$ be a semisimple Lie algebra, then $L$ contains also the nilpotent and semisimple parts of its elements.

In particular, in any representation $\phi$ the diagonalizable elements of $L$ will be diagonalizable in $\phi(L)$. This will help us later present any module $V$ as a direct sum of eigenspaces to this elements.
Proof. We recall a result from linear algebra that says that the semisimple and nilpotent parts, \( x_s \) and \( x_n \) respectively, of \( x \) can be expressed as polynomials in \( x \) with 0 constant coefficient. This in particular shows that if \( xA \subset B \) for any subspaces \( B \subset A \), then \( x_A, x_s A \subset B \) also. In order to use this fact we will attempt to create subspaces \( A \) and \( B \), such that \( xA \subset B \) if and only if \( x \in L \).

For any \( W \) \( L \)-submodule of \( V \), let \( L_W = \{ y \in \mathfrak{gl}(V) | y(W) \subset W, Tr(y|_W) = 0 \} \). Since we showed as a consequence to the direct sum decomposition of \( L \) that \( [LL] = L \), we have that every element is of the form \([x, y]\) for some \( x, y \in L \) and has trace 0. Moreover we have that \( x_s \) and \( x_n \) are also in \( L_W \): since \( x(W) \subset W \) we have \( x_s(W) \subset W \) and \( x_n(W) \subset W \) and also \( Tr(x_n|_W) = 0 \) since \( x_n \) is nilpotent and hence \( Tr(x_s) = Tr(x) = 0 \) on \( W \) also. So \( L, x_s, x_n \subset L_W \) for all \( L \)-submodules \( W \). Let \( N \subset \mathfrak{gl}(V) \) be the space of \( z \in \mathfrak{gl}(V) \), such that \([xL] \subset L \). We also have that \( x_s, x_n \in N \): since we showed that \( \text{ad} x_n \) and \( \text{ad} x_s \) are the nilpotent and semisimple parts of \( \text{ad} x : L \to L \), we have that \( \text{ad} x_n(L) \subset L \) and \( \text{ad} x_s(L) \subset L \) also, i.e. \([x_n, L] \subset L \) and \([x_s, L] \subset L \). So if we show that the intersection \( L' = N \cap (\cap LW) \subset L \) we will be done.

Observe that \( L' \) is an \( L \)-module under the bracket action since for every \( x \in L \) and \( y \in L_W \) we have \([x, y](W) \subset x(y(W)) + y(x(W)) \subset W \) as \( y(W), x(W) \subset W \) and then since it is a commutator its trace is 0; \( N \) is clearly an \( L \) module as \([LN] \subset L \subset N \). Now comes the big moment of applying Weyl's theorem. Since \( L' \) is an \( L \)-module and it contains \( L \) we have \( L' = L + M \) where \( M \) is an \( L \)-submodule of \( L' \) and \( L \cap M = 0 \). But since \( L' \subset N \), then \([L'] \subset [NL] \subset L \), so \( L + [LM] = [LL] + [LM] = [L'] \subset L \), so \([LM] \subset L \). But as an \( L \)-module under \([ \] \) we have \([LM] \subset M \), so \([LM] = 0 \) and in particular the elements of \( M \) are endomorphisms commuting with \( L \). So if \( y \in M \) and \( W \) is an irreducible \( L \)-submodule of \( V \), then since \( M \subset LW \) we have \( y(W) \subset W \) and \( Tr(y|_W) = 0 \). By Schur's lemma \( y \) should act as a scalar on \( W \) and since its trace is 0 it should be 0. Since this is true for any irreducible \( W \) and we can write \( V \) as a direct sum of such \( W \)'s by Weyl's theorem (again!), then \( y \) acts as 0 on all of them, hence on \( V \) and so \( y = 0 \). Therefore \( M = 0 \), so \( L' = L \) and \( x_s, x_n \in L \).

2.3 Representations of \( \mathfrak{sl}(2, F) \)

The special linear Lie algebra of dimension 2 is not just an easy example we start with. As we have already noticed dimension induction has been a key method in proving facts about Lie algebras, so it won’t be a surprise to see that the general case of \( \mathfrak{sl}(n, F) \) will involve the special case of \( \mathfrak{sl}(2, F) \).

Definition 7. The Lie algebra \( \mathfrak{sl}(n, F) \) is defined as the vector space of \( n \times n \) matrices over \( F \) of trace 0 with a bracket operator given by \([X, Y] = XY - YX\).

In particular we see that \( \dim \mathfrak{sl}(n, F) = n^2 - 1 \) and of course that \( \mathfrak{sl}(n, F) \) is well defined as a Lie algebra, since \( Tr(XY) = Tr(YX) \) for any \( n \times n \) matrices \( X, Y \). We will go back to the general case later, but for now let \( n = 2 \). Clearly \( \mathfrak{sl}(2, F) \) can be generated as a vector space by

\[
x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\]

We can determine \( \mathfrak{sl}(2, F) \) completely by listing the bracket relations among these generators. We have that

\[
[x, y] = h, \quad [h, x] = 2x, \quad [y, h] = 2y,
\]

so in particular any of \( x, y, h \) would generate \( \mathfrak{sl}(2, F) \) as a Lie algebra.

It would have been pointless to discuss semisimple Lie algebras if our main example wasn’t such. So we will show briefly that \( \mathfrak{sl}(2, F) \) is not only semisimple, but in fact simple. For any element \( z = ax + by + ch \in \mathfrak{sl}(2, F) \) with \( a, b, c \in F \) we have that \([z, x] = a0 + b[y, x] + c[h, x] = -bh + 2cx \) and then \([h, z, x] = -6.0 + 2c[h, x] = 4cx\). Similarly \([h, z, y] = h, (ah - 2cy)] = -4cy \) and \([x, [z, h]] = [x, (-2ax + 2by)] = 2bh \) and \([y, [z, h]] = 2ah \). So we see that if \( \text{char} F = 0 \) and at least one of \( a, b, c \) is not 0, i.e. \( z \neq 0 \), we can get one of \( h \) or \( x \) (or \( y \)) by applying the bracket to \( z \) and from them we can get any element, so the ideal generated by a nonzero \( z \) is just \( \mathfrak{sl}(n, F) \) and so \( \mathfrak{sl}(n, F) \) is simple.

Now that we know the structure of \( \mathfrak{sl}(2, F) \) so explicitly we can consider its irreducible representations. Let \( V \) be any \( \mathfrak{sl}(2, F) \) module, \( \phi : \mathfrak{sl}(2, F) \to \mathfrak{gl}(V) \). Since \( h \) is semisimple (diagonal) then \( \phi(h) \) is diagonal
Theorem 8. If \( v \in V \) then we employ our fundamental equality to get 
\[
\lambda.\nu = (\lambda + 2).\nu, \quad \nu = 0
\]
where \( \lambda \) is the eigenvalue of \( h \) with respect to some basis of \( V \). The eigenvalues \( \lambda \) are called \textbf{weights} and the corresponding spaces \( V_\lambda \) - \textbf{weight spaces}. These agree with a more general definition which we will give when we consider representations of semisimple Lie algebras in general.

Now we should see how \( L \) acts on these \( V_\lambda \). If \( v \in V_\lambda \), then we employ our fundamental equality to get 
\[
h.(x.v) = [h,x]v + x(hv) = 2x.v + x.\lambda v = (\lambda + 2).x.v,
\]
so \( x.v \) is an eigenvector of \( h \) with eigenvalue \( \lambda + 2 \), and \( x.\nu \in V_{\lambda+2} \). Similarly \( h(y.v) = (\lambda - 2)y.v \), so \( y(V_\lambda) \subset V_{\lambda+2} \). Since \( V \) is finite dimensional and since \( V = \bigoplus V_\lambda \) as a vector space we have only finitely many \( \lambda \), in particular, there must be a \( \lambda \) such that \( V_{\lambda+2} = 0 \) (otherwise we get all \( \lambda + 2\mathbb{Z} \)), in particular we’ll have \( x(V_\lambda) = 0 \). We call any \( v \in V_\lambda \) for which \( x.v = 0 \) a maximal vector of weight \( \lambda \).

Now let \( V \) be irreducible. Choose one of these maximal vectors, \( v_0 \in V_\lambda \). Since \( x, y, h \) generate \( L \) both as a vector space and algebra, it is natural to think that the vector space generated by successive application of \( x, y, h \) on \( v_0 \) will be stable under \( L \). We have \( h.v_0 = \lambda.v_0, x.v_0 = 0 \) and let \( v_i = 1/\mathfrak{d}! y^i v_0 \). By induction we will show that

**Lemma 3.**

1. \( h.v_i = (\lambda - 2i)v_i \),
2. \( y.v_i = (i + 1)v_{i+1} \),
3. \( x.v_i = (\lambda - i + 1)v_{i-1} \).

**Proof.** We have that \( y.v_0 = v_1 \) by definition and in fact the second part of the lemma follows straight from the definition of \( v_i, v_{i+1} = 1/(i + 1)!y^{i+1}v_0 = 1/(i + 1)y.(1/\mathfrak{d}!y^iv_0) = 1/(i + 1)v_i \). As for the rest, if (1) holds for \( v_i \), then \( v_i \in V_{\lambda-2i} \) and so as we remarked \( y.v_i \in V_{\lambda-2i} \), so \( h.(y.v_i) = (\lambda - 2(i + 1))y.v_i \), i.e. \( (i + 1)h.v_{i+1} = (i + 1)(\lambda - 2i)v_{i+1} \) from (2) and so (1) follows by induction. We use (2) and our fundamental equality again to show (3):

\[
\begin{align*}
  x.v_{i+1} &= x(1/(i + 1)y.v_i) = 1/(i + 1)(x.y.v_i) = 1/(i + 1)\{x.y.v_i + x.y.v_i\} = \\
  &= 1/(i + 1)(h.v_i + y(\lambda - i + 1)v_{i-1}) = 1/(i + 1)((\lambda - 2i)v_i + (\lambda - i + 1)y.v_{i-1}) = \\
  &= 1/(i + 1)((\lambda - 2i)v_i + (\lambda - i + 1)i.v_i) = 1/(i + 1)(\lambda - i + 1)v_i - i^2v_i = (\lambda - i)v_i,
\end{align*}
\]

which is what we needed to prove.

From this recurrences we can show the following main theorem which describes the irreducible representations of \( \mathfrak{sl}(2, F) \) completely.

**Theorem 8.** If \( V \) is an irreducible \( \mathfrak{sl}(2, F) \)-module, then

1. It is a direct sum of weight spaces \( V_\mu \) for \( \mu = m, m - 2, \ldots, -(m - 2), -m \), such that each \( V_\mu \) is one dimensional and is spanned by an eigenvector of \( h \) of eigenvalue \( \mu \).
2. \( V \) has a unique maximal vector, whose weight is \( m \), called the highest weight of \( V \).
3. The action of \( L \) on \( V \) is determined by the recurrences in the lemma, so up to isomorphism there is a unique irreducible \( \mathfrak{sl}(2, F) \)-module of dimension \( m + 1 \).

**Proof.** From part (1) of lemma (3) we have that the \( v_i \)’s are linearly independent being eigenvectors for different eigenvalues. Now since \( \dim V < \infty \), the \( v_i \)’s must be finitely many, so there is a minimal \( m \), such that \( v_{m+1} = 0 \), so \( v_m \neq 0 \). By the recurrence (2) we have of course that \( v_i = 0 \) for all \( i > m \) and \( v_i \neq 0 \) for all \( i = 0, \ldots, m \). The subspace of \( V \) spanned by \( \{v_0, v_1, \ldots, v_m\} \) is an \( \mathfrak{sl}(2, F) \)-submodule of \( V \), since it is fixed by \( x, y, h \) which generate \( \mathfrak{sl}(2, F) \) as a vector space. Since \( V \) is irreducible then we must have \( V = \text{span} \{v_0, v_1, \ldots, v_m\} \) and then \( \dim V = m + 1 \).

Moreover, since \( v_{m+1} = 0 \), relation (3) of lemma (3) shows that \( 0 = x.v_{m+1} = (\lambda - m)v_m \) and since \( v_m \neq 0 \) we must have \( \lambda = m \). From relation (1) then we have that \( h.v_i = (m - 2i)v_i \), so the eigenvalues are precisely \( m, m - 2, \ldots, m - 2m = -m \), in particular symmetric around 0. The maximal vector is naturally
\( v_0 \) and is unique up to scalar as the only one in \( V_m \cap V \). The uniqueness of \( V \) follows from the fact that the choice of \( m \) determines it completely by the recurrence relations of (3) and the fact that, again, \( x, y, z \) generate \( \mathfrak{sl}(2, F) \) (so that these relations determine the action of \( \mathfrak{sl}(2, F) \)).

By Weyl’s theorem we have that any \( \mathfrak{sl}(2, F) \)-module \( W \) is completely reducible and so every eigenvector of \( h \) would be in some irreducible module and then by the theorem it will be in one of the \( V_i \)-s, in particular its eigenvalue will be an integer. We also have from the theorem that in any irreducible \( V \) we have either 0 or 1 as an eigenvalue (if \( m \) is even or odd respectively). If \( W = \oplus V_i \) for some irreducible \( V_i \)’s, then we have \( W_\mu = \oplus V_i^\mu, \) in particular \( W_0 = \oplus V_0^i \) and \( W_1 = \oplus V_1^i \) and so \( \dim W_0 = \sum_{i | V_0^i \neq 0} 1 \) and \( \dim W_1 = \sum_{i | V_1^i \neq 0} 1 \) and since each \( V^i \) falls in exactly one of these categories, then \( \dim W_0 + \dim W_1 \) is the number of \( V_i \)’s in \( W \), i.e. number of summands.

Apart from existence we have essentially described the irreducible representations of \( \mathfrak{sl}(2, F) \). We will soon show the general case of \( \mathfrak{sl}(n, F) \).

## 3 Lie Groups and their Lie algebras

So far we have considered Lie algebras from the axiomatic point of view, but that didn’t motivate neither their existence nor our current interest in them. To motivate their existence we will introduce a much more natural structure.

**Definition 8.** A Lie group \( G \) is a \( \mathbb{C}^\infty \) manifold with differentiable maps multiplication \( \times : G \times G \to G \) and inverse \( \iota : G \to G \) which turn \( G \) into a group.

Naturally a morphism between Lie groups would be a map that is both differentiable and a group homomorphism. A Lie subgroup would be a subgroup that is also a closed submanifold.

The main example of a Lie group to our interest is the general linear group \( GL_n \mathbb{R} \) (or over \( \mathbb{C} \)) of invertible \( n \times n \) matrices. The map \( \times : GL_n \times GL_n \to GL_n \) is given by the usual matrix multiplication and is clearly differentiable, the inverse map can be given by Cramer’s rule and so is also differentiable. Other Lie groups of interest happen to be subgroups of \( GL_n \), for example the special linear group \( SL_n \), the group \( B_n \) of upper triangular matrices, the special orthogonal group \( SO_n \) of transformations of \( \mathbb{R}^n \) preserving some bilinear form.

A representation of a Lie group \( G \) is defined as the usual representations of groups, i.e. a morphism \( G \to GL(V) \).

It turns out the tangent spaces to Lie groups are Lie algebras. To see how this works we consider maps \( \rho : G \to H \) and we aim to show that they are completely described by their differentials at the identity, i.e. \( d\rho_e : T_e G \to T_e H \). This will reduce our studies to the studies of Lie algebras.

So let \( \rho : G \to H \) be a morphism of Lie groups. For any \( g \in G \) we have its right action on \( G \) is respected by \( \rho \). However it does not fix the identity, so we go further and consider conjugation by \( g \). Define \( \Psi_g : G \to G \) by \( \Psi_g(x) = g.x.g^{-1} \) for \( x \in G \). Now we have that \( \Psi_{\rho(g)}(\rho(x)) = \rho(g)\rho(x)\rho(g)^{-1} = \rho(g.x.g^{-1}) = \rho \circ \Psi_g(x) \), i.e. the following diagram commutes

\[
\begin{array}{ccc}
G & \xrightarrow{\rho} & H \\
\Psi_g & & \Psi_{\rho(g)} \\
\downarrow & & \downarrow \\
G & \xrightarrow{\rho} & H.
\end{array}
\]

In particular we have that \( \Psi : G \to \text{Aut}(G) \) and \( \Psi_g(e) = e \). \( \Psi_g \) is also a Lie group morphism, in particular differentiable, so we can consider its differential at \( e \).

**Definition 9.** The representation \( \text{Ad} : G \to \text{Aut}(T_e G) \) given by \( \text{Ad}(g) = (d\Psi_g)_e : T_e G \to T_e G \) is called the adjoint representation of \( G \).
Since \( \rho \circ \Psi_g = \Psi_{\rho(g)} \circ \rho \) we have that for any arc \( \gamma(t) \) in \( G \) starting at \( e \) and \( \frac{d\gamma}{dt}\big|_{t=0} = v \) we have \( \rho(\Psi_g(\gamma(t))) = \Psi_{\rho(g)}(\rho(\gamma(t))) \) and differentiating with respect to \( t \) we get \( d\rho(d\Psi_g(d\gamma(t)/dt)) = d\Psi_{\rho(g)}(d\rho(d\gamma(t)/dt)) \) and evaluating at \( t = 0 \) we get \( d\rho(d\Psi_g(v)) = d\Psi_{\rho(g)}(d\rho(v)) \), in other words the following diagram commutes:

\[
\begin{array}{ccc}
T_eG & \xrightarrow{(dp)_e} & T_eH \\
\text{Ad}(g) \downarrow & & \downarrow \text{Ad}(\rho(g)) \\
T_eG & \xrightarrow{(dp)_e} & T_eH.
\end{array}
\]

We want to get rid of the map \( \rho \) and leave only \( d\rho \) in order to consider only maps between tangent spaces, so we take again a differential. We define

\[ \text{ad} = d\text{Ad} : T_eG \to \text{End}(T_eG), \]

since the tangent space of \( \text{Aut}(T_eG) \subset \text{End}(T_eG) \) is just the space of endomorphisms of \( T_eG \). Then we have that \( \text{ad}(X) \in \text{End}(T_eG) \), so for any \( Y \in T_eG \), we can define

\[ [X,Y] := \text{ad}(X)(Y), \]

which is a bilinear map. Since \( \rho \) and \( (dp)_e \) respect \( \text{Ad} \), we have that \( d\rho_e(\text{ad}(X)(Y)) = \text{ad}(d\rho_e(X))(d\rho_e(Y)) \), i.e. the following diagram commutes

\[
\begin{array}{ccc}
T_eG & \xrightarrow{(dp)_e} & T_eH \\
\text{ad}(X) \downarrow & & \downarrow \text{ad}(\rho(X)) \\
T_eG & \xrightarrow{(dp)_e} & T_eH,
\end{array}
\]

so the differential \( d\rho_e \) respects the bracket and moreover we have arrived at a situation involving only differentials and tangent spaces and in particular no actual map \( \rho \).

The notation \([X,Y]\) is not coincidental with previous section, it is in fact the same bracket and \( T_eG \) is a Lie algebra. So that things become explicit we will consider the case when \( G \) is a subgroup of \( GL_n(\mathbb{C}) \), in which case we know exactly what \( \text{Ad}(g) \) is. Let \( \beta(t) \) and \( \gamma(t) \) be two arcs in \( G \) with \( \beta(0) = \gamma(0) = e \) and \( \frac{d\beta}{dt}\big|_{t=0} = X \) and \( \frac{d\gamma}{dt}\big|_{t=0} = Y \). Then \( \Psi_g(\gamma(t)) = g\gamma(t)g^{-1} \), so \( \text{Ad}(g)(Y) = d\left(g\gamma(t)g^{-1}\right)/dt\big|_{t=0} = gYg^{-1} \). Then

\[
\begin{align}
\text{ad}(X)(Y) &= \frac{d}{dt}\text{Ad}(\beta(t))(Y) = \frac{d}{dt}\big|_{t=0}(\beta(t)Y\beta(t)^{-1}) = \\
&= \frac{d\beta(t)}{dt}\big|_{t=0}X\beta(t)^{-1} + \beta(t)^{-1}\frac{d\beta(t)}{dt}\big|_{t=0}X = XY + eY(-X) = XY - YX.
\end{align}
\]

This definition of the bracket explains why we defined the action of a Lie algebra on a tensor product the way we did. Let \( V \) and \( W \) be representations of a Lie group \( G \), then the usual definition of the representation \( V \otimes W \) is given by \( g(v \otimes w) = g(v) \otimes g(w) \). If \( \beta(t) \) is an arc in \( G \) with \( \beta(0) = e \) and \( \frac{d\beta}{dt}(0) = X \), then the action of \( X \in T_eG \) on any space \( V \) is given by \( X(v) = \frac{d}{dt}\big|_{t=0}\beta(t).v \), so the action on \( V \otimes W \) will be naturally given by differentiation by parts, i.e.

\[ X(v \otimes w) = \frac{d}{dt}(\beta(t)v \otimes \beta(t)w) = \frac{d}{dt}\big|_{t=0}\beta(t)v \otimes w + v \otimes \frac{d}{dt}\big|_{t=0}\beta(t)w = X(v) \otimes w + v \otimes X(w). \]

**Definition 10.** We define the Lie algebra associated to a Lie group \( G \) to be the tangent space to \( G \) at the origin, i.e. \( T_eG \).
To make our discussion complete and justify our study of Lie algebras (as opposed to Lie groups) we will prove the following two principles

**Theorem 9** (First Principle). Let $G$ and $H$ be Lie groups, $G$ connected. A map $\rho : G \to H$ is uniquely determined by its differential $d\rho_e : T_eG \to T_eH$ at $e$-the identity.

**Theorem 10** (Second Principle). Let $G$ and $H$ be Lie groups with $G$ connected and simply connected. A linear map $\delta : T_eG \to T_eH$ is the differential of a homomorphism $\rho : G \to H$ if and only if the map preserves the bracket operation in the sense that $\delta([X,Y]) = [\delta(X), \delta(Y)]$.

We are going to prove these theorems later, after we establish the relationship between a Lie group $G$ and its Lie algebra. For that we will consider one parameter subgroups.

Let $g \in G$ and let $m_g : G \to G$ be the map given by left multiplication by $g$. Let $X \in L = T_eG$, we define a vector field $v_X$ on $G$ by

$$v_X(g) = (m_g)_*(X).$$

(5)

We can integrate this field, i.e. there is a unique differentiable map ("the integral") $\phi : I \to G$, where $0 \in I$, $I$ is open and $\phi(0)$ is any given point, such that $\phi'(t) = v_X(\phi(t))$. If $\alpha(t) = \phi(s) \circ \phi(t)$ for $s, t \in I$ and $\beta(t) = \phi(s+t)$, then $\beta'(t) = \phi'(s+t) = v_X(\phi(s+t)) = \alpha(\beta(s+t))$ and on the other hand $\alpha'(t) = \frac{d}{dt}(m_{\phi(s)} \phi(t)) = (m_{\phi(s)})_* (\phi'(t)) = (m_{\phi(s)})_* (v_X(\phi(t))) = (m_{\phi(s)})_* (m_{\phi(t)})_* (X) = (m_{\phi(s), \phi(t)})_* X = v_X(\phi(s) \phi(t)) = v_X(\alpha(t))$.

Since $\alpha(0) = \beta(0)$ the uniqueness of such maps implies $\alpha(t) = \beta(t)$, in other words $\phi(s+t) = \phi(s) \phi(t)$, so thus we defined a homomorphism $\phi_X : \mathbb{R} \to G$. This Lie group map $\phi_X$ is called the **one-parameter subgroup** of $G$ with tangent vector $X$.

If we had a homomorphism $\phi : \mathbb{R} \to G$, such that $\phi'(0) = X$, then we would uniquely determine $\phi$ (Exercise 8.31 in [1]). This follows because if we fix $s$ and take $\frac{d}{dt} \phi(s+t) = \frac{d}{dt} \phi(t)$ we will have from the left-hand side $\phi'(s+t)$ and from the right-hand side $(m_{\phi(s)})_*(\phi'(t))$. Taking $t = 0$ we get that $\phi'(s) = (m_{\phi(s)})_*(\phi(0)) = v_X(\phi(s))$, which as we mentioned is uniquely defined. If then $\psi : G \to H$ is map of Lie groups, then $\psi \circ \phi_X : \mathbb{R} \to H$ is a homomorphism. We have that $\psi_*(X)$ is a tangent vector of $H$ at the identity, so $\phi_{\psi, X}$ is the unique homomorphism with such tangent vector. However the homomorphism $\psi \circ \phi_X$ has tangent vector at 0 $\frac{d}{dt} \psi(\phi_X(t))|_{t=0} = \psi_*(\phi_X'(0)) = \psi_*(X)$, so they must coincide, i.e.

$$\phi_{\psi, X} = \psi \circ \phi_X.$$  

(6)

Let $L$ be the Lie algebra of $G$, i.e. $L = T_eG$.

**Definition 11.** The exponential map $\exp : L \to G$ is given by $\exp(X) = \phi_X(1)$.

We have that $\phi_{\lambda X}$ is the unique homomorphism such that $\phi'(0) = \lambda X$. Let $\psi(t) = \phi_X(\lambda t)$. Then $\psi$ is still a homomorphism and $\psi'(0) = \lambda \phi_X(0) = \lambda X$ by the chain rule, so necessarily we must have $\phi_{\lambda X}(t) = \phi_X(\lambda t)$, so the exp maps restricted to the projective space of lines through the origin gives the one-parameter subgroups of $G$.

We can apply (6) to the exponential map $\exp : L \to G$ and get that exp is the unique map from $L$ to $G$ taking $0$ to the identity $e$ of $G$ with differential at the origin the identity and whose restrictions to lines through the origin are the one-parameter subgroups of $G$. Again thanks to (6) we have that for any map $\psi : G \to H$ $\phi_{\psi, X} = \psi(\phi_X)$, i.e. $\exp(\psi_*(X)) = \psi(\exp(X))$ by evaluating at $t = 1$.

Since the differential of exp at the origin is an isomorphism, in particular it’s invertible, the inverse mapping theorem tells us that exp is invertible in a neighborhood $U$ of $e$, so in particular its image will contain that neighborhood. Now let $G$ be connected and $H$ be a subgroup of $G$ generated by elements in a neighborhood $U$ of $e$ (Exercise 8.1 in [1]), then $H$ is an open subset of $G$. Suppose $H \neq G$. All of its cosets $aH$ being translates of $H$ will also be open and besides disjoint from each other, so $(\cup_{aH \neq H} aH) \cap H = 0$, $H \cup (\cup_{aH \neq H}) = G$ and hence $G$ is the union of two disjoint open sets. These sets are also closed in $G$ being complements to open, so $G$ is the disjoint union of two closed sets and hence it is not connected. So $G$ is generated as a group by the elements of $U$, in particular $G$ is generated by $\exp(L)$. This proves the first principle: since for any map $\psi : G \to H$ of Lie groups we have the commutative diagram
Consider the product \( L \) since it is clearly a vector space and the group \((X, \alpha)\) is a linear group, which follows from another fact that every Lie algebra \( X \) is determined uniquely by the differential \( \psi_x = (d\psi)_x \).

The reason the exponential map is called that becomes apparent when we consider the case of \( GL_n(\mathbb{R}) \). For any \( X \in \text{End}(V) \) we can set

\[
\exp(X) = 1 + X + \frac{X^2}{2} + \frac{X^3}{6} + \ldots,
\]

which converges because if \( a \) is the maximal (absolute value) entry in \( X \), then the maximal entry in \( X^k \) is by induction less than \( n|a|^k \) \((n = \dim V)\). This map has an inverse \( \exp(-X) \). Since the differential of this map at \( X = 0 \) is the identity and the map \( \phi_X(t) = \exp(tX) \) satisfies \( \phi_X'(t) = \exp'(tX) = \exp(tX)X = (m_{\phi_X(t)})X = v_X(\phi_X(t)) \), we have that \( \exp \) restricted to any line through the origin gives the one-parameter subgroups, so it is the same \exp we defined above.

We want to show now something stronger than that the \( \exp \) generate \( G \), namely that the group structure of \( G \) is encoded in its Lie algebra. We want to express \( \exp(X) \) and \( \exp(Y) \) as \( \exp(Z) \) where \( Z \in L. \) The difficulty arises from the fact that \( \exp(X) \exp(Y) \) contains products \( X,Y \), which don’t have to belong to \( L \). However the bracket \([X,Y] \in L \), so we aim to express \( \exp(X) \exp(Y) \) as sum of bracket operations. We have that

\[
\log(g) = (g - I) - \frac{(g - I)^2}{2} + \frac{(g - I)^3}{3} + \ldots \quad \text{is the formal power series of the inverse of } \exp.
\]

We define then

\[
X \ast Y = \log(\exp(X).\exp(Y)) \in L, \text{ i.e. } \exp(X \ast Y) = \exp(X) \exp(Y)
\]

and we have that \( X \ast Y = \log(I + (X + Y) + (X^2/2 + X.Y + Y^2/2) + \ldots) \).

**Theorem 11** (Campbell-Hausdorff formula). The quantity \( \log(\exp(X) \exp(Y)) = X + Y + \frac{1}{2}[X,Y] + \ldots \) can be expressed as a convergent sum of \( X,Y \) and the bracket operator. In particular it does depend only on the Lie algebra.

Using this theorem and the assumption that every Lie group may be realized as a subgroup of the general linear group, which follows from another fact that every Lie algebra \( L \) is a subalgebra of \( \mathfrak{gl}_n(\mathbb{R}) \) (we showed it for semisimple Lie algebras) and \( G \) is generated by \( \exp(L) \), we can show that \( \exp(X) \exp(Y) \in \exp(L) \) for \( X,Y \in L \). So if we have a Lie subalgebra \( L_h \) of \( L \), then its image under the exponential map is locally closed and hence the group \( H \) generated by \( \exp(L_h) \) is an immersed subgroup of \( G \) with tangent space \( T_eH = L_h \).

Now let \( H \) and \( G \) be Lie groups with \( G \) simply connected and let \( L_h \) and \( L_g \) be their Lie algebras. Consider the product \( G \times H \). Its Lie algebra is \( L_g \oplus L_h \). If \( \alpha : L_g \rightarrow L_h \) is a map of Lie algebras, then we can consider \( L_J \subset L_g \oplus L_h \), the graph of \( \alpha \) \((L_J = \{(X,\alpha(X))|X \in L_g\})\). Then \( L_J \) is a Lie subalgebra of \( L_g \times L_h \), since it is clearly a vector space and \([[(X,\alpha(X)),(Y,\alpha(Y))]] = [(X,Y),[\alpha(X),\alpha(Y)]] = ([X,Y],\alpha([X,Y]))\). Then by what we just noticed the group \( J \) generated by \( \exp(L_J) \) is an immersed Lie subgroup of \( G \times H \) with tangent space \( L_J \). Now we have a projection map \( \pi : J \rightarrow G \), whose differential \( d\phi : L_J \rightarrow L_g \) should agree with the projection of the inclusion map \( L_J \hookrightarrow L_g \times L_h \rightarrow L_g \), i.e. this is the map \( (X,\alpha(X)) \rightarrow X \), so \( d\phi \) is an isomorphism, so again by the inverse function theorem \( \phi \) must be an isogeny and since \( G \) is simply connected the image of \( \pi \) must be \( G \), so \( \pi \) is an isomorphism. The projection on the second factor composed with this isomorphism \( \eta = \pi_2 \circ \pi^{-1} : G \rightarrow H \) is a map of Lie groups, whose differential is \( d\eta = d\pi_2 \circ d\pi^{-1} = \alpha \circ id = \alpha \).

This proves the second principle that a linear map of Lie algebras is the differential of a map of their Lie groups if and only if it is a map of Lie algebras.
4 Representations of semisimple Lie algebras

Returning back to our study of representations, let's take a look at what we did with \( \mathfrak{sl}(2, F) \). We found elements \( x, y, h \), such that \( x \) and \( y \) where eigenvectors with respect to \( \text{ad}(h) \), \( h \) was semisimple and \( \mathfrak{sl}(2, F) = h \oplus x \oplus y \). Thanks to these elements we could decompose an irreducible representation into eigenspaces with respect to \( h \) with \( x \) and \( y \) sending one to another. We would like to do something similar for any semisimple Lie algebra \( L \). We shouldn’t expect to find elements corresponding to \( h, x \) and \( y \), but rather subalgebras, and if we want to decompose \( L \) into a direct sum of subalgebras the most natural approach is to look at the adjoint action of certain elements of \( L \) on \( L \).

The crucial fact about \( h \) is that it is semisimple, acts diagonally. Suppose we can find a subalgebra \( H \subset L \) of semisimple elements, which is abelian and acts diagonally on \( L \). Then by the action of \( H \) on \( L \) we can decompose \( L \) into eigenspaces, i.e. for \( \alpha \) a linear functional on \( H \) let \( L_\alpha = \{ x \in L | h(x) = \alpha(x)h, h \in H \} \), then \( L = H \oplus (\oplus_{\alpha \in H} L_\alpha) \), where we will have \( H = L_0 \). Since \( L \) is finite dimensional we will have a finite collection of \( \alpha \), the weights of the adjoint representation, we call them the roots of \( L \), and the corresponding spaces \( L_\alpha \) are called the root spaces of \( L \). The set of roots will be denoted \( \mathcal{R} \). We will show that the adjoint action of \( L_\alpha \) carries \( L_\beta \) into \( L_{\alpha + \beta} \) and so the roots will form a lattice, \( \Lambda_R \subset H^* \). It will be of rank the dimension of \( H \). Each \( L_\alpha \) will be one dimensional and \( \mathcal{R} \) will be symmetric about 0.

Now if \( V \) is any irreducible representation of \( L \), then similarly to the above decomposition we can decompose it into a direct sum \( V = \oplus V_\alpha \), where \( V_\alpha = \{ v \in V | h(v) = \alpha(h)v, h \in H \} \) are the eigenspaces with respect to \( H \). These \( \alpha \) in \( H^* \) are called the weights of \( V \) and \( V_\alpha \) - the weight spaces, \( \dim V_\alpha \) - the multiplicity of the weight \( \alpha \). We will have that for any root \( \beta, L_\beta \) send \( V_\alpha \) into \( V_{\alpha + \beta} \). Then the subspaces \( \oplus_{\beta \in \Lambda_R} V_{\alpha + \beta} \) will be an invariant subspace of \( V \) and since \( V \) is irreducible the weights should all be translates of one another by \( \Lambda_R \).

4.1 Root space decomposition

We will consider the case of any representation later, for the moment we will focus on the root space decomposition. Our presentation here will follow [2].

We call a subalgebra toral if it consists of semisimple elements. \( L \) has such subalgebras, as otherwise all its elements would be nilpotent and so \( L \) will be nilpotent by Engel’s theorem, hence solvable and hence not semisimple. Any toral subalgebra is abelian by the following reasoning. Let \( T \) be toral, \( x \in T \), so \( \text{ad} x \) is semisimple and so over an algebraically closed field it is diagonalizable. So if \( \text{ad} x \) has only 0 eigenvalues, then \( \text{ad} x = 0 \). Suppose it has an eigenvalue \( \alpha \neq 0 \), i.e. there is a \( y \in T \), such that \( [x, y] = \alpha y \).

Since \( y \) is also semisimple, so is \( \text{ad} y \) and it has linearly independent eigenvectors \( y_1 = y, \ldots, y_n \) (since \( \text{ad}(y)(y) = 0 \) of eigenvalues 0, \( b_2, \ldots, b_n \), we can write \( x \) in this basis as \( x = a_1 y_1 + \ldots + a_n y_n \). Then \( -\alpha y = \text{ad} y(x) = 0.y + b_2a_2y_2 + \ldots \), i.e. \( y \) is a linear combination of the other eigenvectors, which is impossible. So \( \alpha = 0 \) and \( \text{ad} x = 0 \) for all \( x \in T \), i.e. \( [x, y] = 0 \) for all \( x, y \in T \).

Let \( H \) be a maximal toral subalgebra of \( L \), i.e. not included in any other. For any \( h_1, h_2 \in H \), we have \( \text{ad} h_1 \circ \text{ad} h_2(x) = [h_1, [h_2, x]] = -[h_2, [x, h_1]] = -[x, [h_1, h_2]] = [h_2, [h_1, x]] = \text{ad} h_2 \circ \text{ad} h_1(x) \) by the Jacobi identity, so \( \text{ad} H \) consists of commuting semisimple endomorphisms and by a standard theorem in linear algebra these are simultaneously diagonalizable. So we can find eigenspaces \( L_\alpha = \{ x \in L | [h x] = \alpha(h)x \} \) for all \( h \in H \) for \( \alpha \in H^* \), such that they form a basis of \( L \), i.e. we can write the root space decomposition:

\[
L = C_L(H) \oplus \prod_{\alpha \in \Phi} L_\alpha, \tag{7}
\]

where \( \Phi \) is the set of \( \alpha \in H^* \) for which \( L_\alpha \neq 0 \). Here \( C_L(H) \) is just \( L_0 \). Our goal is to prove that \( C_L(H) = H \).

We will show first that

\footnote{Let \( A \) and \( B \) be two diagonalizable commuting matrices, then for any eigenvalue \( \alpha \) of \( B \) and \( v \) such that \( Bv = \alpha v \) we have \( \alpha(Av) = ABv = B(Av) \), i.e. \( A \) fixes eigenspaces \( B_\alpha \) of \( B \). Since then \( A | B_\alpha \) is still semisimple we can diagonalize and obtain an eigenbasis which is necessarily an eigenbasis for \( B \) also.}
Lemma 4. For all $\alpha, \beta \in H^*$ we have $[L_{\alpha}L_{\beta}] \subset L_{\alpha+\beta}$. For any $x \in L_{\alpha}$ of $\alpha \neq 0$, we have that $\operatorname{ad} x$ is nilpotent and if $\beta \neq -\alpha$ then the spaces $L_{\alpha}, L_{\beta}$ are orthogonal with respect to the Killing form $\kappa$.

Proof. Let $x \in L_{\alpha}$, $y \in L_{\beta}$ and $h \in H$, then by the Jacobi identity we have $\operatorname{ad} h([x, y]) = [h, [x, y]] = [x, [h, y]] + [h, [x, y]]$. Let $\kappa(h, x) = \alpha(h)x, \kappa(h, y) = \beta(h)y$, so $\kappa(x, [h, y]) = (\alpha + \beta)(h)x$, $\kappa(y, [x, h]) = -\alpha(h)x$. But $\kappa(x, [h, y]) = 0$, $\kappa(y, [x, h]) = 0$, so $\alpha = 0$.

Next, let $h \in H$, such that $(\alpha + \beta)(h) \neq 0$. Since the Killing form is associative, we have for $x \in L_{\alpha}, y \in L_{\beta}$ that $\kappa(x, [h, y]) = \kappa(x, [h, y]) = (\beta(h) - \alpha(h))x$, so $\kappa(x, [h, y]) = 0$ from one hand and from another $\kappa(x, [h, y]) = -\kappa(x, [y, h]) = -\kappa(x, [y, h]) = 0$. So $\kappa(x, [y, h]) = 0$.

As a corollary to this lemma we have that the restriction of $\kappa$ to $C_L(H)$ is nondegenerate. Since $\kappa$ is nondegenerate on $L$ and $L_0 = C_L(H)$ is orthogonal to $L$ for all $\alpha \neq 0$, then there is no $z \in L_0$, such that $\kappa(z, L) = 0$. This will help us prove the next theorem.

Theorem 12. Let $H$ be a maximal toral subalgebra of $L$. Then $H = C_L(H)$.

Proof. Our goal is to show that $H$ contains all semisimple elements of $C_L(H)$ and that $C_L(H)$ cannot have any nilpotent ones.

(1) First of all since $C_L(H)$ maps $H$ via $\operatorname{ad}$ to 0, then the nilpotent and semisimple parts of each $\operatorname{ad} x, x \in C_L(H)$ being polynomials in $\operatorname{ad} x$ map $H$ into 0 also. But $(\operatorname{ad} x)_s = \operatorname{ad} x_s$ and $(\operatorname{ad} x)_n = \operatorname{ad} x_n$, so $x_s, x_n \in C_L(H)$ too. Now suppose $x \in C_L(H)$ is semisimple. Then the subalgebra generated by $H$ and $x$ is still toral, as $x$ commutes with everything in $H$ and sum of semisimple elements is still semisimple which follows for example from the simultaneous diagonalizability. Since $H$ is maximal with respect to this property we must have $x \in H$.

(2) If $x \in C_L(H)$ is nilpotent then $\operatorname{ad} x$ is nilpotent, $[x, H] = 0$ and if $y \in H$, then $\operatorname{ad} x \circ \operatorname{ad} y = \operatorname{ad} y \circ \operatorname{ad} x$, so $(\operatorname{ad} x \circ \operatorname{ad} y)^k = (\operatorname{ad} y)^k \circ (\operatorname{ad} x)^k$ for any $k$ and finally $\operatorname{ad} x \circ \operatorname{ad} y$ is nilpotent, i.e. $\kappa(x, y, h) = 0$. Since $\kappa$ is nondegenerate on $C_L(H)$ we must then have $\kappa(h, H) = 0$ for $h$ semisimple, i.e. in $H$. So $\kappa$ is nondegenerate on $H$ also.

(3) We also have that $C_L(H)$ is nilpotent by Engel's theorem as follows. For any $x \in C_L(H), x = x_s + x_n$ and $\operatorname{ad} x_n$ is clearly nilpotent and $\operatorname{ad} x_s = 0$ on $C_L(H)$ since by point (1) $x_s \in H$. Since again by $x_s \in H$ and $x_n \in C_L(H)$, $[x_s, x_n] = 0$, we have that $\operatorname{ad} x$ is the sum of two commuting nilpotents and so is nilpotent.

(4) If $x \in C_L(H)$ is nilpotent then $\operatorname{ad} x$ is nilpotent, $[x, H] = 0$ and if $y \in H$, then $\operatorname{ad} x \circ \operatorname{ad} y = \operatorname{ad} y \circ \operatorname{ad} x$, so $(\operatorname{ad} x \circ \operatorname{ad} y)^k = (\operatorname{ad} y)^k \circ (\operatorname{ad} x)^k$ for any $k$ and finally $\operatorname{ad} x \circ \operatorname{ad} y$ is nilpotent, i.e. $\kappa(x, y, h) = 0$. Since $\kappa$ is nondegenerate on $C_L(H)$ we must then have $\kappa(h, H) = 0$ for $h$ semisimple, i.e. in $H$. So $\kappa$ is nondegenerate on $H$ also.

(5) Point (4) tells us that $H \cap [C_L(H)C_L(H)] = 0$ as follows. For $h \in H, x, y \in C_L(H)$ we have $\kappa(h, [x, y]) = \kappa([h, x], y) = 0$. If $[x, y] \in H$ this contradicts nondegeneracy, so $[x, y] \notin H$ and hence $H \cap [C_L(H)C_L(H)] = 0$.

(6) Suppose $[C_L(H)C_L(H)] \neq 0$, by lemma 1 we have that $C_L(H)$ acts on $[C_L(H), C_L(H)]$ via the adjoint representation, so there is a common eigenvector of eigenvalue 0, i.e. $[C_L(H), C_L(H)] \cap Z(C_L(H)) \neq \emptyset$. Let $z$ be such element, by (2) and then (5) $z$ cannot be semisimple, so $z_n \neq 0$ and $z_n \in C_L(H)$ by (1) and since $[z_n, z_s, c] = 0$ for all $c \in C_L(H)$ and $[z_s, c] = 0$ since $z_s \in H$ we must have $[z_n, c] = 0$ for all $c \in C_L(H)$. Then $\operatorname{ad} z_n \circ \operatorname{ad} c$ is nilpotent by commutativity and so $\kappa((z_n, C_L(H))) = 0$ contradicting nondegeneracy. So $[C_L(H)C_L(H)] = 0$ and is abelian.

(7) Finally, since $C_L(H)$ is abelian for any nilpotent element $x_n$ of it we will have $\kappa(x_n, C_L(H)) = 0$ as $x_n, c$ commute for all $c \in C_L(H)$. This contradicts the nondegeneracy of $\kappa$ over $C_L(H)$, so $C_L(H)$ has no nilpotent elements. But then it must have only semisimple ones by (1) and then by (2) they are all in $H$, so $H = C_L(H)$.

Together with the previous arguments this shows
Theorem 13 (Cartan decomposition). Let $L$ be semisimple. If $H$ is a maximal toral subalgebra of $L$ then there are eigenspaces $L_{\alpha} = \{ x \in L | hx = \alpha(h)x \ \text{for all} \ h \in H \}$ for $\alpha \in H^*$, such that $L = H \oplus (\bigoplus_{\alpha \in H} L_{\alpha})$.

From the above proof we have in particular that $\kappa$ is nondegenerate on $H$, then we can find an orthonormal basis with respect to $\kappa, h_1, \ldots, h_l$. Then every $\phi \in H^*$ is uniquely determined by its values on $h_i$, say $\phi(h_i) = \phi_i$. If $t_\phi = \phi h_1 + \ldots + \phi h_l$, then $\kappa(t_\phi, h_i) = \phi_i$, so by linearity $\kappa(t_\phi, h) = \phi(h)$ for every $h \in H$.

We will prove some useful facts about the Cartan decomposition.

**Theorem 14.** 1. The set $\Phi$ of roots of $L$ relative to $H$ spans $H^*$.

2. If $\alpha \in \Phi$, then $-\alpha \in \Phi$.

3. Let $\alpha \in \Phi$ and let $t_\alpha$ be the element of $H$ as defined above, i.e. $\alpha(h) = \kappa(t_\alpha, h)$. If $x \in L_\alpha$, $y \in L_{-\alpha}$, then $[x, y] = \kappa(x, y)t_\alpha$.

4. If $\alpha \in \Phi$ then $[L_\alpha L_{-\alpha}]$ is one dimensional with basis $t_\alpha$.

5. $\alpha(t_\alpha) \neq 0$ for $\alpha \in \Phi$.

6. If $\alpha \in \Phi$ and $x_\alpha \neq 0$ is in $L_\alpha$, then there is a $y_\alpha \in L_{-\alpha}$ such that the elements $x_\alpha, y_\alpha$ and $h_\alpha = [x_\alpha y_\alpha]$ span a three dimensional subalgebra of $L$ isomorphic to $\mathfrak{sl}(2, F)$ via $x_\alpha \rightarrow x$, $y_\alpha \rightarrow y$ and $h_\alpha \rightarrow h$.

7. $[x_\alpha y_\alpha] = h_\alpha = \frac{2\mu_\alpha}{\kappa(t_\alpha, t_\alpha)}$ and $h_{-\alpha} = -h_{-\alpha}$.

**Proof.** (1) If $\Phi$ does not span $H^*$ then there is an $h \in H$ such that $\alpha(h) = 0$ for all $\alpha \in \Phi$, i.e. for any $x \in L_\alpha$, $\text{ad} h(x) = 0$, i.e. $[h, L_\alpha] = 0$. But since $[hH] = 0$ by $H$ being abelian we have $[hL] = 0$, i.e. $h \in Z(L)$. But $Z(L) = 0$ by $L$ - semisimple, so we reach a contradiction.

(2) We have from lemma 4 that $L_\alpha$ and $L_\beta$ are orthogonal unless $\beta = -\alpha$. If $-\alpha \notin \Phi$, then by the same lemma $\kappa(L_\alpha, L) = 0$ contradicting the nondegeneracy of $\kappa$ on $L$.

(3) Consider for any $h \in H \ \kappa([x, y] - \kappa(x, y)t_\alpha, h) = \kappa([x, y], h) - \kappa(x, y)\kappa(t_\alpha, h) = \kappa(x, [y, h]) - \alpha(h)\kappa(x, y) = \kappa(x, -(-\alpha(h)y)) - \alpha(h)\kappa(x, y) = 0$ by the associativity of $\kappa$. Since $\kappa$ is nondegenerate on $H$ we must have $[x, y] - \kappa(x, y)t_\alpha = 0$ (note that $[x, y] \in H$ by lemma 4).

(4) Is a direct corollary of (3), provided that $[L_\alpha L_{-\alpha}] \neq 0$. If it where, then $\kappa(L_\alpha, L) = 0$ by lemma 4, so since $L_\alpha \neq 0$ this contradicts the nondegeneracy of $\kappa$ on $L$.

(5) Suppose $\alpha(t_\alpha) = 0$. Then if $x \in L_\alpha$ and $y \in L_{-\alpha}$, such that $[x, y] \neq 0$ (by (4) it is possible), we will have that $[t_\alpha x] = \alpha(t_\alpha)x = 0$ and similarly $[t_\alpha y] = 0$, so the subspace $S$ spanned by $x$, $y$ and $t_\alpha$ is a subalgebra of $L$. It acts on $L$ via the adjoint representation and we see that $S \simeq \text{ad}_L S \subset \mathfrak{gl}(L)$. Since $\text{ad} s$ is nilpotent for all $s \in S$ (easy check that $S$ is nilpotent), then we check that $t_\alpha \in [SS]$ is nilpotent and since $S$ is solvable we’ve shown that $\text{ad}_L t_\alpha$ is nilpotent. But since $t_\alpha$ is semisimple, $\text{ad}_L t_\alpha$ is also, so $\text{ad}_L t_\alpha = 0$, i.e. $\text{ad}_L t_\alpha \in Z(L) = 0$, contradiction.

(6) Take $x_\alpha \in L_\alpha$ and let $y_\alpha \in L_{-\alpha}$ be such that $\kappa(x_\alpha, y_\alpha) = \frac{2}{\kappa(t_\alpha, t_\alpha)} = \frac{2}{\alpha(t_\alpha)}$, possible because of (4) and the nondegeneracy of $\kappa$. We set $h_\alpha = \frac{2\mu_\alpha}{\kappa(t_\alpha, t_\alpha)}$, then $[x_\alpha, y_\alpha] = h_\alpha$ by (3) and $[h_\alpha x_\alpha] = \alpha(\frac{2\mu_\alpha}{\alpha(t_\alpha)})x_\alpha = 2x_\alpha$ by linearity of $\alpha$ and similarly since $[t_\alpha y_\alpha] = -\alpha(t_\alpha)y_\alpha$ we have $[h_\alpha y_\alpha] = -2$, the same relations as in equation (1) for $\mathfrak{sl}(2, F)$, so the three-dimensional algebra spanned by $x_\alpha, y_\alpha, h_\alpha$ is isomorphic to $\mathfrak{sl}(2, F)$.

(7) By definition of $t_\alpha$ we have that $\kappa(t_\alpha, h) = \alpha(h)$, so also $\kappa(t_{-\alpha}, h) = -\alpha(h)$, so by linearity of $\kappa$ $\kappa(t_\alpha + t_{-\alpha}, h) = 0$ and by its nondegeneracy on $H$ we must have $t_\alpha = -t_{-\alpha}$ which shows (7).

In view of what we proved in the (6)th part of this proposition, let $S_\alpha \simeq \mathfrak{sl}(2, F)$ be a subalgebra of $L$, generated by $x_\alpha, y_\alpha$ and $h_\alpha$. For a fixed $\alpha \in \Phi$, let $M = \text{span}\{H_1, L_{\alpha} | \text{all}\ \alpha \in \Phi\}$, where $H_1 \subset H$ is generated by $h_\alpha$ and $\text{Ker} \ \alpha \subset H$. Lemma 4 shows that $\text{ad}(S_\alpha)(M) \subset M$, i.e. it is a $S_\alpha$ submodule of $L$, i.e. it is a representation of $S_\alpha$. 17
By our analysis about $\mathfrak{sl}(2,F)$ we have that all weights should be integers, so the ones occurring here being among 0 and 2c (as $\alpha(h_\alpha) = 2$), where cs are among the ones for which $L_{\alpha} \neq 0$. So in particular we must have $c$ an integral multiple of 1/2. Now $M$ is not actually irreducible. In fact, $S_\alpha$ is an $S_\alpha-$submodule of $M$ (it is a subspace already) and Ker $\alpha$ is also a $S_\alpha$ submodule, since for every $h \in \text{Ker} \alpha$ we have $\text{ad} x_\alpha(h) = -\alpha(h)x = 0$, same for $y_\alpha$ and $\text{ad} h_\alpha(h) = [h_\alpha, h]$ since $H$ is abelian. So $M = \text{Ker} \alpha \oplus S_\alpha \oplus M_1$ as an $S_\alpha$ module. Note that a weight 0 of $h_\alpha$ could not occur in $M_1$, as the weights there are the $c\alpha(h_\alpha)$ for $c \neq 0$. So we have a total of 2 occurrences of 0, exhausted by Ker $\alpha$ and $S_\alpha$ and so the even weights are only in Ker $\alpha$ and $S_\alpha$ and these are respectively 0 and -2,0,+2. In particular then for $c = 2$ we cannot have $2\alpha$ as a root (as it will give weight 4, not accounted for). So if twice a root is not a root, then $1/2\alpha$ cannot be a root either (as else 2.1/2$\alpha$ won’t be). So 1 is not a root for $h_\alpha$ and so there are only even weights, namely the 0 and -2,0,+2, so $\text{ad}(\alpha)$ is also an $S_\alpha$ submodule. But $L_\alpha \subset M$, so $L_\alpha = L_\alpha \cap (\text{Ker} \alpha + S_\alpha) = L_\alpha \cap S_\alpha = \text{span}\{x_\alpha\}$ as a vector space. In particular dim $L_\alpha = 1$. Thus we proved the following theorem.

**Theorem 15.** For any $\alpha \in \Phi$ we have that dim $L_\alpha = 1$ and the only roots multiples of $\alpha$ are $\pm \alpha$.

We will show now some more facts revealing the structure of the root space.

**Theorem 16.** Let $\alpha, \beta \in \Phi$, then

1. $\beta(h_\alpha) \in \mathbb{Z}$ and $\beta - \beta(h_\alpha)\alpha \in \Phi$, the numbers $\beta(h_\alpha)$ are called the **Cartan integers**.

2. if also $\alpha + \beta \in \Phi$, then $[L_\alpha L_\beta] = L_{\alpha + \beta}$.

3. if $\beta \neq \pm \alpha$, then if $r$ and $q$ are the largest integers for which $\beta - \alpha r, \beta + q \alpha$ are roots, then all $\beta + q \alpha \in \Phi$ for $p = -r, -r + 1, \ldots, q - 1, q$ and $\beta(h_\alpha) = r - q$.

4. $L$ is generated as a Lie algebra by the root spaces $L_\alpha$.

**Proof.** Consider the action of $S_\alpha$ on $L_\beta$ for $\beta \neq \pm \alpha$. Let $K = \oplus_{i \in \mathbb{Z}} L_{\beta + ia}$. We have that $S_\alpha$ acts on $K$ as $\text{ad}(x_\alpha)(L_{\beta + ia}) \subset L_{\beta + (i+1)a}$, similarly for $y_\alpha$ and $h_\alpha$ fixes them. We also have that since $i\alpha$ is not a root unless $i = 0, \pm 1$ and so $\beta \neq -i\alpha$ being a root, i.e. $\beta + i\alpha \neq 0$. Anyway, $K$ is an $S_\alpha$ submodule of $L$, so its weights must be integral. On the other hand, the weights are given by the action of $h_\alpha$ on the one dimensional (from the previous theorem) spaces $L_{\beta + ia}$, if these are spanned by $z_{\beta + ia}$ respectively, then $\text{ad} h_\alpha(z_{\beta + ia}) = (\beta + i\alpha)(h_\alpha)(z_{\beta + ia})$, so the weights are $\beta(h_\alpha) + i\alpha(h_\alpha) = \beta(h_\alpha) + 2i$, in particular $\beta(h_\alpha) \in \mathbb{Z}$. We also see that these integers have at most one of 0 or 1 among them (once), so $K$ must be irreducible by theorem 8. By the same theorem then if $\beta(h_\alpha) + 2q$ and $\beta(h_\alpha) - 2r$ are the highest, resp lowest weights, all $\beta(h_\alpha) + 2p$ in between must occur. These are weights if and only if $L_{\beta + pa} \neq 0$, i.e. we have $\beta + pa$ is a root for $p = -r, \ldots, q$, i.e. they form a string called the **$\alpha-$string through $\beta$**. Again by the same theorem the weights should be symmetric about 0, so $\beta(h_\alpha) - 2r = -(\beta(h_\alpha) + 2q)$, i.e. $\beta(h_\alpha) = r - q$. In particular then since $-r \leq -r + q \leq q$, the weight $\beta + (q - r)\alpha = \beta - \beta(h_\alpha)\alpha$ also appears. Last, if $\alpha + \beta \in \Phi$, then since $\text{ad}(x_\alpha)(L_\beta) \subset L_{\alpha + \beta}$ and since we showed that since $\alpha + \beta$ is a root then $L_{\alpha + \beta}$ it is a summand of $K$ - irreducible, so $x_\alpha(L_\beta) \neq 0$ as there is no other way to reach $L_{\alpha + \beta}$ through $L_\beta$ by the action of $S_\alpha$. Since $L_{\alpha + \beta}$ is one dimensional then we must have $[L_\alpha L_\beta] = L_{\alpha + \beta}$. 

Finally, let $\alpha_1, \ldots, \alpha_l \in \Phi$ be a basis for $H^*$. For any $\beta \in \Phi$, we can write it as $\beta = \sum_i c_i \alpha_i$ and we want to see what the $c_i$s look like. We can obtain an inner product on $H^*$ from the Killing form, namely if $\delta, \gamma \in H^*$, define $(\gamma, \delta) = \kappa(t_\delta, t_\gamma)$. We have $(\gamma, \gamma) = \kappa(t_\gamma, t_\gamma) = \gamma(t_\gamma) \neq 0$ for $\gamma \in \Phi$ as we already proved and since $\Phi$ spans $H^*$, then we have $(u, u) \neq 0$ for all $u \neq 0 \in H^*$, i.e. this inner product is nondegenerate. This inner product, apart from arising naturally from the Killing form, is useful because by theorem 16, part (1) we have $\gamma(h_\delta) \in \mathbb{Z}$, so by definition of $h_\delta$ this amount to $\gamma(\frac{2t_\delta}{\kappa(t_\delta)}) = 2\gamma(t_\delta) = 2\kappa(t_\delta, t_\delta) = 2\kappa(t_\delta, t_\delta) \in \mathbb{Z}$. So now using this inner product we can write $(\beta, \alpha_j) = \sum_i c_i (\alpha_i, \alpha_j)$ and we see that if we consider this as a system with variables $c_i$, the coefficients are all in $\mathbb{Q}$, solving this system using simple linear algebra we see that the solutions are rational functions in the coefficients, so $c_i \in \mathbb{Q}$. 

18
Moreover, we will show that \((\alpha, \beta) \in \mathbb{Q}\). From linear algebra we have that \(\text{Tr}(AB) = \sum C \text{Tr}(AC)\text{Tr}(CB)\) (\(C\) span the space of matrices, e.g. \(E_{ij}\)), so in particular we have \(\kappa(t_\gamma, t_\delta) = \text{Tr}(\text{ad} t_\gamma \text{ad} t_\delta) = \sum_\alpha \kappa(t_\gamma, t_\alpha)\kappa(t_\delta, t_\alpha)\), in particular then \((\beta, \beta) = \sum_\alpha (\beta, \alpha)^2\). We know already \(\frac{(\beta, \alpha)}{(\beta, \beta)} \in \mathbb{Q}\), so dividing both sides by \((\beta, \beta)^2\) we get that \(\frac{1}{(\beta, \beta)} = \sum_\alpha \frac{(\alpha, \beta)^2}{(\beta, \beta)^2} \in \mathbb{Q}\), so \((\beta, \beta) \in \mathbb{Q}\) and then also \((\alpha, \beta) \in \mathbb{Q}\). The last equality also shows that \((\beta, \beta) > 0\), i.e. our inner product is positive definite. Thus we have shown the following theorem.

**Theorem 17.** If \(\alpha_1, \ldots, \alpha_t\) span \(\Phi\), then every \(\beta \in \Phi\) can be written as \(\sum_i c_i \alpha_i\) with \(c_i \in \mathbb{Q}\) and the inner product on \(H^*\) defined as the extension of

\[(\gamma, \delta) = \kappa(t_\gamma, t_\delta)\]

is positive definite and is rational restricted to \(\Phi\), i.e. \((\beta, \alpha) \in \mathbb{Q}\) for all \(\alpha, \beta \in \Phi\).

### 4.2 Roots and weights; the Weyl group

Let \(\rho : L \to \mathfrak{gl}(V)\) be a representation of \(L\). We know that \(\rho(H)\) would still be an abelian subalgebra of semisimple matrices, i.e. we can write a decomposition \(V = \oplus V_\alpha\) where \(\alpha \in H^*\) and \(V_\alpha = \{v \in V| h.v = \alpha(h)v\text{ for all } h \in H\}\). We will drop the \(\rho\) as it will be clear what we mean by \(h.v = \rho(h).v\). These eigenvalues \(\alpha\) are called **weights**, so for example the weights of the adjoint representation are the roots.

If now \(\beta\) is a root of \(L\) we can consider the action of \(L_\beta\) over \(V_\alpha\). Let \(y \in L_\beta\) and \(v \in V_\alpha\). Then by our fundamental computations we have \(h.(y.v) = y.(h.v) + [h, y].v = y.(\alpha(h)v) + \beta(h).y.v = (\alpha + \beta)(h).y.v\) for any \(h \in H\), so \(L_\beta : V_\alpha \to V_{\alpha + \beta}\). In particular, if \(V\) is irreducible then the weights must be congruent to one another modulo the root lattice \(\Lambda_R = \Lambda_\Phi\) as we showed in the beginning of the Root space decomposition section.  

We showed that for any roots \(\beta\) and \(\alpha\) that \(\beta(h_\alpha) = 0\) from theorem 16. Now suppose \(\beta\) is a weight of \(V\), we claim that the same is true. It follows in the same way we proved it for the roots. Again consider the subalgebra \(S_\alpha \simeq \mathfrak{sl}(2, F)\) and its action of \(V_\beta\), let \(K = \oplus V_{\beta + i\alpha}\) and assume \(\beta \not\equiv m\alpha\) for any \(m \in \mathbb{Z}\), then none of the \(\beta + i\alpha\) would be 0. Then \(K\) is a \(S_\alpha\) submodule of \(V\) and again by what we know about the representations of \(\mathfrak{sl}(2, F)\), we get an \(\alpha\)–string through \(\beta\) and the weights must be integral. The weights are given by the action of \(h_\alpha\) on \(V_{\beta + i\alpha}\), i.e. \(\iota\alpha(h_\alpha) + \beta(h_\alpha)\), \(\alpha\) has still the same meaning, i.e. it is a root, so \(\alpha(h_\alpha) = 2\) (by definition of \(h_\alpha\)) and so \(\beta(h_\alpha) \in \mathbb{Z}\). So we have just shown a lemma.

**Lemma 5.** All weights of all representations assume integer values on the \(h_\alpha\)s.

Now let \(\Lambda_W\) be the set of linear functionals on \(H\), which are integer valued on the \(h_\alpha\)s. Then \(\Lambda_W\) is a lattice and all weights will lie in it, in particular the roots also.

We already noticed that the set of roots is invariant under addition, so we are going to explore further what group action keeps the weights invariant.

**Definition 12.** The Weyl group \(W\) is the group generated by the involutions \(W_\alpha\) given by

\[W_\alpha(\beta) = \beta - \beta(h_\alpha)\alpha = \beta - 2\frac{(\beta, \alpha)}{(\alpha, \alpha)}\alpha,\]

where the inner product is the extension of the one given by \((\alpha, \beta) = \kappa(t_\alpha, t_\beta)\) when \(\alpha, \beta\) are roots.

The \(W_\alpha\) is in fact a reflection and the plane it fixes is given by \(\beta(h_\alpha) = 0\), i.e. it is \(\Omega_{\alpha} = \{\beta \in H^*| \beta(h_\alpha) = 0\}\). With respect to the inner product then we will have for every \(\beta \in \Omega_{\alpha}\) that \((\beta, \alpha) = \kappa(t_\alpha, t_\beta)\) proportional to \(\kappa(t_\beta, h_\alpha) = \beta(h_\alpha) = 0\), so \(\Omega_\alpha \perp \alpha\). We also have \(W_\alpha = \alpha - 2\alpha = -\alpha\), so the \(W_\alpha\)s are indeed reflections. We can say then that the Weyl group is generated by the reflections in the planes perpendicular to the roots.

As a consequence of theorem 16 we have that \(W_\alpha(\Phi) \subset \Phi\) for all \(\alpha \in \Phi\), so

**Lemma 6.** The set of weights of a representation is invariant under the Weyl group.

---

9We will denote the same object - set of roots by both \(R\) and \(\Phi\) as the notations in [2] and [1] differ.
Again from theorem 16, we have that $\beta - r\alpha, \ldots, \beta + q\alpha$ appear, if we let $\gamma = \beta - r\alpha$, then the weights will be $\gamma, \ldots, \gamma + (q + r)\alpha$ and under $h_\alpha$ the corresponding weights in $\mathfrak{sl}(2, F)$ will be $\gamma(h_\alpha), \ldots, \gamma(h_\alpha) + 2(p + q)$, which should be symmetric about 0, so we have $\gamma(h_\alpha) = p + q$, i.e. the length of the $\alpha$--string is equal to the value(99,1444),(877,1546) of the first weight evaluated at $h_\alpha$.

We also have that the multiplicities of the weights are invariant under the Weyl group, again restricting our analysis to any string of weights, i.e. to the action of $S_\alpha$ and applying theorem 8.

In our study of $\mathfrak{sl}(2, F)$ we started off with choosing the "farthest" nonzero $V_\lambda$, but this was possible since the integers have natural ordering. We would like to apply similar analysis and for this purpose we need to introduce an ordering on the roots.

Now let $l : \mathbb{R}_\Phi \to \mathbb{R}$ be any linear functional and let $R^+ = \{\alpha \in \Phi | l(\alpha) > 0\}$ be the set of positive roots with respect to $l$ and $R^- = \{\beta \in \Phi | l(\beta) < 0\}$ - the negative ones. Note that we can choose $l$ so that no root lies on its kernel. So we have $\Phi = R^+ \cup R^-$. Note that the positivity/negativity depends on the choice of functional.

**Definition 13.** Let $V$ be a representation of $L$. A nonzero vector $v \in V$, which is both an eigenvector of the action of $H$ and is in the kernel of all $L_\alpha$ for $\alpha \in R^+$ (for some choice of a functional $l : H^* \to \mathbb{R}$), is called a highest weight vector of $V$. If $V$ is irreducible then we call the weight $\alpha$ the highest weight of the representation.

We will prove now one of the main facts.

**Theorem 18.** Let $L$ be a semisimple Lie algebra. Then

1. Every finite dimensional representation possesses a highest weight vector.

2. The subspace $W$ of $V$ generated by the images of a highest weight vector $v$ under successive applications of the root spaces $L_\beta$ for $\beta \in R^-$ is an irreducible subrepresentation.

3. Given the choice of $l$, i.e. fixed positive roots, the highest weight vector of an irreducible representation is unique up to scalar.

**Proof.** Let $\beta$ be the weight of $V$ which maximizes $l$. Let then $v \in V_\beta$. If $\gamma \in R^+$ is a positive root, then $l(\gamma + \beta) = l(\beta) + l(\gamma) > l(\beta)$, so by our choice $\gamma + \beta$ is not a weight, i.e. $V_{\gamma + \beta} = 0$. But $L_\gamma V_\beta \subset V_{\gamma + \beta} = 0$, so $L_\gamma$ kills $v$ for all positive roots $\gamma$, proving the first part.

We will show part (2) by induction on the number of applications of negative root spaces. Namely, let $W_n$ be the subspace spanned by all $w_n, v$, such that $w_n$ is a word of length at most $n$ of elements of negative root spaces. Let $x \in L_\alpha$ for $\alpha \in R^+$ and consider any $w_n, v$. We have that $w_n = y.w_{n-1}$, where $w_{n-1} \in W_{n-1}$ and $y \in L_\beta$, for $\beta \in R^-$. Then we have $x.(w_n, v) = x.y.(w_{n-1}, v) = y.x.(w_{n-1}, v) + [x, y](w_{n-1}, v)$. By induction $x.w_{n-1}.v \in W_{n-1}$, so $y.x.w_{n-1}.v \in W_n$. Also we know that if $\beta \neq -\alpha$, then $[x, y] = 0$, so we are done in this case. In the case $\beta = -\alpha$, we have $[x, y] = ut_\alpha$ from theorem 14, where $u$ is some scalar. But then $[x, y] \in H$, i.e. it fixes the weight spaces, so $x.y.(w_{n-1}, v) \in W_{n-1}$. Finally we have that $x(w_n, v) \in W_n$. This shows that the space $W = \sum_n W_n$, generated by successive applications of only negative roots to $v$, is fixed by positive roots. As it is also fixed by negative roots, we have that it is a $L$--submodule of $V$. It is clearly irreducible since it is a space generated by $L.v$.

From here it follows in particular that if $V$ is irreducible then $\dim V_\alpha = 1$. If otherwise there were two vectors $v, w \in V_\alpha$ not scalar multiples of each other, then we cannot possibly obtain $w$ from $v$ by applying only negative roots, as $L_{\beta_1}(L_{\beta_2} \ldots v) \subset V_{\alpha + \beta_1 + \ldots} \neq V_\alpha$ fore $l(\alpha + \beta_1 + \ldots) < l(\alpha)$. So $w$ is not in the irreducible subrepresentation generated by $v$ as in part (2) and so we must have $\dim V_\alpha = 1$. Uniqueness then follows from the uniqueness of $\alpha$.

As a corollary to this theorem, or more precisely - to the proof of part (2), we get exercise 14.15 from [?] by considering the adjoint representation. If $\alpha_1, \ldots, \alpha_k$ are roots of $L$, then by induction we can show that the subalgebra $L'$ generated by $H, L_{\alpha_1}, \ldots, L_{\alpha_k}$ is in fact $H \oplus (\oplus L_\alpha) = L''$ for all $\alpha$ which are in $\mathbb{N}\{\alpha_1, \ldots, \alpha_k\} \cap \Phi$. The containment $L' \subset L''$ follows similarly to the proof of theorem 18 part (2) by
induction on the number of summands $\alpha_i$, or it is simply obvious and any $L_{\alpha_i}$ sends $L''$ into itself. The containment $L'' \subset L'$ follows from theorem 16 since $[L_\alpha L_\beta] = L_{\alpha + \beta}$ if $\alpha + \beta$ is a root.

In view of this result we can make the following definition.

**Definition 14.** We call a root $\alpha$ simple if it cannot be expressed as a sum of two positive or two negative roots.

We see that every negative root is a sum of simple negative roots and in view of the preceding paragraph the subalgebra generated by $H$ and $L_\beta$ for $\beta$ running over the negative simple roots is in fact the direct sum of $H \oplus (\oplus_{\beta \in R^-} L_\beta)$. So we can rephrase the last theorem by replacing negative roots with primitive negative roots and get that any irreducible representation is generated as a vector space by the images of a highest weight vector under successive applications of negative simple roots.

We are now going to prove a theorem that ties together the Weyl group and weights of irreducible representations.

**Theorem 19.** Let $V$ be an irreducible representation and $\alpha$ a highest weight. The vertices of the convex hull of the weights of $V$ are conjugate to $\alpha$ under the Weyl group. The set of weights of $V$ are exactly the weights that are congruent to $\alpha$ modulo the root lattice $\Lambda_R$ and that lie in the convex hull of the images of $\alpha$ under the Weyl group $W$.

**Proof.** In view of our observations so far we have that since $V$ is generated by successive applications of negative roots to the highest weight vector $v$, then all the weights appearing in $V$ are in fact of the form $\alpha + \beta_1 + \beta_2 + \ldots$ for $\beta_i \in R^-$, so all weight lie in the positive cone $\alpha + C^-$, where $C^-$ is the positive cone spanned by $R^-$. Note that by our definitions, $C^-$ is entirely contained in one half-plane (the one where $l$ is negative).

Since $\alpha$ is a weight, we showed that if $\alpha + \beta$ is also a weight, then so must be the whole string $\alpha, \alpha + \beta, \ldots, \alpha - \alpha(h_\beta)\beta$. Since $\alpha$ is maximal, it is necessarily the case that $\beta \in R^-$. It also follows that $\alpha - \alpha(h_\beta)\beta = W_\beta(\alpha)$ is a weight. Any vertex of the convex hull of the weights of $V$ will lie on a line passing through $\alpha$ and $\alpha + \beta$ for some $\beta \in R^-$. We then must necessarily have that this vertex is $W_\beta(\alpha)$ as we showed earlier (it is the other end of the string, geometrically - line, containing weights). As we also know the weights of the form $\gamma + n\delta$ for $\gamma$-weight and $\delta \in R$ must form a string, then for any two weights of $V$ the intersection of the line segment connecting them and $\alpha + \Lambda_R$ (i.e. all possible weights) must be the segment itself. So the set of weights is convex, i.e. it is the intersection of $\Lambda_R$ and its convex hull.

Note that if $\alpha - \alpha(h_\beta)\beta$ is weight then its ordering must be "lower" than $\alpha$, i.e. we must have $l(\alpha) > l(\alpha - \alpha(h_\beta)\beta) = l(\alpha) - \alpha(h_\beta)l(\beta)$ and since $l(\beta) < 0$ we must have $\alpha(h_\beta) < 0$ or alternatively $\alpha(-h_\beta) < 0$ i.e. $\alpha(h_\beta) > 0$ where now by the symmetry of roots $-\beta$ runs over $R^+$.

**Definition 15.** A Weyl chamber is the locus $W$ of points in $H^*$, such that $\gamma(h_\beta) \geq 0$ for all $\beta \in R^+$. Alternatively, it is a connected region (polytope) with faces the planes $\Omega_\alpha$.

The Weyl group acts transitively on the Weyl chambers. The importance of the Weyl chamber is revealed in the following theorem.

**Theorem 20.** Let $W$ be the Weyl chamber associated to some ordering of the roots. Let $\Lambda_W$ be the weight lattice. Then for any $\alpha \in W \cap \Lambda_W$, there exists a unique irreducible finite-dimensional representation $\Gamma_\alpha$ of $L$ with highest weight $\alpha$.

This theorem gives a bijection between the points in $W \cap \Lambda_W$ and the irreducible representation. In view of the previous theorem we have that the weights of $\Gamma_\alpha$ will be the weights in the polytope whose vertices are the images of $\alpha$ under the Weyl group $W$. Note moreover that these are all the irreducible representations - for every irreducible representation its highest weight will lie in the Weyl chamber.

**Proof.** We will show uniqueness first. If $V$ and $W$ are two irreducible finite-dimensional representations of $L$ with highest weight vectors $v$ and $w$ of weight $\alpha$, then consider the vector $(v, w) \in V \oplus W$. Its weight is again
α and in general the weight space decomposition is still \( \oplus(V_\beta \oplus W_\beta) \), so the weights are the same and α is again highest (note that the weights are the same because they are uniquely determined by α from the previous theorem). Let \( U \subset V \oplus W \) be the irreducible subrepresentation generated by \( (v, w) \) and consider the nonzero projection maps \( \pi_1 : U \to V \) and \( \pi_2 : U \to W \), nonzero since they map \( (v, w) \) to \( v \) and \( w \) respectively. Then \( \exists \pi_i \) is a nonzero submodule of the irreducible \( V \) (or \( W \)) and \( \ker \pi_i \) a submodule of \( U \), so we must have both \( \pi_i \) surjective and injective, i.e. isomorphism, showing in particular that \( \pi_2 \circ \pi_1^{-1} : V \to W \) is an isomorphism and so the uniqueness is proved.

Existence is much harder and involves concepts we have not introduced. We will try to describe it here though. It involves the concept of a universal enveloping algebra, that is an associative algebra \( \mathfrak{g} \) over \( F \), such that there is a linear map \( i : L \to \mathfrak{g} \), such that \( i([x, y]) = i(x)i(y) - i(y)i(x) \), and which has the universal property of every other such algebra arising from a homomorphism with \( \mathfrak{g} \). Such algebra can be explicitly constructed as \( \mathfrak{g}(L) = T(L)/J \), where \( T(L) = \bigoplus_1^\infty T^iL \) is the tensor algebra and \( J \) is the ideal generated by \( x \otimes y - y \otimes x - [x, y] \) for \( x, y \in L \). Now we can construct a so called standard cyclic module \( Z(\alpha) \) of weight \( \alpha \) as follows. If we choose nonzero \( x = e_{ij} \in L \) for \( \beta \in R^+ \) and let \( I(\alpha) \) be the left ideal of \( \mathfrak{g}(L) \) generated by all these \( e_{ij} \) and all \( h_i - \alpha(h_i) \) for all \( i \in \Phi = R \). The choice of such ideal makes sense as its elements annihilate a highest weight vector of weight \( \alpha \), so we can choose \( Z(\alpha) = \mathfrak{g}(L)/I(\alpha) \), where the coset of \( 1 \) corresponds to the line through our highest weight vector \( v \), i.e. \( Z(\alpha) \cong \mathfrak{g}v \), where \( v \) is our highest weight vector. Next by some facts about standard cyclic modules \( Z(\alpha) \) would have a unique maximal submodule \( Y \) and the \( V(\alpha) = Z(\alpha)/Y \) will be our irreducible representation of weight \( \alpha \). To summarize, the point was to construct a module of highest weight \( \alpha \) explicitly and take its irreducible submodule containing the highest weight vector we started with.

Just like with the definition of simple roots, we can define the fundamental weights \( \omega_1, \ldots, \omega_n \) - the weights such that any weight in the Weyl chamber can be expressed uniquely as nonnegative integral combination of these fundamental weights. These are in fact the first weights along the edges of the Weyl chamber, i.e. \( \omega_i(h_{ij}) = \delta_{ij} \), where \( \omega_i \)s are the simple roots. Then every weight \( \alpha \) in the Weyl chamber (also called dominant weight) is an integral nonnegative sum \( \alpha = \sum a_i \omega_i \) and then we can write the corresponding irreducible representation as \( \Gamma_{a_1, \ldots, a_n} = \Gamma_{a_1 \omega_1 + \cdots + a_n \omega_n} \).

5 The special linear Lie algebra \( \mathfrak{sl}(n, \mathbb{C}) \) and the general linear group \( GL_n(\mathbb{C}) \)

It is time now that we apply all the developed theory to our special case of interest - the linear group and corresponding Lie algebra. We will first reveal the structure of \( L = (n, \mathbb{C}) \), i.e. find its Cartan subalgebra \( H \), roots and \( L_\Phi \) and then we will proceed to finding the weights, the Weyl chamber and constructing the irreducible representations. We will also see at the end how all this helps in finding the representations of \( GL_n(\mathbb{C}) \).

5.1 Structure of \( \mathfrak{sl}(n, \mathbb{C}) \)

We will first find (a) maximal toral subalgebra of \( L = \mathfrak{sl}(n, \mathbb{C}) \). The obvious choice would be the to consider the set of all diagonal matrices. It is clearly a toral subalgebra, the only question is whether it is maximal. Suppose that \( H \) is not maximal and \( H \subsetneq T \), for some other toral subalgebra. Then since \( T \) still must be abelian by what we showed in the beginning of the section on root space decomposition, so we have that for every \( t \in T \) and \( h \in H \), \( [h, t] = 0 \), i.e. \( T \subset C_L(H) \). The question then is to show that for our particular choice the matrices that commute with any diagonal matrix of trace 0 are only the diagonal matrices, i.e. \( H \subsetneq C_L(H) = H \), so \( T = H \). Well, let \( e_{ij} \) be the matrix with only nonzero entry at \( (i, j) \) and this entry is equal to 1. If \( h = \text{diag}(a_1, \ldots, a_n) \), then \( [h, e_{ij}] = a_i e_{ij} - a_j e_{ij} = (a_i - a_j)e_{ij} \). Let \( x = \sum x_{ij} e_{ij} \), i.e. \( x \) is a matrix with entries \( x_{ij} \). Then \( [x, h] = \sum x_{ij} (a_i - a_j)x_{ij} e_{ij} \), and \( x \in C_L(H) \) iff \( (a_i - a_j)x_{ij} = 0 \) for all \( i, j \) and all \( a_1, \ldots, a_n \) such that \( a_1 + \cdots + a_n = 0 \). So fixing \( x \) for every \( i \neq j \) we can take \( a_j = 1 \), \( a_j = -1 \) and \( a_k = 0 \) otherwise and force \( x_{ij} = 0 \), showing that \( x \) must be diagonal.
Now that we have our maximal toral subalgebra \( H \) we need to find the root spaces \( L_\alpha \) and for that we first need to find the roots. We just saw that if \( h \in H \), i.e. \( h = \text{diag}(a_1, \ldots, a_n) \), then \( \{ h_{\alpha ij} \} = (a_i - a_j) e_{ij} \), so the matrices \( e_{ij} \) are eigenvectors for \( H \). If we then let \( \alpha_{ij} : H \to \mathbb{C} \) be given by \( \alpha_{ij}(\text{diag}(a_1, \ldots, a_n)) = a_i - a_j \), we see that \( L_{\alpha_{ij}} = \{ x \in L | [h] = \alpha_{ij}(h)x \text{ for all } h \in H \} \) for \( i \neq j \) is nonempty as it contains \( \mathbb{C} e_{ij} \). In fact counting dimensions, the nilpotent part of \( L \), i.e. the complementary to \( H \) in the Cartan decomposition, is spanned as a vector space by the \( n(n-1) \) vectors \( e_{ij} \) for \( i \neq j \) and on the other hand it contains the \( n(n-1) \) disjoint nonempty spaces \( L_{\alpha_{ij}} \), so we must have

\[
L = H \oplus (\oplus_{i \neq j} L_{\alpha_{ij}})
\]

and clearly \( \alpha_{ij} \) are the roots of \( L = \mathfrak{sl}(n, \mathbb{C}) \). Either from theorems 14 and 16 or simply from dimension count again, since \( \dim H = n-1 \) (codimension 1 in \( \mathbb{C}^n \)) and so \( \dim(\oplus_{i \neq j} L_{\alpha_{ij}}) \leq \dim L - \dim H = n^2 - 1 - (n-1) = n(n-1) \), we must have \( \dim L_{\alpha_{ij}} = 1 \), so \( L_{\alpha_{ij}} = \mathbb{C} e_{ij} \).

For the sake of simplicity, define the functionals on the space of diagonal matrices \( L_{\text{diag}}(\text{diag}(b_1, \ldots, b_n)) = b_i \), then we see \( \alpha_{ij} = l_i - l_j \) and so we express the roots as pairwise differences of \( n \) functionals. On the space \( H \) since the sum of diagonal entries is 0, these \( n \) linear functionals are not independent, but we have \( l_1 + \cdots + l_n = 0 \). This is the only relation they need to satisfy, so we can picture them in \( n-1 \)-dimensional space.

In order to draw a realistic picture we need to determine the inner product, i.e. the Killing form on \( \mathfrak{sl}(n, \mathbb{C}) \). We will determine it from the definition as \( \kappa(x, y) = Tr(ad x ad y) \).

Since \( \kappa \) is linear we can determine the values \( \kappa(e_{ii}, e_{jj}) \) and extend them over \( H \) by linearity. We have that

\[
\kappa(e_{ii}, e_{jj}) = \kappa(\text{ad} e_{ii}, \text{ad} e_{jj}) \quad \text{and for any basis element of } \mathfrak{sl}_n \text{ we have } \kappa(e_{ii}, \text{ad} e_{jj}(e_{pr})) = [e_{ii}, [e_{jj}, e_{pr}]] = [e_{ii}, (\delta_{pj} - \delta_{rq})e_{pr}] = (\delta_{pj} - \delta_{rq})[e_{ii}, e_{pr}] = (\delta_{pj} - \delta_{rq})(\delta_{pi} - \delta_{ri})e_{pr}. \quad \text{So } e_{pr} \text{ are eigenvectors and by dimension count are all possible eigenvectors, so}
\]

\[
Tr(\text{ad} e_{ii}, \text{ad} e_{jj}) = \sum_p \sum_r (\delta_{pj} - \delta_{rq})(\delta_{pi} - \delta_{ri}) = \sum_p \sum_r (\delta_{pi} - \delta_{ri})^2 = \sum_{p \neq i} 1 + \sum_r (1 - \delta_{ri})^2 = 2(n-1) \quad \text{for } i = j
\]

\[
\kappa(h, g) = \sum_{i,j} a_i b_j \kappa(e_{ii}, e_{jj}) = \sum_{i \neq j} (-2)a_i b_j + \sum_{i = j} 2(n-1)a_i b_j = \sum_{i,j} -2a_i b_j + \sum_{j} 2na_i b_i = -2(\sum_{j} a_j)(\sum_{j} b_j) + 2n \sum_{i} a_i b_i = 2n \sum_{i} a_i b_i.
\]

Having figured out \( \kappa \) we need to find the inner product on \( H^* \) given by \( \kappa(\alpha, \beta) = \kappa(t_\alpha, t_\beta) \) and for that we need to find \( t_\alpha \). For \( h = \text{diag}(a_1, \ldots, a_n) \in H \) we have \( l_i(h) = a_i \), and if \( l_i = \text{diag}(t_1, \ldots, t_n) \), we must have by definition of \( t_\alpha \) that \( a_i = l_i(h) = \kappa(t_\alpha, h) = 2n \sum k a_k \) and that for all \( a_i = 0 \) with \( i \) sum \( 0 \), so the conditions are just enough to determine \( t_i \) uniquely as \( \frac{1}{2n} e_i \). Then we can say that the inner product on \( H^* \) is determined by \( \langle l_i, l_j \rangle = \kappa(t_i, t_j) = \frac{1}{2n} \delta_{ij} \). In particular the roots given by \( \alpha_{ij} = l_i - l_j \) will have inner products \( \langle \alpha_{ij}, \alpha_{kl} \rangle = \frac{1}{2n} (\delta_{ik} - \delta_{il} - \delta_{jk} + \delta_{jl}) \). And in general the inner product on \( H^* = \mathbb{C} \{ l_1, \ldots, l_n \} / (l_1 + \cdots + l_n) = \{ a_1 l_1 + \cdots + a_n l_n | a_1 + \cdots + a_n = 0 \} \) is given by

\[
\langle \sum a_i l_i, \sum b_i l_i \rangle = \frac{1}{2n} \langle \sum a_i b_i \rangle.
\]

We see that our inner product makes the \( l_i \) orthogonal, if we want to picture the root lattice then we can consider the space \( \mathbb{C}^n \) with unit vectors given by the \( l_i \)'s (e.g. the usual coordinates \( e_i = (0, \ldots, 1, 0, \ldots) \) with 1 at position \( i \)). Then \( H^* \) will be the plane given by \( \sum x_i = 0 \) and the roots \( \alpha_{ij} = l_i - l_j \) are the vertices of a certain kind of polytope (if \( n = 3 \) this is a planar hexagon for example), called very creatively the root polytope. The root lattice then will simply be

\[
\Lambda_R = \{ \sum a_i l_i | a_i \in \mathbb{Z}, \sum a_i = 0 \} / (\sum l_i = 0),
\]

where the quotient is just to emphasize we are in \( H^* \), i.e. the plane defined by \( \sum l_i = 0 \).
Now we will look at the weight lattice $\Lambda_\mathcal{W}$. It is the set of linear functionals $\beta \in H^*$ which assume integer values on $h_{\alpha_i}$. So we need to find the $h_{\alpha_i}$'s. But we already showed that $h_j = \frac{2a_j}{\alpha_j} = \frac{2a_j}{\alpha_j}$ and we found $t_{n_j} = e_{ii} - e_{jj}$, so $c_{ij}(t_{n_j}) = 1 - (-1) = 2$ and so $h_{n_j} = t_{n_j} = e_{ii} - e_{jj}$. Then if $\beta = \sum b_i l_i \in H^*$ is a weight, we must have that $\beta(h_{\alpha_i}) = \beta(e_{ii} - e_{jj}) = b_i - b_j \in \mathbb{Z}$, so all coefficients are congruent one another modulo $\mathbb{Z}$, so since we have $\sum l_i = 0$ we can assume they are in $\mathbb{Z}$, so

$$\Lambda_\mathcal{W} = \mathbb{Z}\{l_1, \ldots, l_n\}/(\sum l_i = 0).$$

We can now determine the Weyl group also, it is given by reflections $W_{\alpha_i}(\beta) = \beta - \beta(h_{\alpha_i})\alpha = \sum b_i l_k - (b_i - b_j)(l_i - l_j) = b_k l_k + (b_i - b_j) l_i + (b_i - b_j) l_j = b_i l_i + \cdots + b_j l_j + \cdots$, i.e. it exchanges $l_i$ with $l_j$, fixing everything else. So we have that $\mathfrak{W} \approx S_n$ - the symmetric group on the $n$ elements $l_i$.

We can now pick a direction, divide the roots into positive and negative and find a Weyl chamber. For $\beta \in H^*$, i.e. $\beta = \sum a_i l_i$ a linear functional will be determined by its values on $l_i$ (plus the condition that these values sum to 0), so if $\phi : H^* \rightarrow \mathbb{R}$ is our linear functional with $\phi(l_i) = c_i$, we have $\phi(\beta) = \sum b_i c_i$ and $\sum c_i = 0$. On the roots we have $\phi(\alpha_{ij}) = c_i - c_j$, which suggest a natural ordering given that $c_1 > c_2 > \cdots > c_n$, then $\phi(\alpha_{ij}) = c_i - c_j > 0$ iff $i < j$. So $R^+ = \{l_i - l_j, i < j\}$ and $R^- = \{l_i - l_j, i > j\}$. We can easily determine the simple roots also, as for any $i$ we have $l_i - l_{i+1} = (l_i - l_{i-1}) + (l_{i-1} - l_{i+1}) + \cdots + (l_{j-1} - l_j)$ is a sum of positive roots if $i + 1 < j$. So the simple (primitive) positive roots must be among $l_i - l_{i+1}$ and it is easy to see that these cannot be expressed as sums of other positive roots.

The Weyl chamber associated to this ordering will be given by the $\beta \in H^*$, such that $\beta(h_{n_j}) \geq 0$ for every $\gamma \in R^+$, i.e. if $\beta = \sum b_i l_i$, then we will have $\beta(h_{\alpha_i}) = b_i - b_j \geq 0$ for all $i < j$, i.e. then the Weyl chamber is $\mathcal{W} = \{\sum b_i l_i | b_i \geq b_j, \sum b_i \geq \sum b_n\}$. So we can see how choosing a different order will get us a disjoint and geometrically equal Weyl chamber, as we expect by the action of $\mathfrak{W}$. If we let $a_i = b_i - b_{i+1}$ with $a_n = b_n$, we see that $\beta \in \mathcal{W}$ is equivalent to $a_i \geq 0$ for all $i$. Also $b_i = \sum_{j \geq i} a_i$ and so we have that $\beta = \sum_i (\sum_{j \geq i} a_i l_i) = \sum_j a_j (\sum_{i \leq j} l_i)$, so in particular the Weyl chamber is the cone between the rays given by $l_1, l_1 + l_2, \ldots, l_1 + \cdots + l_{n-1}$. The faces of $\mathcal{W}$ are always given by the hyperplanes $\Omega_{\alpha_{ij}}$ in this case $\Omega_{\alpha_{ij}} = \{\sum b_i l_i | b_i = b_j\}$, perpendicular to the roots $\alpha_{ij}$, but we see that given our ordering $b_1 \geq \cdots \geq b_n$ in $\Omega_{\alpha_{ij}} \cap \mathcal{W}$ we must have $b_i = b_{i+1} = \cdots$. This means that the codimension 1 faces of $\mathcal{W}$ must be the $\Omega_{\alpha_{i,i+1}} = \{\sum b_i l_i | b_i = b_{i+1}\}$, the ones perpendicular to the simple roots.

We see here explicitly what the fundamental weights are, namely $\omega_i = l_1 + \cdots + l_i$, which are the first on the edges of the Weyl chamber. Then the irreducible representations will correspond to any $(n-1)$-tuple of nonnegative integers $(a_1, \ldots, a_{n-1})$, corresponding to the weight $\alpha = \sum a_i \omega_i = (a_1 + \cdots + a_{n-1})l_1 + \cdots a_{n-1} l_{n-1}$, $\Gamma_{a_1, \ldots, a_{n-1}}$.

### 5.2 Representations of $\mathfrak{sl}(n, \mathbb{C})$

We saw that $\alpha = (a_1 + \cdots + a_{n-1})l_1 + \cdots a_{n-1} l_{n-1}$ is in the Weyl chamber $\mathcal{W}$ and for $a_i \in \mathbb{N}$ it is also clearly in $\Lambda_\mathcal{W}$, so it is a weight. We want to show that the representation corresponding to it, $\Gamma_{\alpha}$, does in fact exist.

Let $V = \mathbb{C}^n$ be the standard representation of $\mathfrak{sl}(n, \mathbb{C})$. What are its weights? For $h = \text{diag}(b_1, \ldots, b_n)$ with $b_i = 0$, its eigenvalues as a matrix are simply its diagonal entries $b_1, \ldots, b_n$. So if $e_i$ are the usual basis vectors of $V$, then $h.e_i = b_i.e_i$, so for every $h$ we have $h.e_i = l_i(h).e_i$, so $e_i \in V_l$ - weight space for weight $l_i$ and clearly $V = \oplus_l V_l$. This shows that the weights of this representation are the $l_i$s. It is clearly irreducible, for example by checking that the matrix with 1s above the main diagonal and 1 at the lower left corner and all other entries 0 sends $e_n$ into $e_{n-1}$, then $e_{n-2}$, $e_1$ to $e_n$, i.e. acts transitively on the basis vectors.

Now consider $\wedge^i V$, the $i$-th exterior power of $V$. Its elements are $e_1 \wedge \cdots \wedge e_i$ and the action of an element $u \in L$ is given by the usual action on a tensor product. We have $u(e_1 \wedge \cdots \wedge e_i) = \sum_j u(e_j) \wedge \cdots \wedge e_i$, with the wedge product we just replace $\otimes$ with $\wedge$. Now $\wedge^i V$ has basis induced from the tensor product basis of $i$-tuples of unordered basis vectors, i.e. $e_{j_1} \wedge \cdots \wedge e_{j_i}$, such that $j_1 < j_2 \cdots < j_i$. Consider again the action of $h = \text{diag}(b_1, \ldots, b_n)$ on a basis vector, we have $h(e_{j_1} \wedge \cdots \wedge e_{j_i}) = \sum_k e_{j_k} \wedge \cdots \wedge (b_j.e_{j_k}) \wedge \cdots e_{j_i} = (\sum_k b_{j_k})e_{j_1} \wedge \cdots \wedge e_{j_i}$. So we see that $h(e_{j_1} \wedge \cdots \wedge e_{j_i}) = (l_{j_1} + \cdots + l_{j_i})(h)e_{j_1} \wedge \cdots \wedge e_{j_i}$ and so every basis vector of $\wedge^i V$
spans a one-dimensional weight space $V_{l_1 + \cdots + l_i}$. The weights of this representation are given by all possible sums of $i$ different $l_j$s, i.e. from $l_1 + \cdots + l_i$ to $l_{n-1} + \cdots + l_n$. In particular we see that every weight is in the image of the Weyl group acting on $l_1 + \cdots + l_i$, namely if $l_1 + \cdots + l_i$ is a weight then if $\sigma$ is the permutation sending $\sigma(k) = j_k$ for $k = 1, \ldots, i$ and everything else in what remains, then since $\sigma \in S_n \cong \mathfrak{W}$ we have $\sigma(l_1 + \cdots + l_i) = l_{j_1} + \cdots + l_{j_i}$ as the elements of the Weyl group are all possible transpositions $l_r \leftrightarrow l_p$. So since we saw how $l_1 + \cdots + l_i \in \mathfrak{W}$, the Weyl chamber, we then have that $\bigwedge^i V = \Gamma_{l_1 + \cdots + l_i}$ as follows. First of all, $l_1 + \cdots + l_i$ is certainly the highest weight of $\bigwedge^i V$ and if it had any subrepresentations, then one of them would have to account for the highest weight and so be $\Gamma_{l_1 + \cdots + l_i}$. But then $\Gamma_{l_1 + \cdots + l_i}$ would also account for the weights that are images of $l_1 + \cdots + l_i$ under $\mathfrak{W}$, so all the weights of $\bigwedge^i V$, counting multiplicities as these are all 1, and so it must be equal to $\bigwedge^i V$.

By theorem 20 we must have an irreducible representation $\Gamma_\alpha$ for every $\alpha = (a_1 + \cdots + a_{n-1})l_1 + \cdots + a_{n-1}l_{n-1}$, where $a_i \in \mathbb{N}$. If we rewrite $\alpha$ as $\alpha = a_1l_1 + a_2(l_1 + l_2) + \cdots + a_{n-1}(l_1 + \cdots + l_{n-1})$, this can suggest where to look for the desired representation in view of the preceding paragraph. Namely, since $\bigwedge^i V$ has highest weight $l_1 + \cdots + l_i$, and a highest weight vector $v$, then $\text{Sym}^k \bigwedge^i V$ will have highest weight $(l_1 + \cdots + l_i)$ and highest weight vector $v_1 \cdots v$. We will show this explicitly. For $I = \{j_1, \ldots, j_k\}$, let $e_I = e_{j_1} \wedge \cdots \wedge e_{j_k}$, and

$$l_I = \sum_{p \in I} l_p,$$

then $\text{Sym}^k \bigwedge^i V$ is spanned by the commutative products $e_{l_1} \cdots e_{l_k}$ and again by the action of a Lie algebra on tensor products we have that for $h \in H$, $h(e_{l_1} \cdots e_{l_k}) = \sum_{r=1}^k l_r e_{l_1} \cdots e_{l_{r-1}} e_{l_r} e_{l_{r+1}} \cdots e_{l_k}$: So each of these elements $e_{l_1} \cdots e_{l_k}$ span a one dimensional weight space for $\text{Sym}^k \bigwedge^i V$ for the weight $\sum_{r} \sum_{p \in I_r} l_p$ and we readily see by the arrangement of weights $l_1 > l_2$ as long as $i < j$ that the maximal weight is $k.l_1 + \cdots + l_i$ (by maximizing each $l_r$ to be $\{1, \ldots, i\}$), keeping in mind that $l_I$ should contain $i$ different numbers. Now since $\text{Sym}^k \bigwedge^i V$ has maximal weight $k(l_1 + \cdots + l_i)$, it must also contain the unique irreducible representation $\Gamma_{k(l_1 + \cdots + l_i)}$ (i.e. it must contain an irreducible representation generated by a highest weight vector for the given highest weight and since there is only one irreducible representation of a given highest weight, it must be our $\Gamma$).

Now the last step is to show that the representation $S = \text{Sym}^{a_1} V \otimes \text{Sym}^{a_2} \bigwedge^2 V \otimes \cdots \otimes \text{Sym}^{a_{n-1}} (\bigwedge^{n-1} V)$ has highest weight $\alpha = a_1l_1 + a_2(l_1 + l_2) + \cdots + a_{n-1}(l_1 + \cdots + l_{n-1})$ and by the discussion at the end of the preceding paragraph we will necessarily have that it contains the irreducible representation $\Gamma_\alpha$. But this actually follows by similar reasoning as in the previous paragraph. Again our representation $S$ is generated by all tensor products $v_1 \otimes \cdots v_{n-1}$, where $v_1$ run through the generators of $\text{Sym}^{a_1} \bigwedge^i V$ as in the preceding paragraph. By it we have $h(v_1) = m_i(h)v_1$, where if $v_1 = e_{l_1} \cdots e_{l_i}$ then $m_i = \sum_{r=1}^k l_r$. So $h(v_1 \otimes \cdots \otimes v_{n-1}) = \sum_{p=1}^{n-1} v_1 \otimes \cdots \otimes h(v_p) \otimes \cdots \otimes v_{n-1} = (\sum_{m=1}^l m_p)(h)(v_1 \otimes \cdots \otimes v_{n-1})$ and the weights of $S$ must be exactly these sums $\sum_{p=1}^{n-1} m_p$. In order to maximize this sum, since the $m_p$s are independent, we maximize each $m_p$, which by the preceding paragraph gives us exactly $a_p(l_1 + \cdots + l_p)$, so at the end the maximal weight is $\sum_{p=1}^{n-1} a_p(l_1 + \cdots + l_p) = \alpha$, which is what we needed to show. Since the $a$s run through all possible maximal weights we have exhausted all possible irreducible representations and so we’ve proven the following theorem.

**Theorem 21.** The irreducible representations of $\mathfrak{sl}(n, \mathbb{C})$ are the $\Gamma_{a_1, a_2, \ldots, a_{n-1}}$ of highest weight $(a_1 + \cdots + a_{n-1})l_1 + \cdots + a_{n-1}l_{n-1}$ and it appears as a subrepresentation of $\text{Sym}^{a_1} V \otimes \text{Sym}^{a_2} \bigwedge^2 V \otimes \cdots \otimes \text{Sym}^{a_{n-1}} (\bigwedge^{n-1} V)$.

### 5.3 Weyl’s construction, tensor products and some combinatorics

Our goal here is to exhibit the irreducible representations $\Gamma_{a_1, \ldots, a_{n-1}}$ as explicitly as possible as subspaces of the $d$–th tensor power $V^\otimes d$, where $d$ is the dimension of $\text{Sym}^{a_1} V \otimes \text{Sym}^{a_2} \bigwedge^2 V \otimes \cdots \otimes \text{Sym}^{a_{n-1}} (\bigwedge^{n-1} V) \subset V^\otimes d$, i.e. $d = \sum a_i$.

Observe that since $\mathfrak{sl}(n, \mathbb{C})$ acts on $V$, and $S_d$ acts on $V^\otimes d$ by permuting coordinates, then the two actions commute. Moreover, we know from both [1] and [3] what the representations of $S_d$ are, in particular they are nicely indexed by partitions $\lambda \vdash d$, with $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n$, and in fact so are the $\Gamma_{a_1, \ldots, a_{n-1}}$ if we write the partition of $d$ as $(\sum_{i=1}^{n-1} a_i, \sum_{i=2}^{n} a_i, \ldots, a_{n-1}, 0)$ in bijection with any sequence of nonnegative integers.
We can then use the representations of $S_n$ to study the representations of $\mathfrak{sl}(n, \mathbb{C})$ and for that we will digress a little bit in reviewing representations of $S_n$.

A standard way of studying representations of $S_n$ is as shown in [1] the use of Young symmetrizers. Consider the group algebra $\mathbb{C}S_d$ and let $T$ be a Young tableau of shape $\lambda$, where $\lambda \vdash d$. Then let $P_\lambda = \{ g \in S_d : g \text{ preserves each row of } T \}$ and $Q_\lambda = \{ g \in S_d : g \text{ preserves each column of } T \}$. The choice of $T$ does not matter as long as we are consistent (use the same one) as the sets would be conjugate. Now let $a_\lambda = \sum_{g \in P_\lambda} e_g$ and $b_\lambda = \sum_{g \in Q_\lambda} \text{sgn}(g) e_g$, where $e_g$ is the element in the group algebra corresponding to $g$.

The define the Young symmetrizer $c_\lambda = a_\lambda b_\lambda \in \mathbb{C}S_d$. The main theorem in the representation theory of $S_d$ states that $c_\lambda^2 = n_\lambda c_\lambda$ for $n_\lambda \in \mathbb{N}$ and the image of $c_\lambda$ by right multiplication on $\mathbb{C}S_d$ is an irreducible representation $V_\lambda$ and every irreducible representation is obtained in this way.

Going back to $V \otimes^d$ we have the action of $S_d$ on it given by $(v_1 \otimes \cdots \otimes v_d) \cdot \sigma = (v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(d)})$ for $\sigma \in S_d$ and so we can consider the image of $c_\lambda$ on $V \otimes^d$ and denote it by $\Sigma V = \text{Im}(c_\lambda |_{V \otimes^d}) = V \otimes^d c_\lambda = V \otimes^d \otimes \mathbb{C}S_d V_\lambda$, which will be a subrepresentation of $V \otimes^d$ for $GL_n$ and hence for $SL_n$ and taking the differential of the representation map we obtain a corresponding irreducible representation of $\mathfrak{sl}(n, \mathbb{C})$.

**Theorem 22.** The representations $\Sigma V(\mathbb{C}^n)$ is the irreducible representation of $\mathfrak{sl}(n, \mathbb{C})$ with highest weight $\lambda_1 l_1 + \cdots + \lambda_n l_n$. It is in fact the representation $\Gamma_{\lambda_1, \ldots, \lambda_n}$ given by $a_i = \lambda_i - \lambda_i - 1$.

**Proof.** We need to see how the representation of $SL_n$ relates to a corresponding representation of $\mathfrak{sl}(n, \mathbb{C})$. There is no natural map from a Lie group to its Lie algebra, but there is certainly a map from a Lie algebra to a Lie group, the exponential. If $\rho : SL(V) \to SL(W)$ is a representation, the induced representation will be $d\rho : \mathfrak{sl}(V) \to \mathfrak{sl}(W)$ and the following diagram will commute

\[
\begin{array}{ccc}
\mathfrak{sl}(V) & \xrightarrow{d\rho} & \mathfrak{sl}(W) \\
\exp \downarrow & & \exp \downarrow \\
SL(V) & \xrightarrow{\rho} & SL(W)
\end{array}
\]

So in particular if $V' = \oplus V'_i$ is a weight space decomposition, i.e. for every $w \in V'_i$ and $h \in H$ we have $\rho(h) w = \alpha(h) w$, then by commutativity we will have that $\rho(\exp(h)) = \exp((d\rho)(h))$. Since we are dealing with matrices we have $(d\rho)(h) = \frac{d}{dt} \rho(\exp(th)) \big|_{t=0} = \frac{d}{dt} (\exp((d\rho)(h))^t) = e^{(d\rho)(h)} w = e^{\alpha(h) w}$, so its eigenvalues are $e^{\alpha(h)}$ with the same eigenvectors. Therefore we have that if $A \in SL(V)$ with $A = \exp(h)$, then $Tr(\rho(A)) = \sum_{\alpha \in W} e^{\alpha(h)}$.

Now if $h = \text{diag}(y_1, \ldots, y_n) \in \mathfrak{sl}(n, \mathbb{C})$ then $A = \text{diag}(e^{y_1}, \ldots, e^{y_n})$, then we have for $\beta = \text{diag}(b_1, \ldots, b_n) \in W$ (i.e. a weight of $V$, so $b_i \in \mathbb{Z}$), we have $\beta(h) = b_1 y_1 + \cdots + b_n y_n$, so

\[
Tr(A) = \sum_{\beta \in W} e^{\beta(h)} = \sum_{\beta \in W} e^{b_1 y_1 + \cdots + b_n y_n} = \sum_{\beta \in W} (e^{y_1})^{b_1} (e^{y_2})^{b_2} \cdots (e^{y_n})^{b_n}.
\]

Now if we assume for a moment some of the representation theory of $GL_n$, which will be shown later in appendix A, we can use a theorem from the theory of Schur functors giving us a formula for the trace of $A$ when $V' = \Sigma V$. The theorem states that for any $A \in GL(V)$, the trace of $A$ on $\Sigma V$, which is now a representation of $GL(V)$, is the value of the Schur polynomials (see ??) on the eigenvalues $x_1, \ldots, x_n$ of $A$ on $V$, i.e. $Tr_{\Sigma V}(A) = s_\lambda(x_1, \ldots, x_n)$.

We know from ?? that $s_\lambda = m_\lambda + \sum_{\mu < \lambda} K_{\lambda \mu} m_\mu$, where $m_\mu$ are the monomial symmetric functions, i.e. $m_\mu = \sum_{i_1, i_2, \ldots, i_n} \text{distinct } x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$, and $K_{\lambda \mu}$ are the Kostka coefficients, whose combinatorial representation is given by the number of semi-standard Young tableaux of shape $\lambda$ and type $\mu$. So for any diagonal matrix $A \in GL(V)$, which we can always write as $A = \text{diag}(e^{y_1}, \ldots, e^{y_n})$, we have that

\[
Tr(A) = s_\lambda(e^{y_1}, \ldots, e^{y_n}) = m_\lambda(e^{y_1}, \ldots, e^{y_n}) + \sum_{\mu < \lambda} m_\mu(e^{y_1}, \ldots, e^{y_n}).
\]
Now the last equation (12) can be expanded as a sum of monomials in \(e^{y_1}, \ldots, e^{y_n}\), but so is the formula (11). As these must be equal for all \(y_1, \ldots, y_n\), we must have that the monomials coincide, so the weights \(\beta \in W\) must be of the kind \(\mu_1 l_1 + \cdots + \mu_n l_n\) for some permutation \(i_1, \ldots, i_n\) of \((\mu_1, \ldots, \mu_n)\) for some \(\mu \leq \lambda\) and must occur with multiplicity \(K_{\lambda \beta}\). In particular we have that the weight \(\lambda_1 l_1 + \cdots + \lambda_n l_n\) has multiplicity exactly one and is highest among all appearing weights\(^{10}\). By the uniqueness of irreducible representations with given highest weight it follows that the representation \(S_\lambda(C^n)\) is the irreducible representation of highest weight \(\lambda_1 l_1 + \cdots + \lambda_n l_n\).

As a consequence to this theorem we are enabled thanks to combinatorics rules to determine the decomposition of any tensor power into irreducible representations. Specifically, if \(V_1, \ldots, V_k\) are representations, then the trace of a matrix \(A \in GL_n\) on \(V_1 \otimes V_2 \otimes \cdots \otimes V_k\) is \(Tr|_{V_1 \otimes \cdots \otimes V_k}(A) = Tr|_{V_1}(A) \cdots Tr|_{V_k}(A)\). We now know that the traces are polynomials (necessarily symmetric!) in the eigenvalues and the traces of the irreducible representations are the Schur polynomials. So if \(f(x_1, \ldots, x_n)\) is the trace of our representation of interest, if we write \(f(x_1, \ldots, x_n) = \sum \lambda c_\lambda s_\lambda(x_1, \ldots, x_n)\), then we will have that \(c_\lambda\) is the multiplicity of the irreducible representation \(S_\lambda(C^n)\).

An important example of this is the decomposition of \(S_\mu(C^n) \otimes S_\nu(C^n)\) into irreducible representations. The trace polynomial will be \(s_\mu(x_1, \ldots, x_n)s_\nu(x_1, \ldots, x_n)\) and we know, e.g. by [3], that \(s_\mu s_\nu = \sum \lambda c_{\mu \nu}^\lambda s_\lambda\), where the coefficients \(c_{\mu \nu}^\lambda\) are called the Littlewood-Richardson coefficients and since they necessarily are equal to the multiplicity of an irreducible representation they must be nonnegative integers. They can also be computed by the Littlewood-Richardson rule, saying that \(c_{\mu \nu}^\lambda\) is the number of skew SSYTs of shape \(\lambda\), \(\mu\), \(\nu\), and such that their reverse reading word is a lattice permutation.

5.4 Representations of \(GL_n(C^n)\)

We certainly have that the \(S_\lambda(C^n)\)'s are irreducible representations of \(GL_n\) as shown in the appendix, the question is whether these are all the irreducible representations.

The commutative diagram (??), which comes from the theory we developed about maps between Lie groups and their Lie algebras, shows that when the Lie group is simply connected then there is one-to-one correspondence between the irreducible representations of the Lie group and the irreps of its Lie algebra. One can show this by arguing as follows: if \(V_1 \subset V\) is a subrepresentation of the group / algebra then the same is true for the algebra/group by the commutativity of exp and the representation map.

Since \(SL_n(C)\) is simply connected we have such one-to-one correspondence with the irreducible representations \(\Gamma_{a_1, \ldots, a_{n-1}}\) of \(\mathfrak{s}(n, C)\). \(GL_n\) however is not simply connected and \(\mathfrak{gl}_n\) is not even semisimple, so we will try to build the irreducible representations from the special groups.

Denote \(C^n - V\) and let \(D_k\) be the one-dimensional representation of \(GL_nC\) given by the \(k\)th power of the determinant, so if \(k \geq 0\) we have \(D_k = (\wedge^k V) \otimes k\). Given a representation of \(SL_n\) we can lift it to \(CL_n\) by tensoring with \(D_k\). If for any \(a = (a_1, \ldots, a_n)\) we denote by \(\Phi_a\) the subrepresentation of \(\text{Sym}^{a_1} V \otimes \cdots \otimes \text{Sym}^{a_n}(\wedge^k V)\), spanned by \(v = (e_1)^{a_1}(e_1 \wedge e_2)^{a_2} \cdots (e_1 \wedge \cdots \wedge e_n)^{a_n}\), we see that passing to the Lie algebra corresponds to the irrep of highest weight \(a_1 l_1 + \cdots + (a_1 + \cdots + a_n) l_n\) and in particular restricting to \(SL_n\) gives \(\Gamma_{(a_1, \ldots, a_{n-1})}\). Or we can set \(\lambda = (a_1 + \cdots + a_n, a_2 + \cdots + a_n, \ldots, a_n)\) and consider \(S_\lambda V\) as a representation of \(\Psi_\lambda\) of \(GL_n\). We see that in the first case we have \(\Phi_{a_1, \ldots, a_{n-1} + k} = \Phi_{a_1, \ldots, a_n} \otimes D_k\) and in the second \(P \Sigma_{a_1 + k, \ldots, a_n + k} = \Psi_\lambda \otimes D_k\). Moreover these two representations are actually isomorphic, as their restrictions to \(SL_n\) agree and their restriction to the center \(C^* \subset GL_n\) are both \(D_{a_n}\), so agree everywhere.

We will show finally that these are all the irreducible representations of \(GL_n\).

Theorem 23. Every irreducible complex representation of \(GL_nC\) is isomorphic to \(\Psi_\lambda\) for a unique index \(\lambda = (\lambda_1, \ldots, \lambda_n)\) with \(\lambda_1 \geq \lambda_2 \cdots \geq \lambda_n\).

Proof. Consider the corresponding Lie algebras. We have that \(\mathfrak{gl}_nC = \mathfrak{sl}_nC \times C\), where \(C\) is the ideal formed by matrices \(aI\) with \(a \in C\) and \(I\) - the identity matrix. In particular \(C\) is the radical of \(\mathfrak{gl}_n\) and \(n\) is

\(^{10}\)Follows by the ordering of the weights \(l_i > l_{i+1}\) and by the fact that any other weight appearing is of the form \(\mu_1 l_1 + \cdots + \mu_n l_n\) with \(\mu < \lambda\)
the semisimple part. The following reasoning shows that every irreducible representation of \( \mathfrak{gl}_n \) is a tensor product of an irreducible representation of \( \mathfrak{sl}_n \) and a one-dimensional representation.

Let \( U \) be an irreducible representation of \( \mathfrak{gl}_n \). We have that by Lie’s theorem since \( \text{Rad}(\mathfrak{gl}_n) = \mathbb{C} \) is solvable and since \( \text{Rad}(\mathfrak{gl}_n) \subset \mathfrak{gl}(U) \) then \( U \) contains a common eigenvector for all elements in \( \text{Rad}(\mathfrak{gl}_n) \) and so there is a linear functional \( \alpha \), such that \( W = \{ w \in U | x.w = \alpha(x)w \text{ for every } x \in \text{Rad}(L) \} \) is nonempty. But then since every element of \( \mathfrak{gl}_n \) can be written as \( x + y \) with \( x \in \text{Rad}(\mathfrak{gl}_n) \) and \( y \in \mathfrak{sl}_n \) we have that for \( w \in \mathfrak{gl} \) \( x.y.w = x.y + x[y]w = \alpha(x)\alpha(y)w \) since in our case \( \text{Rad}(\mathfrak{gl}_n) \) commutes with every element. Then we must have \( y.w \in \mathfrak{sl} \), so \( \mathfrak{gl}_n \rightarrow \mathfrak{sl}_n \) and hence \( W \) is a subrepresentation of \( U \), so must be all of \( U \). Now if we extend \( \alpha \) to \( \mathfrak{sl}_n \) and let \( D \) be the one-dimensional representation given by \( \alpha \), i.e. for every \( z \in D \) and \( x \in \mathfrak{gl}_n \) we have \( x.z = \alpha(x)z \) and so \( U \otimes D^* \) is trivial on \( \text{Rad}(\mathfrak{gl}_n) \) and so is a necessarily irreducible representation of \( \mathfrak{sl}_n \).

So if \( W_\lambda = \mathcal{S}_\lambda(\mathbb{C}^n) \) is the irreducible representation of \( \mathfrak{sl}_n(\mathbb{C}) \), given by the partition \( \lambda_1, \ldots, \lambda_n \), we can extend it to \( \mathfrak{gl}_n = \mathfrak{sl}_n \times \mathbb{C} \) by acting trivially on the second factor. By the above consideration we have that every irreducible representation of \( \mathfrak{gl}_n \) is isomorphic to \( W_\lambda \otimes L(w) \), where \( L(w) \) is the one dimensional representation given by multiplication by \( w \in \mathbb{C} \) on the second factor of \( \mathfrak{sl}_n \mathbb{C} \times \mathbb{C} \).

So passing back to Lie groups, the same will be true for the simply connected \( SL_n \mathbb{C} \times \mathbb{C} \) (the simply-connectedness we take for granted here). We now need to pass from \( SL_n \mathbb{C} \times \mathbb{C} \) to \( GL_n \mathbb{C} \) and for that there is a natural map \( \rho : SL_n \mathbb{C} \times \mathbb{C} \rightarrow GL_n \mathbb{C} \) given by \( \rho(g, z) = e^{z}g \cdot e^{-z} \). The kernel of this map is generated by \( e^{z}I \times (-s) \) and for \( e^{z}I \in SL_n \mathbb{C} \) we need to have \( s = 2\pi i/n \). So \( SL_n \mathbb{C} \times \mathbb{C} = GL_n \mathbb{C} \times \ker \rho \). Now an irreducible representation of \( GL_n \mathbb{C} \) can be lifted to an irreducible representation on \( SL_n \mathbb{C} \) by letting it be 0 on \( \ker \rho \), so actually this defines a one-to-one correspondence of irreducible representations and we need to find which of \( W_\lambda \times L(w) \) are 0 on \( \ker \rho \). We have that \( e^{z}.I \) acts on \( \mathcal{S}_\lambda \) by multiplication by \( e^{zd} \), where \( d = \sum \lambda_i \), since that’s how it acts on \( V^\otimes d \). \(-s \) acts on \( L(W) \) by multiplication by \( e^{-sw} \), so the action on \( \ker \rho \) is given by multiplication by \( e^{zd-sw} \), which is trivial iff \( sd - sw \in \mathbb{Z} \) and since \( s = 2\pi i/n \), this is equivalent to \( w = \sum \lambda_i + kn \) for \( k \in \mathbb{Z} \).

Then we have that the irreducible representations are \( W_\lambda \otimes L(\sum \lambda_i + kn) \), which restricted to \( GL_n \mathbb{C} \) is \( \mathcal{S}_{\lambda_1 + k, \ldots, \lambda_n + k} \) and the pullback of \( \Psi \) is \( W_\lambda \otimes L(\sum \lambda_i + kn) \), so we have exhausted all irreducible representations of \( GL_n \mathbb{C} \) and so are done.

### 5.5 Some explicit constructions, generators of \( \mathcal{S}_\lambda V \)

For a sequence of nonnegative integers \( a = (a_1, \ldots, a_k) \), let \( A^a = \text{Sym}^{a_1} V \otimes \cdots \otimes \text{Sym}^{a_k} V \), as we saw the irreducible representation \( \Gamma_a \subset A^a. \) Here we will describe \( A^a \) and exhibit a basis for \( \mathcal{S}_\lambda \).

A way to describe \( A^a \) is to look at the right action of the symmetric group on \( V^\otimes d \) and note that on one hand we have

\[
\mathcal{S}_\lambda(V) = V^\otimes d \subset A^a \subset \bigotimes^k V \otimes \cdots \otimes (V)^{a_1} = V^\otimes d b_\lambda.
\]

These inclusions suggest that we can probably find a \( a' \in \mathbb{C} S_d \), such that \( A^{a'} = V^\otimes d . a b_\lambda \). Finding this \( a' \) is not that hard at all if we note that \( [k, k, k-1, \ldots, k-1, 1, \ldots, 1] \) is in fact the conjugate partition of \( \lambda' \), i.e. it is the list of lengths of the columns of \( \lambda \). The wedges \( \bigwedge^i V \) are just \( V^\otimes d b_{i(1)} \), where \( b_{i(1)} = \sum_{\sigma} \text{permut} \text{e} \text{d} \text{ column of length } i \text{ } \text{sgn}(\sigma) \text{ } \text{sgn} \text{ } \text{a} \text{ } \text{column of length } l \). On the other hand \( \text{Sym}^a(\bigwedge^i V) \) is actually taking all these columns of length \( i \) and symmetrizing them, i.e. permuting them in on all possible ways, i.e. \( \bigwedge^i V)^{a_1, a_2, \ldots, a_k} \), where \( (a_1, a_2, \ldots, a_k) \) is the \( \lambda_a \) for \( \lambda = a_i \) (don’t confuse the two different kinds of \( a \)’s!). So we’d like over the columns of different lengths gives us that \( A^a V = (V^\otimes k b_1) \otimes a_2 a_3 \otimes \cdots \otimes (V)^{a_1} a_2 a_3 \), where \( a' = \sum_{r \in S_{a_k} \times \cdots \times S_{a_1}} e_r \), is the sum over the permutations \( r \) which permute the columns of same length between each other, i.e. the group \( S_{a_k} \times \cdots \times S_{a_1} \). Note also that \( a' \) and \( b_\lambda \) commute for the same reason that \( a_2 \) and \( b_\lambda \) do. It is also clear that \( a_2 = a_2 ' a' \), where \( a'' = \sum_{r \in (S_{a_1} \times \cdots \times S_{a_k})/(S_{a_k} \times \cdots \times S_{a_1})} e_r \) over the coset representatives.

28
Set $A^*(V) = \bigoplus_a A^a(V) = \operatorname{Sym}^*(V \oplus A^2 V \oplus \cdots \oplus A^n V)$, which makes a graded ring out of the $a^aV$s. Then we can define a graded ring $S^*$ as the quotient of $A^*(V)$ by the two-sided ideal $I^*$ generated for all $p \geq q \geq r \geq 1$ by all elements of the form (called the "Plucker relations") $(v_1 \wedge \cdots \wedge v_p), (v_1 \wedge \cdots \wedge v_q) - \sum (v_1 \wedge \cdots \wedge v_1 \wedge \cdots \wedge w_r \wedge \cdots \wedge v_q), (v_1, \wedge v_2, \wedge \cdots \wedge v_r, \wedge w_{r+1} \wedge \cdots \wedge w_q)$, where the sum is over all $1 \leq i_1 < i_2 < \cdots < i_p \leq p$ and the $w_1, \ldots, w_r$ are inserted in the places of $v_1, \ldots, v_q$.

We are now going to exhibit the coordinates of $S^*$ in $A^*$. Let $e_1, \ldots, e_n$ be a basis for $V$ and let $T$ be a semistandard Young tableaux of shape $\lambda$ filled with the integers $1, \ldots, n$, i.e. the rows are nondecreasing and the columns strictly increasing. Let $T(i, j)$ denote the entry of $T$ in row $i$ and column $j$ and define

$$e_T = \prod_{j=1}^i e_{T(1, j)} \wedge e_{T(2, j)} \wedge \cdots \wedge e_{T(\lambda', j)};$$

i.e. we wedge together the entries in columns and multiply the results in $S^*$. We are going to prove the following theorem.

**Theorem 24.** The $e_T$ for $T$ a semistandard tableaux on $\lambda$ form a basis for $S^2(V)$. The projection from $A^a(V)$ to $S^a(V)$ maps the subspace $S_\lambda$ isomorphically onto $S^a(V)$.

**Proof.** We will first show that the elements $e_T$ generate $S^a$.

To see that the $e_T$ span $S^a$, note that clearly $S^a$ is spanned by all $e_T$ for all tableaux $T$ with strictly increasing columns (not just semistandard) as these span $A^a$ itself. Introduce a reverse lexicographic ordering on the $T$s by comparing the listing of their elements column by column, and within each column list top to bottom and we compare starting from the last entry going forward until we find two differing entries. Then for example we have that the tableaux $T_0$ which has $T(i, j) = i$ for all $i, j$ is the smallest one. Suppose now that a tableaux $T$ is not semistandard, i.e. there are two adjacent columns indexed $i$ and $i + 1$, for which there is a row $r$, such that $T(r, j) > T(r, j + 1)$ and let $r$ be the smallest such. Then let $v_1, \ldots, v_p$ be the column of index $j$ and $w_1, \ldots, w_q$ the one of index $j + 1$ (i.e. $v_i = T(i, j), w_i = T(i, j + 1)$). For the $r$ we just picked then we have by the Plucker relations in $I^*$ that $(v_1 \wedge \cdots \wedge v_p), (v_1 \wedge \cdots \wedge w_q) = \sum (v_1 \wedge \cdots \wedge v_1 \wedge \cdots \wedge w_r \wedge \cdots \wedge v_q), (v_1 \wedge \cdots \wedge v_q, \wedge w_{r+1} \wedge \cdots \wedge w_q)$. Multiplying the remaining columns of $T$ we get that $e_T = \sum e_T'$, where the $T'$ are obtained from $T$ by interchanging the first $r$ elements from column $j + 1$ of $T$ with some $r$ elements of column $j$ of $T$. We see then that the listing of entries of $T'$ is the same as the one for $T$ starting from the back until we encounter the position $(i, j + 1)$, where the entry of $T'$ will be $v_i = T(i, j) \geq T(r, j) > T(r, j + 1)$, so in the reverse lexicographic ordering we will have $T' > T$. In this way we see that by the Plucker relations $e_T = \sum e_T'$ with $T' > T$ as long as $T$ is not semistandard. Continuing this process iteratively we see that we can express $e_T$ for every non-semistandard tableaux $T$ with a sum of $e_T$ with $T' > T$ and so we will only be increasing in the lexicographic order. Since there are finitely many tableaux on $n$ elements of shape $\lambda$ we will reach a moment where $e_T$ is a sum (of sums of sums etc) of tableaux that must be semistandard (otherwise we can keep expressing as sums of larger ones). This shows that $e_T$ for $T$ -semistandard generate $S^a(V)$.

We need to show now that they are linearly independent. We will show this indirectly by showing that $S_\lambda(V) \subset S^a$. If the projection of $S_\lambda$ into $S^a$ is not zero, then their intersection is not 0 and since both $S_\lambda$ and $S^a$ are representations of $GL(V)$ with $S_\lambda$ irreducible we must have that $S_\lambda$ is a submodule of $S^a$. For every representation we have that its dimension is the trace (character) at the identity of the group, then in our particular case we have that $\dim S_\lambda = Tr(I) = s_\lambda(1, \ldots, 1)$ (from appendix A). On the other hand from the combinatorial interpretation of Schur functions ([3]) we have $s_\lambda(x_1, \ldots, x_n) = \sum_{T[\text{shape}(T) = \lambda]} x^{T^T}$, so

---

11 We have that $S^*$ is in fact the graded algebra of the $S_\lambda$'s. A way to see this is by noticing that $S_\lambda = (V^{S_\lambda})^\omega | S_\lambda$ is equivalent to averaging the elements of $A^*$ which are obtained by one another via a coset representative $p \in (S_\lambda \times \cdots \times S_\lambda)/(S_{\lambda_1} \times \cdots \times S_{\lambda_1})$. In particular the given relations restrict to the case of only two columns, represented by $(v_1 \wedge v_2)$ and $(w_1 \wedge w_2)$, in which we average out over all $r$ over all possible picks of rows $i_1 < \cdots < i_r$, for which we exchange the columns in rows $i_1, \ldots, i_r$ to get the first column to be $(v_1 \wedge \cdots \wedge w_1 \wedge \cdots \wedge w_r)$ and the second $(w_1 \wedge \cdots \wedge v_1 \wedge \cdots \wedge v_r)$. In the original relations we move elements from positions $(i_1, \ldots, i_r)$ to $(1, \ldots, r)$ two times, i.e. conjugate by a permutation preserving columns, and since we apply it twice we don’t change the sign (think of permutations appearing in $b_\lambda$). In any case we have that $A^*/I^* = S^*$. 29
$s_\lambda(1, \ldots, 1) = \sum_{T : \text{shape}(T) = \lambda} 1$ is the number of semistandard tableaux $T$ of shape $\lambda$ filled with the numbers $1, 2, \ldots, n$. But this is the same number of possible generators $e_T$ for $S^a$ - the semistandard tableaux $T$ of shape $\lambda$ and filled with $1, \ldots, n$, so counting dimensions we get $\# \{e_T\}$ generating $S^a$, i.e. $T$-semistandard $\geq \dim S^a \geq \dim S_\lambda = \# \{T \text{ semistandard}\}$, so all inequalities are equalities showing that the $e_T$-s are linearly independent and that $S^a \simeq S_\lambda$.

So we need to show only that the image of $S_\lambda$ in $S^\bullet$ is nonempty. We first of all have that each $e_T$ is in fact in $S_\lambda$ (one can show this by seeing that $e_T.e_\lambda = e_T$, for some scalar $c$). Next we show that at least one of them has a nonzero image in $S^\bullet$. We will check this for $e_{T_0}$ with $T_0(i, j) = i$ for all $i, j$. Now $T_0$ is definitely the smallest in the reverse lexicographic order and moreover, applying the Plucker relations to any two of its columns results in the same tableaux (we have that $w_1 = 1, \ldots, w_r = r$ and $v_i = i_1, \ldots, v_r = i_r$ and the wedge $v_1 \wedge \ldots w_1 \wedge \ldots \wedge v_p = 0$ for having two equal elements unless we replaced $w_1, \ldots, w_r$ with $v_1, \ldots, v_r$, which leads to the trivial relation $e_{T_0} - e_{T_0}$, in particular we cannot have it as a summand in any Plucker relation as reversing the exchange of $r$ elements between two columns of $T_0$ can’t result into anything but $T_0$. So $e_{T_0} \not\in I^\bullet$, showing that its image under the projection is not 0 and so that $e_{T_0} \in S_\lambda \cap S^a$.

\end{proof}

6 Conclusion

In this paper we developed the basic theory of Lie algebras, classifying them through properties like solvability, nilpotency, simplicity and showing criteria for establishing such properties. We studied the structure of semisimple Lie algebras via roots and root space decomposition and used that to study the irreducible representations of semisimple Lie algebras. We showed in particular a theorem that established a bijection between irreducible representations and highest weights (i.e. points in the intersection of the Weyl chamber with the root lattice), which gave us a nice way of listing all irreducible representations. We showed the theory in practice by studying the structure and representations of $\mathfrak{gl}(n, \mathbb{C})$. We established connections between Lie groups and Lie algebras, which enabled us in particular to limit the irreducible representations of the Lie group $GL(V)$ by the irreducible representations of its Lie algebra $\mathfrak{gl}(V)$, thereby proving that the irreducible representations $S_\lambda$ are in fact all the irreducible representations of $GL(V)$. We showed some explicit constructions in the intersection of combinatorics and algebra.

A Schur functors, Weyl’s construction and representations of $GL_n(\mathbb{C})$

Consider the right action of $S_d$ on $V^{\otimes d}$ for any vector space $V$ given by $(v_1 \otimes \cdots \otimes v_d).\sigma = v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(d)}$. This action commutes with the left action of $GL(V)$ acting coordinatewise on the tensor product. Let $\lambda$ be a partition of $d$ and let $c_\lambda$ be the Young symmetrizer in the group algebra of $S_d$ defined as $c_\lambda = a_\lambda b_\lambda$, where $a_\lambda = \sum_{\sigma \in P} e_{\sigma}$ and $b_\lambda = \sum_{\sigma \in Q} \text{sgn}(\sigma) e_{\sigma}$, where $P$ is the set of permutations fixing the rows of a Young tableaux of shape $\lambda$ filled with the elements $\{1, \ldots, d\}$ and $Q$ is the set of permutation fixing the columns of the same tableaux. Let $S_\lambda V = V^{\otimes d}.c_\lambda$ be the so called Schur functor (it is a functor on the set of vector spaces $V$ and linear maps between them). It is clearly a representation of $GL(V)$. We are going to prove the following theorem.

Theorem 25. For any $g \in GL(V)$ the trace of $g$ on $S_\lambda V$ is the value of the Schur polynomial on the eigenvalues $x_1, \ldots, x_n$ of $g$ on $V$, $Tr|_{S_\lambda V}(g) = s_\lambda(x_1, \ldots, x_n)$. Each $S_\lambda V$ is an irreducible representation of $GL(V)$.

In order to prove this theorem we will first prove some general lemmas. Let $G$ be any finite group, in our case its role will be played by $S_d$ and $A = \mathbb{C}G$ be its group algebra. For $U$ a right module over $A$, let $B = \text{Hom}_A(U, U) = \{\phi : U \to U | \phi(v, g) = \phi(v), g \text{ for all } v \in U \text{ and } g \in G\}$. If $U$ decomposes into a direct sum of irreducible modules via $U = \bigoplus U_i^{\oplus n_i}$, then since by Schur’s lemma a map between two irreducible modules is either 0 or scalar multiplication, we have that $\text{Hom}_A(U_i, U_j) = 0$ if $i \neq j$ and $\text{Hom}_A(U_i^{\oplus n_i}, U_i^{\oplus n_i}) = M_{n_i}(\mathbb{C}), n_i \times n_i$ matrices. If then $W$ is a left $A$ module, then the tensor product $U \otimes_A W$ is left $B$ module as for $\phi \in B$ and $a \in A$ we have $\phi.(u.a \otimes w) = \phi(u.a) \otimes w = \phi(u).a \otimes w = \phi(u) \otimes a.w$.

30
Lemma 7. Let $U$ be a finite-dimensional $A$–module. Then

(i) For any $c \in A$ the map $U \otimes_A Ac \to UC$ is an isomorphism of left $B$–modules.

(ii) If $W = Ac$ is an irreducible left $A$–module, then $U \otimes_A W = UC$ is an irreducible left $B$–module.

(iii) If $W_i = Ac_i$ are the distinct irreducible left $A$–modules, with $m_i = \dim W_i$, then 
$U \simeq \oplus_i (Uc_i)^{\otimes m_i}$, is the decomposition of $U$ into irreducible $B$–modules.

Proof. First of all, $U \otimes_A Ac \to UC$ is clearly surjective, as $UC$ is the image of $U \otimes_A Ac$. So we need to show it is injective. Now $Ac$ is a submodule of $A$ and $A$ is a representation of a finite group, so $A$ is completely reducible and in particular we have $A = Ac \oplus A'$, for some submodule $A'$. So $U = U \otimes_A A = U \otimes_A (Ac + A') = (U \otimes_A Ac) \oplus (U \otimes_A A')$. So if $\phi : U \otimes_A Ac \to UC$ is not injective, then so is the inclusion $U \otimes_A Ac \to UC \hookrightarrow U$, contradicting the fact that $U \otimes_A Ac$ is direct summand of $U$. This proves part (i).

For part (ii) suppose first that $U$ is irreducible. From the representation theory of finite groups we have that if $W_i$ are all the irreducible representations of $G$, then from one hand we have $\sum (\dim W_i)^2 = |G|$ and from the other if we map the group algebra into $End(\oplus_i W_i) = \oplus_i End(W_i) = \oplus_i M_{\dim W_i} \mathbb{C}$ since the elements of $G$ are automorphisms of $\oplus_i W_i$, by counting dimensions we must have that this inclusion map is an isomorphism. So we can view $A = \oplus M_{u_i}(\mathbb{C})$, where $u_i$ are the dimensions of its irreducible representations. Since $W = Ac$ is also a right $A$–ideal, it must be a minimal ideal of $A$ (as every ideal is a $A$–module), and so must be a minimal ideal in the matrix ring $M_{u_i}$ for some $i$. Since if a matrix that has only one nonzero column of index $j$, when multiplied on the left by any matrix, results again in a matrix with only one nonzero column of the same index, and since every matrix can be multiplied on the left by a certain matrix (of only nonzero entry at $(j,j)$) to get such a matrix of one nonzero column, we must have that minimal ideal consists of matrices with only nonzero column at some $j$. Similarly since $U$ is a right $A$–ideal, it must have only one nonzero row, say at $i$, for its inclusions in $M_{\dim W_k}$. Finally for every element of $U \otimes_A W$ its part in $M_{\dim W_i}$, is either 0 if one of $U$ of $W$ are there, or is of the form $C \otimes_A D$, for matrices $C$, $D$, we can write $C = IC$ and $C \in A$, so $C \otimes_A D = I \otimes C.D$. But for every matrix $C$ with only nonzero row $i$ and $D$ with only nonzero column $j$, we have that $C.D$ is matrix with only nonzero entry at $(i,j)$. This means that $U \otimes_A W$ is one dimensional and hence necessarily irreducible.

Now suppose $U = \oplus_i U_i \oplus (W_i)$, with $U - i$ irreducible, then $U \otimes_A W = \oplus_i (U_i \otimes W) \oplus m_i$. The $U_i$s being irreducible by the above paragraph must be in some $M_{dim W_i}$ and so that $U_i \otimes A W$ is not 0, we consider only the $U_i \subset M_{\dim W_i}$ for $W \subset M_{\dim W_k}$. However all irreducible modules of a matrix algebra must be isomorphic, so we have that finally $U \otimes_A W = (U_k \otimes_A W)^{\oplus m_k} = \mathbb{C}^{\oplus m_k}$, which is clearly irreducible over $B = \oplus M_{n_j}(\mathbb{C})$, proving part (ii).

For part (iii) we have that since $A = \oplus_i M_{u_i}$ (by for example the beginning of the second paragraph) then for $U$ as an $A$–module we have $U \simeq U \otimes_A A = U \otimes_A (\oplus_i W_i \oplus m_i) = \oplus_i (U \otimes_A W_i)^{\oplus m_i}$. Since by (ii) $U \otimes_A W_i$ is irreducible over $B$ and by (i) it is isomorphic to $Uc_i$, we get that $U \simeq \oplus_i (Uc_i)^{\oplus m_i}$, and by uniqueness of irreducible submodules decomposition, it must be the decomposition of $U$ into irreducible $B$–modules.

To prove theorem 25 we apply this lemma with $A = \mathbb{C}S_d$ and $U = V \otimes d$ viewed as a right $A$–module. Then since we know, by for example [1] or [3], what the irreducible representations of $S_d$ are, namely $V_{\lambda} = Ac_{\lambda}$ with $\lambda \vdash d$ and $c_{\lambda}$ - the Young symmetrizer, we can decompose $U \simeq U \otimes (Uc_{\lambda})$, as a left $B$–module. The question now is to show that the algebra $B$ is spanned by $End(V)$.

Lemma 8. The algebra $B$ is spanned as a linear subspace of $End(V \otimes d)$ by $End(V)$. A subspace of $V \otimes d$ is a $B$–submodule if and only if it is invariant under $GL(V)$.

Proof. One can show that $Sym^d W$ is spanned by $w^d = d!w \otimes \cdots \otimes w$ as $w$ runs through $W$. One can show this for example by induction on $d$, starting with the fact that $v \otimes w + w \otimes v$, the generic element of $Sym^2 W$, can be written as $(v + w) \otimes (v + w) - (v \otimes v) - (w \otimes w)$ and then proceed as if dealing with the elementary symmetric polynomials. We can consider $End(V) = V^* \otimes V$, we have that $End(V \otimes d) = (V \otimes d)^* \otimes (V \otimes d) = (V^* \otimes V \otimes d)^* \otimes (V \otimes d) = (End(V))^\otimes d$. Now since $B = Hom_{S_d}(V \otimes d, V \otimes d) \subset End(V \otimes d) = End(V)^\otimes d$ and it is invariant under $S_d$ it must in fact be a subset of $Sym^d(End(V))$, which is spanned by $End(V)$ (as $\phi^d$). But $GL(V)$ is dense in $End(V)$, so this implies the second statement of the lemma.

31
This lemma now proves that a $B$–module decomposition is the same as a $GL(V)$ decomposition in our case, showing that $V^{⊗d} = \bigoplus_λ (V^{⊗d}c_λ)$ is the irreducible decomposition over $GL(V)$ and of course $S_λV = V^{⊗d}c_λ$, showing the second part of theorem 25.

For the first part now we will use the fact that both the complete symmetric polynomials $h_λ$ and the Schur polynomials span the space of symmetric polynomials. We don’t know what exactly $V^{⊗d}a_λ$ is, but we know this for $V^{⊗d}c_λ$, it is in fact isomorphic to $\text{Sym}^{λ_1}V \otimes \text{Sym}^{λ_2}V \otimes \cdots \otimes \text{Sym}^{λ_κ}V \simeq V^{⊗d} \otimes A_λ a_λ$, as we symmetrize the rows of the Young tableaux of shape $λ$. From the representations of $S_d$ we have that $Aa_λ = \bigoplus_{μ|d} K_{μ,λ} A_μ$ is the decomposition into irreducible representations of $S_d$. From our lemma then we have that as $GL(V)$-modules

$$\text{Sym}^{λ_1}V \otimes \text{Sym}^{λ_2}V \otimes \cdots \otimes \text{Sym}^{λ_κ}V = V^{⊗d}a_λ = \bigoplus_μ (V^{⊗d}c_λ) ⊗ K_{μ,λ}. \quad (13)$$

Let $g \in GL(V)$ be a matrix with eigenvalues $x_1, \ldots, x_n$. Then the trace of $g$ on the left-hand side can be computed by the standard way of dealing with tensor products, i.e.

$$\text{Tr}|_{V^{⊗d}a_λ}(g) = \prod_i \text{Tr}|_{\text{Sym}^{λ_i}V}(g),$$

and we can compute the trace of $g$ on $\text{Sym}^{p}V$ by observing that if $g$ is diagonal(or diagonalizable), then if $v_{i_1}, \ldots, v_{i_p}$ are eigenvectors, $g(v_{i_1} \ldots v_{i_p}) = x_{i_1} \ldots x_{i_p}(v_{i_1} \ldots v_{i_p})$ in $\text{Sym}^{p}V$ and by dimension count going over all $i_1 ≤ \cdots ≤ i_p$, these are all eigenvectors of $g$ acting on $\text{Sym}^{p}V$, so $\text{Tr}(g) = \sum_{i_1 ≤ \cdots ≤ i_p} x_{i_1} \ldots x_{i_p} = h_p(x_1, \ldots, x_n)$ - the complete symmetric function. Using the fact that the semisimple matrices are a dense set in $GL(V)$ we can extend this result for any $g$ of the given eigenvalues. We then have that $\text{Tr}|_{V^{⊗d}a_λ}(g) = \prod_i h_λ(x_1, \ldots, x_n) = h_λ(x_1, \ldots, x_n)$ by definition of $h_λ$.

Now taking $\text{Tr}(g)$ on both sides of equation (13) gives us that

$$h_λ(x_1, \ldots, x_n) = \sum_μ K_{μ,λ} \text{Tr}|_{V^{⊗d}c_μ}(g), \quad (14)$$

which coincides with the expression for $h_λ$ in terms of $s_λ$. Since we know that the matrix with entries $K_{μ,λ}$ is invertible, after writing a system of equations (14) for all $λ \vdash d$, we can multiply by the inverse of the $K_{μ,λ}$ matrix and get that the vector of entries $\text{Tr}|_{S_λV}(g)$ running over $μ \vdash d$ is equal to $K^{-1} h$, where $h$ is the vector of entries $h_λ$. Since we know $s = K^{-1} h$, where $s$ is the vector of entries the Schur functions $s_μ$, we must have

$$\text{Tr}|_{S_λV}(g) = s_λ(x_1, \ldots, x_n),$$

as desired.

References

