Notes on Linear Algebra

Mark Reeder

December 19, 2015

Contents

I The Tensor Algebra and its quotients 3

1 The Tensor Algebra 3

1.1 Filtered and graded algebras ........................................... 4

2 The Symmetric Algebra 5

2.1 Polynomials and differential operators ................................. 5

3 The Exterior Algebra 6

3.1 Free modules ............................................................... 7

3.2 Determinants ............................................................... 9

3.3 The trace ................................................................. 9

3.4 Interpolating between trace and determinant ...................... 10

3.5 Vector spaces ............................................................ 11

3.5.1 The characteristic polynomial .................................... 11

3.5.2 A topological interpretation ...................................... 13

3.5.3 Grassmannians ...................................................... 14

3.5.4 Oriented Grassmannians and the characteristic polynomial .. 15
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Clifford algebras</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>Orthogonal bases and the center of the Clifford algebra</td>
<td>17</td>
</tr>
<tr>
<td>4.2</td>
<td>Orthogonal groups</td>
<td>18</td>
</tr>
<tr>
<td>4.3</td>
<td>The Clifford group</td>
<td>19</td>
</tr>
<tr>
<td>4.4</td>
<td>Clifford conjugation and the spinor norm</td>
<td>20</td>
</tr>
<tr>
<td>4.5</td>
<td>Spin groups</td>
<td>21</td>
</tr>
<tr>
<td>4.6</td>
<td>Summary</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>Lie algebras and their enveloping algebras</td>
<td>23</td>
</tr>
<tr>
<td>5.1</td>
<td>Origin and examples: Lie groups</td>
<td>24</td>
</tr>
<tr>
<td>5.2</td>
<td>Universal enveloping algebras</td>
<td>24</td>
</tr>
<tr>
<td>5.3</td>
<td>Representations and $U(L)$-modules</td>
<td>25</td>
</tr>
<tr>
<td>5.4</td>
<td>Representations of $\mathfrak{sl}_2$</td>
<td>26</td>
</tr>
<tr>
<td>II</td>
<td>Modules over Principal Ideal Domains</td>
<td>27</td>
</tr>
<tr>
<td>5.5</td>
<td>Modules over Noetherian Rings</td>
<td>28</td>
</tr>
<tr>
<td>5.6</td>
<td>Smith Normal Form: statement and consequences</td>
<td>29</td>
</tr>
<tr>
<td>5.7</td>
<td>Orbits of $\text{SL}_n(R)$ on $R^n$</td>
<td>30</td>
</tr>
<tr>
<td>5.8</td>
<td>Existence of Smith Normal Form</td>
<td>32</td>
</tr>
<tr>
<td>5.9</td>
<td>Uniqueness of Smith Normal Form</td>
<td>34</td>
</tr>
<tr>
<td>5.10</td>
<td>Summary and variations</td>
<td>35</td>
</tr>
<tr>
<td>6</td>
<td>Rational canonical form and regular transformations</td>
<td>35</td>
</tr>
<tr>
<td>6.1</td>
<td>Cyclic modules and companion matrices</td>
<td>36</td>
</tr>
<tr>
<td>6.2</td>
<td>Rational canonical form</td>
<td>37</td>
</tr>
<tr>
<td>6.3</td>
<td>Finding the rational canonical form</td>
<td>37</td>
</tr>
<tr>
<td>6.4</td>
<td>Centralizers and regular transformations</td>
<td>39</td>
</tr>
<tr>
<td>6.5</td>
<td>Conjugacy vs stable conjugacy</td>
<td>40</td>
</tr>
</tbody>
</table>
Part I

The Tensor Algebra and its quotients

I assume the reader has learned about tensor products, as found (for example) in Dummit and Foote [DF hereafter]. Throughout, \( R \) is a commutative ring with identity and \( M \) is an \( R \)-module. We will discuss four important \( R \)-algebras constructed from \( M \) and perhaps some additional structure on \( M \). All of these algebras will be quotients of a universal algebra, the **tensor algebra** of \( M \). All tensor products are taken in the category of \( R \)-modules and we write \( \otimes \) for \( \otimes_R \). After each of the four constructions I will try to show why they are important and interesting.

1 The Tensor Algebra

Let \( T^k M \) denote the \( k \)-fold tensor product of \( M \) with itself. Thus,

\[
T^1 M = M, \quad T^2 M = M \otimes M, \quad T^3 M = M \otimes M \otimes M, \quad \cdots.
\]

Define an \( R \)-module by

\[
TM = \bigoplus_{k=0}^{\infty} T^k M.
\]

On \( TM \) we define a product by

\[
(x_1 \otimes \cdots \otimes x_k) \cdot (y_1 \otimes \cdots \otimes y_\ell) = x_1 \otimes \cdots \otimes x_k \otimes y_1 \otimes \cdots \otimes y_\ell.
\]

One checks that this is well-defined. Thus \( TM \) becomes an \( R \)-algebra. It is a graded algebra, in the sense that

\[
T^k M \cdot T^{\ell} M \subset T^{k+\ell} M.
\]

We have a natural inclusion \( \iota : M = T^1 M \hookrightarrow TM \).

The mapping property for \( TM \) is as follows. Suppose \( A \) is an \( R \)-algebra. Then any \( R \)-module map \( \phi : M \to A \) extends to a unique \( R \)-module map \( \tilde{\phi} : TM \to A \), such that \( \tilde{\phi} \circ \iota = \phi \). This follows from the mapping property of tensor products.
If $M$ is a free $R$-module with basis $\{x_i : i \in I\}$, then $TM$ is also a free $R$-module, with basis $\{x_i \otimes x_k : \text{all } i, j \in I\}$. We can think of elements of $TM$ as noncommuting polynomial expressions in the $x_i$'s.

1.1 Filtered and graded algebras

An $R$-algebra $A$ is **graded** by a semigroup $K$ if

$$A = \bigoplus_{k \in K} A^k$$

where the $A^k$ are $R$-submodules of $A$ such that $A^j \cdot A^k \subset A^{j+k}$ for all $j, k \in K$.

For example the tensor algebra $T(V)$ is evidently graded by $\mathbb{N} = \{0, 1, 2, \ldots\}$. Gradings by $\mathbb{Z}/m\mathbb{Z}$ (for various $m$) also arise in nature.

An $\mathbb{N}$-graded algebra $A$ is **filtered** if there are $R$-submodules $A_n \subset A$, for $n \in \mathbb{N}$, such that $A_0 \subset A_1 \subset A_2 \subset \cdots$ with

$$A = \bigcup_{n \in \mathbb{N}} A_n,$$

such that $A_m \cdot A_n \subset A_{m+n}$ for all $m, n \in \mathbb{N}$.

Every $\mathbb{N}$- graded algebra $A = \oplus_{k \in \mathbb{N}} A^k$ is filtered via

$$A_n = \bigoplus_{k=0}^n A^k.$$  \hfill (1)

There are filtered algebras that do not arise in this way. However, given a filtered algebra $A = \bigcup A_n$, we can define the $R$-module

$$\text{gr}(A) = \bigoplus_{k \in \mathbb{N}} \text{gr}(A)^k,$$

where $\text{gr}(A)^0 = A_0$ and $\text{gr}(A)^k = A_k/A_{k-1}$ for $k > 0$. One checks that the product in $A$ induces a well-defined product in $\text{gr}(A)$, making $\text{gr}(A)$ into an $\mathbb{N}$-graded algebra, called the **associated graded algebra** of the filtered algebra $A$.

In what follows we will consider quotients of the tensor algebra $T(V)$ by certain two-sided ideals. Though $T(V)$ is graded, these quotients will not always be graded, but they will always be filtered.

Let $A = \oplus A^k$ be an $\mathbb{N}$-graded algebra, let $I \subset A$ be a (two-sided) ideal in $A$, and let $\pi : A \to A/I$ be the projection. The $R$-algebra $A/I$ is filtered by the image of the filtration of $A$:

$$(A/I)_n = \pi(A_n),$$

with $A_n$ as in (1).

We say that $I$ is **homogeneous** if $I$ is generated by elements lying in the graded subspaces $A^k$. This means $I = \oplus_k (I \cap A^k)$, so $A/I$ inherits the grading

$$(A/I)^k = A^k/I \cap A^k, \quad k \in \mathbb{N}. \hfill (2)$$
In the first of our two quotients of $A = T(V)$ below, the ideal $I$ will be homogeneous, so these quotients will be graded.

However in the third and fourth quotients, $I$ will not be homogeneous: there will be elements $a \in I$ of the form $a = a_k + a_{k-1} + \cdots + a_0$ with all $a_j \in A^j$, but not all $a_j \in I$. Assume $a_k \neq 0$. The leading term $\sigma(a) = a_k$ is called the symbol of $a$. Let $\sigma(I)$ be the ideal generated by $\{\sigma(a) : a \in I\}$. Then $\sigma(I)$ is homogenous, so the quotient $A/\sigma(I)$ is naturally graded, as in (2), and we have

$$A/\sigma(I) \simeq \text{gr}(A/I),$$

as graded algebras, via the map sending the class of $a \in A^j$ to the class of $a + I \in (A/I)_j$.

2 The Symmetric Algebra

This is the first of our four quotients of $T(M)$. Let $I_{\text{sym}}(M)$ be the (two-sided) ideal in $T(M)$ generated by $\{x \otimes y - y \otimes x : x, y \in M\}$. The Symmetric Algebra of $M$ is defined as

$$\text{Sym}(M) = T(M)/I_{\text{sym}}(M).$$

The relations in $I_{\text{sym}}(M)$ force $\text{Sym}(M)$ to be commutative. Since $I_{\text{sym}}(M)$ is homogeneous, the quotient $\text{Sym}(M)$ is graded:

$$\text{Sym}(M) = \bigoplus_{k \geq 0} \text{Sym}^k(M),$$

where $\text{Sym}^k(M)$ is the image of $T^k(M)$ in $\text{Sym}(M)$. We write $x_1x_2\cdots x_k \in \text{Sym}^k(M)$ for the image of $x_1 \otimes \cdots \otimes x_k \in T(M)$.

The mapping property for $\text{Sym}(M)$ is as follows. Suppose $A$ is a commutative $R$-algebra. Then any $R$-module map $\phi : M \to A$ extends to a unique $R$-module map $\tilde{\phi} : \text{Sym}(M) \to A$, such that $\tilde{\phi} \circ \iota = \phi$. This follows from the mapping property of $T(M)$. For $x_1, \ldots, x_k \in M$ we have

$$\tilde{\phi}(x_1x_2\cdots x_k) = \phi(x_1)\phi(x_2)\cdots \phi(x_k).$$

In this sense, $\text{Sym}(M)$ is the largest commutative quotient of $T(M)$.

2.1 Polynomials and differential operators

Suppose $M$ is free over $R$ of rank $n$ and that $\{x_1, \ldots, x_n\}$ is an $R$-basis of $M$. Let $R[X_1, \ldots, X_n]$ be the polynomial ring in indeterminants $X_i$. Since $M$ is free, there is a unique $R$-module mapping

$$\phi : M \to R[X_1, \ldots, X_n], \quad x_i \mapsto X_i.$$
By the mapping property this extends to an isomorphism of graded algebras:

\[ \tilde{\phi} : \text{Sym}(M) \to R[X_1, \ldots, X_n]. \]

This is surjective because the \( X_i \) generate \( R[X_1, \ldots, X_n] \) and is injective because the monomials \( X_1^{i_1} \cdots X_n^{i_n} \) form an \( R \)-basis of \( R[X_1, \ldots, X_n] \). It follows that \( \text{Sym}(M) \) is free over \( R \) with basis \( \{x_1^{i_1} \cdots x_n^{i_n} : i_j \geq 0\} \). Thus, \( \text{Sym}(M) \) is a coordinate-free version of the polynomial ring.

To express \( \text{Sym} \) in terms of polynomial functions, it is more natural to pass to the dual space \( \check{M} = \text{Hom}_R(M, R) \) of \( R \)-linear functionals on \( M \). Let \( \mathcal{F}(M) \) be the \( R \)-algebra of all \( R \)-valued functions on \( M \). The inclusion \( \iota : \check{M} \hookrightarrow \mathcal{F}(M) \) extends, by the mapping property, to an \( R \)-algebra homomorphism \( \iota : S(\check{M}) \to \mathcal{F}(M) \). A polynomial function on \( M \) is a element \( f \in \mathcal{F}(M) \) of the form \( f = \iota(s) \) for some \( s \in S(\check{M}) \). Note that the mapping \( \iota \) need not be injective. For example, if \( R = \mathbb{F}_p = M \), and \( 0 \neq x \in \check{M} \), then \( \iota(x^p - x) = 0 \). Thus, a nonzero polynomial can give the zero function.

However, if \( R \) is an infinite field, then \( \iota \) is injective, so we can identify \( \text{Sym}(\check{M}) \) with polynomial functions on \( M \). In this interpretation, \( \text{Sym}(M) \) becomes a ring of differential operators on \( \text{Sym}(\check{M}) \), as follows. Each \( x \in M \) may be viewed as a linear function \( x : \check{M} \to R \lambda \mapsto \langle \lambda, x \rangle \). This extends to the linear map

\[ \partial_x \in \text{End}_R(T(\check{M})) \]

given by

\[ \partial_x(\lambda_1 \otimes \cdots \otimes \lambda^k) = \sum_{i=1}^k \langle \lambda_i, x \rangle \lambda_1 \otimes \cdots \otimes \lambda_{i-1} \otimes \lambda_{i+1} \otimes \lambda_k. \]

Thus, \( \partial_x \) is the unique extension satisfying the product rule. Such an endomorphism of an algebra is called a derivation, because \( \partial_x \) is analogue of the directional derivative in Calculus.

One checks that for \( x, y \in M \) we have \( \partial_x \partial_y = \partial_y \partial_x \). By the mapping property, \( x \mapsto \partial_x \) extends to an algebra homomorphism

\[ \text{Sym}(M) \to \text{End}_R(\text{Sym}(\check{M})), \quad x_1^{i_1} \cdots x_n^{i_n} \mapsto \partial_{x_1}^{i_1} \cdots \partial_{x_n}^{i_n}. \]

Thus \( \text{Sym}(M) \) acts on \( \text{Sym}(\check{M}) \) via differential operators. Of course \( \text{Sym}(\check{M}) \) acts on itself by multiplication. These two actions do not commute with each other; one checks that

\[ \partial_x \lambda - \lambda \partial_x = \langle \lambda, x \rangle \in \text{End}(\text{Sym}(\check{M})). \]

This is an algebraic version of the Uncertainty Principle of Physics, in which \( \lambda \) detects position and \( \partial_x \) detects momentum. The subalgebra of \( \text{End}(\text{Sym}(\check{M})) \) generated by \( \text{Sym}(M) \) and \( \text{Sym}(\check{M}) \) is called the Weyl Algebra. The Weyl algebra can also be described directly as a quotient of a tensor algebra, but we shall not do this.

### 3 The Exterior Algebra

The exterior algebra \( \Lambda(M) \) is the \( R \)-algebra obtained as the quotient of \( T(M) \) by the homogeneous ideal \( I_{\text{alt}}(M) \) generated by \( \{m \otimes m : m \in M\} \). Each graded component \( \Lambda^k(M) \) is generated as an
Let $\Lambda^k(M) = T^k(M)/A^k(M)$, where $A^k(M)$ is the submodule of $T^k(M)$ generated by all $k$-fold tensors $m_1 \otimes \cdots \otimes m_k$ where $m_i = m_j$ for some $1 \leq i < j \leq k$. Hence we have the following mapping property:

If $N$ is an $R$-module and $f : T^k(M) \rightarrow N$ is an $R$-module map vanishing on pure $k$-tensors with two equal components then there is a unique $R$-module map

\[ \tilde{f} : \Lambda^k M \rightarrow N \text{ such that } \tilde{f}(m_1 \wedge \cdots \wedge m_k) = f(m_1 \otimes \cdots \otimes m_k). \]

By definition, we have $m \wedge m = 0$ for all $m \in M$. Expanding $0 = (m + m') \wedge (m + m')$ shows that

\[ m \wedge m' = -m' \wedge m \]

for all $m, m' \in M$. It follows that switching any pair $m_i, m_j$ in a pure wedge changes the sign:

\[ m_1 \wedge \cdots \wedge m_i \wedge \cdots \wedge m_j \wedge \cdots \wedge m_k = -m_1 \wedge \cdots \wedge m_j \wedge \cdots \wedge m_i \wedge \cdots \wedge m_k. \]

If $-1 = 1$ in $R$ then $A(M)$ properly contains the ideal in $T(M)$ generated by all expressions $m \otimes m' + m' \otimes m$ so these do not suffice to define the relations in $\Lambda(M)$.

### 3.1 Free modules

If $M$ and $M'$ are free $R$-modules with bases $\{m_i\}$ and $\{m'_j\}$ respectively, then $M \otimes_R M'$ is again free, with basis $\{m_i \otimes m'_j\}$.

Assume $M$ is free of rank $n$, and let $\{m_1, \ldots, m_n\}$ be a basis of $M$. By induction, it follows that $T^k(M)$ is free with basis $\{m_{i_1} \otimes \cdots \otimes m_{i_k}\}$ indexed by $[1, n]^k$. Since $m_{i_1} \otimes \cdots \otimes m_{i_k} \in A^k(M)$ whenever $i_p = i_q$ for $p \neq q$, it follows that $\Lambda^k(M)$ is spanned by elements of the form $m_{i_1} \wedge \cdots \wedge m_{i_k}$ for $1 \leq i_1 < i_2 < \cdots < i_k \leq n$. In particular $\Lambda^k(M) = 0$ for $k > n$ and $\Lambda^n(M)$ is a cyclic $R$-module generated by the $n$-fold wedge product $m_1 \wedge \cdots \wedge m_n$.

**Lemma 3.1** Let $M$ be a free $R$-module of rank $n$. Then $\Lambda^n(M)$ is a free $R$-module of rank one, generated by $m_1 \wedge \cdots \wedge m_n$, where $\{m_i\}$ is any basis of $M$.

**Proof:** Our remarks above show that $m_1 \wedge \cdots \wedge m_n$ spans $\Lambda^n(M)$. We will prove the lemma by constructing an $R$-module map $\varphi : \Lambda^n(M) \rightarrow R$ such that $\varphi(m_1 \wedge \cdots \wedge m_n) = 1$. \(^2\)

Let $\{\lambda_1, \ldots, \lambda_n\} \subset \text{Hom}_R(M, R)$ be the dual basis of $\{m_i\}$, defined by $\lambda_i(m_i) = 1$ and $\lambda_i(m_j) = 0$ if $i \neq j$. Then we have a well-defined $R$-module map $\phi : T^n(M) \rightarrow R$ given by

\[ \phi(x_1 \otimes \cdots \otimes x_n) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \lambda_1(x_{\sigma 1}) \cdots \lambda_n(x_{\sigma n}). \]

\(^2\)This idea comes from P. Garrett’s notes http://www.math.umn.edu/~garrett/m/algebra/notes/28.pdf
We show that \( \phi \) vanishes on \( A^n(M) \): Suppose \( x_i = x_j \) for some \( i \neq j \), and let \( \tau \in S_n \) be the transposition switching \( i \) and \( j \). Then the sum over \( S_n \) is a sum over pairs \( \sigma, \tau\sigma \):

\[
\text{sgn}(\sigma)\lambda_1(x_{\sigma_1}) \cdots \lambda_n(x_{\sigma_n}) + \text{sgn}(\tau\sigma)\lambda_1(x_{\sigma_1}) \cdots \lambda_n(x_{\sigma_n}).
\]

(4)

Define \( p, q \in [1, n] \) by \( \sigma p = i \) and \( \sigma q = j \). Then (4) becomes

\[
\text{sgn}(\sigma) \prod_{\ell \neq p, q} \lambda_{\ell}(x_{\sigma_{\ell}}) \cdot [\lambda_p(x_{\sigma p})\lambda_q(x_{\sigma q}) - \lambda_p(x_{\tau\sigma p})\lambda_q(x_{\tau\sigma q})]
\]

and the term in \([\cdots]\) is

\[
\lambda_p(x_i)\lambda_q(x_j) - \lambda_p(x_j)\lambda_q(x_i) = 0,
\]

since \( x_i = x_j \). Hence \( \phi \) induces a well-defined map \( \varphi : \Lambda^n(M) \to R \), given by

\[
\varphi(x_1 \wedge \cdots \wedge x_n) = \sum_{\sigma \in S_n} \text{sgn}(\sigma)\lambda_1(x_{\sigma_1}) \cdots \lambda_n(x_{\sigma_n}).
\]

Taking each \( x_i = m_i \), every term in the sum vanishes except for \( \sigma = e \), so

\[
\varphi(m_1 \wedge \cdots \wedge m_n) = 1
\]

as claimed.

We set \([1, n] = \{1, 2, \ldots, n\} \), let \( k \in [1, n] \) and let \([1, n]_k \) be the set of \( k \)-element subsets of \([1, n] \). Given a set of vectors \( \{v_1, \ldots, v_n\} \) in \( V \) and a subset \( I \subset [1, n]_k \), let \( v_I = v_{i_1} \wedge \cdots \wedge v_{i_k} \), where \( i_1 < i_2 < \cdots < i_k \) are the elements of \( I \) listed in increasing order.

**Proposition 3.2** If \( \{m_1, \ldots, m_n\} \) is a basis of \( M \) then \( \{m_I : I \in [1, n]_k\} \) is a basis of \( \Lambda^k(M) \). In particular \( \Lambda^k(M) \) is free of rank \( \binom{n}{k} \).

**Proof:**

Since the \( k \)-fold pure tensor products of the \( m_i \) span \( T^k(M) \), the corresponding wedge products span \( \Lambda^k(M) \). Since these wedge products may be permuted up to sign, it follows that \( \{M_I : I \in [1, n]_k\} \) spans \( \Lambda^k(M) \).

Suppose we have a dependence relation

\[
\sum_{I \in [1, n]_k} c_Im_I = 0.
\]

Choose \( J \in [1, n]_k \) arbitrarily and let \( J' \in [1, n]_{n-k} \) be its complement. Then \( m_{J'} \wedge m_J = \pm m_1 \wedge \cdots \wedge m_n \), and for \( I \neq J' \) there exists \( i \in I \cap J \) so we have \( m_J \wedge m_I = 0 \). Therefore

\[
0 = m_{J'} \wedge \sum_{I \in [1, n]_k} c_Im_I = \pm c_Jm_1 \wedge \cdots \wedge m_n.
\]

Since \( m_1 \wedge \cdots \wedge m_n \neq 0 \) by Lemma 3.1, it follows that \( c_J = 0 \). This shows that \( \{m_I : I \in [1, n]_k\} \) is linearly independent.

\[\text{The proof of this result in [DF] is not valid if } 2 = 0 \text{ in } R. \text{ The proof here works in any characteristic.}\]
3.2 Determinants

The exterior product $\Lambda^k(\cdot)$ is a functor. That is, given two $R$-modules $M, N$ and an $R$-module map $T : M \to N$ we get for each $k \geq 1$ a well-defined $R$-module map

$$\Lambda^k(T) : \Lambda^k(M) \to \Lambda^k(N) \quad m_1 \wedge \cdots \wedge m_k \mapsto T(m_1) \wedge \cdots \wedge T(m_k).$$

In particular each $T \in \text{End}_R(M)$ gives a series of $R$-linear maps $\Lambda^k(T) \in \text{End}_R(\Lambda^k(M))$, for $k \geq 1$.

If $M$ is free of rank $n$ over $R$, then according to Lemma 3.1, $\Lambda^n(M)$ is a cyclic $R$-module generated by the $n$-fold wedge product

$$m_1 \wedge \cdots \wedge m_n$$

of any basis $\{m_1, \ldots, m_n\}$. Therefore $\Lambda^n(T)$ acts on $\Lambda^n(M)$ via multiplication by some element of $R$, called the determinant of $T$, denoted by $\det(T)$. That is,

$$\Lambda^n(T) = \det(T) \cdot \text{id} \quad \text{on} \quad \Lambda^n(M).$$

This definition makes it clear that $\det(T)$ is independent of any basis of $M$ and that

$$\det(ST) = \det(S) \cdot \det(T), \quad \text{for} \ S, T \in \text{End}_R(M).$$

In particular we have

$$T \in \text{Aut}_R(M) \quad \text{if and only if} \quad \det(T) \in R^\times.$$

Finally if $A_T = [a_{ij}]$ is the matrix of $T$ with respect to the basis $\{m_i\}$ of $M$, then

$$\det(T) \cdot m_1 \wedge \cdots \wedge m_n = T(m_1) \wedge \cdots \wedge T(m_n)$$

$$= \left( \sum_{i=1}^n a_{1i} m_i \right) \wedge \cdots \wedge \left( \sum_{i=1}^n a_{ni} m_i \right)$$

$$= \left( \sum_{\sigma \in S_n} \text{sgn}(\sigma)a_{1\sigma(1)} \cdots a_{n\sigma(n)} \right) \cdot m_1 \wedge \cdots \wedge m_n,$$

recovering the usual matrix formula for $\det(T)$.

3.3 The trace

Just as we defined the determinant canonically so we can define the trace $\text{tr}(T)$ of an endomorphism $T \in \text{End}_R(M)$.

Let $\partial_T : \Lambda^n(M) \to \Lambda^n(M)$ be the (well-defined) $R$-module map given, for any vectors $x_1, \ldots, x_n \in M$ by

$$\partial_T(x_1 \wedge \cdots \wedge x_n) = \sum_{j=1}^n x_1 \wedge \cdots \wedge T(x_j) \wedge \cdots \wedge x_n,$$
where in the \( j \)th term \( T \) is applied to \( x_j \) and all other \( x_i \) are unchanged. Then we define \( \text{tr}(T) \) to be the scalar by which \( \partial T \) acts on the rank-one \( R \)-module \( \Lambda^n(M) \):

\[
\partial T = \text{tr}(T) \cdot \text{id} \quad \text{on} \quad \Lambda^n(M).
\]

This definition of \( \text{tr}(T) \) does not require a choice of basis of \( M \). However, if \( A_T = [a_{ij}] \) is the matrix of \( T \) with respect to a basis \( \{m_1, \ldots, m_n\} \) of \( M \) then for each \( j \) we have

\[
m_1 \wedge \cdots \wedge T(m_j) \wedge \cdots \wedge m_n = \sum_{i=1}^{n} a_{ij} m_1 \wedge \cdots \wedge m_i \wedge \cdots \wedge m_n = a_{jj} m_1 \wedge \cdots \wedge m_j \wedge \cdots \wedge m_n.
\]

It follows that

\[
\text{tr}(T) = \sum_{i=1}^{n} a_{ii}
\]

is the sum of the diagonal entries of \( A_T \), recovering the usual formula for the trace.

If \( n = 1 \) then \( \text{tr}(T) = \det(T) \). Hence for general \( n \geq 1 \) we have

\[
\text{tr}(\Lambda^n(T)) = \det(T).
\]

### 3.4 Interpolating between trace and determinant

Let \( M \) be a free \( R \)-module of rank \( n \) and let \( T \in \text{End}_R(M) \). We know explicit formulas for \( \text{tr}(\Lambda^1(T)) = \text{tr}(T) \) and \( \text{tr}(\Lambda^n(T)) = \det(T) \). In this section we interpolate between these two to give an explicit formula for \( \text{tr}(\Lambda^k(T)) \) for any \( 1 \leq k \leq n \).

Choose a basis \( \{m_i\} \) of \( M \) and let \( A = [a_{ij}] \) be the matrix of \( T \). Given a \( k \)-element subset \( I = \{i_1 < \cdots < i_k\} \subset [1, n] \), let

\[
M_I = \{m_i : i \in I\} \quad \text{and} \quad M'_I = \{m_j : j \notin I\}
\]

so that \( M = M_I \oplus M'_I \) with corresponding inclusion and projection maps

\[
M_I \hookrightarrow M \xrightarrow{\pi} M_I.
\]

Define \( T_I \) to be the composition

\[
T_I : M_I \hookrightarrow M \xrightarrow{T} M \xrightarrow{\pi} M_I.
\]

For example if \( k = 1 \) and \( I = \{i\} \) then \( T_I = a_{ii} \cdot \text{id} \) on \( M_I = \text{R}m_i \).

Let \( \{m_I : I \in [1, n]_k\} \) be the basis of \( \Lambda^k(M) \) from Prop. 3.2.

\[
\Lambda^k(T)m_I = T(m_{i_1}) \wedge \cdots \wedge T(m_{i_k})
\]

\[
= T_I(m_{i_1}) \wedge \cdots \wedge T_I(m_{i_k}) + \sum_{j \notin I} a_{jm_j} m_j
\]

\[
= \det(T_I) m_I + \sum_{j \notin I} c_{jm_I}
\]

10
for some coefficients $c_J \in \mathbb{R}$. It follows that

$$\text{tr}(\Lambda^k(T)) = \sum_{I \in [1,n]_k} \det(T_I).$$

When $k = 1$ this is $\text{tr}(T)$ and when $k = n$ the sum consists of one term, namely $\det(T)$.

## 3.5 Vector spaces

Now let $F$ be a field and let $V$ be a vector space over $F$ of dimension $\dim V = n$. We write $\text{End}(V)$ for $\text{End}_F(V)$.

**Proposition 3.3** A subset $\{v_1, \ldots, v_k\} \subset V$ is linearly independent if and only if $v_1 \wedge \cdots \wedge v_k \neq 0$.

**Proof:** If $\{v_1, \ldots, v_k\} \subset V$ is linearly dependent then some $v_i$ is a linear combination of the other $v_j$. By multilinearity $v_1 \wedge \cdots \wedge v_k$ is a sum of terms each with repeated factors hence is zero.

If $\{v_1, \ldots, v_k\} \subset V$ is linearly independent then it is contained in a basis $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ of $V$. By Lemma 3.1 we have

$$v_1 \wedge \cdots \wedge v_k \wedge v_{k+1} \wedge \cdots \wedge v_n \neq 0,$$

so $v_1 \wedge \cdots \wedge v_k \neq 0$. ■

**Proposition 3.4** Let $\{u_1, \ldots, u_k\}$ and $\{w_1, \ldots, w_k\}$ be two linearly independent subsets of $V$, spanning subspaces $U$ and $W$ respectively, of $V$. Then $U = W$ if and only if $u_1 \wedge \cdots \wedge u_k$ and $w_1 \wedge \cdots \wedge w_k$ are proportional in $\Lambda^k(V)$.

**Proof:** If $U = W$ the map $T : U \to U$ given by $T(u_i) = w_i$ is an isomorphism which induces an isomorphism $\Lambda^k(T) : \Lambda^k(U) \to \Lambda^k(U)$, such that $u_1 \wedge \cdots \wedge u_k = \det(T) \cdot u_1 \wedge \cdots \wedge u_k$.

Conversely, suppose $w_1 \wedge \cdots \wedge w_k = c \cdot u_1 \wedge \cdots \wedge u_k$ for some $c \in F$. Then for each $i$, $u_i \wedge w_1 \wedge \cdots \wedge w_k = 0$, so $\{u_i, w_1, \ldots, w_k\}$ is dependent, by Prop. 3.3. Since the $w_j$ are independent, it follows that $u_i$ belongs to the span of the $w_j$, so $u_i \in W$. As $i$ was arbitrary, we have $U \subset W$ so $U = W$.

### 3.5.1 The characteristic polynomial

An eigenvector for $T \in \text{End}(V)$ is a nonzero vector such that $T(v)$ is a scalar multiple of $V$, say $T(v) = \lambda v$. This scalar $\lambda$ is an eigenvalue for $T$. Eigenvectors with eigenvalue $\lambda$ exist if and only if $\ker(\lambda I_V - T) \neq 0$, in other words, when $\det(\lambda I_V - T) = 0$. Therefore the eigenvalues of $T$ are the roots of the characteristic polynomial $\det(x I_V - T)$. Here $x I_V - T$ is an endomorphism of the free
Proposition 3.5 The the expansion of the characteristic polynomial in powers of $x$ is given by

$$\det(xI_V - T) = \sum_{k=0}^{n} (-1)^{n-k} \text{tr}(\Lambda^k(T)) \cdot x^{n-k}. $$

Proof: It is equivalent to prove that

$$\det(xI_V + T) = \sum_{k=0}^{n} \text{tr}(\Lambda^k(T)) \cdot x^{n-k}. $$

Choose an $F$-basis $\{v_1, \ldots, v_n\}$ of $V$, so that the $n$-form $\omega = v_1 \wedge \cdots \wedge v_n$ generates the one-dimensional $F$-vector space $\Lambda^n(V)$. Identifying $1 \otimes v_i$ with $v_i$, we may also regard $\{v_i\}$ as an $F[x]$-basis of $F[x] \otimes_F V$ and $\omega$ as an $F[x]$-module generator of $\Lambda^n(F[x] \otimes_F V)$. Computing in the latter exterior power, we have

$$\det(xI_V + T) \cdot \omega = (xv_1 + T(v_1)) \wedge \cdots \wedge (xv_n + T(v_n)) $$

$$= x^n \omega + x^{n-1} \left( \sum_{i=1}^{n} v_1 \wedge \cdots \wedge T(v_i) \wedge \cdots \wedge v_n \right) $$

$$+ x^{n-2} \left( \sum_{1 \leq i < j \leq n} v_1 \wedge \cdots \wedge T(v_i) \wedge \cdots \wedge T(v_j) \wedge \cdots \wedge v_n \right) + \cdots + \det(T) \cdot \omega $$

Consider a general term corresponding to $I = \{i_1 < \cdots < i_k\}$. In the notation of section 3.4 we have

$$v_1 \wedge \cdots \wedge T(v_{i_1}) \wedge \cdots \wedge T(v_{i_k}) \wedge \cdots \wedge v_n = v_1 \wedge \cdots \wedge T_I(v_{i_1}) \wedge \cdots \wedge T_I(v_{i_k}) \wedge \cdots \wedge v_n = \det(T_I) \cdot \omega. $$

From Prop. 5 it follows that in the rank-one $F[x]$-module $\Lambda^n(F[x] \otimes_F V)$ we have

$$\det(xI_V + T) \cdot \omega = \sum_{k=0}^{n} x^{n-k} \cdot \sum_{I \in [1,n]_k} \det(T_I) \cdot \omega = \sum_{k=0}^{n} x^{n-k} \cdot \text{tr}(\Lambda^k(T)) \cdot \omega, $$

so

$$\det(xI_V + T) = \sum_{k=0}^{n} x^{n-k} \cdot \text{tr}(\Lambda^k(T)), $$

as claimed.  

Note the left side is an exterior power of an $F[x]$-module and the right side is an exterior power of an $F$-module, with scalars extended to $F[x]$. 

---

4Note the left side is an exterior power of an $F[x]$-module and the right side is an exterior power of an $F$-module, with scalars extended to $F[x]$. 

---
3.5.2 A topological interpretation

Let $V$ be a real vector space of dimension $n$. A lattice in $V$ is a subgroup $L \subset V$ generated by a basis $\{v_i\}$ of $V$. The quotient $X = V/L$ is a torus (a product of $n$-copies of the circle $S^1$), whose $k$th homology group (with real coefficients) $H_k(X)$ is isomorphic to $\Lambda^k(V)$. Explicitly, for each $I \subset [1, n]_k$ the image of $V_I$ in $X$ is a subtorus $X_I$ of dimension $k$ and the fundamental classes of the $X_I$ for $I \in [1, n]_k$ form a basis of $H_k(X)$. Any continuous mapping $f : X \to X$ determines a linear map $h_k(f) : H_k(X) \to H_k(X)$. The Lefschetz number of $f$ is defined as the alternating sum of traces:

$$L(f) = \sum_{k=0}^{n} (-1)^{n-k} \text{tr}(h_k(f)).$$

For example if $f = \text{id}$ then $L(f)$ is the Euler characteristic (which is zero for $X$).

Suppose now that $f$ is smooth, with finite fixed-point set $X^f$. At each fixed point $p \in X^f$, the derivative $f'_p$ gives a linear endomorphism $f'_p \in \text{End}(V)$ and the Lefschetz fixed-point theorem asserts that

$$L(f) = \sum_{p \in X^f} \text{sgn} [\det(I - f'_p)].$$

Any endomorphism $T \in \text{End}(V)$ satisfying $T(L) \subset L$ gives a smooth map $f_T : X \to X$ via $f_T(v + L) = T(v) + L$. Identifying $H_k(X) = \Lambda^k(V)$, we have

$$h_k(f_T) = \Lambda^k(T), \quad \text{so} \quad L(f_T) = \sum_{k=0}^{n} (-1)^{n-k} \text{tr}(\Lambda^k(T)) = \det(I_V - T). \quad (6)$$

The fixed-points of $f_T$ are the $v + T$ such that $T(v) - v \in L$. This fixed-point set is finite if and only if $\det(I_V - T) \neq 0$ (exercise), in which case it is given by

$$X^{f_T} = (I_V - T)^{-1}L/L \simeq L/(I_V - T)L. \quad (7)$$

The derivative $f'_T = T$, so $\det(I - f'_T)_p = \det(I_V - T)$ at every fixed-point $p$. So the Lefschetz fixed-point theorem becomes the equality

$$L(f_T) = |X^{f_T}| \cdot \text{sgn}[\det(I - T)].$$

Comparing with (6) and (7), we get the purely algebraic result

$$|\det(I_V - L)| = [L : (I_V - T)L]. \quad (8)$$

We will shortly give an algebraic proof of equation (8).

---

5 It is customary to define the Lefschetz number in terms of cohomology. Using Poincaré duality it may be phrased in terms of homology, which is simpler in our context.

6 The tangent spaces of $X$ may be identified with $V$. 

---
3.5.3 Grassmannians

The Grassmannian of $k$-planes is the set $G_k(V)$ of $k$-dimensional subspaces of $V$. This generalizes the projective space $\mathbb{P}(V) = G_1(V)$ of lines in $V$. According to Prop. 3.4 we have an injective mapping

$$\lambda : G_k(V) \hookrightarrow \mathbb{P}(\Lambda^k V),$$

sending a $k$-plane $U \subset V$ to the line $[U] \in \mathbb{P}(\Lambda^k V)$ through the $k$-fold wedge of any basis of $U$. This is the Plücker embedding of the Grassmannian. As a subset of $\mathbb{P}(\Lambda^k V)$, it turns out that $G_k(V)$ is an intersection of quadrics, called Plücker relations.

Example: Planes in four dimensions If $k = 1$ or $n - 1$ then the Grassmannian is just a projective space. The first new case, the one originally considered by Plücker, is $k = 2$ and $\dim V = 4$. Assume $2 \neq 0$ in $F$. In this case we have a quadratic form

$$q : \Lambda^2 V \longrightarrow \Lambda^4 V, \quad q(\omega) = \frac{1}{2} \omega \wedge \omega$$

defining a quadric

$$Q = \{ [\omega] \in \mathbb{P}(\Lambda^2 V) : q(\omega) = 0 \}.$$

If $\omega$ is a pure wedge $\omega = u \wedge v$, then $q(u \wedge v) = 0$, so via the embedding $u \wedge v \mapsto [u \wedge v] \in \mathbb{P}(\Lambda^2 V)$, we have

$$G_2(V) \subset Q.$$

We will show this is an equality.

For any ordered basis $(e_1, e_2, e_3, e_4)$ of $V$ we write a nonzero 2-form as

$$\omega = \sum_{i<j} \omega_{ij} e_i \wedge e_j,$$

with coefficients $\omega_{ij} \in F$. We identify $\Lambda^4 V = F$, via the basis element $e_{1234}$. One checks that

$$q(\omega) = \omega_{12} \omega_{34} - \omega_{13} \omega_{24} + \omega_{14} \omega_{23}.$$

We must show that if $q(\omega) = 0$ then $\omega = u \wedge v$ for some $u, v \in V$. Consider the map

$$\omega \wedge : V \longrightarrow \Lambda^3 V, \quad v \mapsto \omega \wedge v.$$

Using the ordered bases $(e_1, e_2, e_3, e_4)$ and $(e_{234}, e_{134}, e_{124}, e_{123})$, the matrix of $\omega \wedge$ is

$$M_\omega = \begin{bmatrix}
0 & \omega_{34} & -\omega_{24} & \omega_{23} \\
\omega_{34} & 0 & -\omega_{14} & \omega_{13} \\
\omega_{24} & -\omega_{14} & 0 & \omega_{12} \\
\omega_{23} & -\omega_{13} & \omega_{12} & 0
\end{bmatrix}$$

Since $\omega \neq 0$, some diagonal $2 \times 2$ minor $\omega_{ij}^2 \neq 0$. It follows that $\dim \ker M_\omega \leq 2$. Remarkably, the $3 \times 3$ minors of $\det M_\omega$ are $(-1)^{i+j} \omega_{ij} q(\omega)$, so that $\det M_\omega = q(\omega)^2$. It follows that $q(\omega) = 0$ if and only if $\dim \ker M_\omega = 2$. In this case we may choose our basis $\{ e_i \}$ so that $\ker M_\omega = \{ e_1, e_2 \}$, making first two columns of $M_\omega$ equal to zero. This implies that $\omega = \omega_{12} e_1 \wedge e_2$. Thus we have proved that $G_2(V) = Q$, as claimed.
3.5.4 Oriented Grassmannians and the characteristic polynomial

We now take $F = \mathbb{R}$, the field of real numbers. Let $U$ be an $\mathbb{R}$-vector space of dimension $k$. Then the set of nonzero $k$-forms $\omega \in \Lambda^k U$ is a disjoint union of two open half-lines. An orientation on $U$ is a choice of one of these half-lines. We write $(U, \omega)$ to denote the oriented vector space $U$ where $\omega \in \Lambda^k U$ belongs to the chosen orientation. An ordered basis $(e_1, \ldots, e_k)$ of $U$ is an oriented basis for $(U, \omega)$ if $e_1 \wedge e_2 \wedge \cdots \wedge e_k$ belongs to the half-line containing $\omega$. An element $g \in \text{GL}(U)$ acts on $\Lambda^k U$ by the scalar $\det g$, so any two oriented bases of $U$ differ by transformation $g \in \text{GL}(U)$ with $\det g > 0$.

Now let $V$ be an $\mathbb{R}$-vector space of dimension $n$. The oriented Grassmannian $\tilde{G}_k(V)$ is the set of pairs $(U, \omega)$ where $U$ is a $k$-plane in $V$ and $\omega$ is an orientation on $U$.

Suppose we also have a positive definite quadratic form $q : V \to \mathbb{R}$. The restriction $q_U$ of $q$ to any subspace $U \subset V$ is again a positive definite quadratic form on $U$. Each orientation $\omega$ of $U$ may be taken to be of the form $\omega = e_1 \wedge \cdots \wedge e_k$ where $(e_1, \ldots, e_k)$ is an oriented orthonormal basis of $U$. Since any two such bases differ by a transformation from $\text{SO}(U, q_U)$, it follows that we have a well-defined embedding

$$\tilde{G}_k(V) \hookrightarrow \Lambda^k V, \quad (U, \omega) \mapsto \omega,$$

(9)

which covers the Plücker embedding $G_k(V) \hookrightarrow \mathbb{P}(\Lambda^k V)$.

If $k = 1$ this identifies $\tilde{G}_1(V)$ with the unit sphere $S = \{ v \in V : |v| = 1 \}$, whereby an oriented line $(L, \omega)$ in $V$ corresponds to the point $\omega \in S$ where the line exits $S$. This embedding gives a geometric interpretation of the trace of a transformation $T \in \text{End}(V)$, as follows. Associated to $q$ we have the bilinear form $\langle u, v \rangle = \frac{1}{2} [q(u + v) - q(u) - q(v)]$. Then we have

$$\frac{\text{tr}(T)}{n} = \int_S \langle T\omega, \omega \rangle,$$

(10)

where the integral is taken with respect to the unique $\text{SO}(V)$-invariant measure on $S$ such that $\int_S 1 = 1$. That is, the average of the eigenvalues of $T$ on $V$ is the average over all $\omega \in S$ of the signed lengths of the orthogonal projections of $T\omega$ back to the oriented line through $\omega$.

We will prove (10) algebraically, by showing how the integral can be computed by differentiation. The integrand is the restriction to $S$ of the new quadratic form $q_T(v) := \langle Tv, v \rangle$ and $\int_S$ is an $\text{SO}(V, q)$-invariant linear functional on the space $Q$ of quadratic forms on $V$. On the other hand, if $\{e_i\}$ is an $q$-orthonormal basis then the Laplacian $\Delta = \sum_i \partial^2_{e_i}$ is also an $\text{SO}(V, q)$ invariant functional on $Q$. I claim that all $\text{SO}(V, q)$ invariant functionals on $Q$ are proportional. Granting this, we have $\int_S f = c \Delta f$ for every $f \in Q$. Taking $f = q$, we find $c = 1/2n$. It follows that

$$\int_S \langle T\omega, \omega \rangle = \frac{\Delta q_T}{2n}.$$

If $[t_{ij}]$ is the matrix of $T$ with respect to $\{e_i\}$, one checks that

$$q_T = \sum_{i,j} t_{ij} x_i x_j,$$
where \( \{x_i\} \) is the dual basis, and that
\[
\Delta q_T = 2 \sum_i t_{ii} = 2 \text{tr}(T),
\]
and (10) follows. It remains to prove that all \( \text{SO}(V, q) \)-invariant functionals on \( Q \) are proportional. Using \( q \) we identify \( V = \tilde{V} \), whence \( Q = \tilde{Q} \), so it is equivalent to prove that all \( \text{SO}(V, q) \)-invariant elements in \( Q \) are proportional. If \( p \) is such, then \( p \) is constant on \( S \), say \( p = c \) on \( S \). Then \( p - cq = 0 \) on \( S \). Since \( p \) and \( q \) are homogenous, we have \( p = cq \) on \( V \), as desired, so the proof of (10) is complete.

The analogue of (10) for \( k = n \) is well-known: \( \det(T) \) is the signed volume of \( T_\omega \), where the sign is + iff \( T \) is orientation-preserving.

The interpolation of these results for \( 1 \leq k \leq n \) is as follows. The oriented Grassmannians \( \tilde{G}_k(V) \) are compact submanifolds of \( \Lambda^k V \) generalizing the unit sphere for \( k = 1 \). The special orthogonal group \( \text{SO}(V, q) \) preserves oriented orthonormal bases, hence acts on \( \tilde{G}_k(V) \). In fact this action is transitive, so there is a unique \( \text{SO}(V) \)-invariant measure assigning unit volume to \( \tilde{G}_k(V) \). Using this measure, we can compute the average of continuous functions on \( \tilde{G}_k(V) \).

Given \( T \in \text{End}(V) \) and an oriented \( k \)-plane \( \omega \in \tilde{G}_k(V) \), we define a number \( \langle T_\omega, \omega \rangle \) as follows. If \( \dim T_\omega < k \) then \( \langle T_\omega, \omega \rangle = 0 \). Assume \( \dim T_\omega = k \) and give \( T_\omega \) the orientation transferred to \( \omega \) by \( T \). Then \( \langle T_\omega, \omega \rangle \) is \( \pm \) the \( k \)-dimensional volume of the orthogonal projection of \( T_\omega \) back to \( \omega \), where we choose \( + \) if \( T : \omega \to T_\omega \) is orientation preserving. Thus, for any \( T \in \text{End}(V) \) we have defined a function \( \omega \mapsto \langle T_\omega, \omega \rangle \) on \( \tilde{G}_k(V) \). Then we have
\[
\left( \frac{n}{k} \right)^{-1} \cdot \text{tr}(T, \Lambda^k V) = \int_{\tilde{G}_k(V)} \langle T_\omega, \omega \rangle.
\]
That is, the average of the eigenvalues of \( T \) on \( \Lambda^k V \) is the average of the signed volumes \( \langle T_\omega, \omega \rangle \) over all oriented \( k \)-planes in \( V \). This result can be proved in the same way as (10), by analyzing the action of \( \text{SO}(V, q) \) on the space of quadratic forms on \( \Lambda^k V \).

## 4 Clifford algebras

This is our third quotient of the tensor algebra. Let \( F \) be an infinite field. Let \( V \) be a vector space over \( F \), with \( \text{dim}_F(V) = n < \infty \). A **quadratic form** is a polynomial function \( q : V \to F \) of degree two. That is, \( q \in S^2(V) \). Let \( I(V, q) \) be the ideal in the tensor algebra \( T(V) \) generated by \( \{ v \otimes v - q(v) : v \in V \} \). The **Clifford algebra** is the \( F \)-algebra
\[
C(V, q) := T(V)/I(V, q).
\]

We see immediately that \( C(V, q) \) is a deformation of the exterior algebra \( \Lambda(V) = C(V, 0) \), but more is true: Since \( I(V, q) \) is not homogeneous, the Clifford algebra is filtered but not graded, see section 1.1. From (3) the associated graded algebra of \( C(V, q) \) is determined by the ideal \( \sigma(I(V, q)) \) generated by all symbols \( \sigma(v \otimes v - q(v)) = v \otimes v \), we have
\[
\sigma(I(V, q)) = I_{\text{alt}}(V),
\]
so that
\[ \text{gr } C(V, q) \cong \Lambda(V), \]
as graded algebras. It follows that if \( \{e_1, \ldots, e_n\} \) is a basis of \( V \) then the products \( \{e_I : I \subset [1, n]\} \) form an \( F \)-basis of \( C(V, q) \), where
\[ e_I = \prod_{i \in I} e_i, \]
product taken in increasing order of the subscripts, with \( e_\varnothing = 1 \).

Though the terms in each generator of \( I(V, q) \) do not have the same degree, their degrees do have the same parity (even). This implies that \( C(V, q) \) does have a \( \mathbb{Z}/2\mathbb{Z} \)-grading. Namely, for \( a \in \mathbb{Z}/2\mathbb{Z} \), let \( C_a(V, q) \) be the projection to \( C(V, q) \) of all \( T^i(V) \) with \( i \equiv a \mod 2 \). We then have
\[ C_a(V, q) \cdot C_b(V, q) \subset C_{a+b}(V, q). \]

Note that \( C_0(V, q) \) is a subalgebra of \( C_0(V, q) \), with \( F \)-basis \( \{e_I : |I| \text{ even}\} \), and that \( C_1(V, q) \) is a \( C_0(V, q) \)-submodule of \( C(V, q) \), with \( F \)-basis \( \{e_I : |I| \text{ odd}\} \).

Associated to \( q \) is the bilinear form
\[ b(u, v) = q(u + v) - q(u) - q(v). \]

In \( C(V, q) \) we have the relation
\[ q(u + v) = (u + v)^2 = q(u) + uv + vu + q(v), \]
hence the relation
\[ uv + vu = b(u, v). \]
In particular, we have \( uv = -vu \) if \( b(u, v) = 0 \).

### 4.1 Orthogonal bases and the center of the Clifford algebra

From now on we also assume that \( 2 \neq 0 \) in \( F \) and that \( q \) is non-degenerate, meaning that if \( u \in V \) and \( b(u, V) = \{0\} \) then \( u = 0 \).

Under these conditions, one can show that \( V \) has an orthogonal basis \( \{e_1, \ldots, e_n\} \), meaning that \( b(e_i, e_j) = 0 \) if \( i \neq j \). Non-degeneracy implies that \( q(e_i) \neq 0 \) for all \( i \). From our remarks above, the vectors \( \{e_I : I \subset [1, n]\} \) form an \( F \)-basis of \( C(V, q) \). By induction on \( I \), one checks that
\[ e_I e_J = (-1)^{|I \cap J|} e_J e_I, \]
where \( |I \cap J| = |I| \cdot |J| + |I \cap J| \).

Of particular importance is the element
\[ \omega := e_{[1,n]} = e_1 e_2 \cdots e_n. \]
One checks that
\[ \omega^2 = D, \]
where \( D \) is the discriminant of the quadratic form \( q \), given by
\[ D = (-1)^{n(n-1)/2} q(e_1)q(e_2) \cdots q(e_n). \]
(This is independent of the choice of orthogonal basis.) Thus, inside of \( C(V,q) \) we have a quadratic algebra
\[ F[\omega] = \{ a + b \omega : a, b \in F \} \cong F[x]/(x^2 - D). \]
This algebra is closely related to the center of \( C(V,q) \), as follows.

Let \( Z(V,q) \) be the center of the algebra \( C(V,q) \) and let \( Z_0(V,q) \) be the center of the even subalgebra \( C_0(V,q) \). From (12) it follows that
\[ Z(V,q) = \begin{cases} F[\omega] & \text{for } n \text{ odd} \\ F & \text{for } n \text{ even} \end{cases}, \quad Z_0(V,q) = \begin{cases} F & \text{for } n \text{ odd} \\ F[\omega] & \text{for } n \text{ even} \end{cases}. \tag{13} \]
In all cases we have
\[ Z(V,q) \cap C_0(V,q) = F. \tag{14} \]

An \( F \)-algebra \( A \) is called central-simple if its center is just \( F \). The Skolem-Noether theorem asserts that every automorphism of a central simple \( F \)-algebra is conjugation by a unit in \( A^\times \).

Each formula in (13) accords with the Skolem-Noether theorem, as follows. The algebra \( C(V,q) \) has a unique automorphism \( \alpha \) such that
\[ \alpha(v_1 \cdots v_k) = (-1)^k v_1 \cdots v_k \quad \text{for all } v_1, \ldots, v_k \in V. \tag{15} \]
When \( n \) is even then \( C(V,q) \) is central-simple and \( \alpha \) is conjugation by \( \omega \in C(V,q)^\times \). When \( n \) is odd then \( C_0(V,q) \) is central-simple and both \( \alpha \) and conjugation-by-\( \omega \) are trivial automorphisms of \( C_0(V,q) \).

### 4.2 Orthogonal groups

We have seen that the group \( \text{SO}_3 \) of rotations of three-dimensional Euclidean space has the two-fold covering \( S^3 \to \text{SO}_3 \) by the unit sphere \( S^3 \) in the quaternion algebra \( \mathbb{H} \). This generalizes to higher dimensions and general fields \( F \), in the setting of Clifford algebras \( C(V,q) \). First we need to prove that orthogonal groups are generated by reflections.

We continue to assume that \( 2 \neq 0 \) in \( F \) and that \( q \) is nondegenerate. The **orthogonal group** is the linear isometry group of \( q \):
\[ \text{O}(V,q) = \{ g \in \text{GL}(V) : q(gv) = q(v) \ \forall v \in V \}. \]
If \( v \in V \) and \( q(v) \neq 0 \) then \( V = Fv \oplus v^\perp \), where \( v^\perp = \{ u \in V : b(v,u) = 0 \} \) is the hyperplane orthogonal to \( v \). There is a unique element \( r_v \in \text{O}(V,q) \) such that \( r_v(v) = -v \) and \( r_v(u) = u \) for all \( u \in v^\perp \). We call \( r_v \) the reflection about \( v^\perp \). An explicit formula is
\[ r_v(x) = x - \frac{\langle x, v \rangle}{q(v)} v, \quad \forall x \in V. \]
Proposition 4.1 The group $O(V, q)$ is generated by reflections. In fact each $g \in O(V, q)$ can be expressed as a product of at most $n$ reflections.

Proof: We only prove the first statement here, since that is all we need. This proof is from Dieudonné “sur les groupes classiques”. It follows an earlier proof of É. Cartan’s in the real case. We only prove the first statement here.

We argue by induction on $n = \dim V$. If $n = 1$ then $O(V, q) = \{\pm 1\}$ and $-1$ is the unique reflection on $V$.

Let $g \in O(V, q)$. Since $q$ is nondegenerate there exists $v \in V$ with $q(v) \neq 0$.

Case 1: Suppose $gv = v$. Then $g$ preserves $v^\perp$. The restriction of $q$ to $v^\perp$ is nondegenerate and the restriction $\bar{g}$ of $g$ to $v^\perp$ belongs to $O(v^\perp, q)$. By induction, $\bar{g} = \bar{r}_1 \cdots \bar{r}_k$, for some reflections $\bar{r}_i \in O(v^\perp, q)$. For each $i$ define $r_i \in O(V, q)$ by $r_i(v) = v$ and $r_i = \bar{r}_i$ on $v^\perp$. Then $r_1 \cdots r_k$ acts on $v$ and $v^\perp$ as $g$ does, so $g = r_1 \cdots r_k$.

From now on we assume $gv \neq v$. We consider the vectors $v \pm gv$. Since

$$q(v \pm gv) = 2q(v) \pm \langle v, gv \rangle$$

and $q(v) \neq 0$, at least one of $q(v \pm gv)$ is nonzero.

Case 2: Suppose $q(v - gv) \neq 0$. Let

$$u = v - gv, \quad u' = v + gv.$$  

Then $\langle u, u' \rangle = 0$, and we have $r_u(u) = -u$ and $r_u(u') = u'$. Let $x = r_u(v)$ and $y = r_u(gv)$. Then

$$gv - v = -u = r_u(u) = x - y \quad \text{and} \quad v + gv = u' = r_u(u') = x + y.$$  

Solving these equations for $x, y$ we get $x = gv, y = v$. Thus, $r_u$ interchanges $v$ and $gv$. Hence $r_ug$ fixes $v$ and we are reduced to Case 1.

Case 3: Suppose $q(v + gv) \neq 0$. Letting $u = v + gv, u' = v - gv$ and proceeding as in Case 2, we find that $r_u(gv) = -v$. Hence $r_ug$ fixes $v$, returning us to Case 1.

In Case 3 a more intricate argument is needed to control the number of reflections required to express $g$. If $F = \mathbb{R}$ and $q$ is positive definite, then Case 3 never arises, so in this case the above proof shows that $g$ is a product of at most $n$ reflections.

4.3 The Clifford group

The Clifford group is defined as

$$\Gamma(V, q) = \{g \in C(V, q)^\times : gV \alpha(g)^{-1} \subset V\},$$
where $\alpha$ is the involution defined in (15). Here we are identifying $V$ with its image in $C(V,q)$, and the product

$$R_g(v) := gv\alpha(g)^{-1}$$

is taken in $C(V,q)$. Thus, we have a group homomorphism

$$R : \Gamma(V,q) \longrightarrow \text{GL}(V) \quad g \mapsto R_g.$$  

I claim that

$$\ker R = F^\times.$$  

It is clear that $F^\times \subset \ker R$. Suppose $g = g_0 + g_1 \in \ker R$, with $g_a \in C_a(V,q)$ for $a \in \mathbb{Z}/2\mathbb{Z}$. This means for every $v \in V$ we have

$$(g_0 + g_1)v = v(g_0 - g_1).$$

In odd degree, this says $g_0v = v_0$, so $g_0 \in F^\times$ by (14). In even degree, this says $g_1v = -vg_1$, so $g_1 = 0$ by (12). Hence $\ker R = F^\times$, as claimed.

I next claim that the image of $R$ lies in $O(V,q)$. For if $g\alpha(g)^{-1} \in V$ then it is negated by $\alpha$ and we have

$$-g\alpha(g)^{-1} = \alpha(g\alpha(g)^{-1}) = \alpha(g)\alpha(v)^{-1} = -\alpha(g)v^{-1},$$

so

$$q(g\alpha(g)^{-1}) = (g\alpha(g)^{-1})^2 = g\alpha(g)^{-1} \cdot \alpha(g)v^{-1} = gv^2g^{-1} = q(v).$$

Therefore $R_g \in O(V,q)$ for all $g \in \Gamma(V,q)$, as claimed. If $q(u) \neq 0$ then $R_u = r_u$ is the reflection negating $u$ and fixing the hyperplane $u^\perp$. By Prop. 4.1, the group $O(V,q)$ is generated by such reflections $r_u$. It follows that $R$ is surjective, and also that $\Gamma(V,q)$ is generated, as a group, by $\{v \in V : q(v) \neq 0\}$. Summarizing, we have exact sequences

$$1 \longrightarrow F^\times \longrightarrow \Gamma(V,q) \xrightarrow{R} O(V,q) \longrightarrow 1$$

$$1 \longrightarrow F^\times \longrightarrow \Gamma_0(V,q) \xrightarrow{R} SO(V,q) \longrightarrow 1,$$

where $\Gamma_0(V,q) = \Gamma(V,q) \cap C_0(V,q)^\times$. Every element of $\Gamma(V,q)$ may be expressed as a product $v_1 \cdots v_k$ of vectors $v_i \in V$ having all $q(v_i) \neq 0$, and $\Gamma_0(V,q)$ consists of such products with $k$ even.

### 4.4 Clifford conjugation and the spinor norm

Recall that the quaternion algebra $\mathbb{H}$ has a conjugation operation $x \mapsto \bar{x}$, satisfying $xy = \bar{y}\bar{x}$. An endomorphism of an algebra that reverses the order of multiplication like this is called an antiautomorphism.

On a general Clifford algebra $C(V,q)$, **Clifford conjugation**, denoted $x \mapsto x^*$, is the unique antiautomorphism which is $-1$ on $V$. On any product of vectors we have

$$(v_1 \cdots v_k)^* = (-1)^k v_k \cdots v_1.$$
For example if \( n = 3 \), \( F = \mathbb{R} \) and \( \{e_1, e_2, e_3\} \) is a \( q \)-orthonormal basis of \( \mathbb{R}^3 \), then \( \mathbb{H} \cong C_0(V,q) \), via
\[
i \mapsto e_1e_2, \quad j \mapsto e_2e_3, \quad k \mapsto e_3e_1,
\]
and each \( e_p e_q \) is negated by \( * \). Thus, Clifford conjugation on \( C(V,q) \) restricts to quaternionic conjugation on \( C_0(V,q) \), in this case.

Returning to the general case, we define the **norm**
\[
N : \Gamma(V,q) \to F^\times, \quad N(g) = gg^*.
\]
To see that this makes sense, recall that every element \( g \in \Gamma(V,q) \) can be written as
\[
g = v_1 \cdots v_k
\]
where all \( v_i \in V \) have \( q(v_i) \neq 0 \). We have
\[
N(g) = gg^* = q(v_1) \cdots q(v_k) \in F^\times.
\]
In particular each product \( gg^* \) is central in \( \Gamma(V,q) \). It follows that \( N \) is a group homomorphism
\[
N : \Gamma(V,q) \to F^\times.
\]
If \( g \in F^\times \) then \( N(g) = g^2 \in F^{\times2} \), so \( N \) induces a group homomorphism
\[
N : O(V,q) \to F^\times/F^{\times2}
\]
called the **spinor norm**. The spinor norm is determined by the formula
\[
N(r_1 \cdots r_k) = q(v_1) \cdots q(v_k),
\]
where each \( r_i = r_{v_i} \) is a reflection for some \( v_i \in V \) with \( q(v_i) \neq 0 \).

If \( gg^* = c^2 \in F^{\times2} \) then \( (g/c)(g/c)^* = 1 \). This means the subgroup of \( \Gamma(V,q) \) mapping to \( F^{\times2} \) under \( N \) is \( F^\times \cdot \text{Pin}(V,q) \), where
\[
\text{Pin}(V,q) = \{g \in \Gamma(V,q) : N(g) = 1\}.
\]
Since \( F^\times \cap \text{Pin}(V,q) = \{\pm 1\} \), it follows that we have an exact sequence
\[
1 \to \{\pm 1\} \to \text{Pin}(V,q) \xrightarrow{R} O(V,q) \xrightarrow{N} F^\times/F^{\times2}.
\]  

(17)

### 4.5 Spin groups

Restricting further to even elements, we define
\[
\text{Spin}(V,q) = \{g \in \Gamma_0(V,q) : N(g) = 1\} = \text{Pin}(V,q) \cap \Gamma_0(V,q).
\]
Then (17) restricts to the exact sequence

$$1 \rightarrow \{\pm 1\} \rightarrow \text{Spin}(V, q) \xrightarrow{R} \text{SO}(V, q) \xrightarrow{N} F^\times / F^\times 2.$$  

(18)

For example, if $q$ is a positive definite quadratic form on $\mathbb{R}^n$ then $N$ is trivial. In this case, (18) becomes an exact sequence

$$1 \rightarrow \{\pm 1\} \rightarrow \text{Spin}(V, q) \xrightarrow{R} \text{SO}(V, q) \xrightarrow{} 1,$$

(19)

which is the universal covering of $\text{SO}(V, q)$, in the sense of topology. For $n = 3$ we have $C_0(V, q) = \mathbb{H}$ and $\text{Spin}(V, q) = S^3$; we recover the exact sequence

$$1 \rightarrow \{\pm 1\} \rightarrow S^3 \xrightarrow{R} \text{SO}_3 \rightarrow 1,$$

obtained previously.

Likewise if $F$ is algebraically closed then $F^\times = F^\times 2$, so we again obtain (19), which is the universal covering of $\text{SO}(V, q)$, in the sense of algebraic groups. For $n = 3$ we have $C_0(V, q) = M_{2\times 2}(F)$ and $\text{Spin}(V, q) = \text{SL}_2(F)$; we get the exact sequence

$$1 \rightarrow \{\pm 1\} \rightarrow \text{SL}_2(F) \xrightarrow{R} \text{SO}_3(F) \rightarrow 1.$$

For general fields $F$ and quadratic forms $q$, we have the **Witt decomposition** of $V$. This is a direct-sum decomposition

$$V = W + V_1 + W',$$

where $q$ is identically zero on $W$ and $W'$ and is never zero on $V_1 - \{0\}$. A Witt decomposition is unique up to a transformation from $O(V, q)$. Note that nondegeneracy of $q$ implies that the bilinear form $b$ induces an isomorphism $W' \simeq \text{Hom}(W, F)$. Hence, given a basis $\{w_1, \ldots, w_r\}$ of $W$, there exists a basis $\{w_1, \ldots, w_r\}$ of $W'$ such that $b(w_i, w_j) = 1$ if $i = j$, zero otherwise. Set $\omega_i = w_iw_{-i} \in C_0(V, q)$. One checks that $\omega_i^2 = \omega_i$, that for each $t \in F^\times$ the element

$$a_i(t) := 1 + (t - 1)\omega_i \in C_0(V, q)$$

actually belongs to $\text{Spin}(V, q)$, and that $a_i(t)a_i(s) = a_i(ts)$ for all $t, s \in F^\times$. Thus for each $i \in [1, r]$ we have an injective group homomorphism

$$a_i : F^\times \rightarrow \text{Spin}(V, q).$$

One further checks that

$$a_i(t) \cdot w_j \cdot a_i(t)^{-1} = \begin{cases} 
    t^{\pm 1}w_i & \text{if } i = \pm j \\
    w_j & \text{else}.
\end{cases}$$

Thus we have an abelian subgroup of $\text{Spin}(V, q)$:

$$A = \langle a_i(t) : i \in [1, m], \ t \in F^\times \rangle$$

which acts diagonalizably on $V$, trivially on $V_1$. 

22
4.6 Summary

All of the above groups are summarized by the following commutative diagram of maps:

\[
\begin{array}{cccccc}
1 & \longrightarrow & F^\times & \longrightarrow & F^\times & \longrightarrow & 1 \\
2 & \uparrow & N & \uparrow & N & & \\
1 & \longrightarrow & F^\times & \longrightarrow & \Gamma(V, q) & \longrightarrow & O(V, q) & \longrightarrow & 1 \\
& & & \uparrow & \uparrow & \uparrow & \uparrow & & \\
1 & \longrightarrow & \{\pm 1\} & \longrightarrow & \text{Pin}(V, q) & \longrightarrow & O(V, q) & N & \longrightarrow & F^\times / F^\times 2 \\
& & & & \uparrow & \uparrow & \uparrow & \uparrow & & \\
1 & \longrightarrow & \{\pm 1\} & \longrightarrow & \text{Spin}(V, q) & \longrightarrow & SO(V, q) & N & \longrightarrow & F^\times / F^\times 2 \\
\end{array}
\]

In this diagram the rows are exact sequences and unlabelled arrows are the obvious injections or surjections.

5 Lie algebras and their enveloping algebras

Enveloping algebras will be our fourth and last quotients of the tensor algebra. Recall that Clifford algebras are deformations of exterior algebras, depending on a quadratic form. We will see that enveloping algebras are deformations of symmetric algebras, depending on a Lie bracket.

A **Lie algebra** over a field \( F \) is an \( F \)-vector space \( L \) and a map

\[ L \wedge L \longrightarrow L, \quad x \wedge y \mapsto [x, y] \]

such that for all \( x \in L \), the endomorphism of \( L \) sending \( y \mapsto [x, y] \) is a derivation.

Explicitly, these conditions mean that the **bracket** \([x, y]\) is bilinear and alternating and that

\[ [x, [y, z]] = [[x, y], z] + [y, [x, z]] \]

for all \( x, y, z \in L \). This last condition is often written as

\[ [x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0, \]

and is called the **Jacobi identity**.

We could have \([x, y] = 0\) for all \( x, y \in L \). Such Lie algebras are called **abelian**. Every one-dimensional Lie algebra is abelian.

A fundamental example of a Lie algebra is \( L = \text{End}(V) \), where \( V \) is a vector space over \( F \) and the bracket is given by \([X, Y] = XY - YX\) for \( X, Y \in \text{End}(V) \). More generally, any associative \( F \)-algebra \( A \) can be made into a Lie algebra, with bracket \([a, b] = ab - ba\), for \( a, b \in A \).

A **homomorphism** between Lie algebras \( L, L' \) is a linear mapping \( \phi : L \rightarrow L' \) such that \( \phi([x, y]) = [\phi(x), \phi(y)] \) for all \( x, y \in L \).
5.1 Origin and examples: Lie groups

Lie algebras arose classically, and acquired their name, as follows. Let $M$ be a smooth manifold and let $S(M)$ be the $\mathbb{R}$-algebra of smooth functions $f : M \to \mathbb{R}$. A vector field on $M$ is a derivation of $S(M)$. This means $X : S(M) \to S(M)$ is a linear endomorphism satisfying the product rule: $X(fg) = X(f)g + fX(g)$ for all $f, g \in S$. A linear combination of vector fields $X, Y$ is again a vector field, but the composition $XY = X \circ Y$ need not be a vector field. However, one can check that $XY - YX$ is a vector field, and that the vector space $\mathfrak{X}(M)$ of vector fields on $M$ becomes a Lie algebra with bracket $[X, Y] = XY - YX$. If $\dim M > 0$ then $\mathfrak{X}(M)$ is infinite-dimensional.

Suppose now that $G$ is a Lie group. This means $G$ is a smooth manifold with a smooth group structure. We have seen many examples of Lie groups already, such as $\text{GL}(V)$, $\text{SL}(V)$, $\text{PGL}(V)$, $\text{SO}(V, q)$ and $\text{Spin}(V, q)$ where in all cases $V$ is a finite-dimensional vector space over the fields $\mathbb{R}$ or $\mathbb{C}$.

By definition, for each $g \in G$ we have a smooth map $L_g : G \to G$ given by $x \mapsto gx$. A vector field $X \in \mathfrak{X}(G)$ is left-invariant if

$$X(f \circ L_g) = (Xf) \circ L_g, \quad \forall f \in S(G).$$

The subspace $\mathfrak{Lie}(G) \subset \mathfrak{X}(G)$ of left-invariant vector fields is closed under the bracket, hence is a Lie algebra. As a linear space, $\mathfrak{Lie}(G)$ is canonically isomorphic to $T_e(G)$, the tangent space to $G$ at the identity element $e$, so $\mathfrak{Lie}(G)$ is a finite dimensional Lie algebra. Also, any basis of $\mathfrak{Lie}(G)$ freely generates $\mathfrak{X}(G)$ as an $S(G)$-module.  

For the Lie groups above, we have the table

<table>
<thead>
<tr>
<th>$G$</th>
<th>$\text{Lie}(G)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{GL}(V)$</td>
<td>$\text{End}(V)$</td>
</tr>
<tr>
<td>$\text{SL}(V)$</td>
<td>$\mathfrak{sl}(V)$</td>
</tr>
<tr>
<td>$\text{PGL}(V)$</td>
<td>$\mathfrak{sl}(V)$</td>
</tr>
<tr>
<td>$\text{SO}(V, q)$</td>
<td>$\mathfrak{so}(V, q)$</td>
</tr>
<tr>
<td>$\text{Spin}(V, q)$</td>
<td>$\mathfrak{so}(V, q)$</td>
</tr>
</tbody>
</table>

where

$$\mathfrak{sl}(V) = \{ X \in \text{End}(V) : \text{tr}(X) = 0 \},$$

$$\mathfrak{so}(V, q) = \{ X \in \text{End}(V) : b(Xu, v) + b(u, Xv) = 0 \quad \forall u, v \in V \},$$

where $b$ is the bilinear form associated to $q$. We note that different Lie groups $G, G'$ can have $\mathfrak{Lie}(G) \simeq \mathfrak{Lie}(G')$. This happens precisely when there is a smooth homomorphism $G \to G'$ with discrete kernel; for such a map induces an isomorphism $T_e(G) \cong T_e(G')$ on tangent spaces.

5.2 Universal enveloping algebras

We return to an arbitrary field $F$. Given a Lie algebra $L$ there is a canonical associative $F$-algebra $U(L)$ containing $L$ such that the bracket in $L$ is given by $[x, y] = xy - yx$ where the products are taken

---

7For manifolds which are not Lie groups, $\mathfrak{X}(M)$ need not be free over $S(M)$; one example is the two-sphere $M = S^2$. However, $\mathfrak{X}(M)$ is always a projective module over $S(M)$.  

---
in $U(L)$.

The algebra $U(L)$ is the **universal enveloping algebra** of $L$, and is constructed as the quotient

$$U(L) = T(L)/I_{\text{lie}}(L),$$

where $T(L)$ be the tensor algebra of the $F$-vector space $L$ and $I_{\text{lie}}(L)$ is the ideal generated by

$$\{ x \otimes y - y \otimes x - [x, y] : x, y \in L \}.$$ 

If $L$ is abelian then $I_{\text{lie}}(L) = I_{\text{sym}}(L)$ is the ideal defining the symmetric algebra $\text{Sym}(L)$, so

$$U(L) = \text{Sym}(L)$$

is just the symmetric algebra of $L$. For any bracket, the symbol ideal of $\sigma(I_{\text{lie}}(L)) = I_{\text{sym}}(L)$, so we have

$$\text{gr}(U(L)) = \text{Sym}(L),$$

as graded algebras. This implies that if $\{x_1, \ldots, x_n\}$ is an ordered basis of $L$ then $U(L)$ has a basis consisting of monomials

$$x_1^{i_1}x_2^{i_2} \cdots x_n^{i_n},$$

where each $i_j \in \{0, 1, 2, 3, \ldots\}$.

In particular the inclusion $L \subset T(L)$ gives an injection $\iota : L \rightarrow U(L)$, by which we identify $L$ with a Lie subalgebra of $U(L)$, where the latter has the bracket $[x, y] = xy - yx$, arising from its structure as associative algebra.

The mapping property of $U(L)$ is as follows. If $A$ is an associative $F$-algebra and

$$\phi : L \rightarrow A$$

is a Lie algebra homomorphism then there is a unique $F$-algebra homomorphism $\tilde{\phi} : U(L) \rightarrow A$ such that $\tilde{\phi} \circ \iota = \phi$.

### 5.3 Representations and $U(L)$-modules

A **representation** of a Lie algebra $L$ is a Lie algebra homomorphism

$$\rho : L \rightarrow \text{End}(V),$$

where $V$ is a $F$-vector space. By the mapping property for enveloping algebras, $\rho$ extends to give a $U(L)$-module

$$\tilde{\rho} : U(L) \rightarrow \text{End}(V).$$

We say the $\rho$ is **irreducible** if the corresponding $U(L)$-module is simple.

The image of $\tilde{\phi}$ is the subalgebra of $\text{End}(V)$ generated by $\phi(L)$, hence the term “enveloping” algebra. The advantage of $U(L)$ over $L$ is that $U(L)$ gives more endomorphisms of $V$. For example, $U(L)$ can have a large center, which helps to analyze the structure of the representation $(\rho, V)$.

---

8The existence of such a basis is known as the **PBW theorem**, after H. Poincaré, G. Birkhoff and E. Witt.
Any representation $\rho : L \rightarrow \text{End}(V)$ gives rise to more representations, via the natural Lie algebra homomorphisms

$$
\begin{align*}
\text{End}(V) & \rightarrow \text{End}(T(V)) \\
\text{End}(V) & \rightarrow \text{End}(\text{Sym}(V)) \\
\text{End}(V) & \rightarrow \text{End}(\Lambda(V)),
\end{align*}
$$

sending $X \mapsto \partial_X$, where $\partial_X$ is the derivation of $T(V), \text{Sym}(V)$ or $\Lambda(V)$ such that

$$
\partial_X(v_1v_2\cdots v_k) = \sum_{i=1}^{k} v_1v_2\cdots v_{i-1}(Xv_i)v_{i+1}\cdots v_k, \tag{21}
$$

the products being taken in $T(V), \text{Sym}(V)$ or $\Lambda(V)$.

### 5.4 Representations of $\mathfrak{sl}_2$

We now consider the most fundamental of all Lie algebras, $\mathfrak{sl}_2$. This is the Lie algebra with $\mathbb{C}$-basis $\{e, h, f\}$ and bracket

$$
[h, e] = 2e, \quad [e, f] = h, \quad [f, h] = 2f. \tag{22}
$$

The enveloping algebra $U(\mathfrak{sl}_2)$ has the basis

$$\{ f^i h^j e^k : i, j, k = 0, 1, 2, \ldots \}.$$

To give a representation of $\mathfrak{sl}_2$, or a module for $U(\mathfrak{sl}_2)$ it is equivalent to give a vector space $V$ over $F$ and three endomorphisms $E, H, F \in \text{End}(V)$ satisfying the corresponding relations (22). By the mapping property we will then have a homomorphism $\tilde{\rho} : U(\mathfrak{sl}_2) \rightarrow \text{End}(V)$ such that $\tilde{\rho}(f^i h^j e^k) = F^i H^j E^k$.

We will classify representations $\rho : \mathfrak{sl}_2 \rightarrow \text{End}(V)$ where $\dim V$ is finite. It is a fact, which we shall not prove, that every such representation is a direct sum (in the evident way) of irreducible representations. Therefore it is enough to classify irreducible representations.

The standard representation of $\mathfrak{sl}_2$ has a two-dimensional vector space $V_1$ with ordered basis $\{u, v\}$ and operators

$$
\begin{align*}
E_1 & = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, & H_1 & = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, & F_1 & = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},
\end{align*}
$$

in terms of the basis. To see this is irreducible, note that any proper submodule would have to be one-dimensional over $\mathbb{C}$, and preserved by $h$. Since $h$ has distinct eigenvalues $\pm 1$ on $V_1$, the submodule would have be either $\mathbb{C}u$ or $\mathbb{C}v$. But these lines are not preserved by both $e$ and $f$.

Now suppose $\rho : \mathfrak{sl}_2 \rightarrow \text{End}(V)$ is an arbitrary irreducible representation with $\dim \mathbb{C} V$ finite, and let $E = \rho(e), H = \rho(h), F = \rho(f)$, as above. For every $\lambda \in \mathbb{C}$, let

$$
V(\lambda) = \{ v \in V : Hv = \lambda v \}$$

26
The relations (22) imply that
\[ E^k V(\lambda) \subset V(\lambda + 2k), \quad F^i V(\lambda) \subset V(\lambda - 2i). \]

Since \( V \) is finite dimensional and \( \mathbb{C} \) is algebraically closed, every operator in \( \text{End}(V) \) has an eigenvalue, so \( V(\lambda) \neq 0 \) for some \( \lambda \in \mathbb{C} \). But only finitely many \( V(\lambda) \) can be nonzero, since eigenspaces for distinct eigenvalues are linearly independent. Hence there is some \( \lambda \in \mathbb{C} \) with \( V(\lambda) \neq 0 \), but \( EV(\lambda) = 0 \).

Choose a nonzero vector \( w \in V(\lambda) \) again since \( \dim V \) is finite, there is an integer \( n \) such that \( F^nw \neq 0 \), but \( F^{n+1}w = 0 \). Let \( W \) be the span of \( \{w, Fw, F^2w, \ldots, F^nw\} \). We clearly have \( FW \subset W \). Also each \( F^iw \) belongs to \( V(\lambda - 2i) \), hence is an eigenvector for \( H \), so \( HW \subset W \). I claim that \( EW \subset W \), but this is less clear. In \( U(sl_2) \) the middle relation of (22) can be written as \( ef = fe + h \). By induction we have
\[ ef^i = f^i e + f^{i-1}h + f^{i-2}hf + \cdots + fhf^{i-2} + hf^{i-1}, \]
in \( U(sl_2) \). It follows that for any \( i \geq 1 \) we have
\[ EF^i w = i(\lambda - i + 1)F^{i-1}w, \]
in \( \text{End}(V) \). So \( W \) is indeed preserved by \( E \), so \( W = V \), by irreducibility. Since \( F^{n+1}w = 0 \), this also shows that
\[ 0 = EF^{n+1}w = (n + 1)(\lambda - n)F^n w, \]
and since \( F^n w \neq 0 \), we must have \( \lambda = n \). It follows that the actions of \( E, H, F \) are given in terms of the ordered basis \( \{w, Fw, \ldots, F^n w\} \) by the formulas
\[ EF^i w = i(n - i + 1)F^{i-1}w, \quad HF^i w = (n - 2i)F^{i}w, \quad FF^i w = F^{i+1}w, \quad (23) \]
and \( F^{n+1}w = 0 \). We have shown that for each integer \( n \geq 1 \) there is at most one simple \( U(sl_2) \)-module \( V_n \) with \( \dim V_n = n + 1 \). It remains to actually construct such a module. This could be done by checking that the formulas (23) satisfy the relations (22), but there is a more natural way to construct \( V_n \), namely
\[ V_n = \text{Sym}^n V_1, \]
where \( V_1 \) is the standard representation described above and the operators \( E_1, H_1, F_1 \) are extended to \( V_n \) as in (21). In this picture, the ordered basis \( \{w, Fw, \ldots, F^n w\} \) becomes \( \{u^n, u^{n-1}v, \ldots, uv^{n-1}, v^n\} \).
Note that we could even take \( n = 0 \), in which case \( V_0 = \mathbb{C} \) with \( e, h, f \) acting by zero.

To summarize: The Lie algebra \( sl_2 \) has exactly one irreducible complex representation, up to isomorphism, in each finite dimension. They can be realized on the symmetric powers \( V_n \) of the standard two-dimensional representation \( V_1 \).

Finite dimensional representations of \( sl_2 \) occur throughout mathematics. We shall see them again in the classification of conjugacy classes in general linear groups.
Part II

Modules over Principal Ideal Domains

In this section we consider the structure of finitely generated $R$-modules, under the assumption that $R$ is a Principal Ideal Domain. This will lead to a classification of orbits in $M_{m\times n}(R)$ under $\text{GL}_m(R) \times \text{GL}_n(R)$. Taking $R = F[X]$, a polynomial ring over a field $F$, this will lead to a classification of conjugacy classes in $\text{GL}_n(F)$.

5.5 Modules over Noetherian Rings

In this section $R$ is just a commutative ring with $1$. An $R$-module $M$ is Noetherian if every $R$-submodule of $N$ is finitely generated.

**Proposition 5.1** An $R$-module $M$ is Noetherian if and only if every ascending chain of submodules $M_1 \subset M_2 \subset \cdots$ eventually stabilizes.

**Proof:** The proof is similar to the analogous result for Noetherian rings, and is omitted. ■

**Lemma 5.2** Let $M$ be an $R$-module and let $N \subset M$ be an $R$-submodule. Then $M$ is Noetherian if and only if $N$ and $M/N$ are Noetherian.

**Proof:** Suppose $N$ and $M/N$ are Noetherian and let $M_1 \subset M_2 \subset \cdots$ be a chain of submodules of $M$. Let $\bar{M}_i$ be the image of $M_i$ in $M/N$. By assumption there is an integer $n$ such that $M_i \cap N = M_{i+1} \cap N$ and $\bar{M}_i = \bar{M}_{i+1}$ for all $i \geq n$. If $m \in M_{i+1}$ then there is $m' \in M_i$ such that $m \equiv m' \mod N$. This means $m - m' \in M_{i+1} \cap N = M_i \cap N$, so $m \in M_i$. Hence $M_i = M_{i+1}$ and the chain $M_1 \subset M_2 \subset \cdots$ stabilizes.

Suppose $M$ is Noetherian. Every chain of submodules of $N$ is also a chain in $M$, hence must stabilize. Every chain of submodules of $M/N$ is the image of a chain of submodules of $M$ which stabilizes, so any chain in $M/N$ must stabilize. ■

**Proposition 5.3** If $R$ is Noetherian then every finitely generated $R$-module is Noetherian.

**Proof:** If $M$ is cyclic then $M = R/I$ for some ideal $I$ in $R$, so the result follows from $R$ being Noetherian. Using Lemma 5.2, the result follows by induction on the minimum number of generators of $M$. ■

Assume that $R$ is Noetherian and let $M$ be a finitely generated $R$-module, with generating set $\{x_1, \ldots, x_m\}$. Thus we have a surjective $R$-module map

$$\phi : R^m \longrightarrow M, \quad (r_1, \ldots, r_m) \mapsto r_1 x_1 + \cdots + r_m x_m.$$
By Prop. 5.3 we have that $R^m$ is Noetherian, so the submodule $\ker \phi \subset R^m$ has a finite generating set $\{y_1, \ldots, y_n\}$, giving an $R$-module homomorphism

$$\psi : R^n \longrightarrow R^m, \quad (r_1, \ldots, r_n) \mapsto r_1y_1 + \cdots + r_ny_n,$$

whose image is $\ker \phi$. Let $A \in M_{m \times n}$ be the matrix of $\psi$ with respect to the standard bases of $R^n$ and $R^m$. Thus we have

$$M \simeq R^m/\ker \phi = R^m/\operatorname{im} \psi = R^m/AR^n$$

so the $R$-module $M$ is completely determined by the matrix $A$. In fact, for any $g \in \operatorname{GL}_m(R)$ and $h \in \operatorname{GL}_n(R)$ we have

$$M \simeq R^m/AR^n = R^m/AhR^n \simeq R^m/gAhR^n$$

so the isomorphism class of the $R$-module $M$ depends only on the orbit of $A$ under the group

$$\Gamma_{m \times n} = \operatorname{GL}_m(R) \times \operatorname{GL}_n(R)$$

acting on $M_{m \times n}(R)$ by left and right multiplication. Thus, the classification of $R$-modules with $n$ generators is becomes the problem of classifying the $\Gamma_{m \times n}$ orbits on $M_{m \times n}(R)$ for $m \geq 1$.

We write $A \sim B$ if two matrices $A, B \in M_{m \times n}(R)$ belong to the same $\Gamma_{m \times n}$-orbit. This means there is $g \in \operatorname{GL}_m(R)$ and $h \in \operatorname{GL}_n(R)$ such that $gAh = B$.

### 5.6 Smith Normal Form: statement and consequences

From now on $R$ is a PID. This means $R$ is a commutative integral domain with identity and every ideal in $R$ is generated by just one element.

**Theorem 5.4 (Smith Normal Form)** If $A \in M_{m \times n}(R)$ then there is $k \leq m$ and nonzero elements $f_1, \ldots, f_k$ of $R$, where $f_1 \mid f_2 \mid \cdots \mid f_k$, such that

$$A \sim \begin{bmatrix} f_1 & 0 & 0 & \cdots & 0 \\ 0 & f_2 & 0 & \cdots & 0 \\ 0 & 0 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & & 0 & f_k \end{bmatrix}.$$  

Moreover $r$ is unique and the elements $f_i$ are unique up to multiplication by units in $R$.

The elements $f_1, \ldots, f_k$ are the **invariant factors** of $A$.

In terms of modules, the theorem asserts that if $N$ is a submodule of $R^m$ then there is a basis $v_1, \ldots, v_m$ of $R^m$ and a unique sequence $f_1 \mid f_2 \mid \cdots \mid f_k \in R$ (for some unique $k \leq m$) such that $\{f_1v_1, \ldots, f_kv_k\}$ is a basis of $N$.
Corollary 5.5 If $R$ is a PID and $M$ is an $R$-module generated by $m$ elements. Then

1. There is a unique chain of ideals $I_1 \supset I_2 \supset \cdots \supset I_m$ of $R$ such that
   
   $$M \cong R/I_1 \oplus R/I_2 \oplus \cdots \oplus R/I_m.$$ 

2. The torsion submodule of $M$ is
   
   $$M_{\text{tor}} \cong R/I_1 \oplus R/I_2 \oplus \cdots \oplus R/I_k,$$
   
   where $k = \max\{j : I_j \neq 0\}$; hence $M \cong M_{\text{tor}} \oplus R^{m-k}$.

3. The minimal number of generators of $M$ is $m - \ell$, where $\ell = \max\{j : I_j \neq R\}$.

4. The annihilator of $M_{\text{tor}}$ is the ideal $I_k$.

Proof: We have seen that $M \cong R^m/AR^n$ for some integer $n$ and matrix $A \in M_{m \times n}(R)$. By 5.4, we have $M \cong R^n/BR^m$ where $B$ is the Smith Normal Form of $A$. Taking $I_j = Rf_j$, items 1, 3 and 4 are immediate. Item 3 will be proved in section 5.9.

The number $r = m - k$ is called the rank of $M$. It is intrinsically determined as follows. Let $F$ be the fraction field of $R$. Since $F \otimes_R M_{\text{tor}} = 0$, the rank $r$ is the dimension of the $F$-vector space $F \otimes_R M$, hence is uniquely determined by $M$.

Corollary 5.6 Let $V$ be a finite dimensional real vector space and let $L$ be a lattice in $V$. If $A \in \text{GL}(V)$ preserves $L$ then $[L : AL] = |\det(A)|$.

Proof: With respect to a basis $v_1, \ldots, v_n$ of $L$ we have $A \in M_{n \times n}(\mathbb{Z})$. By Thm. 5.4, there are $g, h \in \text{GL}_n(\mathbb{Z})$ such that $gAh$ is a diagonal matrix with diagonal entries $(f_1, \ldots, f_n)$, where the $f_i$ are the invariant factors of the $\mathbb{Z}$-module $L$. We have

$$L/AL \cong L/gAhL = \bigoplus_{i=0}^{n} \mathbb{Z}/f_i\mathbb{Z},$$

so $[L : AL] = |f_1 \cdots f_n| = |\det(gAh)| = |\det(A)|$.

Applying this result to $I_V - A$ gives an algebraic proof of equation (8) above.

5.7 Orbits of $\text{SL}_n(R)$ on $R^n$

In this section we prove the $m = 1$ case of Thm. 5.4. Recall that $R$ is a PID. Given a vector $v = (v_1, \ldots, v_n) \in R^n$ we let $(v)$ denote the ideal in $R$ generated by $(v_1, \ldots, v_n)$. Since $R$ is a PID we have $(v) = Rd$ for some $d \in R$. 

30
Proposition 5.7 Assume that \( n \geq 2 \). Then for \( u, v \in \mathbb{R}^n \) the following are equivalent:

1. \( u, v \) belong to the same \( \text{SL}_n(\mathbb{R}) \)-orbit;
2. \( u, v \) belong to the same \( \text{GL}_n(\mathbb{R}) \)-orbit;
3. \( (u) = (v) \).

Proof: The implication \((1 \Rightarrow 2)\) is clear.

\((2 \Rightarrow 3)\): If \( u = gv \) for some \( g \in \text{GL}_n(\mathbb{R}) \) then the components of \( gv \) are \( \mathbb{R} \)-linear combinations of the components of \( v \), so \( (gv) \subset (v) \). Replacing \( g \) by \( g^{-1} \) shows the other containment, so we have \( (gv) = (v) \).

\((3 \Rightarrow 1)\): It suffices to show that any \( v \in \mathbb{R}^n \) lies in the orbit of \((d, 0, \ldots, 0)\), where \( Rd = (v) \). Dividing each component of \( v \) by \( d \), we may assume \( d = 1 \). The problem now is to show that if \( (v) = R \) then there is a matrix in \( \text{SL}_n(\mathbb{R}) \) whose first column is \( v \).

Suppose \( n = 2 \), and let \( v = (v_1, v_2) \) have \( (v) = R \). Choose \( x_1, x_2 \) such that \( x_1v_1 + x_2v_2 = 1 \). Then the matrix

\[
\begin{bmatrix}
v_1 & -x_2 \\
v_2 & x_1
\end{bmatrix}
\]

belongs to \( \text{SL}_2(\mathbb{R}) \), proving the result for \( n = 2 \). We now argue by induction. \(^9\)

Assume the result holds for \( n \geq 2 \), and let \( v \in \mathbb{R}^{n+1} \) have \( (v) = R \). Set \( v' = (v_1, \ldots, v_n) \) and let \( (v') = Rd \). Then \( v' = du \) where \( u = (u_1, \ldots, u_n) \in \mathbb{R}^n \) and \( (u) = R \). By induction, there exists a matrix \( h \in \text{SL}_n(\mathbb{R}) \) whose first column is \( u \):

\[
h = [u \; h'],
\]

where \( h' \in M_{n \times n-1}(\mathbb{R}) \). Since \( (d, v_{n+1}) = (v) = R \) there are \( s \) and \( t \) in \( R \) such that \( td - sv_{n+1} = 1 \). Set

\[
g = \begin{bmatrix}
v' & h & su \\
v_{n+1} & 0 & t
\end{bmatrix} = \begin{bmatrix}
du & h' & su \\
v_{n+1} & 0 & t
\end{bmatrix}.
\]

Here \( v' = du \) and \( su \) have size \((n - 1) \times 1\), and \( 0 \) is \( 1 \times (n - 2) \), so \( g \) is an \( n \times n \) matrix whose first column is \( v \). We compute

\[
\det(g) = t \cdot \det[du \; h'] + (-1)^n v_{n+1} \det[h' \; su] = td \cdot \det(h) + (-1)^{n+1} sv_{n+1} \cdot \det(h) = 1.
\]

Hence \( g \) is a matrix in \( \text{SL}_{n+1}(\mathbb{R}) \) whose first column is \( v \). \( \blacksquare \)

Note that the same result holds with the first column replaced by any column. Taking transposes, the result also holds for any row. We have proved Thm. 5.4 for \( m = 1 \).

5.8 Existence of Smith Normal Form

In this section let $A \sim B$ means that $A, B \in M_{m\times n}(R)$ belong to the same $SL_m(R) \times SL_n(R)$-orbit.

**Lemma 5.8** Assume $m \geq 2$. If $A \in M_{m\times n}(R)$ then

$$A \sim \begin{bmatrix} b & 0' \\ 0'' & B \end{bmatrix},$$

where $(\text{row}_1(A), \text{col}_1(A)) \subset Rb$, $B \in M_{m-1,n-1}(R)$, $0'$ is a row vector in $R^{n-1}$, $0''$ is a column vector in $R^{m-1}$, and the ideal $(\text{row}_1(A), \text{col}_1(A)) \subset Rb$.

**Proof:** In this proof all matrices will be partitioned as in the statement of the lemma and $*, *', *''$ will denote variable matrices over $R$ of the correct sizes.

By induction on $m$, we may assume $\text{row}_1(A)$ is nonzero. Applying a permutation matrix $h_0 \in SL_n(R)$ (changing one nonzero entry to $-1$ if necessary to make $\det(h_0) = 1$) we have the matrix

$$A_0 := Ah_0 = \begin{bmatrix} a_0 & *' \\ *'' & * \end{bmatrix}, \quad a_0 \neq 0.$$

We now apply Prop. 5.7 repeatedly:

There is $g_1 \in SL_m(R)$ such that

$$A_1 := g_1A_0 = \begin{bmatrix} a_1 & *' \\ 0'' & * \end{bmatrix}, \quad \text{with } Ra_0 \subset Ra_1 = (\text{col}_1(A_0)).$$

Next there is $h_2 \in SL_n(R)$ such that

$$A_2 := A_1h_2 = \begin{bmatrix} a_2 & 0' \\ *'' & * \end{bmatrix}, \quad \text{with } Ra_1 \subset Ra_2 = (\text{row}_1(A_1)).$$

Next there is $g_2 \in SL_m(R)$ such that

$$A_3 := g_3A_2 = \begin{bmatrix} a_3 & *' \\ 0'' & * \end{bmatrix}, \quad \text{with } Ra_2 \subset Ra_3 = (\text{col}_1(A_2)).$$

Repeating this over and over we get a chain of ideals:

$$Ra_0 \subset Ra_1 \subset Ra_2 \subset \cdots \subset R$$

which must stabilize because $R$ is Noetherian. Thus, there is $i \geq 0$ such that $a_{i+1} = ua_i$ for some unit $u \in R^\times$.

Suppose $i$ is odd. Then

$$A_i = \begin{bmatrix} a_i & *' \\ 0'' & * \end{bmatrix},$$
where now $a_i$ divides each entry of $*'.$ The matrix $h = \begin{bmatrix} 1 & -*'/a_i \\ 0'' & I \end{bmatrix}$ lies in $\text{SL}_n(R)$ and $A_i h$ is of the form asserted in the lemma.

If $i$ is even then

$$A_i = \begin{bmatrix} a_i & 0' \\ *'' & * \end{bmatrix},$$

where $a_i$ divides each entry of $*''$. The matrix $g = \begin{bmatrix} 1 & 0' \\ -*''/a_i & I \end{bmatrix}$ lies in $\text{SL}_m(R)$ and $gA_i$ is of the form asserted in the lemma.

The lemma is proved. ■

**Lemma 5.9** Suppose we have a matrix $\begin{bmatrix} b & 0' \\ 0'' & B \end{bmatrix}$ as in Lemma 5.8, such that some entry of $B$ is not divisible by $b$. Then

$$\begin{bmatrix} b & 0' \\ 0'' & B \end{bmatrix} \sim \begin{bmatrix} c & 0' \\ 0'' & C \end{bmatrix},$$

where $Rb \subsetneq Rc$.

**Proof:** Suppose $v = \text{row}_i(B)$ is a row of $B$ containing an entry not divisible by $b$. Let $g \in \text{SL}_m(R)$ be the matrix sending the standard basis vector $e_i \mapsto e_1 + e_i$ and fixing all the other $e_j$‘s. Then

$$g \begin{bmatrix} b & 0' \\ 0'' & B \end{bmatrix} = \begin{bmatrix} b & v \\ 0'' & B \end{bmatrix}.$$ 

By Lemma 5.8 we have

$$\begin{bmatrix} b & v \\ 0'' & B \end{bmatrix} \sim \begin{bmatrix} c & 0' \\ 0'' & C \end{bmatrix},$$

where $Rb$ and $(v)$ are contained in $Rc$. Since $(v)$ is not contained in $Rb$ we have $Rb \subsetneq Rb + (v) \subsetneq Rc$. ■

Now let $A \in M_{m \times n}(R)$. Applying Lemmas 5.8 and 5.9 alternately, we have

$$A \sim \begin{bmatrix} b & 0' \\ 0'' & B \end{bmatrix} \sim \begin{bmatrix} c & 0' \\ 0'' & C \end{bmatrix} \sim \begin{bmatrix} d & 0' \\ 0'' & D \end{bmatrix} \sim \ldots$$

with proper containments of ideals $Rb \subsetneq Rc \subsetneq Rd \ldots$.

Since $R$ is Noetherian, eventually the situation of Lemma 5.9 cannot occur. Thus we arrive at

$$A \sim \begin{bmatrix} f_1 & 0' \\ 0'' & F_1 \end{bmatrix},$$

where $f_1$ is a linear combination of the entries of $A$ and $f_1$ divides every entry of the matrix $F_1$. Repeating all of this with $A$ replaced by $F_1$ we have

$$F_1 \sim \begin{bmatrix} f_2 & 0' \\ 0'' & F_2 \end{bmatrix},$$
where $f_2$ divides every entry of $F_2$ and is a linear combination of the entries of $F_1$, so $f_1 \mid f_2$.

Repeating this process we arrive at the invariant factors $f_1 \mid f_2 \mid \cdots \mid f_k$ and we have proved the existence of Smith Normal form.

### 5.9 Uniqueness of Smith Normal Form

The aim of this section is to prove that each $\text{SL}_m(R) \times \text{SL}_n(R)$-orbit in $M_{m \times n}$ has exactly one set of invariant factors $f_1 \mid f_2 \mid \cdots \mid f_k$. In fact we will show how the $f_i$ may be computed directly from $A$. This eliminates the need for row and column operations, which are fine for proving existence, but quite boring to implement in practice.

Let $A \in M_{m \times n}(R)$. Choose a set of $j$ columns and $j$ rows of $A$ and take the determinant of the resulting $j \times j$ matrix. This determinant is a $j$-minor of $A$. From our discussion of exterior powers, we see that the $j$-minors of $A$ are the matrix entries of the exterior power matrix $\Lambda^j(A) \in \text{End}(\Lambda^j(R^n))$. Let $I_j(A)$ be the ideal in $A$ generated by all $j$-minors of $A$. We note that

$$I_j(A) = I_1(\Lambda^j(A)).$$

Since each minor is an $R$-linear combination of $j - 1$-minors, we have

$$I_1(A) \supset I_2(A) \supset \cdots \supset I_\rho(A),$$

where $\rho = \rho(A)$ is the largest value of $j$ such that $I_j(A) \neq 0$. Equivalently, $\rho(A)$ is the largest value of $j$ such that $\Lambda^j(A)$ is a nonzero matrix.  

**Lemma 5.10** If $A \sim B \in M_{m \times n}(R)$ then $I_j(A) = I_j(B)$ for all $j$.

**Proof:** Writing $B = gAh$ we clearly have $I_1(A) \supset I_1(B)$. By symmetry we have $I_1(A) = I_1(B)$. Now $\Lambda^j(B) = \Lambda^j(g)\Lambda^j(A)\Lambda^j(h)$, so

$$I_j(A) = I_1(\Lambda^j(A)) = I_1(\Lambda^j(B)) = I_j(B).$$

Now if

$$A \sim \begin{bmatrix} f_1 & 0 & 0 & \cdots & 0 \\ 0 & f_2 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & f_k & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

---

10 We can also characterize $\rho(A)$ as the rank of the image $AR^n$. There is an unfortunate conflict of terminology here, because we have been speaking of the “rank” $r$ of $R^n/AR^n$. The two notions of rank are related by: $\rho(A) + r = m$.  

34
with \( f_1 \mid f_2 \mid \cdots \mid f_k \) then \( Rf_1 \cdots f_j = I_j(A) \) for each \( j \), so each \( f_j \) is uniquely determined, up to a unit in \( R \), by the minors of \( A \). This proves uniqueness of the Smith Normal Form, as well as the converse to Lemma 5.10:

**Corollary 5.11** We have \( A \sim B \in M_{m \times n}(R) \) if and only if \( I_j(A) = I_j(B) \) for all \( j \).

### 5.10 Summary and variations

When a group \( G \) acts on a set \( X \) with many orbits, one tries to classify the orbits in two steps:

i) Decompose \( X = \bigsqcup_{i \in I} X_i \) into \( G \)-orbits \( X_i \), where \( i \) ranges over some index set \( I \).

ii) Find functions \( \phi_j : X \to I \), \( 1 \leq j \leq m \) whose values on any given \( x \in X \) determine the \( G \)-orbit \( X_i \) containing \( x \). These functions are the *basic invariants* for the \( G \)-orbits.

#### Example 1: The group \( \text{SO}_n(\mathbb{R}) \) acts on \( \mathbb{R}^n \). The orbit decomposition is

\[
\mathbb{R}^n = \bigsqcup_{r \geq 0} S_r
\]

where \( S_r \) of radius \( r \geq 0 \). We need only one basic invariant: the quadratic form \( q(x) = \sum x_i^2 \), and \( S_r = \{ x \in \mathbb{R}^n : q(x) = r^2 \} \).

#### Example 2: Let \( R \) be a PID. The group \( G = \text{SL}_n(R) \) acts on \( R^n \). Again we need only one basic invariant, which assigns to \( v \in \mathbb{R}^n \) the ideal \((v) \subset R \) generated by the components of \( v \). The orbits are parametrized by the set \( I \) of ideals of \( R \) and

\[
R^n = \bigsqcup_{I \in I} R^n_I,
\]

where \( R^n_I = \{ v \in R^n : (v) = I \} \).

#### Example 3: \( R \) is again a PID. Now \( G = \text{SL}_m(R) \times \text{SL}_n(R) \), acting on \( M_{m \times n}(R) \), and the orbit decomposition is

\[
M_{m \times n}(R) = \bigsqcup_{I \in I} M_{m \times n}(R)_I,
\]

where \( I \) is the set of chains of ideals \( I = (I_1 \supset I_2 \supset \cdots \supset I_m) \), and \( A \in M_{m \times n}(R)_I \) iff each invariant factor \( f_j \) of \( A \) generates \( I_j \) for \( 1 \leq j \leq m \). The ideals \( Rf_1, \ldots, Rf_m \) are the basic invariants of \( A \), and they can be found by computing the minors of \( A \).

### 6 Rational canonical form and regular transformations

Let \( V \) be a vector space of dimension \( n \) over a field \( F \). In this section we classify the orbits of \( \text{GL}(V) \) acting on \( \text{End}_F(V) \), by conjugation. This will follow from Cor. 5.5, applied to the polynomial ring
Each ideal in $I \subset F[X]$ has a canonical generator, namely the unique monic polynomial $f \in I$ of minimal degree.

For each $T \in \text{End}_F(V)$ we have a homomorphism
\[
\varphi_T : F[X] \longrightarrow \text{End}_F(V) \quad f(X) \mapsto f(T)
\] (24)
making $V$ into an $F[X]$-module which we denote by $V_T$. The minimal polynomial of $T$ is the canonical generator of the ideal $\ker \varphi_T$. Since $\dim(V) < \infty$ the $F[X]$-module $V_T$ has rank zero and coincides with its torsion submodule: $V_T = (V_T)_{\text{tor}}$.

One checks that for $T, S \in \text{End}_F(V)$ we have $T \sim S$ under $\text{GL}(V)$ if and only if $V_T \simeq V_S$ as $F[X]$-modules. Therefore the classification of $\text{GL}(V)$-orbits in $\text{End}_F(V)$ is equivalent to the classification of torsion $F[X]$-modules.

By Cor. 5.5 there is an integer $s \in [1, n]$ and unique ideals $I_1 \supset I_2 \supset \cdots \supset I_s \neq 0$ such that
\[
V_T \simeq \bigoplus_{j=1}^s F[X]/I_j.
\] (25)
Under this isomorphism the action of $T$ on $V_T$ corresponds to multiplication by $X$ on each summand $F[X]/I_j$.

### 6.1 Cyclic modules and companion matrices

To make (25) more explicit, we first study the case of one summand:
\[
V_T \simeq F[X]/I,
\]
where $I$ is a nonzero ideal in $F[X]$. These are precisely the torsion $F[X]$-modules which are cyclic, meaning they are generated by one element.

The ideal $I$ is also a cyclic $F[X]$ module, and it has a canonical generator, namely the monic polynomial $f \in I$ of minimal degree. We write
\[
f = X^d - a_1X^{d-1} - a_2X^{d-2} - \cdots - a_{d-1}X - a_d.
\]
Note that $f$ is the minimal polynomial of the endomorphism $T_f \in \text{End}_F(F[X]/I)$ given by multiplication by $X$.

As an $F$-vector space, $F[X]/I$ has the ordered basis $\{x^0, x, \ldots, x^{d-1}\}$, where $x$ is the image of $X$ in $F[X]/I$, and $x^0 = 1 + I$. With respect to this basis, the matrix of $T_f$ is given by
\[
A_f = \begin{bmatrix}
0 & 0 & 0 & \ldots & a_d \\
1 & 0 & 0 & \ldots & a_{d-1} \\
0 & 1 & 0 & \ddots & \vdots \\
\vdots & \ddots & 0 & a_2 \\
0 & 0 & \ldots & 1 & a_1
\end{bmatrix},
\]
since in $F[X]/I$ we have the relation

$$x^d = a_1x^{d-1} + a_2x^{d-2} + \cdots + a_{d-1}x + a_0.$$  

We call $A_f$ the companion matrix of $f$. Using $A_f$ one checks that $f$ is also the characteristic polynomial of $T_f$ and that the invariant factors of $T_f$ are $1, 1, \ldots, 1, f$.

### 6.2 Rational canonical form

Returning to (25), let $f_j$ be the canonical generator of $I_j$. Then we have

$$f_1 \mid f_2 \mid \cdots \mid f_s, \quad \sum_{j=1}^{s} \deg(f_j) = n,$$

and $f_s(x)$ is the minimal polynomial of $T$; this is the monic polynomial generating the annihilator of $V_T$, that is, the kernel of the map (24).

The matrix $A_T$ of $T$ on $V$ with respect to the canonical basis in each summand is a diagonal block matrix of companion matrices

$$A_T = \begin{bmatrix} A_{f_1} & 0 & \cdots & 0 \\ 0 & A_{f_2} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{f_s} \end{bmatrix}.$$  

This matrix $A_T$ is the rational canonical form of $T$. In summary, we have

**Proposition 6.1** For two endomorphisms $S, T \in \text{End}_F(V)$ the following are equivalent.

1. $S$ is conjugate to $T$ by an element of $\text{GL}(V)$.
2. $V_S \simeq V_T$ as $F[X]$-modules.
3. $A_S = A_T$.

### 6.3 Finding the rational canonical form

In this section we show how to find the rational canonical form of any transformation $T \in \text{End}_F(V)$.

**Proposition 6.2** The rational canonical form of $T \in \text{End}_F(V)$ is the block matrix $\oplus_j A_{f_j}$, where the $f_j$ are the nonconstant invariant factors of the transformation $T_X := X \cdot I_V - T \in \text{End}_{F[X]}(F[X] \otimes_F V)$.  

37
Proof: By reverse-engineering. Suppose $f \in F[x]$ is a monic polynomial of degree $n > 0$ with companion matrix $A_f$. Consider the matrix

$$X \cdot I_V - A_f \in M_n(F[X]).$$

For $j < n$ this matrix has a $j$-minor equal to 1 and $\det(X \cdot I_V - A_f) = f$. Hence the invariant factors of $X \cdot I_V - A_f$ are $1, \ldots, 1, f$. This shows that $f$ is the unique nonconstant invariant factor of $X \cdot I_V - A_f$.

Now suppose $T \in \text{End}_F(V)$ has rational canonical form $A_T = \bigoplus A_{f_j}$. Then the matrix of $T_X$ in the same basis is the block matrix

$$\bigoplus_{j=1}^s [X \cdot I_V - A_{f_j}]$$

whose nonconstant invariant factors are $f_1, \ldots, f_s$.

In particular the characteristic polynomial $P_T(X) = \det(T_X)$ is the product of the invariant factors:

$$P_T(X) = \prod_{j=1}^s f_j.$$

It follows that the minimal polynomial $f_s$ divides the characteristic polynomial $P_T(X)$. This proves the Cayley-Hamilton theorem:

$$P_T(T) = 0.$$

As an example, we compute the rational canonical form of the matrix

$$T = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

The matrix

$$T_X = \begin{bmatrix} x - 1 & -2 & -3 \\ -4 & x - 5 & -6 \\ -7 & -8 & x - 9 \end{bmatrix}$$

has $d_1 = d_2 = 1$ and $d_3 = \det(xI - T) = x^3 - 15x^2 - 18x$ is the unique invariant factor. Hence the rational canonical form of $T$ is the matrix

$$A_T = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 18 \\ 0 & 1 & 15 \end{bmatrix}.$$

It was not obvious at the outset that $T$ would be a single companion matrix with respect to some basis. In the next section we will characterize such transformations.
6.4 Centralizers and regular transformations

Given \( T \in \text{End}_F(V) \), we consider the centralizer

\[ Z(T) = \{ S \in \text{End}_F(V) : ST = TS \}. \]

This is an \( F \)-subalgebra of \( \text{End}_F(V) \). Equivalently,

\[ Z(T) = \text{End}_{F[X]}(V_T) \]

is the algebra of \( F[X] \)-module endomorphisms of \( V_T \).

The isomorphism (25) expresses \( V_T \) as direct sum of \( F[X] \)-modules \( F[X]/(f_j) \), where the \( f_j \) are the monic invariant factors of \( T \). Each of these summands is actually a commutative \( F \)-algebra which also commutes with the action of \( T \) on all of \( V_T \). Hence we have

\[ \prod_{j=1}^s F[X]/(f_j) \subset Z[T], \]

as \( F \)-algebras. In particular \( \dim_F Z[T] \geq \sum_j \deg f_j = n \). We say \( T \) is regular if \( \dim_F Z(T) = n \).

**Proposition 6.3** For \( T \in \text{End}_F(V) \) the following are equivalent.

1. \( T \) is regular;
2. \( T_X \) has just one nonconstant invariant factor;
3. \( V_T \) is a cyclic \( F[X] \)-module;
4. The minimal and characteristic polynomials of \( T \) coincide;
5. There is a basis of \( V \) for which the matrix of \( T \) has the form

\[
\begin{bmatrix}
0 & 0 & 0 & \ldots & a_n \\
1 & 0 & 0 & \ldots & a_{n-1} \\
0 & 1 & 0 & \ddots & \vdots \\
\vdots & \ddots & 0 & a_2 \\
0 & 0 & \ldots & 1 & a_1 \\
\end{bmatrix},
\]

for some elements \( a_1, \ldots, a_n \in F \).

**Proof:** We prove 1 \( \iff \) 2; the rest is easily seen from our discussion so far.

If there is a pair nonconstant invariant factors then \( f_{s-1}, f_s \) would be one such pair. Since \( f_{s-1} \mid f_s \), we have an evident surjective map of \( R \)-modules

\[ F[X]/(f_s) \longrightarrow F[X]/(f_{s-1}). \]
Extending this map by zero on the summands $F[X]/(f_j)$ for $j < s$, we get an endomorphism in $\text{End}_{F[X]}(V_T)$ which does not preserve each summand, so cannot lie in $\prod_j F[X]/(f_j)$. It follows that $\dim Z(T) > n$ so $T$ is not regular.

For any commutative ring $R$ with identity and ideal $I \subset R$, we have $\text{End}_R(R/I) = R/I$. More precisely, the endomorphism ring of the $R$-module $R/I$ is iso the $R$-algebra $R/I$. Now if $V_T$ only one nonconstant invariant factor $f$ then $V_T \cong F[X]/(f)$, so

$$Z(T) = \text{End}_{F[X]}(V_T) = F[X]/(f).$$

Since $\deg f = \dim V = n$ it follows that $T$ is regular. □

The set $\text{End}_F(V)_{\text{reg}}$ of regular transformations in $\text{End}_F(V)$ is a union of $\text{GL}(V)$-orbits under conjugation. In part 5 of Prop. 6.3, the numbers $a_1, \ldots, a_n$ are the coefficients of the characteristic polynomial of $T \in \text{End}_F(V)_{\text{reg}}$, hence the $a_k$ are obtained directly from $T$ via

$$a_k = (-1)^{k-1} \text{tr}(\Lambda^k T),$$

where $\text{tr}(\Lambda^k T)$ is the trace of the natural map induced by $T$ on the exterior power $\Lambda^k V$. It follows that we have a bijection

$$\text{End}_F(V)_{\text{reg}}/\text{GL}(V) \sim \rightarrow F^n,$$

sending

$$T \mapsto (\text{tr}(T), -\text{tr}(\Lambda^2 T), \ldots, (-1)^{n-1} \text{tr}(\Lambda^n T)).$$

The inverse bijection is obtained by choosing an ordered basis $(e_1, \ldots, e_n)$ of $V$ so as to identify $\text{End}_F(V) = M_n(F)$ and $\text{GL}(V) = \text{GL}_n(F)$, and sending $(a_1, \ldots, a_n) \in F^n$ to the the corresponding companion matrix in $M_n(F)$. Passing to $\text{GL}_n(F)$-orbits makes this independent of the choice of basis.

The companion matrices form an affine subspace $\mathcal{A} \subset M_n(F)$ which meets each regular orbit of $\text{GL}_n(F)$ in $M_n(F)$ in exactly one point, namely the companion matrix of the orbit. This affine space $\mathcal{A}$ has a canonical basepoint, namely

$$J_n = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ddots & \vdots \\ \vdots & \ddots & 0 & 0 \\ 0 & 0 & \ldots & 1 & 0 \end{bmatrix}$$

which is the unique matrix in $\mathcal{A}$ with minimal(=characteristic) polynomial equal to $X^n$.

### 6.5 Conjugacy vs stable conjugacy

**Corollary 6.4** Let $K$ be a field containing $F$ and suppose $S, T \in \text{End}(V)$ are conjugate by an element of $\text{GL}(K \otimes_F V)$. Then $S$ and $T$ are conjugate by an element of $\text{GL}(V)$. 

40
Proof: Let \( f_1, \ldots, f_r \) be the invariant factors of \( T \), so that
\[
V_T \simeq \bigoplus_{i=1}^r F[x]/(f_i).
\]
Since
\[
(K \otimes V)_T \simeq K \otimes V_T \simeq \bigoplus_{i=1}^r K[x]/(f_i),
\]
the matrix \( A_T \) with respect to the canonical basis is the same on \( V \) and on \( K \otimes V \). Since the same holds for \( S \) which is conjugate to \( T \) by an element of \( \text{GL}(K \otimes V) \), it follows that \( A_S = A_T \), so \( S \) is conjugate to \( T \) by an element of \( \text{GL}(V) \), by Prop. 6.1.

Example: Let \( F = \mathbb{R} \) and consider the two matrices \( S = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \) and \( T = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \) in \( \text{SL}_2(\mathbb{R}) \). Then \( S \) is conjugate to \( T \) by \( \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \) in \( \text{SL}_2(\mathbb{C}) \) and by \( \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \) in \( \text{GL}_2(\mathbb{R}) \), but \( S \) and \( T \) are not conjugate in \( \text{SL}_2(\mathbb{R}) \), as you can easily check. We say that \( S \) and \( T \) are stably conjugate but not rationally conjugate in \( \text{SL}_2(\mathbb{R}) \). Cor. 6.4 shows this phenomenon does not occur in \( \text{GL}_n(F) \): there stable and rational conjugacy coincide.

7 Jordan Canonical Form

We have seen that Rational Canonical Form of an endomorphism \( T \in \text{End}(V) \) is based on the invariant factors of the \( F[x] \)-module \( V_T \), and this canonical form is unchanged if we extend scalars to a larger field.

However, the structure of the summands \( F[X]/(f_i) \) does depend on \( F \). Indeed, take a monic polynomial \( f \in F[x] \) of positive degree, with factorization \( f = p_1^{m_1} \cdots p_k^{m_k} \) where \( p_1, \ldots, p_k \) are distinct irreducible monic polynomials in \( F[x] \). By the Chinese Remainder Theorem, we have
\[
F[x]/(f) = \bigoplus_{i=1}^k F[x]/(p_i^{m_i}),
\]
so the structure of this \( F[x] \)-module depends on the factorization of \( f \) in \( F[x] \). The simplest situation is when \( F \) is large enough to contain all of the roots of \( f \). If \( f = (x - \lambda_1)^{m_1} \cdots (x - \lambda_k)^{m_k} \) where \( \lambda_1, \ldots, \lambda_k \in F \) are the distinct roots of \( f \) then
\[
F[x]/(f) = \bigoplus_{i=1}^k F[x]/(x - \lambda_i)^{m_i}.
\]
As a basis of \( F[x]/(x - \lambda)^m \) we may take
\[
\{1, x - \lambda, (x - \lambda)^2, \ldots, (x - \lambda)^{m-1}\}
\]
the matrix of multiplication by $x - \lambda$ is then all zeros except 1’s below the main diagonal. It follows that the matrix of $X$ is the Jordan matrix

$$J_m(\lambda) = \begin{bmatrix} \lambda & 0 & 0 & \cdots & 0 \\ 1 & \lambda & 0 & \cdots & 0 \\ 0 & 1 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \cdots & 0 \\ 0 & 0 & \cdots & 1 & \lambda \end{bmatrix} = \lambda I + J_m, \quad (27)$$

where $J_m$ is the matrix from (26).

Now let $f_1 \mid f_2 \mid \cdots \mid f_r$ be the invariant factors of $T \in \text{End}(V)$, and suppose all the roots of $f_r$ belong to the field $F$. Let $m_i(\lambda)$ be the multiplicity of $\lambda$ as a root of $f_i$. Then

$$F[x]/(f) = \bigoplus_{j=1}^{r} \bigoplus_{i=1}^{j} F[x]/(x - \lambda)^{m_i(\lambda)},$$

where $\lambda$ runs over the roots of $f_r$ (that is, the eigenvalues of $T$). So there is a basis of $V$ putting the matrix of $T$ in the form

$$J_T = \bigoplus_{\lambda} J_T(\lambda), \quad (28)$$

Where

$$J_T(\lambda) = \bigoplus_{i=1}^{r} J_{m_i(\lambda)}(\lambda)$$

is a direct sum of Jordan matrices (27). The decomposition (28) is the Jordan canonical form of $T$. For each eigenvalue $\lambda$ of $T$ we have

$$m_1(\lambda) + m_2(\lambda) + \cdots + m_r(\lambda) = m(\lambda),$$

where $m(\lambda)$ is the multiplicity of $\lambda$ as a root of $\det(xI - T)$. Since $f_1 \mid f_2 \mid \cdots \mid f_r$, we have

$$m_1(\lambda) \leq m_2(\lambda) \leq \cdots \leq m_r(\lambda);$$

such a sequence is a partition of $m(\lambda)$.

Thus, the sequences

$$\mu_T(\lambda) = (m_1(\lambda), \ldots, m_r(\lambda))$$

(for each eigenvalue $\lambda$ of $T$) determine the conjugacy class of $T$. More precisely:

**Proposition 7.1** Let $S, T \in \text{End}(V)$ and let $K \supset F$ be a field containing the eigenvalues of $S$ and $T$. Then $S$ and $T$ are conjugate by an element of $\text{GL}(V)$ if and only if $\mu_T(\lambda) = \mu_S(\lambda)$ for all $\lambda \in K$.

**Proof:** The partitions $\mu_T(\lambda)$ and $\mu_S(\lambda)$ determine the invariant factors of $S$ and $T$ in $V_K$, so $S$ and $T$ are conjugate by an element of $\text{GL}(K \otimes_F V)$. By Cor. 6.4, it follows that $S$ and $T$ are conjugate by an element of $\text{GL}(V)$.

42
7.1 Semisimple and nilpotent endomorphisms

An endomorphism \( N \in \text{End}(T) \) is \textbf{nilpotent} if some power of \( N \) is zero.

**Proposition 7.2** For \( N \in \text{End}(T) \) the following are equivalent.

1. \( N \) is nilpotent.
2. The eigenvalues of \( N \) are all zero.
3. \( \text{tr}(\Lambda^k(N)) = 0 \) for all \( 1 \leq k \leq n \).

**Proof:** Suppose \( N^r = 0 \), and \( \lambda \) is an eigenvalue of \( N \) in some field \( K \) containing \( F \). There is a nonzero vector \( v \) such that \( Nv = \lambda v \). Then \( N^rv = \lambda^rv \), so \( \lambda^r = 0 \), which implies that \( \lambda = 0 \). Conversely, if the eigenvalues of \( N \) are then all zero, so \( N \) is conjugate to a sum
\[
J_{n_1}(0) + \cdots + J_{n_r}(0)
\]
of strictly upper-triangular Jordan matrices, which is nilpotent. Finally, having all eigenvalues zero means that \( \det(xI - N) = x^n \), so 2 and 3 are equivalent because \( \pm \text{tr}(\Lambda^k(N)) \) is the coefficient of \( x^k \) in \( \det(xI - N) \).

**Remark:** If \( F = \mathbb{R} \) or \( \mathbb{C} \), there is also a topological condition for nilpotency: \( N \) is nilpotent if and only if the conjugacy class of \( N \) contains 0 in its closure in \( \text{End}(V) \). In fact this holds for any algebraically closed field \( F \) with the Zariski topology on \( \text{End}(V) \).

The invariant factors of a nilpotent element \( N \in \text{End}(V) \) are \( x^{m_1}, x^{m_2}, \ldots, x^{m_r} \), where \( m_1 \leq m_2 \leq \cdots \leq m_r \) is a partition of \( n \), and the Jordan form \( N \) is
\[
J_N = \bigoplus_{i=1}^r J_{m_i}(0).
\]

In particular we have

**Proposition 7.3** The set of nilpotent conjugacy classes in \( \text{End}(V) \) is in bijection with the set of of partitions of \( n = \text{dim} V \). In particular there are only finitely many nilpotent classes.

At the other extreme we say that an endomorphism \( S \in \text{End}(V) \) is \textbf{semisimple} if \( V_S \) is a semisimple \( F[X] \)-module, meaning that \( V_S \) is a direct sum of simple \( F[X] \)-modules. Now the submodules of the \( F[X] \)-module \( F[X]/(f) \) are just the ideals in the ring \( F[X]/(f) \). It follows that \( F[X]/(f) \) is simple if and only if \( f \) is irreducible in \( F[X] \).

In the following result we assume \( F \) is a \textit{perfect field}. These are the fields \( F \) for which irreducible polynomials are \textit{separable}, meaning they have distinct roots in an algebraic closure of \( F \). Perfect fields include all fields of characteristic zero, as well as algebraic extensions of finite fields. \(^{11}\)

\(^{11}\)The simplest example of an imperfect field is \( \mathbb{F}_p(t) \), the rational functions in one variable over \( \mathbb{F}_p \). Here the polynomial \( f = X^p - t \) is irreducible, by Eisenstein’s criterion, but \( f = (X - \tau)^p \) over the field \( K = F(\tau) \), where \( \tau \) is a root of \( f \).
Proposition 7.4 Assume $F$ is a perfect field. For $S \in \text{End}_F(V)$ the following are equivalent.

1. $S$ is semisimple;
2. The minimal polynomial of $S$ is separable;
3. If $K \supset F$ is any field containing all the roots of $\det(xI - S)$ then $K \otimes_F V$ is a direct sum of eigenlines for $S$.

Proof: (1 $\iff$ 2): Let $f_1 | \cdots | f_k$ be the invariant factors of $S$, so that $f_k$ is the minimal polynomial of $S$ and

$$V_S = \bigoplus_{i=1}^k F[x]/(f_i).$$

Factoring each $f_i$, the Chinese Remainder Theorem shows that $V_S$ is a direct sum of $F[X]$-modules of the form $F[X]/(g^e)$, where $g \in F[X]$ is irreducible and $e \geq 1$. Now $F[X]/(g^e)$ has a unique simple submodule, generated by the image of $g^{e-1}$, hence can only be semisimple if $e = 1$, in which case it is simple. It follows that each invariant factor $f_i$ factors into distinct irreducible polynomials, which are separable since $F$ is perfect. In particular the minimal polynomial $f_k$ is separable. Conversely if $f_k$ is separable, then each invariant factor $f_i$, which divides $f_k$, is also separable. Factoring $f_i$ into irreducibles, we see that $V_S$ is a direct sum of simple $F[X]$-modules, so $S$ is semisimple.

(2 $\iff$ 3): We have remarked that if $f_k$ is separable then so is every invariant factor $f_i$. Let $\{\lambda_{ij}\}$ be the roots of $f_i$ in $K$. Then $(K \otimes_F V)_S$ is a direct sum of modules $K[x]/(x - \lambda_{ij})$ which are one-dimensional over $K$, and are eigenlines for $S$. Conversely, if 2 fails then some irreducible factor $g$ of $f_r$ appears with multiplicity $e > 1$. It follows the Jordan form of $S$ contains a Jordan matrix $J_e(\lambda)$, which has only one eigenline, so 3 fails.

Remark: If $F = \mathbb{R}$ or $\mathbb{C}$, there is also a topological condition for semisimplicity: $S$ is semisimple if and only if the conjugacy class of $S$ is closed in $\text{End}(V)$. Again, this holds for any algebraically closed field $F$ with the Zariski topology on $\text{End}(V)$.

Proposition 7.5 Let $S, S' \in \text{End}(V)$ be semisimple. Then $S$ and $S'$ are conjugate under $\text{GL}(V)$ if and only if $\det(xI - S) = \det(xI - S')$.

Proof: Assume that $\det(xI - S) = \det(xI - S')$. Let $K$ be a field containing the eigenvalues $\lambda_1, \ldots, \lambda_s$ of $S$ and $S'$ and let $m_j$ be the multiplicity of $\lambda_j$. By Prop 7.4, $S$ and $S'$ are both conjugate via $\text{GL}(K \otimes FV)$ to the diagonal matrix

$$\begin{pmatrix} \lambda_1, \ldots, \lambda_1, \lambda_2, \ldots, \lambda_2, \ldots, \lambda_s, \ldots, \lambda_s \end{pmatrix}_{m_1 m_2 \cdots m_s}.$$

From Cor. 6.4 we conclude that $S, S'$ are conjugate via $\text{GL}(V)$.
7.2 The Jordan decomposition

A Jordan matrix $J_m(\lambda)$ can be written as the sum

$$J_m(\lambda) = \lambda I + J_m(0),$$

a scalar matrix $\lambda I$ plus a nilpotent matrix $J_m(0)$, and these two matrices commute. It follows that the Jordan form has a basis-free interpretation:

**Proposition 7.6 (Jordan decomposition)** Every endomorphism $T \in \text{End}(V)$ may be uniquely written as $T = S + N$, where $S$ is semisimple, $N$ is nilpotent, and $SN = NS$.

The last condition means that $N$ belongs to the centralizer

$$\text{End}_S(V) := \{ A \in \text{End}(V) : AS = SA \} = \text{End}_{F[x]}(V_S),$$

so we have a partition

$$\text{End}(V) = \coprod S \left[ S + \text{End}_S(V)_{\text{nil}} \right],$$

where $S$ ranges over the semisimple endomorphisms in $\text{End}(V)$ and $\text{End}_S(V)_{\text{nil}}$ is the set of nilpotent endomorphisms of $\text{End}_S(V)$. Note that $S$ is the unique semisimple element in $S + \text{End}_S(V)_{\text{nil}}$.

If $S$ is a given semisimple conjugacy class in $\text{End}(V)$ then the set

$$\text{End}(V)_S := \coprod_{S \in S} [S + \text{End}_S(V)_{\text{nil}}]$$

consists of finitely many conjugacy classes, namely those those with a given characteristic polynomial.

In geometric terms, we have a mapping

$$\chi : \text{End}(V) \longrightarrow F^n, \quad \chi(T) = (\text{tr}(T), \text{tr}(\Lambda^2T), \ldots, \text{tr}(\Lambda^nT)),$$

which assigns to $T \in \text{End}(V)$ the vector of coefficients (up to signs) in the characteristic polynomial $\det(xI - T)$. The sets $\text{End}(V)_S$ are the fibers of $\chi$. If $F = \mathbb{R}$ or $\mathbb{C}$ or is algebraically closed, then $S$ is the unique closed conjugacy class in $\text{End}(V)_S$. Moreover $\text{End}(V)_S$ is itself the closure of a unique conjugacy class, namely, that of the companion matrix $C_f$, where $f$ is the characteristic polynomial for $S$.

7.3 Nilpotent elements and $\mathfrak{sl}_2$

Now assume the ground field $F = \mathbb{C}$ (or any algebraically closed field of characteristic zero).

For each partition $n = m_1 + m_2 + \cdots + m_r$, with $1 \leq m_1 \leq \cdots \leq m_r$, we also have a representation

$$V = V_{m_1-1} \oplus V_{m_2-1} \oplus \cdots \oplus V_{m_r-1}$$
of $\mathfrak{sl}_2$, where as in section 5.4, $V_d = \text{Sym}^d V_1$ and $V_1$ is the standard two-dimensional representation. If $d \geq 1$ the representation $\mathfrak{sl}_2 \rightarrow \text{End}(V_d)$ is injective. It follows that the representation $\mathfrak{sl}_2 \rightarrow \text{End}(V)$ is injective if $m_r \geq 2$. In each summand, the element $f \in \mathfrak{sl}_2$ acts via the Jordan matrix $J_{m_k}$, in the $f$-cyclic basis. Hence each nonzero nilpotent element $N \in \text{End}(V)$ is the image of $f$ under an injective representation $\rho_N : \mathfrak{sl}_2 \rightarrow \text{End}(V)$. Moreover the conjugacy class of $N$ and the isomorphism class of $\rho_N$ determine each other. Thus we have bijections

$$\text{End}(V)_{\text{nil}} / \text{GL}(V) \cong \text{Irr}_n(\mathfrak{sl}_2)$$

$$N \mapsto \rho_N$$

$$\rho(f) \mapsto \rho$$

between the set of $\text{GL}(V)$-conjugacy classes of nilpotent endomorphisms in of $V$ and the set $\text{Irr}_n(\mathfrak{sl}(2))$ of isomorphism classes of $n$-dimensional representations of $\mathfrak{sl}_2$. For example, the unique regular nilpotent conjugacy class corresponds to the unique $n$-dimensional irreducible representation of $\mathfrak{sl}_2$.

From the results of section 5.4, we note that the $e, h, f$ act on $V$ via operators with trace zero. Hence the above images of $\mathfrak{sl}_2$ actually appear in $\mathfrak{sl}(V)$, the trace-zero subalgebra of $\text{End}(V)$. The bijections above are a special case (for $L = \mathfrak{sl}(V)$) of the Jacobson-Morozov Theorem for semisimple Lie algebras $L$ over $F$. See Bourbaki Groupes et Algèbres de Lie book VIII, chapter 11.