

Willoughby Seago

Notes from

Lie Theory

September 26th, 2024

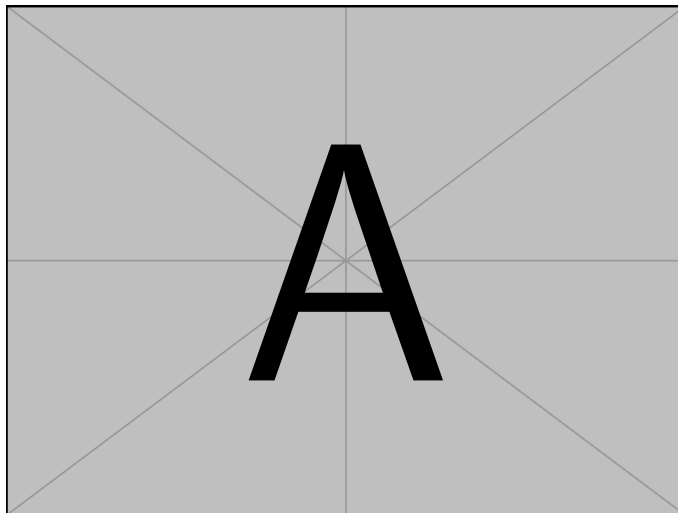
UNIVERSITY OF GLASGOW

Lie Theory

Willoughby Seago

September 26th, 2024

These are my notes from the master's course *Lie Theory* taught at the University of Glasgow by Dr Dinakar Muthiah. These notes were last updated at 17:44 on September 26, 2024.



Chapters

	Page
Chapters	ii
Contents	iii
1 Introduction	1
2 Direct Sums and Diagonalisation	3

Contents

	Page
Chapters	ii
Contents	iii
1 Introduction	1
1.1 Notation	1
2 Direct Sums and Diagonalisation	3
2.1 Direct Sums	3
2.1.1 Binary Direct Sums	3
2.1.2 Finite Direct Sums	4
2.1.3 Arbitrary Direct Sums	5
2.2 Diagonalisable Operators and the Direct Sum Decomposition . . .	6
2.2.1 Basis Independent Definition	7
2.2.2 Decomposition Theorem	8

One

Introduction

In this course we will study Lie algebras. The motivation for the definition of a Lie algebra is as a tool for studying Lie groups. Despite this we will rarely mention Lie groups, since the study of Lie groups requires a knowledge of differential geometry. We will stick to the study of Lie algebras, which can be defined purely algebraically.

Roughly speaking, the story is thus: Lie groups are both smooth manifolds and groups in a compatible way. By this, we mean that if G is a Lie group then the maps

- $m : G \times G \rightarrow G$ given by $m(x, y) = xy$
- $i : G \rightarrow G$ given by $i(g) = g^{-1}$

are smooth and these maps satisfy the obvious requirements to make G a group.

Lie groups were first used by Sophus Lie, who lends his name to these objects. Lie is pronounced Lee by the way, not lye. He was looking at the continuous symmetries possessed by the solutions to differential equations, and found that these naturally formed what we now know as a Lie group. The problem with studying Lie groups, apart from all of the topology, is that they are generally very non-linear. We can think of Lie algebras as being a linear approximation of Lie groups, given by expanding about the identity and discarding non-linear terms.

To get a Lie algebra out of a Lie group we take the tangent space at the identity, $T_e G$, which is a vector space. It turns out that the group commutator, $[g, h] = ghg^{-1}h^{-1}$, induces a natural operation on the tangent space, called the Lie bracket. A Lie algebra is then this tangent space equipped with this Lie bracket.

It is possible to abstract the properties of this Lie bracket to define a Lie algebra without reference to a Lie group. It is also possible to go in reverse, given a Lie algebra there is always a corresponding Lie group. It should be noted that this assignment is not unique, in general many different Lie groups can have the same Lie algebra. The assignment does become invertible if we restrict ourselves to simply connected Lie groups, but then we're really getting into topology in a way that this course hopes to avoid. This assignment of a Lie algebra to a Lie group is functorial.

1.1 Notation

Throughout \mathbb{k} will denote a field. Often we will place further restrictions on \mathbb{k} , such as being algebraically closed. Most of the time we'll be interested in the case

$\mathbb{k} = \mathbb{C}$, with occasional use of $\mathbb{k} = \mathbb{R}$. Unless stated otherwise vector spaces are assumed to be vector spaces over \mathbb{k} .

We will denote by 0 the zero vector space (for a given field), $\{0\}$, which consists of only the zero vector.

If we have a category, \mathcal{C} , we will denote by $\mathcal{C}(A, B) = \text{Hom}_{\mathcal{C}}(A, B)$ the set of all morphisms $A \rightarrow B$, or simply $\text{Hom}(A, B)$ if \mathcal{C} is clear from context. This is just notation, we won't make much use of categories apart from borrowing some of the language. The main category we'll make use of is $\mathbb{k}\text{-Vect}$, the category of vector spaces over \mathbb{k} and linear maps.

We will mostly denote Lie algebras with lowercase fraktur letters, or so called `\mathfrak` letters. For reference, here's the alphabet in `\mathfrak`:

`\mathfrak{a b c d e f g h i j k l m n o p q r s t u v w x y z}`

Two

Direct Sums and Diagonalisation

Our goal in this section is to state a definition of a diagonalisable linear operator in such a basis free way. We will then generalise this to get an “almost diagonal form” for arbitrary linear operators when \mathbb{k} is algebraically closed. To this end we will recap some linear algebra in this first section.

2.1 Direct Sums

2.1.1 Binary Direct Sums

Definition 2.1.1 — Binary Direct Sum Let V be a vector space over \mathbb{k} , and let $U_1, U_2 \subseteq V$ be subspaces. Then we say that $V = U_1 \oplus U_2$, that is V is the (internal) **direct sum** of U_1 and U_2 , if

- $V = U_1 + U_2 := \{u_1 + u_2 \mid u_1 \in U_1 \text{ and } u_2 \in U_2\}$; and
- $U_1 \cap U_2 = 0$.

If U_1 and U_2 are vector spaces over \mathbb{k} then we can construct a vector space $V = U_1 \oplus U_2$, also over \mathbb{k} , by defining

$$V = \{(u_1, u_2) \mid u_1 \in U_1 \text{ and } u_2 \in U_2\} \quad (2.1.2)$$

and defining addition and scalar multiplication by

$$(u_1, u_2) + (u'_1, u'_2) = (u_1 + u'_1, u_2 + u'_2) \quad \text{and} \quad \lambda(u_1, u_2) = (\lambda u_1, \lambda u_2) \quad (2.1.3)$$

for all $u_1, u'_1 \in U_1, u_2, u'_2 \in U_2$, and $\lambda \in \mathbb{k}$.

After constructing the external direct sum we may identify U_1 with the subspace consisting of elements of the form $(u_1, 0)$ with $u_1 \in U_1$, and U_2 with the subspace of elements of the form $(0, u_2)$ with $u_2 \in U_2$. Then the external direct sum coincides with the internal direct sum. For this reason we won't distinguish internal and external direct sums, making such identifications as necessary.

The following lemma gives an alternative characterisation of the direct sum.

Lemma 2.1.4 If $V = U_1 \oplus U_2$ then every vector $v \in V$ can be written *uniquely* as a sum $v = u_1 + u_2$ with $u_1 \in U_1$ and $u_2 \in U_2$. Conversely, if every v has a unique decomposition as $v = u_1 + u_2$ with $u_1 \in U_1$ and $u_2 \in U_2$ then $V = U_1 \oplus U_2$.

Proof. Suppose that $V = U_1 \oplus U_2$. From the definition of a direct sum we know that $v \in V = U_1 + U_2$ and as such $v = u_1 + u_2$ for some $u_1 \in U_1$ and $u_2 \in U_2$ because this is how elements of $U_1 + U_2$ are defined. We need only prove uniqueness. Suppose that $v = u_1 + u_2$ and $v = u'_1 + u'_2$ with $u_1, u'_1 \in U_1$ and $u_2, u'_2 \in U_2$ are two decompositions of v . Then we have $u_1 + u_2 = u'_1 + u'_2$, which we can rearrange to get $(u_1 - u'_1) + (u_2 - u'_2) = 0$. This means that $u_1 - u'_1 = u'_2 - u_2 =: w$. Now, $u_1 - u'_1 \in U_1$, since it's a linear combination of elements of U_1 , and similarly $u'_2 - u_2 \in U_2$. Thus, $w \in U_1 \cap U_2 = 0$ and so $w = 0$.

Suppose instead that every $v \in V$ has a unique decomposition as $v = u_1 + u_2$. Then clearly v corresponds to $(u_1, u_2) \in U_1 \oplus U_2$ and if $v = u_1 + u_2$ and $v' = u'_1 + u'_2$ and $\lambda \in \mathbb{k}$ then $v + \lambda v' = (u_1 + u_2) + \lambda(u'_1 + u'_2)$ corresponds to $(u_1, u_2) + \lambda(u'_1, u'_2)$, but also $v + \lambda v' = (u_1 + \lambda u'_1) + (u_2 + \lambda u'_2)$ corresponds to $(u_1 + \lambda u'_1, u_2 + \lambda u'_2)$. This shows that this correspondence defines a linear map. Clearly this correspondence is invertible, and thus we have an isomorphism $V \cong U_1 \oplus U_2$. \square

Thus, the direct sum may be characterised as giving a unique decomposition of each vector into a pair of vectors from two subspaces with only the zero vector in common.

2.1.2 Finite Direct Sums

Direct sums of two vector spaces generalise to direct sums of a finite number of vector spaces in an obvious way.

Definition 2.1.5 — Finite Direct Sum Let V be a vector space over \mathbb{k} , and let $U_1, \dots, U_r \subseteq V$ be subspaces. Then we say that

$$V = U_1 \oplus \dots \oplus U_r = \bigoplus_{i=1}^r U_i, \quad (2.1.6)$$

that is V is the (internal) **direct sum** of the U_i , if

- $V = U_1 + \dots + U_r := \{u_1 + \dots + u_r \mid u_i \in U_i \text{ for all } i = 1, \dots, r\}$; and
- for all $i = 1, \dots, r$ we have $U_i \cap (U_1 + \dots + U_{i-1} + U_{i+1} + \dots + U_r) = 0$ where the sum is over all subspaces apart from U_i .

Note that we can define the external direct sum, but care has to be taken as if we define $U_1 \oplus U_2 \oplus U_3$ to consist of elements of the form (u_1, u_2, u_3) then this is not the same as defining $U_1 \oplus (U_2 \oplus U_3)$ to consist of elements of the form $(u_1, (u_2, u_3))$ and $(U_1 \oplus U_2) \oplus U_3$ to consist of elements of the form $((u_1, u_2), u_3)$. However these spaces are naturally (in the technical sense) isomorphic, and as such we will identify them with each other. A complicated way to put this is that the direct sum

is associative up to natural isomorphism. An even more complicated way to put this, along with the fact that $V \oplus 0 \cong V \cong 0 \oplus V$, is that $(\mathbb{k}\text{-Vect}, \oplus, 0)$ is a monoidal category.

The same characterisation of the direct sum giving a unique decomposition of $v \in V$ carries over to finite direct sums.

Lemma 2.1.7 If $V = U_1 \oplus \cdots \oplus U_r$ then for all $v \in V$ there exist unique $u_i \in U_i$ such that $v = u_1 + \cdots + u_r$.

Proof. We proceed by induction on r . The case $r = 2$ is [Lemma 2.1.4](#). Suppose that the result holds some $k \geq 2$ and that $V = U_1 \oplus \cdots \oplus U_k \oplus U_{k+1}$. Take $v \in V$. Writing $V = (U_1 \oplus \cdots \oplus U_k) \oplus U_{k+1}$ we see that there are unique vectors $u \in U_1 \oplus \cdots \oplus U_k$ and $u_{k+1} \in U_{k+1}$ such that $v = u + u_{k+1}$. By the induction hypothesis since $U_1 \oplus \cdots \oplus U_k$ is a k -fold direct sum we can uniquely decompose u as $u = u_1 + \cdots + u_k$. Then $v = u_1 + \cdots + u_k + u_{k+1}$ gives a unique decomposition of v . Thus, by induction the result holds for any finite direct sum. \square

2.1.3 Arbitrary Direct Sums

We can further generalise direct sums to arbitrary collections of spaces. One thing we have to be cautious about is that infinite sums of vectors are not defined, and we get around this by considering sums over infinite sets, but requiring that only finitely many of the vectors in the sum are not zero. More formally, we can view a sequence $(u_i)_{i \in I}$ of vectors as a function $u : I \rightarrow V$ with $u(i) = u_i$ and then we require that u has finite support.

Definition 2.1.8 — Arbitrary Direct Sums Let V be a vector space over \mathbb{k} , and let $\{U_i\}_{i \in I}$ be an indexed family of subspaces, that is $U_i \subseteq V$ is a subspace for all $i \in I$. Then we say that

$$V = \bigoplus_{i \in I} U_i, \quad (2.1.9)$$

that is V is the (internal) **direct sum** of the U_i , if

- we have

$$V = \sum_{i \in I} U_i := \left\{ \sum_{i \in I} u_i \mid u_i \in U_i \text{ and } u_i \neq 0 \text{ for only finitely many terms} \right\};$$

- for all $i \in I$ we have $U_i \cap \sum_{j \in I \setminus \{i\}} U_j = 0$.

We can define the external direct sum by replacing sequences such as (u_1, u_2, \dots) with functions $I \rightarrow \bigcup_{j \in I} U_j$ with finite support, the idea being that $u(i)$ corresponds to u_i in the case where $I = \{1, \dots, r\}$.

Lemma 2.1.10 If $V = \bigoplus_{i \in I} U_i$ then for all $v \in V$ there exist unique $u_i \in U_i$ such that $v = \sum_{i \in I} u_i$ with only finitely many of the u_i being nonzero.

Proof. The proof is essentially the same as before, but now we work with formally infinite sums in which most terms are zero, and all we have to do is check that at each stage this is still the case, which it is, since adding two such sums cannot result in infinitely many nonzero terms. \square

Example 2.1.11 Let $V = \mathbb{k}[x]$ be the vector space of polynomials in x with coefficients in \mathbb{k} . That is

$$V = \left\{ \sum_{i=0}^N a_i x^i \mid N \in \mathbb{N} \text{ and } a_i \in \mathbb{k} \right\}. \quad (2.1.12)$$

Addition is just addition of polynomials, and scaling is simply scaling each coefficient, which we can phrase as multiplying by the corresponding constant polynomial. We have the subspaces

$$U_i = \text{span}\{x^i\} := \{ax^i \mid a \in \mathbb{k}\} \quad (2.1.13)$$

and it's easy to see that

$$V = \bigoplus_{i \in \mathbb{N}} U_i. \quad (2.1.14)$$

Specifically, if we have a polynomial, $f(x) \in \mathbb{k}[x]$, then

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_N x^N \quad (2.1.15)$$

and $a_0 \in U_0$, $a_1 x \in U_1$, $a_2 x^2 \in U_2$, and so on up to $a_N x^N \in U_N$, so we get a finite decomposition of $f(x)$, and clearly this is unique, showing that $\mathbb{k}[x]$ is indeed this direct sum as claimed.

Remark 2.1.16 The most general definition of the direct sum is as the co-product in the category $\mathbb{k}\text{-Vect}$.

2.2 Diagonalisable Operators and the Direct Sum Decomposition

In this section we'll give the standard definition of a diagonalisable operator, which should be familiar. We'll then give an equivalent definition which makes no mention of a basis.

Let V be a finite dimensional vector space over \mathbb{k} with basis $\mathcal{B} = \{v_1, \dots, v_n\}$. Recall that if $T: V \rightarrow V$ is a linear operator then its matrix in basis \mathcal{B} is the matrix

$$[T]_{\mathcal{B}} := \begin{bmatrix} \uparrow & & \uparrow \\ Tv_1 & \dots & Tv_n \\ \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix} \quad (2.2.1)$$

where c_{ij} is the coefficient of v_i in Tv_j when expressed in the basis \mathcal{B} . That is, the i th column of $[T]_{\mathcal{B}}$ is Tv_i as a column vector in basis \mathcal{B} .

Definition 2.2.2 — Diagonalisable Operator Let V be a finite dimensional vector space over \mathbb{k} . Then a linear operator, $T: V \rightarrow V$, is **diagonalisable** if there exists some basis, \mathcal{B} , in which the matrix $[T]_{\mathcal{B}}$ is diagonal.

Notice that if $T: V \rightarrow V$ is a linear operator diagonalised by the basis $\mathcal{B} = \{v_1, \dots, v_n\}$ then

$$[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \quad (2.2.3)$$

for some $\lambda_i \in \mathbb{k}$. We can also express the basis vectors v_i in this basis, they are simply the standard basis, in which $[v_i]_{\mathcal{B}} = e_i$ is the column vector with 1 in the i th position and 0 everywhere else. Then, for example, we have

$$[Tv_1]_{\mathcal{B}} = [T]_{\mathcal{B}}[v_1]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \vdots \\ 0 \end{bmatrix} = \lambda_1 \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix} = \lambda_1 [v_1]_{\mathcal{B}}. \quad (2.2.4)$$

Thus, we have $Tv_1 = \lambda_1 v_1$, that is, v_1 is an eigenvector of T with eigenvalue λ_1 . In general, v_i will be an eigenvector of T with eigenvalue λ_i .

Now make the following observations. We may define the subspaces $U_i = \text{span}\{v_i\}$. The fact that $V = \bigoplus_{i=1}^n U_i$ follows immediately from (and is equivalent to) the fact that \mathcal{B} is a basis. Each subspace U_i is **T -invariant**, meaning that $T(U_i) \subseteq U_i$, which simply means that if $u \in U_i$ then $Tu \in U_i$ also, so it's not possible to leave U_i just by the action of T . This is clear in this case because T acts on each subspace, U_i , by scalar multiplication, specifically, by multiplication by λ_i . Because of this it makes sense to consider the restricted linear map $T|_{U_i}: U_i \rightarrow U_i$, defined by $T|_{U_i}(u) = T(u)$ for $u \in U_i$. By definition $T(u) = \lambda_i u$ for $u \in U_i$, and thus we have $T|_{U_i} = \lambda_i \text{id}_{U_i}$ where $\text{id}_W: W \rightarrow W$ is the identity linear map.

2.2.1 Basis Independent Definition

We can now move towards making a basis independent definition of diagonalisability. To do so we need the notion of an eigenspace.

Definition 2.2.5 — Eigenspace Let \mathbb{k} be a field and let V be a vector space over V . Let $T: V \rightarrow V$ be a linear operator. Then for $\alpha \in \mathbb{k}$ define the **eigenspace** corresponding to α to be

$$W_{\alpha} := \{w \in V \mid T(w) = \alpha w\}. \quad (2.2.6)$$

Note that W_α can alternatively be characterised as

$$W_\alpha = \{w \in V \mid (T - \alpha)(w) = 0\} = \ker(T - \alpha) \quad (2.2.7)$$

where we perform the common abuse of notation writing $T - \alpha$ when we mean $T - \alpha \text{id}_V$.

For example, in the case of a diagonalisable operator where the eigenvalues are all distinct, as discussed at the end of the previous section, the subspaces U_i are exactly the eigenspaces W_{λ_i} . Also, $W_\alpha = 0$ if $\alpha \neq \lambda_i$ for any i . This is a general fact, if α is not an eigenvalue then there will be no $w \in V$ satisfying $T(w) = \alpha w$.

The problem is that eigenvalues needn't be distinct, in fact, eigenvalues are distinct if and only if the nonzero eigenspaces have dimension 1. Regardless of this problem, we can still perform the direct sum decomposition from the last part, replacing the U_i with the eigenspaces W_α . In fact, this gives us an equivalent definition of diagonalisability without reference to a basis.

Definition 2.2.8 — Diagonalisable Operator Let V be a finite dimensional vector space over \mathbb{k} . Then a linear operator, $T: V \rightarrow V$, is **diagonalisable** if

$$V = \bigoplus_{\alpha \in \mathbb{k}} W_\alpha. \quad (2.2.9)$$

2.2.2 Decomposition Theorem

We now ask if there is a generalisation of this decomposition to not-necessarily-diagonalisable operators. We'll start with an example.

Example 2.2.10 — Non-diagonalisable Operator Consider the case $\mathbb{k} = \mathbb{C}$, $V = \mathbb{C}^2$, and

$$T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad (2.2.11)$$

in the standard basis

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (2.2.12)$$

To find the eigenvalues we solve the characteristic polynomial, which is

$$\chi_T(t) = \det(\lambda - T) = \det \begin{bmatrix} \lambda & -1 \\ 0 & \lambda \end{bmatrix} = \lambda^2. \quad (2.2.13)$$

The only solution to $\lambda^2 = 0$ is $\lambda = 0$, and thus we have one repeated eigenvalue. This means that T is not diagonalisable. For $\alpha \neq 0$ we have $W_\alpha = 0$, and we have

$$W_0 = \{w \in V \mid Tw = 0\} = \ker(T - 0) = \ker T = \text{span}\{e_1\}. \quad (2.2.14)$$

This follows simply by considering the action of T on the basis vectors, $Te_1 = 0$ and $Te_2 = e_1 \neq 0$. This means that

$$\bigoplus_{\alpha \in \mathbb{C}} W_\alpha = W_0 = \text{span}\{e_1\} \neq \mathbb{C}^2. \quad (2.2.15)$$

From this calculation we see that $T^2e_2 = 0$, and so if we instead consider the operator T^2 then the eigenspaces of this cover all of \mathbb{C}^2 , so this suggests that we should look not just at T , but also powers of T in our decomposition.

Remark 2.2.16 The matrix

$$T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = E \quad (2.2.17)$$

is important for the Lie algebra $\mathfrak{sl}(2, \mathbb{C})$, which is one of the most important Lie algebras. The fact that this matrix is not diagonalisable is important in that it effects whether E commutes with other matrices. Recall that if two matrices can be simultaneously diagonalised then they commute, an important fact in quantum mechanics.

Following on from the idea that we should look at powers of T we replace the eigenspace with the following definition.

Definition 2.2.18 — Generalised Eigenspace Let \mathbb{k} be a field and let T be a vector space over \mathbb{k} . Let $T: V \rightarrow V$ be a linear operator. Then for $\alpha \in \mathbb{k}$ define the **generalised eigenspace** corresponding to α to be

$$V_\alpha := \{w \in V \mid (T - \alpha)^N(w) = 0 \text{ for some } N \in \mathbb{N}\}. \quad (2.2.19)$$

We can then often make a decomposition into a direct sum of generalised eigenspaces. There's just one hitch, this only works if \mathbb{k} is algebraically closed.

Definition 2.2.20 — Algebraically Closed Field A field, \mathbb{k} , is algebraically closed if every nonconstant polynomial $f(t) \in \mathbb{k}[t]$ has a root, $\alpha \in \mathbb{k}$, such that $f(\alpha) = 0$.

The complex numbers, \mathbb{C} , are algebraically closed, this is the fundamental theorem of algebra. The real numbers, \mathbb{R} , are not algebraically closed, for example $t^2 + 1$ has no (real) roots.

The following proof relies on three standard results, which we state without proof.

Theorem 2.2.21 — Cayley–Hamilton. If $\chi_T(t) = t^n + a_{n-1}t^{n-1} + \dots + a_1t + a_0$ is the characteristic polynomial of the linear operator $T: V \rightarrow V$ then the

linear operator

$$T^n + a_{n-1}T^{n-1} + \cdots + a_1T + a_0\text{id}_V : V \rightarrow V \quad (2.2.22)$$

is the zero map. That is, T satisfies its own characteristic polynomial, $\chi_T(T) = 0$.

Lemma 2.2.23 — Bézout's Lemma If R is a principal ideal domain and $x, y \in R$ have greatest common divisor d then there exist $a, b \in R$ such that $ax + by = d$.

Lemma 2.2.24 If \mathbb{k} is a field then $\mathbb{k}[x]$ is a PID, and in particular Bézout's lemma applies to polynomials.

Theorem 2.2.25 — Decomposition Theorem. Let \mathbb{k} be an algebraically closed field, V a finite dimensional vector space, and $T : V \rightarrow V$ a linear operator. Then

$$V = \bigoplus_{\alpha \in \mathbb{k}} V_\alpha \quad (2.2.26)$$

where V_α is the generalised eigenspace associated with α and only finitely many of the V_α are nonzero.

Proof. We start by showing that $V_\alpha \cap \sum_{\beta \in \mathbb{k} \setminus \{\alpha\}} V_\beta = 0$. Take $v \in V_\alpha \cap \sum_{\beta \in \mathbb{k} \setminus \{\alpha\}} V_\beta$, then we have $v \in V_\alpha$ and $v \in V_{\beta_1} + \cdots + V_{\beta_r}$ for some $\beta_1, \dots, \beta_r \in \mathbb{k} \setminus \{\alpha\}$. This means that there exist some $N, N_1, \dots, N_r \in \mathbb{N}$ such that

$$(T - \alpha)^N v = (T - \beta_1)^{N_1} \cdots (T - \beta_r)^{N_r} v = 0. \quad (2.2.27)$$

Since $\alpha \neq \beta_i$ we know that the greatest common factor of $(t - \alpha)^N$ and $(t - \beta_1)^{N_1} \cdots (t - \beta_r)^{N_r}$ is 1. Thus, by Bézout's lemma there exist polynomials $f(t), g(t) \in \mathbb{k}[t]$ such that

$$f(t)(t - \alpha)^N + g(t)(t - \beta_1)^{N_1} \cdots (t - \beta_r)^{N_r} = 1. \quad (2.2.28)$$

Evaluating at T and applying this to v we then have

$$f(T)(T - \alpha)^N v + g(T)(T - \beta_1)^{N_1} v \cdots (T - \beta_r)^{N_r} v = v, \quad (2.2.29)$$

and we know that these operators acting on v both give zero, so the left hand side vanishes, and thus $v = 0$, and so $V_\alpha \cap \sum_{\beta \in \mathbb{k} \setminus \{\alpha\}} V_\beta = 0$.

We now show that $\sum_{\alpha \in \mathbb{k}} V_\alpha = V$. Factorise the characteristic polynomial, $\chi_T(t)$, as follows

$$\chi_T(t) = (t - \alpha_1)^{N_1} \cdots (t - \alpha_s)^{N_s}, \quad (2.2.30)$$

with $\alpha_i \neq \alpha_j$ for $i \neq j$. Note that the existence of such a factorisation relies on \mathbb{k} being algebraically closed. We claim that $v \in V_{\alpha_1} + \dots + V_{\alpha_s}$. This can be proven by induction on s . The basis case, $s = 1$, has the characteristic polynomial factorise as

$$\chi_T(t) = (t - \alpha_1)^{N_1}. \quad (2.2.31)$$

Then by the Cayley–Hamilton theorem we have

$$(T - \alpha_1)^{N_1} v = 0 \quad (2.2.32)$$

for all $v \in V$, and thus $V_{\alpha_1} = V$. For the inductive step suppose that $s > 1$ and the statement is true for $s - 1$. The highest common factor of $(t - \alpha_s)^{N_s}$ and $(t - \alpha_1)^{N_1} \dots (t - \alpha_{s-1})^{N_{s-1}}$ is 1. Thus, by Bézout’s lemma there exist $f(t), g(t) \in \mathbb{k}[t]$ such that

$$f(t)(t - \alpha_s)^{N_s} + g(t)(t - \alpha_1)^{N_1} \dots (t - \alpha_{s-1})^{N_{s-1}} = 1. \quad (2.2.33)$$

Evaluating at T and applying the map to $v \in V$ we have

$$f(T)(T - \alpha_s)^{N_s} v + g(T)(T - \alpha_1)^{N_1} \dots (T - \alpha_{s-1})^{N_{s-1}} v = v. \quad (2.2.34)$$

Define

$$v' = f(T)(T - \alpha_s)^{N_s} v, \quad \text{and} \quad v_s = g(T)(T - \alpha_1)^{N_1} \dots (T - \alpha_{s-1})^{N_{s-1}} v, \quad (2.2.35)$$

so we have $v = v' + v_s$. Note that we have $(T - \alpha_s)^{N_s} v_s = 0$ and $(T - \alpha_1)^{N_1} \dots (T - \alpha_{s-1})^{N_{s-1}} v' = 0$. This is simply the Cayley–Hamilton theorem applied to the linear map T restricted to the subspaces corresponding to v_s and v' respectively. Define

$$V' = \{u \in V \mid (T - \alpha_1)^{N_1} \dots (T - \alpha_{s-1})^{N_{s-1}} u = 0\}. \quad (2.2.36)$$

Then $v_s \in V_{\alpha_s}$ and $v' \in V'$. The characteristic polynomial of the restricted map, $T|_{V'}$, has all of its roots in $\{\alpha_1, \dots, \alpha_{s-1}\}$. By induction, we therefore have $v' = v_1 + \dots + v_{s-1}$ with $v_i \in V'_{\alpha_i} = V' \cap V_{\alpha_i}$. Thus, $v = v_1 + \dots + v_{s-1} + v_s \in V_{\alpha_1} + \dots + V_{\alpha_s}$. Since $v \in V$ was arbitrary this shows that $V = V_{\alpha_1} + \dots + V_{\alpha_s}$, proving that $V = V_{\alpha_1} \oplus \dots \oplus V_{\alpha_s}$. This gives the final result by realising that if $\beta \neq \alpha_i$ then $V_\beta = 0$ since $V_\beta \cap (V_{\alpha_1} + \dots + V_{\alpha_s}) = 0$ so $V_\beta \cap V = 0$, so $V_\beta = 0$. Thus, $V = \bigoplus_{\alpha \in \mathbb{k}} V_\alpha$ and V_α is only nonzero for the finite index set $\{\alpha_1, \dots, \alpha_s\}$. \square

Note that both conditions, the finite dimension of V and the algebraic closure of \mathbb{k} , are required here. Dropping either condition will result in cases where this decomposition doesn’t exist.

Example 2.2.37 Consider again the example of $\mathbb{k} = \mathbb{C}$, $V = \mathbb{C}^2$, and

$$T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \quad (2.2.38)$$

As we saw before (Example 2.2.10) T is not diagonalisable, and $\bigoplus_{\alpha \in \mathbb{C}} W_\alpha = W_0 = \text{span}\{e_1\} \neq \mathbb{C}^2$. We also saw there that $T^2 e_2 = 0$, meaning that when considering generalised eigenspaces we also have $e_2 \in V_0$, and thus

$$\bigoplus_{\alpha \in \mathbb{C}} V_\alpha = V_0 = \mathbb{C}^2. \quad (2.2.39)$$

Definition 2.2.40 — Jordan Block and Jordan Normal Form A **Jordan block** is a matrix of the form

$$\begin{bmatrix} \alpha & 1 & & & \\ & \alpha & 1 & & \\ & & \ddots & \ddots & \\ & & & \alpha & 1 \\ & & & & \alpha \end{bmatrix}. \quad (2.2.41)$$

That is, we have some $\alpha \in \mathbb{k}$ on the diagonal and 1s above the diagonal, and 0 everywhere else.

A matrix, X , is a sum of Jordan blocks, or is in **Jordan normal form** if it is given by the block diagonal matrix

$$X = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{bmatrix} \quad (2.2.42)$$

with the J_i Jordan blocks.

Corollary 2.2.43 — Existence of Jordan Normal Form Let \mathbb{k} be an algebraically closed field, let V be a finite dimensional vector space, and let $T: V \rightarrow V$ be a linear map. There exists a basis of V such that $[T]_{\mathcal{B}}$ is in Jordan normal form.

Proof. Consider the decomposition of V into a sum of generalised eigenspaces of T . Each generalised eigenspace is invariant under T , and thus if $v_\alpha \in V_\alpha$ then $Tv_\alpha = c_1 v_1 + \cdots + c_r v_r$ where $\{v_1, \dots, v_r\}$ is a basis for V_α . We can further choose this basis by scaling and taking linear combinations until $Tv_i = \alpha v_i + v_{i+1}$, which can be shown by induction on the size of the dimension of V_α . Then T acts on V_α in this basis as a Jordan block, and so T acts on V , in the basis formed by the union of these bases, as a matrix in Jordan normal form. \square