Complex Variables and Vector Spaces

Dominik Szablonski

February 13, 2025

Contents

Chapter 1

Vector Spaces

We wish to generalise the idea of a vector and field. Let us first define a field,

Definition 1: Fields

A field $\mathbb{F}$ is a set with 2 binary operations defined on it, addition $(+)$ and multiplication $(\cdot)$. The following axioms hold $\forall a, b, c \in \mathbb{F}$,

1. *Associativity*,

$$a + (b + c) = (a + b) + c \qquad a \cdot (b \cdot c) = (a \cdot b) \cdot c \tag{1.1}$$

2. *Commutativity*,

$$a + b = b + a \qquad a \cdot b = b \cdot a \tag{1.2}$$

3. *Identity.* $\exists 0, 1 \in \mathbb{F}$ such that,

$$a + 0 = a \qquad a \cdot 1 = a \tag{1.3}$$

4. *Additive inverse.* $\forall a \in \mathbb{F}, \exists - a \in \mathbb{F}$ such that,

$$a + (-a) = 0. \tag{1.4}$$

5. *Multiplicative inverse.* $\forall a \in \mathbb{F}, \exists a^{-1} \in \mathbb{F}$ such that,

$$a \cdot a^{-1} = 1. \tag{1.5}$$

We can then define a vector space,

Definition 2: Vector Space

Let $\mathbb{F}$ be a field. A vector space V over $\mathbb{F}$ is a set of objects $\mathbf{u}, \mathbf{v}, \mathbf{w}, \ldots$ which satisfy,

1. *Addition.* The set is closed under addition, such that $\mathbf{u}, \mathbf{v} \in V \implies \mathbf{w} = \mathbf{u} + \mathbf{v} \in V$. This operation is commutative and associative.

2. *Scalar multiplication.* The set is closed under multiplication by a scalar, i.e., $\mathbf{u} \in V \implies \lambda \mathbf{u} \in V$ for $\lambda \in \mathbb{F}$. Scalar multiplication is associative and distributive.

3. *Null vector.* $\exists \mathbf{0}, \mathbf{u} + \mathbf{0} = \mathbf{u}$.

4. *Negative vector.* $\forall \mathbf{u} \in V, \exists - \mathbf{u} \in V$ such that,

$$\mathbf{u} + (-\mathbf{u}) = 0. \tag{1.6}$$

1.1 Linear Independence

If vectors are linearly independent, then they cannot be written as a combination of each other. Let us write down the formal definition,

Definition 3: Linear Independence

A set of vectors $\{\mathbf{u}_i \text{ for } i = 1, 2, \ldots, n\}$ is linearly independent if the equation,

$$\sum_j^n \lambda_j \mathbf{u}_j = \mathbf{0} \tag{1.7}$$

has only 1 solution, $\forall i : \lambda_i = 0$.

1.2 Postulate of Dimensionality and Basis Vectors

Definition 4: Dimensionality

A vector space V has dimensions N if it can accommodate no more than N linearly independent vectors $\mathbf{u}_j$.

We often denote N dimensional vector spaces over a field $\mathbb{F}$ as $\mathbb{F}^N$, or more generally V_N. We are often also interested in the *span* of a vector space.

Definition 5: Span

The span of a set of vectors $\{\mathbf{u}_i, fori = 1, 2, \ldots, n\}$ is the set of all vectors which can be written as a linear combination of $\mathbf{u}_i$.

The above definition naturally leads to the below theorem,

Theorem 1: I

an N-dimensional vector space V_N, any vector $\mathbf{u}$ can be written as a linear combination of N linearly independent basis vectors $\mathbf{e}_j$.

Proof. Since there are no more than N linearly independent vectors, the set of vectors $\{\mathbf{e}_i\}_{i=1}^N + \mathbf{u}$ must be linearly dependent. Therefore, there must be a relation of the form,

$$\sum_{i=1}^N \lambda_i \mathbf{e}_i + \lambda_0 \mathbf{u} = \mathbf{0}, \tag{1.8}$$

where $\mathbf{u} \in V_N$ is an arbitrary vector and $\exists \lambda_i \neq 0$. From the definition of linear dependence, we require $\lambda_0 \mathbf{u}_0 \neq 0$, so,

$$\mathbf{u} = -\frac{1}{\lambda_0} \sum_{i=1}^N \lambda_i \mathbf{e}_i = \sum_i^N u_i \mathbf{e}_i \tag{1.9}$$

where $u_i = -\frac{\lambda_i}{\lambda_0}$. □

From the above theorem, we are able to define the **basis** of a vector space,

Definition 6: Basis

Any set of N linearly independent vectors in V_n is called a **basis**, and then **span** V_N, or synonymously, they are **complete** if N is finite.

This allows us to write any vector $\mathbf{v} \in V_N$ as,

$$\mathbf{v} = \sum_{i}^{N} v_i \mathbf{e}_i \tag{1.10}$$

where $\mathbf{e}_i$ is any complete basis.

1.3 Linear Subspaces

We can consider a subspace of V_N as a vector space spanned by a set of $M < N$ linearly independent vectors. The subspace V_M must satisfy the following properties,

1. It must contain the zero vector $\mathbf{0}$.
2. It must be closed under addition and scalar multiplication.

An example of a subspace would be the subspace of $\mathbb{R}^3$ which is the set of vectors $(x, y, 0)$, where $x, y \in \mathbb{R}$ which define the xy-plane in $\mathbb{R}^3$. This is a case of a more general result,

Theorem 2: Subspaces

Any set of M ($M \leq N$) linearly independent vectors $\{\mathbf{e}_i\}_{i=1}^{M}$ in V_N span a subspace V_M of V_N.

However, counterexamples do exist such as the set of vectors lying within a unit circle $\{(x, y) : x^2 + y^2 \leq 1\}$ which cannot be a subspace of $\mathbb{R}^3$ This is because we can choose a λ such that λx_1 or $\lambda y_1 > 1$ lies outside of the unit circle, and thus is not closed under multiplication.

1.4 Normed Spaces

We wish to now generalise length in order to define the closeness of vectors. We do this by defining a *norm*.

Definition 7: Norm

Give a vector space V over a field $\mathbb{F}$, a norm on V is a real-valued function $p : V \to \mathbb{R}$ with the following properties,

1. **Triangle Inequality**, $p(\mathbf{x} + \mathbf{y}) \leq p(\mathbf{x}) + p(\mathbf{y}), \forall \mathbf{x}, \mathbf{y} \in V$
2. **Absolute Homogeneity**, $p(sx) = |s| p(\mathbf{x}), \forall \mathbf{x} \in V, \forall s \in \mathbb{R}$.
3. **Positive Definiteness**, $\forall \mathbf{x} \in V, p(x) \geq 0; p(x) = 0 \iff x = 0$.

For a vector space V_N and two vectors $\mathbf{u}, \mathbf{v} \in V_N$, the distance between them is given by $\|\mathbf{u} - \mathbf{v}\|$. There are different types of norms, some of which are defined in sections below.

1.4.1 Supremum Norm

$\forall \mathbf{x} \in V_N$ where x_i are the components in a given basis, the we define the *supremum* or *infinity* norm.

Definition 8: Supremum Norm

$$\|\mathbf{x}\|_S = \|\mathbf{x}\|_\infty = \max_i |x_i|. \tag{1.11}$$

It can be shown that, since $|a + b| \leq |a| + |b|$ $\forall a, b \in \mathbb{R}$ or $\forall a, b \in \mathbb{C}$,

$$\begin{aligned} \|\mathbf{x} + y\| = \max_i |x_i + y_i| &\leq \max_i (|x_i| + |y_i|) \\ &\leq \max_i |x_i| + \max_j |y| \end{aligned} \tag{1.12}$$

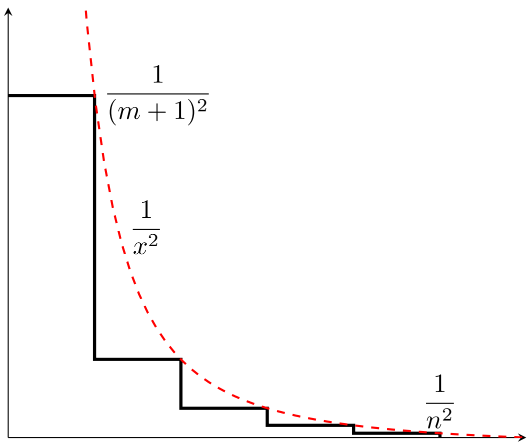

Figure 1.1: Graphical proof used in example 1.

1.4.2 1-Norm

$\forall \mathbf{x} \in V_N$ where x_i are the components of $\mathbf{x}$, we define the 1-norm,

Definition 9: 1-Norm

$$\|x\|_1 = \sum_{i=1}^{N} |x_i|. \tag{1.13}$$

1.5 Completeness

1.5.1 Cauchy Sequences

Definition 10: Cauchy Sequence

A sequence $\{a_n\}_{n=0}^{\infty}$, $a_n \in V$ and V is a normed vector space is Cauchy if $\forall \epsilon > 0, \exists N > 0$ such that $\forall n, m > N, \|a_n - a_m\| < \epsilon$.

Let us consider some sequences and show if they are Cauchy.

Sequences over $\mathbb{R}$

Example 1: $a_n = \sum_{i=1} \frac{1}{i^2}$

A sequence in $\mathbb{R}$ with $\|a\| = |a|$ is

$$a_n = \sum_{i=1}^{n} \frac{1}{i^2}. \tag{1.14}$$

Is this sequence Cauchy?

For $n > m$, let us write,

$$|a_n - a_m| = \sum_{i=m=1}^{n} \frac{1}{i^2} \tag{1.15}$$

If we consider the sum as the integral over a series of step functions, then we can consider an approximation of this integral as $\frac{1}{x^2}$, as in figure 1.1. Thus,

$$\begin{aligned}\sum_{i=m+1}^{n} \frac{1}{i^2} &\leq \int_m^n \frac{1}{x^2}\,\mathrm{d}x \\ &= \frac{1}{n} - \frac{1}{m} \leq \frac{1}{n} \leq \frac{1}{N}.\end{aligned} \tag{1.16}$$

Let us now choose $N > \frac{1}{\epsilon}$, so that we find,

$$|a_n - a_m| < \epsilon \tag{1.17}$$

thus the sequence is Cauchy. $\square$

Example 2: $a_n = n$

onsider a sequence $a_n = n$. Is this sequence Cauchy?

Let us choose $\epsilon = 1$, $n = N + 1$, and $m = N + 3$

$$|a_n - a_m| = 2 > \epsilon \tag{1.18}$$

so the sequence is not Cauchy. $\square$

Cauchy sequences of functions

We can also apply similar proofs to functions.

Example 3: $f : [0, 1] \to \mathbb{R}$, $f_n(x) = \frac{x}{n}$.

Consider $f : [0, 1] \to \mathbb{R}$ where $f_n(x) = \frac{x}{n}$. Is this function Cauchy?

Let $n > m$,

$$\begin{aligned}\|f_n - f_m\|_1 &= \int_0^1 \left|\frac{x}{n} - \frac{x}{m}\right| \mathrm{d}x \\ &= \left|\frac{1}{n} - \frac{1}{m}\right| \int_0^1 x\,\mathrm{d}x \\ &= \frac{1}{2}\left|\frac{1}{n} - \frac{1}{m}\right| \leq \frac{1}{2}\left(\left|\frac{1}{n}\right| + \left|\frac{1}{m}\right|\right) \leq \frac{1}{2}\frac{2}{N} = \frac{1}{N}.\end{aligned} \tag{1.19}$$

Choose $N > 1/\epsilon \implies \|f_n - f_m\| < \epsilon$, so f is Cauchy.

1.5.2 Cauchy Sequences and Convergence

Every convergent sequence is Cauchy, because if $a_n \to x \implies \|a_m - a_n\| \leq \|a_m - x\| + \|x - a_n\|$ both of which go to zero. Whether every Cauchy sequence is convergent gives rise to the following definition,

Definition 11: Completeness

A field is complete if every Cauchy sequence in the field converges to an element of the field.

Example 4: Completeness of $\mathbb{Q}$

Consider $a_n = \frac{a_{n-1}}{2} + \frac{1}{a_{n-1}}$. Let us assume a_∞ exists.

$$a_\infty = \frac{a_\infty}{2} + \frac{1}{a_\infty} \tag{1.20}$$

$\implies \frac{1}{2}a_\infty^2 = 1 \implies a_\infty = \sqrt{2} \notin \mathbb{Q} \therefore \mathbb{Q}$ is not complete. $\square$

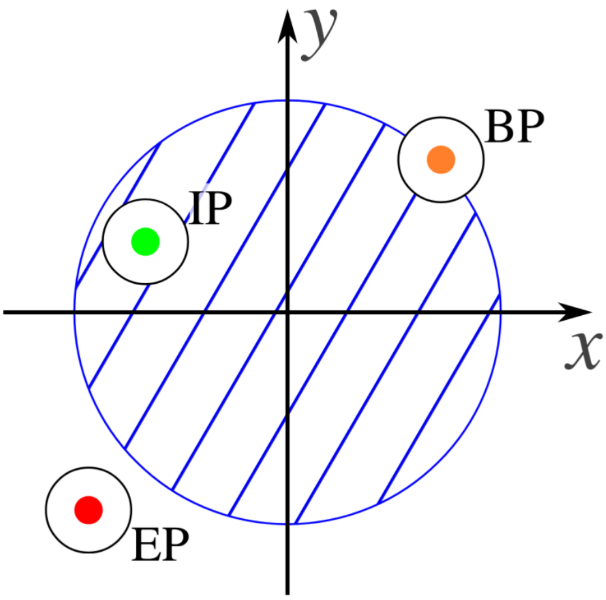

Figure 1.2: Interior point (IP), exterior point (EP), and boundary point (BP).

1.6 Open and Closed Sets

Now that we have defined completeness, let us look at the difference between open and closed sets, particularly on the 2D plane. We will be considering a ball in the 2D plane, defined,

Definition 12: Ball

A ball of radius ϵ around a point $\mathbf{r}_0$ is the set of all points $\mathbf{r}$ such that $\|\mathbf{r} - \mathbf{r}_0\|$.

A sphere is the points where $\|\mathbf{r} - \mathbf{r}_0\| = \epsilon$. Let us denote the set of the sphere S. We will consider three types of points, visualised in figure 1.2,

- **Exterior point**, for some ϵ, all $\mathbf{r} \notin S$.
- **Interior point**, for some ϵ, all $\mathbf{r} \in S$.
- **Boundary point**, for some ϵ, some of the neighbourhood of $\mathbf{r} \in S$ and some $\mathbf{r} \notin S$.

We can then define closed and open sets.

Definition 13: Closed Set

A set that contains all its boundary points is closed.

An example of this is a set of points $|r| \leq 1$, as $|r| = 1$ is a boundary point, and also belongs to the set.

Definition 14: Open Set

A set that only includes interior points is open.

We must furthmore define,

Definition 15: Connected Set

Sets for which any two points can be joined by a continuous path.

If a set is connected and open, we call it a *region*.

Example 5

The function $f(z) = \frac{1}{(1-z)}$ has a defined Taylor series for $z \neq 1$,

$$f(z) = \sum_{i=0}^{\infty} z^i. \tag{1.21}$$

For what complex numbers is this series Cauchy? Is this an open or closed set?

We will consider the cases $|z| < 1$ and $|z| > 1$ separately, with $|z| = 1$ as a boundary case. Let us define,

$$a_n = \sum_{i=0}^{n} z^i. \tag{1.22}$$

For any $z \neq 1$, assuming $n > m$,

$$|a_n - a_m| = \left| \sum_{i=m+1}^{n} z_i \right| = \left| \frac{z^{m+1} - z^{n+1}}{1-z} \right|. \tag{1.23}$$

For $|z| < 1$,

$$|a_n - a_m| = \frac{|z|^m}{|1-z|} \left|1 - z^{n-m+1}\right| \leq \frac{2}{|1-z|} |z|^m \tag{1.24}$$

and since $|z|^m$ is decreasing as a function of m, the series is Cauchy. For $|z| > 1$,

$$|a_n - a_m| = \frac{|z|^n}{|1-z|} \left|1 - z^{-n+m+1}\right| \geq \frac{2}{|1-\frac{1}{z}|} |z|^n = z^{n+1} \tag{1.25}$$

and since $|z|^n$ is an increasing function of n, the series is not Cauchy. Thus the series is Cauchy in the open set $|z| < 1$.

Chapter 2

Inner Product Space

An inner product space is a vector space with an inner product, which is a generalisation of the scalar product.

Definition 16: Inner product, $\langle \mathbf{a}, \mathbf{b} \rangle$

Given a vector space V_N over $\mathbb{F}$, the inner product between two vectors $\mathbf{a}, \mathbf{b} \in V_N$ is a function such that $V \times V \to \mathbb{F}$. If $\mathbb{F} \subset \mathbb{C}$, the following properties hold,

1. **Linearity**. If $\mathbf{w} = \lambda \mathbf{u} + \mu \mathbf{v}$ then $\langle \mathbf{a}, \mathbf{w} \rangle = \lambda \langle \mathbf{a}, \mathbf{u} \rangle + \mu \langle \mathbf{a}, \mathbf{u} \rangle$.
2. **Conjugation Symmetry.** $\overline{\langle \mathbf{w}, \mathbf{a} \rangle} = \langle \mathbf{a}, \mathbf{w} \rangle$
3. **Positive Definiteness.** $\forall \mathbf{x} \neq 0, \langle \mathbf{x}, \mathbf{x} \rangle > 0$.

From our definition of the inner product, we can define the 2-norm,

$$\|\mathbf{a}\|^2 = \langle \mathbf{a}, \mathbf{a} \rangle \geq 0. \tag{2.1}$$

2.1 Orthogonality

Definition 17: Orthogonality

$\forall \mathbf{a}, \mathbf{b} \neq 0 \in V_N$ if $\langle \mathbf{a}, \mathbf{b} \rangle = 0$ then $\mathbf{a}$ and $\mathbf{b}$ are orthogonal.

This allows us to then define an orthonormal basis.

Definition 18: Orthonormal basis

The set basis vectors $\{\mathbf{e}_i\}_{i=1}^N \in V_N$ is orthogonal if,

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = A_i \delta_{ij}. \tag{2.2}$$

and $A_i \neq 0$. The set of basis vectors is orthonormal for $A_i = 1, \forall i \in [1, N]$.

Given we can decompose any vector $\mathbf{a} \in V_N$ if given a complete set of basis vectors, we can define a general inner product for V_N over $\mathbb{F} \subset \mathbb{C}$. Let us begin by writing the decomposition of two vectors $\mathbf{a}, \mathbf{b} \in V_N$ into a set of basis vectors $\{\mathbf{e}_j\}_{j=1}^N$,

$$\mathbf{a} = \sum_{j=1}^N a_j \mathbf{e}_j \qquad \mathbf{b} = \sum_{j=1}^N b_j \mathbf{e}_j. \tag{2.3}$$

Then, using linearity,

$$\begin{aligned}\langle \mathbf{a}, \mathbf{b}\rangle &= \sum_{j,k=1}^{N} \overline{a}_j \langle \mathbf{e}_j, \mathbf{e}_k\rangle b_k \\ &= \sum_{i,j=1}^{N} \overline{a}_j \delta_{jk} b_k \\ &= \sum_{j=1}^{N} \overline{a}_j b_j .\end{aligned} \tag{2.4}$$

NOTE: This only holds when using an orthonormal basis.

We can obtain further insight into the decomposition of a vector by considering the inner product,

$$\mathbf{a} = \sum_{j=1}^{N} a_j \mathbf{e}_j \implies \langle \mathbf{e}_k, \mathbf{a}\rangle = \sum_{j=1}^{N} a_j \underbrace{\langle \mathbf{e}_j, \mathbf{e}_k\rangle}_{\delta_{jk}} = a_k. \tag{2.5}$$

We often refer to $a_k = \langle \mathbf{e}_k, \mathbf{a}\rangle$ as the *projection* of $\mathbf{a}$ onto $\mathbf{e}_k$ as it gives the component of $\mathbf{a}$ in the $\mathbf{e}_k$ direction.

2.2 Gram-Schmidt Orthonormalisation

Definition 19: Gram-Schmidt Algorithm

Given a basis $\{\mathbf{v}_j\}_{j=1}^{N} \in V_N$,

1. Define

$$\mathbf{e}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} \tag{2.6}$$

2. Define

$$\mathbf{u}_2 = \mathbf{v}_2 - \langle \mathbf{e}_1, \mathbf{e}_2\rangle \mathbf{e_1} \qquad\qquad \mathbf{e}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} \tag{2.7}$$

⋮

m. Define,

$$\mathbf{u}_m = \mathbf{v}_m - \sum_{j=1}^{m-1} \langle \mathbf{e}_j, \mathbf{v}_m\rangle \mathbf{e}_j \tag{2.8}$$

thus,

$$\mathbf{e}_m = \frac{\mathbf{u}_m}{\|\mathbf{u}_m\|} \tag{2.9}$$

up to N.

The Gram-Schmidt process is able to take any set of basis vectors and turn it into a set of orthonormal basis vectors. The idea behind it is that given 2 vectors $\mathbf{v}, \mathbf{u}$ such that $\|\mathbf{u}\| = 1$, then we wish to define a vector $\mathbf{v}' = \mathbf{v} - \langle \mathbf{u}, \mathbf{v}\rangle \mathbf{u}$. The inner product with $\mathbf{u}$ and this new vector is then,

$$\langle \mathbf{u}, \mathbf{v}'\rangle = \langle \mathbf{u}, \mathbf{v}\rangle - \langle \mathbf{u}, \mathbf{v}\rangle \cancelto{1}{\langle \mathbf{u}, \mathbf{u}\rangle} = 0. \tag{2.10}$$

So, we essentially are removing the non-orthonormal components from each subsequent basis vector, based on the first basis vector in the set.

2.3 Inequalities of Inner Product Space

Theorem 3: Cauchy-Schwartz Inequality

$\forall \mathbf{a}, \mathbf{b} \in V_N, |\langle \mathbf{a}, \mathbf{b}\rangle| \leq \|\mathbf{a}\|\|\mathbf{b}\|$.

Proof. Consider $\mathbf{u} = \mathbf{a} - \lambda\mathbf{b}$,

$$\|\mathbf{a}\|^2 = \|\mathbf{a}\|^2 + |\lambda|^2\|\mathbf{b}\|^2 - \overline{\lambda}\langle \mathbf{b}, \mathbf{a}\rangle - \lambda\langle \mathbf{a}, \mathbf{b}\rangle \geq 0. \tag{2.11}$$

Choose,

$$\lambda = \frac{\langle \mathbf{b}, \mathbf{a}\rangle}{\|\mathbf{b}\|^2}. \tag{2.12}$$

Thus,

$$\|\mathbf{u}\|^2 = \|\mathbf{a}\|^2\|\mathbf{b}\|^2 - \frac{|\langle \mathbf{a}, \mathbf{b}\rangle|}{\|\mathbf{b}\|^2} \geq 0 \tag{2.13}$$

$\implies |\langle \mathbf{a}, \mathbf{b}\rangle| \leq \|\mathbf{a}\|\|\mathbf{b}\|$. □

Theorem 4: Triangle Inequality

$\forall \mathbf{a}, \mathbf{b} \in V_N, \|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|$

Proof.

$$\begin{aligned} \|\mathbf{a} + \mathbf{b}\|^2 &= \langle \mathbf{a} + \mathbf{b}, \mathbf{a} + \mathbf{b}\rangle \\ &\leq \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 + 2|\langle \mathbf{a}, \mathbf{b}\rangle| \\ &\leq \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 + 2\|\mathbf{a}\|\|\mathbf{b}\| = (\|\mathbf{a}\| + \|\mathbf{b}\|)^2 \end{aligned} \tag{2.14}$$

$\implies \|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|$ □

Chapter 3

Linear Operators

Definition 20: Map

A map M is a function which takes $V \to W$ where V, W are vector spaces. This is such that M acting on $\mathbf{c} \in V$ produces a different vector $\mathbf{c}' \in W$.

We are often interested in linear maps, or linear operators as they are often called in physics.

Definition 21: Linear Operator

An operator $\hat{A}$ is linear if, for a vector $\mathbf{c} = \lambda\mathbf{a} + \mu\mathbf{b}$ where $\forall\lambda, \mu \in \mathbb{F}$ and $\forall\mathbf{a}, \mathbf{b} \in V$ over $\mathbb{F}$,

$$\mathbf{c}' = \hat{A}\mathbf{c} = \mu(\hat{A}\mathbf{a}) + \lambda(\hat{A}\mathbf{b}) \tag{3.1}$$

In physics, but not generally, linear operators always map $V \to V$.

3.1 Algebra of operators

We can define the following properties of linear operators,

1. $(\hat{A} + \hat{B}) \equiv \hat{A}\mathbf{v} + \hat{B}\mathbf{v}$.
2. $()\lambda\hat{A})\mathbf{v} = \lambda(\hat{A}\mathbf{v})$.
3. $\exists\hat{\mathbb{1}} : \hat{\mathbb{1}}\mathbf{v} = \mathbf{v}, \forall\mathbf{v} \in V$.
4. $\exists\hat{0} : \hat{0}\mathbf{v} = \mathbf{0}, \forall\mathbf{v} \in V$.
5. $\forall\mathbf{v} \in V, (\hat{B}\hat{A})\mathbf{v} = \hat{B}(\hat{A}\mathbf{v})$, generally $\hat{A}\hat{B} \neq \hat{B}\hat{A}$

3.2 Matrix Representation of Linear Operators

Consider V_N over $\mathbb{F}$. Assume $\{\mathbf{e}_j\}_{j=1}^N$ is an orthonormal basis. $\forall\mathbf{v} \in V_N$,

$$\mathbf{v} = \sum_{j=1}^{N} v_j\mathbf{e}_j \tag{3.2}$$

where $\langle\mathbf{e}_k, \mathbf{v}\rangle = v_k$. Given a linear operator $\hat{A}$,

$$\hat{A}\mathbf{v} = \hat{A}\left(\sum_{j=1}^{N} v_j\mathbf{e}_j\right) = \sum_{j=1}^{N} v_j\hat{A}\mathbf{e}_j \tag{3.3}$$

by linearity. We have $\hat{A}\mathbf{e}_j \in V_N$, thus,

$$\hat{A}\mathbf{e}_j = \sum_{i=1}^{N} \mathbf{e}_i \underbrace{(\hat{A}\mathbf{e}_j)}_{A_{ij}} \tag{3.4}$$

so,

$$\begin{aligned} \hat{A}\mathbf{e}_j &= \sum_{i=1}^{N} \mathbf{e} A_{ij} \\ \implies A_{ij} &= \langle \mathbf{e}_i, \hat{A}\mathbf{e}_j \rangle. \end{aligned} \tag{3.5}$$

So,

$$\begin{aligned} \hat{A}\mathbf{v} &= \sum_{j=1}^{N} v_j \sum_{i=1}^{N} (A_i j \mathbf{e}_i) \\ &= \sum_{i=1}^{N} \left(\sum_{j=1}^{N} A_{ij} v_j \right) \mathbf{e}_i \end{aligned} \tag{3.6}$$

thus,

$$(\hat{A}\mathbf{v})_i = \sum_{j=1}^{N} A_{ij} v_j. \tag{3.7}$$

Interestingly, we have obtained matrix multiplication from linearity and orthonormality, rather than having to define it.

3.3 Inverse Operator

Definition 22: Inverse operator

Given an operator $\hat{A}$ from $V \to V$, $\forall \mathbf{v} \in V$, $\exists \hat{B}$ such that, $\hat{B}(\hat{A}\mathbf{v}) = \mathbf{v}$, $(\hat{B}\hat{A})\mathbf{v} = \mathbf{v} \implies \hat{B}\hat{A} = \mathbb{1}$. We call $\hat{B} = \hat{A}^{-1}$.

NOTE: Not all operators have an inverse.

3.4 Dyadic (Outer) Product

Recall $\langle u| \hat{A} |v\rangle \to \mathbf{u}^\dagger \hat{A}\mathbf{v}$, or $\langle \mathbf{u}, \hat{A}\mathbf{v} \rangle$. We can clearly see that $\mathbf{c}\langle \mathbf{a}, \mathbf{b} \rangle$ is a vector. Let us define,

$$\mathbf{c}' = \mathbf{c}(\mathbf{a}^\dagger \mathbf{b}) = \underbrace{(\mathbf{c}\mathbf{a}^\dagger)}_{\substack{\text{Linear} \\ \text{Operator}}} \mathbf{b} \tag{3.8}$$

this linear operator is known as a *dyad* or the *outer product* of $\mathbf{a}$ and $\mathbf{c}$. Let us define it more formall,y

Definition 23: Dyad/Outer product

For $V_N \subset \mathbb{C}^N$, and vectors $\mathbf{a}, \mathbf{b} \in V_N$, their outer product is defined $\mathbf{a}\mathbf{b}^\dagger$. It has the following properties,

1. **Linearity in the first argument**. If $\mathbf{c} = \lambda \mathbf{a} + \mu \mathbf{b}$, then $\mathbf{c}\mathbf{d}^\dagger = \lambda \mathbf{a}\mathbf{d}^\dagger + \mu \mathbf{b}\mathbf{d}^\dagger$.
2. **Antilinearity in the second argument**. If $\mathbf{d} = \lambda \mathbf{a} + \mu \mathbf{b}$, then $\mathbf{c}\mathbf{d}^\dagger = \bar{\lambda}\mathbf{c}\mathbf{a}^\dagger + \bar{\mu}\mathbf{c}\mathbf{b}^\dagger$.

3.5 Projection Operator and the Completeness Relation

If $\hat{P}^2\mathbf{v} = \hat{P}\mathbf{v}$, $\forall \mathbf{v} \in V_N$, then $\hat{P}$ is a projection operator. For a given orthonormal basis $\{\mathbf{e}_j\}_{j=1}^N \in V_N$, $\hat{P}_j = \mathbf{e}_j\mathbf{e}_j^\dagger$. A consequence of this is the completeness relation,

$$\sum_{j=1}^{N} \hat{P}_j = \sum_{j=1}^{N} \mathbf{e}_j\mathbf{e}_j^\dagger = \hat{\mathbb{1}}. \tag{3.9}$$

3.6 The Adjoint or Hermitian Conjugate

Definition 24: Adjoint

$$\forall \hat{A}, \forall \mathbf{u}, \mathbf{v} \in V_N : \langle \mathbf{u}, \hat{A}^\dagger \mathbf{v}\rangle = \overline{\langle \mathbf{v}, \hat{A}\mathbf{u}\rangle} \tag{3.10}$$

In matrix notation,

$$(\hat{A}^\dagger)_{ij} = \overline{A_{ji}}. \tag{3.11}$$

An operator is self adjoint if $\hat{A} = \hat{A}^\dagger$.

Definition 25: Unitary Operator

$$\exists \hat{U} : \hat{U}^\dagger\hat{U} = \hat{U}\hat{U}^\dagger = \hat{\mathbb{1}} \tag{3.12}$$

i.e.,

$$\hat{U}^{-1} = \hat{U}^\dagger. \tag{3.13}$$

3.7 Representations

3.7.1 Representation of a Vector

We know that a vector can be decomposed into orthonormal basis vectors. A vector space can have more than one set of orthonormal basis vectors which can represent a given vector. From the completeness relation, we find,

$$\begin{aligned} \mathbf{v} = \hat{\mathbb{1}}\mathbf{v} &= \sum_{j=1}^{N} \langle \mathbf{e}_j, \mathbf{v}\rangle \mathbf{e}_j \\ &= \sum_{j=1}^{N} \langle \mathbf{f}_j, \mathbf{v}\rangle \mathbf{f}_j \end{aligned} \tag{3.14}$$

i.e., changing basis does not change the abstract concept of a vector.

Invariance of the Inner Product

We can emphasise the idea above by considering the inner product of a vector. We can write some vector $\mathbf{v} \in V_N$ as,

$$\mathbf{v} = \sum_{j=1}^{N} a_i \mathbf{e}_i = \sum_{j=1}^{N} b_i \mathbf{f}_i. \tag{3.15}$$

Computing the inner product,

$$\begin{aligned}\langle \mathbf{v}, \mathbf{v}\rangle &= \sum_{j=1}^{N} (a_j \mathbf{e}_j)^\dagger \sum_{k=1}^{N} a_k \mathbf{e}_k \\ &= \sum_{j=1}^{N}\sum_{k=1}^{N} \overline{a_j} a_k \underbrace{\mathbf{e}_j^\dagger \mathbf{e}_k}_{\delta_{jk}} \\ &= \sum_{j=1}^{N} |a_j|^2\end{aligned} \tag{3.16}$$

and repeating this using the $\{\mathbf{f}_j\}_{j=1}^N$ representation,

$$\begin{aligned}\langle \mathbf{v}, \mathbf{v}\rangle &= \sum_{j=1}^{N} (b_j \mathbf{f}_j)^\dagger \sum_{k=1}^{N} b_k \mathbf{f}_k \\ &= \sum_{j=1}^{N} |b_j|^2\end{aligned} \tag{3.17}$$

and we find $\sum_{j=1}^N |a_j|^2 = \sum_{j=1}^N |b_j|^2$, so the inner product is invariant under a basis change.

3.7.2 Representation of Linear Operators

Consider a vector space V_N with basis $\{\mathbf{e}_j\}_{j=1}^N$. Consider $\mathbf{b}, \mathbf{c} \in V_N$ and an operator $\hat{A}$, such that $\mathbf{c} = \hat{A}\mathbf{b}$. We can find its representation by considering,

$$\begin{aligned}\mathbb{1}\mathbf{c} &= \mathbb{1}\hat{A}\mathbb{1}\mathbf{b} \\ \sum j = 1^N \langle \mathbf{e}_j, \mathbf{c}\rangle \mathbf{e}_j &= \sum_{j,k=1}^{N} \left(\mathbf{e}_j \mathbf{e}_j^\dagger\right) \hat{A} \left(\mathbf{e}_k \mathbf{e}_k^\dagger\right) \mathbf{b} \\ &= \sum_{j,k=1}^{N} \mathbf{e}_j \underbrace{\left(\mathbf{e}_j^\dagger \hat{A} \mathbf{e}_k\right)}_{A_{ij}} \underbrace{\left(\mathbf{e}_k^\dagger \mathbf{b}\right)}_{b_k}\end{aligned} \tag{3.18}$$

Thus, we find that we can denote the entries of a matrix $\underline{\underline{A}}$ representing a linear operator $\hat{A}$ are $\left(\underline{\underline{A}}\right)_{ij}$. We can denote this representation by,

$$\hat{A} \xrightarrow{\{\mathbf{e}\}_{j=1}^N} \underline{\underline{A}}. \tag{3.19}$$

3.7.3 Changing Representation of a Vector

Consider a vector $\mathbf{v} \in V_N$, and orthonormal bases $\{\mathbf{e}_j\}_{j=1}^N$ and $\{\mathbf{f}_j\}_{j=1}^N$. $\mathbf{v}$ can be decomposed as,

$$\mathbf{v} = \sum_{j=1}^{N} a_i \mathbf{e}_i = \sum_{j=1}^{N} b_i \mathbf{f}_i. \tag{3.20}$$

The l^{th} element of the vector is given by,

$$a_l = \langle \mathbf{e}_l, \mathbf{v}\rangle = \sum_{k=1}^{N} b_k \langle \mathbf{e}_l, \mathbf{f}_k\rangle = \sum_{k=1}^{N} \left(\underline{\underline{U}}_{lk}\right) b_k \tag{3.21}$$

where we have found the matrix $(\underline{\underline{U}})_{lk}$ which represents a unitary operator that takes a vector representation from one basis and transforms it into another. We can denote this,

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix} = \begin{pmatrix} \langle \mathbf{e}_1, \mathbf{f}_1\rangle & \cdots & \langle \mathbf{e}_1, \mathbf{f}_N\rangle \\ \langle \mathbf{e}_2, \mathbf{f}_1\rangle & \cdots & \langle \mathbf{e}_2, \mathbf{f}_N\rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{e}_N, \mathbf{f}_1\rangle & \cdots & \langle \mathbf{e}_N, \mathbf{f}_N\rangle \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix} \tag{3.22}$$

3.7.4 Changing Representation of a Matrix

We can write a linear operator as,

$$\hat{A} = \mathbb{1}\hat{A}\mathbb{1} = \sum_{j,k=1}^{N} A_{jk}\mathbf{e}_j\mathbf{e}_k^\dagger = \sum_{l,m=1} \tilde{A}_{lm}\mathbf{f}_l\mathbf{f}_m^\dagger \tag{3.23}$$

we then find that,

$$A_{jk} = \sum_{l,m=1}^{N} \langle \mathbf{e}_j, \mathbf{f}_l\rangle \tilde{A}_{lm}\langle \mathbf{f}_m, \mathbf{e}_k\rangle. \tag{3.24}$$

Thus, we find that we can change the matrix representation of a linear operator by,

$$\underline{\underline{A}} = \underline{\underline{U\tilde{A}U}}^\dagger. \tag{3.25}$$

3.8 Eigenvalue Problems

Definition 26: Eigen Equation

For a linear operator $\hat{A}$, the eigenequation is,

$$\hat{A}\mathbf{u} = \lambda\mathbf{u} \tag{3.26}$$

where λ and $\mathbf{u}$ are the eigenvalue and right eigenvector of $\hat{A}$ respectively, such that $\mathbf{u} \neq \mathbf{0}$.

A left eigenvector will satisfy,

$$\mathbf{v}^\dagger\hat{A} = \lambda\mathbf{v}^\dagger. \tag{3.27}$$

In order to obtain the eigenvalues, we rearrange to get,

$$\left(\hat{A} - \lambda\hat{\mathbb{1}}\right)\mathbf{u} = 0 \tag{3.28}$$

and solve,

$$\det\left(\hat{A} - \lambda\hat{\mathbb{1}}\right) \tag{3.29}$$

which generates a polynomial of degree N, with N solutions for $\lambda in\mathbb{C}$. Each distinct eigenvalue will correspond to a distinct eigenvector. However, if we have a repeated root, we get repeated eigenvectors. For $m > 1$ repeated eigenvectors, there may be up to, but at least 1, m linearly independent eigenvectors corresponding to the degenerate eigenvalue.

3.8.1 Eigenvalues and eigenvectors of Hermitian operators

Theorem 5: Eigenvalues of Hermitian Operators

Consider the eigenvalues $\{\lambda_i\}_{i=1}^N$ and eigenvectors $\{\mathbf{u}\}_{i=1}^N$ of $\hat{A}$. $\forall\hat{A} : \hat{A}^\dagger = \hat{A} \implies \forall i : \lambda_i \in \mathbb{R}$, $\forall i \neq j : \mathbf{u}_i^\dagger\mathbf{u}_j = 0$.

Proof. The eigenvalue equation is given by,

$$\hat{A}\mathbf{u}_i = \lambda_i\mathbf{u}_i. \tag{3.30}$$

Without loss of generality, consider the matrix elements of $\hat{A}$ with respect to two eigenvectors,

$$\mathbf{u}_j^\dagger\hat{A}\mathbf{u}_k = \lambda_k\langle\mathbf{u}_j, \mathbf{u}_k\rangle. \tag{3.31}$$

Using the Hermitian property,

$$\mathbf{u}_j^\dagger\hat{A}\mathbf{u}_k = \mathbf{u}_j^\dagger\hat{A}^\dagger\mathbf{u}_k = \overline{\mathbf{u}_k^\dagger\hat{A}\mathbf{u}_j} = \overline{\lambda_j}\left\langle\mathbf{u}_j, \mathbf{u}_k\right\rangle. \tag{3.32}$$

Equating equations (3.31) and (3.32),

$$(\lambda_k - \overline{\lambda}_k)\langle \mathbf{u}_j, \mathbf{u}_k\rangle = 0. \tag{3.33}$$

Let us consider two cases,

1. $k = j$, since $\langle \mathbf{u}_j, \mathbf{u}_j\rangle > 0$, we require $\lambda_j = \overline{\lambda}_j$.
2. $k \neq j$, since λ_j, λ_k are distinct, we require $\langle \mathbf{u}_j, \mathbf{u}_k\rangle = 0$.

□

3.8.2 Diagonalisation

Let $\underline{\underline{U}}$ be the matrix with orthonormal eigenvectors $\mathbf{u}_j$ as columns, such that $U_{ij} = (u_j)_i$, and $\hat{U}$ be the corresponding linear operator. By orthogonality, $\hat{U}$ is unitary. For a linear operator $\hat{A}$, we can diagonalise it (construct a matrix with $\hat{A}$'s eigenvalues on the diagonal) by,

$$\left[\hat{U}^\dagger \hat{A}\hat{U}\right]_{ij} = \lambda_i \delta_{ij}. \tag{3.34}$$

To obtain the original operator form its diagonalised form Λ_{ij}, we perform,

$$A_{ij} = U_{ik}\Lambda_{kl}\overline{U}_{jl}. \tag{3.35}$$

Eigenvalues and eigenvectors of unitary operators

We use unitary operators in physics to describe time evolution in quantum mechanics. We can do this because they preserve the norm, and thus the inner product, of a vector. i.e., $\langle \hat{U}\mathbf{u}, \hat{U}\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$.

Theorem 6: Eigenvalues of unitary operators

Consider a unitary operator $\hat{U} : V \to V$ with eigenvalue equation,

$$\hat{U}\mathbf{u}_j = \lambda_j \mathbf{u}_j. \tag{3.36}$$

The eigenvalues satisfy the following condition:

1. $|\lambda_j| = 1 \implies \lambda_j = e^{i\theta_j}, \theta \in \mathbb{R}$.
2. $\lambda_j \neq \lambda_k \implies \langle \mathbf{u}_j, \mathbf{u}_k\rangle = 0$.
3. The eigenvectors of $\hat{U}$ form an orthonormal basis for V.

Proof. 1.

□

Spectral Representation

We can diagonalise a general operator $\hat{A}$ which is not necessarily Hermitian, under the assumption that it has a complete set of eigenvectors. The left and right eigenvector equations are,

$$\mathbf{v}_j^\dagger \hat{A} = \lambda_j \mathbf{v}_j^\dagger \qquad \hat{A}\mathbf{u}_j = \lambda_j \mathbf{u}_j. \tag{3.37}$$

We then have,

$$\mathbf{v}_j^\dagger \hat{A}\mathbf{u}_k = \lambda_k \langle \mathbf{v}_j, \mathbf{u}_k\rangle = \lambda_j \langle \mathbf{v}_j, \mathbf{u}_k\rangle \tag{3.38}$$

thus for $\lambda_j \neq \lambda_k$, we require $\langle \mathbf{v}_j, \mathbf{u}_k\rangle = 0$.

We choose a normalisation $\langle \mathbf{v}_j, \mathbf{u}_k\rangle$, the completeness relation is given by,

$$\sum_{j=1}^{N} \mathbf{u}_j \mathbf{v}_j^\dagger = \hat{\mathbb{1}} \tag{3.39}$$

and we call $\mathbf{v}_j$ the dual basis to $\mathbf{u}_j$. We can thus write the operator $\hat{A}$,

$$\boxed{\hat{A} = \hat{A}\mathbb{1} = \sum_{j=1}^{N} \hat{A}\mathbf{u}_j\mathbf{v}_j^\dagger = \sum_{j=1}^{N} \lambda_j \mathbf{u}_j \mathbf{u}_j^\dagger} \tag{3.40}$$

which is known as the *spectral representation* of $\hat{A}$. Given an orthonormal basis $\{\mathbf{e}_j\}_{j=1}^{N}$ in which we can write the matrix representation of $\hat{A}$ which we denote $\underline{\underline{A}}$, we can diagonalise it by applying,

$$\underline{\underline{A}}^{\mathrm{diag}} = \underline{\underline{SAT}}^\dagger \tag{3.41}$$

where we define,

$$\left(\underline{\underline{S}}\right)_{jk} = \langle \mathbf{v}_j, \mathbf{e}_k \rangle \qquad \left(\underline{\underline{T}}\right)_{jk} = \langle \mathbf{u}_j, \mathbf{e}_k \rangle \tag{3.42}$$

which are unitary only if $\hat{A}$ is Hermitian.

Diagonalisation of commuting operators

For two commuting operators $\hat{A}$ and $\hat{B}$, when,

$$\hat{B}\mathbf{u} = \lambda \mathbf{u} \tag{3.43}$$

we have,

$$\begin{aligned} \hat{A}\hat{B}\mathbf{u} &= \lambda \hat{A}\mathbf{u} \\ \implies \hat{B}\hat{A} &= \lambda \left(\hat{A}\mathbf{u}\right) \end{aligned} \tag{3.44}$$

and thus $\hat{A}\mathbf{u}$ is also an eigenvector of $\hat{B}$ with eigenvalue λ. Below the implications of this statement are stated,

- **Non-degenerate:** If λ is non-degenerate, $\forall \lambda, \exists \mathbf{u} : \hat{A}\mathbf{u} = \mu \mathbf{u} \implies \mathbf{u}$ is a simultaneous eigenvector of $\hat{B}$ and $\hat{A}$.
- **Degenerate:** If λ has degeneracy $m > 1$, with eigenvectors $\{\mathbf{u}_j\}_{j=1}^{m}$, then a linear combination of these eigenvectors is an eigenvector of $\hat{A}$. $\hat{B}$ will act in this subspace, and it is possible to find a basis which simultaneously diagonalises $\hat{A}$ and $\hat{B}$.