Vector Spaces and Linear Algebra for Physics

v1.0.0

Hao-Yang Yen

NTHU, department of physics

Contents

Preface

This note is supplementary material for the 2025 Theoretical Mechanics course, where I am serving as a TA. Vectors are the most elementary objects in physics, especially for second-year majors beginning mechanics and electromagnetism, which is why many textbooks open with vector analysis. Yet, despite the existence of a mature and elegant framework in linear algebra, standard physics texts (e.g., Marion & Thornton; Griffiths) often retain older, less formal treatments, likely reflecting when they were written.

Alongside teaching, I am conducting research in mathematical physics, and I have come to appreciate how linear algebra underpins virtually all of theoretical physics, including classical mechanics. Within our department, however, there is no single course that systematically develops this beautiful and powerful theory. That gap motivates me to teach core linear-algebra ideas to students in Theoretical Mechanics.

In this note, I present a modern approach to linear algebra and a little bit of abstract algebra, emphasizing the concepts most useful in physics and illustrating them with physics examples and physically motivated examples. Besides teaching, preparing these materials also allows me to revisit some foundational ideas. I aim to share a more rigorous and contemporary perspective on vectors and linear algebra—one that helps students solve physical problems more effectively and, I hope, inspires some to pursue mathematical research in the future.

Chapter 1

Introduction to Algebra

1.1 From Operations to Algebraic Structures

In mathematics, the word **algebra** does not merely refer to solving equations with unknowns, but more broadly to the study of **sets equipped with operations**. Instead of working only with numbers, algebra allows us to abstract and study the rules of operations themselves, and to classify the structures that arise. This perspective is especially important in linear algebra and physics, where we encounter many different kinds of operations: vector addition, matrix multiplication, composition of transformations, and more.

Definition 1.1: Binary Operations

Let S be a set. A **binary operation** on S is a function

$$S \times S \to S, \quad (a, b) \mapsto a \star b,$$

where $\star$ denotes the operation. That is, a binary operation takes two elements of S and produces another element of S.

Definition 1.2: Associative and Commutative

Let S be a set and $\star$ be a binary operation on S. We say that $\star$ is

- **associative** if $a \star (b \star c) = (a \star b) \star c$ for all $a, b, c \in S$.
- **commutative** if $a \star b = b \star a$ for all $a, b \in S$.

Definition 1.3: Identity

If there exists $e \in S$ such that

$$e \star a = a \star e = a \quad \text{for all } a \in S,$$

then e is called an **identity** for the binary operation $\star$.

These definitions provide the foundation for the algebraic structures that underlie linear algebra. For example:

- Vector addition is both associative and commutative, with the zero vector as its identity.
- Matrix multiplication is associative but not commutative, with the identity matrix as its identity element.

By abstracting operations in this way, algebra offers a unified language to describe the structures that appear throughout physics and mathematics.

Once we have the notion of a binary operation, we can study sets together with such operations. Depending on which properties the operation satisfies, we arrive at different **algebraic structures**. These structures are the building blocks of linear algebra.

1.2 Groups: The Simplest Algebraic Structures

We start from the simplest algebraic structure called **group** (which is also a useful mathematical tool in physics to study symmetry and permutation) to see the basic idea of algebra.

Definition 1.4: Group

A **group** is a set G equipped with a binary operation $\star$ such that:

- **Associativity:** $(a \star b) \star c = a \star (b \star c)$ for all $a, b, c \in G$.
- **Identity:** There exists $e \in G$ such that $e \star a = a \star e = a$ for all $a \in G$.
- **Inverses:** For each $a \in G$ there exists $a^{-1} \in G$ such that $a \star a^{-1} = a^{-1} \star a = e$.

If, in addition, $\star$ is commutative, the group is called an **abelian group**.

Example 1.1: Integers under addition

The set of integers $\mathbb{Z}$ with addition $+$ is a group:

- Associative: $(a + b) + c = a + (b + c)$.
- Identity: 0.
- Inverse: for each a, the inverse is $-a$.
- Commutative: $a + b = b + a$.

This is the simplest example of an **abelian group**.

Example 1.2: Rotations in the plane

Consider rotations of the plane $\mathbb{R}^2$ about the origin by an angle θ.

- Composition of rotations is associative.
- The identity rotation corresponds to $\theta = 0$.
- Each rotation by θ has an inverse rotation by $-\theta$.

These rotations actually form the **special orthogonal group** $SO(2)$, which is crucial in describing angular momentum in physics.

Example 1.3: Reflection symmetries of a square

The set of all rotations and reflections that map a square onto itself forms a group, denoted D_4 (the dihedral group of order 8). Physically, this describes the **symmetry group** of the square.

Example 1.4: Pauli matrices under multiplication

The Pauli matrices

$$\sigma^x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma^y = \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix}, \quad \sigma^z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

together with $\pm I$ form a finite group under matrix multiplication (the **Pauli group**). This appears in quantum mechanics when describing spin-$1/2$ systems.

1.2.1 Application: Permutations and Levi-Civita Symbol

If we have n indistinguishable particles, exchanging their labels corresponds to an element of the **permutation group** S_n. The distinction between **even** and **odd** permutations leads to the classification of particles into bosons and fermions.

An element σ in the symmetric group S_n is a **permutation** of the labels $\{1, 2, \ldots, n\}$. In "two-line notation", it is written as

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix},$$

which means that

$$\sigma(i) = a_i, \quad i \in \{1, 2, \ldots, n\}.$$

Example 1.5

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

means that label 1 is sent to 2, label 2 is sent to 1, and label 3 stays fixed. Physically, this corresponds to **exchanging particle** 1 **and particle** 2 while leaving particle 3 untouched.

However, physicists often prefer a shorter way to write the same permutation, called **cycle notation**. The above example can be written as

$$\sigma = (12),$$

which denotes a swap (a **transposition**) of particles 1 and 2. More generally:

- (12) means exchanging particle 1 and 2.
- (123) means sending $1 \to 2$, $2 \to 3$, $3 \to 1$ (a 3-cycle).
- The identity permutation (do nothing) is denoted by ( ) or simply e.

Physical relevance. When particles are indistinguishable, their quantum states must transform consistently under such permutations.

- For bosons, wavefunctions are symmetric: $\Psi(\ldots, x_i, \ldots, x_j, \ldots) = \Psi(\ldots, x_j, \ldots, x_i, \ldots)$.

- For fermions, wavefunctions are antisymmetric: exchanging two labels multiplies the state by -1.

Thus, the abstract notation of permutations in S_n directly encodes the physical rules of particle exchange.

In physics, the **Levi-Civita symbol** ε_{ijk} is defined in three dimensions by

$$\varepsilon_{ijk} = \begin{cases} +1 & \text{if } (i,j,k) \text{ is an even permutation of } (1,2,3), \\ -1 & \text{if } (i,j,k) \text{ is an odd permutation of } (1,2,3), \\ 0 & \text{if any two indices are equal.} \end{cases}$$

This symbol encodes the orientation of permutations of indices. The set of all permutations of three objects $\{1,2,3\}$ forms the **symmetric group** S_3. Each element of S_3 is a bijection of $\{1,2,3\}$, and the group operation is the composition of permutations.

Connection to groups.

- Every permutation $\sigma \in S_3$ has a **parity**: it is either even or odd depending on whether it can be written as a product of an even or odd number of transpositions (swaps).
- The Levi-Civita symbol ε_{ijk} assigns $+1$ to even permutations and -1 to odd permutations, precisely encoding the **sign of a permutation**, which is a group homomorphism
$$\text{sgn} : S_3 \to \{\pm 1\}.$$
- As a consequence, the Levi-Civita symbol actually can be represented in terms of the sign of permutations
$$\varepsilon_{ijk} = \begin{cases} \text{sgn}(i,j,k) & \text{if } (i,j,k) \in S_3, \\ 0 & \text{otherwise.} \end{cases}$$

Example 1.6: Example in vector algebra

The cross product of two vectors $u, v \in \mathbb{R}^3$ can be written using ε_{ijk} as

$$(u \times v)_i = \sum_{j,k=1}^{3} \varepsilon_{ijk}\, u_j v_k.$$

Here, the antisymmetry of ε_{ijk} reflects the group-theoretic fact that swapping indices corresponds to applying an odd permutation, which flips the sign. Thus, the Levi-Civita symbol provides a bridge between concrete vector operations in physics and the abstract concept of permutation groups in algebra.

1.2.2 Application: The Determinant

The determinant of an $n \times n$ matrix $A = (a_{ij})$ can be expressed using the Levi-Civita symbol:

$$\det(A) = \sum_{i_1, i_2, \ldots, i_n = 1}^{n} \varepsilon_{i_1 i_2 \cdots i_n}\, a_{1i_1} a_{2i_2} \cdots a_{ni_n}.$$

Here, the Levi-Civita symbol $\varepsilon_{i_1 i_2 \cdots i_n}$ encodes the sign of the permutation $(i_1, i_2, \ldots, i_n)$ of $(1, 2, \ldots, n)$.

Group-theoretic interpretation.

- The set of all permutations of n objects forms the **symmetric group** S_n.
- Each term in the determinant expansion corresponds to a permutation $\sigma \in S_n$.
- The Levi-Civita symbol contributes $\mathrm{sgn}(\sigma)$, the group homomorphism

  $$\mathrm{sgn} : S_n \to \{\pm 1\},$$

  which distinguishes even and odd permutations.

Example 1.7: Determinant of three dimensional matrices

$$\det(A) = \sum_{\sigma \in S_3} \operatorname{sgn}(\sigma)\, a_{1\sigma(1)} a_{2\sigma(2)} a_{3\sigma(3)}.$$

This is exactly the same as the Levi-Civita expression

$$\det(A) = \sum_{i,j,k=1}^{3} \varepsilon_{ijk}\, a_{1i} a_{2j} a_{3k}.$$

In three dimensions, the determinant of the matrix whose columns are vectors u, v, w equals the scalar triple product

$$\det([u\ v\ w]) = u \cdot (v \times w),$$

which geometrically measures the oriented volume of the parallelepiped spanned by u, v, w. The appearance of ε_{ijk} in both the determinant and the cross product reflects the same underlying group-theoretic structure: the sign representation of the permutation group.

1.3 Rings and Fields

So far, groups involved only **one** binary operation (like addition in $\mathbb{Z}$ or composition in permutations). But in many parts of physics and linear algebra, we work with objects that naturally have **two** operations: addition and multiplication.

Example 1.8: Integers

Integers can be added and multiplied.

- Under addition, $\mathbb{Z}$ is an abelian group.
- Multiplication is associative and distributes over addition.

This motivates a more general algebraic structure having **two** operations called **rings**.

Definition 1.5: Ring

A **ring** is a set R equipped with two binary operations, usually called addition $(+)$ and multiplication $(\cdot)$, such that:

- R is an abelian group with the binary operation $(+)$.
- Multiplication is associative: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b, c \in R$.
- Multiplication distributes over addition:

$$a \cdot (b + c) = a \cdot b + a \cdot c, \quad (a + b) \cdot c = a \cdot c + b \cdot c.$$

Example 1.9: Real numbers

Real numbers can also be added and multiplied, but they go further:

- There is a multiplicative identity 1.
- Every nonzero element has a multiplicative inverse $1/a$.

This extra property defines **fields**.

In short:

- A **ring** is a set with addition and multiplication satisfying the distributive law.
- A **field** is a ring where division by nonzero elements is always possible.

Definition 1.6: Field

A **field** is a ring F in which:

- Multiplication has an identity element $1 \neq 0$.

- Every nonzero element $a \in F$ has a multiplicative inverse a^{-1} with $a \cdot a^{-1} = 1$.
- Multiplication is commutative.

Example 1.10

Integers $\mathbb{Z}$ form a ring but not a field (since not every nonzero integer has a multiplicative inverse). Real numbers $\mathbb{R}$ and complex numbers $\mathbb{C}$ are fields.

In linear algebra, vector spaces, which are defined over fields, are the central mathematical entities that we are interested in. That is, vectors form an **abelian group** under addition, and they can be scaled by elements of a field (which are called **scalars**). This algebraic framework provides the foundation for studying linear transformations, matrices, and eigenvalues—concepts that appear throughout physics.

Chapter 2

Vector Spaces

2.1 Abstract Algebraic Structures of Vectors

So far, we have introduced algebraic structures step by step:

- Groups: one operation (like addition of integers).
- Rings: two operations, addition and multiplication (like integers $\mathbb{Z}$).
- Fields: rings where every nonzero element is invertible under multiplication (like $\mathbb{R}$ or $\mathbb{C}$).

The next step is to study objects that can be **added together** and also **scaled** by elements of a field. This leads to the concept of a **vector space**, which is an abstract algebraic viewpoint of the traditional vectors in $\mathbb{R}^n$.

Definition 2.1: Vector Space

Let F be a field. A **vector space** over F is a set V equipped with two operations:

- **Vector addition:** $V \times V \to V$, $(u, v) \mapsto u + v$,

- **Scalar multiplication:** $F \times V \to V$, $(a, v) \mapsto a \cdot v$, sometimes denoted by av.

such that:

1. $(V, +)$ is an abelian group (with additive identity 0 and inverses $-v$).
2. Scalar multiplication is compatible with field multiplication:
$$a \cdot (b \cdot v) = (a \cdot b) \cdot v, \quad 1 \cdot v = v.$$
3. Scalar multiplication distributes over both vector addition and field addition:
$$a \cdot (u + v) = a \cdot u + a \cdot v, \quad (a + b) \cdot v = a \cdot v + b \cdot v.$$

Example 2.1

- Certainly, $\mathbb{R}^n$ is a vector space over $\mathbb{R}$, with componentwise addition and scalar multiplication. This is the motivation for defining the abstract vectors from an algebraic viewpoint.
- The set of all complex-valued wavefunctions $\psi : \mathbb{R}^3 \to \mathbb{C}$ forms a vector space over $\mathbb{C}$.
- The set of $m \times n$ real matrices is a vector space over $\mathbb{R}$.

Vector spaces are the central objects of linear algebra. They provide the natural framework for describing states and transformations in physics—from classical vectors in mechanics to Hilbert spaces in quantum mechanics.

Remark 2.1: Modules

In algebra, there is a more general concept called a **module**. A module is defined just like a vector space, except that the scalars come from a **ring** instead of a field.

- Vector space = module over a field.
- Module = generalization where scalars may not have multiplicative inverses.

Example 2.2

- $\mathbb{Z}^n$ is a module over $\mathbb{Z}$, but not a vector space (since integers are not a field).
- Any abelian group can be seen as a module over $\mathbb{Z}$.

For our purposes in linear algebra and physics, we will restrict to the case of fields (usually $\mathbb{R}$ or $\mathbb{C}$), since vector spaces provide the right setting for describing states, observables, and transformations.

2.2 Vector Subspaces

As we just discussed, $\mathbb{R}$ and $\mathbb{R}^2$ are both vector spaces, which means that we can consider the $x-$axis as a vector space contained in another vector space $\mathbb{R}^2$. This is called a **vector subspace**—a vector space is contained in another vector space.

Definition 2.2: Vector subspaces

Let V be a vector space over F. A subset $W \subseteq V$ is said to be a **subspace** of V if W is itself a vector space over F under the addition and scalar multiplication of V.

Example 2.3

- View $\mathbb{R}$ as a vector space over $\mathbb{R}$. $\mathbb{Z} \subseteq \mathbb{R}$ is **not** a subspace of $\mathbb{R}$ since the scalar multiplication is not closed (e.g., $\sqrt{2} \cdot 1 = \sqrt{2} \notin \mathbb{Z}$).
- View $\mathbb{R}^2$ as a vector space over $\mathbb{R}$. $\mathbb{R} \times \{0\} \subseteq \mathbb{R}^2$ is a subspace of $\mathbb{R}^2$.

Theorem 2.1: Subspace test

Let V be a vector space over F and $W \subseteq V$ be a subset. Then W is a subspace of V if and only if W is nonempty and closed under addition and the scalar multiplication of V.

Proof. Exercise. $\square$

Definition 2.3: Internal direct sum

Let U and W be subspaces of a vector space V. We say that V is the **direct sum** of U and W, written

$$V = U \oplus W,$$

if every $v \in V$ can be written uniquely as

$$v = u + w, \quad u \in U,\, w \in W.$$

Equivalently,

$$V = \{v + w \mid v \in V,\ w \in W\} := U + W$$

and $U \cap W = \{0\}$.

Definition 2.4: External direct sum

Let V and W be vector spaces over the same field F. Their **direct sum** is the Cartesian product

$$V \oplus W := V \times W,$$

with vector addition and scalar multiplication defined componentwise:

$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2,\ w_1 + w_2), \quad \lambda(v, w) = (\lambda v,\ \lambda w).$$

2.3 Linear Combinations and Bases

The central operation in a vector space is forming **linear combinations**. Let V be a vector space over a field F, and let $S \subseteq V$ be any (possibly infinite) set of vectors. The **span** of S, denoted $\mathrm{Span}(S)$, is defined **algebraically** as the set of all **finite** linear combinations of elements of S:

$$\mathrm{Span}(S) := \left\{ \sum_{j=1}^{m} a_j v_j \;\middle|\; m \in \mathbb{N},\ v_j \in S,\ a_j \in F \right\}.$$

Equivalently,

$$\mathrm{Span}(S) = \bigcup_{\substack{T \subseteq S \\ T \text{ finite}}} \mathrm{Span}(T).$$

In particular, if $v_1, v_2, \ldots, v_k \in V$ are finitely many vectors and $a_1, \ldots, a_k \in F$ are scalars, a linear combination of them is

$$a_1 v_1 + a_2 v_2 + \cdots + a_k v_k.$$

The span of them is the set of all possible linear combinations of $\{v_1, \ldots, v_k\}$, denoted

$$\mathrm{Span}\{v_1, \cdots, v_k\} = \{a_1 v_1 + a_2 v_2 + \cdots + a_k v_k \mid a_1, \ldots, a_k \in F\}.$$

Remark 2.2: Subspace generated by vectors

The span of any collection of vectors is always a subspace of V. It is the "smallest" subspace that contains all of them.

Remark 2.3: Finite and infinite sums

In the abstract definition of a vector space, **linear combinations are always finite sums**. Some algebra texts adopt the convention that any infinite linear combination is simply defined to be zero, so that only finite sums are genuine objects.

In physics and analysis, however, infinite sums like $\sum_{n=1}^{\infty} a_n v_n$ are ubiquitous (e.g., Fourier series, mode expansions). To make sense of them, one needs extra structure (norms, inner products, topology), so that **convergence** can be defined.

Example 2.4: Polynomials as an infinite generating set

The vector space $\mathbb{R}[x]$ consists of all real polynomials:

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n.$$

The set $\{1, x, x^2, x^3, \dots\}$ is infinite, and every polynomial uses only finitely many of them. Nevertheless, the whole infinite collection is required

$$\mathrm{Span}\{1, x, x^2, x^3, \dots\} = \mathbb{R}[x].$$

So $\mathbb{R}[x]$ is an example of an **infinite-dimensional vector space**: it cannot be generated by any finite set of vectors.

Example 2.5: Superposition Principle

In physics, linear combinations express the **superposition principle**:

- In mechanics, any displacement of a particle in the plane $\mathbb{R}^2$ can be written as $x\hat{\mathbf{i}} + y\hat{\mathbf{j}}$, that is,

$$\mathrm{Span}\left\{ \hat{\mathbf{i}},\ \hat{\mathbf{j}} \right\} = \mathbb{R}^2.$$

Moreover,

$$\operatorname{Span}\left\{\, \hat{\mathbf{i}} \,\right\} = x\text{−axis}, \quad \operatorname{Span}\left\{\, \hat{\mathbf{j}} \,\right\} = y\text{−axis}.$$

- In wave physics, two standing waves on a string can be added to produce a new vibration pattern. Therefore, the span of $\{\sin(n\pi x/L)\}_{n=1}^{\infty}$ contains all possible shapes of the string satisfying the boundary conditions, that is,

$$\operatorname{Span}\left\{ \sin\left(\frac{\pi x}{L}\right), \ \sin\left(\frac{2\pi x}{L}\right), \ \cdots \right\} = \text{all possible shapes of the string.}$$

- In quantum mechanics, every spin state $|\psi\rangle$ can be written as $a\,|{\uparrow}\rangle + b\,|{\downarrow}\rangle$ for a spin-1/2 particle, so

$$\operatorname{Span}\{|{\uparrow}\rangle\,, |{\downarrow}\rangle\} = \text{all spin superposition state.}$$

Definition 2.5: Basis

Let V be a vector space over a field F. A (possibly infinite) set of vectors $\beta = \{v_i\}_{i\in I} \subseteq V$ is called a **basis** of V if it satisfies:

1. **Linear independence.** For every finite subset $\{i_1, \dots, i_n\} \subseteq I$,

$$a_1 v_{i_1} + \cdots + a_n v_{i_n} = 0 \iff a_1 = \cdots = a_n = 0, \quad (a_j \in F).$$

2. **Spanning (finite linear combinations).** Every $v \in V$ can be written as a **finite** linear combination of elements of β, i.e., there exist indices $i_1, \dots, i_n \in I$ and scalars $a_1, \dots, a_n \in F$ such that

$$v = a_1 v_{i_1} + \cdots + a_n v_{i_n}.$$

Equivalently,

$$\operatorname{Span}\beta = V.$$

Proposition 2.1: Uniqueness of coordinates

If β is a basis, then the coefficients in the finite expansion of any $v \in V$ with respect to β are unique.

Proof. Exercise. $\square$

Example 2.6: Bases in Physics

- $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ is a basis of $\mathbb{R}^3$, representing directions in space.

  Proof. Obviously, $\text{Span}\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\} = \mathbb{R}^2$, and since

  $$a_1\,\hat{\mathbf{i}} + a_2\,\hat{\mathbf{j}} + a_3\,\hat{\mathbf{k}} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = 0 \iff a_1 = a_2 = a_3 = 0,$$

  $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ is linearly independent. Therefore, $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ is indeed a basis for $\mathbb{R}^3$. $\square$

- $\{\sin\left(\frac{n\pi x}{L}\right)\}_{n=1}^{\infty}$ form a basis for the vibrating string with fixed ends.

Remark 2.4: Coefficients as Physical Amplitudes

In physics, the coefficients of a linear combination often have direct meaning:

- Components of displacement along x, y, z directions.
- Fourier coefficients a_n giving the strength of each normal mode.
- Probability amplitudes a, b in a quantum state.

2.3.1 Application: Vibrating String

Consider a string of length L fixed at both ends. The transverse displacement $y(x,t)$ satisfies the wave equation

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2}, \quad y(0,t) = y(L,t) = 0.$$

At a fixed time t, the shape of the string is described by a function

$$f(x) = y(x,t), \quad f : [0,L] \to \mathbb{R}.$$

The set of all such functions f forms a vector space:

- If f_1 and f_2 are possible shapes, then any linear combination $af_1 + bf_2$ is also a possible shape.
- The natural scalar field is $\mathbb{R}$, since the displacements are real numbers.

In fact, the normal modes of vibration are sine functions

$$f_n(x) = \sin\left(\frac{n\pi x}{L}\right), \quad n = 1, 2, 3, \ldots$$

and they form a basis for this function space. Any string shape can be expressed as a linear combination (Fourier sine series):

$$f(x) = \sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi x}{L}\right).$$

2.3.2 Application: Electromagnetic Plane Waves

Maxwell's equations in free space admit plane-wave solutions of the form

$$\mathbf{E}(\mathbf{r},t) = \mathbf{E}_0 \mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{r} - \omega t)}, \quad \mathbf{B}(\mathbf{r},t) = \frac{1}{c}\hat{\mathbf{k}} \times \mathbf{E}(\mathbf{r},t),$$

with wave vector $\mathbf{k}$ and frequency $\omega = c|\mathbf{k}|$.

At a fixed time t, the electric field $\mathbf{E}(\mathbf{r})$ is a function $\mathbb{R}^3 \to \mathbb{C}^3$. The set of all finite linear combinations of such plane waves forms a vector space:

$$a_1\,\mathbf{E}_1(\mathbf{r}) + a_2\,\mathbf{E}_2(\mathbf{r}), \quad a_1, a_2 \in \mathbb{C}.$$

This vector space is infinite-dimensional. In fact, any well-behaved solution of Maxwell's equations in free space can be expressed as a superposition (Fourier integral) of plane waves:

$$\mathbf{E}(\mathbf{r}, t) = \int \tilde{\mathbf{E}}(\mathbf{k})\,\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{r} - \omega t)}\,\mathrm{d}^3 k.$$

The superposition principle of vector spaces explains interference and diffraction of light: overlapping plane waves simply add up to form more complicated field patterns.

2.4 Algebras over Fields

Up to now, we have built vector spaces over a field F, where vectors add and scalars from F act linearly. The next natural step is to **add an internal multiplication on the same space** that is compatible with the linear structure. This leads to the notion of an **algebra over** F.

Definition 2.6: Algebra over a Field

Let F be a field. An **algebra over** F is a vector space A over F equipped with a **vector multiplication** $(*)$

$$A \times A \;\to\; A, \quad (a, b) \mapsto a * b,$$

such that:

1. Multiplication is left and right distributive

$$(a+b)*c = a*c + b*c, \quad a*(b+c) = a*b + a*c.$$

2. Multiplication is compatible with scalars

$$(\lambda a)*(\mu b) = (\lambda\mu)a*b, \quad \lambda, \mu \in F.$$

If multiplication is associative (i.e. $(a*b)*c = a*(b*c)$), we call A an **associative algebra**. If there exists an element $1 \in A$ with $1*a = a*1 = a$, we say A is a **unital algebra**.

Example 2.7

- **Matrix algebra:** $M_n(\mathbb{R})$ (or $M_n(\mathbb{C})$) with usual matrix addition and multiplication is an associative unital F–algebra.
- **Polynomial algebra:** $F[x]$ with the usual operations is an associative unital algebra over F.

Example 2.8: Physics connections

- **Operators in QM:** On a finite-dimensional Hilbert space H, the linear operators $\mathrm{End}_{\mathbb{C}}(H)$ form an associative unital algebra over $\mathbb{C}$ under composition. Using the commutator $[A, B] = AB - BA$ as the product gives a (non-associative) **Lie algebra** structure on the same vector space.
- **Classical mechanics:** On a phase space (the space of position and momentum) with coordinates (q, p), the smooth functions C^∞ form a commutative and unital algebra over $\mathbb{R}$ under pointwise multiplication.

Also, the **Poisson bracket**

$$\{f, g\} = \frac{\partial f}{\partial q}\frac{\partial g}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial g}{\partial q}$$

is a bilinear, antisymmetric operation, making it an algebra over $\mathbb{R}$ called **Poisson algebra** (commutative algebra and Lie bracket compatible by Leibniz rule).

- **Products in $\mathbb{R}^3$:** The dot product $\mathbf{u} \cdot \mathbf{v} \in \mathbb{R}$ does **not** preserve the result in $\mathbb{R}^3$, so it does **not** make $\mathbb{R}^3$ an algebra over $\mathbb{R}$. But if we take the cross product $\mathbf{u} \times \mathbf{v} \in \mathbb{R}^3$, then it does make $\mathbb{R}^3$ an algebra over $\mathbb{R}$.

Thus, algebras provide a common language that marries linear structure with multiplication—precisely the mix that appears throughout physics.

2.4.1 Application: Angular Momentum as a Lie Algebra

Angular momentum provides a unifying example where algebraic structures appear both in classical and quantum mechanics. In each case, the angular momentum components generate a Lie algebra, but the product is given by different operations (Poisson bracket and commutator).

Example 2.9: Classical mechanics

In classical mechanics, the angular momentum components

$$L_x = yp_z - zp_y, \quad L_y = zp_x - xp_z, \quad L_z = xp_y - yp_x$$

satisfy the Poisson bracket relations

$$\{L_x, L_y\} = L_z, \quad \{L_y, L_z\} = L_x, \quad \{L_z, L_x\} = L_y.$$

Thus, the vector space spanned by $\{L_x, L_y, L_z\}$ forms a Lie algebra under the Poisson bracket. This algebraic structure is called $\mathfrak{so}(3)$, the algebra of infinitesimal rotations in three-dimensional space.

Example 2.10: Quantum mechanics

In quantum mechanics, the angular momentum operators J_x, J_y, J_z obey the commutation relations

$$[J_x, J_y] = \mathrm{i}\hbar J_z, \quad [J_y, J_z] = \mathrm{i}\hbar J_x, \quad [J_z, J_x] = \mathrm{i}\hbar J_y.$$

Here, the operator algebra with the commutator as product forms a Lie algebra. The structure is again $\mathfrak{so}(3)$ (in quantum mechanics, we like to call it $\mathfrak{su}(2)$ due to some physical reasons), whose group representation theory (a mathematical theory used to connect the group theory and linear algebra) explains the possible spin values.

Remark 2.5: Classical–quantum correspondence

The algebraic structures of angular momentum in classical and quantum mechanics are parallel:

$$\{L_i, L_j\} = \varepsilon_{ijk} L_k \quad \longleftrightarrow \quad \frac{1}{\mathrm{i}\hbar}[J_i, J_j] = \varepsilon_{ijk} J_k.$$

This illustrates how Lie algebras capture the essence of rotational symmetry in both settings, with the commutator playing the role of the quantized Poisson bracket.

Chapter 3

Linear Transformations

3.1 Linearity

Physical systems often involve operations that map vectors to other vectors in a way that respects addition and scalar multiplication.

Example 3.1: Motivating examples

- In mechanics, the force on a mass-spring system can be written as $F = Kx$, where K is a stiffness matrix. Doubling the displacement doubles the force, and forces from two displacements add linearly.
- In calculus, the derivative operator D is linear: $D(af + bg) = aD(f) + bD(g)$.
- In quantum mechanics, operators such as the Hamiltonian H act linearly: $H(a\psi + b\phi) = aH\psi + bH\phi$.

Definition 3.1: Linear transformation

Let V and W be vector spaces over a field F. A map $T : V \to W$ is called a **linear transformation** if

$$T(u+v) = T(u) + T(v), \quad T(\lambda v) = \lambda T(v),$$

for all $u, v \in V$, $\lambda \in F$.

Example 3.2: Linearity

Decide whether the following maps are linear transformations.

1. $T : \mathbb{R}^2 \to \mathbb{R}^2, \quad (x, y) \mapsto (x + y,\, 2y)$.
2. $S : \mathbb{R}^2 \to \mathbb{R}^2, \quad (x, y) \mapsto (x^2,\, y)$.
3. $D : C^1(\mathbb{R}) \to C(\mathbb{R}), \quad f \mapsto f'$, the derivative operator.
4. $U : \mathbb{R}^2 \to \mathbb{R}^2, \quad (x, y) \mapsto (x + 1,\, y)$.

Solution.

1. Linear: $T(u + v) = T(u) + T(v)$ and $T(\lambda u) = \lambda T(u)$.
2. Not linear: the square x^2 breaks additivity, since $(x_1 + x_2)^2 \neq x_1^2 + x_2^2$ in general.
3. Linear: the derivative operator satisfies $D(af + bg) = aD(f) + bD(g)$.
4. Not linear: the constant shift $(x + 1, y)$ does not satisfy $U(0, 0) = 0$.

□

Example 3.3: Linearity in physics examples

Check whether the following maps are linear transformations.

1. $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$, rotation by angle θ:

$$R_\theta \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

2. $K : \mathbb{R}^n \to \mathbb{R}^n$, Hooke's law for a coupled spring system:

$$Kq = \text{force vector},$$

   where K is a stiffness matrix and q is the displacement vector.

3. $N : \mathbb{R}^2 \to \mathbb{R}^2$, normalization map:

$$N(x, y) = \frac{1}{\sqrt{x^2 + y^2}}(x, y), \quad (x, y) \neq (0, 0).$$

Solution.

1. Linear: R_θ is given by matrix multiplication, so it preserves addition and scalar multiplication.
2. Linear: Kq is linear in q, since K is a fixed matrix.
3. Not linear: normalization fails linearity (e.g. $N(2x, 2y) = N(x, y) \neq 2N(x, y)$).

□

Remark 3.1

Most elementary operators in mechanics and quantum physics are linear (rotations, Hamiltonians, stiffness matrices, derivatives). Nonlinear maps (e.g., the renormalization group, fluid mechanics) also appear and provide much interesting physics, but they lie outside the linear algebra framework.

So far we have studied linear maps, which are linear in one variable. In physics, however, many operations depend linearly on **two** inputs.

Example 3.4

- The work done by a force, $W = \mathbf{F} \cdot \mathbf{x}$, is linear in both force and displacement.
- The dot product $\mathbf{u} \cdot \mathbf{v}$ is linear in $\mathbf{u}$ if $\mathbf{v}$ is fixed, and linear in $\mathbf{v}$ if $\mathbf{u}$ is fixed.

This motivates the following definition of **bilinear transformation**.

Definition 3.2: Bilinear transformation

Let V, W, U be vector spaces over a field F. A map $B : V \times W \to U$ is called **bilinear** if it is linear in each argument:

$$B(v_1 + v_2,\, w) = B(v_1, w) + B(v_2, w), \quad B(\lambda v, w) = \lambda B(v, w),$$
$$B(v,\, w_1 + w_2) = B(v, w_1) + B(v, w_2), \quad B(v,\, \mu w) = \mu B(v, w).$$

Equivalently, bilinearity captures the idea of linearity in each variable separately.

Example 3.5: Bilinearity

- **Dot product:** $(\mathbf{u}, \mathbf{v}) \mapsto \mathbf{u} \cdot \mathbf{v}$ is bilinear on $\mathbb{R}^n$.
- **Matrix multiplication:** $(A, B) \mapsto AB$ is bilinear in $M_n(\mathbb{R})$.
- **Lie brackets:** all Lie brackets are certainly bilinear from the definition.

We have seen that bilinear maps are maps that are linear in each of two arguments. This idea generalizes naturally.

Definition 3.3: Multilinear map

Let $V_1, \ldots, V_k$ and U be vector spaces over a field F. A map

$$f : V_1 \times V_2 \times \cdots \times V_k \ \to \ U$$

is called **multilinear** if it is linear in each argument separately. That is, fixing all but one variable, the map is linear in the remaining variable.

Example 3.6: Multilinearity

- **Triple product:** In $\mathbb{R}^3$, the scalar triple product

$$(\mathbf{u}, \mathbf{v}, \mathbf{w}) \ \mapsto \ \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$$

  is multilinear (in fact, trilinear). It represents the signed volume of the parallelepiped spanned by $\mathbf{u}, \mathbf{v}, \mathbf{w}$.

- **Determinant:** The determinant

$$\det : (\mathbb{R}^n)^n \to \mathbb{R}, \quad (v_1, \ldots, v_n) \mapsto \det[v_1 \ \cdots \ v_n]$$

  is multilinear in its n vector arguments. This property is fundamental in geometry and physics (e.g., volume forms).

3.1.1 Determinant

In linear algebra, the **determinant** of linear transformations provides a measure of how a linear transformation changes *volumes* and *orientation* in space.

(1) Volume scaling. Consider a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ represented by a matrix A. If we apply T to a unit cube in $\mathbb{R}^n$, the image is a parallelepiped whose volume is scaled by $|\det A|$. Thus,

$$\text{Volume after transformation} = |\det A| \times (\text{original volume}).$$

When $|\det A| = 1$, the transformation preserves volume.

(2) Orientation. The *sign* of $\det A$ carries geometric meaning:

$$\begin{cases} \det A > 0 & \Rightarrow \text{orientation is preserved,} \\ \det A < 0 & \Rightarrow \text{orientation is reversed (mirror reflection).} \end{cases}$$

For example, in $\mathbb{R}^3$, an orthogonal matrix with $\det Q = +1$ represents a proper rotation, while one with $\det Q = -1$ represents a reflection or inversion.

3.2 Isomorphisms

As we have discussed, in algebra, we mainly focus on the "algebraic structures", i.e., the operations on the sets. Therefore, we are more interested in how the operations send the elements of the sets, rather than in the names of the elements.

Often, two vector spaces look different on the surface, but they carry exactly the same algebraic structure. To characterize these spaces, we define a kind of special linear transformation.

Definition 3.4: Isomorphism

A linear map $T : V \to W$ between vector spaces is called an **isomorphism** if it is bijective (one-to-one and onto). In this case we say that V and W are **isomorphic**, written $V \cong W$.

Example 3.7: Different representations of the same structure

- $\mathbb{R}^3$ and the set of all degree-2 polynomials $\mathbb{R}_2[x]$ the identification

$$a\hat{\mathbf{i}} + b\hat{\mathbf{j}} + c\hat{\mathbf{k}} \mapsto a + bx + cx^2$$

  is an isomorphism. (Certainly, this is true for $\mathbb{R}^n$ and $\mathbb{R}_{n-1}[x]$)

- $\mathbb{R}^4$ and the space $M_2(\mathbb{R})$ of 2×2 real matrices: the correspondence

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto (a, b, c, d)$$

  defines an isomorphism.

- In quantum mechanics, any 2-dimensional Hilbert space is isomorphic to $\mathbb{C}^2$, even though one may think of the space as spin states, or polarization states, or qubit states.

Example 3.8: Constructing an isomorphism

Let $M_2(\mathbb{R})$ be the space of all 2×2 real matrices, and let $\mathbb{R}^4$ be the usual 4-dimensional vector space.

1. Define a map

$$T : M_2(\mathbb{R}) \to \mathbb{R}^4, \quad T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (a, b, c, d).$$

   Show that T is linear.

2. Show that T is bijective.

3. Conclude that $M_2(\mathbb{R}) \cong \mathbb{R}^4$.

Solution.

1. For linearity:

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}\right) = (a+a', b+b', c+c', d+d') = T(A) + T(A'),$$

   and $T(\lambda A) = \lambda T(A)$ for any $\lambda \in \mathbb{R}$.

2. T is one-to-one because $T(A) = T(B)$ implies all entries are equal, hence $A = B$. T is onto because any $(a, b, c, d) \in \mathbb{R}^4$ is the image of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

3. Therefore, T is an isomorphism, so $M_2(\mathbb{R})$ and $\mathbb{R}^4$ are structurally the same vector space.

$\square$

Remark 3.2: Physics meaning

This shows that the choice of representation is not essential: whether we write a 4D vector as a matrix or as a 4-tuple, the underlying vector space structure is identical. In physics, this principle allows us to change between different but equivalent descriptions (e.g., spin states as column vectors, or as functions, or as abstract kets).

If V is a finite-dimensional vector space, then we can write it in the form of Choosing a basis for a vector space V defines an isomorphism $V \cong F^n$. This allows us to represent abstract vectors as coordinate vectors, and linear transformations as matrices.

Proposition 3.1: Dimension criterion

Two finite-dimensional vector spaces V and W are isomorphic if and only if $\dim V = \dim W$. Thus, the dimension completely characterizes finite-dimensional spaces up to isomorphism.

Proof. Since V and W are finite-dimensional, say $n-$dimensional, they have bases $\{v_1, \cdots, v_n\}$ and $\{w_1, \cdots, w_n\}$. Consider a linear map

$$\Phi : V \to W, \quad v_i \mapsto w_i.$$

Since for all $w \in W$, it an be written as

$$w = \sum_{i=1}^{n} a_i w_i = \sum_{i=1}^{n} a_i \Phi(v_i) = \Phi\left(\sum_{i=1}^{n} a_i v_i\right),$$

Φ is onto.

Assume that $u_1, u_2 \in V$ such that $\Phi(u_1) = \Phi(u_2)$,

□

Example 3.9: Rotation in $\mathbb{R}^2$

In the standard basis $\{e_1, e_2\}$, the rotation R_θ is represented by

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

The operator is abstract, but the matrix depends on the choice of basis.

Remark 3.3: Physics connection

In quantum mechanics, a Hilbert space H is abstract. Once we pick a basis (e.g. $\{|\uparrow\rangle, |\downarrow\rangle\}$ for a spin-$\frac{1}{2}$ system), we obtain an isomorphism $H \cong \mathbb{C}^2$. Operators such as the Hamiltonian or spin observables then have concrete matrix representations, like the Pauli matrices:

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

3.3 Matrix Representation

3.4 Image, Kernel, and Rank-Nullity Theorem

Consider a linear transformation $T : V \to W$ between vector spaces. A natural question is: for which vectors $v \in V$ does T “kill” them, i.e., send them to zero?

Definition 3.5: Kernel (Null Space)

The **kernel** (or **null space**) of T is

$$\ker(T) := \{\, v \in V \mid T(v) = 0 \,\}.$$

Equivalently, for a matrix $A \in M_{m\times n}(F)$,

$$\ker(A) = \{\, x \in F^n \mid Ax = 0 \,\}.$$

Example 3.10: Linear equations

The solutions of the homogeneous system $Ax = 0$ form exactly $\ker(A)$. This is always a subspace of F^n.

Example 3.11: Physics intuition

- For a stiffness matrix K in mechanics, $\ker(K)$ corresponds to rigid-body motions with zero restoring force.
- In electromagnetism, divergence-free vector fields (like magnetic fields with $\nabla \cdot \mathbf{B} = 0$) are kernels of the divergence operator.

If the kernel describes what is "lost" under a linear map, the image describes what can actually be "reached".

Definition 3.6: Image (Range)

Let $T : V \to W$ be a linear transformation. The **image** (or **range**) of T is

$$\operatorname{im}(T) := \{\, T(v) \mid v \in V \,\} \subseteq W.$$

For a matrix $A \in M_{m\times n}(F)$,

$$\operatorname{im}(A) = \{\, Ax \mid x \in F^n \,\},$$

which is the span of the column vectors of A.

Example 3.12: Linear equations

In the system $Ax = b$, the equation has a solution precisely when $b \in \operatorname{im}(A)$. Thus the image tells us which right-hand sides are possible outputs.

Example 3.13: Physics intuition

- In mechanics, for a stiffness matrix K, the set of all possible force vectors is $\operatorname{im}(K)$.
- In electromagnetism, any magnetic field $\mathbf{B}$ that can be written as $\mathbf{B} = \nabla \times \mathbf{A}$ lies in the image of the curl operator.

Remark 3.4: Terminology

In mathematics texts, one often sees "kernel" and "image". In applied mathematics and physics, it is common to say "null space" instead of kernel, and "range" or "column space" instead of image. All these terms describe the same objects in finite dimensions.

Definition 3.7: Rank and Nullity

Let V and W be finite-dimensional vector spaces and $T : V \to W$ be a linear transformation:

- The **rank** of T is $\dim(\mathrm{im}(T))$.
- The **nullity** of T is $\dim(\ker(T))$.

Theorem 3.1: Rank–Nullity Theorem

Let V and W be finite-dimensional vector spaces and $T : V \to W$ be a linear transformation, then

$$\dim V = \mathrm{rank}(T) + \mathrm{nullity}(T).$$

Proof. □

Remark 3.5: Geometric meaning

The dimension of the input space V splits into two parts:

$$\underbrace{\dim(\ker T)}_{\text{lost directions}} + \underbrace{\dim(\mathrm{im} T)}_{\text{accessible outputs}} .$$

Example 3.14: Rank–Nullity in action

Consider the matrix

$$A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 3 & 2 & 1 \end{bmatrix}.$$

1. Compute the rank of A.
2. Use the rank–nullity theorem to determine the nullity of A.
3. Describe the solution space of the homogeneous system $Ax = 0$.

Solution. Row reduction gives

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 3 & 2 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Thus $\operatorname{rank}(A) = 3$.
Since A maps $F^4 \to F^3$, the input space has dimension 4. By rank–nullity,

$$\dim(\ker A) = 4 - 3 = 1.$$

So the kernel is one-dimensional: all solutions of $Ax = 0$ form a line through the origin in $\mathbb{R}^4$.
Explicitly, solving $Ax = 0$ gives the solution space, that is, the kernel of A

$$\ker A = \left\{ x = t \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \quad t \in \mathbb{R} \right\},$$

is indeed a one-dimensional vector space. □

Remark 3.6: Physical interpretation

The rank tells us there are 3 independent constraints, while the nullity tells us the system has 1 free mode of motion (a one-dimensional space of solutions). In mechanics, this would correspond to a system with one rigid-body motion that costs no energy.

Example 3.15: Physics connection

- In solving $Ax = b$, the nullity tells us how many free parameters the solution space has.
- For a stiffness matrix K, the nullity counts rigid-body modes with zero restoring force, while the rank counts the number of independent force directions.
- In electromagnetism, $\nabla \cdot \mathbf{B} = 0$ says magnetic fields lie in the kernel of the divergence operator; the rank–nullity balance ensures the right count of independent degrees of freedom.

Chapter 4

Diagonalization and Eigenvectors

Chapter
5

Inner Product

5.1 Inner Product and Norms

In classical mechanics, the concept of a *dot product* arises naturally when computing the work done by a force. If a force $\mathbf{F}$ acts on a particle that undergoes a displacement $\mathbf{d}$, the work performed is given by

$$W = \mathbf{F} \cdot \mathbf{d} = |\mathbf{F}|\,|\mathbf{d}|\cos\theta.$$

This formula measures how much of one vector "lies in the direction" of another. In other words, the dot product quantifies the *similarity of direction* between two vectors.

The same idea extends far beyond geometric vectors in $\mathbb{R}^3$. In physics and mathematics, we often deal with more general vector spaces: spaces of functions, signals, or even abstract states in quantum mechanics. To measure the "angle" or "projection" between such generalized vectors, we need a more abstract version of the dot product, called the **inner product**.

An inner product provides a way to define lengths, angles, and orthogonality in any vector space. It is a bridge between algebraic and geometric viewpoints—allowing us to treat vectors, functions, and even physical states under the same unified framework.

Definition 5.1: Inner product space

Let V be a vector space over the field F, where F is either the field of real numbers $\mathbb{R}$ or complex numbers $\mathbb{C}$. An **inner product** on V is a function

$$\langle \cdot, \cdot \rangle : V \times V \to F$$

that assigns to each pair of vectors (u, v) a scalar $\langle u, v \rangle$ satisfying:

1. **Positivity:** For all $v \in V$,
$$\langle v, v \rangle \geq 0,$$
and $\langle v, v \rangle = 0$ if and only if $v = 0$.
2. **Conjugate Symmetry:** For all $u, v \in V$,
$$\langle u, v \rangle = \overline{\langle v, u \rangle}.$$
(In the real case, this simply means $\langle u, v \rangle = \langle v, u \rangle$.)
3. **Linearity in the First Argument:** For all $u, w, v \in V$ and all scalars $a, b \in F$,
$$\langle au + bw, v \rangle = a\langle u, v \rangle + b\langle w, v \rangle.$$

A vector space equipped with such an operation is called an **inner product space**. If $F = \mathbb{C}$, we often call it a **Hermitian inner product space**.

Remark 5.1: Why conjugate symmetry?

The reason why we do not require the inner product to be symmetric $\langle u, v \rangle = \langle v, u \rangle$ but conjugate symmetric $\langle u, v \rangle = \overline{\langle v, u \rangle}$ is that, as we require conjugate symmetry, we immediately have $\langle u, u \rangle \in \mathbb{R}$, which implies that we can define the notion of "length" of a vector in this space by $\sqrt{\langle u, u\rangle}$, we will see this more explicitly later.

Remark 5.2: Different definition

Two notational conventions are common for inner products:

1. **Mathematics convention** uses $\langle u, v\rangle$ and is **linear in the first argument**:
$$\langle au_1 + bu_2, v\rangle = a\langle u_1, v\rangle + b\langle u_2, v\rangle.$$
2. **Physics (Dirac) convention** often writes $\langle u \mid v\rangle$ and is**linear in the second argument**:
$$\langle u \,|\, av_1 + bv_2\rangle = a\langle u \,|\, v_1\rangle + b\langle u \,|\, v_2\rangle.$$

These conventions are related by

$$\langle u \mid v\rangle_{\text{phys}} \;=\; \langle v, u\rangle_{\text{math}}.$$

In this note we adopt the *mathematics convention* (linear in the first argument). If you prefer the physics convention, swap the roles of the two arguments in all formulas.

Example 5.1

- For vectors $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ in $\mathbb{R}^n$, define
$$\langle \mathbf{x}, \mathbf{y}\rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$
This is the familiar dot product from geometry. It satisfies all four properties above, and thus makes $\mathbb{R}^n$ an inner product space. In $\mathbb{C}^n$, due

to the conjugate symmetric, we need to modify this to

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n}.$$

- Let $\mathcal{C}[a,b]$ denote the space of continuous real-valued functions on the interval $[a,b]$. Define

$$\langle f, g \rangle = \int_a^b f(x)g(x)\,\mathrm{d}x.$$

This inner product is analogous to the standard inner product in $\mathbb{R}^n$ by considering the variable x in the function $\psi(x)$ as the "infinite index" of the vector, that is, a vector v_i but $i \in [a,b]$, which measures how "similar" two functions are on the interval $[a,b]$. Certainly, in the continuous complex-valued function space, the definition becomes

$$\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}\,\mathrm{d}x.$$

In physics, such inner products are ubiquitous — for example, in quantum mechanics:

$$\langle \psi \mid \phi \rangle = \int_{-\infty}^{\infty} \psi^*(x)\,\phi(x)\,\mathrm{d}x,$$

where ψ and ϕ are wavefunctions and ψ^* denotes the complex conjugate of ψ.

- Let $M_{n\times m}(\mathbb{R})$ be the set of all $n \times n$ real matrices. Define

$$\langle A, B \rangle = \operatorname{Tr} AB^{\mathsf{T}}.$$

This definition is motivated by viewing the matrices as compositions of several row vectors, so just as in previous cases, we need to modify the definition in $M_{n\times n}(\mathbb{C})$ to

$$\langle A, B \rangle = \operatorname{Tr} A\overline{B}^{\mathsf{T}}.$$

Just as the dot product, we can introduce many concepts of geometry into abstract vector spaces through the inner product structure. The most basic geometric quantity–length–is defined by what we call **norm**.

Definition 5.2: Norm

Let V be an inner product space over $F = \mathbb{C}, \mathbb{R}$. For $u \in V$, the **norm** of u, denoted by $\|u\|$, is defined as

$$\|u\| = \sqrt{\langle u, u \rangle}.$$

These concepts generalize the familiar geometric ideas from Euclidean space. They form the foundation for many powerful geometric inequalities.

Theorem 5.1: Triangle inequality

Let V be an inner product space over $F = \mathbb{R}$ or $F = \mathbb{C}$. Then for all $u, v \in V$

$$\|u + v\| \leq \|u\| + \|v\|.$$

Proof.

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle = \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle \\ &= \langle u, u \rangle + 2\operatorname{Re}\langle u, v \rangle + \langle v, v \rangle \\ &\leq \|u\|^2 + 2|\langle u, v \rangle| + \|v\|^2 \\ &\leq \|u\|^2 + 2\|u\| \cdot \|v\| + \|v\|^2 = (\|u\| + \|v\|)^2. \end{aligned}$$

Thus,

$$\|u + v\| = \|u\| + \|v\|.$$

$\square$

Theorem 5.2: Cauchy-Schwarz inequality

Let V be an inner product space over $F = \mathbb{R}$ or $F = \mathbb{C}$. Then for all $u, v \in V$

$$|\langle u, v\rangle| \leq \|u\| \cdot \|v\|.$$

Proof. If $v = 0$, then we are done. So suppose that $v \neq 0$. For all $c \in F$, we have

$$\begin{aligned} 0 \leq \|u - cv\|^2 &= \langle u - cv, u - cv\rangle \\ &= \langle u, u\rangle - \langle u, cv\rangle - \langle cv, u\rangle + \langle cv, cv\rangle \\ &= \langle u, u\rangle - \bar{c}\langle u, v\rangle - c\langle v, u\rangle + c\bar{c}\langle v, v\rangle \end{aligned}$$

If we take $c = \langle u, v\rangle / \langle v, v\rangle$, then

$$\begin{aligned} 0 &\leq \langle u, u\rangle - \frac{\overline{\langle u, v\rangle}}{\langle v, v\rangle}\langle u, v\rangle - \frac{\langle u, v\rangle}{\langle v, v\rangle}\langle v, u\rangle + \frac{\langle u, v\rangle}{\langle v, v\rangle}\frac{\overline{\langle u, v\rangle}}{\langle v, v\rangle}\langle v, v\rangle \\ &= \langle u, u\rangle - \frac{\overline{\langle u, v\rangle}}{\langle v, v\rangle}\langle u, v\rangle. \end{aligned}$$

Thus,

$$\frac{\overline{\langle u, v\rangle}}{\langle v, v\rangle}\langle u, v\rangle \leq \langle u, u\rangle \implies |\langle u, v\rangle| \leq \|u\| \cdot \|v\|.$$

□

As we have the notion of length in vector spaces, we can certainly use it to define the unit vectors.

Definition 5.3: Normal vectors

A vector v in an inner product space is said to be **normalized** if it has $\|v\| = 1$.

5.2 Orthogonality of Vectors

As long as we can define the notion of length in a vector space, we can define the concept of "angles" and, certainly, orthogonality, as well.

Definition 5.4: Angles and orthogonality in vector spaces

Let V be an inner product space, and $u, v \in V$ are nonzero vectors. An angle between two vectors is defined as

$$\theta = \arccos\left(\frac{\langle u, v\rangle}{\|u\| \cdot \|v\|}\right).$$

If the angle between two vectors is zero, i.e., $\langle u, v\rangle = 0$, then these vectors are said to be **orthogonal**.

Example 5.2: Orthogalnality of trigonometric functions

Let $V = \mathcal{C}[-\pi, \pi]$ be all continuous function on the interval $[-\pi, \pi]$ equipped with the inner product $\langle f, g\rangle = \int_{-\pi}^{\pi} f(x)g(x)\,\mathrm{d}x$. Let m and n be two different integers, then

$$\begin{aligned}
\langle \cos(m\theta), \cos(n\theta)\rangle &= \int_{-\pi}^{\pi} \cos(m\theta)\cos(n\theta)\,\mathrm{d}\theta \\
&= \int_{-\pi}^{\pi} \frac{1}{2}\big[\cos\big((m+n)\theta\big) + \cos\big((m-n)\theta\big)\big]\,\mathrm{d}\theta \\
&= \frac{1}{2(m+n)}\int_{-(m+n)\pi}^{(m+n)\pi} \cos\phi\,\mathrm{d}\phi \\
&\quad + \frac{1}{2(m-n)}\int_{-(m-n)\pi}^{(m-n)\pi} \cos\phi\,\mathrm{d}\phi = 0,
\end{aligned}$$

which means that $\cos(m\theta)$ and $\cos(n\theta)$ are orthogonal.

Definition 5.5: Orthogonal set

Let V be an inner product space. A subset $S \subseteq V$ is said to be an **orthogonal set** if any two distinct vectors in S are orthogonal.

Then, just as the standard basis in the Euclidean space, which is orthogonal and has unit lengths, we use orthogonality and normalization to define the **orthonormal set**.

Definition 5.6: Orthonormal basis

Let V be an inner product space. A subset $S \subseteq V$ is said to be an **orthonormal set** if any two distinct vectors in S are orthogonal and normalized.

5.2.1 Application: Special Relativity

5.2.2 Application: Fourier Series

5.3 Gram–Schmidt Theorem

Proposition 5.1

Let V be an inner product space. If S is an orthogonal subset of V consisting of nonzero vectors, then S is linearly independent.

Proof. Suppose $u_1, \ldots, u_n \in S$ and $a_1u_1 + \cdots + a_nu_n = 0$. Take the inner product with u_k on both sides:

$$\langle a_1u_1 + \cdots + a_nu_n, u_k\rangle = \langle 0, u_k\rangle = 0.$$

Then

$$a_1\langle u_1, u_k\rangle + \cdots + a_k\langle u_k, u_k\rangle + \cdots + a_n\langle u_n, u_k\rangle = 0.$$

Since the set is orthogonal, $\langle u_i, u_k\rangle = 0$ for $i \neq k$, so

$$a_k\langle u_k, u_k\rangle = 0 \quad \Rightarrow \quad a_k = 0.$$

Hence $a_1 = \cdots = a_n = 0$, and S is linearly independent. $\square$

By the above theorem, if $\dim V = n$, then n nonzero orthogonal vectors $u_1, \ldots, u_n$ will form a basis for V.

Such a basis has the following advantage: Given any $u \in V$, it is easy to compute coefficients $a_1, \ldots, a_n \in F$ such that

$$a_1u_1 + \cdots + a_nu_n = u.$$

For a real inner product space, one can compute a_i geometrically. Let θ be the angle between u and u_1:

$$a_1u_1 = \frac{\langle u, u_1\rangle}{\langle u_1, u_1\rangle}u_1.$$

If $F = \mathbb{C}$, the geometric interpretation may not hold, but the formula remains valid.

Proposition 5.2

Let V be an inner product space and $S = \{u_1, \ldots, u_n\}$ be an orthogonal basis of V consisting of nonzero vectors. If $u \in \operatorname{span}(S)$, then

$$u = \sum_{i=1}^{n} \frac{\langle u, u_i\rangle}{\langle u_i, u_i\rangle}u_i.$$

Proof. Write $u = a_1u_1 + \cdots + a_nu_n$. Taking the inner product with u_i gives:

$$\langle u, u_i\rangle = a_i\langle u_i, u_i\rangle \quad \Rightarrow \quad a_i = \frac{\langle u, u_i\rangle}{\langle u_i, u_i\rangle}.$$

$\square$

Now, we naturally want to ask a question. Does an orthogonal basis always exist? If yes, how can we find it? The answer is given by the **Gram–Schmidt orthogonalization process**.

Given any linearly independent set of vectors $\{w_1, \ldots, w_n\}$ in an inner product space, the Gram–Schmidt process is a systematic procedure that produces an orthogonal set $\{v_1, \ldots, v_n\}$ such that:

$$\operatorname{span}\{v_1, \ldots, v_k\} = \operatorname{span}\{w_1, \ldots, w_k\}, \quad k = 1, \ldots, n.$$

Theorem 5.3: Gram–Schmidt orthogonalization process

Let V be an inner product space, and let $\{w_1, \ldots, w_n\}$ be a linearly independent subset of V. Then the set of vectors $\{v_1, \ldots, v_n\}$ defined by

$$\begin{aligned}
v_1 &= w_1, \\
v_2 &= w_2 - \frac{\langle w_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1, \\
v_3 &= w_3 - \frac{\langle w_3, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle w_3, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2, \\
&\vdots \\
v_k &= w_k - \sum_{j=1}^{k-1} \frac{\langle w_k, v_j \rangle}{\langle v_j, v_j \rangle} v_j
\end{aligned}$$

is an orthogonal set.

Proof. For $n = 1$, the result is trivial. Assume the statement holds for $n-1$ vectors $\{v_1, \ldots, v_{n-1}\}$ forming an orthogonal set such that

$$\operatorname{span}\{v_1, \ldots, v_k\} = \operatorname{span}\{w_1, \ldots, w_k\}, \quad k = 1, \ldots, n-1.$$

We define

$$v_n = w_n - \sum_{i=1}^{n-1} \frac{\langle w_n, v_i \rangle}{\langle v_i, v_i \rangle} v_i.$$

To show $v_n \neq 0$: If $v_n = 0$, then $w_n \in \operatorname{span}\{v_1, \ldots, v_{n-1}\} = \operatorname{span}\{w_1, \ldots, w_{n-1}\}$, contradicting the linear independence of $\{w_i\}$. Hence $v_n \neq 0$.

For each $i < n$,

$$\langle v_n, v_i \rangle = \langle w_n, v_i \rangle - \frac{\langle w_n, v_i \rangle}{\langle v_i, v_i \rangle} \langle v_i, v_i \rangle = 0,$$

so v_n is orthogonal to $\{v_1, \ldots, v_{n-1}\}$.

To show $\operatorname{span}\{v_1, \ldots, v_n\} = \operatorname{span}\{w_1, \ldots, w_n\}$: Since $v_n = w_n - \sum_{i=1}^{n-1} \frac{\langle w_n, v_i \rangle}{\langle v_i, v_i \rangle} v_i$, we see $v_n \in \operatorname{span}\{w_1, \ldots, w_n\}$ and $w_n \in \operatorname{span}\{v_1, \ldots, v_n\}$. Hence the spans are equal. □

Example 5.3: Legendre polynomials

Let $V = \mathbb{R}[x]$ with the inner product

$$\langle f, g \rangle = \int_{-1}^{1} f(x) g(x) \, dx.$$

Apply the Gram–Schmidt process to $\{1, x, x^2, \ldots\}$:

$$\begin{aligned}
w_1 &= 1, \quad v_1 = w_1 = 1, \\
w_2 &= x, \quad v_2 = w_2 - \frac{\langle w_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = x - \frac{\int_{-1}^{1} x \, \mathrm{d}x}{\int_{-1}^{1} 1 \, \mathrm{d}x} = x, \\
w_3 &= x^2, \quad v_3 = w_3 - \frac{\langle w_3, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle w_3, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2.
\end{aligned}$$

The integrals are:

$$\int_{-1}^{1} x^2 \, \mathrm{d}x = \frac{2}{3}, \quad \int_{-1}^{1} 1 \, \mathrm{d}x = 2, \quad \int_{-1}^{1} x^3 \, \mathrm{d}x = 0.$$

Hence, we have an orthonormal basis

$$v_3 = x^2 - \frac{1/3}{1} = x^2 - \frac{1}{3}.$$

The polynomials $v_1, v_2, v_3, \ldots$ are orthogonal, and their span equals $\mathbb{R}_n[x]$. These are the (unnormalized) **Legendre polynomials**.

5.4 Adjoint Operator

The adjoint of a linear operator is not the same as the "classical adjoint" of a matrix.

The definition of the adjoint is more involved; we need some preliminary results.

Proposition 5.3

Let V be an inner product space. By the definition of the inner product, if we fix a vector $v \in V$ and consider the function

$$\langle \cdot, v \rangle : V \to F, \quad u \mapsto \langle u, v \rangle,$$

then this function is linear.

Theorem 5.4: Riesz Representation Theorem, finite-dimensional case

Let V be a finite-dimensional inner product space over F (where $F = \mathbb{R}$ or $\mathbb{C}$), and let $g : V \to F$ be a linear functional. Then there exists a unique vector $v_g \in V$ such that

$$g(u) = \langle u, v_g \rangle \quad \forall u \in V.$$

This means that any linear functional on V can be represented uniquely by an inner product with some vector. Pick an orthonormal basis $\beta = \{u_1, \ldots, u_n\}$ for V, and let $\sigma = \{1\}$ be the basis for F. Then

$$[g]_\beta^\sigma = (g(u_1), g(u_2), \ldots, g(u_n)).$$

If $h : V \to V$ is defined by $h(u_i) = \sum a_i u_i$, then

$$[h]_\beta^\beta = (a_1, a_2, \dots, a_n).$$

Thus $g = h$ if and only if $[g]_\beta^\sigma = [h]_\beta^\beta$, which implies $a_i = g(u_i)$ for each i. Hence

$$v_g = \sum_{i=1}^{n} g(u_i) u_i$$

is the unique vector such that $g(u) = \langle u, v_g \rangle$.

Example 5.4

Let $V = \mathbb{R}^3$ and define $g : \mathbb{R}^3 \to \mathbb{R}$ by

$$g(x, y, z) = 3x + 2y + 5z.$$

Then

$$g(x, y, z) = \langle (x, y, z), (3, 2, 5) \rangle,$$

so $v_g = (3, 2, 5)$.

Remark 5.3

The Riesz representative v_g of a linear functional $g : V \to F$ is the unique vector in V satisfying

$$g(u) = \langle u, v_g \rangle.$$

Proposition 5.4

Let V and W be finite-dimensional inner product spaces, and let $T : V \to W$ be a linear transformation. Then there exists a unique linear transformation

$T^* : W \to V$ such that

$$\langle T(v), w \rangle = \langle v, T^*(w) \rangle \quad \forall v \in V,\ w \in W.$$

Proof. For each fixed $w \in W$, define the function

$$g_w(v) = \langle T(v), w \rangle, \quad v \in V.$$

This g_w is linear. By the Riesz Representation Theorem, there exists a unique vector $T^*(w) \in V$ such that

$$g_w(v) = \langle v, T^*(w) \rangle \quad \forall v \in V.$$

Thus, T^* is defined by the correspondence $w \mapsto T^*(w)$.

□

Proposition 5.5

The function $T^* : W \to V$ defined above is linear.

Proof. Let $w_1, w_2 \in W$. Then for any $v \in V$,

$$\begin{aligned}
\langle v, T^*(w_1 + w_2) \rangle &= \langle T(v), w_1 + w_2 \rangle \\
&= \langle T(v), w_1 \rangle + \langle T(v), w_2 \rangle \\
&= \langle v, T^*(w_1) \rangle + \langle v, T^*(w_2) \rangle \\
&= \langle v, T^*(w_1) + T^*(w_2) \rangle.
\end{aligned}$$

Hence, $T^*(w_1 + w_2) = T^*(w_1) + T^*(w_2)$. Similarly,

$$\langle v, T^*(cw_1) \rangle = \langle T(v), cw_1 \rangle = c\langle T(v), w_1 \rangle = \langle v, cT^*(w_1) \rangle,$$

so $T^*(cw_1) = cT^*(w_1)$. Therefore, T^* is linear. □

Definition 5.7

The linear map $T^* : W \to V$ in the previous proposition is called the **adjoint** of $T : V \to W$.

Remark 5.4

We also have

$$\langle u, T(v)\rangle = \langle T^*(u), v\rangle = \overline{\langle v, T^*(u)\rangle}, \quad \forall u \in W,\ v \in V,$$

which expresses the symmetry of the adjoint relation.

5.5 Orthogonal Transformations

In mechanics, a **rigid body** is an idealized object that does not deform under motion. Every point of the body moves so that the distance between any two points remains constant:

$$\|\mathbf{x}' - \mathbf{y}'\| = \|\mathbf{x} - \mathbf{y}\|.$$

Such motions include translations and rotations. If we focus on motions that keep the origin fixed (pure rotations or reflections), the transformation of position vectors can be represented by a matrix Q such that

$$\mathbf{x}' = Q\,\mathbf{x}.$$

The condition that all distances are preserved implies

$$\langle Q\,\mathbf{x}, Q\,\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle,$$

which, in matrix form, becomes

$$Q^{\mathsf{T}} Q = I.$$

However, this intuitive notion can not be realized in more abstract vector spaces (e.g., 4D Euclidean space or function space). So we need to define this concept in abstract vector spaces, which leads to the **orthogonal transformation**.

Definition 5.8: Orthogonal transformations

Let V be a vector space over $\mathbb{R}$ and $T : V \to V$ a linear transformation. T is called an **orthogonal transformation** if it preserves the inner product:

$$\langle T(u), T(v)\rangle = \langle u, v\rangle \quad \text{for all } u, v \in V.$$

In matrix form, if T is represented by a real matrix Q with respect to an orthonormal basis, this condition becomes

$$Q^{\mathsf{T}} Q = I.$$

Hence, Q is called an **orthogonal matrix**.

Orthogonal transformations correspond exactly to rigid rotations and reflections in Euclidean space — they preserve distances and angles. Moreover, these matrices with the same dimension form a group called **orthogonal group**, describing the symmetry of a rigid body.

Definition 5.9: Orthogonal group

The set of all orthogonal matrices of dimension n forms the **orthogonal group**,

$$O(n) = \{Q \in \mathbb{R}^{n\times n} \mid Q^{\mathsf{T}} Q = I\}.$$

Elements of $O(n)$ represent rotations and reflections of a rigid body.

However, one may think that the "reflection" part of the orthogonal transformations is non-physical (since a rigid body can not really be inverted in the real world), we define another important subset of $O(n)$, which is also a group and describes the physical part of $O(n)$.

Definition 5.10: Special orthogonal group

An orthogonal transformation with determinant $+1$ is called a **special orthogonal transformation**. The set of all such transformations in an $n-$dimensional real vector space in matrix representation forms the **special orthogonal group**,

$$SO(n) = \{Q \in \mathbb{R}^{n\times n} \mid Q^{\mathsf{T}}Q = I,\ \det Q = 1\} \subseteq O(n).$$

Elements of $SO(n)$ represent proper rotations — i.e., rigid-body motions without reflection or inversion.

Physically, $SO(3)$ describes all possible orientations of a rigid body in three-dimensional space, and thus represents all possible motions of a rigid body that leave the origin fixed, and plays a central role in rotational dynamics.

Particularly, we mainly focus on the rigid body in 3D space in physics, i.e., we mainly consider $SO(3)$ in physics, which is called **Euler's angles**.

5.5.1 Application: Rigid-Body Rotations

Definition 5.11: Euler's angles

Euler proved that in three-dimensional space, any rotation matrix R can be represented (except for singular cases) by three successive rotations known as the **Euler angles** (α, β, γ):

$$R(\alpha,\beta,\gamma) = R_z(\alpha)\,R_y(\beta)\,R_z(\gamma), \quad \alpha,\gamma \in [0,2\pi),\ \beta \in [0,\pi].$$

Here, the basic rotation matrices are

$$R_z(\alpha) = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad R_y(\beta) = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix},$$

which satisfy:

$$R_z(\alpha)R_z(\alpha)^\mathsf{T} = R_y(\beta)R_y(\beta)^\mathsf{T} = I, \quad \det R_z(\alpha) = \det R_y(\beta) = 1.$$

Therefore, in modern mathematical language, we say that $R \in SO(3)$.
The three Euler angles describe a rotation as (Fig. 5.1):

- A rotation by γ about the fixed (space) z-axis.
- A rotation by β about the new y-axis.
- A rotation by α about the final (body) z-axis.

Together, they specify the orientation of a rigid body relative to a laboratory reference frame. In addition, if we view $SO(3)$ as a "Lie group" (loosely speaking, a smooth group that we can do calculus on), we say that $SO(3)$ is of dimension 3.

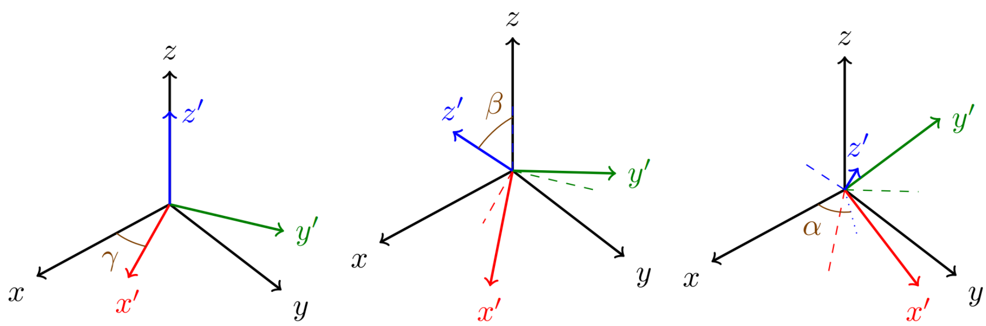

Figure 5.1: The rotational process of $R_z(\alpha)R_z(\beta)R_z(\gamma)$.

Example 5.5: Rrotating the z-axis to a direction n

Let $\mathbf{n}$ be a normal vector:

$$\mathbf{n} = (\sin\theta\cos\phi,\ \sin\theta\sin\phi,\ \cos\theta).$$

Then for any ψ, we have

$$R(\phi, \theta, \psi)\,\mathbf{e}_z = \mathbf{n}.$$

Hence, (ϕ, θ) determines the direction of $\mathbf{n}$ on the unit sphere, while ψ represents a rotation of the body about that axis (its "spin" around $\mathbf{n}$).

Remark 5.5

- At $\theta = 0$ or π, the angles ϕ and ψ become indistinguishable (the so-called **gimbal lock**).
- The instantaneous angular velocity can be expressed as

$$\boldsymbol{\omega} = \dot{\phi}\,\hat{\mathbf{z}} + \dot{\theta}\,\hat{\mathbf{y}}' + \dot{\psi}\,\hat{\mathbf{z}}'',$$

  where $\hat{\mathbf{y}}'$ and $\hat{\mathbf{z}}''$ are the intermediate rotation axes after the first and second steps, respectively.

The group of all proper rotations in three-dimensional space is certainly $SO(3)$, which is a "Lie group", meaning that it has a differentiable group structure.

The corresponding **Lie algebra $\mathfrak{so}(3)$** is defined as the "tangent space" to $SO(3)$ at the identity (i.e., the collection of tangent vectors at $I \in SO(3)$):

$$\mathfrak{so}(3) = \{\Omega \in \mathbb{R}^{3\times 3} \mid \Omega^{\mathsf{T}} = -\Omega\},$$

the set of all real skew-symmetric matrices.

For any time-dependent rotation $R(t) \in SO(3)$, differentiating the orthogonality condition

$$R^{\mathsf{T}}(t)R(t) = I$$

gives

$$\dot{R}^{\mathsf{T}} R + R^{\mathsf{T}} \dot{R} = 0 \implies R^{\mathsf{T}} \dot{R} \in \mathfrak{so}(3).$$

Hence, there exists a unique skew-symmetric matrix $\Omega(t) \in \mathfrak{so}(3)$ such that

$$\dot{R}(t) = R(t)\,\Omega(t).$$

This $\Omega(t)$ is called the **angular velocity matrix**, and it encodes the instantaneous rate of rotation of the body.

Explicitly, any element of $\mathfrak{so}(3)$ can be written in terms of a real vector $\boldsymbol{\omega} \in \mathbb{R}^3$ as

$$\Omega(\boldsymbol{\omega}) = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}, \quad \text{so that} \quad \Omega(\boldsymbol{\omega})\,\mathbf{v} = \boldsymbol{\omega} \times \mathbf{v}.$$

Therefore, the vector $\boldsymbol{\omega}$ can be viewed as the *coordinate representation* of the Lie algebra element Ω.

The vector cross product in $\mathbb{R}^3$ becomes the Lie bracket

$$[\boldsymbol{\omega}_1, \boldsymbol{\omega}_2] = \boldsymbol{\omega}_1 \times \boldsymbol{\omega}_2$$

corresponds to the Lie bracket of $\mathfrak{so}(3)$.

In summary, we have the relations:

- The Lie group $SO(3)$ describes finite rotations.
- The Lie algebra $\mathfrak{so}(3)$ describes *infinitesimal* rotations.
- The angular velocity $\boldsymbol{\omega}$ is an element of $\mathfrak{so}(3)$, specifying the tangent direction of motion on $SO(3)$.

Thus, angular velocity is not merely a vector in space — it is a **Lie algebra element** generating motion along the Lie group of rotations.

5.6 Unitary Transformations

Since we have defined the concept of length in complex vector spaces, as in real vector spaces, we can also extend the concept of a norm-preserving map, i.e., the concept of orthogonal transformations, to complex vector spaces. These transformations are said to be **unitary**.

Definition 5.12: Unitary transformations

Let V be a vector space over $\mathbb{C}$, the analogue of an orthogonal transformation in V is a **unitary transformation**.
A linear map $T : V \to V$ is called a **unitary transformation** if it preserves the Hermitian inner product:

$$\langle T(u), T(v)\rangle = \langle u, v\rangle \quad \text{for all } u, v \in V.$$

In matrix form, if T is represented by a matrix U, this means

$$U^\dagger U = I,$$

where $U^\dagger = \overline{U}^\mathsf{T}$ is the conjugate transpose of U.

Unitary transformations preserve both the length and phase relations of complex vectors, making them the natural generalization of rotations to complex Hilbert spaces, and are very useful in quantum mechanics.

Definition 5.13: Special unitary transformations

A unitary transformation with determinant of modulus one and $\det U = 1$ is called a **special unitary transformation**. The corresponding group is

$$SU(n) = \{U \in \mathbb{C}^{n\times n} \mid U^\dagger U = I,\ \det U = 1\}.$$

Elements of $SU(n)$ represent "proper" rotations in complex vector spaces and appear frequently in quantum mechanics and particle physics (for instance,

$SU(2)$ in spin systems and $SU(3)$ in quantum chromodynamics).

5.6.1 Application: Unitary Time Evolution

In quantum mechanics, the state of a finite n-level system is represented by a complex vector

$$|\psi\rangle = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{bmatrix} \in \mathbb{C}^n, \quad \langle\psi|\psi\rangle = 1.$$

According to the Schrödinger equation,

$$\mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H\,|\psi(t)\rangle,$$

where H is the Hamiltonian operator (self-adjoint matrix, $H^\dagger = H$). The formal solution is

$$|\psi(t)\rangle = U(t)\,|\psi(0)\rangle, \quad U(t) = \mathrm{e}^{-\mathrm{i}Ht/\hbar}.$$

Since H is self-adjoint, $U(t)$ satisfies

$$U^\dagger(t)U(t) = I,$$

so $U(t) \in U(n)$. This means that a unitary transformation always represents the time evolution of a closed quantum system.

In this picture, $U(n)$ describes all possible "motions" of an n-state quantum system in its Hilbert space, analogous to how $SO(3)$ describes rigid-body rotations in three-dimensional space. For example, for a two-level atom (a qubit), $U(2)$ describes all possible reversible evolutions of its quantum state.

5.6.2 Application: Spin-$1/2$ system and $SU(2)$

In quantum mechanics, the state of a spin-$\frac{1}{2}$ particle (such as an electron) is described by a normalized vector

$$|\psi\rangle = \begin{bmatrix} a \\ b \end{bmatrix}, \quad a, b \in \mathbb{C}, \quad |a|^2 + |b|^2 = 1,$$

which represents a point on the Bloch sphere.

Any physically allowed transformation of this spin state must preserve its norm, i.e.,

$$\langle\psi'|\psi'\rangle = \langle\psi|\psi\rangle.$$

Therefore, the transformation is unitary:

$$|\psi'\rangle = U|\psi\rangle, \quad U^\dagger U = I.$$

For spin-$\frac{1}{2}$ systems, such transformations form the group $SU(2)$, which can be parametrized by three real parameters (similar to Euler angles) as

$$U(\alpha, \beta, \gamma) = \mathrm{e}^{-\mathrm{i}\alpha\sigma_z/2}\, \mathrm{e}^{-\mathrm{i}\beta\sigma_y/2}\, \mathrm{e}^{-\mathrm{i}\gamma\sigma_z/2},$$

where $\sigma_x, \sigma_y, \sigma_z$ are the **Pauli matrices**:

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

This $SU(2)$ rotation acts as a “quantum rotation” of the spin state and corresponds to a real-space rotation described by $SO(3)$. Indeed, $SU(2)$ is the **double cover** of $SO(3)$ — meaning that a 2π rotation in space changes the sign of the spinor, while a 4π rotation brings it back to its original state.

Physically, this explains phenomena such as the spinor phase shift observed in the **neutron interferometry experiment** or the behavior of electrons under magnetic fields in the Stern–Gerlach setup.

5.6.3 Application: Rotations in Four-Dimensional Space

Moreover, since we have defined the special orthogonal group $SO(n)$ in general dimension n, we can consider the higher-dimensional rotations. The most useful one is four-dimensional rotations in relativity.

In 3D space, any rotation can be characterized by a single axis and an angle. However, in four dimensions, rotations are fundamentally different — they can occur **simultaneously in two independent planes**.

A general element of $SO(4)$ can be viewed as a rotation by angles (θ_1, θ_2) in two orthogonal planes. For example, consider coordinates (x_1, x_2, x_3, x_4) in $\mathbb{R}^4$. A rotation in the x_1x_2-plane by θ_1 and in the x_3x_4-plane by θ_2 is given by

$$R(\theta_1, \theta_2) = \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 & 0 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 & 0 \\ 0 & 0 & \cos\theta_2 & -\sin\theta_2 \\ 0 & 0 & \sin\theta_2 & \cos\theta_2 \end{bmatrix}, \quad R(\theta_1, \theta_2) \in SO(4).$$

When $\theta_1 = \theta_2$, the rotation is called an **isoclinic rotation**, which preserves all angles between lines through the origin.

Interestingly, $SO(4)$ has a special algebraic property:

$$SO(4) \cong \frac{SU(2)_L \times SU(2)_R}{\mathbb{Z}_2}.$$

This means that any 4D rotation can be represented by a pair of unit quaternions (or equivalently, a pair of $SU(2)$ matrices). The two independent $SU(2)$ factors correspond to rotations in two mutually orthogonal planes — the "left" and "right" isoclinic rotations.

Physically, such higher-dimensional rotations appear in:

- The mathematical study of spacetime symmetries in Euclideanized relativity (Wick rotation).

- The representation of spinor fields in four-dimensional spacetime, where the Lorentz group $SO(3,1)$ is closely related to $SL(2,\mathbb{C})$, a complex extension of $SU(2)$.

Thus, $SO(4)$ serves as the natural generalization of the familiar $SO(3)$ rotations into four dimensions, linking geometric intuition with algebraic symmetry.

Chapter 6

Tensor Product

Bibliography

[1] S. H. Friedberg, A. J. Insel, and L. E. Spence, *Linear Algebra*, 4th ed., Prentice Hall, 2003.

[2] Axler, Sheldon. *Linear algebra done right*. Springer Nature, 2024.

[3] Loring W. Tu, *An Introduction to Manifolds*, 2nd ed., Springer, New York, 2011.

[4] W. Keith Nicholson, *Introduction to Abstract Algebra*, 3rd ed., John Wiley & Sons, 2019.

[5] Roman, Steven. Advanced linear algebra. New York, NY: Springer New York, 2005.

[6] Arnol'd, Vladimir Igorevich. Mathematical methods of classical mechanics. Vol. 60. Springer Science & Business Media, 2013.

[7] Abraham, Ralph, and Jerrold E. Marsden. Foundations of mechanics. No. 364. American Mathematical Soc., 2008.