Geometry I: Vector Conics and Surfaces

Gudfit

Contents

Chapter 1

Ideas & Motivations

Welcome to Geometry I: Vector Conics and Surfaces by me (Gudfit). The goal here is to build a clean bridge between classical Euclidean geometry and linear algebra, without jumping straight into abstraction. Instead of treating vectors as formal objects from the start, we develop them geometrically: as displacements, directions, and tools for encoding familiar geometric ideas.

I aim for each set of notes to be max 100 pages [1], as rigorous as possible, and far-reaching too. That means I'll cover the axioms and proofs of the most interesting stuff, plus I'll pull in other subjects we've already touched on to show how math builds on itself like funky Lego. These notes build on my existing informal logic, algebra I and Geometry I notes, and they're aimed at keeping the proofs, ideas, and build-up of geometry as informal as possible.

It'll be a mix of quick ideas and concepts, but in the appendix section, I'll be rigorous with the key axioms pulled from a bunch of books not covered in the main text. For those theorems and ideas in the appendix, everything will be proved without algebra (since that's coming in the next book).

The original idea was a fully rigorous intro like Euclid's Elements, but that felt too grindy. Why slog through it when you can just read other people's notes, papers, or books? So this'll be more efficient, not totally deductive, assuming you've got some mathematical rigor. Either way, let's dive in and enjoy!

[1] Figs took it over 100 pages, but i'm okay with 104 pages.

Chapter 2

Vectors in the Plane

In previous studies of geometry, points on the plane were represented by pairs of real numbers called coordinates, and algebraic operations were carried out on these individual coordinates to discover properties of geometric configurations. For instance, given two points P and Q represented by (a, b) and (c, d) respectively, the straight line passing through P and Q consists of points X whose coordinates (x, y) satisfy a polynomial equation derived from the coordinates of P and Q. In such instances, algebraic operations are not performed on the points P, Q, and X themselves, but rather on their constituent coordinates.

There are, however, approaches where algebra is applied directly to the geometric entities. For example, the correspondence between points of the plane and complex numbers allows one to use a single symbol to refer to a pair of coordinates. This enables direct operation on the complex numbers representing points, placing the properties of the complex number system at our disposal. While powerful, this method encounters limitations when attempting to extend to three-dimensional space, as no analogous system of three-dimensional numbers exists.

In this chapter, we explore a robust algebraic approach to geometry using vectors. Vectors share the advantage of using a single symbol to represent a multi-component entity. Crucially, the notion of vectors in the plane lends itself to seamless generalisation to spaces of higher dimensions, forming the foundation of linear algebra.

2.1 The Definition of a Vector

In elementary physics, a vector is often described as a quantity possessing both magnitude and direction (such as displacement, velocity, or force), distinguished from scalar quantities like mass or temperature. While it is possible to formalise vectors strictly through magnitude and direction, generalising 'direction' to higher dimensions can prove cumbersome. Instead, we adopt an algebraic definition based on ordered sets of numbers, deriving the geometric properties of magnitude and direction as consequences.

Definition 2.1.1. ***Vector***. A vector in the plane is an ordered pair of real numbers. The numbers comprising the pair are called the components of the vector. We denote vectors by bold-faced lower-case letters, and enclose the components in square brackets. Thus, if a_1, a_2 are real numbers, the vector $\mathbf{a}$ is given by:

$$\mathbf{a} = [a_1, a_2]$$

Remark. In handwritten work, where bold typeface is impractical, it is customary to indicate a vector by placing an arrow over the letter ($\overrightarrow{a}$), boldfaced ($\boldsymbol{a}$) or a bar beneath it ($\underline{a}$).

Equality between vectors is defined component-wise. Two vectors $\mathbf{a} = [a_1, a_2]$ and $\mathbf{b} = [b_1, b_2]$ are equal, written $\mathbf{a} = \mathbf{b}$, if and only if $a_1 = b_1$ and $a_2 = b_2$.

Visualisation

Vectors in the plane may be visualised as arrows in the Cartesian plane. To represent the vector $\mathbf{a} = [a_1, a_2]$ graphically, we draw a directed line segment from the origin O to the point A with coordinates (a_1, a_2). Consequently, every vector corresponds to a unique arrow originating at O, and conversely, every arrow starting at O represents a vector.

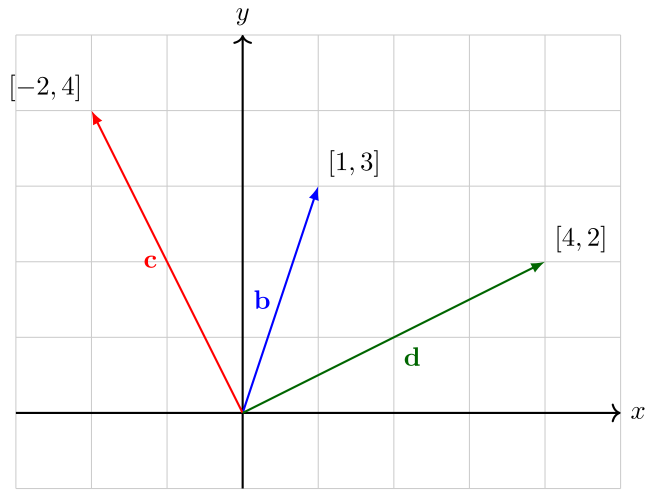


Figure 2.1: Graphical representation of vectors $\mathbf{b} = [1, 3]$, $\mathbf{c} = [-2, 4]$, and $\mathbf{d} = [4, 2]$ as directed segments from the origin.

An ordered pair of real numbers can fundamentally represent two distinct concepts: a geometric point or an algebraic vector. While the distinction may seem subtle initially, the entities are treated differently in practice. Points are geometric locations; vectors are algebraic quantities often representing displacements or shifts. We adopt the following notational convention to distinguish them.

Remark. (Point vs Vector).

- A *point* is denoted by a capital letter with coordinates in parentheses: $A = (a_1, a_2)$.
- A *vector* is denoted by a bold lower-case letter with components in brackets: $\mathbf{a} = [a_1, a_2]$.

Magnitude

Given a vector $\mathbf{a} = [a_1, a_2]$, the directed segment from the origin O to the point $A(a_1, a_2)$ has a length determined by Pythagoras' theorem. We define this length as the magnitude of the vector.

Definition 2.1.2. *Magnitude*. The magnitude (or norm) of a vector $\mathbf{a} = [a_1, a_2]$, denoted by $|\mathbf{a}|$, is the non-negative real number given by:

$$|\mathbf{a}| = \sqrt{a_1^2 + a_2^2}$$

For the vectors illustrated in Figure 2.1, we compute:

$$|\mathbf{b}| = \sqrt{1^2 + 3^2} = \sqrt{10}, \quad |\mathbf{c}| = \sqrt{(-2)^2 + 4^2} = \sqrt{20}, \quad |\mathbf{d}| = \sqrt{4^2 + 2^2} = \sqrt{20}.$$

Vectors with magnitude 1 are termed *unit vectors*. Examples include $[1/\sqrt{2}, -1/\sqrt{2}]$ and $[\cos\theta, \sin\theta]$. The standard unit coordinate vectors are denoted $\mathbf{e}_1 = [1, 0]$ and $\mathbf{e}_2 = [0, 1]$.

***Theorem* 2.1.1. Zero Vector Magnitude.** Let $\mathbf{a}$ be a vector. Then $|\mathbf{a}| = 0$ if and only if $\mathbf{a} = [0, 0]$.

Proof. Since $a_1^2 \geq 0$ and $a_2^2 \geq 0$ for any real numbers, the sum $a_1^2 + a_2^2 = 0$ implies $a_1 = 0$ and $a_2 = 0$. Conversely, if $a_1 = a_2 = 0$, the magnitude is clearly zero. ■

The vector $[0, 0]$ is called the *zero vector*, denoted by $\mathbf{0}$. Geometrically, it corresponds to a degenerate segment where the initial and terminal points coincide at the origin. All other vectors are non-zero vectors, possessing a well-defined positive magnitude and a specific direction.

2.2 The Vector Space $\mathbb{R}^2$

The primary advantage of vector geometry lies in the algebraic system formed by vectors, which mirrors properties of complex numbers. We define two fundamental operations: addition and scalar multiplication.

Definition 2.2.1. ***Vector Addition***. Let $\mathbf{a} = [a_1, a_2]$ and $\mathbf{b} = [b_1, b_2]$ be vectors. The sum $\mathbf{a} + \mathbf{b}$ is defined as:

$$\mathbf{a} + \mathbf{b} = [a_1 + b_1, a_2 + b_2]$$

Definition 2.2.2. ***Scalar Multiplication***. Let $\mathbf{a} = [a_1, a_2]$ be a vector and r be a real number (referred to as a scalar in this context). The scalar multiple $r\mathbf{a}$ is defined as:

$$r\mathbf{a} = [ra_1, ra_2]$$

We also define the negative of a vector as $-\mathbf{a} = (-1)\mathbf{a} = [-a_1, -a_2]$, and vector subtraction as $\mathbf{b} - \mathbf{a} = \mathbf{b} + (-\mathbf{a})$.

Geometric Interpretation

Vector addition adheres to the Parallelogram Law. If vectors $\mathbf{a}$ and $\mathbf{b}$ are represented by directed segments $\overrightarrow{OA}$ and $\overrightarrow{OB}$, their sum $\mathbf{c} = \mathbf{a} + \mathbf{b}$ corresponds to $\overrightarrow{OC}$, where $OACB$ forms a parallelogram.

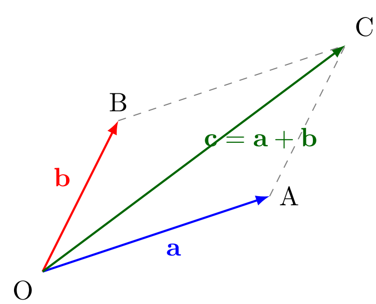


Figure 2.2: The Parallelogram Law of vector addition.

Scalar multiplication $r\mathbf{a}$ corresponds to scaling the length of the vector by a factor of $|r|$. If $r > 0$, the direction remains unchanged; if $r < 0$, the direction is reversed. If $r = 0$, the result is the zero vector.

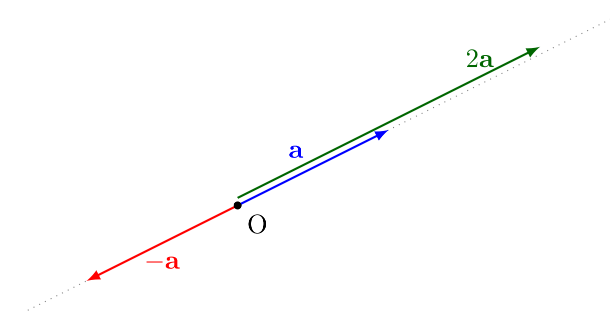


Figure 2.3: Scalar multiples of a vector **a**.

Definition 2.2.3. ***Vector Space*** $\mathbb{R}^2$. The set of all vectors in the plane, combined with the operations of vector addition and scalar multiplication, is denoted by $\mathbb{R}^2$ and is called the vector space of the plane.

The algebraic structure of $\mathbb{R}^2$ is governed by the following fundamental properties.

Theorem **2.2.1. Properties of** $\mathbb{R}^2$. Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be vectors in $\mathbb{R}^2$ and let r, s be scalars. Then:

(i) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ (Commutativity)
(ii) $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$ (Associativity)
(iii) There exists a unique vector $\mathbf{0}$ such that $\mathbf{a} + \mathbf{0} = \mathbf{a}$. (Additive Identity)
(iv) For every $\mathbf{a}$, there exists a unique vector $-\mathbf{a}$ such that $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$. (Additive Inverse)
(v) $(rs)\mathbf{a} = r(s\mathbf{a})$
(vi) $(r + s)\mathbf{a} = r\mathbf{a} + s\mathbf{a}$
(vii) $r(\mathbf{a} + \mathbf{b}) = r\mathbf{a} + r\mathbf{b}$
(viii) $1\mathbf{a} = \mathbf{a}$

The proof of these properties follows directly from the properties of real numbers applied to the components. For example, $\mathbf{a} + \mathbf{b} = [a_1 + b_1, a_2 + b_2] = [b_1 + a_1, b_2 + a_2] = \mathbf{b} + \mathbf{a}$.

The correspondence between complex numbers and vectors is notable here. Identifying $\mathbf{a} = [a_1, a_2]$ with $z = a_1 + ia_2$, vector addition corresponds to complex addition, and scalar multiplication corresponds to multiplying a complex number by a real number. However, unlike complex numbers, $\mathbb{R}^2$ does not inherently possess a vector-by-vector multiplication operation in this context.

We now demonstrate the power of these axioms by proving algebraic theorems in two ways: first using components (concrete), and second using only the axioms (abstract). This dual approach highlights that the results hold for any system satisfying the axioms of Theorem 2.2.1, not just $\mathbb{R}^2$.

Theorem **2.2.2. Zero Product Law**. Let $\mathbf{a}$ be a vector and r be a scalar. Then $r\mathbf{a} = \mathbf{0}$ if and only if $r = 0$ or $\mathbf{a} = \mathbf{0}$.

Proof 1 (Using Components). Let $\mathbf{a} = [a_1, a_2]$. Then $r\mathbf{a} = [ra_1, ra_2]$. If $r = 0$, then $r\mathbf{a} = [0 \cdot a_1, 0 \cdot a_2] = [0, 0] = \mathbf{0}$. If $\mathbf{a} = \mathbf{0}$, then $a_1 = 0, a_2 = 0$, so $r\mathbf{a} = [r \cdot 0, r \cdot 0] = \mathbf{0}$. Conversely, suppose $r\mathbf{a} = \mathbf{0}$. Then $[ra_1, ra_2] = [0, 0]$, implying $ra_1 = 0$ and $ra_2 = 0$. If $r \neq 0$, we must have $a_1 = 0$ and $a_2 = 0$, which implies $\mathbf{a} = \mathbf{0}$. ■

Proof 2 (Axiomatic). We use the properties from Theorem 2.2.1. First, to show $0\mathbf{a} = \mathbf{0}$:

$$0\mathbf{a} = (0 + 0)\mathbf{a} \overset{(6)}{=} 0\mathbf{a} + 0\mathbf{a}.$$

Adding $-(0\mathbf{a})$ to both sides implies $\mathbf{0} = 0\mathbf{a}$. Second, to show $r\mathbf{0} = \mathbf{0}$:

$$r\mathbf{0} = r(\mathbf{0} + \mathbf{0}) \overset{(7)}{=} r\mathbf{0} + r\mathbf{0}.$$

Adding $-(r\mathbf{0})$ implies $\mathbf{0} = r\mathbf{0}$. Conversely, assume $r\mathbf{a} = \mathbf{0}$. If $r = 0$, we are done. If $r \neq 0$, then $1/r$ exists.

$$\mathbf{a} \overset{(8)}{=} 1\mathbf{a} = \left(\frac{1}{r} \cdot r\right)\mathbf{a} \overset{(5)}{=} \frac{1}{r}(r\mathbf{a}) = \frac{1}{r}\mathbf{0} = \mathbf{0}.$$

Thus, if $r \neq 0$, $\mathbf{a}$ must be $\mathbf{0}$. ■

Theorem **2.2.3. Negative Scalars**. Let $\mathbf{a}$ be a vector and r be a scalar. Then:

$$(-r)\mathbf{a} = r(-\mathbf{a}) = -(r\mathbf{a})$$

Proof 1 (Using Components). Let $\mathbf{a} = [a_1, a_2]$.

$$(-r)\mathbf{a} = [-ra_1, -ra_2]$$

$$r(-\mathbf{a}) = r[-a_1, -a_2] = [-ra_1, -ra_2]$$
$$-(r\mathbf{a}) = -[ra_1, ra_2] = [-ra_1, -ra_2]$$

All three expressions yield the same components. ∎

Proof 2 (Axiomatic). The vector $-(r\mathbf{a})$ is defined as the unique additive inverse of $r\mathbf{a}$. Thus, it suffices to show that both $(-r)\mathbf{a}$ and $r(-\mathbf{a})$ act as inverses to $r\mathbf{a}$. For $(-r)\mathbf{a}$:

$$r\mathbf{a} + (-r)\mathbf{a} \overset{(6)}{=} (r + (-r))\mathbf{a} = 0\mathbf{a} = \mathbf{0}.$$

For $r(-\mathbf{a})$:

$$r\mathbf{a} + r(-\mathbf{a}) \overset{(7)}{=} r(\mathbf{a} + (-\mathbf{a})) = r\mathbf{0} = \mathbf{0}.$$

Since inverses are unique (Property 4), $(-r)\mathbf{a} = r(-\mathbf{a}) = -(r\mathbf{a})$. ∎

2.3 Linear Combinations and Independence

We have established that vector addition and scalar multiplication are the fundamental operations within the vector space $\mathbb{R}^2$. By combining these operations, we arrive at the concept of a linear combination, which forms the bedrock of linear algebra.

Definition 2.3.1. *Linear Combination*. Let $\mathbf{a}$ and $\mathbf{b}$ be vectors in $\mathbb{R}^2$, and let r and s be scalars. The vector $\mathbf{v}$ defined by

$$\mathbf{v} = r\mathbf{a} + s\mathbf{b}$$

is called a linear combination of $\mathbf{a}$ and $\mathbf{b}$.

In terms of components, if $\mathbf{a} = [a_1, a_2]$ and $\mathbf{b} = [b_1, b_2]$, the linear combination is expressed as:

$$r\mathbf{a} + s\mathbf{b} = [ra_1 + sb_1,\ ra_2 + sb_2].$$

This construction generalises the basic operations: the sum $\mathbf{a} + \mathbf{b}$ is a linear combination with $r = s = 1$, and the scalar multiple $r\mathbf{a}$ is a linear combination with $s = 0$.

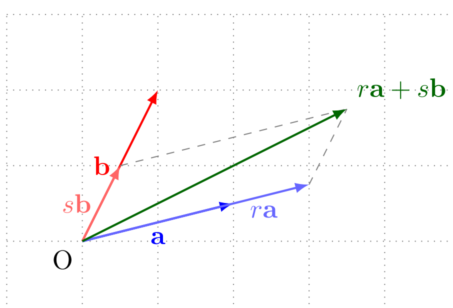


Figure 2.4: The vector $r\mathbf{a} + s\mathbf{b}$ represented via the Parallelogram Rule.

The unit coordinate vectors $\mathbf{e}_1 = [1, 0]$ and $\mathbf{e}_2 = [0, 1]$ occupy a distinguished role. Any vector $\mathbf{a} = [a_1, a_2]$ can be decomposed uniquely as:

$$\mathbf{a} = [a_1, 0] + [0, a_2] = a_1[1, 0] + a_2[0, 1] = a_1\mathbf{e}_1 + a_2\mathbf{e}_2.$$

Because every vector in $\mathbb{R}^2$ can be expressed as a linear combination of $\mathbf{e}_1$ and $\mathbf{e}_2$, we say that these vectors *generate* (or span) the vector space $\mathbb{R}^2$.

Linear Independence

While $\mathbf{e}_1$ and $\mathbf{e}_2$ generate the plane, they also satisfy a property of non-redundancy: neither is a scalar multiple of the other. If $\mathbf{e}_1 = r\mathbf{e}_2$, then $[1, 0] = [0, r]$, implying $1 = 0$, a contradiction. This concept is formalised as linear independence.

Definition 2.3.2. ***Linear Dependence and Independence***. Two vectors $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^2$ are said to be *linearly dependent* if one is a scalar multiple of the other. That is, either $\mathbf{a} = r\mathbf{b}$ or $\mathbf{b} = s\mathbf{a}$ for some scalars r, s. If neither vector is a scalar multiple of the other, they are said to be *linearly independent.*

Note. If $\mathbf{a} = \mathbf{0}$, then $\mathbf{a} = 0\mathbf{b}$, making the pair linearly dependent. Similarly, any vector is linearly dependent with itself. Thus, linear independence is a property relevant to distinct, non-zero vectors.

Geometrically, two non-zero vectors are linearly dependent if and only if they are collinear with the origin; that is, the directed segments $\overrightarrow{OA}$ and $\overrightarrow{OB}$ lie on the same straight line passing through O.

***Theorem* 2.3.1. Determinant Criterion for Dependence**. Let $\mathbf{a} = [a_1, a_2]$ and $\mathbf{b} = [b_1, b_2]$. The vectors $\mathbf{a}$ and $\mathbf{b}$ are linearly dependent if and only if

$$a_1b_2 - a_2b_1 = 0.$$

Proof. Suppose $a_1b_2 - a_2b_1 = 0$. If $\mathbf{a} = \mathbf{0}$, the vectors are dependent. Assume $\mathbf{a} \neq \mathbf{0}$; then at least one component, say a_1, is non-zero. From $a_1b_2 = a_2b_1$, we have $b_2 = \frac{a_2}{a_1}b_1$. We can express $\mathbf{b}$ as:

$$\mathbf{b} = [b_1, b_2] = \left[b_1, \frac{a_2}{a_1}b_1\right] = \frac{b_1}{a_1}[a_1, a_2] = \left(\frac{b_1}{a_1}\right)\mathbf{a}.$$

Thus $\mathbf{b}$ is a scalar multiple of $\mathbf{a}$. A similar argument holds if $a_2 \neq 0$. Conversely, suppose $\mathbf{a}$ and $\mathbf{b}$ are linearly dependent. If $\mathbf{a} = r\mathbf{b}$, then $a_1 = rb_1$ and $a_2 = rb_2$. Substituting these into the expression:

$$a_1b_2 - a_2b_1 = (rb_1)b_2 - (rb_2)b_1 = r(b_1b_2 - b_2b_1) = 0.$$

The same result follows if $\mathbf{b} = s\mathbf{a}$. ∎

The expression $a_1b_2 - a_2b_1$ is called the *determinant* of the pair of vectors. It plays a role analogous to the discriminant in quadratic equations, providing a purely algebraic test for a geometric property (We will see more of this in the next notes).

We may also characterise linear dependence through linear combinations yielding the zero vector.

***Theorem* 2.3.2. Algebraic Criterion for Dependence**. Two vectors $\mathbf{a}$ and $\mathbf{b}$ are linearly dependent if and only if there exist scalars r and s, not both zero, such that

$$r\mathbf{a} + s\mathbf{b} = \mathbf{0}.$$

Proof. Suppose such scalars exist. Without loss of generality, assume $r \neq 0$. Then we can rearrange the equation:

$$r\mathbf{a} = -s\mathbf{b} \implies \mathbf{a} = \left(-\frac{s}{r}\right)\mathbf{b}.$$

Thus $\mathbf{a}$ is a scalar multiple of $\mathbf{b}$, implying dependence. Conversely, if $\mathbf{a}$ and $\mathbf{b}$ are dependent, then either $\mathbf{a} = k\mathbf{b}$ or $\mathbf{b} = k\mathbf{a}$. In the first case, $1\mathbf{a} + (-k)\mathbf{b} = \mathbf{0}$ (where $r = 1 \neq 0$). In the second, $(-k)\mathbf{a} + 1\mathbf{b} = \mathbf{0}$ (where $s = 1 \neq 0$). ∎

We summarise these findings in two equivalent statements, distinguishing the dependent and independent cases.

***Theorem* 2.3.3. Conditions for Linear Dependence**. Let $\mathbf{a} = [a_1, a_2]$ and $\mathbf{b} = [b_1, b_2]$ be vectors in $\mathbb{R}^2$. The following statements are equivalent:

(i) **a** and **b** are linearly dependent.
(ii) The points $O(0,0)$, $A(a_1,a_2)$, and $B(b_1,b_2)$ are collinear.
(iii) $a_1b_2 - a_2b_1 = 0$.
(iv) There exist scalars r, s, not both zero, such that $r\mathbf{a} + s\mathbf{b} = \mathbf{0}$.

***Theorem* 2.3.4. Conditions for Linear Independence**. Let $\mathbf{a} = [a_1, a_2]$ and $\mathbf{b} = [b_1, b_2]$ be vectors in $\mathbb{R}^2$. The following statements are equivalent:

(i) **a** and **b** are linearly independent.
(ii) **a** and **b** are non-zero, and the points O, A, B are distinct and not collinear.
(iii) $a_1b_2 - a_2b_1 \neq 0$.
(iv) The equation $r\mathbf{a} + s\mathbf{b} = \mathbf{0}$ implies $r = 0$ and $s = 0$.

Remark. One should observe the distinction between the two types of proofs presented in this chapter.

(i) **Component-based proofs** (e.g., Theorem 2.2.1, Theorem 2.3.1) rely on the explicit definition of a vector as an ordered pair of numbers. These are specific to $\mathbb{R}^2$.
(ii) **Axiomatic proofs** (e.g., Theorem 2.2.2, Theorem 2.3.2) rely only on the algebraic properties (axioms) of vector addition and scalar multiplication. These proofs are more powerful as they apply to any vector space, regardless of dimension or the nature of the vectors involved.

2.4 Exercises

In the following exercises, vectors are strictly elements of $\mathbb{R}^2$. When asked to prove a statement, explicitly state whether you are relying on the component-wise definition of a vector or the axiomatic properties of the vector space.

Part I: Component Arithmetic and Geometry

1. Let $\mathbf{a} = [8, -2]$, $\mathbf{b} = [3, 4]$, and $\mathbf{c} = [0, 5]$. Compute the following vectors:

(a) $2\mathbf{a} - \mathbf{b}$
(b) $\frac{1}{2}(\mathbf{c} - 3\mathbf{a})$
(c) The vector $\mathbf{x}$ such that $2\mathbf{a} + \mathbf{x} = \mathbf{b} - \mathbf{c}$.

2. Find all real numbers k such that the following pairs of vectors are linearly dependent.

(a) $[1, k]$ and $[k, 4]$
(b) $[k, 1-k]$ and $[2, 3]$
(c) $[1, 1]$ and $[k^2, k]$

3. Determine the magnitude of the vector $\mathbf{v} = [3, -4]$. Find a scalar r such that $r\mathbf{v}$ is a unit vector pointing in the opposite direction to $\mathbf{v}$.

4. Collinearity. Use the conditions for linear dependence to determine if the points $A(-1, -2)$, $B(2, 4)$, and $C(5, 10)$ are collinear.

Remark. Consider the vectors $\mathbf{u}$ representing the segment $\overrightarrow{AB}$ and $\mathbf{v}$ representing $\overrightarrow{AC}$. Are they dependent?

5. Express the vector $[5, 2]$ as a linear combination of:

(a) The standard basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$.
(b) The vectors $\mathbf{u} = [1, 1]$ and $\mathbf{v} = [1, -1]$.

Part II: Axiomatic Structure

The following problems should be solved using *only* the axioms provided in Theorem 2.2.1 and the subsequent theorems proved in the text. Do not use components.

6. **The Cancellation Law.** Prove that for any vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$, if $\mathbf{a} + \mathbf{b} = \mathbf{a} + \mathbf{c}$, then $\mathbf{b} = \mathbf{c}$.

7. **Uniqueness of the Additive Identity.** Prove that the zero vector $\mathbf{0}$ is unique. That is, if there exists a vector $\mathbf{z}$ such that $\mathbf{a} + \mathbf{z} = \mathbf{a}$ for all $\mathbf{a}$, prove that $\mathbf{z} = \mathbf{0}$.

8. **Uniqueness of the Additive Inverse.** Prove that for a given vector $\mathbf{a}$, the vector $-\mathbf{a}$ is unique. That is, if $\mathbf{a} + \mathbf{b} = \mathbf{0}$ and $\mathbf{a} + \mathbf{c} = \mathbf{0}$, then $\mathbf{b} = \mathbf{c}$.

9. Prove that if $\mathbf{a} \neq \mathbf{0}$ and $r\mathbf{a} = s\mathbf{a}$, then $r = s$.

10. Prove that $-(\mathbf{a} + \mathbf{b}) = (-\mathbf{a}) + (-\mathbf{b})$.

Part III: Independence and Bases

11. **Uniqueness of Representation.** Let $\mathbf{a}$ and $\mathbf{b}$ be linearly independent vectors. Prove that if

$$r\mathbf{a} + s\mathbf{b} = p\mathbf{a} + q\mathbf{b},$$

then $r = p$ and $s = q$.

Remark. This result allows us to treat the coefficients (r, s) as "coordinates" relative to the basis $\{\mathbf{a}, \mathbf{b}\}$.

12. Let $\mathbf{a}$ and $\mathbf{b}$ be linearly independent vectors. Determine whether the following pairs of vectors are linearly dependent or independent.
    (a) $\mathbf{a} + \mathbf{b}$ and $\mathbf{a} - \mathbf{b}$.
    (b) $\mathbf{a} - 2\mathbf{b}$ and $3\mathbf{a} - 6\mathbf{b}$.
    (c) $\mathbf{a} + 2\mathbf{b}$ and $2\mathbf{a} + \mathbf{b}$.

13. **The Determinant Map.** Let $D(\mathbf{a}, \mathbf{b}) = a_1 b_2 - a_2 b_1$. Prove the following algebraic properties of the determinant:
    (a) *Anti-commutativity:* $D(\mathbf{b}, \mathbf{a}) = -D(\mathbf{a}, \mathbf{b})$.
    (b) *Linearity in the first argument:* $D(\mathbf{a} + \mathbf{c}, \mathbf{b}) = D(\mathbf{a}, \mathbf{b}) + D(\mathbf{c}, \mathbf{b})$ and $D(r\mathbf{a}, \mathbf{b}) = rD(\mathbf{a}, \mathbf{b})$.
    (c) $D(\mathbf{a}, \mathbf{a}) = 0$.

14. **Change of Basis.** Suppose $\mathbf{u} = r\mathbf{a} + s\mathbf{b}$ and $\mathbf{v} = p\mathbf{a} + q\mathbf{b}$. Prove that $\mathbf{u}$ and $\mathbf{v}$ are linearly independent if and only if $\mathbf{a}$ and $\mathbf{b}$ are linearly independent and $rq - sp \neq 0$.

15. $\star$ **The Centroid.** Let A, B, C be three non-collinear points representing the vertices of a triangle. Let their positions be given by vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$.
    (a) The midpoint M_{AB} of the side AB is given by $\frac{1}{2}(\mathbf{a} + \mathbf{b})$. Prove that the vector from the vertex C to the midpoint M_{AB} (the median) is given by $\frac{1}{2}(\mathbf{a} + \mathbf{b}) - \mathbf{c}$.
    (b) Prove that the point G defined by the vector $\mathbf{g} = \frac{1}{3}(\mathbf{a} + \mathbf{b} + \mathbf{c})$ lies on the median connecting C to M_{AB}.
    (c) Conclude that all three medians of the triangle intersect at the single point G.

16. $\star$ **Three is a Crowd.** Prove that any set of three vectors in $\mathbb{R}^2$ must be linearly dependent.

Remark. Let the vectors be $\mathbf{a}, \mathbf{b}, \mathbf{c}$. If $\mathbf{a}, \mathbf{b}$ are dependent, the proof is trivial. If they are independent, use Theorem 2.3.1 to show that $\mathbf{c}$ can be solved as a combination of $\mathbf{a}$ and $\mathbf{b}$ by solving the resulting system of linear equations.

Chapter 3

Plane Geometry

We have now developed sufficient algebraic machinery to revisit plane geometry through the lens of vectors. Rather than reformulating the entirety of Euclidean geometry, we establish a dictionary between the geometric properties of the plane and the algebraic structure of the vector space $\mathbb{R}^2$. This correspondence allows us to prove geometric theorems with algebraic rigour.

3.1 Position and Displacement Vectors

Notation 3.1.1. Position Vector Let $O = (0,0)$ be the origin of the Cartesian plane. For any point $A = (a_1, a_2)$, the vector $\mathbf{a} = [a_1, a_2]$ represented by the directed segment $\overrightarrow{OA}$ is called the position vector of A.

Through this convention, every point in the plane is uniquely associated with a vector. While position vectors are "bound" to the origin, we also consider vectors between arbitrary points.

Definition 3.1.1. ***Displacement Vector***. Let A and B be distinct points with position vectors $\mathbf{a}$ and $\mathbf{b}$ respectively. The directed segment $\overrightarrow{AB}$ is represented by the vector

$$\overrightarrow{AB} = \mathbf{b} - \mathbf{a}.$$

This vector is called the displacement vector from A to B.

This definition is consistent with the triangle rule for vector addition: $\overrightarrow{OA} + \overrightarrow{AB} = \overrightarrow{OB}$, which implies $\mathbf{a} + \overrightarrow{AB} = \mathbf{b}$.

Fundamental geometric properties can now be translated into vector notation.

- **Distance:** The length of the segment AB is the magnitude of the displacement vector: $d(A,B) = |\mathbf{b} - \mathbf{a}|$.
- **Collinearity:** Three distinct points A, B, C are collinear if and only if the displacement vectors $\overrightarrow{AB}$ and $\overrightarrow{AC}$ are linearly dependent. That is, $\mathbf{b} - \mathbf{a} = k(\mathbf{c} - \mathbf{a})$ for some scalar k.
- **Parallelograms:** A quadrilateral $ABCD$ is a parallelogram if and only if $\overrightarrow{AB} = \overrightarrow{DC}$, which implies $\mathbf{b} - \mathbf{a} = \mathbf{c} - \mathbf{d}$.

***Proposition* 3.1.1.** *Midpoint Formula*. Let A and B be points with position vectors $\mathbf{a}$ and $\mathbf{b}$. The midpoint M of the segment $\overline{AB}$ has the position vector

$$\mathbf{m} = \frac{1}{2}(\mathbf{a} + \mathbf{b}).$$

Proof. Let M be the midpoint. Then $\overrightarrow{AM} = \overrightarrow{MB}$. In terms of position vectors:

$$\mathbf{m} - \mathbf{a} = \mathbf{b} - \mathbf{m} \implies 2\mathbf{m} = \mathbf{a} + \mathbf{b} \implies \mathbf{m} = \frac{1}{2}(\mathbf{a} + \mathbf{b}).$$ ■

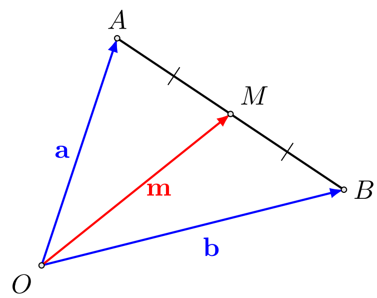


Figure 3.1: The position vector of the midpoint M is the average of $\mathbf{a}$ and $\mathbf{b}$.

Using these tools, we can prove classical geometric theorems efficiently.

***Theorem* 3.1.1. Parallelogram Diagonals**. The diagonals of a parallelogram bisect each other.

Proof. Let $ABCD$ be a parallelogram with position vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$. Since $ABCD$ is a parallelogram, $\overrightarrow{AB} = \overrightarrow{DC}$, so $\mathbf{b} - \mathbf{a} = \mathbf{c} - \mathbf{d}$, which implies $\mathbf{a} + \mathbf{c} = \mathbf{b} + \mathbf{d}$.

Let M_1 be the midpoint of diagonal AC, and M_2 be the midpoint of diagonal BD. Their position vectors are:

$$\mathbf{m}_1 = \frac{1}{2}(\mathbf{a} + \mathbf{c}) \quad \text{and} \quad \mathbf{m}_2 = \frac{1}{2}(\mathbf{b} + \mathbf{d}).$$

Since $\mathbf{a} + \mathbf{c} = \mathbf{b} + \mathbf{d}$, it follows that $\mathbf{m}_1 = \mathbf{m}_2$. Thus, the midpoints of the diagonals coincide, and the diagonals bisect each other. ■

The Parametric Equation of a Line

We established in the previous section that a point X lies on the line passing through A and B if the vectors $\overrightarrow{AX}$ and $\overrightarrow{AB}$ are linearly dependent.

***Theorem* 3.1.2. Parametric Line Equation**. Let A and B be distinct points with position vectors $\mathbf{a}$ and $\mathbf{b}$. A point X lies on the line L_{AB} if and only if its position vector $\mathbf{x}$ satisfies

$$\mathbf{x} = (1 - t)\mathbf{a} + t\mathbf{b}$$

for some scalar $t \in \mathbb{R}$.

Proof. If X is on the line, $\overrightarrow{AX}$ is parallel to $\overrightarrow{AB}$. Thus $\overrightarrow{AX} = t\overrightarrow{AB}$ for some scalar t.

$$\mathbf{x} - \mathbf{a} = t(\mathbf{b} - \mathbf{a})$$

$$\mathbf{x} = \mathbf{a} + t\mathbf{b} - t\mathbf{a} = (1 - t)\mathbf{a} + t\mathbf{b}.$$

Conversely, any vector of this form satisfies $\mathbf{x} - \mathbf{a} = t(\mathbf{b} - \mathbf{a})$, implying X lies on the line determined by A and B. ■

The scalar t is the parameter. The position of X relative to A and B is determined by t:

- $t = 0 \implies X = A$.
- $t = 1 \implies X = B$.
- $0 < t < 1 \implies X$ lies strictly between A and B (on the segment $\overline{AB}$).
- $t > 1 \implies B$ lies between A and X.
- $t < 0 \implies A$ lies between X and B.

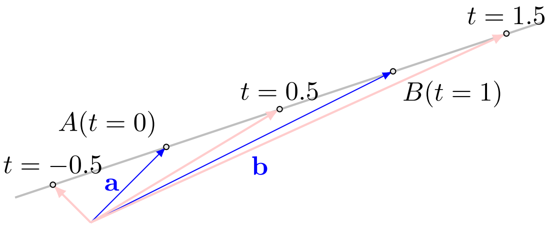


Figure 3.2: Points on the line L_{AB} determined by the parameter t.

Concurrency and the Centroid

The parametric representation is particularly powerful for proving concurrency theorems, such as the existence of the centroid.

***Theorem* 3.1.3. Concurrency of Medians**. The medians of a triangle are concurrent at a point called the centroid.

Proof. Let ABC be a triangle with position vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$. Let D, E, F be the midpoints of sides BC, CA, AB respectively. The position vectors of these midpoints are:

$$\mathbf{d} = \frac{1}{2}(\mathbf{b} + \mathbf{c}), \quad \mathbf{e} = \frac{1}{2}(\mathbf{c} + \mathbf{a}), \quad \mathbf{f} = \frac{1}{2}(\mathbf{a} + \mathbf{b}).$$

Consider the median AD. Any point on AD has the position vector:

$$\mathbf{x} = (1 - t)\mathbf{a} + t\mathbf{d} = (1 - t)\mathbf{a} + \frac{t}{2}(\mathbf{b} + \mathbf{c}).$$

If we choose $t = 2/3$, we obtain a specific point $\mathbf{g}$:

$$\mathbf{g} = \left(1 - \frac{2}{3}\right)\mathbf{a} + \frac{1}{2} \cdot \frac{2}{3}(\mathbf{b} + \mathbf{c}) = \frac{1}{3}\mathbf{a} + \frac{1}{3}\mathbf{b} + \frac{1}{3}\mathbf{c} = \frac{1}{3}(\mathbf{a} + \mathbf{b} + \mathbf{c}).$$

By the symmetry of the expression $\frac{1}{3}(\mathbf{a}+\mathbf{b}+\mathbf{c})$, this point $\mathbf{g}$ must also lie on the median BE (with parameter $2/3$ from B) and the median CF (with parameter $2/3$ from C). Since $\mathbf{g}$ lies on all three medians, the medians are concurrent at G, the centroid of the triangle. ■

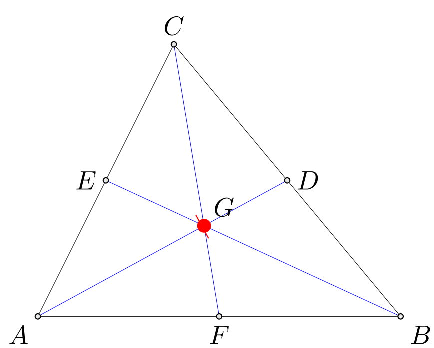


Figure 3.3: The medians intersect at the centroid G, which divides each median in a $2:1$ ratio.

Affine Dependence and Menelaus' Theorem

To handle more complex incidence theorems, we generalise the condition for collinearity.

***Theorem* 3.1.4. Affine Dependence**. Three points X, Y, Z with position vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are collinear if and only if there exist scalars u, v, w, not all zero, such that:

$$u\mathbf{x} + v\mathbf{y} + w\mathbf{z} = \mathbf{0} \quad \text{and} \quad u + v + w = 0.$$

Proof. Suppose X, Y, Z are collinear. Then Z lies on the line through X and Y. By Theorem 3.1.2, $\mathbf{z} = (1-t)\mathbf{x} + t\mathbf{y}$. Rearranging gives:

$$(1-t)\mathbf{x} + t\mathbf{y} - 1\mathbf{z} = \mathbf{0}.$$

Let $u = 1-t$, $v = t$, and $w = -1$. Then $u\mathbf{x} + v\mathbf{y} + w\mathbf{z} = \mathbf{0}$ and $u + v + w = (1-t) + t - 1 = 0$. Note that $w \neq 0$, so not all scalars are zero.

Conversely, suppose $u\mathbf{x} + v\mathbf{y} + w\mathbf{z} = \mathbf{0}$ and $u + v + w = 0$ with, say, $w \neq 0$. Then $w = -(u+v)$.

$$u\mathbf{x} + v\mathbf{y} - (u+v)\mathbf{z} = \mathbf{0} \implies (u+v)\mathbf{z} = u\mathbf{x} + v\mathbf{y}.$$

Since $w \neq 0$, $u + v \neq 0$, so we can divide:

$$\mathbf{z} = \frac{u}{u+v}\mathbf{x} + \frac{v}{u+v}\mathbf{y}.$$

Let $t = \frac{v}{u+v}$. Then $1 - t = 1 - \frac{v}{u+v} = \frac{u}{u+v}$. Thus $\mathbf{z} = (1-t)\mathbf{x} + t\mathbf{y}$, proving X, Y, Z are collinear. ■

We apply this to prove Menelaus' Theorem. First, we define the directed ratio.

Definition 3.1.2. *Directed Ratio*. Let A and B be distinct points. For any point Z on the line L_{AB} (where $Z \neq B$), the directed ratio of Z relative to A, B, denoted $\mathrm{dr}(A, B; Z)$, is the unique scalar r such that

$$\overrightarrow{AZ} = r\overrightarrow{ZB}.$$

Remark. If Z is between A and B, $r > 0$. If Z is outside the segment, $r < 0$. In vector terms, $\mathbf{z} - \mathbf{a} = r(\mathbf{b} - \mathbf{z})$, which implies $(1+r)\mathbf{z} = \mathbf{a} + r\mathbf{b}$.

***Theorem* 3.1.5. Menelaus' Theorem**. Let X, Y, Z be points on the lines containing sides BC, CA, AB of a triangle ABC, respectively. The points X, Y, Z are collinear if and only if

$$\mathrm{dr}(B, C; X) \cdot \mathrm{dr}(C, A; Y) \cdot \mathrm{dr}(A, B; Z) = -1.$$

Proof. Let $r = \mathrm{dr}(B, C; X)$, $s = \mathrm{dr}(C, A; Y)$, and $t = \mathrm{dr}(A, B; Z)$. From the definition of directed ratio:

$$\overrightarrow{BX} = r\overrightarrow{XC} \implies (1+r)\mathbf{x} = \mathbf{b} + r\mathbf{c}$$
$$\overrightarrow{CY} = s\overrightarrow{YA} \implies (1+s)\mathbf{y} = \mathbf{c} + s\mathbf{a}$$
$$\overrightarrow{AZ} = t\overrightarrow{ZB} \implies (1+t)\mathbf{z} = \mathbf{a} + t\mathbf{b}$$

Assume $rst = -1$. We form a linear combination of $\mathbf{x}, \mathbf{y}, \mathbf{z}$ to satisfy Theorem 3.1.4. Consider the combination:

$$st(1+r)\mathbf{x} + (1+s)\mathbf{y} - s(1+t)\mathbf{z}.$$

Substituting the expressions for $\mathbf{x}, \mathbf{y}, \mathbf{z}$:

$$\begin{aligned} &st(\mathbf{b} + r\mathbf{c}) + (\mathbf{c} + s\mathbf{a}) - s(\mathbf{a} + t\mathbf{b}) \\ &= st\mathbf{b} + str\mathbf{c} + \mathbf{c} + s\mathbf{a} - s\mathbf{a} - st\mathbf{b} \\ &= (str + 1)\mathbf{c}. \end{aligned}$$

Since $rst = -1$, we have $str + 1 = 0$. Thus the linear combination is the zero vector. We must now check the sum of the coefficients:

$$\Sigma = st(1+r) + (1+s) - s(1+t) = st + str + 1 + s - s - st = str + 1.$$

Again, since $rst = -1$, the sum of coefficients is 0. Therefore, by Theorem 3.1.4, points X, Y, Z are collinear. ■

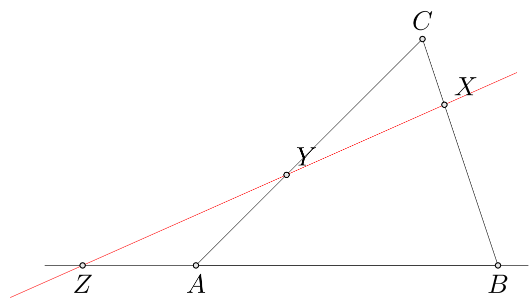


Figure 3.4: Menelaus' Theorem: Points X, Y, Z are collinear.

The Dot Product

In the preceding sections, we employed vector algebra to investigate affine properties of plane geometry — concepts such as parallelism, collinearity, and ratios of segments. These properties are independent of any specific unit of measurement. However, to discuss *metric* geometry (concepts involving lengths of segments and measures of angles), we require a stronger algebraic tool. This tool is the dot product.

Definition 3.1.3. ***Dot Product***. Let $\mathbf{a} = [a_1, a_2]$ and $\mathbf{b} = [b_1, b_2]$ be vectors in $\mathbb{R}^2$. The dot product (also known as the scalar product or inner product) of $\mathbf{a}$ and $\mathbf{b}$ is the real number defined by:

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2.$$

It is crucial to observe that while the operation takes two vectors as input, the output is a scalar. By examining the definition alongside our previous definition of Magnitude, we immediately observe a fundamental link between the dot product and the magnitude of a vector:

$$\mathbf{a} \cdot \mathbf{a} = a_1^2 + a_2^2 = |\mathbf{a}|^2.$$

Thus, the magnitude of a vector is the square root of the dot product of the vector with itself: $|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}}$.

The dot product satisfies several fundamental algebraic laws which justify its manipulation in equations.

***Theorem* 3.1.6. Algebraic Properties of the Dot Product**. For any vectors $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^2$ and any scalar $r \in \mathbb{R}$, the following hold:

(i) **Symmetry:** $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$.

(ii) **Bilinearity:** The dot product is linear in both arguments.

- $(r\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = r(\mathbf{a} \cdot \mathbf{c}) + \mathbf{b} \cdot \mathbf{c}$.
- $\mathbf{a} \cdot (r\mathbf{b} + \mathbf{c}) = r(\mathbf{a} \cdot \mathbf{b}) + \mathbf{a} \cdot \mathbf{c}$.

(iii) **Positive Definiteness:** $\mathbf{a} \cdot \mathbf{a} \geq 0$, with equality if and only if $\mathbf{a} = \mathbf{0}$.

Proof. These properties follow directly from the properties of real numbers. For instance, symmetry holds because $a_1 b_1 + a_2 b_2 = b_1 a_1 + b_2 a_2$. Bilinearity is verified by expanding the components. Positive definiteness follows from the fact that squares of real numbers are non-negative. ■

Remark. The vector space $\mathbb{R}^2$ equipped with this specific dot product is often referred to as the *Euclidean plane*. This structure bridges the gap between abstract vector spaces and Euclidean geometry.

Metric Theorems

The interaction between the dot product and the magnitude leads to several powerful inequalities and identities.

***Theorem* 3.1.7. Fundamental Metric Identities**. Let $\mathbf{a}$ and $\mathbf{b}$ be vectors in $\mathbb{R}^2$.

(i) **Homogeneity:** $|r\mathbf{a}| = |r||\mathbf{a}|$.
(ii) **Parallelogram Law:** $|\mathbf{a}+\mathbf{b}|^2 + |\mathbf{a}-\mathbf{b}|^2 = 2(|\mathbf{a}|^2 + |\mathbf{b}|^2)$.
(iii) **Cauchy-Schwarz Inequality:** $|\mathbf{a}\cdot\mathbf{b}| \le |\mathbf{a}||\mathbf{b}|$.
(iv) **Triangle Inequality:** $|\mathbf{a}+\mathbf{b}| \le |\mathbf{a}| + |\mathbf{b}|$.
(v) **Law of Cosines:** $|\mathbf{a}-\mathbf{b}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2(\mathbf{a}\cdot\mathbf{b})$.

Proof.

(i) This follows immediately from the definition of magnitude.
(ii) We expand the norms using the dot product properties:

$$|\mathbf{a}+\mathbf{b}|^2 = (\mathbf{a}+\mathbf{b})\cdot(\mathbf{a}+\mathbf{b}) = |\mathbf{a}|^2 + 2(\mathbf{a}\cdot\mathbf{b}) + |\mathbf{b}|^2$$
$$|\mathbf{a}-\mathbf{b}|^2 = (\mathbf{a}-\mathbf{b})\cdot(\mathbf{a}-\mathbf{b}) = |\mathbf{a}|^2 - 2(\mathbf{a}\cdot\mathbf{b}) + |\mathbf{b}|^2$$

Adding these two equations yields the result. Geometrically, this states that the sum of the squares of the diagonals of a parallelogram equals the sum of the squares of its sides.
(iii) If $\mathbf{a} = \mathbf{0}$, the inequality holds trivially. Assume $\mathbf{a} \neq \mathbf{0}$. Consider the vector $\mathbf{v}(x) = x\mathbf{a} + \mathbf{b}$ for any scalar x. By positive definiteness, $|\mathbf{v}(x)|^2 \ge 0$. Expanding this:

$$|x\mathbf{a}+\mathbf{b}|^2 = x^2|\mathbf{a}|^2 + 2x(\mathbf{a}\cdot\mathbf{b}) + |\mathbf{b}|^2 \ge 0.$$

This is a quadratic polynomial in x. Since it is non-negative for all real x, its discriminant must be non-positive:

$$\Delta = 4(\mathbf{a}\cdot\mathbf{b})^2 - 4|\mathbf{a}|^2|\mathbf{b}|^2 \le 0 \implies (\mathbf{a}\cdot\mathbf{b})^2 \le |\mathbf{a}|^2|\mathbf{b}|^2.$$

Taking the square root yields $|\mathbf{a}\cdot\mathbf{b}| \le |\mathbf{a}||\mathbf{b}|$.
(iv) Using the expansion from part (ii) and the Cauchy-Schwarz inequality:

$$|\mathbf{a}+\mathbf{b}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2 + 2(\mathbf{a}\cdot\mathbf{b}) \le |\mathbf{a}|^2 + |\mathbf{b}|^2 + 2|\mathbf{a}||\mathbf{b}| = (|\mathbf{a}| + |\mathbf{b}|)^2.$$

(v) This is simply the expansion of $|\mathbf{a}-\mathbf{b}|^2 = (\mathbf{a}-\mathbf{b})\cdot(\mathbf{a}-\mathbf{b})$.

■

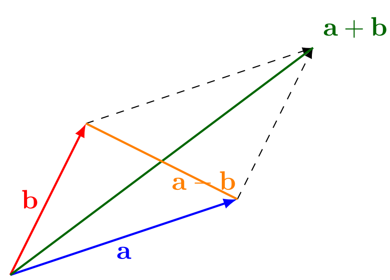


Figure 3.5: The Parallelogram Law involves the diagonals $\mathbf{a}+\mathbf{b}$ and $\mathbf{a}-\mathbf{b}$.

Angle and Orthogonality

To appreciate why Identity (v) in Theorem 3.1.7 is commonly referred to as the Cosine Law for vectors, consider a triangle OAB in the plane. Let the lengths of the sides be OA, OB, and AB, and let θ be the angle at the vertex O. The classical Law of Cosines from trigonometry states:

$$AB^2 = OA^2 + OB^2 - 2(OA)(OB)\cos\theta.$$

We can translate this geometric statement directly into vector notation. Let $\mathbf{a} = \overrightarrow{OA}$ and $\mathbf{b} = \overrightarrow{OB}$. Then the side AB corresponds to the magnitude of the displacement vector $\mathbf{b} - \mathbf{a}$ (or equivalently $\mathbf{a} - \mathbf{b}$), so $AB = |\mathbf{a} - \mathbf{b}|$. Substituting the vector magnitudes into the trigonometric formula yields:

$$|\mathbf{a} - \mathbf{b}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2|\mathbf{a}||\mathbf{b}|\cos\theta.$$

Now, compare this with identity (v) derived purely algebraically from the properties of the dot product:

$$|\mathbf{a} - \mathbf{b}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2(\mathbf{a} \cdot \mathbf{b}).$$

By equating the two expressions for $|\mathbf{a} - \mathbf{b}|^2$, we see that the term containing the angle must correspond to the dot product term:

$$-2|\mathbf{a}||\mathbf{b}|\cos\theta = -2(\mathbf{a} \cdot \mathbf{b}).$$

Simplifying this, we obtain the fundamental relationship connecting the algebraic dot product to the geometric angle:

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos\theta \quad \text{or} \quad \cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}.$$

This relationship allows us to define the angle between two vectors rigorously. However, for the definition to be valid, the ratio on the right-hand side must lie within the range of the cosine function, $[-1, 1]$. Recall the Cauchy-Schwarz inequality established earlier:

$$|\mathbf{a} \cdot \mathbf{b}| \leq |\mathbf{a}||\mathbf{b}|.$$

This inequality ensures that:

$$-1 \leq \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|} \leq 1.$$

Thus, the value is always valid for the arccosine function, permitting the following definition.

Definition 3.1.4. ***Angle Between Vectors***. Let $\mathbf{a}$ and $\mathbf{b}$ be non-zero vectors. The angle θ between them is the unique number in the interval $[0, \pi]$ such that

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}.$$

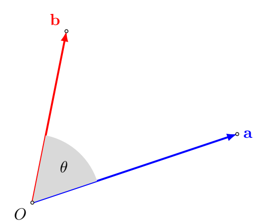


Figure 3.6: The angle θ between vectors $\mathbf{a}$ and $\mathbf{b}$ is defined by their dot product.

Example 3.1.1. Calculating Angles. Let $\mathbf{a} = [1, 0]$, $\mathbf{b} = [-1, 1]$, and $\mathbf{c} = [5, -5\sqrt{3}]$. First, we compute the magnitudes:

$$|\mathbf{a}| = 1, \quad |\mathbf{b}| = \sqrt{(-1)^2 + 1^2} = \sqrt{2}, \quad |\mathbf{c}| = \sqrt{25 + 25(3)} = \sqrt{100} = 10.$$

We calculate the angles between the various pairs:

(i) **Angle between a and b**:

$$\mathbf{a} \cdot \mathbf{b} = (1)(-1) + (0)(1) = -1.$$

$$\cos\theta = \frac{-1}{(1)(\sqrt{2})} = -\frac{1}{\sqrt{2}} \implies \theta = \frac{3\pi}{4}.$$

(ii) **Angle between a and c**:

$$\mathbf{a} \cdot \mathbf{c} = (1)(5) + (0)(-5\sqrt{3}) = 5.$$

$$\cos\theta = \frac{5}{(1)(10)} = \frac{1}{2} \implies \theta = \frac{\pi}{3}.$$

(iii) **Angle between b and c**:

$$\mathbf{b} \cdot \mathbf{c} = (-1)(5) + (1)(-5\sqrt{3}) = -5 - 5\sqrt{3}.$$

$$\cos\theta = \frac{-5(1+\sqrt{3})}{10\sqrt{2}} = -\frac{1+\sqrt{3}}{2\sqrt{2}} = -\frac{\sqrt{2}+\sqrt{6}}{4}.$$

This corresponds to $\theta = 11\pi/12$.

The dot product also provides a simple algebraic test for perpendicularity. If vectors $\mathbf{a}$ and $\mathbf{b}$ are perpendicular, the angle between them is $\pi/2$. Since $\cos(\pi/2) = 0$, this implies $\mathbf{a} \cdot \mathbf{b} = 0$. We formalise this as orthogonality, extending the concept to include the zero vector.

Definition 3.1.5. ***Orthogonality***. Two vectors $\mathbf{a}$ and $\mathbf{b}$ are said to be orthogonal if $\mathbf{a} \cdot \mathbf{b} = 0$. We denote this by $\mathbf{a} \perp \mathbf{b}$.

Note. The zero vector is orthogonal to every vector in $\mathbb{R}^2$.

Orthogonal Projection

A fundamental problem in geometry and physics is the resolution of a vector into distinct components relative to a reference direction. Specifically, given a vector $\mathbf{x}$ and a non-zero reference vector $\mathbf{a}$, we wish to decompose $\mathbf{x}$ into a sum $\mathbf{x} = \mathbf{p} + \mathbf{b}$, where $\mathbf{p}$ is parallel to $\mathbf{a}$ and $\mathbf{b}$ is orthogonal to $\mathbf{a}$. One may approach this problem using elementary algebra. Let $\mathbf{x} = [x_1, x_2]$ and $\mathbf{a} = [a_1, a_2] \neq \mathbf{0}$. We seek scalars t and u such that

$$\mathbf{x} = t\mathbf{a} + u\mathbf{a}^{\perp},$$

where $\mathbf{a}^{\perp} = [a_2, -a_1]$ is a vector orthogonal to $\mathbf{a}$. This leads to the system of linear equations:

$$a_1 t + a_2 u = x_1$$
$$a_2 t - a_1 u = x_2$$

Multiplying the first equation by a_1 and the second by a_2, then adding them, eliminates u:

$$(a_1^2 + a_2^2)t = a_1 x_1 + a_2 x_2.$$

Recognising the terms as dot products and magnitudes, we obtain $t|\mathbf{a}|^2 = \mathbf{x} \cdot \mathbf{a}$, which implies $t = (\mathbf{x} \cdot \mathbf{a})/|\mathbf{a}|^2$. This algebraic result motivates the following vector-based formulation.

***Theorem* 3.1.8. Orthogonal Decomposition**. Let $\mathbf{a} \in \mathbb{R}^2$ be a non-zero vector. For any vector $\mathbf{x} \in \mathbb{R}^2$, there exists a unique scalar t and a unique vector $\mathbf{b}$ such that

$$\mathbf{x} = t\mathbf{a} + \mathbf{b} \quad \text{and} \quad \mathbf{b} \perp \mathbf{a}.$$

The vector $\mathbf{p} = t\mathbf{a}$ is called the orthogonal projection of $\mathbf{x}$ onto $\mathbf{a}$, denoted $\text{proj}_{\mathbf{a}}\mathbf{x}$.

Proof. We seek a scalar t such that the vector $\mathbf{b} = \mathbf{x} - t\mathbf{a}$ is orthogonal to $\mathbf{a}$.

$$(\mathbf{x} - t\mathbf{a}) \cdot \mathbf{a} = 0 \implies \mathbf{x} \cdot \mathbf{a} - t(\mathbf{a} \cdot \mathbf{a}) = 0.$$

Since $\mathbf{a} \neq \mathbf{0}$, $\mathbf{a} \cdot \mathbf{a} \neq 0$, yielding the unique solution:

$$t = \frac{\mathbf{x} \cdot \mathbf{a}}{\mathbf{a} \cdot \mathbf{a}} = \frac{\mathbf{x} \cdot \mathbf{a}}{|\mathbf{a}|^2}.$$

The projection vector is therefore:

$$\text{proj}_{\mathbf{a}}\mathbf{x} = \left(\frac{\mathbf{x} \cdot \mathbf{a}}{|\mathbf{a}|^2}\right)\mathbf{a}.$$

■

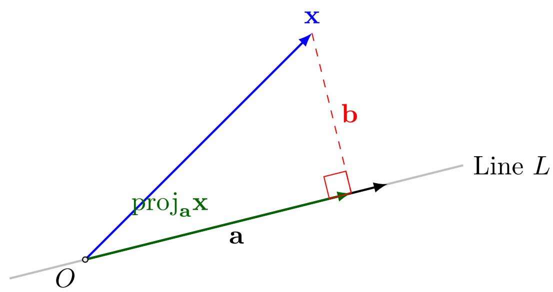


Figure 3.7: The orthogonal decomposition of $\mathbf{x}$ onto the line generated by $\mathbf{a}$.

Graphically, the term $\mathbf{b} = \mathbf{x} - \text{proj}_{\mathbf{a}}\mathbf{x}$ represents the "error" or the perpendicular distance from the tip of $\mathbf{x}$ to the line spanned by $\mathbf{a}$.

***Corollary* 3.1.1.** *Special Cases*

(i) $\mathbf{x}$ and $\mathbf{a}$ are linearly dependent if and only if $\text{proj}_{\mathbf{a}}\mathbf{x} = \mathbf{x}$ (i.e., $\mathbf{b} = \mathbf{0}$).
(ii) $\mathbf{x}$ and $\mathbf{a}$ are orthogonal if and only if $\text{proj}_{\mathbf{a}}\mathbf{x} = \mathbf{0}$.

The length of the projection vector is given by:

$$|\text{proj}_{\mathbf{a}}\mathbf{x}| = \left|\frac{\mathbf{x}\cdot\mathbf{a}}{|\mathbf{a}|^2}\mathbf{a}\right| = \frac{|\mathbf{x}\cdot\mathbf{a}|}{|\mathbf{a}|^2}|\mathbf{a}| = \frac{|\mathbf{x}\cdot\mathbf{a}|}{|\mathbf{a}|}.$$

This formula is particularly useful for finding the distance between points projected onto a line.

Example 3.1.2. Projection of a Segment. Let $X = (-1, 3)$, $Y = (3, 0)$, $A = (2, 4)$, and $B = (1, -2)$ be four points in the plane. We wish to find the length of the orthogonal projection of the segment XY onto the straight line passing through A and B.

First, we determine the displacement vectors representing the segment and the line direction. Let

$$\mathbf{z} = \overrightarrow{XY} = [3 - (-1), 0 - 3] = [4, -3].$$

$$\mathbf{c} = \overrightarrow{AB} = [1 - 2, -2 - 4] = [-1, -6].$$

The length of the projection of the segment XY onto the line AB is equivalent to the magnitude of the projection of vector $\mathbf{z}$ onto vector $\mathbf{c}$:

$$\text{Length} = \frac{|\mathbf{z}\cdot\mathbf{c}|}{|\mathbf{c}|}.$$

Computing the dot product and magnitude:

$$\mathbf{z}\cdot\mathbf{c} = (4)(-1) + (-3)(-6) = -4 + 18 = 14.$$

$$|\mathbf{c}| = \sqrt{(-1)^2 + (-6)^2} = \sqrt{1 + 36} = \sqrt{37}.$$

Thus, the length of the projection is $14/\sqrt{37}$.

3.2 The Equation of a Straight Line

In elementary coordinate geometry, a straight line is defined as the locus of points $X = (x, y)$ satisfying a linear equation of the form

$$ax + by + c = 0,$$

where a and b are not both zero. By translating this algebraic constraint into the language of vectors, we uncover the geometric significance of the coefficients a and b.

The Point-Normal Form

Let $\mathbf{n} = [a, b]$ and let $\mathbf{x} = [x, y]$ be the position vector of a point X. The linear term $ax + by$ is precisely the dot product $\mathbf{n} \cdot \mathbf{x}$. Consequently, the equation of the line may be rewritten as:

$$\mathbf{n} \cdot \mathbf{x} + c = 0.$$

To interpret this geometrically, let P be a fixed point on the line with position vector $\mathbf{p}$. Since P lies on the line, its coordinates satisfy the equation, so $\mathbf{n} \cdot \mathbf{p} + c = 0$, which implies $c = -\mathbf{n} \cdot \mathbf{p}$. Substituting this back into the general equation yields:

$$\mathbf{n} \cdot \mathbf{x} - \mathbf{n} \cdot \mathbf{p} = 0 \implies \mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0.$$

The vector $\mathbf{x} - \mathbf{p}$ represents the displacement $\overrightarrow{PX}$ along the line. The condition $\mathbf{n} \cdot \overrightarrow{PX} = 0$ implies that the vector $\mathbf{n}$ is orthogonal to the direction of the line.

Definition 3.2.1. ***Normal Vector***. A non-zero vector $\mathbf{n}$ is called a normal vector to a straight line L if $\mathbf{n}$ is orthogonal to the displacement vector between any two distinct points on L.

This leads to the vector characterisation of a straight line.

***Theorem* 3.2.1. Point-Normal Form**. The straight line passing through a specific point P with position vector $\mathbf{p}$, and perpendicular to a non-zero normal vector $\mathbf{n}$, consists of all points X with position vectors $\mathbf{x}$ satisfying:

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0.$$

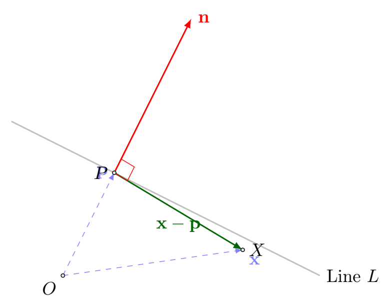


Figure 3.8: The line L through P with normal $\mathbf{n}$. The displacement $\mathbf{x} - \mathbf{p}$ is orthogonal to $\mathbf{n}$.

Example 3.2.1. Constructing a Line. Find the equation of the line passing through $A(3, 2)$ and perpendicular to the vector $\mathbf{n} = [2, -1]$. Using the point-normal form:

$$[2, -1] \cdot [x - 3, y - 2] = 0$$

$$2(x - 3) - 1(y - 2) = 0$$

$$2x - 6 - y + 2 = 0 \implies 2x - y - 4 = 0.$$

This formulation allows us to easily find the perpendicular bisector of a segment.

Example 3.2.2. Let $A = (-1, 3)$ and $B = (5, 1)$. The perpendicular bisector of $\overline{AB}$ passes through the midpoint M of the segment and has $\overrightarrow{AB}$ as a normal vector.

(i) Midpoint $\mathbf{m} = \frac{1}{2}(\mathbf{a} + \mathbf{b}) = [\frac{-1+5}{2}, \frac{3+1}{2}] = [2, 2]$.

(ii) Normal vector $\mathbf{n} = \mathbf{b} - \mathbf{a} = [5 - (-1), 1 - 3] = [6, -2]$. For simplicity, we can use the parallel vector $[3, -1]$.

(iii) Equation: $[3, -1] \cdot [x - 2, y - 2] = 0 \implies 3(x - 2) - (y - 2) = 0 \implies 3x - y - 4 = 0$.

Distance from a Point to a Line

The vector approach provides an elegant derivation for the distance from a point to a line, a result that is often tedious to prove using classical coordinates.

***Theorem* 3.2.2. Distance to a Line**. Let L be the line defined by $\mathbf{n} \cdot \mathbf{x} + c = 0$. The perpendicular distance from an arbitrary point X_0 (with position vector $\mathbf{x}_0$) to L is given by:

$$d(X_0, L) = \frac{|\mathbf{n} \cdot \mathbf{x}_0 + c|}{|\mathbf{n}|}.$$

Proof. Let P be any point on the line L. Then $\mathbf{n} \cdot \mathbf{p} + c = 0$, so $c = -\mathbf{n} \cdot \mathbf{p}$. The distance from X_0 to L is the length of the orthogonal projection of the displacement vector $\overrightarrow{PX_0}$ onto the normal vector $\mathbf{n}$. Using the projection formula from Theorem 3.1.8:

$$d = |\text{proj}_{\mathbf{n}}(\mathbf{x}_0 - \mathbf{p})| = \left| \left(\frac{(\mathbf{x}_0 - \mathbf{p}) \cdot \mathbf{n}}{|\mathbf{n}|^2} \right) \mathbf{n} \right| = \frac{|(\mathbf{x}_0 - \mathbf{p}) \cdot \mathbf{n}|}{|\mathbf{n}|}.$$

Expanding the dot product in the numerator:

$$d = \frac{|\mathbf{n} \cdot \mathbf{x}_0 - \mathbf{n} \cdot \mathbf{p}|}{|\mathbf{n}|}.$$

Substituting $-\mathbf{n} \cdot \mathbf{p} = c$, we obtain:

$$d = \frac{|\mathbf{n} \cdot \mathbf{x}_0 + c|}{|\mathbf{n}|}.$$

■

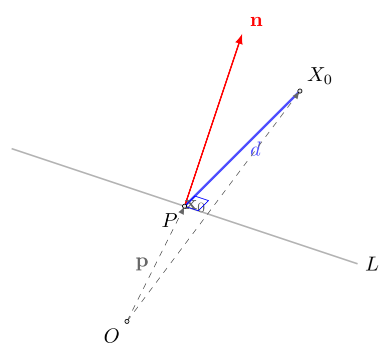


Figure 3.9: The distance d is the magnitude of the projection of $\overrightarrow{PX_0}$ onto $\mathbf{n}$.

This theorem suggests a canonical way to represent straight lines. If we divide the equation $\mathbf{n} \cdot \mathbf{x} + c = 0$ by the magnitude $|\mathbf{n}|$, we obtain the Hessian normal form:

$$\frac{\mathbf{n}}{|\mathbf{n}|} \cdot \mathbf{x} + \frac{c}{|\mathbf{n}|} = 0.$$

Defining the unit normal $\hat{\mathbf{n}} = \mathbf{n}/|\mathbf{n}|$ and $p = c/|\mathbf{n}|$, the function $f(\mathbf{x}) = \hat{\mathbf{n}} \cdot \mathbf{x} + p$ has the property that $|f(\mathbf{x}_0)|$ gives the exact distance from X_0 to the line.

Example 3.2.3. Distance Calculation. Find the distance from $X_0(2, -1)$ to the line $3x + 4y - 12 = 0$. Here $\mathbf{n} = [3, 4]$, so $|\mathbf{n}| = \sqrt{3^2 + 4^2} = 5$.

$$d = \frac{|3(2) + 4(-1) - 12|}{5} = \frac{|6 - 4 - 12|}{5} = \frac{|-10|}{5} = 2.$$

Angle Between Lines

The angle between two straight lines is defined as the angle between their respective normal vectors. This definition is consistent with the geometric intuition that if two lines intersect, the angle between them is preserved if both are rotated by 90 degrees (transforming the lines into their normals).

***Theorem* 3.2.3. Angle Between Lines**. Let L_1 and L_2 be lines with normal vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ respectively. The angle $\theta \in [0, \pi]$ between the lines is given by:

$$\cos\theta = \frac{\mathbf{n}_1 \cdot \mathbf{n}_2}{|\mathbf{n}_1||\mathbf{n}_2|}.$$

If we are interested in the acute angle ϕ between the lines, we take the absolute value of the right-hand side: $\cos\phi = \frac{|\mathbf{n}_1 \cdot \mathbf{n}_2|}{|\mathbf{n}_1||\mathbf{n}_2|}$.

Example 3.2.4. Find the angle between the lines $x - 2y + 3 = 0$ and $3x + y - 5 = 0$. The normal vectors are $\mathbf{n}_1 = [1, -2]$ and $\mathbf{n}_2 = [3, 1]$.

$$\mathbf{n}_1 \cdot \mathbf{n}_2 = (1)(3) + (-2)(1) = 1.$$

$$|\mathbf{n}_1| = \sqrt{5}, \quad |\mathbf{n}_2| = \sqrt{10}.$$

$$\cos\theta = \frac{1}{\sqrt{5}\sqrt{10}} = \frac{1}{\sqrt{50}} = \frac{1}{5\sqrt{2}}.$$

3.3 Exercises

1. Let $\mathbf{a} = [3, 4]$, $\mathbf{b} = [5, -12]$, and $\mathbf{c} = [1, -3]$.
   - (a) Compute the scalar products $\mathbf{a} \cdot \mathbf{b}$, $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c})$, and $(\mathbf{a} - \mathbf{c}) \cdot \mathbf{b}$.
   - (b) Find the angle between $\mathbf{a}$ and $\mathbf{b}$.
   - (c) Determine the orthogonal projection of $\mathbf{c}$ onto $\mathbf{a}$.
   - (d) Find a scalar k such that $\mathbf{a} + k\mathbf{b}$ is orthogonal to $\mathbf{a}$.

2. Find the equation of the straight line in the form $ax + by + c = 0$ satisfying the following conditions:
   - (a) Passing through $P(-2, 5)$ with normal vector $\mathbf{n} = [3, -1]$.
   - (b) Passing through $A(4, 1)$ and $B(2, -3)$.
   - (c) Passing through $C(1, 2)$ and perpendicular to the line $2x - 5y + 10 = 0$.

3. Consider the lines $L_1 : 3x - 4y + 5 = 0$ and $L_2 : x + y - 2 = 0$.
   - (a) Calculate the perpendicular distance from the point $P(2, -1)$ to L_1.
   - (b) Find the point on L_1 that is closest to P.
   - (c) Determine the cosine of the acute angle between L_1 and L_2.

4. Find the values of k such that the vectors $\mathbf{u} = [k, -2]$ and $\mathbf{v} = [3, k + 5]$ are:
   - (a) Orthogonal.
   - (b) Parallel.
   - (c) Separated by an angle of $\pi/3$.

5. Pythagorean Generalisation. Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^2$. Prove that $|\mathbf{u} + \mathbf{v}|^2 = |\mathbf{u}|^2 + |\mathbf{v}|^2$ if and only if $\mathbf{u}$ and $\mathbf{v}$ are orthogonal. Interpret this geometrically.

6. The Rhombus Identity. Let $\mathbf{u}$ and $\mathbf{v}$ be vectors of equal magnitude (i.e., $|\mathbf{u}| = |\mathbf{v}|$).
   - (a) Prove that the vectors $\mathbf{u} + \mathbf{v}$ and $\mathbf{u} - \mathbf{v}$ are orthogonal.
   - (b) Interpret this result in terms of the diagonals of a rhombus.

7. Apollonius' Theorem. Let ABC be a triangle. Let M be the midpoint of the side BC. Prove, using vector algebra, that $AB^2 + AC^2 = 2(AM^2 + BM^2)$.

Remark. Set the origin at M. Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be the position vectors. Note that $\mathbf{b} = -\mathbf{c}$.

8. **Equality in Cauchy-Schwarz.** We established that $|\mathbf{a}\cdot\mathbf{b}| \le |\mathbf{a}||\mathbf{b}|$. Prove that equality holds, i.e., $|\mathbf{a}\cdot\mathbf{b}| = |\mathbf{a}||\mathbf{b}|$, if and only if $\mathbf{a}$ and $\mathbf{b}$ are linearly dependent.

9. **Area of a Triangle.** Let a triangle have vertices at the origin O and points A, B with position vectors $\mathbf{a}, \mathbf{b}$.
   (a) From elementary geometry, the area is $\frac{1}{2}|\mathbf{a}||\mathbf{b}|\sin\theta$. Use the relation between $\sin\theta$ and $\cos\theta$ to prove that the area can be written as:
   $$\text{Area} = \frac{1}{2}\sqrt{|\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a}\cdot\mathbf{b})^2}.$$
   (b) Verify that if $\mathbf{a} = [a_1, a_2]$ and $\mathbf{b} = [b_1, b_2]$, this expression reduces to $\frac{1}{2}|a_1b_2 - a_2b_1|$.

10. **Vector Angle Bisector.** Let $\mathbf{u}$ and $\mathbf{v}$ be non-zero, non-parallel vectors.
    (a) Show that the vector $\mathbf{w} = |\mathbf{v}|\mathbf{u} + |\mathbf{u}|\mathbf{v}$ bisects the angle between $\mathbf{u}$ and $\mathbf{v}$.
    (b) Hence, or otherwise, find the equation of the angle bisector of the lines $3x-4y = 0$ and $5x+12y = 0$.

11. **Minimisation and Projection.** Let $\mathbf{a}$ and $\mathbf{b}$ be vectors with $\mathbf{a} \neq \mathbf{0}$. Consider the function $f(t) = |\mathbf{b} - t\mathbf{a}|^2$.
    (a) Using calculus or by completing the square, find the value of t that minimizes $f(t)$.
    (b) Show that for this optimal t, the vector $\mathbf{b} - t\mathbf{a}$ is orthogonal to $\mathbf{a}$.
    (c) Explain how this relates to the definition of orthogonal projection given in Theorem 3.1.8.

12. **Reflection Operator.** Let L be a line passing through the origin with a normal vector $\mathbf{n}$. The reflection of a vector $\mathbf{x}$ across the line L is given by the map:
    $$R(\mathbf{x}) = \mathbf{x} - 2\text{proj}_{\mathbf{n}}\mathbf{x}.$$
    (a) Draw a diagram to justify this definition geometrically.
    (b) Prove that R is an isometry, meaning it preserves magnitudes: $|R(\mathbf{x})| = |\mathbf{x}|$.
    (c) Prove that R is a linear map: $R(r\mathbf{x} + s\mathbf{y}) = rR(\mathbf{x}) + sR(\mathbf{y})$.

13. **Locus of Angle Bisectors.** Consider two non-parallel lines given by normal equations:
    $$\mathbf{n}_1\cdot\mathbf{x} + c_1 = 0 \quad \text{and} \quad \mathbf{n}_2\cdot\mathbf{x} + c_2 = 0.$$
    Prove that the locus of points equidistant from these two lines is given by the two perpendicular lines:
    $$\frac{\mathbf{n}_1\cdot\mathbf{x} + c_1}{|\mathbf{n}_1|} = \pm\frac{\mathbf{n}_2\cdot\mathbf{x} + c_2}{|\mathbf{n}_2|}.$$

14. $\star$ **The Circumcentre.** Let ABC be a triangle with position vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ relative to an arbitrary origin.
    (a) The perpendicular bisector of the side AB is the set of points $\mathbf{x}$ such that $|\mathbf{x} - \mathbf{a}| = |\mathbf{x} - \mathbf{b}|$. Show that this is equivalent to $\mathbf{x}\cdot(\mathbf{b} - \mathbf{a}) = \frac{1}{2}(|\mathbf{b}|^2 - |\mathbf{a}|^2)$.
    (b) Suppose the perpendicular bisectors of sides AB and BC intersect at a point O. Show that O must also lie on the perpendicular bisector of AC.
    (c) Conclude that the three perpendicular bisectors of a triangle are concurrent at the circumcentre.

15. $\star$ **The Orthocentre.** Let ABC be a triangle. Let the position vectors of the vertices be $\mathbf{a}, \mathbf{b}, \mathbf{c}$.
    (a) An altitude from vertex A is the line passing through A and perpendicular to BC. Write down the condition for a point $\mathbf{x}$ to lie on this altitude involving the dot product.
    (b) Let the origin be the circumcentre of the triangle (so $|\mathbf{a}| = |\mathbf{b}| = |\mathbf{c}| = R$). Consider the point H defined by the position vector $\mathbf{h} = \mathbf{a} + \mathbf{b} + \mathbf{c}$.
    (c) Compute the dot product $\mathbf{h}\cdot(\mathbf{b} - \mathbf{c})$.
    (d) Deduce that H lies on the altitude from A. By symmetry, conclude that H lies on all three altitudes. This point is called the orthocentre.

Chapter 4

Vectors and Geometry in Space

In the preceding chapters, we established a rigorous correspondence between the geometry of the plane and the algebraic structure of $\mathbb{R}^2$. We now extend this framework to three-dimensional space. While the geometric complexity increases, the algebraic methods developed for the plane, specifically the vector operations, generalise naturally. In this chapter, we construct the vector space $\mathbb{R}^3$ and explore the fundamental interplay between algebraic equations and solid geometry.

4.1 Cartesian Coordinates in Space

To define the position of a point in space algebraically, we require a reference system consisting of three mutually perpendicular lines intersecting at a common point, which we designate as the origin O. We assign a direction to each line, establishing them as the x-axis, the y-axis, and the z-axis.

Orientation

Unlike the plane, where the relative orientation of axes is generally fixed by convention (counter-clockwise from x to y), three-dimensional space offers a choice of orientation. There are $3! = 6$ possible assignments of axes to three mutually orthogonal lines. These assignments partition into two distinct classes.

(i) **Right-Hand Systems:** Assignments that can be rotated into one another.
(ii) **Left-Hand Systems:** Assignments that are mirror images of the first group and cannot be reached via rotation.

Definition 4.1.1. ***Right-Hand Coordinate System.*** A coordinate system is said to be right-handed if the x, y, and z axes correspond respectively to the thumb, index finger, and middle finger of the right hand when extended in mutually perpendicular directions.

We shall exclusively adopt the right-hand coordinate system. This choice ensures consistency with standard conventions in physics and vector calculus, particularly regarding the cross product (to be discussed in subsequent sections).

The Bijection between Points and Triples

We identify each coordinate axis with the real line $\mathbb{R}$. Given an ordered triple of real numbers (a_1, a_2, a_3), we associate it with a unique point A in space via the intersection of three planes:

- The plane perpendicular to the x-axis at a_1.

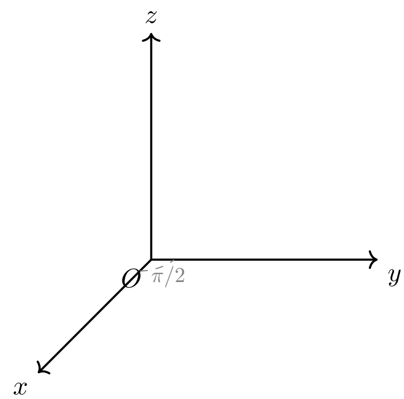


Figure 4.1: A right-handed coordinate system.

- The plane perpendicular to the y-axis at a_2.
- The plane perpendicular to the z-axis at a_3.

Conversely, for any point B in space, we construct planes passing through B perpendicular to the axes, intersecting them at coordinates b_1, b_2, b_3 respectively.

Definition 4.1.2. ***Space Coordinates***. The ordered triple of real numbers (a_1, a_2, a_3) associated with a point A are called the coordinates of A. We write $A = (a_1, a_2, a_3)$. The set of all such triples is denoted by $\mathbb{R}^3$.

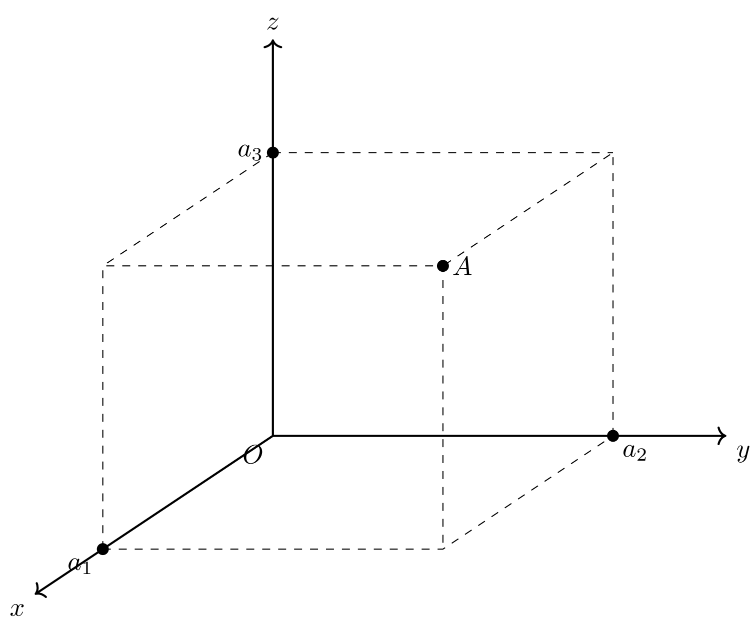


Figure 4.2: The point $A(a_1, a_2, a_3)$ determined by the intersection of planes.

By this correspondence, $\mathbb{R}^3$ becomes an algebraic model for three-dimensional Euclidean space.

The Euclidean Metric in Space

The metric properties of space (specifically distance), are derived directly from the coordinate definitions by iteratively applying Pythagoras' Theorem.

***Theorem* 4.1.1. Distance Formula in $\mathbb{R}^3$.** Let $A = (a_1, a_2, a_3)$ and $B = (b_1, b_2, b_3)$ be points in space. The Euclidean distance between A and B, denoted $|AB|$, is given by:

$$|AB| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2}.$$

Proof. Construct the planes through A and B perpendicular to the coordinate axes. These planes define a rectangular box (cuboid) with diagonal AB. Consider the projection of the segment AB onto the xy-plane. Let $A' = (a_1, a_2, b_3)$ and $C = (b_1, b_2, b_3)$. The points A', C and B all lie on the plane $z = b_3$. In this plane, the distance $|A'C|$ is the hypotenuse of a right-angled triangle with sides $|a_1 - b_1|$ and $|a_2 - b_2|$. Thus:

$$|A'C|^2 = (a_1 - b_1)^2 + (a_2 - b_2)^2.$$

Now consider the triangle ACA'. The segment AC is the vertical edge connecting the plane $z = a_3$ to $z = b_3$, so $|AC| = |a_3 - b_3|$. Since the vertical line AC is perpendicular to the plane containing $A'C$, the triangle ACA' is right-angled at A (or C depending on projection orientation, essentially we are finding the space diagonal). More precisely, let us use the auxiliary point $C' = (b_1, b_2, a_3)$ in the plane $z = a_3$. Then $|AC'|^2 = (a_1 - b_1)^2 + (a_2 - b_2)^2$. The triangle $AC'B$ is right-angled at C' because $C'B$ is parallel to the z-axis.

$$|AB|^2 = |AC'|^2 + |C'B|^2 = \left((a_1 - b_1)^2 + (a_2 - b_2)^2\right) + (a_3 - b_3)^2.$$

Taking the square root yields the result. ■

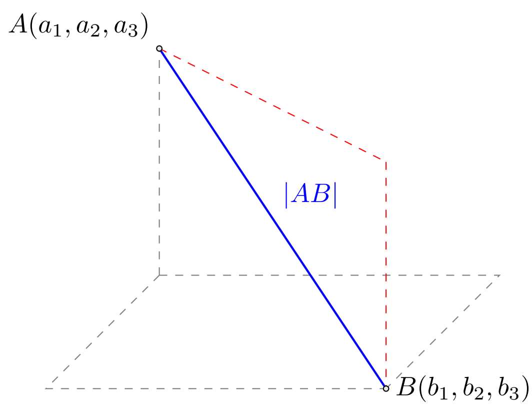


Figure 4.3: The distance between points A and B derived via the space diagonal of a cuboid.

The Equation of a Sphere

The distance formula allows us to characterise the sphere algebraically, perfectly mirroring the definition of a circle in the plane.

Definition 4.1.3. *Sphere*. A sphere is the locus of all points in space at a fixed distance r (the radius) from a fixed point C (the centre).

Let the centre be $C = (c_1, c_2, c_3)$ and let $X = (x, y, z)$ be an arbitrary point on the sphere. By definition, $|CX| = r$. Squaring both sides and applying the distance formula:

$$(x - c_1)^2 + (y - c_2)^2 + (z - c_3)^2 = r^2.$$

Example 4.1.1. Spherical Equation. Determine the set of points satisfying the equation:

$$x^2 + y^2 + z^2 - 2x - 4y + 6z - 11 = 0.$$

We complete the square for each variable:

$$(x^2 - 2x) + (y^2 - 4y) + (z^2 + 6z) = 11$$

$$(x-1)^2 - 1 + (y-2)^2 - 4 + (z+3)^2 - 9 = 11$$
$$(x-1)^2 + (y-2)^2 + (z+3)^2 = 11 + 1 + 4 + 9 = 25.$$

This represents a sphere centered at $(1, 2, -3)$ with radius $\sqrt{25} = 5$.

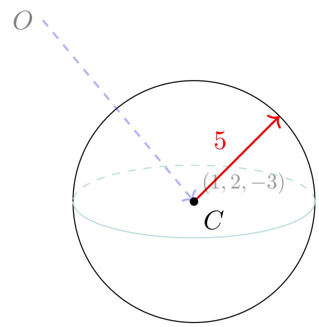


Figure 4.4: The distance between points A and B derived via the space diagonal of a cuboid.

4.2 Vectors in Space

Just as the cartesian plane allowed us to transition from coordinate pairs to two-dimensional vectors, the cartesian space structures the extension to three-dimensional vectors. The algebraic definition generalizes naturally: points and vectors are both represented by ordered triples, yet they play distinct roles in our geometric framework.

Definition 4.2.1. ***Vector in Space***. A vector in space is an ordered triple of real numbers. The individual numbers are called the components of the vector. The set of all such vectors is denoted by $\mathbb{R}^3$.

As in the planar case, an ordered triple (x, y, z) possesses a dual nature: it may specify a fixed geometric location (a point) or a directed magnitude (a vector). To maintain rigour and clarity, we strictly adhere to the notational convention established in the previous chapter.

Notation 4.2.1. Point vs Vector

- A *point* is denoted by a capital letter with coordinates in parentheses: $A = (a_1, a_2, a_3)$.
- A *vector* is denoted by a bold lower-case letter with components in brackets: $\mathbf{a} = [a_1, a_2, a_3]$.

Geometric Representation

Geometrically, the vector $\mathbf{a} = [a_1, a_2, a_3]$ is represented by the directed line segment originating at the origin $O(0, 0, 0)$ and terminating at the point $A(a_1, a_2, a_3)$. In this context, $\mathbf{a}$ is the *position vector* of point A. Conversely, any directed segment from the origin to a point B defines unique components for a vector $\mathbf{b}$.

Three specific vectors form the standard basis for $\mathbb{R}^3$, corresponding to the unit points on the coordinate axes:

$$\mathbf{e}_1 = [1, 0, 0], \quad \mathbf{e}_2 = [0, 1, 0], \quad \mathbf{e}_3 = [0, 0, 1].$$

The zero vector is denoted by $\mathbf{0} = [0, 0, 0]$ and corresponds to the degenerate segment at the origin.

The length of the directed segment representing $\mathbf{a}$ is termed its magnitude. Following the distance formula derived in the previous section, we define:

Definition 4.2.2. ***Magnitude***. The magnitude (or norm) of a vector $\mathbf{a} = [a_1, a_2, a_3]$ is the non-negative real number

$$|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

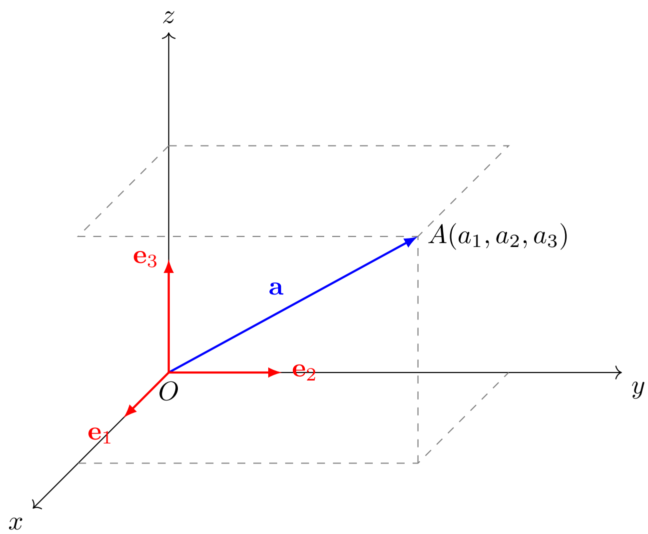


Figure 4.5: The position vector $\mathbf{a}$ and the standard basis vectors in $\mathbb{R}^3$.

Consistent with the properties of $\mathbb{R}^2$, the zero vector is the unique vector with zero magnitude: $|\mathbf{a}| = 0$ if and only if $\mathbf{a} = \mathbf{0}$.

The Vector Space $\mathbb{R}^3$

Points in space are geometric locations; they do not inherently admit addition or multiplication. However, vectors, being algebraic entities, support these operations directly. The definitions of vector addition and scalar multiplication in space are immediate extensions of their planar counterparts.

Definition 4.2.3. ***Operations in*** $\mathbb{R}^3$. Let $\mathbf{a} = [a_1, a_2, a_3]$ and $\mathbf{b} = [b_1, b_2, b_3]$ be vectors, and let r be a scalar.

1. The **vector sum** is $\mathbf{a} + \mathbf{b} = [a_1 + b_1, a_2 + b_2, a_3 + b_3]$.
2. The **scalar multiple** is $r\mathbf{a} = [ra_1, ra_2, ra_3]$.

Geometrically, $\mathbf{a} + \mathbf{b}$ represents the diagonal of the parallelogram defined by $\mathbf{a}$ and $\mathbf{b}$ (which lies in the specific plane containing these two vectors and the origin). Similarly, $r\mathbf{a}$ represents a scaling of the length of $\mathbf{a}$ by a factor of $|r|$, preserving direction if $r > 0$ and reversing it if $r < 0$.

The set of all ordered triples, equipped with these two operations, forms the vector space $\mathbb{R}^3$. The algebraic structure of $\mathbb{R}^3$ is identical to that of $\mathbb{R}^2$; the increase in dimension does not alter the fundamental axioms.

Theorem **4.2.1. Properties of** $\mathbb{R}^3$. Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be vectors in $\mathbb{R}^3$ and let r, s be scalars. The following properties hold:

(i) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ (Commutativity)
(ii) $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$ (Associativity of Addition)
(iii) There exists a unique vector $\mathbf{0}$ such that $\mathbf{a} + \mathbf{0} = \mathbf{a}$. (Additive Identity)
(iv) For every $\mathbf{a}$, there exists a unique vector $-\mathbf{a}$ such that $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$. (Additive Inverse)
(v) $(rs)\mathbf{a} = r(s\mathbf{a})$ (Associativity of Scalar Multiplication)
(vi) $(r + s)\mathbf{a} = r\mathbf{a} + s\mathbf{a}$ (Distributivity over Scalar Sums)
(vii) $r(\mathbf{a} + \mathbf{b}) = r\mathbf{a} + r\mathbf{b}$ (Distributivity over Vector Sums)
(viii) $1\mathbf{a} = \mathbf{a}$ (Scalar Identity)

Proof. These properties are verified component-wise, relying solely on the field properties of real numbers.

For instance, to prove (v):

$$(rs)\mathbf{a} = [(rs)a_1, (rs)a_2, (rs)a_3] = [r(sa_1), r(sa_2), r(sa_3)] = r[sa_1, sa_2, sa_3] = r(s\mathbf{a}).$$

The proofs for the remaining properties follow an identical pattern to those in Theorem 2.2.1. ■

Since the axioms of $\mathbb{R}^3$ match those of $\mathbb{R}^2$, any theorem derived solely from these axioms applies purely to $\mathbb{R}^3$ without need for modification. A prime example is the Zero Product Law.

***Theorem* 4.2.2. Zero Product Law in Space**. Let $\mathbf{a} \in \mathbb{R}^3$ and $r \in \mathbb{R}$. Then $r\mathbf{a} = \mathbf{0}$ if and only if $r = 0$ or $\mathbf{a} = \mathbf{0}$.

Proof. The axiomatic proof provided in Theorem 2.2.2 relies only on the existence of identities and inverses and the distributivity laws. Since Theorem 2.2.1 and the theorem above share the same algebraic structure, the proof remains valid for vectors in space. ■

4.3 Linear Combinations and Independence

The concept of linear combinations, introduced in the context of the plane, extends naturally to $\mathbb{R}^3$. This generalisation forms the foundation for understanding dimension and basis in higher-dimensional vector spaces.

Definition 4.3.1. *Linear Combination*. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ be vectors in $\mathbb{R}^3$ and let $c_1, c_2, \ldots, c_k$ be scalars. The vector $\mathbf{v}$ defined by

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

is called a linear combination of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

Specifically, the linear combinations of a single vector $\mathbf{a}$ constitute the line passing through the origin in the direction of $\mathbf{a}$ (vectors of the form $r\mathbf{a}$). The linear combinations of two non-collinear vectors $\mathbf{a}$ and $\mathbf{b}$ form the plane passing through the origin spanned by these vectors (vectors of the form $r\mathbf{a} + s\mathbf{b}$).

Linear Independence of Two Vectors

The definition of linear independence for two vectors in space mirrors that of the plane.

Definition 4.3.2. *Independence of Two Vectors*. Two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ are linearly independent if neither is a scalar multiple of the other. If one is a scalar multiple of the other, they are linearly dependent.

We provide four equivalent conditions for linear dependence. Note the expansion of the component-wise condition compared to $\mathbb{R}^2$.

***Theorem* 4.3.1. Dependence of Two Vectors**. Let $\mathbf{a} = [a_1, a_2, a_3]$ and $\mathbf{b} = [b_1, b_2, b_3]$. The following statements are equivalent:

(i) $\mathbf{a}$ and $\mathbf{b}$ are linearly dependent.
(ii) The points $O(0, 0, 0)$, $A(a_1, a_2, a_3)$, and $B(b_1, b_2, b_3)$ are collinear.
(iii) There exist scalars r and s, not both zero, such that $r\mathbf{a} + s\mathbf{b} = \mathbf{0}$.
(iv) $a_1b_2 - a_2b_1 = 0, \quad a_2b_3 - a_3b_2 = 0, \quad \text{and} \quad a_3b_1 - a_1b_3 = 0.$

Proof. The equivalence of (i), (ii), and (iii) follows the exact same arguments as provided in Chapter 2 for plane vectors. We focus on the equivalence of (i) and (iv).

Assume (i) holds. Without loss of generality, let $\mathbf{a} = r\mathbf{b}$. Then $a_i = rb_i$ for $i = 1, 2, 3$. Substituting into the first expression of (iv):

$$a_1b_2 - a_2b_1 = (rb_1)b_2 - (rb_2)b_1 = r(b_1b_2 - b_2b_1) = 0.$$

By symmetry, the other two expressions also vanish.

Conversely, assume (iv) holds. If $\mathbf{b} = \mathbf{0}$, dependence is trivial. Assume $\mathbf{b} \neq \mathbf{0}$; then at least one component, say b_1, is non-zero. From $a_1b_2 - a_2b_1 = 0$, we have $a_2 = (a_1/b_1)b_2$. From $a_3b_1 - a_1b_3 = 0$, we have $a_3 = (a_1/b_1)b_3$. Let $r = a_1/b_1$. Then $a_1 = rb_1$, $a_2 = rb_2$, and $a_3 = rb_3$, implying $\mathbf{a} = r\mathbf{b}$. ■

Note. Condition (iv) requires that all three 2×2 determinants (minors) formed by the components of $\mathbf{a}$ and $\mathbf{b}$ must vanish. This anticipates the definition of the cross product; specifically, vectors are dependent if and only if their cross product is the zero vector.

By negating the theorem above, we characterise linear independence.

***Theorem* 4.3.2. Independence of Two Vectors**. Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$. The following are equivalent:

(i) $\mathbf{a}$ and $\mathbf{b}$ are linearly independent.
(ii) O, A, B are distinct and not collinear.
(iii) $r\mathbf{a} + s\mathbf{b} = \mathbf{0}$ implies $r = s = 0$.
(iv) At least one of the expressions $a_1b_2 - a_2b_1$, $a_2b_3 - a_3b_2$, or $a_3b_1 - a_1b_3$ is non-zero.

Linear Independence of Three Vectors

The introduction of a third dimension permits the consideration of three vectors.

Definition 4.3.3. *Independence of Three Vectors*. Three vectors $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ are linearly dependent if at least one of them is a linear combination of the other two. They are linearly independent if none is a linear combination of the others.

For instance, the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are linearly independent. We now formulate the algebraic criterion for dependence involving linear combinations yielding zero.

***Theorem* 4.3.3. Algebraic Criterion for Three Vectors**. Three vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are linearly dependent if and only if there exist scalars r, s, t, not all zero, such that

$$r\mathbf{a} + s\mathbf{b} + t\mathbf{c} = \mathbf{0}.$$

Proof. Suppose the vectors are dependent. Then one vector, say $\mathbf{c}$, is a linear combination of the others: $\mathbf{c} = r\mathbf{a} + s\mathbf{b}$. This rearranges to $r\mathbf{a} + s\mathbf{b} + (-1)\mathbf{c} = \mathbf{0}$. Since the coefficient of $\mathbf{c}$ is $-1 \neq 0$, the scalars are not all zero.

Conversely, suppose $r\mathbf{a} + s\mathbf{b} + t\mathbf{c} = \mathbf{0}$ with at least one non-zero scalar. Without loss of generality, assume $t \neq 0$. We can solve for $\mathbf{c}$:

$$\mathbf{c} = -\frac{r}{t}\mathbf{a} - \frac{s}{t}\mathbf{b}.$$

Thus $\mathbf{c}$ is a linear combination of $\mathbf{a}$ and $\mathbf{b}$, proving dependence. ■

Just as the linear dependence of two vectors implies collinearity, the linear dependence of three vectors implies coplanarity.

***Theorem* 4.3.4. Geometric Criterion for Three Vectors**. The vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are linearly dependent if and only if the points O, A, B, C are coplanar (lie in the same plane).

Proof. Suppose $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are dependent. Then one is a combination of the others, say $\mathbf{c} = r\mathbf{a} + s\mathbf{b}$. If $\mathbf{a}$ and $\mathbf{b}$ are collinear with O, then the plane containing O, A, B is not unique, but $\mathbf{c}$ lies on the line generated by them, so all four points are collinear (and thus coplanar). If $\mathbf{a}$ and $\mathbf{b}$ are linearly independent, they span a unique plane passing through O. The vector $r\mathbf{a} + s\mathbf{b}$ lies strictly in this plane by the parallelogram law. Thus C lies in the plane defined by O, A, B.

Conversely, assume O, A, B, C are coplanar. If O, A, B are collinear, then $\mathbf{a}$ and $\mathbf{b}$ are dependent, satisfying the condition immediately (e.g., $\mathbf{a} = k\mathbf{b} + 0\mathbf{c}$). If O, A, B are not collinear, they form a triangle defining a unique plane. Since C lies in this plane, we can construct lines through C parallel to $\overrightarrow{OA}$ and $\overrightarrow{OB}$ (or extensions thereof). These lines intersect the axes generated by $\mathbf{a}$ and $\mathbf{b}$ at points A' and B' respectively. Then $\overrightarrow{OC} = \overrightarrow{OA'} + \overrightarrow{OB'}$. Since A' is on the line OA, $\overrightarrow{OA'} = r\mathbf{a}$. Since B' is on the line OB, $\overrightarrow{OB'} = s\mathbf{b}$. Thus $\mathbf{c} = r\mathbf{a} + s\mathbf{b}$, proving dependence. ■

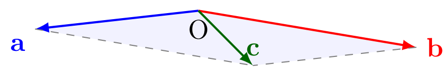


Figure 4.6: Linearly dependent vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are coplanar. Here $\mathbf{c} = \mathbf{a} + \mathbf{b}$.

We summarise the criteria for independence of three vectors.

***Theorem* 4.3.5. Independence of Three Vectors**. Let $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$. The following are equivalent:

(i) $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are linearly independent.
(ii) The points O, A, B, C are distinct and not coplanar (they form a tetrahedron).
(iii) If $r\mathbf{a} + s\mathbf{b} + t\mathbf{c} = \mathbf{0}$, then $r = s = t = 0$.

Remark. One may observe the absence of a component-based condition (iv) in the theorems for three vectors, analogous to the determinant condition $a_1b_2 - a_2b_1 = 0$ for the plane. A straightforward generalisation exists: $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are linearly dependent if and only if the determinant formed by their components is zero:

$$\det(\mathbf{a}, \mathbf{b}, \mathbf{c}) = a_1(b_2c_3 - b_3c_2) - a_2(b_1c_3 - b_3c_1) + a_3(b_1c_2 - b_2c_1) = 0.$$

While this is a necessary and sufficient condition, the rigorous development of determinant theory for 3×3 systems is reserved for the subsequent notes on Linear Algebra. For now, the geometric and algebraic definitions suffice.

General Linear Independence and Dimension

We conclude our investigation of linear independence by formalising the relationship between a set of vectors and its subsets. These results, while implicitly used in previous arguments, merit explicit statement as they apply to vector spaces of any dimension.

***Theorem* 4.3.6. Inheritance of Dependence**. Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be three vectors in $\mathbb{R}^3$. If any one of them is the zero vector, or if any pair of them is linearly dependent, then the set $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ is linearly dependent.

Proof. If $\mathbf{a} = \mathbf{0}$, then $1\mathbf{a} + 0\mathbf{b} + 0\mathbf{c} = \mathbf{0}$, satisfying the condition for dependence. If a pair, say $\mathbf{a}$ and $\mathbf{b}$, is linearly dependent, there exist scalars r, s not both zero such that $r\mathbf{a} + s\mathbf{b} = \mathbf{0}$. This implies $r\mathbf{a} + s\mathbf{b} + 0\mathbf{c} = \mathbf{0}$, satisfying the condition for the dependence of the three vectors. ■

Conversely, the property of linear independence is inherited by subsets.

***Theorem* 4.3.7. Inheritance of Independence**. Let $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ be a linearly independent set of vectors. Then they are distinct non-zero vectors, and any pair of them is linearly independent.

Proof. This is the contrapositive of the previous theorem. If the set is independent, no non-trivial linear combination yields zero. If one were zero, or a pair were dependent, a non-trivial combination would exist. ■

It is natural to extend our inquiry to sets of four or more vectors. We begin with the general definition.

Definition 4.3.4. ***General Linear Dependence***. Let $n > 1$ be a natural number. A set of n vectors is said to be linearly dependent if at least one of them can be written as a linear combination of the remaining $n - 1$ vectors. Otherwise, they are linearly independent.

While this definition allows for the concept of independence for any number of vectors, the structure of $\mathbb{R}^3$ imposes a strict limit. It turns out that any set of four vectors in space is necessarily linearly dependent.

***Theorem* 4.3.8. Dependence of Four Vectors**. Any four vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ in $\mathbb{R}^3$ are linearly dependent.

Proof. We seek scalars r, s, t, u, not all zero, such that:

$$r\mathbf{a} + s\mathbf{b} + t\mathbf{c} + u\mathbf{d} = \mathbf{0}.$$

Let $\mathbf{a} = [a_1, a_2, a_3]$, $\mathbf{b} = [b_1, b_2, b_3]$, etc. The vector equation corresponds to the following system of three simultaneous linear equations:

$$\begin{aligned} a_1r + b_1s + c_1t + d_1u &= 0 \\ a_2r + b_2s + c_2t + d_2u &= 0 \\ a_3r + b_3s + c_3t + d_3u &= 0 \end{aligned}$$

This is a homogeneous system of linear equations. A fundamental result of algebra states that a homogeneous system with more unknowns (here, 4) than equations (here, 3) always possesses a non-trivial solution. Thus, there exist scalars r, s, t, u, not all zero, satisfying the equation. Consequently, at least one vector is a linear combination of the others. ■

By an immediate generalisation of Theorem 4.3.6, if any four vectors are dependent, then any set of five or more vectors in $\mathbb{R}^3$ must also be dependent.

This observation reveals a fundamental characteristic of the vector space. We have seen that there exist three linearly independent vectors in $\mathbb{R}^3$ (for example, the standard basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$), but no set of four vectors can be independent. The number 3 represents a maximum threshold for independence in this space.

Definition 4.3.5. *Dimension*. The dimension of a vector space is the maximum number of linearly independent vectors that can exist within it.

- The vector space $\mathbb{R}^3$ has dimension 3.
- The vector space $\mathbb{R}^2$ has dimension 2.

This definition aligns perfectly with our geometric intuition of "three-dimensional space" and "two-dimensional plane," providing a rigorous algebraic foundation for these terms. In the subsequent study of linear algebra, we will see that this number also corresponds to the number of vectors required to form a *basis* — a set that generates the entire space.

4.4 Exercises

In the following exercises, vectors are elements of $\mathbb{R}^3$ unless otherwise stated.

Part I: Cartesian Coordinates and Vector Arithmetic

1. Let $\mathbf{a} = [2, 2, -1]$, $\mathbf{b} = [4, 2, 5]$, and $\mathbf{c} = [-1, -3, 2]$. Compute:

(a) $3(\mathbf{a} + \mathbf{b}) - 2(\mathbf{b} - \mathbf{c})$
(b) $2(\mathbf{a} + \mathbf{b} + \mathbf{c}) + 3(\mathbf{a} - \mathbf{b} - \mathbf{c})$
(c) The vector $\mathbf{x}$ such that $\mathbf{a} + 2\mathbf{x} = \mathbf{b} - \mathbf{c} - \mathbf{x}$.
(d) The vector $\mathbf{y}$ such that $2(\mathbf{a} - \mathbf{b} + \mathbf{y}) = 3(\mathbf{c} - \mathbf{y})$.

2. Find the values of a and b such that the vector $[a, -8, b]$ is a linear combination of $[6, 6, 3]$ and $[5, 4, 0]$.

3. Let $A(1, 2, 3)$, $B(4, 0, 5)$, and $C(3, 6, 4)$ be points in space.

(a) Determine whether the triangle ABC is isosceles, right-angled, or equilateral.

(b) Find the coordinates of the point D such that $ABCD$ forms a parallelogram.

4. Describe the geometric locus of points $P(x, y, z)$ satisfying the following equations:

(a) $x^2 + y^2 + z^2 - 4x + 6y = 12$
(b) $x^2 + y^2 + z^2 \leq 4z$
(c) $x^2 + y^2 = 9$ (in $\mathbb{R}^3$)

5. Describe the set of points (x, y, z) satisfying the following equations or inequalities.

(a) $x^2 + y^2 = 1$ (in $\mathbb{R}^3$).
(b) $z \geq \sqrt{x^2 + y^2}$.
(c) $xyz = 0$.

6. Given two non-zero vectors $\mathbf{a}$ and $\mathbf{b}$. By drawing suitable diagrams or using algebraic properties, find a necessary and sufficient condition for each of the following equalities to hold.

(a) $|\mathbf{a}| + |\mathbf{b}| = |\mathbf{a} + \mathbf{b}|$
(b) $|\mathbf{a}| + |\mathbf{b}| = |\mathbf{a} - \mathbf{b}|$
(c) $|\mathbf{a} + \mathbf{b}| = |\mathbf{a} - \mathbf{b}|$

7. Determine the equation of the sphere satisfying the given conditions.

(a) Centred at $(-1, 2, 4)$ with radius $\sqrt{7}$.
(b) Passing through the points $A(4, 0, 0)$, $B(0, 4, 0)$, and $C(0, 0, 4)$, with centre at the origin.
(c) Defined by the equation $x^2 + y^2 + z^2 - 4x + 6y - 2z + 5 = 0$. Find its centre and radius.

8. Locus Problems.

(a) Find the equation of the set of points equidistant from $A(1, 2, 3)$ and $B(-1, 4, 1)$. Describe this geometric object.
(b) Find the equation of the set of points whose distance from the origin is twice their distance from the point $(0, 0, 3)$.

9. Use the definition of vector subtraction $\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b})$ and the axioms of $\mathbb{R}^3$ to prove formally that:

(a) $\mathbf{a} - (\mathbf{b} + \mathbf{c}) = (\mathbf{a} - \mathbf{b}) - \mathbf{c}$.
(b) $r(\mathbf{a} - \mathbf{b}) = r\mathbf{a} - r\mathbf{b}$.

Part II: Linear Independence and Bases

10. Determine whether the following sets of vectors are linearly dependent or independent.

(a) $\mathbf{u} = [1, 2, 3]$, $\mathbf{v} = [0, 2, 1]$.
(b) $\mathbf{u} = [1, -1, 0]$, $\mathbf{v} = [0, 1, -1]$, $\mathbf{w} = [-1, 0, 1]$.
(c) $\mathbf{u} = [1, 1, 0]$, $\mathbf{v} = [1, 0, 1]$, $\mathbf{w} = [0, 1, 1]$.

11. Change of Basis. Given three linearly independent vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$, we define:

$$\mathbf{a} = 2\mathbf{x} + \mathbf{y}, \quad \mathbf{b} = \mathbf{y} - \mathbf{z}, \quad \mathbf{c} = \mathbf{x} + \mathbf{y} + \mathbf{z}.$$

(a) Prove that the set $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ is linearly independent.
(b) Express the vector $\mathbf{d} = \mathbf{x} + 2\mathbf{y} + \mathbf{z}$ as a linear combination of $\mathbf{a}, \mathbf{b}$, and $\mathbf{c}$.

12. Show that for any three vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$ in $\mathbb{R}^3$ and any scalars r, s, t, the three vectors:

$$\mathbf{u} = r\mathbf{x} - s\mathbf{y}, \quad \mathbf{v} = t\mathbf{y} - r\mathbf{z}, \quad \mathbf{w} = s\mathbf{z} - t\mathbf{x}$$

are linearly dependent.

Remark. Try to find a linear combination of $\mathbf{u}, \mathbf{v}, \mathbf{w}$ that equals zero without solving for x, y, z.

13. Prove or give a counter-example for each of the following geometric statements.

(a) If points O, A, B are collinear and points O, B, C are collinear, then O, A, C are collinear. (Assume $B \neq O$).
(b) If points O, A, B, C are coplanar and points O, C, D, E are coplanar, then O, A, B, D, E are coplanar. (Assume O, C define a unique line or plane as needed for your counter-example).

14. Determinant Condition. Let $\mathbf{a} = [a_1, a_2, a_3]$, $\mathbf{b} = [b_1, b_2, b_3]$, and $\mathbf{c} = [c_1, c_2, c_3]$. Prove that for these vectors to be linearly dependent, it is necessary that:

$$a_1(b_2c_3 - b_3c_2) - a_2(b_1c_3 - b_3c_1) + a_3(b_1c_2 - b_2c_1) = 0.$$

Remark. Assume dependence, i.e., $\mathbf{a} = r\mathbf{b} + s\mathbf{c}$. Substitute components into the expression and show it vanishes.

Part III: Advanced Vector Geometry

15. Generalised Centroid. Let A, B, C, D be four non-coplanar points in space.

(a) The centroid of the triangle BCD is $G_A = \frac{1}{3}(\mathbf{b} + \mathbf{c} + \mathbf{d})$. The line segment AG_A is a median of the tetrahedron $ABCD$.
(b) Show that the point P with position vector $\mathbf{p} = \frac{1}{4}(\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d})$ lies on the median AG_A.
(c) Deduce that the four medians of a tetrahedron are concurrent at the point P.

16. Coplanarity of Midpoints. Let $ABCD$ be a quadrilateral in space (the vertices need not be coplanar). Let P, Q, R, S be the midpoints of sides AB, BC, CD, DA respectively.

(a) Express the vector $\overrightarrow{PQ}$ in terms of $\mathbf{a}, \mathbf{b}, \mathbf{c}$.
(b) Express the vector $\overrightarrow{SR}$ in terms of $\mathbf{a}, \mathbf{d}, \mathbf{c}$.
(c) Prove that $\overrightarrow{PQ} = \overrightarrow{SR}$. What geometric shape is $PQRS$?
(d) Conclude that even for a skew quadrilateral, the midpoints always form a plane figure.

17. $\star$ **Regular Tetrahedrons.** A regular tetrahedron is a solid with four equilateral triangular faces. Let the vertices be at the origin O and points A, B, C.

(a) If the edge length is L, express the magnitudes $|\mathbf{a}|, |\mathbf{b}|, |\mathbf{c}|$ and the dot products $\mathbf{a} \cdot \mathbf{b}, \mathbf{b} \cdot \mathbf{c}, \mathbf{c} \cdot \mathbf{a}$.
(b) Use the algebraic criterion for independence to prove that $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are linearly independent.
(c) Find the angle between the edge OA and the median of the face OBC.

18. $\star$ **The Spherical Cosine Rule..** Let A, B, C be three points on the surface of a unit sphere centred at the origin. The "sides" of the spherical triangle ABC are the angles subtended at the centre: $a = \angle BOC$, $b = \angle AOC$, $c = \angle AOB$. By considering the vector identity $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{a} \times \mathbf{c})$ (or otherwise using dot products of normal vectors to planes OAB and OAC), one can derive the relationship for the angle α between the planes OAB and OAC (the spherical angle at A).

Remark. This question invites you to research or derive how the dot product structure on $\mathbb{R}^3$ induces geometry on the sphere S^2.

19. $\star$ **Gram-Schmidt Process (Introductory).** Given two linearly independent vectors $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^3$.

(a) Construct a vector $\mathbf{u}_1$ parallel to $\mathbf{a}$ with magnitude 1.
(b) Construct a vector $\mathbf{v}_2 = \mathbf{b} - \text{proj}_{\mathbf{a}}\mathbf{b}$. Prove $\mathbf{v}_2$ is orthogonal to $\mathbf{u}_1$.
(c) Normalise $\mathbf{v}_2$ to get $\mathbf{u}_2$.
(d) Given a third vector $\mathbf{c}$ independent of $\mathbf{a}$ and $\mathbf{b}$, suggest a formula to construct a vector $\mathbf{v}_3$ orthogonal to both $\mathbf{u}_1$ and $\mathbf{u}_2$.

20. $\star$ **Change of Basis (Standard).** Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be the standard basis. Let $\mathbf{v}_1 = \mathbf{e}_1 + \mathbf{e}_2$, $\mathbf{v}_2 = \mathbf{e}_2 + \mathbf{e}_3$, and $\mathbf{v}_3 = \mathbf{e}_3 + \mathbf{e}_1$.

(a) Show that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a linearly independent set.
(b) Express the vector $\mathbf{x} = [1, 1, 1]$ as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

Chapter 5

Solid Geometry

We now apply the algebraic framework of $\mathbb{R}^3$ to investigate the geometry of three-dimensional space. By exploiting the correspondence between points and their position vectors, we can extend the vector methods of plane geometry to prove theorems about spatial figures with elegance and efficiency.

5.1 Displacement Vectors in Space

Just as in the plane, the concept of displacement allows us to describe the relative positions of points without reference to the origin.

Definition 5.1.1. ***Displacement Vector***. Let A and B be points in space with position vectors $\mathbf{a}$ and $\mathbf{b}$. The displacement vector of the directed segment AB is defined as:

$$\overrightarrow{AB} = \mathbf{b} - \mathbf{a}.$$

In coordinates, if $A = (a_1, a_2, a_3)$ and $B = (b_1, b_2, b_3)$, then

$$\overrightarrow{AB} = [b_1 - a_1, b_2 - a_2, b_3 - a_3].$$

Many geometric results from the plane generalise directly to space because the algebraic rules governing vector addition and scalar multiplication are identical. For instance:

- **Midpoint Formula:** The midpoint M of a segment AB has position vector $\mathbf{m} = \frac{1}{2}(\mathbf{a} + \mathbf{b})$.
- **Collinearity:** Three points A, B, C are collinear if and only if $\overrightarrow{AC} = k\overrightarrow{AB}$ for some scalar k.
- **Parametric Line:** The line through points A and B consists of points X with position vectors $\mathbf{x} = (1 - t)\mathbf{a} + t\mathbf{b}$. The interpretation of the parameter t remains unchanged (e.g., $0 \leq t \leq 1$ defines the segment AB).

Since a triangle in space always lies in a single plane, properties involving only the triangle itself (such as the concurrency of medians at the centroid) hold true in space just as they do in plane geometry. The proofs are algebraically identical to those presented in Theorem 3.1.2 and subsequent examples.

5.2 The Tetrahedron

Moving beyond planar figures, the simplest truly three-dimensional object is the tetrahedron. Just as a triangle (a 2-simplex) is defined by three non-collinear points, a tetrahedron (a 3-simplex) is defined by four non-coplanar points.

Definition 5.2.1. ***Tetrahedron***. A tetrahedron is the geometric configuration formed by four non-coplanar points A, B, C, D, called the vertices. It consists of:

- **4 Vertices:** A, B, C, D.
- **4 Faces:** The triangles ABC, BCD, CDA, DAB.
- **6 Edges:** The segments AB, AC, AD, BC, CD, BD.

We define specific relationships between these elements:

- **Opposite Edges:** Two edges are opposite if they share no vertices (e.g., AB and CD).
- **Opposite Face:** A vertex and a face are opposite if the vertex is not contained in the face (e.g., vertex A and face BCD).

Analogous to the medians of a triangle, we can define special lines within a tetrahedron.

Definition 5.2.2. ***Medians and Bimedians*.**

- A **median** of a tetrahedron is a line segment connecting a vertex to the centroid of the opposite face.
- A **bimedian** of a tetrahedron is a line segment connecting the midpoints of two opposite edges.

There are four medians (one for each vertex) and three bimedians (one for each pair of opposite edges). Remarkably, all seven of these lines intersect at a single point.

***Theorem* 5.2.1. Centroid of a Tetrahedron.** Let $ABCD$ be a tetrahedron. The three bimedians and the four medians differ, but they all meet at a single point M, called the centroid of the tetrahedron.

Proof. Let the position vectors of the vertices be $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$. Consider the bimedian connecting the midpoint U of edge BC to the midpoint V of the opposite edge AD. The position vectors of these midpoints are:

$$\mathbf{u} = \frac{1}{2}(\mathbf{b} + \mathbf{c}), \quad \mathbf{v} = \frac{1}{2}(\mathbf{a} + \mathbf{d}).$$

The midpoint M of this bimedian UV is given by:

$$\mathbf{m} = \frac{1}{2}(\mathbf{u} + \mathbf{v}) = \frac{1}{2}\left(\frac{\mathbf{b} + \mathbf{c}}{2} + \frac{\mathbf{a} + \mathbf{d}}{2}\right) = \frac{1}{4}(\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d}).$$

Due to the symmetry of the expression $\frac{1}{4}(\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d})$ with respect to the four vertices, the same point M must be the midpoint of the other two bimedians (e.g., connecting the midpoints of AB and CD). Thus, the three bimedians are concurrent at M.

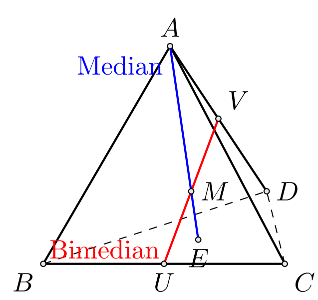


Figure 5.1: The intersection of a bimedian UV and a median AE at the centroid M.

Now consider the median from vertex A to the centroid E of the opposite face BCD. The centroid E has position vector:

$$\mathbf{e} = \frac{1}{3}(\mathbf{b} + \mathbf{c} + \mathbf{d}).$$

We observe that M lies on the segment AE. Specifically, we can express $\mathbf{m}$ as a linear combination of $\mathbf{a}$ and $\mathbf{e}$:

$$\mathbf{m} = \frac{1}{4}\mathbf{a} + \frac{3}{4}\left(\frac{\mathbf{b} + \mathbf{c} + \mathbf{d}}{3}\right) = \frac{1}{4}\mathbf{a} + \frac{3}{4}\mathbf{e}.$$

This equation shows that M lies on the line segment AE, dividing it in the ratio 3 : 1 (closer to the face). Since the expression for M is symmetric, M must essentially lie on all four medians. Therefore, the three bimedians and the four medians intersect at M. ■

Barycentric Coordinates and Ceva's Theorem

In previous chapters and sections, we established that a line through two points A and B is the locus of points whose position vectors are linear combinations of $\mathbf{a}$ and $\mathbf{b}$ where the coefficients sum to unity. We now extend this principle to define planes in space.

***Theorem* 5.2.2. Parametric Representation of a Plane**. Through three non-collinear points A, B, and C, there passes a unique plane E. A point X lies on E if and only if its position vector $\mathbf{x}$ satisfies

$$\mathbf{x} = r\mathbf{a} + s\mathbf{b} + t\mathbf{c} \quad \text{with} \quad r + s + t = 1.$$

Proof. A point X lies on the plane defined by A, B, and C if and only if the displacement vector $\overrightarrow{AX}$ can be expressed as a linear combination of the displacement vectors $\overrightarrow{AB}$ and $\overrightarrow{AC}$. Thus,

$$\mathbf{x} - \mathbf{a} = s(\mathbf{b} - \mathbf{a}) + t(\mathbf{c} - \mathbf{a})$$

Rearranging terms, we obtain:

$$\mathbf{x} = (1 - s - t)\mathbf{a} + s\mathbf{b} + t\mathbf{c}.$$

Letting $r = 1 - s - t$, we have $\mathbf{x} = r\mathbf{a} + s\mathbf{b} + t\mathbf{c}$ where $r + s + t = 1$. ■

The scalars (r, s, t) are known as the *barycentric coordinates* of X relative to the triangle ABC. They have a physical interpretation as the weights required at vertices A, B, and C such that X is the centre of gravity of the system.

This vector representation provides a powerful method for proving incidence theorems. We now prove a cornerstone of classical geometry regarding concurrent lines in a triangle.

***Theorem* 5.2.3. Ceva's Theorem**. Let ABC be a triangle and let F, G, H be points on the sides BC, CA, AB respectively. Define the directed ratios $\lambda = \operatorname{dr}(B, C; F)$, $\mu = \operatorname{dr}(C, A; G)$, and $\nu = \operatorname{dr}(A, B; H)$. The lines AF, BG, and CH are concurrent if and only if

$$\lambda\mu\nu = 1.$$

Proof. Let X be any point in the plane of $\triangle ABC$. Its position vector can be written as

$$\mathbf{x} = \frac{r\mathbf{a} + s\mathbf{b} + t\mathbf{c}}{r + s + t}.$$

Consider a point U on the segment BC with position vector $\mathbf{u} = \frac{s\mathbf{b}+t\mathbf{c}}{s+t}$. The vector $\mathbf{x}$ can be re-expressed as a combination of $\mathbf{a}$ and $\mathbf{u}$:

$$\mathbf{x} = \frac{r\mathbf{a} + (s + t)\mathbf{u}}{r + s + t},$$

implying that X lies on the line AU. To find the ratio in which U divides BC, we note that $(s+t)\mathbf{u} = s\mathbf{b}+t\mathbf{c}$, which yields $s(\mathbf{u} - \mathbf{b}) = t(\mathbf{c} - \mathbf{u})$, or $s\overrightarrow{BU} = t\overrightarrow{UC}$. Thus, $\operatorname{dr}(B, C; U) = t/s$.

If we set $U = F$, then $t/s = \lambda$. Similarly, points on the line BG satisfy $r/t = \mu$, and points on CH satisfy $s/r = \nu$.

Consider the point D with barycentric coordinates proportional to $(1, \nu, 1/\mu)$. For D to lie on AF, we require the ratio of its $\mathbf{b}$ and $\mathbf{c}$ components to be λ; specifically, $\nu/(1/\mu) = \lambda$, or $\lambda\mu\nu = 1$.

More formally, define the point D by the position vector:

$$\mathbf{d} = \frac{1\mathbf{a} + (1/\mu)\mathbf{b} + \lambda(1/\mu)\mathbf{c}}{1 + 1/\mu + \lambda/\mu}.$$

By construction, the ratio of the components of **b** and **c** is λ, so D lies on AF. The ratio of the components of **a** and **c** is $1/(\lambda/\mu) = \mu/\lambda$. If $\lambda\mu\nu = 1$, then $\mu/\lambda = \mu^2\nu$. Under the condition $\lambda\mu\nu = 1$, D lies on all three lines AF, BG, and CH. ∎

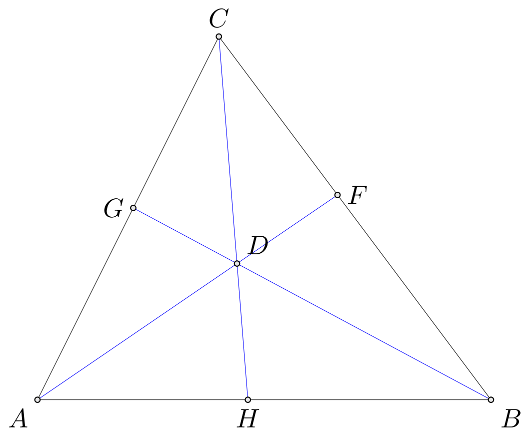


Figure 5.2: Concurrency of Cevians AF, BG, and CH at point D.

Remark. If $\lambda = \mu = \nu = 1$, the points F, G, H are the midpoints of the sides, and AF, BG, CH are the medians. In this case, $\lambda\mu\nu = 1 \cdot 1 \cdot 1 = 1$, verifying the concurrency of medians as a special case of Ceva's Theorem.

The Dot Product

In the vector space $\mathbb{R}^3$, the concepts of length and angle are rigorously defined through the dot product. This operation generalises the planar inner product by incorporating the third dimension.

Definition 5.2.3. ***Dot Product***. Let $\mathbf{a} = [a_1, a_2, a_3]$ and $\mathbf{b} = [b_1, b_2, b_3]$ be vectors in $\mathbb{R}^3$. The dot product of **a** and **b** is the scalar defined by:

$$\mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3.$$

This product shares the fundamental algebraic properties of its two-dimensional counterpart.

Theorem **5.2.4. Properties of the Dot Product**. For any vectors $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ and any scalar r, the following hold:

(i) $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$ (Symmetry).
(ii) $(r\mathbf{a}) \cdot \mathbf{b} = r(\mathbf{a} \cdot \mathbf{b})$ (Homogeneity).
(iii) $(\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c}$ (Distributivity).
(iv) $\mathbf{a} \cdot \mathbf{a} \geq 0$, with equality if and only if $\mathbf{a} = \mathbf{0}$ (Positive Definiteness).

Crucially, the dot product connects algebra to metric geometry via the magnitude. The square of the magnitude is identical to the self-dot product:

$$|\mathbf{a}|^2 = a_1^2 + a_2^2 + a_3^2 = \mathbf{a} \cdot \mathbf{a}.$$

Consequently, the Euclidean length of a segment OA is $|OA| = \sqrt{\mathbf{a} \cdot \mathbf{a}}$, and the distance between points A and B corresponds to the magnitude of the displacement vector: $|AB| = |\mathbf{b} - \mathbf{a}|$.

Metric Theorems in Space

The relationship between the dot product and magnitude yields a suite of inequalities and identities identical to those in the plane.

***Theorem* 5.2.5. Fundamental Metric Identities**. Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ and $r \in \mathbb{R}$. Then:

(i) $|r\mathbf{a}| = |r||\mathbf{a}|$.
(ii) $|\mathbf{a}+\mathbf{b}|^2 + |\mathbf{a}-\mathbf{b}|^2 = 2(|\mathbf{a}|^2 + |\mathbf{b}|^2)$ (Parallelogram Law).
(iii) $|\mathbf{a}\cdot\mathbf{b}| \leq |\mathbf{a}||\mathbf{b}|$ (Cauchy-Schwarz Inequality).
(iv) $|\mathbf{a}+\mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}|$ (Triangle Inequality).
(v) $|\mathbf{a}-\mathbf{b}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2(\mathbf{a}\cdot\mathbf{b})$ (Law of Cosines).

Proof. The proofs are algebraic and rely only on the axioms of the dot product, which are identical for $\mathbb{R}^2$ and $\mathbb{R}^3$. Thus, the proofs provided in Chapter 2 apply here without modification. ■

Angle and Orthogonality

The Law of Cosines identity allows us to define the angle between spatial vectors. For non-zero vectors $\mathbf{a}$ and $\mathbf{b}$, the quantity $\frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}||\mathbf{b}|}$ lies in the interval $[-1, 1]$ due to the Cauchy-Schwarz inequality.

Definition 5.2.4. *Angle Between Vectors*. Let $\mathbf{a}$ and $\mathbf{b}$ be non-zero vectors in $\mathbb{R}^3$. The angle θ between them is the unique value in $[0, \pi]$ such that

$$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}||\mathbf{b}|}.$$

Vectors $\mathbf{a}$ and $\mathbf{b}$ are orthogonal (or perpendicular) if $\mathbf{a}\cdot\mathbf{b} = 0$, denoted $\mathbf{a} \perp \mathbf{b}$.

This definition is consistent with the geometric angle $\angle AOB$ formed by the segments OA and OB in the unique plane containing the triangle OAB.

Orthogonal Projection

We can decompose any vector $\mathbf{a}$ into components along the coordinate axes using the dot product:

$$\mathbf{a} = (\mathbf{a}\cdot\mathbf{e}_1)\mathbf{e}_1 + (\mathbf{a}\cdot\mathbf{e}_2)\mathbf{e}_2 + (\mathbf{a}\cdot\mathbf{e}_3)\mathbf{e}_3.$$

The term $(\mathbf{a}\cdot\mathbf{e}_i)\mathbf{e}_i$ is the projection of $\mathbf{a}$ onto the i-th axis. More generally, we can project a vector $\mathbf{b}$ onto any non-zero vector $\mathbf{a}$.

***Theorem* 5.2.6. Orthogonal Projection**. Let $\mathbf{a} \in \mathbb{R}^3$ be a non-zero vector. For any vector $\mathbf{b}$, there exists a unique decomposition

$$\mathbf{b} = t\mathbf{a} + \mathbf{c}$$

where $\mathbf{c} \perp \mathbf{a}$. The scalar is $t = \frac{\mathbf{b}\cdot\mathbf{a}}{\mathbf{a}\cdot\mathbf{a}}$, and the vector $\text{proj}_{\mathbf{a}}\mathbf{b} = t\mathbf{a}$ is called the orthogonal projection of $\mathbf{b}$ onto $\mathbf{a}$.

Geometrically, $\mathbf{b}$, $\text{proj}_{\mathbf{a}}\mathbf{b}$, and $\mathbf{c}$ form a right-angled triangle in the plane spanned by $\mathbf{a}$ and $\mathbf{b}$. The length of the projection is $|\text{proj}_{\mathbf{a}}\mathbf{b}| = \frac{|\mathbf{b}\cdot\mathbf{a}|}{|\mathbf{a}|}$. Since the hypotenuse is longer than any leg, $|\mathbf{b}| \geq \frac{|\mathbf{b}\cdot\mathbf{a}|}{|\mathbf{a}|}$, which rearranges to the Cauchy-Schwarz inequality $|\mathbf{a}||\mathbf{b}| \geq |\mathbf{b}\cdot\mathbf{a}|$.

Equality in the metric inequalities corresponds to specific linear dependencies.

***Corollary* 5.2.1.** Equality Conditions

Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$.

(i) $|\mathbf{a}\cdot\mathbf{b}| = |\mathbf{a}||\mathbf{b}|$ if and only if $\mathbf{a}$ and $\mathbf{b}$ are linearly dependent (collinear).
(ii) $|\mathbf{a}+\mathbf{b}| = |\mathbf{a}| + |\mathbf{b}|$ if and only if one vector is a non-negative scalar multiple of the other (collinear and codirectional).

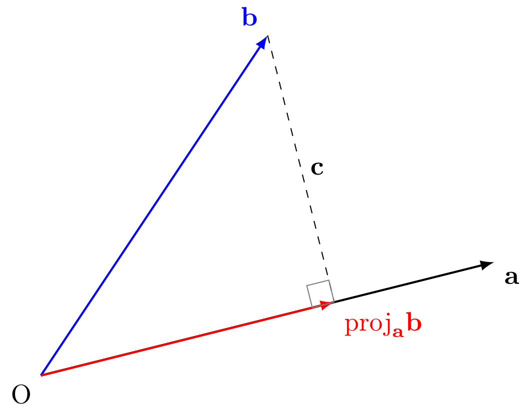


Figure 5.3: Orthogonal projection in space occurs in the plane defined by **a** and **b**.

Proof. For (ii): Squaring both sides of $|\mathbf{a}+\mathbf{b}| = |\mathbf{a}| + |\mathbf{b}|$ yields $|\mathbf{a}|^2 + |\mathbf{b}|^2 + 2(\mathbf{a}\cdot\mathbf{b}) = |\mathbf{a}|^2 + |\mathbf{b}|^2 + 2|\mathbf{a}||\mathbf{b}|$. This simplifies to $\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}|$. This equality requires both $|\mathbf{a}\cdot\mathbf{b}| = |\mathbf{a}||\mathbf{b}|$ (dependence) and $\mathbf{a}\cdot\mathbf{b} \geq 0$ (same direction). ∎

5.3 Planes in Space

In the plane, linear equations define lines. In space, they define planes. Consider the homogeneous linear equation:

$$a_1x + a_2y + a_3z = 0.$$

Using the dot product, this can be written as $\mathbf{a}\cdot\mathbf{x} = 0$, where $\mathbf{a} = [a_1, a_2, a_3]$ and $\mathbf{x} = [x, y, z]$. This equation describes the set of all position vectors $\mathbf{x}$ that are orthogonal to a fixed vector $\mathbf{a}$. Geometrically, this locus is a plane passing through the origin, with $\mathbf{a}$ acting as its normal vector.

To describe a general plane not necessarily passing through the origin, we consider the inhomogeneous equation:

$$a_1x + a_2y + a_3z + c = 0.$$

Let Q be a specific point on the plane with position vector $\mathbf{q}$. Then $\mathbf{a}\cdot\mathbf{q} + c = 0$, or $c = -\mathbf{a}\cdot\mathbf{q}$. Substituting this back into the general equation:

$$\mathbf{a}\cdot\mathbf{x} - \mathbf{a}\cdot\mathbf{q} = 0 \implies \mathbf{a}\cdot(\mathbf{x}-\mathbf{q}) = 0.$$

This form reveals the geometric structure: the plane consists of all points X such that the displacement vector from a fixed point Q to X is orthogonal to the normal vector $\mathbf{a}$.

***Theorem* 5.3.1. Equation of a Plane**. A linear equation $a_1x + a_2y + a_3z + c = 0$ defines a plane in space. The vector $\mathbf{a} = [a_1, a_2, a_3]$ is a normal vector to the plane.

Example 5.3.1. Intercept Form. Find the equation of the plane intersecting the axes at $A(a, 0, 0)$, $B(0, b, 0)$, and $C(0, 0, c)$. We seek an equation satisfied by all three points. The equation

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$$

clearly works (e.g., at A, $a/a + 0 + 0 = 1$). Multiplying by abc to remove fractions:

$$bcx + acy + abz - abc = 0.$$

By Theorem 5.3.1, a normal vector is $\mathbf{n} = [bc, ac, ab]$.

Distance from a Point to a Plane

The problem of finding the shortest distance from a point to a plane is perfectly analogous to the point-line distance in 2D.

***Theorem* 5.3.2. Distance to a Plane.** Let E be the plane defined by $\mathbf{a} \cdot \mathbf{x} + c = 0$. The perpendicular distance from a point P (position vector $\mathbf{p}$) to E is:

$$d(P, E) = \frac{|\mathbf{a} \cdot \mathbf{p} + c|}{|\mathbf{a}|}.$$

Proof. Let R be the foot of the perpendicular from P to E, and let Q be any arbitrary point on the plane. The distance $|RP|$ is the magnitude of the orthogonal projection of the displacement vector $\overrightarrow{QP}$ onto the normal vector $\mathbf{a}$.

$$\overrightarrow{RP} = \text{proj}_{\mathbf{a}}(\mathbf{p} - \mathbf{q}) = \frac{(\mathbf{p} - \mathbf{q}) \cdot \mathbf{a}}{\mathbf{a} \cdot \mathbf{a}}\mathbf{a}.$$

Taking magnitudes:

$$d = |\overrightarrow{RP}| = \frac{|(\mathbf{p} - \mathbf{q}) \cdot \mathbf{a}|}{|\mathbf{a}|} = \frac{|\mathbf{a} \cdot \mathbf{p} - \mathbf{a} \cdot \mathbf{q}|}{|\mathbf{a}|}.$$

Since Q lies on the plane, $\mathbf{a} \cdot \mathbf{q} = -c$. Substituting this yields the formula. ■

If we normalise the equation by dividing by $|\mathbf{a}|$, we obtain the *normal form*: $\hat{\mathbf{n}} \cdot \mathbf{x} + k = 0$, where $|\hat{\mathbf{n}}| = 1$. In this form, the distance is simply $|f(\mathbf{p})|$, where $f(\mathbf{x})$ is the linear polynomial defining the plane.

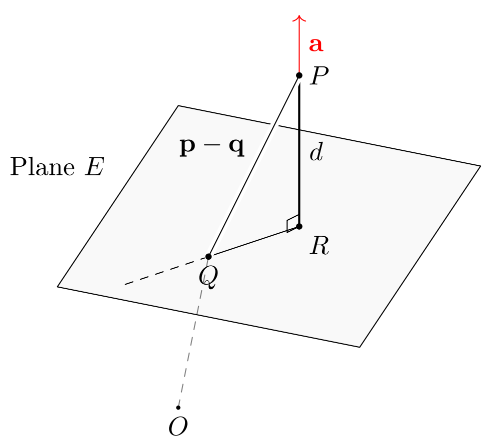


Figure 5.4: The distance from P to the plane is the projection of $\overrightarrow{QP}$ onto the normal $\mathbf{a}$.

Example 5.3.2. Ratio of Division. Let $f(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} + c = 0$ define a plane E. If the line segment connecting P and Q intersects E at R, the directed ratio $\lambda = \text{dr}(P, Q; R)$ is given by:

$$\lambda = -\frac{f(\mathbf{p})}{f(\mathbf{q})}.$$

Proof. Since R divides PQ in ratio λ, $\mathbf{r} = \frac{\mathbf{p} + \lambda \mathbf{q}}{1 + \lambda}$. Since R is on E, $f(\mathbf{r}) = 0$. By linearity of the dot product:

$$\mathbf{a} \cdot \left(\frac{\mathbf{p} + \lambda \mathbf{q}}{1 + \lambda} \right) + c = 0 \implies \mathbf{a} \cdot \mathbf{p} + c + \lambda(\mathbf{a} \cdot \mathbf{q} + c) = 0.$$

Thus $f(\mathbf{p}) + \lambda f(\mathbf{q}) = 0$, which yields the result. ■

5.4 Systems of Planes

We now examine the geometric relationships between multiple planes, specifically parallelism and intersection.

Parallel Planes

Let two planes E and F be defined by the equations:

$$f(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} + c = 0 \quad \text{and} \quad g(\mathbf{x}) = \mathbf{b} \cdot \mathbf{x} + d = 0.$$

The vectors **a** and **b** are normal to E and F respectively. The planes are parallel if and only if their normal vectors are collinear; that is, $\mathbf{b} = r\mathbf{a}$ for some non-zero scalar r. In this scenario, we can rewrite the equation for F as $\mathbf{a} \cdot \mathbf{x} + d/r = 0$ (provided we adjust the constant term). To simplify, assume both planes have the same normal vector **a**.

***Theorem* 5.4.1. Distance Between Parallel Planes**. The distance between two parallel planes $\mathbf{a} \cdot \mathbf{x} + c = 0$ and $\mathbf{a} \cdot \mathbf{x} + d = 0$ is given by

$$\text{dist}(E, F) = \frac{|c - d|}{|\mathbf{a}|}.$$

Proof. Pick an arbitrary point P on E. Then $\mathbf{a} \cdot \mathbf{p} + c = 0$, so $\mathbf{a} \cdot \mathbf{p} = -c$. The distance from P to the plane F is given by Theorem 5.3.2:

$$\text{dist}(P, F) = \frac{|\mathbf{a} \cdot \mathbf{p} + d|}{|\mathbf{a}|} = \frac{|-c + d|}{|\mathbf{a}|} = \frac{|c - d|}{|\mathbf{a}|}.$$ ■

Intersecting Planes and Sheaves

If two planes E and F are not parallel, they intersect in a straight line L. The angle θ between the planes is defined as the angle between their normal vectors **a** and **b**:

$$\cos\theta = \frac{|\mathbf{a} \cdot \mathbf{b}|}{|\mathbf{a}||\mathbf{b}|}.$$

The set of all planes passing through the intersection line L is called a *sheaf* (or pencil) of planes.

***Theorem* 5.4.2. Sheaf of Planes**. Let $f(\mathbf{x}) = 0$ and $g(\mathbf{x}) = 0$ define two intersecting planes. Any plane passing through their line of intersection L can be represented by the equation

$$rf(\mathbf{x}) + sg(\mathbf{x}) = 0,$$

where r and s are not both zero. Conversely, every such equation defines a plane in the sheaf.

Proof. Let H be a plane through L with normal **e**. Since **e** is orthogonal to the direction of L, and the normals $\mathbf{a}, \mathbf{b}$ span the plane orthogonal to L, **e** must be a linear combination $\mathbf{e} = r\mathbf{a} + s\mathbf{b}$. The constant term k in H's equation is determined by any point Q on L, leading to $h(\mathbf{x}) = rf(\mathbf{x}) + sg(\mathbf{x})$. ■

This algebraic structure leads to a profound projective invariant. First, recall the cross-ratio of four collinear points A_1, A_2, A_3, A_4:

$$\text{cr}(A_1, A_2; A_3, A_4) = \frac{\text{dr}(A_1, A_2; A_3)}{\text{dr}(A_1, A_2; A_4)}.$$

***Theorem* 5.4.3. Cross-Ratio Invariance**. Four planes of a sheaf intersect any transversal line in four points having a constant cross-ratio.

Proof. Let the planes be E_i defined by $f(\mathbf{x})+t_i g(\mathbf{x}) = 0$ for $i = 1, 2, 3, 4$. Let a transversal line intersect these planes at A_1, A_2, A_3, A_4. Using the ratio formula from the previous section, $\mathrm{dr}(A_1, A_2; A_3) = -\frac{f(\mathbf{a}_1)+t_3 g(\mathbf{a}_1)}{f(\mathbf{a}_2)+t_3 g(\mathbf{a}_2)}$. Since A_1 lies on E_1, $f(\mathbf{a}_1) = -t_1 g(\mathbf{a}_1)$. Similarly $f(\mathbf{a}_2) = -t_2 g(\mathbf{a}_2)$. Substituting these:

$$\mathrm{dr}(A_1, A_2; A_3) = -\frac{-t_1 g(\mathbf{a}_1) + t_3 g(\mathbf{a}_1)}{-t_2 g(\mathbf{a}_2) + t_3 g(\mathbf{a}_2)} = -\frac{g(\mathbf{a}_1)}{g(\mathbf{a}_2)}\frac{t_3 - t_1}{t_3 - t_2}.$$

The term involving A_4 is identical with t_4 replacing t_3. The ratio of the position-dependent terms $g(\mathbf{a}_1)/g(\mathbf{a}_2)$ cancels out in the cross-ratio:

$$\mathrm{cr}(A_1, A_2; A_3, A_4) = \frac{(t_3 - t_1)(t_4 - t_2)}{(t_3 - t_2)(t_4 - t_1)}.$$

This value depends solely on the parameters t_i defining the planes, not on the transversal line. ■

5.5 Exercises

Part I: Barycentric Coordinates and Incidence

1. Let A, B, C be points with position vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$. Suppose O, A, B are not collinear. Let $\mathbf{c} = s\mathbf{a} + t\mathbf{b}$. Prove that the points A, B, C are collinear if and only if $s + t = 1$.

2. Let $ABCD$ be a quadrilateral in space. Let E, F, G, H be points on the edges AB, AC, DB, DC respectively such that:

$$\frac{|AE|}{|AB|} = \frac{|AF|}{|AC|} = \frac{|DG|}{|DB|} = \frac{|DH|}{|DC|}.$$

Prove that the quadrilateral $EFGH$ is a parallelogram.

3. **Ceva's Theorem in Space.** Consider a tetrahedron $ABCD$. Let X be a point inside the tetrahedron. The lines connecting each vertex to X intersect the opposite face at points A', B', C', D'.

(a) Let the barycentric coordinates of X relative to A, B, C, D be (r, s, t, u) with $r + s + t + u = 1$. Show that the intersection D' of the line DX with the face ABC has barycentric coordinates proportional to (r, s, t) relative to that face.

(b) Let $\mathrm{Area}(PQR)$ denote the area of triangle PQR. Prove that the ratio of the volumes of tetrahedra $XBCD$ and $ABCD$ is equal to the barycentric coordinate r.

4. **Menelaus for Tetrahedra.** Let a plane intersect the edges AB, BC, CD, DA of a tetrahedron $ABCD$ at points P, Q, R, S. Define the directed ratios $r_1 = \mathrm{dr}(A, B; P)$, $r_2 = \mathrm{dr}(B, C; Q)$, $r_3 = \mathrm{dr}(C, D; R)$, and $r_4 = \mathrm{dr}(D, A; S)$. Prove that:

$$r_1 r_2 r_3 r_4 = 1.$$

5. Let A, B, C, D be four points in space. Let A' be the centroid of $\triangle BCD$, B' be the centroid of $\triangle ACD$, C' be the centroid of $\triangle ABD$, and D' be the centroid of $\triangle ABC$.

(a) Prove that the lines AA', BB', CC', DD' are concurrent at the centroid G of the tetrahedron $ABCD$.

(b) Prove that G is also the centroid of the tetrahedron $A'B'C'D'$.

Part II: Metric Properties and Products

6. Let $\mathbf{a} = [2, 3, 4]$ and $\mathbf{b} = [1, 1, 1]$.

(a) Find the cosine of the angle between $\mathbf{a}$ and $\mathbf{b}$.

(b) Find the orthogonal projection of $\mathbf{a}$ onto $\mathbf{b}$.

(c) Find a vector $\mathbf{c}$ such that $\mathbf{c}$ is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$.

7. Given vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ such that $|\mathbf{a}| = 1, |\mathbf{b}| = 2, |\mathbf{c}| = 3$. Suppose $\mathbf{a} \perp \mathbf{b}$, the angle between $\mathbf{a}$ and $\mathbf{c}$ is $\pi/3$, and the angle between $\mathbf{b}$ and $\mathbf{c}$ is $\pi/4$. Calculate $|\mathbf{a} + \mathbf{b} + \mathbf{c}|$.

8. Let $\mathbf{a} = [1, -1, 1]$ and $\mathbf{b} = [3, -4, 5]$. Let $\mathbf{x}(t) = \mathbf{a} + t\mathbf{b}$.

(a) Express $|\mathbf{x}(t)|^2$ as a quadratic in t.
(b) Find the value of t that minimises $|\mathbf{x}(t)|$.
(c) Verify that for this optimal t, the vector $\mathbf{x}(t)$ is orthogonal to $\mathbf{b}$.

9. Orthogonality of Diagonals. Prove that in a tetrahedron $ABCD$, if the opposite edges are perpendicular (i.e., $AB \perp CD$ and $AC \perp BD$), then the third pair of opposite edges is also perpendicular ($AD \perp BC$). Such a tetrahedron is called an *orthocentric tetrahedron.*

Remark. Use position vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$. The condition $AB \perp CD$ implies $(\mathbf{b} - \mathbf{a}) \cdot (\mathbf{d} - \mathbf{c}) = 0$.

10. Lagrange's Identity. Prove that for any vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$:

$$|\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2 = \sum_{1 \le i < j \le 3} (a_i b_j - a_j b_i)^2.$$

This quantity is the square of the area of the parallelogram spanned by $\mathbf{a}$ and $\mathbf{b}$.

Part III: Planes and Distance

11. Find the equation of the plane satisfying the following conditions:

(a) Passing through $(5, 1, 4)$ and parallel to the plane $x + y - 2z = 0$.
(b) Passing through $(2, 3, -1)$ and containing the line of intersection of the planes $x - y + z = 1$ and $x + y - z = 1$.

12. Let Π be the plane defined by $Ax + By + Cz + D = 0$. Let M_1 and M_2 be points not on the plane. If the line segment M_1M_2 intersects the plane at M such that $\overrightarrow{M_1M} = k\overrightarrow{MM_2}$, show that:

$$k = -\frac{Ax_1 + By_1 + Cz_1 + D}{Ax_2 + By_2 + Cz_2 + D}.$$

13. Distance to Intercept Plane. Suppose a plane cuts the coordinate axes at points $(a, 0, 0), (0, b, 0)$, and $(0, 0, c)$, where $a, b, c \neq 0$. Let p be the perpendicular distance from the origin to this plane. Prove that:

$$\frac{1}{p^2} = \frac{1}{a^2} + \frac{1}{b^2} + \frac{1}{c^2}.$$

14. Reflection in a Plane. Let Π be the plane $\mathbf{n} \cdot \mathbf{x} + d = 0$.

(a) Show that the reflection of a point P (position vector $\mathbf{p}$) in the plane Π is given by the vector

$$\mathbf{p}' = \mathbf{p} - 2\left(\frac{\mathbf{n} \cdot \mathbf{p} + d}{|\mathbf{n}|^2}\right)\mathbf{n}.$$

(b) Find the reflection of the point $(1, 2, 3)$ in the plane $2x - y + 2z - 6 = 0$.

15. Tetrahedron Plane Intersection. Given a tetrahedron $ABCD$ with centroid G. A plane passes through G and intersects the edges DA, DB, DC at points P, Q, R respectively. Prove that:

$$\frac{DA}{DP} + \frac{DB}{DQ} + \frac{DC}{DR} = 4.$$

Remark. Set the origin at D. Express $\mathbf{g}$ in terms of $\mathbf{a}, \mathbf{b}, \mathbf{c}$. Use the condition that P, Q, R, G are coplanar.

16. $\star$ **Regular Tetrahedra.** Let $ABCD$ be a regular tetrahedron (all edges have equal length L).

(a) Determine the angle θ between any two faces (the dihedral angle).
(b) Determine the angle ϕ between any two edges meeting at a vertex.
(c) Prove that the distance between the midpoints of opposite edges is $L/\sqrt{2}$.

Chapter 6

The Cross Product

In the previous chapter, we established that a plane in $\mathbb{R}^3$ is uniquely determined by a normal vector $\mathbf{n}$ and a point P. While the point is usually given, finding the normal vector often requires determining a direction orthogonal to two known non-parallel vectors lying within the plane. Unlike in $\mathbb{R}^2$, where a single vector has a unique orthogonal direction (up to scaling), in $\mathbb{R}^3$ a single vector is orthogonal to an entire plane of vectors. However, two non-collinear vectors $\mathbf{a}$ and $\mathbf{b}$ uniquely define a third direction orthogonal to both.

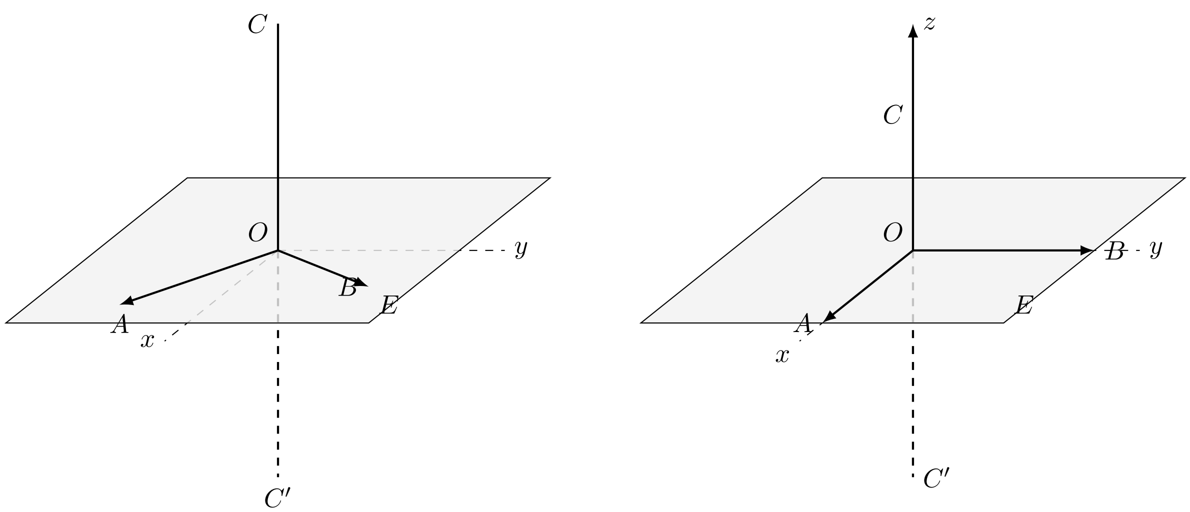


Figure 6.1: Vectors $\mathbf{a}$ and $\mathbf{b}$ in plane E define a unique normal line CC'.

In this chapter, we formalise this geometric necessity into an algebraic operation known as the cross product. Unlike the dot product, which yields a scalar, the cross product of two vectors in $\mathbb{R}^3$ yields a new vector.

6.1 Definition and Derivation

Let $\mathbf{a} = [a_1, a_2, a_3]$ and $\mathbf{b} = [b_1, b_2, b_3]$ be two linearly independent vectors in $\mathbb{R}^3$. We seek a non-zero vector $\mathbf{x} = [x, y, z]$ such that $\mathbf{x} \perp \mathbf{a}$ and $\mathbf{x} \perp \mathbf{b}$. This requirement translates to the system of homogeneous linear equations:

$$\begin{aligned} a_1x + a_2y + a_3z &= 0 \\ b_1x + b_2y + b_3z &= 0 \end{aligned}$$

To solve for the ratios of the components, we eliminate z. Multiplying the first equation by b_3 and the second by $-a_3$, then adding them, yields:

$$(a_1b_3 - a_3b_1)x + (a_2b_3 - a_3b_2)y = 0.$$

A natural solution (which preserves symmetry) is to set $x = a_2b_3 - a_3b_2$ and $y = a_3b_1 - a_1b_3$. Substituting these values back into the original equations allows us to solve for z, obtaining $z = a_1b_2 - a_2b_1$.

This specific solution $\mathbf{x}$ is the foundation of the cross product.

Definition 6.1.1. ***Cross Product***. Let $\mathbf{a} = [a_1, a_2, a_3]$ and $\mathbf{b} = [b_1, b_2, b_3]$ be vectors in $\mathbb{R}^3$. The cross product (or vector product) of $\mathbf{a}$ and $\mathbf{b}$, denoted $\mathbf{a} \times \mathbf{b}$, is the vector:

$$\mathbf{a} \times \mathbf{b} = [a_2b_3 - a_3b_2,\ a_3b_1 - a_1b_3,\ a_1b_2 - a_2b_1].$$

Note. The components correspond to the determinants of the 2×2 minors formed by the components of $\mathbf{a}$ and $\mathbf{b}$. The k-th component corresponds to a cyclic permutation of indices. If (i, j, k) is a cyclic permutation of $(1, 2, 3)$, the k-th component is $a_ib_j - a_jb_i$.

For the standard unit coordinate vectors, this definition yields:

$$\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3, \quad \mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1, \quad \mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2.$$

Algebraic Properties

The algebraic structure of the cross product differs significantly from standard multiplication; most notably, it is not commutative.

Theorem **6.1.1. Properties of the Cross Product**. Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be vectors in $\mathbb{R}^3$ and r be a scalar.

(i) $\mathbf{a} \cdot (\mathbf{a} \times \mathbf{b}) = 0$ and $\mathbf{b} \cdot (\mathbf{a} \times \mathbf{b}) = 0$ (Orthogonality).
(ii) $\mathbf{a} \times \mathbf{b} = -(\mathbf{b} \times \mathbf{a})$ (Anti-commutativity).
(iii) $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$ (Distributivity).
(iv) $(r\mathbf{a}) \times \mathbf{b} = r(\mathbf{a} \times \mathbf{b}) = \mathbf{a} \times (r\mathbf{b})$ (Homogeneity).

Proof. (i) We verify $\mathbf{a} \cdot (\mathbf{a} \times \mathbf{b}) = a_1(a_2b_3 - a_3b_2) + a_2(a_3b_1 - a_1b_3) + a_3(a_1b_2 - a_2b_1)$. Expanding this yields $a_1a_2b_3 - a_1a_3b_2 + a_2a_3b_1 - a_2a_1b_3 + a_3a_1b_2 - a_3a_2b_1$. All terms cancel pairwise, resulting in 0. The proof for $\mathbf{b}$ is identical.

(ii) Swapping $\mathbf{a}$ and $\mathbf{b}$ in the definition simply negates each component (e.g., $b_2a_3 - b_3a_2 = -(a_2b_3 - a_3b_2)$).

(iii) and (iv) follow directly from the linearity of real number arithmetic applied to the components. ∎

Since $\mathbf{a} \times \mathbf{a} = -(\mathbf{a} \times \mathbf{a})$, it follows that $\mathbf{a} \times \mathbf{a} = \mathbf{0}$. This observation leads to a criterion for linear dependence.

Theorem **6.1.2. Criterion for Dependence**. Let $\mathbf{a}$ and $\mathbf{b}$ be vectors in $\mathbb{R}^3$. Then $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ if and only if $\mathbf{a}$ and $\mathbf{b}$ are linearly dependent.

Proof. If $\mathbf{a}$ and $\mathbf{b}$ are dependent, one is a scalar multiple of the other, say $\mathbf{b} = r\mathbf{a}$. Then $\mathbf{a} \times \mathbf{b} = \mathbf{a} \times (r\mathbf{a}) = r(\mathbf{a} \times \mathbf{a}) = \mathbf{0}$. Conversely, if $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ and $\mathbf{a} \neq \mathbf{0}$, explicit expansion of the components shows that the ratios of components of $\mathbf{b}$ to $\mathbf{a}$ must be constant, implying dependence. ∎

Geometric Interpretation

The cross product $\mathbf{a} \times \mathbf{b}$ is a vector, and thus possesses both magnitude and direction. We analyze these separately.

We relate the magnitude of the cross product to the dot product via a fundamental algebraic identity.

***Theorem* 6.1.3. Lagrange's Identity**. For any vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$:

$$|\mathbf{a} \times \mathbf{b}|^2 = |\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2.$$

Proof. We expand the square of the magnitude of the cross product:

$$|\mathbf{a} \times \mathbf{b}|^2 = (a_2b_3 - a_3b_2)^2 + (a_3b_1 - a_1b_3)^2 + (a_1b_2 - a_2b_1)^2.$$

We also expand the right-hand side of the identity:

$$(a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - (a_1b_1 + a_2b_2 + a_3b_3)^2.$$

Upon full expansion of both expressions, the "mixed" squared terms (like $a_1^2b_1^2$) cancel out in the subtraction on the RHS, leaving only terms of the form $a_i^2b_j^2$ and $-2a_ib_ia_jb_j$. A careful term-by-term comparison reveals the expressions are identical. ■

Recall that $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta$, where θ is the angle between the vectors. Substituting this into Lagrange's Identity:

$$|\mathbf{a} \times \mathbf{b}|^2 = |\mathbf{a}|^2|\mathbf{b}|^2 - |\mathbf{a}|^2|\mathbf{b}|^2 \cos^2\theta = |\mathbf{a}|^2|\mathbf{b}|^2(1 - \cos^2\theta) = |\mathbf{a}|^2|\mathbf{b}|^2 \sin^2\theta.$$

Taking the square root (since $\sin\theta \geq 0$ for $\theta \in [0, \pi]$):

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}| \sin\theta.$$

Geometrically, $|\mathbf{a}||\mathbf{b}| \sin\theta$ represents the area of the parallelogram spanned by vectors $\mathbf{a}$ and $\mathbf{b}$.

***Theorem* 6.1.4. Area of a Triangle**. Let A, B, C be points in space. The area of the triangle ABC is given by:

$$\text{Area} = \frac{1}{2}|\overrightarrow{AB} \times \overrightarrow{AC}|.$$

Direction and Planes

Since $\mathbf{a} \times \mathbf{b}$ is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$, it is orthogonal to the plane containing them. The orientation follows the *right-hand rule*: if the fingers of the right hand curl from $\mathbf{a}$ to $\mathbf{b}$, the thumb points in the direction of $\mathbf{a} \times \mathbf{b}$. This confirms that $\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}$ form a right-handed system, analogous to the standard basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$.

This geometric property provides the most efficient method for determining the equation of a plane through three points.

Example 6.1.1. Plane Through Three Points. Find the equation of the plane passing through $A(1, -1, 0)$, $B(2, 3, 1)$, and $C(0, 1, -1)$. We first compute two displacement vectors lying in the plane:

$$\mathbf{u} = \overrightarrow{AB} = [1, 4, 1]$$

$$\mathbf{v} = \overrightarrow{AC} = [-1, 2, -1]$$

A normal vector $\mathbf{n}$ to the plane is given by the cross product $\mathbf{u} \times \mathbf{v}$:

$$\begin{aligned} \mathbf{n} &= [(4)(-1) - (1)(2),\ (1)(-1) - (1)(-1),\ (1)(2) - (4)(-1)] \\ &= [-4 - 2,\ -1 + 1,\ 2 + 4] \\ &= [-6, 0, 6]. \end{aligned}$$

For simplicity, we can use the parallel vector $\mathbf{n}' = [-1, 0, 1]$. The equation of the plane is $\mathbf{n}' \cdot (\mathbf{x} - \mathbf{a}) = 0$:

$$-1(x - 1) + 0(y + 1) + 1(z - 0) = 0 \implies -x + 1 + z = 0 \implies z - x + 1 = 0.$$

6.2 Triple Products

Given three vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$, we can form products involving all three. The position of the brackets and the choice of operators determines the nature of the result.

The Scalar Triple Product

The product $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ results in a scalar.

***Theorem* 6.2.1. Volume of a Parallelepiped.** The absolute value of the scalar triple product $|(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}|$ is equal to the volume of the parallelepiped determined by the edges $\mathbf{a}, \mathbf{b}, \mathbf{c}$.

Proof. The volume of a parallelepiped is the area of the base times the perpendicular height. Let the base be spanned by $\mathbf{a}$ and $\mathbf{b}$.

$$\text{Area} = |\mathbf{a} \times \mathbf{b}|.$$

The height h is the length of the orthogonal projection of $\mathbf{c}$ onto the normal vector of the base, which is $\mathbf{n} = \mathbf{a} \times \mathbf{b}$.

$$h = |\text{proj}_{\mathbf{n}} \mathbf{c}| = \frac{|\mathbf{c} \cdot \mathbf{n}|}{|\mathbf{n}|} = \frac{|\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})|}{|\mathbf{a} \times \mathbf{b}|}.$$

Thus, Volume = Area $\times h = |(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}|$. ∎

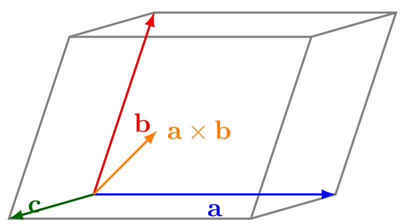


Figure 6.2: The volume is determined by the base area $|\mathbf{a} \times \mathbf{b}|$ and the projection of $\mathbf{c}$ onto the normal.

The scalar triple product satisfies a cyclic invariance property.

***Theorem* 6.2.2. Cyclic Permutation.** Let $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$. Then:

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = (\mathbf{b} \times \mathbf{c}) \cdot \mathbf{a} = (\mathbf{c} \times \mathbf{a}) \cdot \mathbf{b}.$$

It follows that the dot and cross operations may be interchanged:

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}).$$

The Vector Triple Product

The product $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ yields a vector. Note that the cross product is not associative: $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$. For instance, $\mathbf{e}_2 \times (\mathbf{e}_1 \times \mathbf{e}_2) = \mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1$, whereas $(\mathbf{e}_2 \times \mathbf{e}_1) \times \mathbf{e}_2 = (-\mathbf{e}_3) \times \mathbf{e}_2 = -(-\mathbf{e}_1) = \mathbf{e}_1$. However, $\mathbf{e}_1 \times (\mathbf{e}_1 \times \mathbf{e}_2) = -\mathbf{e}_2$ while $(\mathbf{e}_1 \times \mathbf{e}_1) \times \mathbf{e}_2 = \mathbf{0}$.

The vector $\mathbf{v} = \mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ is orthogonal to $\mathbf{b} \times \mathbf{c}$. Since $\mathbf{b} \times \mathbf{c}$ is the normal to the plane spanned by $\mathbf{b}$ and $\mathbf{c}$, the vector $\mathbf{v}$ must lie *in* the plane of $\mathbf{b}$ and $\mathbf{c}$. Thus, $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ is a linear combination of $\mathbf{b}$ and $\mathbf{c}$.

***Theorem* 6.2.3. Vector Triple Product Expansion.** Let $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$. Then:

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}.$$

Proof. We verify the first component. The x-component of $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ is:

$$a_2(\mathbf{b} \times \mathbf{c})_3 - a_3(\mathbf{b} \times \mathbf{c})_2 = a_2(b_1c_2 - b_2c_1) - a_3(b_3c_1 - b_1c_3).$$

Rearranging terms to group by b_1 and c_1:

$$= b_1(a_2c_2 + a_3c_3) - c_1(a_2b_2 + a_3b_3).$$

We add and subtract $a_1b_1c_1$ to complete the dot products:

$$= b_1(a_1c_1 + a_2c_2 + a_3c_3) - c_1(a_1b_1 + a_2b_2 + a_3b_3).$$

$$= b_1(\mathbf{a} \cdot \mathbf{c}) - c_1(\mathbf{a} \cdot \mathbf{b}).$$

This is precisely the first component of $(\mathbf{a}\cdot\mathbf{c})\mathbf{b}-(\mathbf{a}\cdot\mathbf{b})\mathbf{c}$. Symmetry arguments apply to the other components. ■

This identity is often remembered by the mnemonic "BAC - CAB". Similarly, $(\mathbf{a}\times\mathbf{b})\times\mathbf{c} = -(\mathbf{c}\times(\mathbf{a}\times\mathbf{b})) = -((\mathbf{c}\cdot\mathbf{b})\mathbf{a} - (\mathbf{c}\cdot\mathbf{a})\mathbf{b}) = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{b}\cdot\mathbf{c})\mathbf{a}$.

6.3 Lines in Space

The extension of the theory of straight lines from the plane to three-dimensional space is natural and intuitive. A line in space is fundamentally a one-dimensional object, and its algebraic description mirrors that of the plane line equation, albeit with an additional coordinate.

Parametric Representations

Recall that in the plane, a line is determined by a point and a direction. This geometric definition remains valid in $\mathbb{R}^3$.

Definition 6.3.1. ***Point-Direction Form***. Let P be a fixed point with position vector $\mathbf{p}$, and let $\mathbf{d}$ be a non-zero vector. The straight line L passing through P and parallel to $\mathbf{d}$ consists of all points X whose position vectors $\mathbf{x}$ satisfy:

$$\mathbf{x} = \mathbf{p} + t\mathbf{d},$$

where $t \in \mathbb{R}$ is a scalar parameter.

The vector $\mathbf{d}$ is called a *direction vector* of the line. Just as in the plane, this representation is not unique; replacing $\mathbf{d}$ with any non-zero scalar multiple $k\mathbf{d}$ describes the same line.

Example 6.3.1. Line and Plane Intersection. Let L be the line passing through $P(0, 1, 1)$ parallel to the vector $\mathbf{d} = [1, -1, 2]$. The parametric equation of the line is:

$$\mathbf{x} = [0, 1, 1] + t[1, -1, 2] = [t,\ 1 - t,\ 1 + 2t].$$

To find the intersection of this line with the plane defined by $4x+7y-z-1 = 0$, we substitute the components of $\mathbf{x}$ into the plane equation:

$$4(t) + 7(1 - t) - (1 + 2t) - 1 = 0$$

$$4t + 7 - 7t - 1 - 2t - 1 = 0$$

$$-5t + 5 = 0 \implies t = 1.$$

Substituting $t = 1$ back into the line equation yields the point of intersection:

$$\mathbf{x} = [1,\ 1 - 1,\ 1 + 2] = [1, 0, 3].$$

Since two distinct points determine a unique line, we may also define a line using two position vectors.

***Theorem* 6.3.1. Two-Point Form**. Let A and B be distinct points with position vectors $\mathbf{a}$ and $\mathbf{b}$. The line passing through A and B is given by:

$$\mathbf{x} = \mathbf{a} + t(\mathbf{b} - \mathbf{a}) \quad \text{or} \quad \mathbf{x} = (1-t)\mathbf{a} + t\mathbf{b}.$$

This is identical to the affine combination derived in the plane geometry chapter.

Implicit Representation and The Two-Plane Form

Unlike in the plane, a single linear equation $ax + by + cz + d = 0$ does not represent a line in $\mathbb{R}^3$; it represents a plane. However, the intersection of two non-parallel planes defines a line.

Definition 6.3.2. *Two-Plane Form*. A line in space may be defined as the solution set of a system of two linear equations:

$$a_1x + a_2y + a_3z + c = 0$$
$$b_1x + b_2y + b_3z + d = 0$$

provided that the normal vectors $\mathbf{a} = [a_1, a_2, a_3]$ and $\mathbf{b} = [b_1, b_2, b_3]$ are linearly independent.

Geometrically, the direction vector of this line must be orthogonal to both normal vectors $\mathbf{a}$ and $\mathbf{b}$. Therefore, the direction of the line is given by the cross product $\mathbf{a} \times \mathbf{b}$.

Example 6.3.2. Converting Forms. Consider the line H defined by the intersection of two planes:

$$2x + y - z - 1 = 0$$
$$-x + 3y + z - 8 = 0$$

The normals are $\mathbf{n}_1 = [2, 1, -1]$ and $\mathbf{n}_2 = [-1, 3, 1]$. The direction vector of the line is:

$$\mathbf{d} = \mathbf{n}_1 \times \mathbf{n}_2 = [1(1) - (-1)(3),\ (-1)(-1) - 2(1),\ 2(3) - 1(-1)] = [4, -1, 7].$$

To construct the parametric form, we need a single point on the line. We can set one coordinate to a convenient value, say $z = 3$ (as suggested by an inspection or by solving the system). Substitute $z = 3$ into the system:

$$2x + y = 4 \quad \text{and} \quad -x + 3y = 5.$$

Multiplying the second by 2: $-2x + 6y = 10$. Adding to the first: $7y = 14 \implies y = 2$. Then $2x + 2 = 4 \implies x = 1$. Thus, the point $P(1, 2, 3)$ lies on H. The parametric equation is:

$$\mathbf{x} = [1, 2, 3] + t[4, -1, 7] = [1 + 4t,\ 2 - t,\ 3 + 7t].$$

Symmetric Form and Direction Cosines

The direction of a line is often normalized. If $\mathbf{d}$ is a direction vector, the unit vector $\hat{\mathbf{d}} = \mathbf{d}/|\mathbf{d}| = [d_1, d_2, d_3]$ satisfies $d_1^2 + d_2^2 + d_3^2 = 1$. The components d_1, d_2, d_3 are the cosines of the angles between the line and the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. These are termed the *direction cosines* of the line.

If none of the components of the direction vector are zero, we can eliminate the parameter t from the parametric equations $x_i = p_i + td_i$ by writing $t = (x_i - p_i)/d_i$. Equating these expressions yields the Cartesian or symmetric form.

***Theorem* 6.3.2. Symmetric Equations**. Let a line pass through $P(p_1, p_2, p_3)$ with direction vector $\mathbf{d} = [d_1, d_2, d_3]$ where $d_i \neq 0$. The points on the line satisfy:

$$\frac{x - p_1}{d_1} = \frac{y - p_2}{d_2} = \frac{z - p_3}{d_3}.$$

Remark. These equations essentially state that the displacement vector $\mathbf{x} - \mathbf{p}$ is proportional to $\mathbf{d}$. If one component of $\mathbf{d}$ is zero, say $d_3 = 0$, the line is parallel to the xy-plane. The symmetric form becomes $\frac{x-p_1}{d_1} = \frac{y-p_2}{d_2}$, $z = p_3$.

Relative Positions and Distances

In the plane, two lines are either parallel or intersecting. In space, a third possibility arises: lines may be non-parallel yet never intersect. Such lines are called *skew lines*.

Consider two lines L and G defined by:

$$L : \mathbf{x} = \mathbf{p} + t\mathbf{d} \quad \text{and} \quad G : \mathbf{x} = \mathbf{q} + s\mathbf{f}.$$

We classify their relative positions based on the linear independence of their defining vectors.

***Theorem* 6.3.3. Relative Positions of Two Lines**. Let L and G be lines with direction vectors $\mathbf{d}$ and $\mathbf{f}$, passing through P and Q respectively.

(i) **Identical:** $\mathbf{d}$ and $\mathbf{f}$ are parallel (linearly dependent), and the displacement $\overrightarrow{PQ} = \mathbf{q} - \mathbf{p}$ is parallel to $\mathbf{d}$.
(ii) **Parallel (distinct):** $\mathbf{d}$ and $\mathbf{f}$ are parallel, but $\overrightarrow{PQ}$ is not parallel to $\mathbf{d}$.
(iii) **Intersecting:** $\mathbf{d}$ and $\mathbf{f}$ are linearly independent, and $\overrightarrow{PQ}$ lies in the plane spanned by $\mathbf{d}$ and $\mathbf{f}$. Equivalently, the scalar triple product vanishes: $(\mathbf{q} - \mathbf{p}) \cdot (\mathbf{d} \times \mathbf{f}) = 0$.
(iv) **Skew:** $\mathbf{d}$, $\mathbf{f}$, and $\overrightarrow{PQ}$ are linearly independent. Equivalently, $(\mathbf{q} - \mathbf{p}) \cdot (\mathbf{d} \times \mathbf{f}) \neq 0$.

The distance between two lines is defined as the minimum distance between any pair of points, one from each line.

***Theorem* 6.3.4. Distance Between Lines**.

(i) If lines L and G are **parallel** (with common direction $\mathbf{d}$), the distance is the magnitude of the component of $\overrightarrow{PQ}$ orthogonal to $\mathbf{d}$:

$$\operatorname{dist}(L, G) = \frac{|(\mathbf{q} - \mathbf{p}) \times \mathbf{d}|}{|\mathbf{d}|}.$$

(ii) If lines L and G are **skew** (directions $\mathbf{d}, \mathbf{f}$), the distance is the magnitude of the projection of $\overrightarrow{PQ}$ onto the common normal $\mathbf{n} = \mathbf{d} \times \mathbf{f}$:

$$\operatorname{dist}(L, G) = \frac{|(\mathbf{q} - \mathbf{p}) \cdot (\mathbf{d} \times \mathbf{f})|}{|\mathbf{d} \times \mathbf{f}|}.$$

Proof. For parallel lines, the area of the parallelogram spanned by $\overrightarrow{PQ}$ and $\mathbf{d}$ is $|(\mathbf{q} - \mathbf{p}) \times \mathbf{d}|$. Dividing this area by the base length $|\mathbf{d}|$ gives the height, which is the perpendicular distance.

For skew lines, the volume of the parallelepiped spanned by $\overrightarrow{PQ}$, $\mathbf{d}$, and $\mathbf{f}$ is $|(\mathbf{q} - \mathbf{p}) \cdot (\mathbf{d} \times \mathbf{f})|$. The area of the base spanned by $\mathbf{d}$ and $\mathbf{f}$ is $|\mathbf{d} \times \mathbf{f}|$. Since Volume = Area × Height, the height (distance between the planes containing the lines) is given by the ratio of the scalar triple product to the magnitude of the cross product. ∎

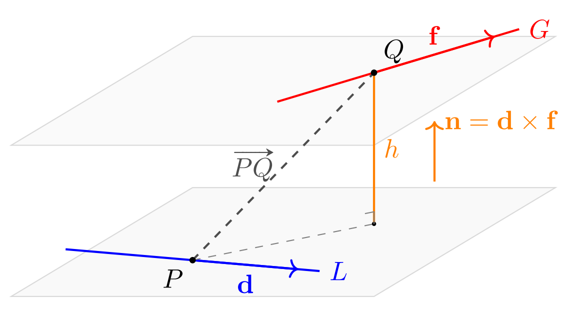


Figure 6.3: The distance between skew lines is the projection of the connector $\overrightarrow{PQ}$ onto the direction orthogonal to both lines.

6.4 Curvilinear Coordinates

While the Cartesian coordinate system (x, y, z) is universal, it is often advantageous to employ coordinate systems that exploit the symmetries of a geometric configuration. For planar geometry, polar coordinates (r, θ) simplify problems involving circular symmetry. In this section, we extend this concept to three dimensions, introducing cylindrical and spherical coordinates.

Cylindrical Coordinates

The cylindrical coordinate system is a direct extension of plane polar coordinates into three-dimensional space by augmenting the polar pair (r, θ) with the Cartesian height z.

Let X be a point in space with Cartesian coordinates (x, y, z). We project X orthogonally onto the xy-plane to obtain the point $X' = (x, y, 0)$. The position of X' can be described by polar coordinates (r, θ), where r is the distance from the origin O to X', and θ is the angle measured counter-clockwise from the positive x-axis to the ray OX'.

Definition 6.4.1. ***Cylindrical Coordinates*.** The cylindrical coordinates of a point X are the triple (r, θ, h), where:

- $r \geq 0$ is the radial distance in the xy-plane: $r = \sqrt{x^2 + y^2}$.
- $\theta \in [0, 2\pi)$ is the azimuthal angle in the xy-plane.
- h is the signed height along the z-axis: $h = z$.

The transformation from cylindrical coordinates to Cartesian coordinates is given by the identities:

$$x = r\cos\theta, \quad y = r\sin\theta, \quad z = h.$$

Conversely, the cylindrical coordinates are derived from Cartesian coordinates by:

$$r = \sqrt{x^2 + y^2}, \quad \tan\theta = \frac{y}{x}, \quad h = z.$$

Geometrically, the set of points with a constant coordinate $r = c$ forms a circular cylinder of radius c centered on the z-axis. This property gives the system its name.

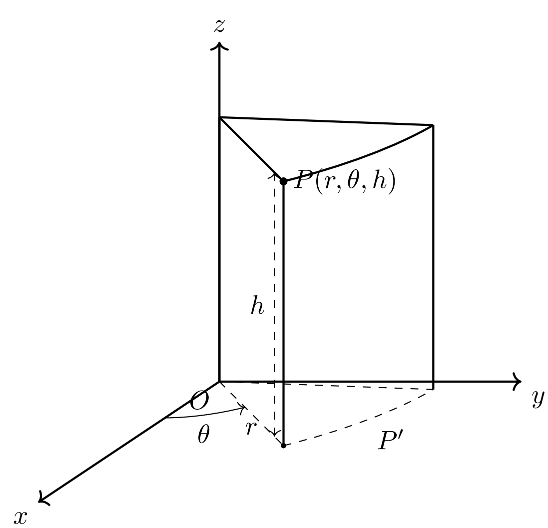


Figure 6.4: Cylindrical coordinates (r, θ, h). The point X lies on a cylinder of radius r.

Spherical Coordinates

A second system, spherical coordinates, is defined by the distance from the origin and two angles. This system is particularly useful for problems possessing spherical symmetry.

Let X be a point in space with Cartesian coordinates (x, y, z). Its distance from the origin is given by $r = \sqrt{x^2 + y^2 + z^2}$, so X lies on a sphere of radius r centred at O. To specify the position of X on this sphere, we use the longitude and the colatitude.

Let X' be the orthogonal projection of X onto the xy-plane. As with cylindrical coordinates, θ denotes the angle of rotation in the xy-plane from the positive x-axis to OX'. This angle is the *longitude*. Furthermore, let φ denote the angle between the positive z-axis and the ray OX. This angle is the *colatitude* (the complement of the latitude).

Definition 6.4.2. ***Spherical Coordinates***. The spherical coordinates of a point X are the triple (r, φ, θ), where:

- $r \geq 0$ is the distance from the origin: $r = \sqrt{x^2 + y^2 + z^2}$.
- $\varphi \in [0, \pi]$ is the polar angle (colatitude) measured from the positive z-axis.
- $\theta \in [0, 2\pi)$ is the azimuthal angle (longitude) in the xy-plane.

To derive the relationship with Cartesian coordinates, observe that the projection of the vector $\overrightarrow{OX}$ onto the z-axis has length $r \cos \varphi$, and its projection onto the xy-plane (the length $|OX'|$) is $r \sin \varphi$. The Cartesian coordinates are thus:

$$x = (r \sin \varphi) \cos \theta, \quad y = (r \sin \varphi) \sin \theta, \quad z = r \cos \varphi.$$

Simplifying, we obtain the transformation identities:

$$x = r \cos \theta \sin \varphi, \quad y = r \sin \theta \sin \varphi, \quad z = r \cos \varphi.$$

Remark. One must be cautious when consulting other literature, particularly in physics, where the symbols θ and φ are often swapped (using θ for the polar angle and ϕ for the azimuth). The convention adopted here aligns with standard mathematical texts.

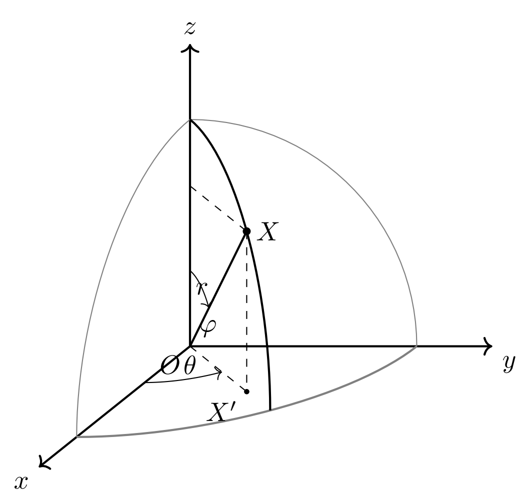


Figure 6.5: Spherical coordinates (r, φ, θ). The meridian arc passes through X and meets the equator at X'.

The coordinate surfaces for spherical coordinates are:

- Spheres ($r =$ constant).
- Cones centred on the z-axis ($\varphi =$ constant).
- Half-planes containing the z-axis ($\theta =$ constant).

6.5 Exercises

Part I: Cross Product Algebra and Geometry

1. Let $\mathbf{a} = [1, 2, -3]$, $\mathbf{b} = [1, 0, -1]$, and $\mathbf{c} = [3, 1, 0]$. Compute:

(a) $\mathbf{a} \times \mathbf{b}$ and $\mathbf{b} \times \mathbf{c}$.
(b) The area of the triangle with edges defined by vectors $\mathbf{a}$ and $\mathbf{b}$.
(c) The scalar triple product $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$.

2. Let $\mathbf{a}$ and $\mathbf{b}$ be vectors such that $|\mathbf{a}| = 2$, $|\mathbf{b}| = 5$, and $\mathbf{a} \cdot \mathbf{b} = -6$.

(a) Use Lagrange's Identity to compute $|\mathbf{a} \times \mathbf{b}|$.
(b) Verify your result by finding the angle θ between $\mathbf{a}$ and $\mathbf{b}$.

3. Find the equation of the plane passing through the three points $A(1, 2, 0)$, $B(3, 0, -3)$, and $C(5, 2, 1)$.

4. **Vector Identity Proofs.** Let $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ be vectors in $\mathbb{R}^3$. Prove the following identities:

(a) $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c})$.
(b) $(\mathbf{a} + \mathbf{b}) \cdot [(\mathbf{b} + \mathbf{c}) \times (\mathbf{c} + \mathbf{a})] = 2\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$.

5. **Heron's Formula.** Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be the vector sides of a triangle such that $\mathbf{a} + \mathbf{b} + \mathbf{c} = \mathbf{0}$. Let $a = |\mathbf{a}|$, $b = |\mathbf{b}|$, $c = |\mathbf{c}|$, and $s = (a + b + c)/2$.

(a) Show that $\mathbf{a} \cdot \mathbf{b} = \frac{1}{2}(c^2 - a^2 - b^2)$.
(b) Use the identity $|\mathbf{a} \times \mathbf{b}|^2 = |\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2$ to derive Heron's formula for the area of the triangle:

$$\text{Area} = \sqrt{s(s-a)(s-b)(s-c)}.$$

6. **Solving Vector Equations.** Given a scalar m and vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ such that $\mathbf{a} \cdot \mathbf{b} \neq 0$ and $\mathbf{b} \cdot \mathbf{c} = 0$, find the unique vector $\mathbf{x}$ satisfying the system:

$$\mathbf{x} \cdot \mathbf{a} = m \quad \text{and} \quad \mathbf{x} \times \mathbf{b} = \mathbf{c}.$$

Remark. Consider the expansion of $\mathbf{a} \times (\mathbf{x} \times \mathbf{b})$ or similar constructions to isolate $\mathbf{x}$.

Part II: Lines and Distances

7. Find the parametric and symmetric equations of the line L passing through $P(2, 3, -1)$ and $Q(1, -5, 1)$. Determine the coordinates of the points where L intersects the three coordinate planes.

8. Find the equation of the line passing through $A(0, 4, 1)$, parallel to $\mathbf{v} = [2, 6, -1]$, and intersecting the line $\mathbf{x} = [6, -5, 4] + t[4, -3, 1]$.

Remark. This requires the two lines to be coplanar. Use this condition to find the specific parameter or geometric setup.

9. Given two skew lines:

$$L_1 : \mathbf{x} = [0, 1, 1] + t[1, -1, 0] \quad \text{and} \quad L_2 : \mathbf{x} = [-1, 1, 0] + s[2, 1, 2].$$

(a) Calculate the shortest distance between L_1 and L_2.
(b) Find the equation of the line that is perpendicular to both L_1 and L_2 and intersects both.

10. **Reflection of a Line.** Find the equation of the line L' which lies in the plane $x + y + z = 0$, passes through the origin, and makes an angle of $\pi/6$ with the line of intersection of the planes $x + y + 2 = 0$ and $x - y + 2 = 0$.

11. **Cylindrical and Spherical Coordinates.**

(a) Convert the point with rectangular coordinates $(2\sqrt{3}, 2, 4)$ into both cylindrical and spherical coordinates.

(b) Find the Cartesian equation for the surface defined in spherical coordinates by $r\cos\varphi = a$. Identify the surface.

12. $\star$ **Jacobi Identity.** Prove that for any vectors $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$:

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{0}.$$

Remark. Use the expansion $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$ for each term.

13. $\star$ **Collinearity Condition.** Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be the position vectors of points A, B, C. Prove that A, B, C are collinear if and only if:

$$\mathbf{a} \times \mathbf{b} + \mathbf{b} \times \mathbf{c} + \mathbf{c} \times \mathbf{a} = \mathbf{0}.$$

Interpret the vector $\mathbf{v} = \mathbf{a} \times \mathbf{b} + \mathbf{b} \times \mathbf{c} + \mathbf{c} \times \mathbf{a}$ geometrically in terms of the area of triangle OAB, OBC, OCA.

14. $\star$ **Distance to a Line Segment.** Let AB be a line segment in space with $\mathbf{u} = \overrightarrow{AB}$ and $\mathbf{v} = \overrightarrow{AC}$. Show that the perpendicular distance from a point C to the line containing AB is given by:

$$d = \frac{\sqrt{|\mathbf{u}|^2|\mathbf{v}|^2 - (\mathbf{u} \cdot \mathbf{v})^2}}{|\mathbf{u}|}.$$

Use this result to show that the sum of distances from any interior point of an equilateral triangle to its three sides is constant.

15. $\star\star$ **Reciprocal Basis.** Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be three linearly independent vectors. Define the vectors:

$$\mathbf{a}' = \frac{\mathbf{b} \times \mathbf{c}}{[\mathbf{a}, \mathbf{b}, \mathbf{c}]}, \quad \mathbf{b}' = \frac{\mathbf{c} \times \mathbf{a}}{[\mathbf{a}, \mathbf{b}, \mathbf{c}]}, \quad \mathbf{c}' = \frac{\mathbf{a} \times \mathbf{b}}{[\mathbf{a}, \mathbf{b}, \mathbf{c}]}$$

where $[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$.

(a) Show that $\mathbf{a} \cdot \mathbf{a}' = 1$ and $\mathbf{a} \cdot \mathbf{b}' = 0$.

(b) Prove that any vector $\mathbf{r}$ can be expanded as:

$$\mathbf{r} = (\mathbf{r} \cdot \mathbf{a})\mathbf{a}' + (\mathbf{r} \cdot \mathbf{b})\mathbf{b}' + (\mathbf{r} \cdot \mathbf{c})\mathbf{c}'.$$

(c) Show that $[\mathbf{a}', \mathbf{b}', \mathbf{c}'] = \frac{1}{[\mathbf{a}, \mathbf{b}, \mathbf{c}]}$.

16. $\star\star$ **Tetrahedral Geometry.** In a tetrahedron $ABCD$, let the areas of the faces opposite vertices A, B, C, D be S_A, S_B, S_C, S_D respectively. Let $\mathbf{n}_A, \mathbf{n}_B, \mathbf{n}_C, \mathbf{n}_D$ be the outward-pointing unit normal vectors to these faces. Prove the "Law of Cosines for Tetrahedra":

$$S_D^2 = S_A^2 + S_B^2 + S_C^2 - 2S_AS_B\cos\theta_{AB} - 2S_BS_C\cos\theta_{BC} - 2S_CS_A\cos\theta_{CA}$$

where θ_{XY} is the internal dihedral angle between face X and face Y.

Remark. Start from the vector identity $\sum S_i\mathbf{n}_i = \mathbf{0}$ (which you may need to prove first).

Chapter 7

Conic

The curves known as conic sections (the ellipse, the hyperbola, and the parabola), constitute the simplest class of curves beyond the circle. Their study boasts a rich history, tracing back to the Greek geometer and astronomer Menaechmus in the 4th century BC. While investigating the Delian problem of duplicating the cube, Menaechmus sought two mean proportionals x and y between two given segments of length a and b:

$$a : x = x : y = y : b.$$

This continued proportion implies the algebraic relations $x^2 = ay$ and $xy = ab$. From these, one deduces $x^3 = axy = a(ab) = a^2b$, leading to the proportion $a^3 : x^3 = a^3 : a^2b = a : b$. If $b = 2a$, then x represents the side of a cube with volume double that of a cube with side a.

Menaechmus recognised that the curves represented by $x^2 = ay$ and $xy = ab$ are obtained by slicing a right circular cone with a plane. Over a century later, Apollonius of Perga synthesised this knowledge in his definitive treatise *The Conics*.

In this chapter, we transition from the vector algebra of lines and planes to the analytic geometry of curves. Rather than relying on the three-dimensional cone intersection definition immediately, we adopt a unified definition based on metric properties in the plane — specifically, distances to fixed points and lines. This approach allows us to deploy the algebraic tools developed in previous chapters.

7.1 Focus, Directrix and Eccentricity

Classically, the three types of conic sections are often treated as distinct entities. However, a more coherent understanding arises from a single geometric definition governed by a parameter known as eccentricity.

Definition 7.1.1. ***Conic Section***. A conic section (or conic) is the locus of a point X in a plane containing a fixed point F and a fixed line D (not containing F), such that the distance from X to F is in a constant ratio e ($e > 0$) to the perpendicular distance from X to D:

$$|XF| = e|XE|,$$

where E is the orthogonal projection of X onto D. The fixed point F is called the focus, the fixed line D is called the directrix, and the constant e is called the eccentricity.

The nature of the curve is determined entirely by the value of the eccentricity e.

***Theorem* 7.1.1. Classification of Conics**. Let e be the eccentricity of a conic section.

- If $e = 1$, the conic is a **parabola**.
- If $0 < e < 1$, the conic is an **ellipse**.
- If $e > 1$, the conic is a **hyperbola**.

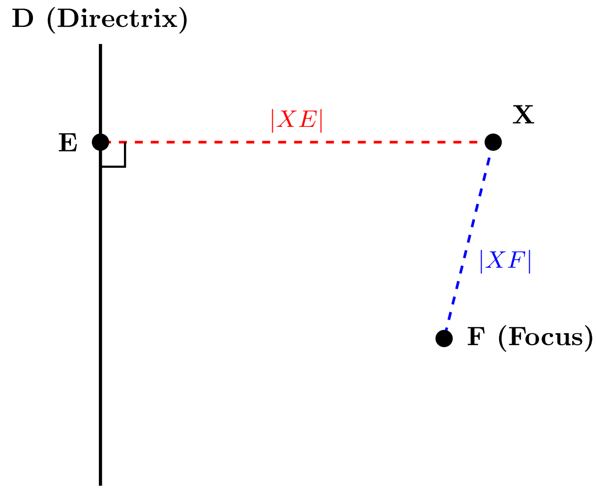


Figure 7.1: The geometric definition of a conic section.

The Parabola

We begin our detailed analysis with the parabola, defined by the eccentricity $e = 1$. By Figure 7.1, the parabola is the locus of points equidistant from the focus and the directrix.

To derive the algebraic equation of the parabola, we must construct an appropriate coordinate system. Let F be the focus and D be the directrix. We drop a perpendicular from F to D, intersecting D at a point H. The midpoint O of the segment FH must lie on the parabola, as it is equidistant from F and D.

We define O as the origin $(0, 0)$. We align the x-axis with the line FH, pointing towards F, and the y-axis perpendicular to it at O. Let the distance $|FH| = 2a$, where $a > 0$. Consequently, the coordinates of the focus are $F(a, 0)$, and the directrix D is the line given by $x = -a$ (or $x + a = 0$).

***Theorem* 7.1.2. Standard Equation of a Parabola**. The locus of points equidistant from the focus $F(a, 0)$ and the directrix $x = -a$ is the curve given by the equation:

$$y^2 = 4ax.$$

Proof. Let $X(x, y)$ be a point on the parabola. The distance from X to the focus F is given by the distance formula (magnitude of displacement vector):

$$|XF| = \sqrt{(x-a)^2 + y^2}.$$

The perpendicular distance from X to the directrix $x = -a$ is simply the horizontal distance:

$$|XE| = |x - (-a)| = |x + a|.$$

By definition of the parabola, $|XF| = |XE|$. Equating the expressions:

$$\sqrt{(x-a)^2 + y^2} = |x + a|.$$

Squaring both sides yields:

$$(x-a)^2 + y^2 = (x+a)^2.$$

Expanding the squares:

$$x^2 - 2ax + a^2 + y^2 = x^2 + 2ax + a^2.$$

Subtracting $x^2 + a^2$ from both sides and rearranging:

$$y^2 = 2ax + 2ax \implies y^2 = 4ax.$$

Conversely, suppose $X(x, y)$ satisfies $y^2 = 4ax$. Since $a > 0$ and $y^2 \geq 0$, we must have $x \geq 0$. Then $|XF|^2 = (x-a)^2 + 4ax = x^2 + 2ax + a^2 = (x+a)^2$. Thus $|XF| = |x + a|$. Since $x \geq 0$ and $a > 0$, $x + a > 0$, so $|x + a| = x + a$, which is exactly the distance to the directrix. ∎

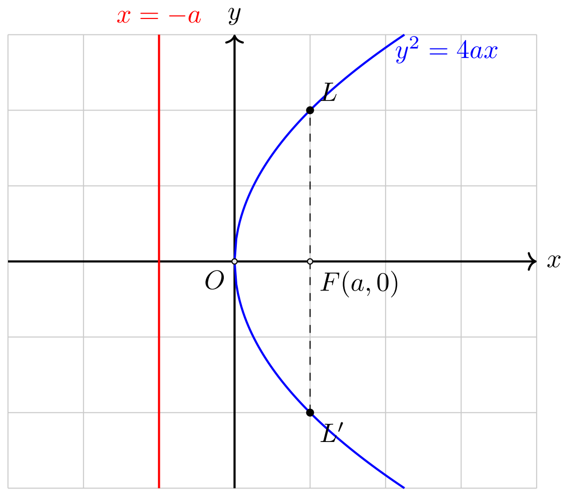


Figure 7.2: The parabola $y^2 = 4ax$ with focus F and directrix $x = -a$.

The geometric features of the parabola lead to specific terminology:

- **Axis:** The line passing through the focus and perpendicular to the directrix. For $y^2 = 4ax$, the axis is the x-axis ($y = 0$). The curve is symmetric about its axis.
- **Vertex:** The point where the parabola intersects its axis. Here, it is the origin O.
- **Latus Rectum:** The chord passing through the focus and perpendicular to the axis. Substituting $x = a$ into $y^2 = 4ax$ gives $y^2 = 4a^2$, so $y = \pm 2a$. The endpoints are $(a, 2a)$ and $(a, -2a)$. The length of the latus rectum is $4a$.

Just as the position vector of a point depends on the origin, the equation of a parabola depends on the orientation of the axes. If the focus is placed at $(-a, 0)$ with directrix $x = a$, the equation becomes $y^2 = -4ax$. Interchanging x and y (placing the focus on the y-axis) yields $x^2 = 4ay$ or $x^2 = -4ay$.

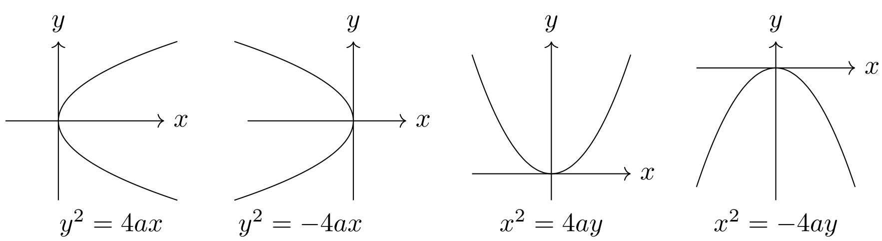


Figure 7.3: Orientations of the parabola with vertex at the origin.

A remarkable property of parabolas is that they lack a fundamental scale; they differ only in "zoom".

***Theorem* 7.1.3. Similarity of Parabolas**. All parabolas are similar.

Proof. Consider two parabolas given by $y^2 = 4ax$ and $Y^2 = 4bX$. We wish to show that the first can be transformed into the second by a scaling transformation. Let $x = \frac{a}{b}X$ and $y = \frac{a}{b}Y$. Substituting these into the first equation:

$$\left(\frac{a}{b}Y\right)^2 = 4a\left(\frac{a}{b}X\right) \implies \frac{a^2}{b^2}Y^2 = \frac{4a^2}{b}X.$$

Multiplying through by $\frac{b^2}{a^2}$ yields $Y^2 = 4bX$. Thus, a simple homothety (scaling) maps any parabola onto any other. ∎

The Ellipse

We turn our attention to the conic sections defined by an eccentricity e satisfying $0 < e < 1$. Such curves are called ellipses. Unlike the parabola, which is an open curve extending to infinity, we shall see that the ellipse is a closed curve possessing a centre of symmetry.

Consider an ellipse with focus F and directrix D. To derive the simplest algebraic description, we anticipate that the curve possesses an axis of symmetry passing through the focus and perpendicular to the directrix.

While one could define the coordinate system with the focus at the origin (as is common in polar coordinates), the Cartesian equation is most elegant when the origin coincides with the geometric centre of the curve. Let us place the focus F at the point $(c, 0)$ on the x-axis, where $c > 0$, and let the corresponding directrix D be the vertical line $x = d$, with $d > c$.

Let $X(x, y)$ be a point on the ellipse. By definition, the ratio of the distance $|XF|$ to the distance from the directrix $|XE|$ is the eccentricity e.

$$|XF|^2 = e^2|XE|^2.$$

Substituting the coordinates:

$$(x - c)^2 + y^2 = e^2(d - x)^2.$$

Expanding and grouping terms by powers of x:

$$x^2 - 2cx + c^2 + y^2 = e^2(d^2 - 2dx + x^2).$$

$$(1 - e^2)x^2 - 2(c - e^2d)x + y^2 + (c^2 - e^2d^2) = 0.$$

To eliminate the linear term in x and centre the equation at the origin, we require the coefficient of x to vanish:

$$c - e^2d = 0 \implies c = e^2d.$$

This relates the position of the focus and the directrix. Let us define a scaling parameter a such that the distance from the centre to the directrix is $d = a/e$. Then the focus is located at $c = e(a/e) = ae$. Substituting $c = ae$ and $d = a/e$ into the constant term:

$$c^2 - e^2d^2 = (ae)^2 - e^2(a/e)^2 = a^2e^2 - a^2 = -a^2(1 - e^2).$$

The equation simplifies to:

$$(1 - e^2)x^2 + y^2 = a^2(1 - e^2).$$

Dividing by $a^2(1 - e^2)$:

$$\frac{x^2}{a^2} + \frac{y^2}{a^2(1 - e^2)} = 1.$$

Since $0 < e < 1$, the term $1 - e^2$ is positive. We define $b^2 = a^2(1 - e^2)$. This yields the standard equation of the ellipse.

***Theorem* 7.1.4. Standard Equation of the Ellipse**. The locus of an ellipse with eccentricity $0 < e < 1$ can be represented in a coordinate system centred at the ellipse's centre by the equation:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

where $a > b > 0$ and $b^2 = a^2(1 - e^2)$.

This derivation reveals the geometric significance of the constants:

- The curve intersects the x-axis at $(\pm a, 0)$ and the y-axis at $(\pm b, 0)$.
- Because x appears only as x^2, the curve is symmetric about the y-axis. Consequently, there must exist a second focus $F'(-ae, 0)$ and a second directrix $x = -a/e$ that generate the same curve.
- The curve is symmetric about the origin, which is called the *centre*.

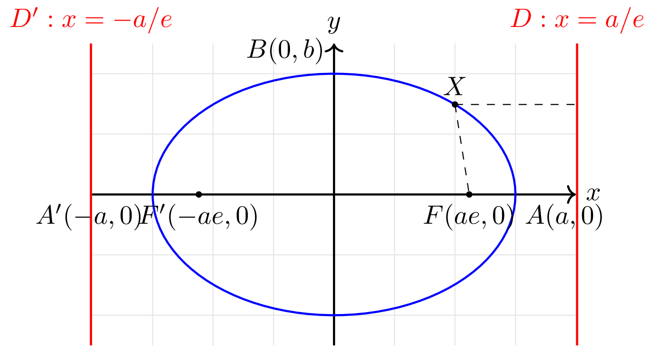


Figure 7.4: The ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ with its foci and directrices.

The chord AA' of length $2a$ lying on the focal axis is called the *major axis*. The chord BB' of length $2b$ perpendicular to the focal axis is called the *minor axis*. The relationship between the parameters is summarised by the Pythagorean identity inherent in the definition:

$$a^2 = b^2 + c^2$$

where $c = ae$ is the distance from the centre to a focus.

Geometry and Similarity

Unlike parabolas, not all ellipses are similar. The shape of an ellipse is governed entirely by its eccentricity.

- As $e \to 0$, the foci converge to the centre ($c \to 0$) and $b \to a$. The ellipse approaches a circle.
- As $e \to 1$, the foci move towards the vertices ($c \to a$) and $b \to 0$. The ellipse becomes increasingly elongated (or flat).

Two ellipses are similar if and only if they have the same eccentricity.

If the equation is given as $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ with $b > a$, the major axis lies along the y-axis. In this case, the eccentricity is determined by $e = \sqrt{1 - a^2/b^2}$, and the foci are located at $(0, \pm be)$.

The Bifocal Property

A distinguishing characteristic of the ellipse is the "string property": the sum of the distances from any point on the curve to the two foci is constant.

***Theorem* 7.1.5. Bifocal Definition.** The locus of a point X in the plane such that the sum of its distances from two fixed points F and F' is a constant $2a$ (where $2a > |FF'|$) is an ellipse with foci at F and F' and major axis $2a$.

Proof. Let the foci be located at $F(c, 0)$ and $F'(-c, 0)$. The condition is:

$$|XF| + |XF'| = 2a.$$

Using coordinates $X(x, y)$:

$$\sqrt{(x-c)^2 + y^2} + \sqrt{(x+c)^2 + y^2} = 2a.$$

Rearranging one radical to the right side and squaring:

$$(x-c)^2 + y^2 = 4a^2 - 4a\sqrt{(x+c)^2 + y^2} + (x+c)^2 + y^2.$$

Simplifying the squared terms:

$$x^2 - 2cx + c^2 + y^2 = 4a^2 - 4a\sqrt{(x+c)^2 + y^2} + x^2 + 2cx + c^2 + y^2.$$
$$-4cx - 4a^2 = -4a\sqrt{(x+c)^2 + y^2}.$$

Dividing by -4 and squaring again:

$$(a^2 + cx)^2 = a^2[(x+c)^2 + y^2].$$
$$a^4 + 2a^2cx + c^2x^2 = a^2(x^2 + 2cx + c^2 + y^2).$$
$$a^4 + 2a^2cx + c^2x^2 = a^2x^2 + 2a^2cx + a^2c^2 + a^2y^2.$$

Collecting terms:

$$a^2x^2 - c^2x^2 + a^2y^2 = a^4 - a^2c^2.$$
$$(a^2 - c^2)x^2 + a^2y^2 = a^2(a^2 - c^2).$$

Since $2a > 2c$, we have $a > c$, so $a^2 - c^2$ is positive. Let $b^2 = a^2 - c^2$.

$$b^2x^2 + a^2y^2 = a^2b^2 \implies \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Conversely, any point satisfying the ellipse equation satisfies the algebraic steps in reverse, provided we respect the non-negativity of the distance functions (which corresponds to the triangle inequality $|XF| + |XF'| \geq |FF'|$). ■

Parametric Equations and the Auxiliary Circle

The ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ can be parameterized by introducing an angle ϕ:

$$x = a\cos\phi, \quad y = b\sin\phi, \quad \text{for } \phi \in [0, 2\pi).$$

This representation has a geometric interpretation involving the *auxiliary circle*, which is the circle of radius a circumscribing the ellipse (centred at the origin).

***Theorem* 7.1.6. Auxiliary Circle Construction**. Let C_a be the circle $x^2 + y^2 = a^2$ and C_b be the circle $x^2 + y^2 = b^2$. Let a ray from the origin at angle ϕ intersect C_a at R and C_b at Q. The point X with the x-coordinate of R and the y-coordinate of Q lies on the ellipse.

Proof. The coordinates of R are $(a\cos\phi, a\sin\phi)$. The coordinates of Q are $(b\cos\phi, b\sin\phi)$. The constructed point X has coordinates $(a\cos\phi, b\sin\phi)$. Substituting into the standard equation:

$$\frac{(a\cos\phi)^2}{a^2} + \frac{(b\sin\phi)^2}{b^2} = \cos^2\phi + \sin^2\phi = 1.$$ ■

The angle ϕ is called the *eccentric angle* of the point X. It is important to distinguish ϕ from the standard polar angle θ; they are related by $\tan\theta = \frac{b}{a}\tan\phi$.

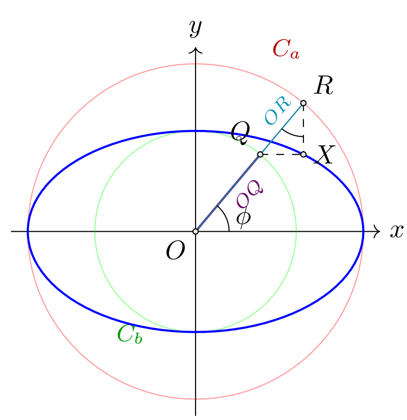


Figure 7.5: Construction of a point X on the ellipse using the eccentric angle ϕ.

The Hyperbola

The final family of conic sections, defined by an eccentricity $e > 1$, is the hyperbola. This curve consists of two disjoint, unbounded branches. While geometrically distinct from the ellipse, the hyperbola shares a remarkably similar algebraic structure, differing primarily by a sign change in the standard equation.

Consider a hyperbola with focus F, directrix D, and eccentricity $e > 1$. By definition, the locus of points X on the hyperbola satisfies $|XF| = e|XE|$, where $|XE|$ is the perpendicular distance to the directrix.

To derive the standard equation, we employ the same coordinate alignment strategy as used for the ellipse. We seek a centre of symmetry at the origin. Let the focus be placed at $F(ae, 0)$ and the directrix be the line $x = a/e$. Since $e > 1$, the focus is further from the origin than the directrix.

For any point $X(x, y)$ on the curve:

$$\sqrt{(x - ae)^2 + y^2} = e\left|x - \frac{a}{e}\right| = |ex - a|.$$

Squaring both sides:

$$(x - ae)^2 + y^2 = (ex - a)^2.$$

Expanding:

$$x^2 - 2aex + a^2e^2 + y^2 = e^2x^2 - 2aex + a^2.$$

Rearranging terms to group x and y:

$$x^2(e^2 - 1) - y^2 = a^2(e^2 - 1).$$

Since $e > 1$, the term $e^2 - 1$ is positive. We define $b^2 = a^2(e^2 - 1)$. Substituting this definition into the equation:

$$x^2\frac{b^2}{a^2} - y^2 = b^2.$$

Dividing by b^2, we obtain the canonical form.

***Theorem* 7.1.7. Standard Equation of the Hyperbola**. The locus of a hyperbola with eccentricity $e > 1$ can be represented in a coordinate system centred at the hyperbola's centre by the equation:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$

where $b^2 = a^2(e^2 - 1)$.

Geometric Properties

The hyperbola possesses two axes of symmetry intersecting at the centre $O(0, 0)$:

- The **transverse axis** (or major axis) contains the foci and intersects the curve. For the standard equation, this is the x-axis. The intersection points $A(a, 0)$ and $A'(-a, 0)$ are called the **vertices**. The distance between them is $2a$.
- The **conjugate axis** (or minor axis) is perpendicular to the transverse axis and does not intersect the curve. For the standard equation, this is the y-axis. The segment connecting $B(0, b)$ and $B'(0, -b)$ has length $2b$.

The relationship between the geometric parameters is given by:

$$c^2 = a^2 + b^2$$

where $c = ae$ is the distance from the centre to each focus. This differs from the ellipse relation ($a^2 = b^2 + c^2$).

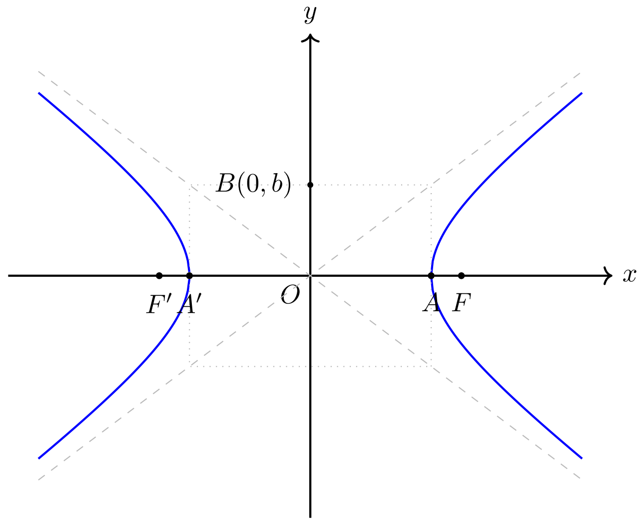


Figure 7.6: The hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ with asymptotes and foci.

If the roles of x and y are interchanged, the equation becomes $\frac{y^2}{a^2} - \frac{x^2}{b^2} = 1$. This hyperbola opens vertically, with vertices at $(0, \pm a)$ and foci at $(0, \pm c)$. Such a hyperbola is said to be *conjugate* to the standard one.

Asymptotes

A unique feature of the hyperbola is the existence of asymptotes—lines that the curve approaches arbitrarily closely as it extends to infinity. Solving for y in the standard equation:

$$y = \pm\frac{b}{a}\sqrt{x^2 - a^2} = \pm\frac{b}{a}x\sqrt{1 - \frac{a^2}{x^2}}.$$

As $|x| \to \infty$, the term $a^2/x^2 \to 0$, and thus $y \approx \pm\frac{b}{a}x$. This suggests that the lines $y = \frac{b}{a}x$ and $y = -\frac{b}{a}x$ are asymptotes.

***Theorem* 7.1.8. Asymptotes of a Hyperbola**. The hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ is bounded by the asymptotes:

$$y = \pm\frac{b}{a}x.$$

Each branch of the hyperbola lies entirely within the region formed by the asymptotes containing the respective vertex.

Proof. Consider a point $X_1(x_1, y_1)$ on the hyperbola in the first quadrant $(x_1 > a, y_1 > 0)$. The vertical distance between the line $y = \frac{b}{a}x$ and the curve is:

$$\Delta y = \frac{b}{a}x_1 - \frac{b}{a}\sqrt{x_1^2 - a^2} = \frac{b}{a}\left(x_1 - \sqrt{x_1^2 - a^2}\right).$$

Multiplying by the conjugate:

$$\Delta y = \frac{b}{a}\frac{(x_1^2 - (x_1^2 - a^2))}{x_1 + \sqrt{x_1^2 - a^2}} = \frac{ab}{x_1 + \sqrt{x_1^2 - a^2}}.$$

As $x_1 \to \infty$, the denominator grows without bound, so $\Delta y \to 0$. Thus, the curve approaches the line asymptotically. ■

The asymptotes provide a convenient method for sketching the curve. By drawing the rectangle with sides $x = \pm a$ and $y = \pm b$, the diagonals of this rectangle extend to form the asymptotes.

The Bifocal Property

Analogous to the ellipse, the hyperbola is the locus of points defined by the difference of distances to the foci.

***Theorem* 7.1.9. Bifocal Definition**. The locus of a point X such that the absolute difference of its distances from two fixed points F and F' is a constant $2a$ (where $2a < |FF'|$) is a hyperbola with foci at F and F' and transverse axis $2a$:

$$||XF| - |XF'|| = 2a.$$

Proof. Let $F(c, 0)$ and $F'(-c, 0)$. The derivation follows the exact algebraic steps as the ellipse proof, but starting with $\sqrt{(x-c)^2 + y^2} - \sqrt{(x+c)^2 + y^2} = \pm 2a$. Squaring twice leads to $(c^2 - a^2)x^2 - a^2y^2 = a^2(c^2 - a^2)$. Since $c > a$ (triangle inequality in $\triangle XFF'$), we set $b^2 = c^2 - a^2$, yielding the standard equation. ■

Parametric Equations

To parameterise the hyperbola, we exploit the hyperbolic trigonometric identity $\cosh^2 t - \sinh^2 t = 1$ or the trigonometric identity $\sec^2 \phi - \tan^2 \phi = 1$. Using the latter, we set:

$$x = a \sec \phi, \quad y = b \tan \phi.$$

This parameterisation covers the hyperbola for $\phi \in (-\pi/2, \pi/2) \cup (\pi/2, 3\pi/2)$. Geometrically, this relates to a tangent construction on the auxiliary circle $x^2 + y^2 = a^2$. Construct a tangent to the auxiliary circle at $R(a \cos \theta, a \sin \theta)$. Let this tangent intersect the x-axis at T. Then the x-coordinate of the hyperbola point is OT.

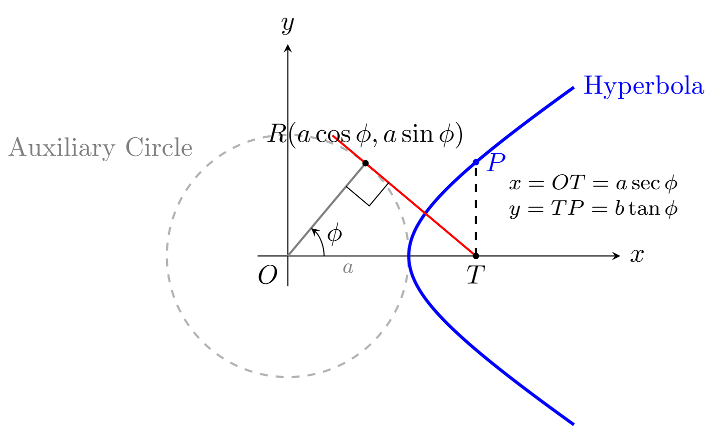


Figure 7.7: Geometric construction of the hyperbola using the auxiliary circle ($x^2 + y^2 = a^2$). The point P is defined by the intersection of the vertical line through the secant point T and the horizontal level defined by $b \tan \phi$.

7.2 Exercises

Part I: Fundamentals and Standard Forms

1. Determine the equation of the conic section satisfying the given conditions.

(a) A hyperbola with vertices at $(0, \pm 1)$ and foci at $(0, \pm 3)$.
(b) An ellipse with foci at $(\pm 2, 0)$ passing through the point $(2, 3)$.
(c) A parabola with focus at $(2, 1)$ and directrix $x = -4$.

2. Find the eccentricity, coordinates of the foci, and equations of the directrices for the following curves:

(a) $9x^2 + 25y^2 = 225$
(b) $x^2 - 4y^2 = 16$
(c) $y^2 = -12x$

3. **The Latus Rectum.** The latus rectum is the chord passing through the focus perpendicular to the axis of symmetry.
   (a) Show that the length of the latus rectum for the parabola $y^2 = 4ax$ is $4a$.
   (b) Show that the length of the latus rectum for the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ is $\frac{2b^2}{a}$.
   (c) Calculate the length of the latus rectum for the hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$.

4. Let P be a point on the hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$.
   (a) Find the perpendicular distance from P to the asymptote $y = \frac{b}{a}x$.
   (b) Find the perpendicular distance from P to the asymptote $y = -\frac{b}{a}x$.
   (c) Deduce that the product of the perpendicular distances from any point on a hyperbola to its asymptotes is constant and equal to $\frac{a^2b^2}{a^2+b^2}$.

5. **Midpoint Loci.**
   (a) Let P be a variable point on the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Find the locus of the midpoint of the segment connecting P to the centre of the ellipse.
   (b) Let P be a variable point on the parabola $y^2 = 4ax$. Find the locus of the midpoint of the segment connecting P to the vertex.

Part II: Geometric Properties and Proofs

6. **The Focal Chord Property.** A chord of the parabola $y^2 = 4ax$ passes through the focus $F(a, 0)$ and intersects the curve at points P and Q. Let SP and SQ denote the lengths of the segments from the focus to the curve. Prove that the semi-latus rectum $2a$ is the harmonic mean of SP and SQ. That is:

$$\frac{1}{|SP|} + \frac{1}{|SQ|} = \frac{1}{a}.$$

7. **Optical Property of the Parabola.** Let $P(at^2, 2at)$ be a point on the parabola $y^2 = 4ax$.
   (a) Show that the tangent to the parabola at P has the equation $ty = x + at^2$.
   (b) Let the tangent intersect the x-axis at T. Prove that $|TF| = |PF|$, where F is the focus.
   (c) Conclude that the triangle TFP is isosceles and that the tangent line bisects the angle formed by the focal radius FP and the line parallel to the axis of the parabola.

   Remark. This explains why parabolic mirrors reflect parallel rays to the focus.

8. **Affine Transformations and Area.**
   (a) Consider the unit circle $u^2 + v^2 = 1$. Apply the transformation $x = au$, $y = bv$. Show that the image of the circle is the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.
   (b) The area of the unit circle is π. Using the property that a linear transformation $(x, y) \mapsto (au, bv)$ scales area by the determinant of the transformation matrix, deduce the area of the ellipse.

9. **Orthogonal Intersection.** Prove that the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ and the hyperbola $\frac{x^2}{a^2-k^2} - \frac{y^2}{k^2-b^2} = 1$ (where $b < k < a$) are confocal (share the same foci) and intersect at right angles.

   Remark. Find the slopes of the tangents at an intersection point and show their product is -1.

10. **The Auxiliary Circle and Eccentric Angle.**
    (a) Let P be a point on the ellipse with eccentric angle ϕ. Show that the equation of the tangent at P is $\frac{x}{a}\cos\phi + \frac{y}{b}\sin\phi = 1$.
    (b) The normal to the ellipse at P is the line perpendicular to the tangent at P. Find the equation of the normal.

(c) Prove that the normal at P does *not* pass through the centre of the ellipse unless the ellipse is a circle or P is a vertex.

11. $\star$ **The Director Circle.** The director circle of a conic is the locus of points from which the two tangent lines drawn to the conic are perpendicular.

(a) Show that the line $y = mx + \sqrt{a^2m^2 + b^2}$ is tangent to the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ for any slope m.
(b) Let $P(h, k)$ be a point from which perpendicular tangents are drawn. Use the tangent equation to form a quadratic in m and apply the condition for perpendicular roots ($m_1m_2 = -1$).
(c) Prove that the director circle of the ellipse is $x^2 + y^2 = a^2 + b^2$.
(d) What is the locus of perpendicular tangents for a parabola?

Part III: Advanced Topics

12. Polar Equations of Conics.

(a) Consider a conic with focus at the origin and directrix $x = -d$. Let (r, θ) be polar coordinates. Using the definition $|OP| = e|PE|$, derive the polar equation of the conic:

$$r = \frac{ed}{1 - e\cos\theta}.$$

(b) Show that the length of the semi-latus rectum is $l = ed$. Hence write the equation as $r = \frac{l}{1-e\cos\theta}$.
(c) Discuss how the range of possible r values changes as e crosses the threshold of 1.

13. $\star$ **Conjugate Diameters.** Two diameters of an ellipse are said to be conjugate if each bisects the chords parallel to the other.

(a) If $y = mx$ and $y = m'x$ are equations of two conjugate diameters of the ellipse $b^2x^2 + a^2y^2 = a^2b^2$, prove that $mm' = -\frac{b^2}{a^2}$.
(b) Let P and D be extremities of two conjugate diameters. Parameterise P as $(a\cos\phi, b\sin\phi)$. Show that D has parameters corresponding to $\phi \pm \frac{\pi}{2}$.
(c) Prove the **First Theorem of Apollonius**: For any pair of conjugate semi-diameters of lengths r_1 and r_2, the sum $r_1^2 + r_2^2$ is constant and equal to $a^2 + b^2$.

14. The Rectangular Hyperbola. A hyperbola is called rectangular if its asymptotes are perpendicular.

(a) State the condition on the axes a and b for the hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ to be rectangular.
(b) Consider the coordinate transformation corresponding to a rotation by 45°: $x = \frac{X+Y}{\sqrt{2}}$, $y = \frac{-X+Y}{\sqrt{2}}$. Apply this to the rectangular hyperbola $x^2 - y^2 = a^2$ and show it reduces to the form $XY = c^2$. Find c in terms of a.
(c) Show that parametric equations $x = ct, y = c/t$ satisfy this equation.

15. $\star$ **Three Normals from a Point.** Consider the parabola $y^2 = 4ax$.

(a) Show that the normal to the parabola at the point $(at^2, 2at)$ is given by $y + tx = 2at + at^3$.
(b) If this normal passes through a fixed point (h, k), show that t must satisfy the cubic equation $at^3 + (2a - h)t - k = 0$.
(c) Deduce that, in general, three normals can be drawn from a given point (h, k) to the parabola.
(d) If the feet of these three normals are points corresponding to parameters t_1, t_2, t_3, prove that $t_1 + t_2 + t_3 = 0$. What does this imply about the centroid of the triangle formed by the feet of the normals?

Part IV: The Hyperbolic Functions

Remark. The hyperbolic cosine and hyperbolic sine are defined as:

$$\cosh t = \frac{e^t + e^{-t}}{2}, \quad \sinh t = \frac{e^t - e^{-t}}{2}.$$

These functions satisfy the fundamental identity $\cosh^2 t - \sinh^2 t = 1$. Geometrically, they parameterise the unit hyperbola $x^2 - y^2 = 1$ just as $\cos\theta$ and $\sin\theta$ parameterise the unit circle. The point $(x, y) = (\cosh t, \sinh t)$ lies on the right branch of the hyperbola. The parameter t represents twice the signed area of the hyperbolic sector bounded by the curve, the x-axis, and the ray from the origin to the point.

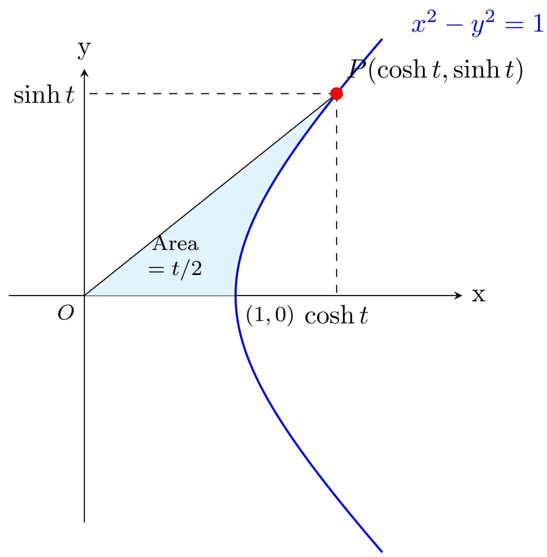


Figure 7.8: The geometric definition of hyperbolic functions using the unit hyperbola.

16. Fundamental Properties.

(a) Using the exponential definitions, verify the identity $\cosh^2 t - \sinh^2 t = 1$.

(b) Prove that $\cosh t + \sinh t = e^t$ and $\cosh t - \sinh t = e^{-t}$.

(c) By solving for t in terms of x and y, show that the inverse parameterisation for a point $P(x, y)$ on the right branch of the unit hyperbola is given by $t = \ln(x + y)$.

(d) Verify algebraically that the point $(a \cosh t, b \sinh t)$ lies on the general hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$.

17. Hyperbolic Identities and De Moivre's Analogue.

(a) Prove the addition formula: $\cosh(a + b) = \cosh a \cosh b + \sinh a \sinh b$.

(b) Derive the corresponding formula for $\sinh(a + b)$.

(c) **Hyperbolic De Moivre's Theorem.** Using the property $e^{nt} = (e^t)^n$, prove that for any integer n:

$$(\cosh t + \sinh t)^n = \cosh(nt) + \sinh(nt).$$

(d) Use this result to express $\cosh 2t$ and $\sinh 2t$ in terms of $\cosh t$ and $\sinh t$.

18. The Tangent Line (Algebraic Approach). In plane geometry, a line is tangent to a conic if it intersects the curve at exactly one point (a "repeated root").

(a) Consider the line L given by equation $x \cosh t - y \sinh t = 1$. Show that the point $P(\cosh t, \sinh t)$ lies on both the unit hyperbola $x^2 - y^2 = 1$ and the line L.

(b) To determine the intersection of L and the hyperbola, substitute x from the line equation into the hyperbola equation. Show that this results in a quadratic equation with a discriminant of zero.

(c) Conclude that $x \cosh t - y \sinh t = 1$ is the tangent to the unit hyperbola at P.

Chapter 8

Plane Sections of a Cone

In the chapter earlier, we defined the ellipse, parabola, and hyperbola through the metric properties of foci and directrices in the plane. We now justify the collective terminology "conic sections" by demonstrating that these curves arise naturally as the intersections of a plane with a double-napped circular cone.

8.1 The Circular Cone

A circular cone is a surface of revolution generated by rotating a line about a fixed axis.

Definition 8.1.1. ***Circular Cone***. Let B be a fixed line (the axis) and A be a fixed point on B (the vertex). Let L be a line passing through A that intersects B at an acute angle θ. The surface generated by rotating L about the axis B is called a right circular cone. The angle θ is the half-angle of the cone. Any line on the surface passing through A is called a generator or ruling.

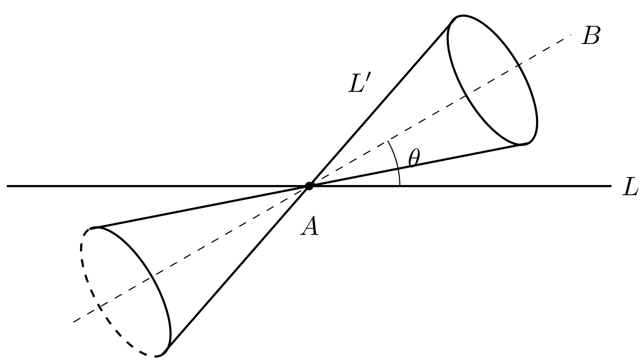


Figure 8.1: A double-napped cone with axis B inclined at angle θ to the horizontal line L.

To derive the algebraic equation of the cone, let $\mathbf{a} = [a_1, a_2, a_3]$ be the position vector of the vertex A, and let $\mathbf{b} = [b_1, b_2, b_3]$ be a unit vector along the axis B. A point X with position vector $\mathbf{x} = [x, y, z]$ lies on the cone if and only if the angle between the displacement vector $\mathbf{x} - \mathbf{a}$ and the axis vector $\mathbf{b}$ is either θ or $\pi - \theta$.

Using the dot product, this condition is expressed as:

$$(\mathbf{x} - \mathbf{a}) \cdot \mathbf{b} = \pm|\mathbf{x} - \mathbf{a}| \cos\theta.$$

Let $\lambda = \cos\theta$. The equation splits into two cases:

$$(\mathbf{x} - \mathbf{a}) \cdot \mathbf{b} = \lambda|\mathbf{x} - \mathbf{a}| \quad \text{and} \quad (\mathbf{x} - \mathbf{a}) \cdot \mathbf{b} = -\lambda|\mathbf{x} - \mathbf{a}|.$$

These correspond to the two *nappes* of the cone, which meet only at the vertex. Squaring both sides eliminates the ambiguity of the sign and yields the general vector equation of the cone:

$$[(\mathbf{x}-\mathbf{a})\cdot\mathbf{b}]^2 = \lambda^2|\mathbf{x}-\mathbf{a}|^2.$$

In terms of coordinates, with $\mathbf{b}$ being a unit vector ($b_1^2+b_2^2+b_3^2=1$):

$$[b_1(x-a_1)+b_2(y-a_2)+b_3(z-a_3)]^2 = \lambda^2[(x-a_1)^2+(y-a_2)^2+(z-a_3)^2].$$

8.2 The Geometry of the Intersection

We investigate the curve Γ formed by the intersection of a cone K and a plane E. To analyse the nature of this curve, we position our coordinate system to simplify the algebra without loss of generality.

Let ϕ be the acute angle of inclination between the axis B of the cone and the cutting plane E. We arrange the axes such that:

1. The cutting plane E coincides with the xy-plane ($z=0$).
2. The axis B of the cone lies in the xz-plane ($y=0$).
3. The angle between B and the positive x-axis is ϕ.

Under this configuration, the unit vector along the axis is $\mathbf{b}=[\cos\phi, 0, \sin\phi]$. Thus, $b_1=\cos\phi$, $b_2=0$, and $b_3=\sin\phi$. Since the vertex A lies on the axis B (which is in the xz-plane), its coordinates are $\mathbf{a}=[a_1, 0, a_3]$.

Substituting these specific values and $z=0$ into the general cone equation yields the equation of the intersection curve Γ in the xy-plane:

$$[b_1(x-a_1)-b_3a_3]^2 = \lambda^2[(x-a_1)^2+y^2+a_3^2].$$

Recall that $\lambda=\cos\theta$, where θ is the half-angle of the cone. This equation describes the locus of points in the plane $z=0$. By expanding and regrouping terms, we can determine the type of conic section based on the relationship between ϕ and θ.

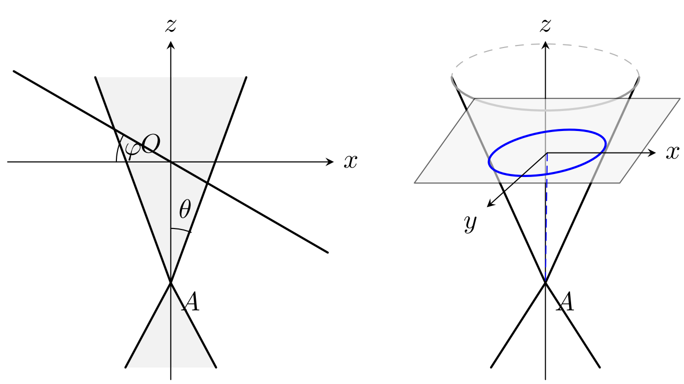


Figure 8.2: Intersection of a cone with a plane. Left: The geometric parameters θ and ϕ. Right: An elliptical section.

Classification of Sections

We assume the cutting plane does not pass through the vertex ($z=0$ does not contain A), so $a_3\neq 0$. We examine the curve based on the angle ϕ relative to θ.

Case 1: The Circle ($\phi=\pi/2$) If $\phi=\pi/2$, the axis of the cone is perpendicular to the cutting plane. Consequently, $b_1=\cos(\pi/2)=0$. Since b is a unit vector, $b_3^2=1$. We may adjust the origin along the x-axis such that $a_1 = 0$. The intersection equation simplifies to:

$$[-b_3 a_3]^2 = \lambda^2[x^2 + y^2 + a_3^2].$$

Since $b_3^2 = 1$:

$$a_3^2 = \lambda^2 x^2 + \lambda^2 y^2 + \lambda^2 a_3^2.$$

Rearranging terms:

$$\lambda^2 x^2 + \lambda^2 y^2 = a_3^2(1 - \lambda^2).$$

Since $\lambda = \cos\theta$ and the vertex is not on the plane ($a_3 \neq 0$), the right-hand side is positive ($a_3^2 \sin^2\theta > 0$). This is the equation of a circle.

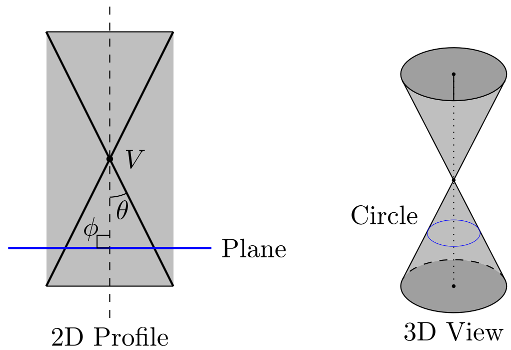


Figure 8.3: The Circle.

Case 2: The Parabola ($\phi = \theta$) If the inclination of the plane equals the half-angle of the cone, the cutting plane is parallel to a generator. Here, $b_1 = \cos\phi = \cos\theta = \lambda$. The coefficient of x^2 in the expansion of the intersection equation involves $b_1^2 - \lambda^2$. Since $b_1 = \lambda$, the x^2 term vanishes. The equation becomes:

$$[b_1(x - a_1) - b_3 a_3]^2 = b_1^2[(x - a_1)^2 + y^2 + a_3^2].$$

Expanding the left side:

$$b_1^2(x - a_1)^2 - 2b_1(x - a_1)b_3 a_3 + b_3^2 a_3^2 = b_1^2(x - a_1)^2 + b_1^2 y^2 + b_1^2 a_3^2.$$

The quadratic terms in x cancel out. The remaining equation is linear in x and quadratic in y. Specifically:

$$b_1^2 y^2 = -2b_1 b_3 a_3(x - a_1) + \text{constant}.$$

By translating the coordinate system to eliminate the constant terms, we arrive at the standard form $y^2 = cx$, representing a parabola.

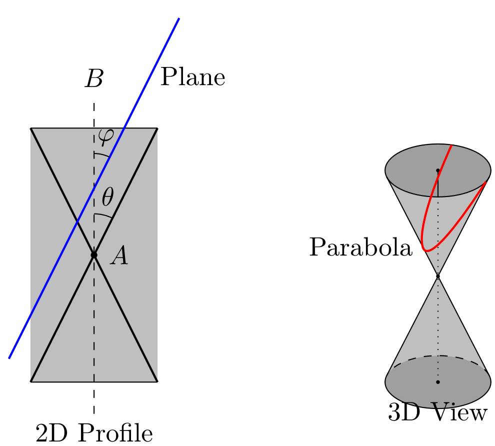


Figure 8.4: The Parabola.

Case 3: The Ellipse ($\phi > \theta$) The plane cuts through a single nappe of the cone. Since $\phi > \theta$, we have $\cos\phi < \cos\theta$, so $b_1 < \lambda$. Let $\mu^2 = \lambda^2 - b_1^2 > 0$. Rearranging the intersection equation to group quadratic terms:

$$(\lambda^2 - b_1^2)x^2 + \lambda^2 y^2 + \cdots = \text{positive constant.}$$

Since the coefficients of x^2 and y^2 are both positive, the curve is an ellipse.

Case 4: The Hyperbola ($\phi < \theta$) The plane cuts both nappes of the cone. Since $\phi < \theta$, we have $\cos\phi > \cos\theta$, so $b_1 > \lambda$. Let $-\mu^2 = \lambda^2 - b_1^2 < 0$. The intersection equation takes the form:

$$-\mu^2 x^2 + \lambda^2 y^2 + \cdots = \text{constant.}$$

The coefficients of x^2 and y^2 have opposite signs. This equation represents a hyperbola.

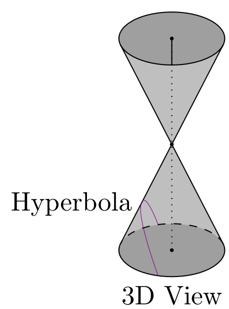


Figure 8.5: The Hyperbola.

***Theorem* 8.2.1. Conic Sections Theorem**. Let K be a circular cone with half-angle θ. Let E be a plane not passing through the vertex, with inclination angle ϕ to the axis of the cone. The intersection of K and E is:

(i) A circle if $\phi = \pi/2$.
(ii) An ellipse if $\phi > \theta$.
(iii) A parabola if $\phi = \theta$.
(iv) A hyperbola if $\phi < \theta$.

Degenerate Conics

The derivation above assumed that the cutting plane does not pass through the vertex of the cone. Algebraically, this ensured that the constant terms in the quadratic equation did not vanish in a way that reduced the solution set to a single point or lines. However, if the plane E passes through the vertex A, the geometric intersection degenerates.

In our coordinate setup, passing through the vertex implies $a_3 = 0$ (since the vertex lies on the x-axis and the plane is $z = 0$). The equation of the intersection becomes homogeneous:

$$(\lambda^2 - b_1^2)u^2 + \lambda^2 y^2 = 0.$$

This represents a "conic" centred at the origin (the vertex). The solution set depends on the sign of the coefficient $(\lambda^2 - b_1^2)$, exactly as in the non-degenerate cases.

(a) **Degenerate Ellipse ($\phi > \theta$):** Here $\lambda^2 - b_1^2 > 0$. The equation is of the form $Ax^2 + Cy^2 = 0$ with $A, C > 0$. The only real solution is the single point $x = 0, y = 0$. Geometrically, the plane cuts the cone only at the vertex.
(b) **Degenerate Parabola ($\phi = \theta$):** Here $\lambda^2 - b_1^2 = 0$. The equation reduces to $\lambda^2 y^2 = 0$, which implies $y = 0$. This solution represents the x-axis (or u-axis). Geometrically, the plane is tangent to the cone along a single generator (ruling).

(c) **Degenerate Hyperbola ($\phi < \theta$):** Here $\lambda^2 - b_1^2 < 0$. Let $A = -(\lambda^2 - b_1^2) > 0$. The equation is $-Ax^2 + Cy^2 = 0$, which factors as $(\sqrt{C}y - \sqrt{A}x)(\sqrt{C}y + \sqrt{A}x) = 0$. This describes two intersecting lines passing through the vertex. Geometrically, the plane cuts through the vertex and intersects both nappes of the cone.

The Cylinder as a Limit

The circular cylinder may be conceptually regarded as a cone whose vertex has receded to infinity, resulting in a half-angle $\theta \to 0$. The generators become parallel to the axis. The intersection of a plane with a cylinder exhibits properties analogous to those of the cone.

Let the axis of the cylinder be the z-axis. The equation of the cylinder is $x^2 + y^2 = r^2$. A cutting plane can be defined by $z = my + k$ (assuming it is not vertical). Substituting this into the cylinder's equation reveals that the projection onto the xy-plane is a circle; thus, the curve on the plane is an ellipse.

If the cutting plane is parallel to the axis of the cylinder (vertical), the intersection geometry mirrors the degenerate cases:

(a) **Empty Set:** If the distance from the axis to the plane is greater than the radius, there is no intersection.
(b) **Degenerate Parabola (Single Line):** If the distance equals the radius, the plane is tangent to the cylinder, intersecting it in a single ruling.
(c) **Degenerate Hyperbola (Parallel Lines):** If the distance is less than the radius, the plane cuts the cylinder in two parallel rulings.

Thus, the classification of curves into ellipses, parabolas, and hyperbolas (and their degenerate line-pair forms) provides a complete description of all possible planar sections of quadratic surfaces of revolution.

8.3 The Dandelin Spheres

In the preceding sections, we employed algebraic methods to demonstrate that the intersection of a plane and a cone (or cylinder) yields a conic section. A more elegant, purely geometric proof was discovered by the Belgian mathematician Germinal Dandelin (1794-1847). His method constructs the foci and directrices of the conic directly using spheres inscribed within the cone or cylinder, tangent to the cutting plane.

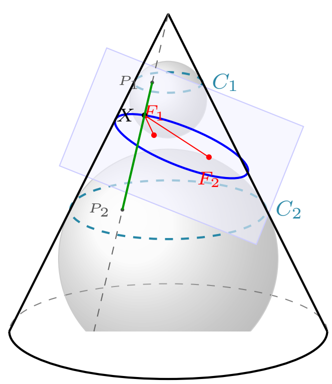


Figure 8.6: The Dandelin Spheres construction. The foci F_1 and F_2 are the points where the inscribed spheres touch the cutting plane.

Elliptical Section of a Cylinder

Consider a circular cylinder cut by a plane E that is neither parallel nor perpendicular to the cylinder's axis. We assert that the resulting curve is an ellipse.

To visualize this, imagine fitting a sphere inside the cylinder such that it is tangent to the cylinder along a circle and also tangent to the cutting plane E at a single point F_1. Since the plane is oblique, we can fit a second such sphere on the opposite side of the plane, touching E at F_2. These two spheres, known as *Dandelin spheres*, touch the cylinder along two parallel circular rings, C_1 and C_2, which lie in planes perpendicular to the cylinder's axis.

Let X be an arbitrary point on the intersection curve. Through X, there passes a unique generator (vertical line) of the cylinder. Let this generator intersect the circle C_1 at P_1 and the circle C_2 at P_2. The distance $|P_1P_2|$ is constant for all points X; it is simply the distance between the two parallel planes containing C_1 and C_2, say $2a$.

Now consider the geometry relative to the spheres.

- The segments XP_1 and XF_1 are both tangent to the first sphere originating from the point X. By the property of tangents to a sphere from an external point, their lengths are equal: $|XF_1| = |XP_1|$.
- Similarly, the segments XP_2 and XF_2 are tangent to the second sphere, implying $|XF_2| = |XP_2|$.

Summing these distances yields:

$$|XF_1| + |XF_2| = |XP_1| + |XP_2| = |P_1P_2| = 2a.$$

Since the sum of the distances from X to the fixed points F_1 and F_2 is constant, the curve is an ellipse with foci at F_1 and F_2.

Remark. If the cutting plane is perpendicular to the axis, the two spheres become tangent to the plane at the same point (the centre). The foci coincide, and the ellipse becomes a circle.

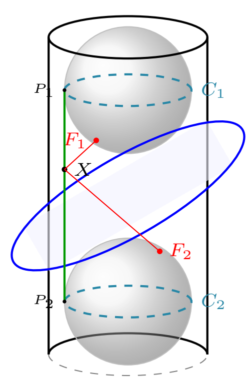


Figure 8.7: Dandelin Spheres for a Cylinder. The length $|P_1P_2|$ is constant (the vertical distance between the tangency circles), implying $|XF_1| + |XF_2|$ is constant.

Conic Sections of a Cone

The Dandelin sphere construction provides even profounder insights for the cone, revealing not just the foci but the directrices as well.

Consider a cone cut by a plane E such that the intersection is an ellipse. We inscribe two spheres within the cone: a small one near the vertex tangent to the plane at F_1, and a larger one further away tangent to the plane at F_2. These spheres touch the cone along horizontal circles C_1 and C_2.

For any point X on the ellipse, the generator passing through X intersects C_1 at P_1 and C_2 at P_2. As with the cylinder, the segments XF_1 and XP_1 are tangents to the first sphere, so $|XF_1| = |XP_1|$. Similarly, $|XF_2| = |XP_2|$. However, unlike the cylinder, the distance $|P_1P_2|$ is not constant along the generator; rather, $|P_1P_2| = |XP_1| + |XP_2| = |XF_1| + |XF_2|$ is constant because the circles C_1 and C_2 cut off a fixed segment length on every generator of the cone. Thus, the curve is an ellipse.

Finding the Directrix

Consider one of the Dandelin spheres, say the one tangent to the plane at F. It touches the cone along a circle C. The plane containing this circle C is perpendicular to the axis of the cone. The cutting plane E intersects the plane of C along a line D. We claim that D is the directrix corresponding to the focus F.

Let X be a point on the ellipse. Let the generator through X intersect the circle C at P. As established, $|XF| = |XP|$. Now, drop a perpendicular XH from X to the plane of circle C, and a perpendicular XE from X to the line D. Consider the triangle XHE. The angle $\angle XEH$ is the angle between the cutting plane E and the plane of C. Since the plane of C is perpendicular to the cone's axis, the angle between the planes is $90° - \phi$, where ϕ is the inclination of E to the axis. Thus:

$$|XH| = |XE| \sin(90° - \phi) = |XE| \cos\phi.$$

On the other hand, consider the vertical cross-section containing the generator XP. The generator makes an angle θ (the cone's half-angle) with the vertical axis. Therefore:

$$|XH| = |XP| \cos\theta.$$

Equating the two expressions for $|XH|$:

$$|XP| \cos\theta = |XE| \cos\phi \implies |XP| = |XE| \frac{\cos\phi}{\cos\theta}.$$

Since $|XF| = |XP|$, we have:

$$|XF| = |XE| \left(\frac{\cos\phi}{\cos\theta} \right).$$

This is precisely the definition of a conic section: the distance to the focus is a constant multiple of the distance to the directrix. The eccentricity is:

$$e = \frac{\cos\phi}{\cos\theta}.$$

This result aligns perfectly with our previous classification:

- If $\phi > \theta$, then $\cos\phi < \cos\theta \implies e < 1$ (Ellipse).
- If $\phi = \theta$, then $\cos\phi = \cos\theta \implies e = 1$ (Parabola).
- If $\phi < \theta$, then $\cos\phi > \cos\theta \implies e > 1$ (Hyperbola).

For the parabola and hyperbola, the sphere construction requires consideration of the unbounded nature of the cone (and the second nappe for the hyperbola), but the fundamental geometric argument linking the tangent lengths to the distances from the planes remains valid. In the case of the parabola, the "second" sphere moves to infinity, reflecting the fact that the second focus is at infinity.

8.4 Exercises

1. **The Cartesian Equation.** Consider a right circular cone with its vertex at the origin $O(0, 0, 0)$ and its axis coinciding with the z-axis.

   (a) Let the axis direction be $\mathbf{b} = [0, 0, 1]$. Using the general vector equation

   $$[(\mathbf{x} - \mathbf{a}) \cdot \mathbf{b}]^2 = \lambda^2 |\mathbf{x} - \mathbf{a}|^2,$$

   derive the standard Cartesian equation of the cone in terms of x, y, z and the half-angle θ.

   (b) Show that for any fixed $z = h$, the cross-section is a circle. Determine its radius in terms of h and θ.

2. **Vector Proof of Circular Sections.** Let K be a cone with vertex $\mathbf{a}$ and axis direction unit vector $\mathbf{b}$. Consider a cutting plane E perpendicular to the axis of the cone at a distance d from the vertex. This plane consists of points $\mathbf{x}$ satisfying $(\mathbf{x} - \mathbf{a}) \cdot \mathbf{b} = d$.

   Substitute this condition into the vector equation of the cone to prove that the intersection is a circle (a locus of points at a constant distance from a fixed centre). Find the centre and the radius of this circle.

3. **Dandelin Spheres for the Hyperbola.** The text demonstrates that for an ellipse, the sum of the distances from a point on the curve to the two foci is constant.

   (a) Consider a double-napped cone cut by a plane E parallel to the axis (i.e., $\phi = 0$), producing a hyperbola. Describe the placement of the two Dandelin spheres S_1 and S_2 relative to the two nappes of the cone.

   (b) Let X be a point on the hyperbola. Let F_1 and F_2 be the points of contact between the plane and the spheres. Let the generator passing through X touch the contact circles of the spheres at P_1 and P_2.

   (c) Using the tangency properties $|XF_1| = |XP_1|$ and $|XF_2| = |XP_2|$, prove that the absolute difference $||XF_1| - |XF_2||$ is constant for all X on the curve.

4. **Geometry of the Cylinder.** A cylinder of radius r is cut by a plane inclined at an angle α to the axis of the cylinder (where $0 < \alpha < \pi/2$). The intersection is an ellipse.

   (a) Explain why the semi-minor axis of this ellipse must be r.

   (b) By considering the geometry of the "wedge" formed by the cylinder and the plane, show that the semi-major axis is $a = r \csc \alpha$.

   (c) Hence, verify the eccentricity formula $e = \cos \alpha$ for the cylindrical limit (recalling that for a cylinder, the "half-angle" θ effectively approaches 0, but the derivation in the text for eccentricity $e = \cos \phi / \cos \theta$ requires careful interpretation; instead, use standard ellipse formulae $b^2 = a^2(1 - e^2)$).

5. **Eccentricity and Angles.** A cone has a half-angle $\theta = \pi/6$. A plane cuts this cone with an inclination angle ϕ to the axis.

   (a) Determine the specific angle ϕ required to produce a parabola.

   (b) Determine the angle ϕ required to produce a *rectangular hyperbola*. (Note: A hyperbola is rectangular if its eccentricity is $\sqrt{2}$).

6. **Focal Distance.** In the derivation of the directrix, we established that $|XF| = |XE|(\cos \phi / \cos \theta)$. Consider the specific case of a parabola, where $\phi = \theta$. Let the Dandelin sphere touch the cone along a circle C of radius R.

   (a) Show that the vertical distance from the vertex of the cone to the plane of the circle C is $R \cot \theta$.

   (b) The focus F lies on the axis of the parabola. Using the geometry of the tangent sphere, determine the distance from the vertex of the parabola to the focus F in terms of R and θ.

Chapter 9

Translation and Classification of Conics

In the preceding chapter, we derived the equations of the conic sections based on their geometric definitions relative to the origin and the coordinate axes. Such configurations are termed *standard position*. For instance, a central conic is in standard position if its centre coincides with the origin and its axes align with the coordinate axes.

However, geometric figures in the plane are rarely confined to such specific locations. To analyse conics in general positions, we must extend our algebraic framework to accommodate the motion of these curves. In this chapter, we restrict our attention to translations, which allow us to move conics while maintaining the parallelism of their axes with the coordinate axes. This investigation leads naturally to the study of the general quadratic equation lacking a cross-product term.

9.1 Parallel Translation

A parallel translation is a rigid motion of the plane that displaces every point by a constant vector.

Definition 9.1.1. ***Translation***. Let $\mathbf{h} = [h, k]$ be a fixed vector in $\mathbb{R}^2$. A translation by $\mathbf{h}$ is the mapping $T_{\mathbf{h}} : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$T_{\mathbf{h}}(\mathbf{x}) = \mathbf{x} + \mathbf{h}.$$

In terms of coordinates, if $\mathbf{x} = [x, y]$, then the image $\mathbf{x}' = [x', y']$ is given by

$$x' = x + h, \quad y' = y + k.$$

The vector $\mathbf{h}$ is called the displacement vector of the translation.

Translations possess an inverse $T_{\mathbf{h}}^{-1} = T_{-\mathbf{h}}$, defined by $\mathbf{x} \mapsto \mathbf{x} - \mathbf{h}$. Geometrically, if a curve Γ is defined by the implicit equation $f(\mathbf{x}) = 0$, the translated curve $\Gamma' = T_{\mathbf{h}}(\Gamma)$ consists of points $\mathbf{x}'$ such that $T_{\mathbf{h}}^{-1}(\mathbf{x}') \in \Gamma$. Substituting the inverse mapping yields the equation for the translated curve:

$$f(\mathbf{x}' - \mathbf{h}) = 0 \quad \text{or} \quad f(x' - h, y' - k) = 0.$$

Translation of Conics

Applying this principle to the conic sections derived previously, we obtain the equations for conics with axes parallel to the coordinate axes but centred at arbitrary points.

***Theorem* 9.1.1. Standard Forms with Translation**. Let $\mathcal{C}$ be a conic section.

(i) **Ellipse:** An ellipse with centre (h, k) and semi-axes a and b is defined by

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1.$$

(ii) **Hyperbola:** A hyperbola with centre (h, k) is defined by

$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1 \quad \text{(horizontal transverse axis)}$$

or

$$\frac{(y-k)^2}{a^2} - \frac{(x-h)^2}{b^2} = 1 \quad \text{(vertical transverse axis).}$$

(iii) **Parabola:** A parabola with vertex (h, k) is defined by

$$(y-k)^2 = 4a(x-h) \quad \text{(opens horizontally)}$$

or

$$(x-h)^2 = 4a(y-k) \quad \text{(opens vertically).}$$

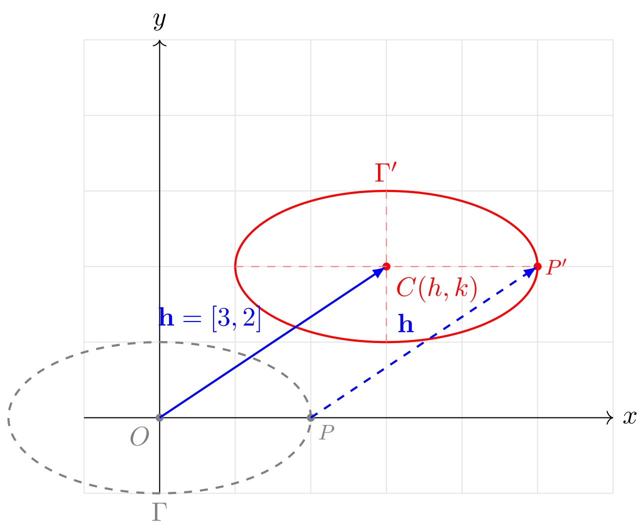


Figure 9.1: Translation of an ellipse Γ by vector $\mathbf{h}$ to position Γ'. The centre shifts from O to (h, k).

The determination of the equation of a translated conic often involves analysing the geometric properties of the focus and vertex to find the displacement vector.

Example 9.1.1. Parabola from Geometric Data. Find the equation of the parabola with focus $F(2, 3)$ and vertex $V(2, -1)$.

Solution. The axis of symmetry passes through V and F. Since both points share the x-coordinate 2, the axis is the vertical line $x = 2$. The parabola opens towards the focus, i.e., in the positive y-direction. The distance from the vertex to the focus is the parameter a:

$$a = |VF| = \sqrt{(2-2)^2 + (3-(-1))^2} = 4.$$

A standard parabola with $a = 4$ opening upwards has the equation $x^2 = 4(4)y \implies x^2 = 16y$. The vertex is shifted from $(0, 0)$ to $(2, -1)$. Applying the translation $\mathbf{h} = [2, -1]$, the equation becomes:

$$(x-2)^2 = 16(y-(-1)) \implies (x-2)^2 = 16(y+1).$$

Expanding this yields $x^2 - 4x - 16y - 12 = 0$.

Example 9.1.2. Ellipse from Geometric Data. Find the equation of the ellipse with foci at $F_1(2, -1)$ and $F_2(2, 7)$, with a major axis of length 10.

Solution. The major axis is vertical because the foci share the x-coordinate. The length of the major axis is $2a = 10$, so $a = 5$. The centre C is the midpoint of the foci:

$$C = \frac{1}{2}(F_1 + F_2) = \left(\frac{2+2}{2}, \frac{-1+7}{2}\right) = (2, 3).$$

The distance between the foci is $2c = |7 - (-1)| = 8$, so $c = 4$. For an ellipse, $b^2 = a^2 - c^2 = 5^2 - 4^2 = 25 - 16 = 9$, so $b = 3$. The standard equation for a vertical ellipse is $x^2/b^2 + y^2/a^2 = 1$. Translating the centre to $(2, 3)$:

$$\frac{(x-2)^2}{9} + \frac{(y-3)^2}{25} = 1.$$

9.2 The General Quadratic Equation

In the previous examples, we expanded the standard forms to obtain equations of the type $Ax^2 + Cy^2 + Dx + Ey + F = 0$. Conversely, we may ask: given an equation of this form, what curve does it represent?

Consider the general quadratic equation in two variables with no cross-product (xy) term:

$$Ax^2 + Cy^2 + Dx + Ey + F = 0 \tag{9.1}$$

where A and C are not both zero. To identify the curve, we employ the method of *completing the square*. This algebraic technique effectively determines the translation vector required to centre the conic (or locate its vertex) at the origin of a new coordinate system.

Completing the Square

We group the terms in x and y separately.

- If $A \neq 0$, we write $Ax^2 + Dx = A(x^2 + \frac{D}{A}x) = A\left(x + \frac{D}{2A}\right)^2 - \frac{D^2}{4A}$.
- If $C \neq 0$, we write $Cy^2 + Ey = C(y^2 + \frac{E}{C}y) = C\left(y + \frac{E}{2C}\right)^2 - \frac{E^2}{4C}$.

Substituting these into (9.1) allows us to classify the curve based on the coefficients A and C.

Example 9.2.1. Identifying a Conic. Classify the curve defined by $4x^2 - 9y^2 - 16x + 18y - 29 = 0$.

Solution. We rearrange and complete the squares:

$$4(x^2 - 4x) - 9(y^2 - 2y) = 29$$

$$4(x^2 - 4x + 4) - 9(y^2 - 2y + 1) = 29 + 4(4) - 9(1)$$

$$4(x-2)^2 - 9(y-1)^2 = 29 + 16 - 9 = 36$$

Dividing by 36:

$$\frac{(x-2)^2}{9} - \frac{(y-1)^2}{4} = 1.$$

This represents a hyperbola centred at $(2, 1)$ with a horizontal transverse axis.

Classification by Coefficients

The nature of the curve is determined primarily by the signs of the quadratic coefficients A and C.

***Theorem* 9.2.1. Classification of Conics**. The graph of the equation $Ax^2 + Cy^2 + Dx + Ey + F = 0$ is determined by the product AC:

(i) If $AC > 0$ (coefficients have the same sign), the curve is an **ellipse** or a degenerate form thereof.
(ii) If $AC < 0$ (coefficients have opposite signs), the curve is a **hyperbola** or a degenerate form thereof.
(iii) If $AC = 0$ (one coefficient is zero), the curve is a **parabola** or a degenerate form thereof.

The term "degenerate" refers to limiting cases where the conic breaks down into lines, points, or the empty set. These arise from the value of the constant term on the right-hand side (RHS) after completing the square.

Case 1: Elliptic Type ($AC > 0$) Let $A, C > 0$. The equation transforms to:

$$A(x-h)^2 + C(y-k)^2 = M.$$

- If $M > 0$, we have a real ellipse.
- If $M = 0$, the equation $A(x-h)^2 + C(y-k)^2 = 0$ implies $x = h$ and $y = k$. The graph is a single point (h, k) (a point-circle).
- If $M < 0$, the sum of non-negative terms cannot be negative. The graph is the empty set.

Case 2: Hyperbolic Type ($AC < 0$) Assume $A > 0$ and $C < 0$. Let $C = -C'$ where $C' > 0$. The equation is:

$$A(x-h)^2 - C'(y-k)^2 = M.$$

- If $M \neq 0$, we have a hyperbola.
- If $M = 0$, we have $A(x-h)^2 = C'(y-k)^2$, which leads to

$$\sqrt{A}(x-h) = \pm\sqrt{C'}(y-k).$$

  This defines two intersecting lines passing through (h, k). This corresponds to the intersection of a plane with the vertex of a double cone.

Case 3: Parabolic Type ($AC = 0$) Suppose $A \neq 0$ and $C = 0$. The equation is $Ax^2 + Dx + Ey + F = 0$. Completing the square for x:

$$A(x-h)^2 = -Ey + K.$$

- If $E \neq 0$, this is a parabola $(x-h)^2 = -\frac{E}{A}(y - \frac{K}{E})$.
- If $E = 0$, the equation is $A(x-h)^2 = K$.
  - If $K/A > 0$, we have $x - h = \pm\sqrt{K/A}$, representing two parallel vertical lines.
  - If $K = 0$, we have $(x-h)^2 = 0$, representing a single vertical line (two coincident lines).
  - If $K/A < 0$, the graph is the empty set.

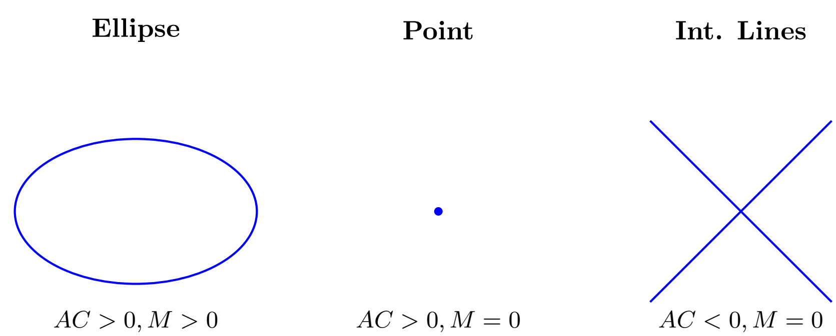


Figure 9.2: Geometric realizations of non-degenerate and degenerate quadratic equations.

The quantity $-4AC$ is often referred to as the *discriminant* of the quadratic form (excluding the linear terms). This simple sign check allows for immediate classification of the conic type before any algebraic manipulation is performed.

Remark. It is worth noting that if the term Bxy were present (i.e., $B \neq 0$), the axes of the conic would be rotated relative to the coordinate axes. In that general case, the relevant discriminant is $B^2 - 4AC$. Since we assumed $B = 0$, our discriminant simplifies to $-4AC$, consistent with the general theory.

9.3 Rotation and the General Conic

In the preceding section, we defined conic sections geometrically through the ratio of distances to a focus and a directrix, or as the intersection of a plane and a cone. Algebraically, these definitions yielded quadratic equations of the form $Ax^2 + Cy^2 + Dx + Ey + F = 0$. However, a generic polynomial of degree two in two variables includes a "cross-product" term xy:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0.$$

The presence of the term Bxy (where $B \neq 0$) indicates that the principal axes of the conic are not parallel to the Cartesian coordinate axes. To analyse such curves we must align our coordinate system with the symmetry axes of the curve. This necessitates the algebraic formalisation of rotation.

Rotation of Coordinates

We consider two distinct interpretations of rotation: rotating the points within a fixed coordinate system, or rotating the coordinate axes themselves while keeping the points fixed. In the context of simplifying algebraic equations, the latter approach (the *change of basis*), is most natural.

Let Oxy be a Cartesian coordinate system. We introduce a new system $Ox'y'$ sharing the same origin O, but rotated by an angle θ counter-clockwise relative to the original axes.

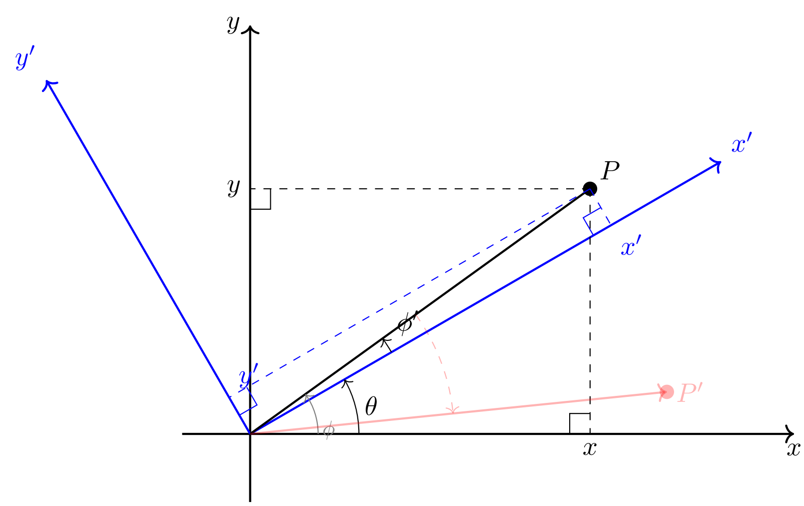


Figure 9.3: Rotation of coordinate axes by angle θ. The faded point P' shows the result of actively rotating P by $-\theta$.

Let P be a point in the plane. Its position is determined by the distance $r = |OP|$ and the angle ϕ relative to the positive x-axis. Thus, its coordinates in the original system are:

$$x = r\cos\phi, \quad y = r\sin\phi.$$

In the rotated system $Ox'y'$, the distance r remains unchanged, but the angle relative to the positive x'-axis is $\phi' = \phi - \theta$. The new coordinates (x', y') are:

$$x' = r\cos(\phi - \theta), \quad y' = r\sin(\phi - \theta).$$

To relate the two systems, we express the old coordinates in terms of the new ones. Using the angle addition formulae for sine and cosine, and noting that $\phi = \phi' + \theta$:

$$\begin{aligned} x &= r\cos(\phi' + \theta) = r(\cos\phi'\cos\theta - \sin\phi'\sin\theta) \\ y &= r\sin(\phi' + \theta) = r(\sin\phi'\cos\theta + \cos\phi'\sin\theta). \end{aligned}$$

Since $x' = r\cos\phi'$ and $y' = r\sin\phi'$, we substitute these back to obtain the transformation laws.

***Theorem* 9.3.1. Rotation of Axes**. Let the coordinate axes be rotated counter-clockwise by an angle θ about the origin. The coordinates (x, y) of a point in the original system are related to the coordinates (x', y') in the rotated system by the linear equations:

$$x = x'\cos\theta - y'\sin\theta$$
$$y = x'\sin\theta + y'\cos\theta$$

The inverse transformation is given by:

$$x' = x\cos\theta + y\sin\theta$$
$$y' = -x\sin\theta + y\cos\theta$$

This transformation allows us to rewrite any curve defined by $f(x, y) = 0$ in terms of the new variables x' and y'.

Remark. One may observe that the structure of these equations suggests a multiplication of the coordinate pair by an array of coefficients. Indeed, this is the foundation of matrix algebra, where the coefficients form an *orthogonal matrix*. While we do not rely on matrix machinery here, it is worth noting that the operation preserves the quantity $x^2 + y^2$, consistent with the geometric fact that rotation is a rigid motion (isometry) and does not alter the distance from the origin ($x^2 + y^2 = x'^2 + y'^2$).

Elimination of the Cross-Product Term

Consider the general second-degree equation:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0.$$

We aim to find a rotation angle θ such that, in the new coordinate system, the coefficient of $x'y'$ vanishes. We substitute the expressions for x and y from Theorem 9.3.1 into the quadratic terms:

$$\begin{aligned}
Ax^2 &= A(x'\cos\theta - y'\sin\theta)^2 \\
&= A(x'^2\cos^2\theta - 2x'y'\cos\theta\sin\theta + y'^2\sin^2\theta) \\
Bxy &= B(x'\cos\theta - y'\sin\theta)(x'\sin\theta + y'\cos\theta) \\
&= B(x'^2\cos\theta\sin\theta + x'y'(\cos^2\theta - \sin^2\theta) - y'^2\sin\theta\cos\theta) \\
Cy^2 &= C(x'\sin\theta + y'\cos\theta)^2 \\
&= C(x'^2\sin^2\theta + 2x'y'\sin\theta\cos\theta + y'^2\cos^2\theta)
\end{aligned}$$

The transformed equation takes the form:

$$A'x'^2 + B'x'y' + C'y'^2 + D'x' + E'y' + F = 0.$$

We are specifically interested in the new coefficient B' of the cross-product term $x'y'$. summing the relevant terms from the expansions above:

$$B' = -2A\sin\theta\cos\theta + B(\cos^2\theta - \sin^2\theta) + 2C\sin\theta\cos\theta.$$

Using the double-angle identities $\sin 2\theta = 2\sin\theta\cos\theta$ and $\cos 2\theta = \cos^2\theta - \sin^2\theta$, this simplifies to:

$$B' = (C - A)\sin 2\theta + B\cos 2\theta.$$

To eliminate the $x'y'$ term, we set $B' = 0$:

$$B\cos 2\theta = (A - C)\sin 2\theta.$$

If $A \neq C$, we divide by $(A - C)\cos 2\theta$ to obtain:

$$\tan 2\theta = \frac{B}{A - C}.$$

If $A = C$, the equation reduces to $B\cos 2\theta = 0$, which implies $2\theta = \pi/2$, or $\theta = \pi/4$. Note that the formula for $\tan 2\theta$ approaches infinity as $A \to C$, consistent with $2\theta \to \pi/2$.

***Theorem* 9.3.2. Elimination of the xy Term**. The general quadratic equation $Ax^2+Bxy+Cy^2+\cdots=0$ can be transformed into the standard form $A'x'^2+C'y'^2+\cdots=0$ by rotating the axes through an angle θ given by:

$$\cot 2\theta = \frac{A-C}{B}.$$

Once θ is determined, the curve can be classified as an ellipse, hyperbola, or parabola based on the signs of the new coefficients A' and C'. However, performing the full substitution can be laborious. Fortunately, there are quantities associated with the quadratic form that remain unchanged under rotation.

***Theorem* 9.3.3. Invariants of Rotation**. Let $Ax^2+Bxy+Cy^2$ be transformed by rotation into $A'x'^2+B'x'y'+C'y'^2$. The following quantities are invariant:

(i) **The Trace:** $A+C=A'+C'$.
(ii) **The Discriminant:** $B^2-4AC=B'^2-4A'C'$.

Proof. (i) From the expansion in the previous section:

$$A' = A\cos^2\theta + B\sin\theta\cos\theta + C\sin^2\theta$$

$$C' = A\sin^2\theta - B\sin\theta\cos\theta + C\cos^2\theta$$

Adding these yields:

$$A'+C' = A(\cos^2\theta+\sin^2\theta)+C(\sin^2\theta+\cos^2\theta) = A+C.$$

(ii) We calculate the difference $B'^2-4A'C'$. While the algebra is tedious to perform directly by substitution, we may observe that if we choose θ such that $B'=0$, the invariant relation asserts:

$$B^2-4AC=-4A'C'.$$

Since A' and C' are the coefficients in the principal axes system, their product determines the nature of the conic. The invariance of the discriminant is a fundamental property of quadratic forms (and is closely related to the determinant of the associated matrix mentioned in our earlier remark). ∎

Classification by the Discriminant

Using the discriminant invariant, we can classify the conic without explicitly finding the rotation angle. Assume we rotate the axes such that $B'=0$. The equation becomes $A'x'^2+C'y'^2+\cdots=0$. The discriminant is $-4A'C'$.

- **Ellipse:** If $B^2-4AC<0$, then $-4A'C'<0$, which implies $A'C'>0$. A' and C' have the same sign.
- **Hyperbola:** If $B^2-4AC>0$, then $-4A'C'>0$, which implies $A'C'<0$. A' and C' have opposite signs.
- **Parabola:** If $B^2-4AC=0$, then either $A'=0$ or $C'=0$ (but not both, as the quadratic terms would vanish entirely).

Example 9.3.1. Rectangular Hyperbola. Consider the curve defined by $xy=1$. Here $A=0, B=1, C=0$. The discriminant is $1^2-4(0)(0)=1>0$, confirming it is a hyperbola. To find the standard equation, we calculate the angle θ:

$$\cot 2\theta = \frac{0-0}{1} = 0 \implies 2\theta = \frac{\pi}{2} \implies \theta = \frac{\pi}{4}.$$

Substituting $\theta=\pi/4$ into the transformation equations:

$$x = \frac{x'-y'}{\sqrt{2}}, \quad y = \frac{x'+y'}{\sqrt{2}}.$$

The equation $xy=1$ becomes:

$$\left(\frac{x'-y'}{\sqrt{2}}\right)\left(\frac{x'+y'}{\sqrt{2}}\right) = 1 \implies \frac{x'^2-y'^2}{2} = 1 \implies x'^2-y'^2=2.$$

This is a hyperbola with vertices on the x'-axis (the line $y = x$) and asymptotes $x' = \pm y'$ (the coordinate axes $x = 0$ and $y = 0$).

Example 9.3.2. General Conic Reduction. Identify the curve defined by $5x^2 - 6xy + 5y^2 - 8 = 0$.

1. **Classification:** $A = 5, B = -6, C = 5$.

$$\Delta = (-6)^2 - 4(5)(5) = 36 - 100 = -64 < 0.$$

   The curve is an ellipse.
2. **Rotation Angle:**

$$\cot 2\theta = \frac{5-5}{-6} = 0 \implies \theta = \frac{\pi}{4}.$$

3. **Coefficients:** We could perform the full substitution, or use the invariants. Since $\theta = \pi/4$, we know $B' = 0$. Trace: $A' + C' = A + C = 10$. Discriminant: $-4A'C' = -64 \implies A'C' = 16$. The numbers A' and C' sum to 10 and multiply to 16. They are the roots of $z^2 - 10z + 16 = 0$, which are 2 and 8. Thus the equation in the rotated system is either $2x'^2 + 8y'^2 = 8$ or $8x'^2 + 2y'^2 = 8$. (The ambiguity corresponds to whether the major axis lies along x' or y'). Choosing the first, we divide by 8:

$$\frac{x'^2}{4} + y'^2 = 1.$$

   This describes an ellipse with semi-major axis $a = 2$ and semi-minor axis $b = 1$.

Translation of Axes

Often, a conic equation contains both cross-product terms and linear terms. The general procedure to reduce such an equation is twofold:

1. **Rotation:** Eliminate the xy term to align the axes.
2. **Translation:** Eliminate the linear terms Dx, Ey (if possible) to center the conic at the origin.

A translation corresponds to the substitution $x = u + h$ and $y = v + k$, where (h, k) is the new origin. Algebraically, this is equivalent to completing the square for the x and y terms separately.

Example 9.3.3. Full Classification. Identify the curve $16x^2 - 24xy + 9y^2 - 130x - 90y = 0$. Discriminant: $\Delta = (-24)^2 - 4(16)(9) = 576 - 576 = 0$. This is a parabola. Since $A = 16, C = 9$, we compute the rotation angle:

$$\cot 2\theta = \frac{16-9}{-24} = -\frac{7}{24}.$$

This implies $\cos 2\theta = -7/25$. Using half-angle formulae:

$$\cos\theta = \sqrt{\frac{1 + (-7/25)}{2}} = \frac{3}{5}, \quad \sin\theta = \sqrt{\frac{1 - (-7/25)}{2}} = \frac{4}{5}.$$

Substituting $x = \frac{3x'-4y'}{5}$ and $y = \frac{4x'+3y'}{5}$ into the original equation eventually yields a form $y'^2 = 2x'$, confirming the parabolic nature. (The reader is invited to verify the arithmetic).

9.4 The General Quadratic Equation

We have established that by an appropriate rotation of axes, the cross-product term Bxy in the general second-degree equation

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

can be eliminated. This transformation simplifies the equation to a form where the principal axes of the conic align with the coordinate axes, allowing for immediate classification. However, calculating the rotation angle θ and subsequently evaluating the new coefficients $A', C', D', \ldots$ is often computationally laborious.

In this section, we develop a more efficient classification method based on algebraic invariants — quantities that remain constant under coordinate rotation. We also explore a specialized algebraic technique for analyzing parabolas without resorting to trigonometric transformations.

Invariants under Rotation

Let the coordinate axes be rotated by an angle θ. As derived in the previous section, the coefficients of the transformed equation

$$A'x'^2 + B'x'y' + C'y'^2 + D'x' + E'y' + F' = 0$$

are related to the original coefficients by:

$$\begin{aligned}
A' &= A\cos^2\theta + B\sin\theta\cos\theta + C\sin^2\theta \\
B' &= B(\cos^2\theta - \sin^2\theta) + 2(C - A)\sin\theta\cos\theta \\
C' &= A\sin^2\theta - B\sin\theta\cos\theta + C\cos^2\theta
\end{aligned}$$

and the linear coefficients transform as:

$$\begin{aligned}
D' &= D\cos\theta + E\sin\theta \\
E' &= -D\sin\theta + E\cos\theta \\
F' &= F.
\end{aligned}$$

While individual coefficients change, specific combinations do not. We identify two primary invariants governing the quadratic form.

***Theorem* 9.4.1. Invariants of the Quadratic Form**. For the general quadratic equation, the following quantities are invariant under any rotation of axes:

(i) The sum of the coefficients of the squared terms: $A + C$.
(ii) The discriminant: $\Delta = B^2 - 4AC$.

Proof.

(i) Adding the expressions for A' and C':

$$A' + C' = A(\cos^2\theta + \sin^2\theta) + C(\sin^2\theta + \cos^2\theta) = A + C.$$

(ii) Rather than expanding $B'^2 - 4A'C'$ directly, consider the auxiliary expression $H = B^2 + (A - C)^2$. Using the double angle identities $\sin 2\theta = 2\sin\theta\cos\theta$ and $\cos 2\theta = \cos^2\theta - \sin^2\theta$, we can rewrite the transformation for B' and the difference $A' - C'$:

$$\begin{aligned}
B' &= B\cos 2\theta - (A - C)\sin 2\theta \\
A' - C' &= (A - C)\cos 2\theta + B\sin 2\theta.
\end{aligned}$$

Squaring and adding these two equations:

$$\begin{aligned}
B'^2 + (A' - C')^2 &= [B\cos 2\theta - (A - C)\sin 2\theta]^2 + [(A - C)\cos 2\theta + B\sin 2\theta]^2 \\
&= [B^2 + (A - C)^2](\cos^2 2\theta + \sin^2 2\theta) \\
&= B^2 + (A - C)^2.
\end{aligned}$$

Thus, $B^2 + (A - C)^2$ is invariant. Now, observe the identity:

$$B^2 - 4AC = B^2 + (A - C)^2 - (A + C)^2.$$

Since $B^2 + (A - C)^2$ and $A + C$ are both invariant, their difference $B^2 - 4AC$ must also be invariant.

■

Classification Theorem

The invariance of the discriminant $\Delta = B^2 - 4AC$ allows us to classify the conic solely from the original coefficients. If we rotate the axes to eliminate the cross-product term ($B' = 0$), the discriminant becomes $\Delta = -4A'C'$. The sign of Δ therefore determines the relative signs of A' and C'.

***Theorem* 9.4.2. Classification of Conics**. The conic defined by $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$ is determined by the discriminant $\Delta = B^2 - 4AC$ as follows:

1. **Ellipse:** If $\Delta < 0$. (Includes circles and degenerate point-ellipses).
2. **Parabola:** If $\Delta = 0$. (Includes degenerate parallel lines).
3. **Hyperbola:** If $\Delta > 0$. (Includes degenerate intersecting lines).

Proof. If $B' = 0$, the equation is $A'x'^2 + C'y'^2 + \cdots = 0$.

- If $\Delta < 0$, then $-4A'C' < 0 \implies A'C' > 0$. A' and C' have the same sign, defining an ellipse.
- If $\Delta > 0$, then $-4A'C' > 0 \implies A'C' < 0$. A' and C' have opposite signs, defining a hyperbola.
- If $\Delta = 0$, then $-4A'C' = 0$. Since A' and C' cannot both be zero (otherwise the equation is not quadratic), exactly one is zero, defining a parabola. ∎

Analysis of the Parabolic Case

When the discriminant vanishes ($B^2-4AC = 0$), the quadratic terms $Ax^2+Bxy+Cy^2$ form a perfect square. This algebraic peculiarity allows us to analyze the parabola without performing a rotation of coordinates.

If $A, C > 0$, then $B = \pm 2\sqrt{AC}$, and we may write:

$$Ax^2 + Bxy + Cy^2 = (\sqrt{A}x \pm \sqrt{C}y)^2.$$

The general equation can thus be rearranged as:

$$(\sqrt{A}x \pm \sqrt{C}y)^2 = -(Dx + Ey + F).$$

If the linear expression on the right-hand side represents a line perpendicular to the line $\sqrt{A}x \pm \sqrt{C}y = 0$, this equation is already in the standard definition form (distance to axis squared equals constant times distance to tangent at vertex). If not, we may introduce an arbitrary constant to adjust the linear terms.

Example 9.4.1. Parabola Decomposition. Identify the vertex and axis of the conic $x^2-4xy+4y^2-4x-2y = 0$.

1. **Check Discriminant:** $A = 1, B = -4, C = 4$.
$$\Delta = (-4)^2 - 4(1)(4) = 16 - 16 = 0.$$
The conic is a parabola.
2. **Factor Quadratic Terms:**
$$(x - 2y)^2 = 4x + 2y.$$
3. **Check Orthogonality:** We have two implicit lines: $L_1 : x - 2y = 0$ and $L_2 : 4x + 2y = 0$. The normal vectors are $\mathbf{n}_1 = [1, -2]$ and $\mathbf{n}_2 = [4, 2]$. Their dot product is $4 - 4 = 0$. Thus, the lines are orthogonal. We can define a new coordinate system aligned with these lines. Let $Y = \frac{x-2y}{\sqrt{5}}$ and $X = \frac{2x+y}{\sqrt{5}}$ (normalizing the vectors). Note that L_2 can be written as $2(2x+y)$. The equation becomes:
$$(\sqrt{5}Y)^2 = 2(\sqrt{5}X) \implies 5Y^2 = 2\sqrt{5}X \implies Y^2 = \frac{2}{\sqrt{5}}X.$$
4. **Result:** The curve is a parabola.
   - **Axis:** $Y = 0 \implies x - 2y = 0$.

- **Tangent at Vertex:** $X = 0 \implies 2x + y = 0$.
- **Vertex:** The intersection of the axis and tangent, which is the origin $(0, 0)$.

However, it is not always the case that the resulting linear terms are immediately orthogonal. Consider the following more complex case where we must use the method of undetermined coefficients.

Example 9.4.2. Method of Undetermined Coefficients. Analyze the curve $x^2+2xy+y^2-2x-10y+5 = 0$.

1. **Check Discriminant:** $B^2 - 4AC = 2^2 - 4(1)(1) = 0$. It is a parabola.
2. **Initial Factoring:**
$$(x + y)^2 = 2x + 10y - 5.$$
The line $x + y = 0$ (normal $[1, 1]$) is not perpendicular to $2x + 10y - 5 = 0$ (normal $[2, 10]$).
3. **Introduce Parameter:** We add a constant λ inside the square term. To maintain equality, we subtract the generated terms on the other side:
$$(x + y + \lambda)^2 = (x + y)^2 + 2\lambda(x + y) + \lambda^2.$$
Substituting the conic equation $(x + y)^2 = 2x + 10y - 5$ into the RHS:
$$(x + y + \lambda)^2 = (2x + 10y - 5) + 2\lambda x + 2\lambda y + \lambda^2.$$
Grouping the linear terms:
$$(x + y + \lambda)^2 = 2(1 + \lambda)x + 2(5 + \lambda)y + (\lambda^2 - 5).$$
4. **Force Orthogonality:** We require the line defined by the LHS, $x+y+\lambda = 0$ (normal $\mathbf{n}_1 = [1, 1]$), to be perpendicular to the line defined by the RHS (normal $\mathbf{n}_2 = [2(1 + \lambda), 2(5 + \lambda)]$).
$$\mathbf{n}_1 \cdot \mathbf{n}_2 = 0 \implies 1 \cdot 2(1 + \lambda) + 1 \cdot 2(5 + \lambda) = 0.$$
$$2 + 2\lambda + 10 + 2\lambda = 0 \implies 4\lambda = -12 \implies \lambda = -3.$$
5. **Substitute $\lambda = -3$:** LHS: $(x+y-3)^2$. RHS: $2(1-3)x+2(5-3)y+(9-5) = -4x+4y+4 = 4(-x+y+1)$. The equation becomes:
$$(x + y - 3)^2 = 4(-x + y + 1).$$
6. **Interpretation:** Let the axis be the line $L_A : x + y - 3 = 0$. Let the tangent at the vertex be $L_T : -x + y + 1 = 0$. The distance from a point to L_A is $Y = \frac{|x+y-3|}{\sqrt{2}}$. The distance from a point to L_T is $X = \frac{|-x+y+1|}{\sqrt{2}}$. Substituting these into the equation $(x + y - 3)^2 = 4(-x + y + 1)$:
$$(\sqrt{2}Y)^2 = 4(\sqrt{2}X) \implies 2Y^2 = 4\sqrt{2}X \implies Y^2 = 2\sqrt{2}X.$$
The length of the semi-latus rectum is $2\sqrt{2}$. The **vertex** is the intersection of L_A and L_T: Adding the equations: $2y - 2 = 0 \implies y = 1$. Substituting back: $x + 1 - 3 = 0 \implies x = 2$. Vertex: $V(2, 1)$.

Example 9.4.3. Degenerate Parabola. Analyze $8x^2 + 24xy + 18y^2 - 14x - 21y + 3 = 0$.

1. **Discriminant:** $24^2 - 4(8)(18) = 576 - 576 = 0$. Parabolic type.
2. **Factoring:** The quadratic part $2(4x^2 + 12xy + 9y^2) = 2(2x + 3y)^2$. The equation is $2(2x + 3y)^2 - 7(2x + 3y) + 3 = 0$.
3. **Solution:** This is a quadratic in the variable $u = 2x + 3y$:
$$2u^2 - 7u + 3 = 0 \implies (2u - 1)(u - 3) = 0.$$
The locus consists of two parallel lines:
$$4x + 6y - 1 = 0 \quad \text{and} \quad 2x + 3y - 3 = 0.$$

This algebraic method for parabolas is elegant and direct. For ellipses and hyperbolas, while the discriminant quickly identifies the type, finding the precise axes and foci usually requires the calculation of the rotation angle θ or the eigenvalues of the associated quadratic form, as discussed in the general theory.

9.5 Tangents to Conic Sections

For a circle, the tangent at a point P is the unique line passing through P perpendicular to the radius OP. This elegant geometric definition relies on the constancy of the radius, a property not shared by other conics. For a general non-degenerate conic, we must adopt a limiting process to define the tangent.

Consider a point P on a conic curve C. Let Q be another point on the curve distinct from P. The line connecting P and Q is a secant. As the point Q moves along the curve and approaches P, the secant line approaches a limiting position. This limiting line is defined as the tangent to the curve at P.

The Slope of the Tangent

Algebraically, if the curve is described by a function $y = f(x)$ (or implicitly), the slope of the tangent at $P(x_0, y_0)$ is the limit of the slope of the secant PQ as $Q \to P$. If $Q = (x_0 + h, y_0 + k)$, the slope is the limit of k/h as $h \to 0$.

Example 9.5.1. Tangent to a Parabola. Consider the parabola $x^2 = 4ay$. Let $P(x_0, y_0)$ be a point on the curve. Let $Q(x_0 + h, y_0 + k)$ be a nearby point on the curve. Since both lie on the parabola:

$$x_0^2 = 4ay_0 \quad \text{and} \quad (x_0 + h)^2 = 4a(y_0 + k).$$

Expanding the second equation:

$$x_0^2 + 2x_0h + h^2 = 4ay_0 + 4ak.$$

Subtracting the first equation from this yields:

$$2x_0h + h^2 = 4ak \implies k = \frac{2x_0h + h^2}{4a}.$$

The slope of the secant PQ is:

$$m_{sec} = \frac{k}{h} = \frac{2x_0 + h}{4a}.$$

Taking the limit as $h \to 0$, the slope of the tangent at P is:

$$m = \frac{2x_0}{4a} = \frac{x_0}{2a}.$$

The equation of the tangent line is $y - y_0 = m(x - x_0)$. Substituting m:

$$y - y_0 = \frac{x_0}{2a}(x - x_0) \implies 2ay - 2ay_0 = x_0x - x_0^2.$$

Since $x_0^2 = 4ay_0$, we substitute this into the equation:

$$2ay - 2ay_0 = x_0x - 4ay_0 \implies 2ay + 2ay_0 = x_0x.$$

Dividing by the common factor gives the symmetric form:

$$x_0x = 2a(y + y_0).$$

This derivation generalises to all standard conics. The equation of the tangent at a point (x_0, y_0) can be obtained from the equation of the conic by applying the following "splitting" substitutions:

- Replace x^2 with x_0x.
- Replace y^2 with y_0y.
- Replace x with $\frac{1}{2}(x + x_0)$.
- Replace y with $\frac{1}{2}(y + y_0)$.

***Theorem* 9.5.1. Tangents at a Point**. The equation of the tangent line to a standard conic at the point $P(x_0, y_0)$ is given by:

Conic	**Equation**	**Tangent at** (x_0, y_0)
Parabola	$y^2 = 4ax$	$y_0 y = 2a(x + x_0)$
Parabola	$x^2 = 4ay$	$x_0 x = 2a(y + y_0)$
Ellipse	$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$	$\frac{x_0 x}{a^2} + \frac{y_0 y}{b^2} = 1$
Hyperbola	$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$	$\frac{x_0 x}{a^2} - \frac{y_0 y}{b^2} = 1$
Circle	$x^2 + y^2 = r^2$	$x_0 x + y_0 y = r^2$

Tangents with Given Slope

A common problem is to find the tangent(s) to a conic that are parallel to a given direction. Let m be the specified slope. We seek a line $y = mx + c$ that intersects the conic at exactly one point (tangency).

Example 9.5.2. Parabola. Find the tangent to $y^2 = 4ax$ with slope m. Substitute $y = mx + c$ into the parabola equation:

$$(mx + c)^2 = 4ax \implies m^2x^2 + (2mc - 4a)x + c^2 = 0.$$

For tangency, this quadratic in x must have a zero discriminant (a repeated root):

$$(2mc - 4a)^2 - 4(m^2)(c^2) = 0.$$

$$4m^2c^2 - 16amc + 16a^2 - 4m^2c^2 = 0.$$

$$-16amc + 16a^2 = 0 \implies amc = a^2.$$

Assuming $a \neq 0$ and $m \neq 0$, we find $c = a/m$. Thus, the tangent with slope m is:

$$y = mx + \frac{a}{m}.$$

For central conics (ellipse and hyperbola), the quadratic condition yields two solutions for c, corresponding to two parallel tangents on opposite sides of the centre.

***Theorem* 9.5.2. Slope Forms of Tangents**. Let m be a real number.

(i) The parabola $y^2 = 4ax$ has a unique tangent with slope $m \neq 0$:

$$y = mx + \frac{a}{m}.$$

(ii) The ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ has two tangents with slope m:

$$y = mx \pm \sqrt{a^2m^2 + b^2}.$$

(iii) The hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ has tangents with slope m only if $|m| > b/a$:

$$y = mx \pm \sqrt{a^2m^2 - b^2}.$$

Remark. For the hyperbola, if $|m| < b/a$, the line intersects the curve but is not tangent (or does not intersect at all, depending on c). If $|m| = b/a$, the line is parallel to an asymptote and intersects the hyperbola at only one point (at infinity, projectively speaking), but is strictly an asymptote, not a tangent in the affine sense.

Tangents from an External Point

To find the tangents from a point $Q(x_1, y_1)$ not on the curve, one can assume the equation of the tangent is $y - y_1 = m(x - x_1)$ and apply the condition of tangency to determine m. This usually leads to a quadratic equation in m, yielding two tangents (real, coincident, or imaginary).

Alternatively, one may use the "chord of contact" method. If tangents from $Q(x_1, y_1)$ touch the conic at P_1 and P_2, the line P_1P_2 is linear in (x_1, y_1) and has the same form as the tangent equation ($T = 0$).

Example 9.5.3. Tangents to an Ellipse. Find the equations of the tangents to the ellipse $x^2 + 2y^2 = 3$ passing through $Q(-1, 2)$. Let the tangent have slope m. Its equation is $y - 2 = m(x + 1) \implies y = mx + (m + 2)$. The condition for tangency to $x^2 + 2y^2 = 3$ (rewritten as $\frac{x^2}{3} + \frac{y^2}{3/2} = 1$) is $c^2 = a^2m^2 + b^2$. Here $a^2 = 3, b^2 = 3/2, c = m + 2$.

$$(m + 2)^2 = 3m^2 + \frac{3}{2}.$$
$$m^2 + 4m + 4 = 3m^2 + 1.5.$$
$$2m^2 - 4m - 2.5 = 0 \implies 4m^2 - 8m - 5 = 0.$$
$$(2m + 1)(2m - 5) = 0.$$

The slopes are $m = -1/2$ and $m = 5/2$. Substitute back into the line equation: 1. $y - 2 = -0.5(x + 1) \implies 2y - 4 = -x - 1 \implies x + 2y - 3 = 0$. 2. $y - 2 = 2.5(x + 1) \implies 2y - 4 = 5x + 5 \implies 5x - 2y + 9 = 0$.

Reflection Properties

The geometric relationship between the tangent and the foci of a conic gives rise to optical reflection properties.

***Theorem* 9.5.3. Reflection Properties.**

(i) **Parabola:** The tangent at any point P bisects the angle between the focal radius FP and the line through P parallel to the axis. Consequently, rays originating from the focus are reflected parallel to the axis.

(ii) **Ellipse:** The tangent at a point P bisects the external angle between the focal radii F_1P and F_2P. Equivalently, the normal bisects the internal angle $\angle F_1PF_2$. Rays from one focus reflect off the ellipse to pass through the other focus.

(iii) **Hyperbola:** The tangent at P bisects the internal angle $\angle F_1PF_2$. Rays directed towards one focus reflect off the hyperbola towards the other.

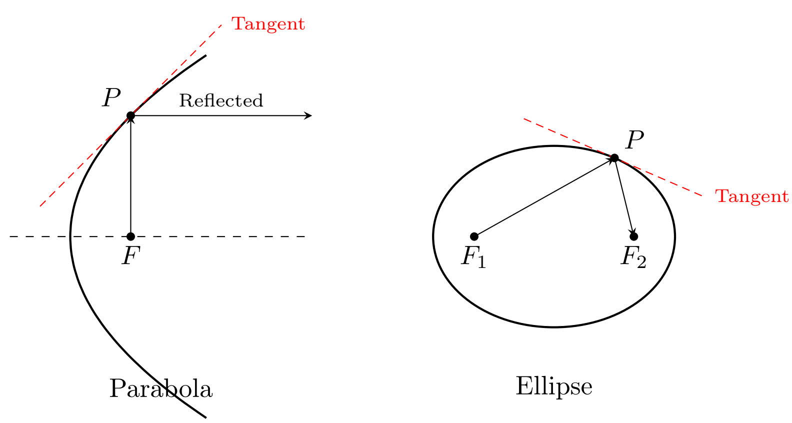


Figure 9.4: Reflection properties of the Parabola and Ellipse.

These properties are fundamental to the design of optical instruments, from parabolic antennas and car headlights to whispering galleries and telescopes.

9.6 Exercises

The following exercises explore the algebraic and geometric properties of conic sections. Unless otherwise stated, assume all coordinates refer to a standard Cartesian system.

Part I: Translations and Standard Forms

1. **Classification by Completion.** Determine the type of conic represented by the following equations. If the conic is non-degenerate, find its centre (or vertex) and the lengths of its semi-axes (or the parameter a).
   (a) $9x^2 + 4y^2 - 18x + 16y - 11 = 0$
   (b) $y^2 - 4y - 8x + 20 = 0$
   (c) $3x^2 - 2y^2 + 6x + 8y - 11 = 0$
   (d) $x^2 + y^2 - 4x + 6y + 13 = 0$

2. **Geometric Construction.** Find the Cartesian equation of the conic described by the following geometric data:
   (a) A parabola with focus at $(3, 2)$ and directrix $x = -1$.
   (b) An ellipse with foci at $(1, 3)$ and $(1, 9)$, passing through the point $(4, 6)$.
   (c) A hyperbola with vertices at $(0, 2)$ and $(4, 2)$, and asymptotes with slopes ± 2.

3. **Rotation of Axes.** Consider the curve defined by $5x^2 - 4xy + 8y^2 = 36$.
   (a) Calculate the rotation angle θ required to eliminate the xy term.
   (b) Transform the equation into the $x'y'$ system.
   (c) Identify the curve and determining the lengths of its semi-axes.

4. **Tangent Construction.**
   (a) Find the equations of the tangents to the parabola $y^2 = 12x$ that pass through the point $(-3, 0)$.
   (b) Find the equations of the tangents to the ellipse $x^2 + 4y^2 = 16$ that are parallel to the line $y = x$.

5. **The Normal Line.** The normal to a curve at a point P is the line perpendicular to the tangent at P. Find the equation of the normal to the hyperbola $x^2 - y^2 = 16$ at the point $(5, 3)$. Show that this normal line does not pass through the origin.

Part II: General Theory and Geometry

6. **The Director Circle.**
   (a) Let $y = mx + \sqrt{a^2m^2 + b^2}$ be a tangent to the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Write down the equation of the tangent with slope $-1/m$ (perpendicular to the first).
   (b) By eliminating m between these two equations, prove that the locus of the intersection of perpendicular tangents is the circle $x^2 + y^2 = a^2 + b^2$.

7. **Product of Perpendiculars.** Let P be any point on the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Let d_1 and d_2 be the perpendicular distances from the foci F_1 and F_2 to the tangent at P. Prove that $d_1 d_2 = b^2$.

   Remark. Use the slope form of the tangent $y = mx + \sqrt{a^2m^2 + b^2}$ and the coordinates of the foci $(\pm ae, 0)$.

8. **Algebraic Invariants.** Consider the general quadratic form $Q(x, y) = Ax^2 + Bxy + Cy^2$. Let the axes be rotated by an angle θ such that the new form is $A'x'^2 + B'x'y' + C'y'^2$. Without using the explicit expressions for A', B', C', prove algebraically that if $B^2 - 4AC = B'^2 - 4A'C'$, then the nature of the roots of the characteristic equation $\lambda^2 - (A + C)\lambda + (AC - B^2/4) = 0$ is preserved. Hence, conclude that the eigenvalues of the quadratic form are invariant under rotation.

9. **Area of the General Ellipse.** Let $Ax^2 + Bxy + Cy^2 = 1$ define an ellipse (so $B^2 - 4AC < 0$ and $A + C > 0$).

(a) Let λ_1 and λ_2 be the coefficients of x'^2 and y'^2 in the standard form obtained by rotation. Express $\lambda_1\lambda_2$ in terms of the discriminant of the original equation.
(b) Recall that the area of an ellipse with semi-axes a and b is πab. Show that the area of the ellipse defined by the general equation is given by:

$$\text{Area} = \frac{2\pi}{\sqrt{4AC - B^2}}.$$

10. Tangents to the Parabola. Let $P(at^2, 2at)$ and $Q(as^2, 2as)$ be two points on the parabola $y^2 = 4ax$.

(a) Show that the tangents at P and Q intersect at the point $R(ast, a(t + s))$.
(b) Prove that if the tangents at P and Q are perpendicular, the intersection point R lies on the directrix of the parabola.

11. Midpoints of Chords. Consider the parabola $y^2 = 4ax$. A set of parallel chords has slope m.

(a) Let a chord intersect the parabola at points (x_1, y_1) and (x_2, y_2). Show that the ordinate of the midpoint is $k = \frac{y_1+y_2}{2}$.
(b) By considering the relationship $y_1^2 - y_2^2 = 4a(x_1 - x_2)$, show that the locus of the midpoints of these chords is the horizontal line $y = \frac{2a}{m}$.

12. Degenerate Conics. Consider the equation $2x^2 + 3xy - 2y^2 - x + 3y - 1 = 0$.

(a) Calculate the discriminant of the quadratic terms. What does this suggest about the conic type?
(b) Attempt to factorise the equation into two linear factors of the form $(ax+by+c)(dx+ey+f) = 0$.
(c) Hence, describe the graph of the equation geometrically.

13. The Rectangular Hyperbola. A hyperbola is called rectangular if its asymptotes are perpendicular.

(a) Show that the hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ is rectangular if and only if $a = b$.
(b) Prove that for a general equation $Ax^2 + Bxy + Cy^2 + \cdots = 0$, the condition for the hyperbola to be rectangular is $A + C = 0$.

14. Common Tangents. Find the equations of the common tangents to the circle $x^2 + y^2 = 8$ and the parabola $y^2 = 16x$.

Remark. Let the tangent to the parabola be $y = mx + 4/m$. Determine m such that the distance from the origin to this line equals the radius of the circle.

15. Focal Chords. A chord of a conic that passes through a focus is called a focal chord. Let PQ be a focal chord of the parabola $y^2 = 4ax$. Let SP and SQ denote the segment lengths from the focus S to the curve. Prove that the semi-latus rectum $2a$ is the harmonic mean of the focal segments:

$$\frac{1}{SP} + \frac{1}{SQ} = \frac{1}{a}.$$

Part III: Advanced Geometric Properties

16. The Chord of Contact. From a point $P(x_1, y_1)$ external to the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, two tangents PA and PB are drawn, touching the ellipse at points A and B.

(a) Assume the contact points are $A(h, k)$ and $B(u, v)$. Write down the equations of the tangent lines at these points.
(b) Since both lines pass through $P(x_1, y_1)$, substitute these coordinates into the tangent equations.
(c) Observe the structure of the resulting equations to prove that the line segment AB (the chord of contact) lies on the line:

$$\frac{x_1 x}{a^2} + \frac{y_1 y}{b^2} = 1.$$

(d) **Generalisation:** This equation is identical to the tangent equation, but P is not on the curve. This line is called the *polar* of P, and P is the *pole*. Prove that if the polar of P passes through a point Q, then the polar of Q must pass through P.

Chapter 10

Quadric Surfaces

In the plane, a linear equation in two variables defines a line, while in space, a linear equation in three variables defines a plane. A line in space is subsequently defined as the intersection of two planes. Moving to the second degree, a quadratic equation in two variables defines a curve in the plane known as a conic section, the geometry of which is derived from the plane sections of a circular cone.

The natural extension of this hierarchy to three dimensions leads us to the study of *quadric surfaces*. A general quadratic equation in three variables,

$$Ax^2 + By^2 + Cz^2 + Dxy + Eyz + Fzx + Gx + Hy + Iz + J = 0,$$

defines a surface in space. Just as plane sections of a cone yield conics, the intersection of a plane with a quadric surface is defined by a quadratic equation in two variables (after coordinate transformation), and is therefore a conic section. This property provides a powerful tool for visualising and analysing the shape of these surfaces.

While there are only three primary types of non-degenerate conics (ellipses, parabolas, and hyperbolas), the classification of quadric surfaces is richer. In this chapter, we classify these surfaces, beginning with those possessing rotational symmetry.

10.1 Surfaces of Revolution

The simplest method of generating a surface with symmetry is to rotate a plane curve about a fixed line lying in the same plane. This line is termed the *axis of revolution*, and the resulting locus is a *surface of revolution*.

Definition 10.1.1. ***Surface of Revolution***. Let C be a curve in the yz-plane defined by the equation $f(y,z) = 0$. The surface generated by rotating C about the z-axis consists of all points $P(x,y,z)$ such that the point $P'(0,r,z)$ lies on C, where $r = \sqrt{x^2+y^2}$ is the perpendicular distance from P to the axis. The equation of the surface is therefore:

$$f\left(\sqrt{x^2+y^2}, z\right) = 0 \quad \text{or} \quad f\left(-\sqrt{x^2+y^2}, z\right) = 0.$$

More compactly, we substitute y^2 with $x^2 + y^2$ in the equation of the curve if f depends only on y^2.

We have already encountered two elementary examples:

1. **The Circular Cylinder:** Rotating the line $y = r$ about the z-axis yields $x^2 + y^2 = r^2$.
2. **The Circular Cone:** Rotating the line $y = az$ about the z-axis yields $x^2 + y^2 = a^2z^2$.

We now apply this construction to the conic sections, assuming the axis of rotation is an axis of symmetry of the conic. The resulting surfaces are known as *quadric surfaces of revolution*.

The Sphere and Spheroids

Rotating a circle $y^2 + z^2 = r^2$ about the z-axis (a diameter) yields the sphere $x^2 + y^2 + z^2 = r^2$. Every plane section is a circle or a point.

If we rotate an ellipse instead of a circle, we obtain an *ellipsoid of revolution*, often called a spheroid. There are two distinct cases depending on which axis of the ellipse is chosen as the axis of revolution. Consider an ellipse in the yz-plane with semi-axes a and b.

Prolate Spheroid Let the ellipse be defined by $\frac{y^2}{b^2} + \frac{z^2}{a^2} = 1$ with $a > b$. The major axis lies along the z-axis. Rotating this curve about the z-axis, we replace y^2 with $x^2 + y^2$:

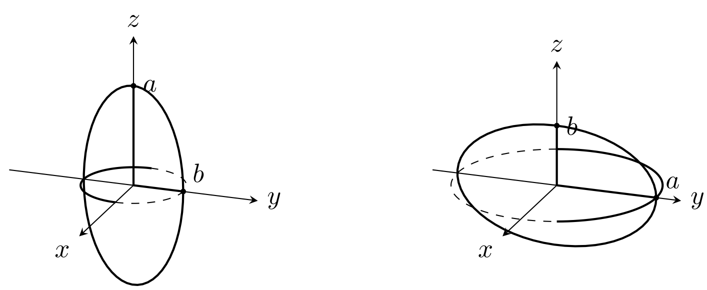
$$\frac{x^2 + y^2}{b^2} + \frac{z^2}{a^2} = 1.$$

The resulting surface is elongated like a rugby ball.

Oblate Spheroid Let the ellipse be defined by $\frac{y^2}{a^2} + \frac{z^2}{b^2} = 1$ with $a > b$. Here, the minor axis lies along the z-axis. Rotation yields:

$$\frac{x^2 + y^2}{a^2} + \frac{z^2}{b^2} = 1.$$

This surface is flattened at the poles, resembling the Earth or a discus.



Prolate ($a > b$) Oblate ($a > b$)

Figure 10.1: Ellipsoids of revolution. The prolate spheroid is generated by rotating an ellipse about its major axis; the oblate by rotating about its minor axis.

In both cases, sections perpendicular to the z-axis are circles (parallels), while sections containing the z-axis are ellipses (meridians).

Hyperboloids of Revolution

Similarly, rotating a hyperbola about an axis of symmetry produces a hyperboloid. Let the hyperbola be centred at the origin.

Hyperboloid of Two Sheets Consider the hyperbola $\frac{z^2}{a^2} - \frac{y^2}{b^2} = 1$ in the yz-plane. The curve intersects the z-axis at $\pm a$ but does not intersect the y-axis. The z-axis is the *transverse axis*. Rotating about this axis yields:

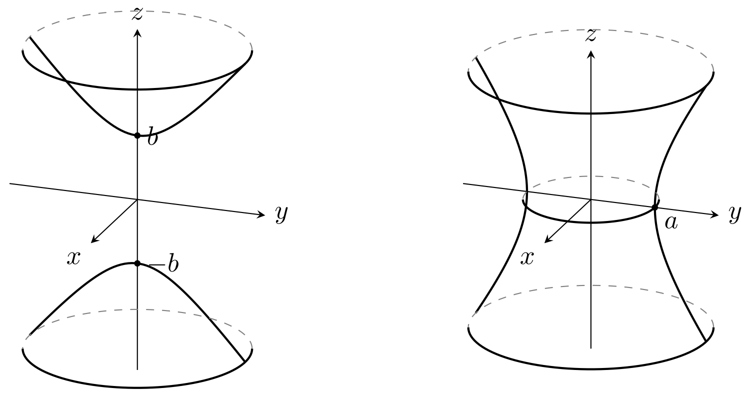
$$\frac{z^2}{a^2} - \frac{x^2 + y^2}{b^2} = 1.$$

The surface consists of two disjoint congruent parts (sheets) separated by the region $-a < z < a$.

Hyperboloid of One Sheet Consider the hyperbola $\frac{y^2}{a^2} - \frac{z^2}{b^2} = 1$. The curve intersects the y-axis but not the z-axis. The z-axis is the *conjugate axis*. Rotation about the z-axis yields:

$$\frac{x^2 + y^2}{a^2} - \frac{z^2}{b^2} = 1.$$

This surface is connected and extends infinitely in the z-direction. It is a ruled surface, famously used in the construction of cooling towers due to its structural stability.



Two Sheets One Sheet

Figure 10.2: Hyperboloids of revolution generated by rotating hyperbolas about the z-axis.

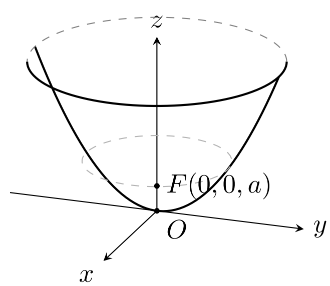
The ellipsoids and hyperboloids possess a centre of symmetry (the origin in these forms) and are collectively termed *central quadrics*.

Paraboloid of Revolution

A parabola has only one axis of symmetry. Consider the parabola $y^2 = 4az$ in the yz-plane. Rotating this about the z-axis yields the *paraboloid of revolution*:

$$x^2 + y^2 = 4az.$$

This surface does not have a centre of symmetry. It is the shape employed in satellite dishes and headlight reflectors due to the property that rays parallel to the axis reflect to the focus.



Paraboloid of Revolution

Figure 10.3: Paraboloid of revolution generated by rotating the parabola $y^2 = 4az$ about the z-axis.

Remark. Not all surfaces generated by rotating curves are quadrics. For instance, rotating a circle $(y - b)^2 + z^2 = a^2$ (where $b > a$) about the z-axis generates a *torus*. The equation is derived by substituting $y = \sqrt{x^2 + y^2}$:

$$(\sqrt{x^2 + y^2} - b)^2 + z^2 = a^2 \implies x^2 + y^2 + z^2 + b^2 - a^2 = 2b\sqrt{x^2 + y^2}.$$

Squaring to remove the radical yields a quartic (degree 4) equation:

$$(x^2 + y^2 + z^2 + b^2 - a^2)^2 = 4b^2(x^2 + y^2).$$

Thus, the torus is not a quadric surface.

10.2 Cylinders and Cones

The circular cylinder, previously defined as a surface of revolution, may be viewed more generally as the locus of a moving line. If a straight line moves along a fixed curve while remaining parallel to a fixed direction, it generates a *cylinder*.

Definition 10.2.1. ***Quadric Cylinder***. Let Γ be a non-degenerate conic in the xy-plane, and let L be a line through a point on Γ parallel to the z-axis. The surface generated by all such lines is a quadric cylinder. Γ is termed the *directrix*, and the lines are the *rulings* (or generators) of the cylinder.

Since the rulings are parallel to the z-axis, the height z does not constrain the coordinates x and y. Consequently, the equation of the cylinder in $\mathbb{R}^3$ is identical to the equation of its directrix in $\mathbb{R}^2$. Depending on the nature of Γ, we classify quadric cylinders into three types:

(i) **Elliptic Cylinder:** $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.
(ii) **Hyperbolic Cylinder:** $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$.
(iii) **Parabolic Cylinder:** $y^2 = 4ax$.

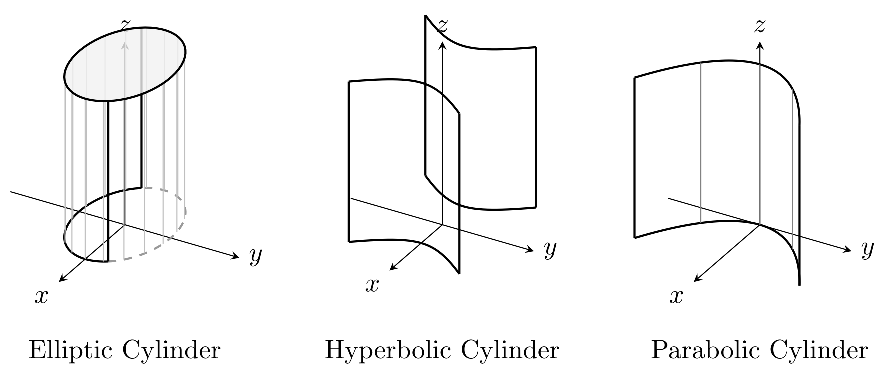


Figure 10.4: The three non-degenerate quadric cylinders viewed in 3D perspective.

Any plane parallel to the z-axis that passes through an axis of symmetry of the directrix is a plane of symmetry for the cylinder. Furthermore, any plane perpendicular to the rulings ($z =$ constant) is a plane of symmetry.

***Theorem* 10.2.1. Plane Sections of Cylinders**. A plane not parallel to the rulings of a quadric cylinder intersects it in a conic of the same type as the directrix. Furthermore, parallel plane sections are congruent conics.

Proof. Let the cylinder Q have a central directrix Γ in the xy-plane defined by $Ax^2 + Bxy + Cy^2 + F = 0$. Let E be a plane intersecting the cylinder at an angle $\alpha \in [0, \pi/2)$. We align the x-axis with the line of intersection of E and the xy-plane.

If we project points $P(x, y, z)$ on the cylinder section Γ' onto the xy-plane, we obtain the directrix Γ. Let (x', y') be coordinates in E. The relationship between the coordinates of a point on Γ' and its projection on Γ is $x = x'$ and $y = y' \cos\alpha$. Substituting these into the equation for Γ:

$$A(x')^2 + B(x')(y' \cos\alpha) + C(y' \cos\alpha)^2 + F = 0.$$

The discriminant of this new quadratic is $D' = (B\cos\alpha)^2 - 4A(C\cos^2\alpha) = (B^2 - 4AC)\cos^2\alpha$. Since $\cos^2\alpha > 0$, the sign of the discriminant is invariant, preserving the conic type. For parallel planes, the equations are identical up to a translation of coordinates, implying congruence. ■

Quadric Cones A general quadric cone is the surface generated by lines connecting a fixed vertex V to points on a non-degenerate conic Γ. Unlike the circular cone, the section perpendicular to the axis of symmetry (if one exists) may be an ellipse.

***Theorem* 10.2.2. Homogeneity of Cones**. A quadric surface is a cone with its vertex at the origin if and only if its equation is homogeneous of degree two.

Proof. A surface is a cone with vertex O if, for every point P on the surface, the entire line OP lies on the surface. Algebraically, if (x, y, z) satisfies the equation $f(x, y, z) = 0$, then (tx, ty, tz) must also satisfy it for all $t \in \mathbb{R}$. A general quadric equation is $f(x, y, z) = \sum a_{ij}x_ix_j + \sum b_ix_i + c = 0$. If f is homogeneous of degree two, then $f(tx, ty, tz) = t^2 f(x, y, z)$. Thus $f(x, y, z) = 0 \implies f(tx, ty, tz) = 0$, satisfying the condition. Conversely, if the line property holds, then $\sum a_{ij}(tx_i)(tx_j) + \sum b_i(tx_i) + c = 0$ for all t. This polynomial in t, $t^2(\sum a_{ij}x_ix_j) + t(\sum b_ix_i) + c = 0$, can only vanish for all t if the coefficients of t^2, t^1, and t^0 are zero. Since this must hold for any point on the surface, it implies the linear and constant terms must be zero, rendering the equation homogeneous. ■

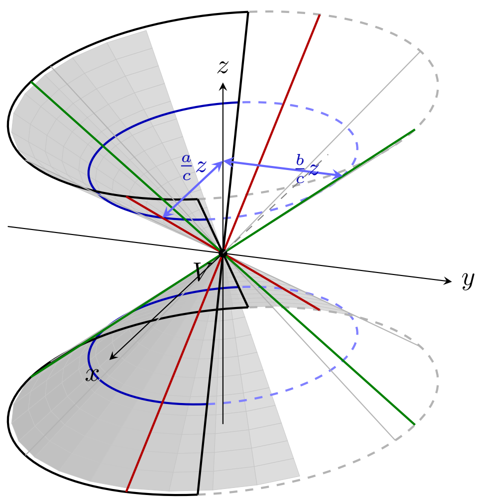


Figure 10.5: A quadric cone with vertex V at the origin, defined by the equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2}$. Cross-sections perpendicular to the z-axis (blue) are ellipses with semi-axes proportional to $|z|$. The red and green lines show the generators lying in the xz- and yz-planes respectively.

Ruled Surfaces Surfaces that can be generated by a moving straight line are known as *ruled surfaces*. Both cylinders and cones are examples of ruled quadric surfaces. Because they are formed by a one-parameter family of lines, such surfaces can be constructed by bending a flat sheet (in the case of cylinders and cones, these are *developable* surfaces).

We shall encounter two other quadric surfaces later: the *hyperboloid of one sheet* and the *hyperbolic paraboloid*. While these are also ruled surfaces, they are not developable; they possess two distinct families of rulings and cannot be formed by bending a plane without stretching or tearing the material.

10.3 The General Ellipsoid

Having explored surfaces generated by revolution and translation, we now approach the classification of quadric surfaces through their general Cartesian definitions. We begin with the surface defined by the sum of three squared terms.

Definition 10.3.1. ***The Ellipsoid***. The quadric surface defined by the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \tag{10.1}$$

where $a, b, c > 0$, is called an *ellipsoid*. The constants a, b, c are the lengths of the semi-axes.

If any two of the semi-axes are equal, the surface is a spheroid (a surface of revolution); if $a = b = c$, it is a sphere. We focus here on the general case where a, b, c are distinct. Without loss of generality, we assume an ordering of magnitude $a > b > c$.

Traces and Geometry

The shape of the ellipsoid is best revealed by its intersections with planes parallel to the coordinate planes, known as *traces*. The intersection with the coordinate plane $z = 0$ is the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. More generally, consider the section by a plane $z = k$. The equation becomes:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 - \frac{k^2}{c^2}.$$

This equation possesses real solutions only if $k^2 < c^2$. For $|k| < c$, the section is an ellipse centred at $(0, 0, k)$ with semi-axes $a\sqrt{1 - k^2/c^2}$ and $b\sqrt{1 - k^2/c^2}$. Geometrically, as k varies from 0 to c, the elliptical cross-section diminishes in size - maintaining constant eccentricity — until it degenerates to a single point $(0, 0, c)$ at the *pole*. The surface does not exist for $|k| > c$.

By symmetry, traces parallel to the x and y planes are also ellipses. Thus, the ellipsoid is a bounded surface, contained entirely within the box $|x| \leq a, |y| \leq b, |z| \leq c$.

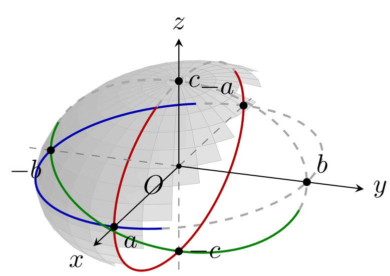


Figure 10.6: The ellipsoid with semi-axes a, b, c. All principal traces are ellipses.

The coordinate planes are planes of symmetry, known as the *principal planes*. They intersect along three mutually perpendicular lines, the *principal axes*, which meet at the *centre* of the ellipsoid [1].

Circular Sections Unlike the spheroid, the general ellipsoid does not possess an axis of rotational symmetry. However, it does possess plane sections that are circular.

***Theorem* 10.3.1. Existence of Circular Sections**. Every ellipsoid possesses two families of parallel planar sections that are circles.

Proof. Let the semi-axes be ordered $a > b > c$. Consider a family of planes passing through the y-axis defined by the equation $z = mx$. Substituting this into the Cartesian equation of the ellipsoid yields:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{m^2x^2}{c^2} = 1 \implies x^2\left(\frac{1}{a^2} + \frac{m^2}{c^2}\right) + \frac{y^2}{b^2} = 1.$$

[1] Claude made this tikz its so nice.

In the plane $z = mx$, let u be the coordinate along the line of intersection with the xz-plane. The distance from the origin to a point on this line is given by $u^2 = x^2 + z^2 = x^2(1+m^2)$. Replacing x^2 with $u^2/(1+m^2)$ in the trace equation gives the equation of the section in the (u, y) plane:

$$\frac{u^2}{1+m^2}\left(\frac{1}{a^2} + \frac{m^2}{c^2}\right) + \frac{y^2}{b^2} = 1.$$

For the elliptical section to be a circle, the coefficients of u^2 and y^2 must be equal:

$$\frac{1}{1+m^2}\left(\frac{1}{a^2} + \frac{m^2}{c^2}\right) = \frac{1}{b^2}.$$

Rearranging to solve for the slope m:

$$b^2(c^2 + a^2m^2) = a^2c^2(1+m^2)$$
$$m^2 = \frac{c^2(a^2 - b^2)}{a^2(b^2 - c^2)}.$$

Given the ordering $a > b > c$, it follows that $(a^2 - b^2) > 0$ and $(b^2 - c^2) > 0$. Thus m^2 is a positive real number, yielding two distinct real solutions for m. These slopes define two central planes whose intersections with the ellipsoid are circles of radius b. Because the equation of any plane parallel to these differs only by a constant term, the resulting sections remain circular, confirming the existence of two families of parallel circular sections. ■

Due to the symmetry of the ellipsoid, there are two such angles (tilted symmetrically with respect to the z-axis). Furthermore, any plane parallel to these specific circular sections will intersect the ellipsoid in a circle (or a point/empty set). This geometric property allows for the construction of physical models of ellipsoids using two sets of interlocking circular discs.

10.4 The Hyperboloid of Two Sheets

We turn now to the quadric surface defined by the presence of two negative terms in its standard equation.

Definition 10.4.1. ***Hyperboloid of Two Sheets*.** The quadric surface defined by the equation

$$\frac{z^2}{c^2} - \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \tag{10.2}$$

is called a *hyperboloid of two sheets*. It is a central quadric symmetric about the origin.

Remark. If the equation were $\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$, the axis of symmetry would be the x-axis. We assume the form in (10.2) for the subsequent geometric description, aligning the axis of the surface with the z-axis.

Geometry and Plane Sections We analyse the surface via its traces:

- **Horizontal Sections ($z = k$):** The equation rearranges to

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{k^2}{c^2} - 1.$$

  If $k^2 < c^2$, the RHS is negative, and there are no real solutions. Thus, the region $-c < z < c$ is empty, indicating a separation between two distinct parts of the surface. If $k = \pm c$, the intersection is the single point $(0, 0, \pm c)$. If $k^2 > c^2$, the sections are ellipses which increase in size as $|k|$ increases.

- **Vertical Sections ($x = 0$ or $y = 0$):** Setting $x = 0$ yields $\frac{z^2}{c^2} - \frac{y^2}{b^2} = 1$, a hyperbola in the yz-plane with vertices at $(0, \pm c)$ and foci on the z-axis. Similarly, setting $y = 0$ yields the hyperbola $\frac{z^2}{c^2} - \frac{x^2}{a^2} = 1$ in the xz-plane.

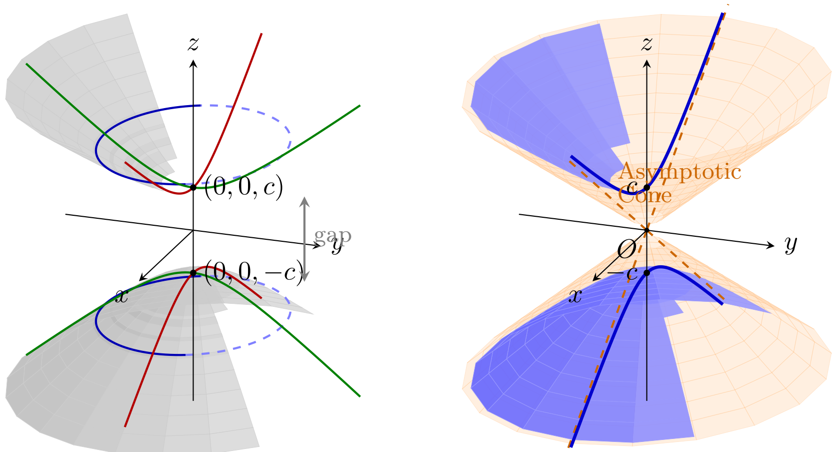
These properties describe two bowl-like sheets opening upwards and downwards along the z-axis, separated by a gap containing the origin. The coordinate planes are the principal planes of symmetry.

The Asymptotic Cone Just as a hyperbola in the plane is intimately related to its asymptotes, the hyperboloid is related to an *asymptotic cone*. Replacing the constant term 1 with 0 in (10.2) yields the homogeneous equation:

$$\frac{z^2}{c^2} - \frac{x^2}{a^2} - \frac{y^2}{b^2} = 0 \implies \frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2}.$$

This equation defines a cone with vertex at the origin.

Consider the trace on the vertical plane $y = 0$. The hyperboloid traces the curve $\frac{z^2}{c^2} - \frac{x^2}{a^2} = 1$, while the cone traces the lines $z = \pm\frac{c}{a}x$. These lines are precisely the asymptotes of the hyperbola. Geometric intuition suggests (and analysis confirms), that as $|z| \to \infty$, the surface of the hyperboloid approaches the surface of the asymptotic cone arbitrarily closely, without ever touching it. The cone fits "inside" the two sheets of the hyperboloid.



Geometry and Sections — Asymptotic Cone

Figure 10.7: Left: The hyperboloid of two sheets showing elliptical cross-sections (blue) and hyperbolic traces in coordinate planes (red, green). The gap $-c < z < c$ contains no points of the surface. Right: The asymptotic cone (orange, transparent) with vertex at the origin. The hyperboloid sheets (blue) approach but never touch the cone as $|z| \to \infty$.

10.5 The Hyperbolic Paraboloid

One of the most visually striking and architecturally significant quadric surfaces is the *hyperbolic paraboloid*, often referred to as a saddle surface.

Definition 10.5.1. ***Hyperbolic Paraboloid***. The quadric surface defined by the equation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = \frac{z}{c} \tag{10.3}$$

where $c > 0$, is called a hyperbolic paraboloid.

It is the only quadric surface (apart from degenerate forms) that is never a surface of revolution, regardless of the choice of a and b. Its geometry is revealed through its traces:

- **Vertical Sections:** The plane $x = 0$ cuts the surface in the parabola $y^2 = -\frac{b^2}{c}z$, which opens downwards. The plane $y = 0$ cuts the surface in the parabola $x^2 = \frac{a^2}{c}z$, which opens upwards. These parabolas share a common vertex at the origin, but lie in perpendicular planes and open in opposite directions, creating the saddle-like shape.

- **Horizontal Sections ($z = k$):** If $k > 0$, the section is a hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = \frac{k}{c}$ with its transverse axis along the x-axis. If $k < 0$, the section is a hyperbola with its transverse axis along the y-axis. If $k = 0$, the intersection consists of two intersecting lines $y = \pm\frac{b}{a}x$ in the xy-plane.

The hyperbolic paraboloid possesses two planes of symmetry ($x = 0$ and $y = 0$) and a single axis of symmetry (the z-axis). The origin is the vertex, a saddle point where the curvature is negative.

Rectilinear Generators Like the hyperboloid of one sheet, the hyperbolic paraboloid is a doubly ruled surface.

***Theorem* 10.5.1. Rulings of the Hyperbolic Paraboloid**. The hyperbolic paraboloid contains two families of straight lines.

Proof. We rewrite (10.3) using the difference of two squares:

$$\left(\frac{x}{a} - \frac{y}{b}\right)\left(\frac{x}{a} + \frac{y}{b}\right) = \frac{z}{c}.$$

We decompose this into two systems of linear equations parameterised by u and v:

$$\text{Family 1:} \quad \frac{x}{a} - \frac{y}{b} = u\frac{z}{c}, \qquad \frac{x}{a} + \frac{y}{b} = \frac{1}{u}.$$

$$\text{Family 2:} \quad \frac{x}{a} - \frac{y}{b} = v, \qquad \frac{x}{a} + \frac{y}{b} = \frac{1}{v}\frac{z}{c}.$$

For any fixed u (or v), the intersection of the two planes defines a line that lies entirely on the surface. (Note: slight modifications handle the $u = 0$ or $v = 0$ cases, typically setting one factor to 0 and the other arbitrary). ■

Geometrically, one family of rulings is parallel to the plane $\frac{x}{a} - \frac{y}{b} = 0$, and the other to $\frac{x}{a} + \frac{y}{b} = 0$. These are the *directrix planes*. This ruled nature makes the hyperbolic paraboloid straightforward to construct with straight beams, explaining its popularity in modern roof structures.

10.6 Coordinate Transformations

The equations derived thus far describe quadric surfaces in *standard position*, where their axes of symmetry align with the coordinate axes. To analyse a general quadratic equation

$$Ax^2 + By^2 + Cz^2 + Dxy + Eyz + Fzx + Gx + Hy + Iz + J = 0,$$

we employ coordinate transformations — translation and rotation — to reduce it to a recognisable standard form.

Translation of Axes A translation shifts the origin to a new point $O'(h, k, l)$ without changing the orientation of the axes. If (x, y, z) are the old coordinates and (x', y', z') are the new coordinates, the relationship is:

$$x = x' + h, \quad y = y' + k, \quad z = z' + l.$$

Substituting these into the quadratic equation allows us to eliminate linear terms (first-degree terms) by completing the square, provided the surface has a centre.

Example 10.6.1. Identifying a Surface by Translation. Consider the equation:

$$2x^2 - 3y^2 - 4z^2 - 4x - 12y - 4z = 0.$$

Grouping terms by variable and completing the squares:

$$2(x-1)^2 - 2(1) - 3(y+2)^2 + 3(4) - 4(z+1/2)^2 + 4(1/4) = 0$$

$$2(x-1)^2 - 3(y+2)^2 - 4(z+1/2)^2 = 2 - 12 - 1 = -11.$$
$$-\frac{2}{11}(x-1)^2 + \frac{3}{11}(y+2)^2 + \frac{4}{11}(z+1/2)^2 = 1.$$

Letting $x' = x-1, y' = y+2, z' = z+1/2$, we have $-A(x')^2 + B(y')^2 + C(z')^2 = 1$ with positive coefficients. This is a hyperboloid of one sheet centred at $(1, -2, -1/2)$.

Rotation of Axes To eliminate the cross-product terms (xy, yz, zx), we must rotate the coordinate axes. A rotation about the origin is described by an orthogonal matrix R. The new coordinates $\mathbf{x}'$ are related to the old $\mathbf{x}$ by $\mathbf{x} = R\mathbf{x}'$. Explicitly, if the new basis vectors are $\mathbf{u}, \mathbf{v}, \mathbf{w}$ (the columns of R), then:

$$x = u_1 x' + v_1 y' + w_1 z'$$
$$y = u_2 x' + v_2 y' + w_2 z'$$
$$z = u_3 x' + v_3 y' + w_3 z'$$

Substitution into the general quadratic form $\mathbf{x}^T Q\mathbf{x}$ yields $(\mathbf{x}')^T(R^T QR)\mathbf{x}'$. The goal is to choose R such that $R^T QR$ is diagonal; this is equivalent to finding the eigenvalues and eigenvectors of the symmetric matrix of coefficients Q.

Example 10.6.2. Simplification by Rotation. Consider the equation $2x^2 - y^2 - z^2 - 2yz - 3 = 0$. The cross-term $-2yz$ involves only y and z. We rotate the y, z axes about the x-axis by an angle θ.

$$y = y'\cos\theta - z'\sin\theta, \quad z = y'\sin\theta + z'\cos\theta.$$

The yz term transforms as:

$$-2(y'\cos\theta - z'\sin\theta)(y'\sin\theta + z'\cos\theta) = -2(y'^2\sin\theta\cos\theta + y'z'(\cos^2\theta - \sin^2\theta) - z'^2\sin\theta\cos\theta).$$

Since the coefficients of y^2 and z^2 are equal, a rotation of $45°$ $(\theta = \pi/4)$ is chosen to eliminate the cross-term. With $\theta = \pi/4$, $y = \frac{1}{\sqrt{2}}(y' - z')$ and $z = \frac{1}{\sqrt{2}}(y' + z')$. Substituting into $-y^2 - z^2 - 2yz$: The sum $y^2 + z^2 = y'^2 + z'^2$ is invariant under rotation. The cross-term $2yz = 2 \cdot \frac{1}{2}(y'^2 - z'^2) = y'^2 - z'^2$. Thus $-(y^2 + z^2) - 2yz = -(y'^2 + z'^2) - (y'^2 - z'^2) = -2y'^2$. The equation becomes $2x^2 - 2y'^2 - 3 = 0$, or $x^2 - y'^2 = 3/2$. This describes a hyperbolic cylinder.

10.7 Exercises

1. **Surface Identification.** Identify and describe the following surfaces by reducing their equations to standard forms via completing the square. Determine the centre (or vertex) and the orientation of the principal axes.
   (a) $4x^2 - y^2 + 4z^2 - 16x + 2y + 11 = 0$
   (b) $x^2 + y^2 + z + 2x - 4y + 5 = 0$
   (c) $9x^2 + 4y^2 + 36z^2 = 36$
   (d) $x^2 - 4y^2 = 0$ (in $\mathbb{R}^3$)
2. **Surfaces of Revolution.**
   (a) Find the equation of the surface generated by rotating the curve $z = e^{-y^2}$ (where $y > 0$) about the z-axis.
   (b) Find the equation of the surface generated by rotating the line $z = 2y$ about the z-axis. Identify the surface.
   (c) Explain why rotating the curve $y = x^3$ about the x-axis generates a surface defined by $y^2 + z^2 = x^6$, but rotating it about the y-axis generates $y = (x^2 + z^2)^{3/2}$.
3. **Traces and Boundaries.** Consider the surface defined by $x^2 - y^2 + z^2 = 1$.
   (a) Sketch the traces in the planes $x = k$, $y = k$, and $z = k$.
   (b) Determine the range of values for x, y, and z for which the surface exists.

(c) Is the surface connected? Is it bounded?

4. **Cylindrical Geometry.**

   (a) Find the equation of the cylinder whose rulings are parallel to the vector $\mathbf{v} = [1, 1, 1]$ and which passes through the circle $x^2 + y^2 = 1$ in the $z = 0$ plane.

   Remark. Let $P(x, y, z)$ be on the cylinder. Then $P - z\mathbf{v}$ must lie in the $z = 0$ plane and satisfy the circle's equation.

   (b) Identify the type of quadric cylinder given by $x^2 - z^2 - 2x + 4z - 2 = 0$.

5. **Coordinate Rotation.** Consider the equation $xy + z^2 = 1$.

   (a) Perform a rotation of axes in the xy-plane by 45° to eliminate the xy term.
   (b) Identify the surface in the new coordinate system.
   (c) Sketch the surface relative to the original axes.

6. **The Focus-Directrix Property in Space.** Let F be a fixed point (the focus) and Π be a fixed plane (the directrix plane). Let $e > 0$ be a scalar (the eccentricity). Consider the locus of points P such that the distance from P to F is e times the perpendicular distance from P to Π.

   (a) Choose coordinates such that Π is the plane $z = -d$ and F is the origin. Show that the equation of the locus is:
   $$x^2 + y^2 + z^2 = e^2(z + d)^2.$$

   (b) Prove that this equation defines:
   - An ellipsoid if $e < 1$.
   - A paraboloid of revolution if $e = 1$.
   - A hyperboloid of revolution if $e > 1$.

7. **Rotating a Skew Line.** Consider the straight line L defined by the equations $x = a$ and $z = by$, where $a, b \neq 0$. This line is skew to the z-axis.

   (a) Let $P(x_0, y_0, z_0)$ be a point on the surface generated by rotating L about the z-axis. Show that P must satisfy $x_0^2 + y_0^2 = a^2 + (z_0/b)^2$.
   (b) Rearrange this equation to identify the surface as a hyperboloid of one sheet.
   (c) Conclude that a hyperboloid of one sheet can be generated by rotating a straight line about an axis skew to it.

8. **Rectilinear Generators (Hyperboloid of One Sheet).** Consider the hyperboloid of one sheet $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$.

   (a) Verify that for any angle θ, the line given by the intersection of the planes
   $$\frac{x}{a} - \frac{z}{c} = \left(1 - \frac{y}{b}\right)\cos\theta \quad \text{and} \quad \left(1 + \frac{y}{b}\right) = \left(\frac{x}{a} + \frac{z}{c}\right)\cos\theta$$
   lies entirely on the hyperboloid.
   (b) By varying the constants, construct a second, distinct family of lines that lie on the surface.
   (c) Prove that through any point on the hyperboloid, there passes exactly one line from each family.

9. **Circular Sections of an Ellipsoid.** Let E be the ellipsoid with semi-axes $a > b > c$ aligned with the x, y, z axes.

   (a) Prove that the planes parallel to $y = 0$ intersect E in ellipses, not circles.
   (b) Show that the planes passing through the y-axis given by $z = x\tan\theta$ intersect the ellipsoid in curves whose projection onto the xz-plane satisfy
   $$\frac{x^2}{a^2} + \frac{x^2\tan^2\theta}{c^2} + \frac{y^2}{b^2} = 1.$$
   (c) Determine the specific values of θ for which this intersection is a circle (i.e., the coefficients of the transformed x^2 and y^2 terms are equal).

10. **Intersection of Quadrics.** Consider two cylinders defined by $x^2 + z^2 = a^2$ and $y^2 + z^2 = a^2$.

    (a) Describe the geometric arrangement of these cylinders.

(b) By subtracting the equations, show that the intersection lies on the pair of planes $x^2 - y^2 = 0$.
(c) Hence, show that the curve of intersection consists of two ellipses.

11. Line Intersection Theorem. Prove that if a quadric surface contains three distinct points lying on a straight line L, then it must contain the entire line L.

Remark. Let the line be $\mathbf{x}(t) = \mathbf{p} + t\mathbf{d}$. Substitute this into the general quadric equation to obtain a polynomial in t. What is the degree of this polynomial? What does it mean if it has three roots?