Chapter 1

Geometric Transformations

Topics :

1. Euclidean 3-Space
2. Linear Transformations
3. Translations and Affine Transformations
4. Isometries
5. Galilean Transformations
6. Lorentz Transformations

1.1 Euclidean 3-Space

The Euclidean space, points, and vectors

Three-dimensional *visual space* $\mathcal{S}$ is often used in mathematics without being formally defined. The "elements" of $\mathcal{S}$ are called *points.* In the usual sense, we introduce *Cartesian coordinates* by fixing a point, called the *origin*, and three (mutually orthogonal) *coordinate axes.* The choice of origin and of axes is arbitrary, but once it has been fixed, three real numbers (or coordinates) p_1, p_2, p_3 can be *measured* to describe the position of each point p.

The one-to-one correspondence

$$p \in \mathcal{S} \mapsto (p_1, p_2, p_3) \in \mathbb{R}^3$$

makes possible the *identification* of $\mathcal{S}$ with the set $\mathbb{R}^3$ of all ordered triplets of real numbers. In other words, instead of saying that three numbers *describe the position* of a point, we define them to *be* the point.

We make the following definition.

1.1.1 Definition. The (standard) **Euclidean 3-space** is the set $\mathbb{R}^3$ together with the *Euclidean distance* between points $p = (p_1, p_2, p_3)$ and $q = (q_1, q_2, q_3)$ given by

$$d(p, q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + (q_3 - p_3)^2}.$$

Note : Euclidean 3-space $\mathbb{R}^3$ is a *model* for the physical space. There are other models for our Universe. The question of what is the most convenient geometry with which to model physical space is an open one, and is the subject of intense contemporary investigation and speculation.

Let $p = (p_1, p_2, p_3)$ and $q = (q_1, q_2, q_3)$ be two points of $\mathbb{R}^3$, and let λ be a *scalar* (real number). The **sum** of p and q is the point

$$p + q := (p_1 + q_1, p_2 + q_2, p_3 + q_3)$$

and the **scalar multiple** of p by λ is the point

$$\lambda p := (\lambda p_1, \lambda p_2, \lambda p_3).$$

Under these two operations (the usual addition and scalar multiplication), $\mathbb{R}^3$ is a *vector space* over $\mathbb{R}$.

NOTE : The origin $o = (0, 0, 0)$ plays the role of *identity* (with respect to addition). The sum $p + (-1)q$ is usually written $p - q$.

We shall consider now the relationship between *points* and *geometric vectors* in Euclidean 3-space $\mathbb{R}^3$.

NOTE : The concept of *vector* originated in physics from such notions as velocity, acceleration, force, and angular momentum. These physical quantities are supplied with length and direction; they can be added and multiplied by scalars.

Intuitively, a geometric vector v in $\mathbb{R}^3$ is represented by a directed line segment (or "arrow") $\overrightarrow{pq}$. Here we take the view that a geometric vector is really the same thing as a *translation* in space.

NOTE : We can also take the view that we can describe an "arrow" (located at some point) by giving the starting point and the *change* necessary to reach its terminal point. This approach leads to the concept of (geometric) *tangent vector* and will be considered in the next chapter.

We make the following definition.

1.1.2 DEFINITION. A (geometric) **vector** in Euclidean 3-space $\mathbb{R}^3$ is a mapping

$$v : \mathbb{R}^3 \to \mathbb{R}^3, \quad p \mapsto v(p)$$

such that for any two points p and q, the midpoint of $\overline{pv(q)}$ is equal to the midpoint of $\overline{qv(p)}$.

Thus, if v is a vector and p, q are two points, then the quadrilateral $\square\, pqv(q)v(p)$ is a *parallelogram* (proper or degenerate).

$\diamond$ **Exercise 1** Show that given two points p and q, there is exactly one vector v such that $v(p) = q$.

This unique vector is denoted by $\overrightarrow{pq}$. A vector $\overrightarrow{pq}$ is sometimes called a **free vector**.

NOTE : An alternative description is the following. Two directed line segments $\overrightarrow{pq}$ and $\overrightarrow{p'q'}$ (or, if one prefers, two ordered pairs of points (p,q) and (p',q')) are *equivalent* if the line segments $\overrightarrow{pq}$ and $\overrightarrow{p'q'}$ are of the same length and are parallel in the same sense. This relation, being reflexive, symmetric, and transitive, is a genuine *equivalence relation.* Such an equivalence class of directed line segments (or, if one prefers, of ordered pairs of points) is a *vector.* We denote the vector $[\overrightarrow{pq}]$ simply by $\overrightarrow{pq}$. If $p = (p_1,p_2,p_3)$ and $q = (q_1,q_2,q_3)$, the *components* of the vector are q_1-p_1, q_2-p_2, and q_3-p_3. Two vectors are *equal* if and only if they have the same components.

◇ **Exercise 2** Show that two directed line segments $\overrightarrow{pq}$ and $\overrightarrow{p'q'}$ are *equivalent* if and only if $p+q'=p'+q$.

If $p = (p_1,p_2,p_3)$ and $q = (q_1,q_2,q_3)$, it is customary to represent the vector $v=\overrightarrow{pq}$ by the 3×1 matrix

$$\begin{bmatrix} q_1-p_1 \\ q_2-p_2 \\ q_3-p_3 \end{bmatrix}.$$

Let o be the origin of the Euclidean 3-space $\mathbb{R}^3$. Any point $p \in \mathbb{R}^3$ can be described by means of the vector $\overrightarrow{op}$ (the *position vector* of the point p). Each point has a unique position vector, and each position vector describes a unique point. Hence we set up a one-to-one correspondence between points and geometric vectors in $\mathbb{R}^3$. It is convenient to *identify*

$$\text{the point } (p_1,p_2,p_3) \text{ with the vector } \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}.$$

NOTE : An element of Euclidean 3-space $\mathbb{E}^3$ can be considered (or represented) either as an ordered triplet of real numbers or as a column 3-matrix with real entries. In other words, *we can think of the Euclidean 3-space as either the set of all its points or the set of all its (geometric) vectors.*

◇ **Exercise 3** Explain why the *identification* of the vector $v=\overrightarrow{pq}$ with the point $q-p$ is legitimate.

The (vector) space $\mathbb{R}^3$ has a built-in standard *inner product* (i.e., a non-degenerate symmetric bilinear form). For $v, w \in \mathbb{R}^3$, the **dot product** of (the vectors) v and w is the number (scalar)

$$v \bullet w := v_1 w_1 + v_2 w_2 + v_3 w_3.$$

The dot product is a *positive definite* inner product; that is, it has the following three properties (for $v, v', w \in \mathbb{R}^3$ and $\lambda, \lambda' \in \mathbb{R}$) :

(IP1) $(\lambda v + \lambda' v') \bullet w = \lambda(v \bullet w) + \lambda'(v' \bullet w)$ (linearity);

(IP2) $v \bullet w = w \bullet v$ (symmetry);

(IP3) $v \bullet v \geq 0$, and $v \bullet v = 0 \iff v = 0$ (positivity).

◇ **Exercise 4** Given $v, w \in \mathbb{R}^3$, show that

$$(v \bullet w)^2 \leq (v \bullet v)(w \bullet w).$$

This inequality is called the *Cauchy-Schwarz inequality.*

Write

$$\|v\| := \sqrt{v \bullet v} = \sqrt{v_1^2 + v_2^2 + v_3^2}$$

and call it the **norm** (or **length**) of (the vector) v. A vector with unit norm is called a **unit vector**.

NOTE : In view of our definition, we can rewrite the Cauchy-Schwarz inequality in the form

$$|v \bullet w| \leq \|v\| \, \|w\|.$$

The norm (more precisely, the *norm function* $v \in \mathbb{R}^3 \mapsto \|v\| \in \mathbb{R}$) has the following properties (for $v, w \in \mathbb{R}^3$ and $\lambda \in \mathbb{R}$) :

(N1) $\|v\| \geq 0$, and $\|v\| = 0 \iff v = 0$ (positivity);

(N2) $\|\lambda v\| = |\lambda| \, \|v\|$ (homogeneity);

(N3) $\|v + w\| \leq \|v\| + \|w\|$ (the triangle inequality).

◇ **Exercise 5** Let $v, w \in \mathbb{R}^3$. Verify the following properties.

(a) *Polarization identity* : $v \bullet w = \frac{1}{4}\left(\|v+w\|^2 - \|v-w\|^2\right)$, which expresses the standard inner product in terms of the norm.

(b) *Parallelogram identity* : $\|v+w\|^2 + \|v-w\|^2 = 2\left(\|v\|^2 + \|w\|^2\right)$. That is, the sum of the squares of the diagonals of a parallelogram equals the sum of the squares of the sides.

◇ **Exercise 6** Given $v, w \in \mathbb{R}^3$, prove the *Pythagorean property*

$$v \bullet w = 0 \iff \|v \pm w\|^2 = \|v\|^2 + \|w\|^2.$$

In terms of the norm we get a compact version of the (Euclidean) distance formula :

$$d(p,q) = \|v-w\| \quad \text{with} \quad v = \overrightarrow{op} \;\text{ and }\; w = \overrightarrow{oq}\,.$$

In other words, $\|v-w\|$ represents the distance between two points with position vectors v and w.

◇ **Exercise 7** Verify that the Euclidean distance satisfies the following properties (the axioms for a *metric*):

(M1) $d(p,q) \geq 0$, and $d(p,q) = 0 \iff p = q$;

(M2) $d(p,q) = d(q,p)$;

(M3) $d(p,r) \leq d(p,q) + d(q,r)$.

Relation (M3) is also known as the *triangle inequality.*

NOTE : Euclidean 3-space $\mathbb{R}^3$ is not only a vector space. It is also a *metric space.* It is important to realize that the Euclidean distance is completely determined by the dot product; indeed,

$$d(p,q) = \sqrt{(q-p)\bullet(q-p)} \qquad (p, q \in \mathbb{R}^3).$$

However, not any distance function is associated with an inner product. A (real) vector space endowed with a specific (positive definite) inner product is called an *inner product space.*

Let v and w be two *nonzero* vectors of $\mathbb{R}^3$. The Cauchy-Schwarz inequality permits us to define the cosine of the *angle* θ, $0 \leq \theta \leq \pi$ between v and w by the equation

$$v \bullet w = \|v\|\,\|w\|\,\cos\theta.$$

Thus the dot product of two vectors is the product of their lengths times the cosine of the angle between them. If $\theta = 0$ or $\theta = \pi$, the vectors v and w are said to be **collinear**, whereas if $\theta = \frac{\pi}{2}$, the vectors are called **orthogonal**.

NOTE : We regard the zero vector as both collinear with and orthogonal to *every* vector. Clearly, vectors v and w are orthogonal if and only if $v \bullet w = 0$.

$\diamond$ **Exercise 8** Given a nonzero vector w, show that vectors v and w are collinear if and only if $v = \lambda w$ for some $\lambda \in \mathbb{R}$.

There is another product on the Euclidean 3-space $\mathbb{R}^3$, second in importance only to the dot product. For $v, w \in \mathbb{R}^3$, the **cross product** of v and w is the vector

$$v \times w := \begin{bmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{bmatrix}.$$

An easy way to remember this formula is to compute the "determinant"

$$v \times w = \begin{vmatrix} e_1 & e_2 & e_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

by *formal* expansion along the first row. Here e_1, e_2, e_3 denote the *standard unit vectors*

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

NOTE : The vectors e_1, e_2, e_3 are lineary independent, and hence form a (orthonormal) *basis* of the vector space $\mathbb{R}^3$. Any vector $v \in \mathbb{R}^3$ can be expressed uniquely as a linear combination of the standard unit vectors e_1, e_2, e_3 :

$$v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = v_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + v_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + v_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = v_1 e_1 + v_2 e_2 + v_3 e_3.$$

Familiar properties of determinants show that the cross product (also called *vector product*) is a skew-symmetric bilinear mapping; that is, it has the following properties (for $v, v', w \in \mathbb{R}^3$ and $\lambda \in \mathbb{R}$) :

(VP1) $(v + v') \times w = v \times w + v' \times w$ (additivity);

(VP2) $\lambda(v \times w) = (\lambda v) \times w$ (homogeneity);

(VP3) $v \times w = -w \times v$ (skew-symmetry).

Hence, in particular, $v \times v = 0$.

◇ **Exercise 9** Show that

$$v \bullet (v \times w) = 0 \quad \text{and} \quad w \bullet (v \times w) = 0.$$

Therefore, the cross product of two vectors is a vector orthogonal to both of them.

◇ **Exercise 10** Verify (by tedious computation) the following formula known as the *Lagrange identity* :

$$\|v \times w\|^2 = \|v\|^2 \|w\|^2 - (v \bullet w)^2.$$

NOTE : The geometric usefulness of the cross product is based mostly on this result. A more intuitive description of the length of a cross product is

$$\|v \times w\| = \|v\| \, \|w\| \sin \theta$$

where θ is the angle between v and w. The *direction* of $v \times w$ on the straight line orthogonal to v and w is given, for practical purposes, by the so-called "right-hand rule": if the fingers of the right hand point in the direction of the shortest rotation of v to w, then the thumb points in the direction of $v \times w$.

◇ **Exercise 11** Show that vectors v and w are collinear if and only if $v \times w = 0$.

Combining the dot and cross product, we get the *triple scalar product* of three vectors $u, v,$ and w : $u \bullet v \times w$. Parantheses are unnecessary : $u \bullet (v \times w)$ is the only possible meaning.

◇ **Exercise 12** Given vectors $u, v,$ and w, show that

$$u \bullet v \times w = v \bullet w \times u = w \bullet u \times v = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

◇ **Exercise 13** Let $v, w \in \mathbb{R}^3$. Show that the only vector $x \in \mathbb{R}^3$ such that $u \bullet x$ is equal to the determinant $\det \begin{bmatrix} u & v & w \end{bmatrix}$ for all $u \in \mathbb{R}^3$ is $x = v \times w$.

◇ **Exercise 14** Given vectors u, v, and w, show that

$$(u \times v) \times w = (u \bullet w)v - (v \bullet w)u.$$

Deduce the *Jacobi identity* :

$$(u \times v) \times w + (v \times w) \times u + (w \times u) \times v = 0.$$

◇ **Exercise 15** Let $v_1, v_2, w_2, w_2 \in \mathbb{R}^3$. Verify the following identities.

(a) $(v_1 \times v_2) \bullet (w_1 \times w_2) = \det \begin{bmatrix} v_1 \times v_2 & w_1 & w_2 \end{bmatrix}$.

(b) $(v_1 \times v_2) \times (w_1 \times w_2) = \det \begin{bmatrix} v_1 & w_1 & w_2 \end{bmatrix} v_2 - \det \begin{bmatrix} v_2 & w_1 & w_2 \end{bmatrix} v_1$.

Geometric transformations

One of the most important concepts in geometry is that of a *transformation.*

NOTE : Moving geometric figures around is an ancient and natural approach to geometry. However, the Greek emphasis on synthetic geometry and constructions and much later the development of analytic geometry overshadowed transformational thinking. The study of polynomials and their roots in the early nineteenth century led to algebraic transformations and abstract groups. At the same time, AUGUST FERDINAND MÖBIUS (1790-1868) began studying geometric transformations. In the late nineteenth century, FELIX KLEIN (1849-1925) and SOPHUS LIE (1842-1899) showed the central importance of both groups and transformations for geometry.

Generally speaking, a geometric transformation is merely a mapping between two sets. However, these sets are assumed to be, in a certain sense, geometrical; they are equipped with some additional structure and are usually referred to as "spaces". We shall find it convenient to use the word *transformation* ONLY IN THE SPECIAL SENSE of a bijective mapping of a set (space) onto itself. *Groups* of transformations form the heart of geometry.

We make the following definition.

1.1.3 DEFINITION. A (geometric) **transformation** on $\mathbb{R}^3$ is a mapping from $\mathbb{R}^3$ to itself that is one-to-one and onto.

NOTE : Hereafter, in this chapter, all the definitions and results hold for $\mathbb{R}^3$ as well as for the *Euclidean plane* $\mathbb{R}^2$. We shall only discuss the case of $\mathbb{R}^3$, and consider the case of $\mathbb{R}^2$ as a special case.

Let T be a transformation on $\mathbb{R}^3$. Then T can be visualized as "moving" (or transforming) *each* point $p \in \mathbb{R}^3$ to its *unique* image $T(p) \in \mathbb{R}^3$. Given two transformations T and S, their *composition* (T followed by S)

$$ST : \mathbb{R}^3 \to \mathbb{R}^3, \quad p \mapsto S(T(p))$$

is called the **product** of S with T.

◇ **Exercise 16** Verify that the product of two transformations is a transformation.

The **identity transformation** I is defined by

$$I : \mathbb{R}^3 \to \mathbb{R}^3, \quad p \mapsto p.$$

For any transformation T on $\mathbb{R}^3$, $TI = IT = T$. Every transformation T has a unique *inverse* T^{-1}.

◇ **Exercise 17** Given two transformations T and S, show that

$$(ST)^{-1} = T^{-1}S^{-1}.$$

The set of all transformations on $\mathbb{R}^3$ is a (transformation) *group*. Various sets of transformations corespond to important geometric properties and also form groups.

NOTE : FELIX KLEIN in his famous *Erlanger Programm* (1872) used groups of transformations to give a definition of geometry : *Geometry is the study of those properties of a set that are preserved under a group of transformations on that set.* KLEIN showed that various non-Euclidean geometries, projective geometry, and Euclidean geometry were closely related, not competing subjects. He realized that we can, for example, investigate the properties of Euclidean geometry by studing *isometries* (i.e., distance-preserving transformations).

1.2 Linear Transformations

Linear transformations (on $\mathbb{R}^3$) are structure-preserving transformations on the *vector space* $\mathbb{R}^3$. The structure that must be preserved is that of vector addition and scalar multiplication (of which the geometric analogues are the parallelograms with one vertex at the origin and straight lines through the origin, respectively).

1.2.1 DEFINITION. A transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a **linear transformation** if, for all $x, y \in \mathbb{R}^3$ and all $\lambda \in \mathbb{R}$,

(L1) $T(x + y) = T(x) + T(y)$;

(L2) $T(\lambda x) = \lambda T(x)$.

NOTE : The terms *function, mapping, map*, and *transformation* are commonly used interchangeably. However, in studying geometric objects (particularly, on smooth manifolds), it is often convenient to make slight distinctions between them. Thus, we will reserve the term "function" for a map whose range is $\mathbb{R}$ (i.e., a real-valued map), whereas the terms "map" or "mapping" can mean any type of map. Furthermore, invertible maps (or mappings) – on some structured sets – will be referred to as "transformations". Typical transformations are structure-preserving bijections on structured sets of a certain kind. (In modern algebraic parlance, such transformations are usually called *automorphisms.*)

Addition and *scalar multiplication* of linear transformations are defined in the usual way. That is, for (linear) transformations S, T and scalar $\lambda \in \mathbb{R}$,

$$\begin{aligned} (S + T)(x) &:= S(x) + T(x) \\ (\lambda T)(x) &:= \lambda T(x). \end{aligned}$$

$\diamond$ **Exercise 18** Is the sum of any two linear transformations a linear transformation ? Justify your answer.

$\diamond$ **Exercise 19** Verify that, under the usual product, the set of all linear transformations on $\mathbb{R}^3$ is a (transformation) *group*.

Let $\{e_1, e_2, e_3\}$ be the *standard basis* of $\mathbb{R}^3$ and let T be a linear transformation on $\mathbb{R}^3$. Then we have, uniquely,

$$T(e_i) = a_{1i}e_1 + a_{2i}e_2 + a_{3i}e_3, \quad i = 1, 2, 3.$$

So we can associate to T a 3×3 matrix with real entries

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Notice that the image of e_i under T is the i^{th} column of the matrix A; that is, $A = \begin{bmatrix} T(e_1) & T(e_2) & T(e_3) \end{bmatrix}$. We can write

$$\begin{aligned}
T(x) &= T(x_1e_1 + x_2e_2 + x_3e_3) = x_1T(e_1) + x_2T(e_2) + x_3T(e_3) \\
&= x_1(a_{11}e_1 + a_{21}e_2 + a_{31}e_3) + x_2(a_{12}e_1 + a_{22}e_2 + a_{32}e_3) + \\
&\quad\; x_3(a_{13}e_1 + a_{23}e_2 + a_{33}e_3) \\
&= (a_{11}x_1 + a_{12}x_2 + a_{13}x_3)e_1 + (a_{21}x_1 + a_{22}x_2 + a_{23}x_3)e_2 + \\
&\quad\; (a_{31}x_1 + a_{32}x_2 + a_{33}x_3)e_3 \\
&= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\
&= Ax.
\end{aligned}$$

Consider now two linear transformations T and S with associated matrices (with respect to the standard basis of $\mathbb{R}^3$) $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ and $B = \begin{bmatrix} b_{ij} \end{bmatrix}$, respectively. Then the product ST is a linear transformation whose associated matrix is $C = BA$ (the matrix product of B and A). Indeed, we have

$$Cx = ST(x) = S\left(T(x)\right) = S(Ax) = B(Ax) = (BA)x.$$

$\diamond$ **Exercise 20** Show that the matrix associated with a linear transformation is *nonsingular* (i.e., invertible).

Let $\mathsf{GL}\,(3, \mathbb{R})$ be the set of all nonsingular 3×3 matrices with real entries. Under the usual matrix multiplication, $\mathsf{GL}\,(3, \mathbb{R})$ is a *group*.

◇ **Exercise 21** Show that the group of all linear transformations on $\mathbb{R}^3$ is *isomorphic* to the group $\mathsf{GL}(3,\mathbb{R})$.

Either one of these groups is called the **general linear group**. Given a matrix $A \in \mathsf{GL}(3,\mathbb{R})$, the transformation T, $x \mapsto Ax$ is the only linear transformation whose associated matrix is A. We say that the matrix A *represents* the linear transformation T. It is convenient to *identify*

the linear transformation T, $x \mapsto Ax$ with the (nonsingular) matrix A.

Henceforth, the same symbol will be used to denote a linear transformation and its associated matrix. Thus, for instance, I will denote the identity transformation $x \mapsto x$ as well as the *identity matrix*

$$\begin{bmatrix}\delta_{ij}\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}.$$

NOTE : The notation Ax stands for both the image of (the point) x under the linear transformation A and the matrix product of the (nonsingular) matrix A by the column matrix (vector) x.

Orthogonal transformations

Recall that the Euclidean 3-space $\mathbb{R}^3$ has a built-in inner product. Inner-product-preserving transformations form an important class of (linear) transformations.

1.2.2 DEFINITION. A linear transformation A, $x \mapsto Ax$ is an **orthogonal transformation** if it preserves the inner-product between any two vectors; that is, for all $x, y \in \mathbb{R}^3$,

$$Ax \bullet Ay = x \bullet y.$$

Let A and B be two orthogonal transformations. Then their product BA is also an orthogonal transformation. Indeed, for all vectors $x, y \in \mathbb{R}^3$,

$$(BA)x \bullet (BA)y = B(Ax) \bullet B(Ay) = Ax \bullet Ay = x \bullet y.$$

◇ **Exercise 22** Verify that the inverse of an orthogonal transformation is also an orthogonal transformation.

The set of all orthogonal transformations on $\mathbb{R}^3$ is a (transformation) *group*.

1.2.3 DEFINITION. A 3×3 matrix (with real entries) A is called **orthogonal** if

$$A^\top A = I,$$

where $A^\top$ is the *transpose* of A.

NOTE : If the matrix $A = \left[a_{ij}\right]$ is orthogonal, then (and only then)

$$a_{1i}a_{1j} + a_{2i}a_{2j} + a_{3i}a_{3j} = \delta_{ij}, \quad i, j = 1, 2, 3.$$

Thus the vectors (the columns of the matrix)

$$a_i := \begin{bmatrix} a_{1i} \\ a_{2i} \\ a_{3i} \end{bmatrix}, \quad i = 1, 2, 3$$

have unit length and are orthogonal to one another :

$$\|a_1\| = \|a_2\| = \|a_3\| = 1 \quad \text{and} \quad a_i \bullet a_j = 0 \quad (i \neq j).$$

(This can be written, in a more compact form, as $a_i \bullet a_j = \delta_{ij}, \quad i, j = 1, 2, 3.$)

Hence $\{a_1, a_2, a_3\}$ is an *orthonormal basis* for $\mathbb{R}^3$.

◇ **Exercise 23** Show that any orthogonal matrix is nonsingular.

◇ **Exercise 24** Let $A \in \mathsf{GL}(3, \mathbb{R})$. Show that

$$A^\top A = I \iff AA^\top = I \iff A^{-1} = A^\top.$$

Let $\mathsf{O}(3)$ be the set of all orthogonal matrices. Thus

$$\mathsf{O}(3) := \{A \in \mathsf{GL}(3, \mathbb{R}) \mid A^\top A = I\}.$$

◇ **Exercise 25** Show that $\mathsf{O}(3)$ is a *subgroup* of the general linear group $\mathsf{GL}(3, \mathbb{R})$.

1.2.4 Proposition. *A linear transformation A, $x \mapsto Ax$ is an orthogonal transformation if and only if the matrix A is orthogonal.*

Proof : ($\Rightarrow$) Suppose the transformation A, $x \mapsto Ax$ is orthogonal. Then we have

$$\begin{aligned} \delta_{ij} &= e_i \bullet e_j = Ae_i \bullet Ae_j \\ &= (Ae_i)^\top Ae_j = e_i^\top (A^\top A)e_j \\ &= (A^\top A)_{ij} \end{aligned}$$

and hence the matrix A is orthogonal.

($\Leftarrow$) Conversely, suppose the matrix A is orthogonal. Then

$$Ax \bullet Ay = (Ax)^\top Ay = x^\top (A^\top A)y = x^\top I y = x \bullet y$$

and thus the transformation $x \mapsto Ax$ is orthogonal. $\square$

The group of all orthogonal transformations is *isomorphic* to the group $\mathsf{O}(3)$. Either one of these groups is called the **orthogonal group**. Those elements of $\mathsf{O}(3)$ which have determinant equal to $+1$ form a subgroup of $\mathsf{O}(3)$, denoted by $\mathsf{SO}(3)$ and called the **special orthogonal group**.

1.2.5 Proposition. *An orthogonal transformation A, $x \mapsto Ax$ preserves the distance between any two points; that is, for all $x, y \in \mathbb{R}^3$,*

$$d(Ax, Ay) = d(x, y).$$

Proof : First we show that A preserves norms. By definition, $\|x\|^2 = x \bullet x$ and hence

$$\|Ax\|^2 = Ax \bullet Ax = x \bullet x = \|x\|^2.$$

Thus $\|Ax\| = \|x\|$ for all (vectors) $x \in \mathbb{R}^3$. Since A is linear, it follows that

$$d(Ax, Ay) = \|Ax - Ay\| = \|A(x-y)\| = \|x - y\| = d(x, y).$$

$\square$

Note : The orthogonal groups $\mathsf{O}(2)$ and $\mathsf{O}(3)$ were first studied, by the number theorists of the eighteenth century, as the groups of transformations preserving the quadratic form $\xi_1^2 + \xi_2^2$ or $\xi_1^2 + \xi_2^2 + \xi_3^2$, respectively.

Rotations and reflections

If $A \in \mathsf{O}(2)$, then the columns of A are unit vectors and are orthogonal to one another. Suppose

$$A = \begin{bmatrix} a_1 & a_3 \\ a_2 & a_4 \end{bmatrix}.$$

Then the point (a_1, a_2) lies on the unit circle $\mathbb{S}^1$ giving

$$a_1 = \cos\theta \quad \text{and} \quad a_2 = \sin\theta$$

for some θ satisfying $0 \leq \theta < 2\pi$. As the vector $\begin{bmatrix} a_3 \\ a_4 \end{bmatrix}$ is at right angles to $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and (as a point) also lies on the unit circle $\mathbb{S}^1$, we have

$$a_3 = \cos\varphi \quad \text{and} \quad a_4 = \sin\varphi$$

where either $\varphi = \theta + \frac{\pi}{2}$ or $\varphi = \theta - \frac{\pi}{2}\cdot$ In the first case we obtain

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

which is an element of $\mathsf{SO}(2)$ and represents a *rotation* about the origin (more precisely, a counterclockwise rotation about the origin through the angle θ). The second case gives

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

which has determinant -1 and represents a *reflection* in a line through the origin (more precisely, a reflection in a line through the origin at angle $\frac{\theta}{2}$ to the positive x_1-axis).

Therefore, *a* 2×2 *orthogonal matrix represents either a rotation of the plane about the origin or a reflection in a line through the origin, and the matrix has determinant* $+1$ *precisely when it represents a rotation.*

NOTE : The group $\mathsf{SO}(2)$ is often referred to as the *rotation group*. $\mathsf{SO}(2)$ is in fact the *unit circle* $\mathbb{S}^1$ in disguise. (Each point on the unit circle has the form $e^{i\theta}$, where $0 \leq \theta < 2\pi$ and hence corresponds to an angle.)

◇ **Exercise 26** Show that the mapping

$$e^{i\theta} \mapsto \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

is an *isomorphism* from $\mathbb{S}^1$ to $\mathsf{SO}\,(2)$. (Here $\mathbb{S}^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ is considered as a subgroup of the multiplicative group $\mathbb{C}^\times$ of complex numbers.)

◇ **Exercise 27** Let

$$A_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \text{and} \quad B_\varphi = \begin{bmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{bmatrix}.$$

(a) Verify that

$$A_\theta A_\varphi = A_{\theta+\varphi}, \quad A_\theta B_\varphi = B_{\theta+\varphi}, \quad B_\theta A_\varphi = B_{\theta-\varphi}, \quad B_\theta B_\varphi = A_{\theta-\varphi}$$

where the angles in the matrices are read modulo 2π. Interpret these results geometrically.

(b) Work out the products

$$A_\theta B_\varphi A_\theta^{-1}, \quad B_\varphi A_\theta B_\varphi, \quad A_\theta B_\varphi A_\theta^{-1} B_\varphi.$$

Evaluate each of these when $\theta = \frac{\pi}{3}$ and $\varphi = \frac{\pi}{2}$.

Rotations of the Euclidean space about any one of the coordinate axes are similar to those of the plane (about the origin). The three basic types are (realized by the following orthogonal matrices) :

$$\begin{aligned} R(e_1,\theta) = R_1(\theta) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \\ R(e_2,\theta) = R_2(\theta) &= \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \\ R(e_3,\theta) = R_3(\theta) &= \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Note : The minus sign appears above the (main) diagonal in R_1 and R_3, but below the diagonal in R_2. This is *not* a "mistake" : it is due to the *orientation* of the positive x_1-axis with respect to the x_2x_3-plane. Clearly, $R_i(\theta) \in \mathsf{SO}(3), \; i = 1,2,3$.

It can be shown that *any rotation* $x \mapsto Ax$ *of* $\mathbb{R}^3$ *which fixes the origin can be written as a product of just three of these elementary rotations :*

$$A = R_1(\theta)R_2(\varphi)R_3(\psi).$$

(The independent parameters θ, φ, ψ are called the *Euler angles* for the given rotation.)

It follows that

1.2.6 Proposition. *Every rotation of* $\mathbb{R}^3$ *which fixes the origin can be represented by a matrix in* $\mathsf{SO}(3)$.

Now suppose that $A \in \mathsf{SO}(3)$. The *characteristic polynomial* $\mathrm{char}_A(\lambda) = \det(\lambda I - A)$ is cubic and therefore must have at least one real root. That is to say, A has a real eigenvalue. As the product of the eigenvalues of a matrix is the determinant of the matrix, we *see* that $+1$ is an eigenvalue of A.

$\diamond$ **Exercise 28** Show that every $A \in \mathsf{SO}(3)$ has an eigenvalue equal to $+1$.

Note : The other two eigenvalues are complex conjugate and have absolute value 1, so they can be written as $e^{i\theta}$ and $e^{-i\theta}$ for some $\theta \in \mathbb{R}$.

If w is a corresponding *eigenvector* (i.e., $Aw = w$), the line through the origin determined by w is invariant under (the linear transformation) A. Also since A preserves right angles, it must send the *plane* which is orthogonal to w, and which contains the origin, to itself.

$\diamond$ **Exercise 29** Check that the set (plane) $w^\perp = \{y \in \mathbb{E}^3 \mid y \bullet w = 0\}$ is invariant under (the orthogonal transformation) A; that is, $A(w^\perp) = w^\perp$.

Construct an *orthonormal* basis for $\mathbb{R}^3$ which has the unit vector $\frac{1}{\|w\|}w$ as first member. The matrix of $x \mapsto Ax$ with respect to this new basis will be of the form

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & a_1 & a_3 \\ 0 & a_2 & a_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix}.$$

Since $R \in \mathsf{SO}(2)$, $x \mapsto Ax$ is a *rotation* with axis determined by w.

Therefore, *each matrix in* $\mathsf{SO}(3)$ *represents a rotation of* $\mathbb{R}^3$ *about an axis which passes through the origin.*

NOTE : Every element (rotation) $A \in \mathsf{SO}(3)$ can be written as

$$\begin{aligned} A &= R(w, \theta) \\ &= P\, R(e_1, \theta)\, P^{-1} \end{aligned}$$

for some w, θ, and $P \in \mathsf{SO}(3)$. (We say that A and $R(e_1, \theta)$ are *conjugate* in $\mathsf{SO}(3)$.) The eigenvector w determines the *axis* of the rotation (i.e., the unique line through the origin which is left fixed). The *angle* of rotation is obtained from the other two eigenvectors. (In fact, θ is given by the eigenvalue $e^{i\theta}$.)

$\diamond$ **Exercise 30** Let $A = R(w, \theta) \in \mathsf{SO}(3)$.

(a) Show that

$$A - A^\top = \begin{bmatrix} 0 & c & b \\ -c & 0 & a \\ -b & -a & 0 \end{bmatrix} \quad \text{and} \quad w \in \ker\left(A - A^\top\right).$$

Hence deduce that (for $\theta \neq 0, \pi$)

$$w = \lambda \begin{bmatrix} -a \\ b \\ -c \end{bmatrix}.$$

(b) Show that

$$\operatorname{tr} A = 1 + 2\cos\theta.$$

(So we can solve for $\cos\theta$ from the trace of A. However, we don't know without further investigation if the rotation is clockwise or counterclockwise about w.)

◇ **Exercise 31** Show that the matrices

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2/3 & 1/3 & 2/3 \\ -2/3 & 2/3 & 1/3 \\ -1/3 & -2/3 & 2/3 \end{bmatrix}$$

both represent rotations, and then find axes and angles for these rotations.

◇ **Exercise 32** Let $A \in \mathsf{SO}(3)$ and $w \in \mathbb{E}^3$ such that $\|w\| = 1$. Show that, for all $x, y \in \mathbb{E}^3$ and $\theta \in \mathbb{R}$,

(a) $Ax \times Ay = A(x \times y)$;

(b) $w \bullet (R(w,\theta)x - x) = 0$;

(c) $A\,R(w,\theta)\,A^{-1} = R(Aw,\theta)$.

(Hint : The cross product $x \times y$ can be characterized as the unique vector such that

$$w \bullet x \times y = \det \begin{bmatrix} w & x & y \end{bmatrix}$$

for every vector w. See **Exercise 13**.)

Note : The group $\mathsf{SO}(3)$ is often referred to as the *rotation group*. $\mathsf{SO}(3)$ and the *sphere* $\mathbb{S}^3$ are *not* the "same" (i.e., they are not isomorphic groups). It is an interesting fact that $\mathbb{S}^0 = \{-1, 1\}$, $\mathbb{S}^1$, and $\mathbb{S}^3$ are the *only* spheres which can be groups.

If A lies in $\mathsf{O}(3)$ but *not* in $\mathsf{SO}(3)$, then $AS \in \mathsf{SO}(3)$ where

$$S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

The matrix S represents a *reflection* in the x_1x_2-plane (identified with the Euclidean plane $\mathbb{R}^2$). We write

$$A = (AS)S.$$

As above, the transformation $x \mapsto ASx$ is a rotation. Consequently, A is a reflection (in the x_1x_2-plane) followed by a rotation.

◇ **Exercise 33** Complete the entries in the matrix

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \cdot \\ 0 & 1 & \cdot \\ -\frac{1}{\sqrt{2}} & 0 & \cdot \end{bmatrix}$$

to give an element of $\mathsf{SO}(3)$, and to give an element of $\mathsf{O}(3) \setminus \mathsf{SO}(3)$. Describe the (linear) trasformations represented by these matrices.

◇ **Exercise 34** Let $c \in \mathbb{R}^3$ such that $\|c\| = 1$. Prove that the correspondence

$$x \mapsto c \times x + (c \bullet x)c$$

defines an orthogonal transformation. Describe its general effect on $\mathbb{R}^3$.

1.3 Translations and Affine Transformations

Let $c \in \mathbb{R}^3$ be a vector and let T_c be the mapping that adds c to every point of $\mathbb{R}^3$. This mapping is one-to-one and onto and hence a transformation.

1.3.1 Definition. The transformation

$$T_c : \mathbb{R}^3 \to \mathbb{R}^3, \quad x \mapsto x + c$$

is called the **translation** by vector c.

Note : A nonidentity translation is *not* a linear transformation.

◇ **Exercise 35** Show that given two points $p, q \in \mathbb{R}^3$, there exists a unique translation T such that $T(p) = q$.

The *inverse* of the translation T_c, $x \mapsto x + v$ is the translation T_c^{-1}, $x \mapsto x - c$. Thus,

$$T_c^{-1} = T_{-c}.$$

◇ **Exercise 36** Verify that the product of two translations is also a translation.

The set of all translations on $\mathbb{R}^3$ is a (transformation) *group*. This group is *isomorphic* to the additive group (also denoted by $\mathbb{R}^3$) of (the vectors of) $\mathbb{R}^3$. Either one of these groups is called the **translation group**.

1.3.2 PROPOSITION. *A translation $T = T_c$, $x \mapsto x + c$ preserves the distance between any two points; that is, for all $x, y \in \mathbb{R}^3$,*

$$d(T(x), T(y)) = d(x, y).$$

PROOF : We have

$$d(T(x), T(y)) = \|T(x) - T(y)\| = \|x + c - (y + c)\| = \|x - y\| = d(x, y).$$

□

1.3.3 DEFINITION. An **affine transformation** F on $\mathbb{R}^3$ is a linear transformation followed by a translation; that is, a transformation of the form

$$F : \mathbb{R}^3 \to \mathbb{R}^3, \quad F = TA$$

where A, $x \mapsto Ax$ is a linear transformation, and $T = T_c$, $x \mapsto x + c$ is a translation. A is called the *linear part* of F, and T the *translation part* of F.

For every $x \in \mathbb{R}^3$,

$$F(x) = Ax + c.$$

NOTE : The pair $(c, A) \in \mathbb{R}^3 \times \mathsf{GL}(3, \mathbb{R})$ *represents* the affine transformation F, $x \mapsto Ax + c$.

Affine transformations F, $x \mapsto Ax + c$, include the linear transformations (with $c = 0$) and the translations (with $A = I$). Let F, $x \mapsto Ax + c$ and G, $x \mapsto Bx + d$ be two affine transformations. Then (for $x \in \mathbb{R}^3$)

$$GF(x) = G(F(x)) = B\,(Ax + c) + d = (BA)x + Bc + d$$

and thus the product of G with F is also an affine transformation.

◇ **Exercise 37** Show that the inverse of an affine transformation is also an affine transformation.

The set of all affine transformations on $\mathbb{R}^3$ is a (transformation) *group*, which contains as subgroups the general linear group $\mathsf{GL}(3, \mathbb{R})$ and the translation group $\mathbb{R}^3$.

NOTE : Affine transformations preserve lines, parallelism, betweeness, and proportions on lines. Affine transformations can distort shapes. However, there is a limit to the amount of distortion : a convex set is always mapped to a convex set. (The converse holds as well : Transformations on $\mathbb{R}^3$ that preserve convexity are affine transformations.)

Any affine transformation (on $\mathbb{R}^3$) is represented by a pair $(c, A) \in \mathbb{R}^3 \times \mathsf{GL}(3, \mathbb{R})$ which we can write further as a 4×4 matrix $\begin{bmatrix} 1 & 0 \\ c & A \end{bmatrix}$. Call such a matrix an **affine matrix**. Let $\mathsf{GA}(3, \mathbb{R})$ be the set of all affine matrices. Thus

$$\mathsf{GA}(3, \mathbb{R}) := \left\{ \begin{bmatrix} 1 & 0 \\ c & A \end{bmatrix} \mid c \in \mathbb{R}^3, \ A \in \mathsf{GL}(3, \mathbb{R}) \right\}.$$

$\diamond$ **Exercise 38** Show that $\mathsf{GA}(3, \mathbb{R})$ is a *group*.

The group of all affine transformations on $\mathbb{R}^3$ is *isomorphic* to the group $\mathsf{GA}(3, \mathbb{R})$. Either of these groups is called the **general affine group**.

NOTE : In $\mathbb{R}^3$, of special interest are the affine transformations $x \mapsto Ax + c$ with $\det A = 1$. These transformations also form a group..

1.4 Isometries

Isometries (on Euclidean 3-space $\mathbb{R}^3$) are distance-preserving transformations on the *metric space* $\mathbb{R}^3$. They do not change the distance between points as the transformations move these points. Isometries are the dynamic counterpart to the Euclidean notion of *congruence*.

1.4.1 DEFINITION. A transformation $F : \mathbb{R}^3 \to \mathbb{R}^3$ is an **isometry** (or **rigid motion**) if it preserves the distance between any two points; that is, for all $x, y \in \mathbb{R}^3$,

$$d(F(x), F(y)) = d(x, y).$$

Orthogonal transformations and translations are isometries. If F is an isometry, then (for $x, y \in \mathbb{R}^3$)

$$d(F^{-1}(x), F^{-1}(y)) = d(FF^{-1}(x), FF^{-1}(y)) = d(x, y)$$

and thus the inverse F^{-1} is also an isometry.

◇ **Exercise 39** Verify that the product of two isometries is also an isometry.

The set of all isometries on $\mathbb{R}^3$ is a (transformation) *group*, which contains as subgroups the orthogonal group $\mathsf{O}(3)$ and the translation group $\mathbb{R}^3$.

1.4.2 PROPOSITION. *If F is an isometry on $\mathbb{R}^3$ such that $F(0) = 0$, then F is an orthogonal transformation.*

PROOF : For any (vector) $x \in \mathbb{R}^3$,

$$\|F(x)\| = d(0, F(x)) = d(F(0), F(x)) = d(0, x) = \|x\|.$$

Let $x, y \in \mathbb{R}^3$. Then we have

$$\|F(x) - F(y)\| = d(F(x), F(y)) = d(x, y) = \|x - y\|$$

which implies

$$(F(x) - F(y)) \bullet (F(x) - F(y)) = (x - y) \bullet (x - y)$$

or

$$\|F(x)\|^2 - 2F(x) \bullet F(y) + \|F(y)\|^2 = \|x\|^2 - 2x \bullet y + \|y\|^2.$$

Thus we have

$$F(x) \bullet F(y) = x \bullet y$$

so that F preserves the inner product of any two vectors.

It remains to prove that F is a linear transformation. Let $x \in \mathbb{R}^3$. Then (with respect to the standard basis)

$$x = x_1 e_1 + x_2 e_2 + x_3 e_3.$$

Since $\{e_1, e_2, e_3\}$ is an orthonormal basis (and F preserves the inner product of any two vectors), it follows that $\{F(e_1), F(e_2), F(e_3)\}$ is also an orthonormal basis so that

$$F(x) = \bar{x}_1 F(e_1) + \bar{x}_2 F(e_2) + \bar{x}_3 F(e_3).$$

Taking the inner product of both sides with $F(e_i)$, we get

$$\bar{x}_i = F(x) \bullet F(e_i) = x \bullet e_i = x_i, \quad i = 1, 2, 3.$$

Hence

$$F(x) = x_1 F(e_1) + x_2 F(e_2) + x_3 F(e_3)$$

and we can easily check the linearity conditions (L_1) and (L_2). $\square$

1.4.3 Theorem. *If F is an isometry on $\mathbb{R}^3$, then there exists a unique orthogonal transformation A, $x \mapsto Ax$ and a unique translation $T = T_c$, $x \mapsto x + c$ such that*

$$F = TA.$$

A is called the orthogonal part *of F, and T the* translation part *of F.*

Proof : Let T be the translation by vector $c = F(0)$. Then T^{-1} is the translation by vector $-c = -F(0)$, and so $T^{-1}F$ is an isometry. Furthermore,

$$T^{-1}F(0) = T^{-1}(F(0)) = F(0) - F(0) = 0.$$

Thus $T^{-1}F$ is an orthogonal transformation, say $T^{-1}F = A$, from which follows immediately $F = TA$.

To prove the required uniqueness, suppose $F = \bar{T}\bar{A}$, where $\bar{T}$ is a translation and $\bar{A}$ is an orthogonal transformation. Then

$$TA = \bar{T}\bar{A}$$

so that $A = T^{-1}\bar{T}\bar{A}$. Since A and $\bar{A}$ are linear transformations, $A(0) = \bar{A}(0) = 0$. It follows that $T^{-1}\bar{T} = I$ (the identity transformation), so that $\bar{T} = T$, which implies $\bar{A} = A$. $\square$

Note : We see that an isometry on $\mathbb{R}^3$ is a special affine transformation. Intermediate between isometries and affine transformations are *similarities*, the transformations corresponding to similar figures. Similarities preserve betweeness, segments, angle measure, and the proportions of all distances. The set of all similarities on $\mathbb{R}^3$ is a subgroup $\mathsf{GE}(3)$ of the general affine group, called the *general Euclidean group*. A similarity that is not an isometry is either a *dilation* or a *dilative rotation (spiral)*.

The group of all isometries on $\mathbb{R}^3$ is (*isomorphic* to) a subgroup of the general affine group $\mathsf{GA}(3,\mathbb{R})$, denoted by $\mathsf{E}(3)$. We have

$$\mathsf{E}(3) = \left\{ \begin{bmatrix} 1 & 0 \\ c & A \end{bmatrix} \mid c \in \mathbb{R}^3, \;\; A \in \mathsf{O}(3) \right\}.$$

Either one of these groups is called the **Euclidean group**.

NOTE : The Euclidean group $\mathsf{E}(3)$ is generated by *reflections*. Each isometry on $\mathbb{R}^3$ is exactly one of the following : *translation, rotation, glide rotation (screw), reflection, glide reflection*, or *rotary reflection*. In the case of the plane, an element of $\mathsf{E}(2)$ is exactly one of the following : *translation, rotation, reflection,* or *glide reflection*.

Orientation

We now come to one of the most interesting and elusive ideas in geometry. Intuitively, it is *orientation* that distinguishes between a right-handed glove and a left-handed glove in ordinary space. We shall not formalize this concept now.

NOTE : To handle the concept of *orientation* mathematically, we replace "gloves" by orthonormal bases (in fact, *frames*) and separate all these orthonormal bases of $\mathbb{R}^3$ into two classes : positively-oriented (or right-handed) and negatively-oriented (or left-handed).

Let F, $x \mapsto Ax + c$ be an isometry on $\mathbb{R}^3$. Since (the matrix) A is orthogonal, its determinant is either $+1$ or -1. We define the **sign** of F to be the determinant of A, with notation

$$\operatorname{sgn} F := \det A.$$

1.4.4 DEFINITION. An isometry F, $x \mapsto Ax + c$ is said to be

- **direct** (or **orientation-preserving**) if $\operatorname{sgn} F = +1$;
- **opposite** (or **orientation-reversing**) if $\operatorname{sgn} F = -1$.

All translations are orientation-preserving. Intuitively this is clear. In fact, the orthogonal part of a translation T is just the identity transformation I, and so $\operatorname{sgn} T = \det I = +1$.

$\diamond$ **Exercise 40** Consider the orthogonal transformation $R_1(\theta)$ represented by the matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}.$$

Show that $R_1(\theta)$ is orientation-preserving.

1.4.5 Example. One can (literally) see reversal of orientation by using a mirror. Suppose the x_2x_3-plane of $\mathbb{R}^3$ is the mirror. If one looks toward the plane, the point $p = (p_1, p_2, p_3)$ appears to be located at the point

$$S(p) = (-p_1, p_2, p_3).$$

The transformation S, $p \mapsto S(p)$ is the *reflection* in the x_2x_3-plane. Evidently, S is an orthogonal transformation represented by the (orthogonal) matrix

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

Thus S is an orientation-reversing isometry, as confirmed by the experimental fact that the mirror image of the right hand is a left hand.

Recall that an isometry is also called a *rigid motion*. If this is the case, a direct isometry is referred to as a **proper rigid motion**. The set of all direct isometries (or proper rigid motions) on $\mathbb{R}^3$ is a (transformation) *group*. This group is (*isomorphic* to) a subgroup of the Euclidean group $\mathsf{E}(3)$, denoted by $\mathsf{SE}(3)$. We have

$$\mathsf{SE}(3) := \left\{ \begin{bmatrix} 1 & 0 \\ c & A \end{bmatrix} \mid c \in \mathbb{R}^3, \;\; A \in \mathsf{SO}(3) \right\}.$$

Either one of these groups is called the **special Euclidean group**.

NOTE : The orientation-preserving isometries on $\mathbb{R}^3$ are precisely the *translations, rotations,* and *glide rotations (screws).* In the case of the plane, the elements of the special Euclidean group $\mathsf{SE}(2)$ are the *translations* and the *rotations.*

1.5 Galilean Transformations

Galilean spacetime

Newtonian mechanics takes place in a *Galilean spacetime.* Let $\mathbb{R}^3$ be the Euclidean 3-space and let $\mathbb{R}\times\mathbb{R}^3$ denote the (standard) **Galilean spacetime**. Elements of $\mathbb{R}\times\mathbb{R}^3$ are called **events**.

NOTE : $\mathbb{R}\times\mathbb{R}^3$ is a *model* for spatio-temporal world of Newtonian mechanics. The Newtonian world is comprised of objects sitting in a Universe (i.e., a Galilean spacetime) and interacting with one another in a way consistent with the *Galilean relativity principle* (which states that for a closed system in Galilean spacetime the governing physical laws are invariant under Galilean transformations). In particular, determinacy principle says that to "see" what will happen in the Universe, one need only specify initial conditions for the ODEs of Newtonian mechanics, and all else follows, at least in principle.

Given two events $\xi = (t,x) = (t,(x_1,x_2,x_3))$ and $\xi' = (t',x') = (t',(x_1',x_2',x_3'))$, the *time* between these events is

$$\mathfrak{t}(\xi,\xi') := t' - t.$$

The *distance* between **simultaneous events** (t,x) and (t,x') is then

$$\mathfrak{d}((t,x),(t,x')) := \|x'-x\| = \sqrt{(x_1'-x_1)^2+(x_2'-x_2)^2+(x_3'-x_3)^2},$$

where $\|\cdot\|$ is the (standard) Euclidean norm on $\mathbb{R}^3$.

NOTE : Distance between events that are *not* simultaneous *cannot* be measured. In particular, it does not make sense to talk about two non-simultaneous events as ocurring in the same place (i.e., as separated by zero distance). The picture one should have in mind for a Galilean spacetime is of it being *a union of simultaneous events, nicely stacked together.* We write

$$\mathbb{R}\times\mathbb{R}^3 = \bigcup_{t\in\mathbb{R}}\{t\}\times\mathbb{R}^3 := \bigcup_{t\in\mathbb{R}}\mathbb{R}^3_t.$$

That one cannot measure distance between non-simultaneous events reflects there being *no* natural direction transverse to the *stratification* by simultaneous events.

Galilean transformations

Galilean transformations are structure-preserving transformations on the Galilean spacetime. They preserve simultaneity of events and do not change the distance between simultaneous events.

1.5.1 Definition. An affine transformation $F : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R} \times \mathbb{R}^3$ is a **Galilean transformation** if it preserves the time between any two events and the distance between any two simultaneous events; that is, for all $\xi, \xi' \in \mathbb{R}\times\mathbb{R}^3$,

$$\mathfrak{t}\,(F(\xi), F(\xi')) = \mathfrak{t}\,(\xi, \xi')$$

and, for all $t \in \mathbb{R}$ and $\xi, \xi' \in \mathbb{R}^3_t$,

$$\mathfrak{d}\,(F(\xi), F(\xi')) = \mathfrak{d}\,(\xi, \xi').$$

Let F, $(t, x) \mapsto A(t, x) + (\zeta, c)$ be a Galilean transformation. Let us write

$$\begin{aligned} A(t,x) + (\zeta, c) &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix} + \begin{bmatrix} \zeta \\ c \end{bmatrix} \\ &= \begin{bmatrix} A_{11}t + A_{12}x + \zeta \\ A_{21}t + A_{22}x + c \end{bmatrix} \\ &= (A_{11}t + A_{12}x + \zeta, A_{21}t + A_{22}x + c) \end{aligned}$$

where $A_{11} \in \mathbb{R}$ and $A_{22} \in \mathsf{GL}\,(3, \mathbb{R})$.

$\diamond$ **Exercise 41** Show that if

$$(t, x) \mapsto (A_{11}t + A_{12}x + \zeta,\ A_{21}t + A_{22}x + c)$$

is a Galilean transformation, then

$$A_{11} = 1, \quad A_{12} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}, \quad A_{21} = v \in \mathbb{R}^{3\times 1}, \quad A_{22} \in \mathsf{O}\,(3).$$

Any Galilean transformation

$$(t, x) \mapsto (t + \zeta,\ Rx + tv + c)$$

where $\zeta \in \mathbb{R},\ c, v \in \mathbb{R}^{3\times 1}$, and $R \in \mathsf{O}\,(3)$, may be written in matrix form as

$$\begin{bmatrix} t \\ x \end{bmatrix} \mapsto \begin{bmatrix} 1 & 0 \\ v & R \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix} + \begin{bmatrix} \zeta \\ c \end{bmatrix}.$$

◇ **Exercise 42** Show that the set of all Galilean transformations is a (transformation) *group*.

The following basic Galilean transformations

- $(t, x) \mapsto (t + \zeta, x + c)$ (*shift* of origin);
- $(t, x) \mapsto (t, x + tv)$ (velocity *boost*);
- $(t, x) \mapsto (t, Rx)$ ("rotation" of reference frame)

can be used to *generate* the whole set (group) of Galilean transformations.

NOTE : The names given to these basic Galilean transformations are suggestive. A shift of the origin (in fact a spacetime translation) may be thought of as moving the origin to a new position and resetting the clock, but maintaining the same orientation in space. A (Galilean) velocity boost means the origin maintains its "orientation" and uses the same clock, but now moves with a certain velocity with respect to the previous origin. Finally, the "rotation" of reference frame (in fact an orthogonal transformation or linear isometry) means the origin stays in the same place and uses the same clock, but rotates the "point of view".

Any Galilean transformation is represented by a quadruple $(\zeta, c, v, R) \in \mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathsf{O}\,(3)$ which we can write further as a 5×5 matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ \zeta & 1 & 0 \\ c & v & R \end{bmatrix}.$$

Let Gal be the set of all such matrices. Thus

$$\mathsf{Gal} := \left\{ \begin{bmatrix} 1 & 0 & 0 \\ \zeta & 1 & 0 \\ c & v & R \end{bmatrix} \mid \zeta \in \mathbb{R},\ c, v \in \mathbb{R}^3,\ R \in \mathsf{O}\,(3) \right\}.$$

◇ **Exercise 43** Show that Gal is a *group*.

The group of all Galilean transformations is *isomorphic* to the group Gal. Either one of these groups is called the **Galilean group**

We saw that the elements of Gal are products of spacetime translations, velocity boosts, and spatial orthogonal transformations (in particular, rotations). Various *subgroups* of Gal are of particular interest in applications (including some familiar transformation groups). For instance,

- the subgroup of *isochronous* Galilean transformations consists of those Galilean transformations (represented by the quadruple (ζ, c, v, R)) for which $\zeta = 0$;
- the subgroup of *unboosted* Galilean transformations consists of those Galilean transformations (represented by the quadruple (ζ, c, v, R)) for which $v = 0$;
- the subgroup of *anisotropic* Galilean transformations consists of those Galilean transformations (represented by the quadruple (ζ, c, v, R)) for which $R = I$;
- the subgroup of *homogeneous* Galilean transformations consists of those Galilean transformations (represented by the quadruple (ζ, c, v, R)) for which $\zeta = 0, \; c = 0$.

◇ **Exercise 44** Identify the following subgroups of Gal.

(a) The subgroup of Gal consisting of those Galilean transformations (represented by the quadruple (ζ, c, v, R)) for which $\zeta = 0, \; v = 0$.

(b) The subgroup of Gal consisting of those Galilean transformations (represented by the quadruple (ζ, c, v, R)) for which $v = 0, \; R = I$.

(c) The subgroup of Gal consisting of those Galilean transformations (represented by the quadruple $\zeta, c, v, R)$) for which $\zeta = 0, \; v = 0, \; R = I$.

(d) The subgroup of Gal consisting of those Galilean transformations (represented by the quadruple (ζ, c, v, R)) for which $c = 0, \; v = 0, \; R = I$.

1.6 Lorentz Transformations

Minkowski spacetime

The geometric setting for EINSTEIN's *Special Theory of Relativity* is provided by *Minkowski spacetime.*

NOTE : A *spacetime* is simply the mathematical version of a universe that, like our own physical universe, has dimensions both of space and of time. A *flat* spacetime is a spacetime with no *gravity*, since gravitation tends to "bend" spacetime. Flat spacetimes are the simplest kind of spacetimes; they stand in the same relation to curved spacetimes as a flat Euclidean plane does to a curved surface.

We make the following definition.

1.6.1 DEFINITION. The (standard) **Minkowski spacetime** $\mathbb{R}^{1,3}$ is the vector space $\mathbb{R}^4$ together with the *Minkowski product* between vectors $v = (v_0, v_1, v_2, v_3)$ and $w = (w_0, w_1, w_2, w_3)$ given by

$$v \odot w := -v_0w_0 + v_1w_1 + v_2w_2 + v_3w_3.$$

The Minkowski product is an inner product; that is, it has the following three properties (for $v, v', w \in \mathbb{R}^{1,3}$ and $\lambda, \lambda' \in \mathbb{R}$) :

(IP1) $(\lambda v + \lambda' v') \odot w = \lambda(v \odot w) + \lambda'(v' \odot w)$;

(IP2) $v \odot w = w \odot v$;

(IP4) $v \odot w = 0$ for all v implies $w = 0$.

We can write (for $v, w \in \mathbb{R}^{1,3}$)

$$v \odot w = v^\top Q w$$

where

$$Q = \operatorname{diag}(-1, 1, 1, 1) = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The elements (vectors) of $\mathbb{R}^{1,3}$ are also called **events**.

NOTE : We can use t in place of v_0 since in *relativity theory* this coordinate is related to the time measurements while the others are related to the spatial ones. Hence we can write elements of the Minkowski spacetime $\mathbb{R}^{1,3}$ in the form

$$\xi = (t, x) = \begin{bmatrix} t \\ x \end{bmatrix} \qquad (t \in \mathbb{R},\ x \in \mathbb{R}^3).$$

Then (for $\xi = (t, x)$ and $\xi' = (t', x')$) $\xi \odot \xi' = -tt' + x \bullet x'$.

Two vectors $v, w \in \mathbb{R}^{1,3}$ are **Minkowski-orthogonal** provided $v \odot w = 0$.

1.6.2 EXAMPLE. Since the Minkowski product is *not* positive definite, there exist nonzero elements (vectors) $v \in \mathbb{R}^{1,3}$ for which $v \odot v = 0$. For instance, such a vector is $v = (1, 0, 1, 0)$. Such vectors are said to be *null* and $\mathbb{R}^{1,3}$ actually has bases which consist exclusively of this type of vector. A *null basis* cannot consist of mutually (Minkowski-)orthogonal vectors, however.

◇ **Exercise 45** Show that two null vectors v, w are Minkowski-orthogonal if and only if they are linearly dependent (i.e., $v = \lambda w$ for some $\lambda \in \mathbb{R}$).

We make the following definitions (this terminology derives from *relativity theory*).

1.6.3 DEFINITION. A nonzero vector $v \in \mathbb{R}^{1,3}$ is called

- **spacelike** provided $v \odot v > 0$;
- **timelike** provided $v \odot v < 0$;
- **null** (or **lightlike**) provided $v \odot v = 0$.

◇ **Exercise 46** Show that if a nonzero vector is Minkowski-orthogonal to a timelike vector, then it must be spacelike.

NOTE : Let $\mathcal{Q}$ denote the quadratic form associated with the Minkowski product on $\mathbb{R}^{1,3}$; that is, the mapping

$$\mathcal{Q} : \mathbb{R}^{1,3} \to \mathbb{R}, \qquad v \mapsto v \odot v.$$

Consider two distinct events ξ and ξ_0 for which the *displacement vector* $v := \xi - \xi_0$ from ξ_0 to ξ is null (i.e., $\mathcal{Q}(\xi - \xi_0) = 0$). Then we can define the *null cone* (or *light cone* $\mathcal{C}_N(\xi_0)$) at ξ_0 by

$$\mathcal{C}_N(\xi_0) := \left\{\xi \in \mathbb{R}^{1,3} \,|\, \mathcal{Q}(\xi - \xi_0) = 0\right\}.$$

$\mathcal{C}_N(\xi_0)$ consists of all those events in $\mathbb{R}^{1,3}$ that are "connectible to ξ_0 by a light ray".

Let $\mathcal{T}$ denote the collection of all timelike vectors in $\mathbb{R}^{1,3}$ and define a relation $\sim$ on $\mathcal{T}$ as follows :

$$v \sim w \iff v \odot w < 0.$$

This is an equivalence relation and hence $\mathcal{T}$ is the union of two disjoint subsets (equivalence classes) $\mathcal{T}^+$ and $\mathcal{T}^-$, called *time cones*, and there is no intrinsic way to distinguish one from the other. We think of the elements of $\mathcal{T}^+$ (and $\mathcal{T}^-$) as having the same time orientation. More specifically, we select (arbitrarily) $\mathcal{T}^+$ and refer to its elements as *future-directed* timelike vectors, whereas the vectors in $\mathcal{T}^-$ we call *past-directed*.

For each ξ_0 in $\mathbb{R}^{1,3}$ we define the *time cone* $\mathcal{C}_T(\xi_0)$, *future time cone* $\mathcal{C}_T^+(\xi_0)$, and *past time cone* $\mathcal{C}_T^-(\xi_0)$ at ξ_0 by

$$\begin{aligned}
\mathcal{C}_T(\xi_0) \quad &:= \quad \left\{\xi \in \mathbb{R}^{1,3} \,|\, \mathcal{Q}(\xi - \xi_0) < 0\right\} \\
\mathcal{C}_T^+(\xi_0) \quad &:= \quad \left\{\xi \in \mathbb{R}^{1,3} \,|\, \xi - \xi_0 \in \mathcal{T}^+\right\} = \mathcal{C}_T(\xi_0) \cap \mathcal{T}^+ \\
\mathcal{C}_T^-(\xi_0) \quad &:= \quad \left\{\xi \in \mathbb{R}^{1,3} \,|\, \xi - \xi_0 \in \mathcal{T}^-\right\} = \mathcal{C}_T(\xi_0) \cap \mathcal{T}^-.
\end{aligned}$$

We picture $\mathcal{C}_T(\xi_0)$ as the interior of the null cone $\mathcal{C}_N(\xi_0)$. It is the (disjoint) union of $\mathcal{C}_T^+(\xi_0)$ and $\mathcal{C}_T^-(\xi_0)$.

The notion of time-orientation can be extended to null vectors. We say that a null vector n is *future-directed* if $n \odot v < 0$ for all $v \in \mathcal{T}^+$ and *past-directed* if $n \odot v > 0$ for all $v \in \mathcal{T}^+$. For any event ξ_0 we define the *future null cone* $\mathcal{C}_N^+(\xi_0)$ and the *past null cone* $\mathcal{C}_N^-(\xi_0)$ at ξ_0 by

$$\begin{aligned}
\mathcal{C}_N^+(\xi_0) \quad &:= \quad \{\xi \in \mathcal{C}_N(\xi_0) \,|\, \xi - \xi_0 \text{ is future-directed}\} \\
\mathcal{C}_N^-(\xi_0) \quad &:= \quad \{\xi \in \mathcal{C}_N(\xi_0) \,|\, \xi - \xi_0 \text{ is past-directed}\}.
\end{aligned}$$

Physically, event ξ is in $\mathcal{C}_N^+(\xi_0)$ if ξ_0 and ξ can be regarded as the emission and reception of a light signal, respectively. Consequently, $\mathcal{C}_N^+(\xi_0)$ may be thought of as the history in spacetime of a spherical electromagnetic wave (photons in all directions) whose emission event is ξ_0.

For a vector $v = (v_0, v_1, v_2, v_3) \in \mathbb{R}^{1,3}$ we write

$$\|v\| := \sqrt{v \odot v} = \sqrt{\left|-v_0^2 + v_1^2 + v_2^2 + v_3^2\right|}$$

and call it the **Minkowski norm** (or **length**) of v. A *unit vector* is a vector v with Minkowski norm 1 : $v \odot v = \pm 1$.

NOTE : This is a funny kind of "length" since null vectors have zero length (even though they are not zero). For any timelike vector v, the Minkowski norm $\|v\|$ is commonly referred to as the *duration* of v. If $v = \xi - \xi_0$ is the displacement vector between two events ξ, ξ_0, then $\|v\|$ is to be interpreted physically as the *time separation* of ξ_0 and ξ (in any admissible frame of reference in which both events occur at the same spatial location).

Many features of Euclidean 3-space $\mathbb{R}^3$ (which is a positive definite inner product space) have counter-intuitive analogues in the Minkowski case. For example, analogues of the basic inequalities (like the Cauchy-Schwarz inequality and the triangle inequality) are generally reversed.

$\diamond$ **Exercise 47** Show that if v and w are timelike vectors, then

$$(v \odot w)^2 \geq (v \odot v)(w \odot w)$$

and equality holds if and only if v and w are linearly dependent.

1.6.4 PROPOSITION. *Let v and w be timelike vectors in the same time cone (i.e., with the same time orientation : $v \odot w < 0$). Then*

$$\|v + w\| \geq \|v\| + \|w\|$$

and equality holds if and only if v and w are linearly dependent.

PROOF : Since $v \odot v < 0$, $v + w \in \mathcal{T}$ and (by **Exercise 47**)

$$\|v\| \, \|w\| \leq -v \odot w.$$

Hence

$$\begin{aligned}
(\|v\| + \|w\|)^2 &= \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 \\
&\leq \|v\|^2 - 2v \odot w + \|w\|^2 \\
&\leq -(v + w) \odot (v + w) \\
&= \|v + w\|^2
\end{aligned}$$

and the equality holds if and only if $\|v\|\,\|w\| = -v \odot w$. The conclusion now follows from **Exercise 47**. $\square$

Lorentz transformations

Lorentz transformations are structure-preserving transformations on the Minkowski spacetime.

1.6.5 Definition. A linear transformation (on $\mathbb{R}^{1,3}$) $L,\ v \mapsto Lv$ is an **orthogonal transformation** if it preserves the Minkowski product between any two vectors; that is, for all $v, w \in \mathbb{R}^{1,3}$,

$$Lv \odot Lw = v \odot w.$$

The set of all orthogonal transformations on $\mathbb{R}^{1,3}$ is a (transformation) *group*.

$\diamond$ **Exercise 48** Let $L : \mathbb{R}^{1,3} \to \mathbb{R}^{1,3}$ be a linear transformation. Then show that the following statements are equivalent :

(a) L is an orthogonal transformation.

(b) L preserves the quadratic form on $\mathbb{R}^{1,3}$ (i.e., $\mathcal{Q}(Lv) = \mathcal{Q}(v)$ for all $v \in \mathbb{R}^{1,3}$).

(c) (The matrix of) L satisfies the condition

$$L^\top Q L = Q$$

where $Q = \operatorname{diag}(-1, 1, 1, 1)$. (Hint : To prove that $(b) \Rightarrow (a)$ compute $L(v+w) \odot L(v+w) - L(v-w) \odot L(v-w)$.)

Any such linear transformation $L,\ v \mapsto Lv$ on $\mathbb{R}^{1,3}$ is called a **general (homogeneous) Lorentz transformation**.

Note : If $L = \begin{bmatrix} l_{ij} \end{bmatrix}$ is a 4×4 matrix such that $L^\top Q L = Q$, where $Q = \operatorname{diag}(-1, 1, 1, 1)$, then its columns are *mutually Minkowski-orthogonal unit vectors*.

Let Lor_{GH} be the set of all such 4×4 matrices (i.e. matrices representing general homogeneous Lorentz transformations). Thus

$$\mathsf{Lor}_{GH} := \{L \in \mathsf{GL}(4, \mathbb{R}) \mid L^\top Q L = Q\}.$$

◇ **Exercise 49** Show that Lor_{GH} is a *subgroup* of the general linear group $\mathsf{GL}(4, \mathbb{R})$.

The group of all (Minkowski-)orthogonal transformations is *isomorphic* to the group Lor_{GH}. Either one of these groups is called the **general (homogeneous) Lorentz group.**

Let $L = \left[l_{ij} \right] \in \mathsf{Lor}_{GH}$ (i.e., $L^\top Q L = Q$). Then, in particular, we have

$$l_{11}^2 = 1 + l_{12}^2 + l_{13}^2 + l_{14}^2 \geq 1$$

so that

$$l_{11} \geq 1 \quad \text{or} \quad l_{11} \leq -1.$$

L is said to be *orthocronous* if $l_{11} \geq 1$ and *nonorthocronous* if $l_{11} \leq -1$. Nonorthocronous Lorentz transformations have certain "unsavory" characteristics; for instance, they always *reverse* time orientation (and so presumably relate reference frames in which someone's clock is running backwards). For this reason, *it is commom practice to restrict attention to the orthocronous elements of* Lor_{GH}.

There is yet one more restriction we would like to impose on our Lorentz transformations.

◇ **Exercise 50** Show that if $L \in \mathsf{Lor}_{GH}$, then

$$\det L = 1 \quad \text{or} \quad \det L = -1.$$

We shall say that a Lorentz transformation $v \mapsto Lv$ is *proper* if $\det L = 1$ and *improper* if $\det L = -1$. The set Lor of all proper, orthocronous Lorentz transformations is a *subgroup* of Lor_{GH}. Generally, we shall refer to Lor simply as the **Lorentz group** and its elements as Lorentz transformations with the understanding that they are all proper and orthocronous.

Note : Ocasionally, it is convenient to enlarge the group Lor to include spacetime translations, thereby obtaining the so-called **inhomogeneous Lorentz group** (or **Poincaré group**). Physically, this amounts to allowing "admissible" observers to use different spacetime origins.

The Lorentz group Lor has an important subgroup consisting of those elements of the form

$$R = \begin{bmatrix} 1 & 0 \\ 0 & A \end{bmatrix}$$

where $A \in \mathsf{SO}\,(3)$ (i.e., $A^\top = A^{-1}$ and $\det A = 1$). Such elements are called (spatial) *rotations* (in Lor).

A Lorentz transformation $v \mapsto L(\beta)v$ of the form

$$L(\beta) := \begin{bmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{bmatrix}$$

where $-1 < \beta < 1$ (and $\gamma := \frac{1}{\sqrt{1-\beta^2}} \geq 1$) is called a **special Lorentz transformation**. The matrix $L(\beta)$ is often called a (Lorentz) *boost* in the x_1-direction.

Note : Likewise, one can define matrices (representing) boosts in the x_2- and x_3-directions. One can also define a boost in an arbitrary direction by first rotating, say, the positive x_1-axis into that direction and then applying $L(\beta)$.

◇ **Exercise 51** Suppose $-1 < \beta_1 \leq \beta_2 < 1$. Show that :

(a) $\left| \dfrac{\beta_1 + \beta_2}{1 + \beta_1\beta_2} \right| < 1$.

(b) $L(\beta_2)L(\beta_1) = L\left(\dfrac{\beta_1 + \beta_2}{1 + \beta_1\beta_2} \right)$.

(Hint : Show that if a is a constant, then the function $x \mapsto \dfrac{x+a}{1+ax}$ is *increasing* for $-1 \leq x \leq 1$.)

It follows from **Exercise 51** that the product of two boosts in the x_1-direction is another boost in the x_1-direction. Since $L(\beta)^{-1} = L(-\beta)$, *the collection of all such special Lorentz transformations forms a subgroup of* Lor. We point out, however, that the product of two boosts in two different directions is, in general, *not* equivalent to a single boost in any direction.

NOTE : A simple computation shows that if we put $\beta = \tanh\theta$, then the Lorentz transformation $L(\beta)$ takes the *hyperbolic form* :

$$L(\theta) = \begin{bmatrix} \cosh\theta & 0 & 0 & -\sinh\theta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\theta & 0 & 0 & \cosh\theta \end{bmatrix}.$$

It is remarkable that *all of the physically interesting behaviour of (proper, orthochronous) Lorentz transformations is exhibited by the special Lorentz transformations* : any element of Lor differs from some $L(\beta)$ only by at most two rotations (in Lor); that is, for $L \in \mathsf{Lor}$ there is some (real) number θ and (spatial) rotations $R_1, R_2 \in \mathsf{Lor}$, such that

$$L = R_1 L(\theta) R_2.$$