The space in which we live is nearly flat everywhere. Its geometry is Euclidean, meaning that Euclid’s axioms and the theorems deduced from them are valid everywhere. After Einstein published his General Relativity in 1916 we know, however, that space is not perfectly flat. In the field of gravity from a massive body, space necessarily curves, but only very little unless the body is a black hole. The kind of physics that is the subject of this book may always be assumed to take place in flat Euclidean space.

One of the consequences of the Euclidean geometry is Pythagoras’ theorem that in the well-known way relates the lengths of the sides of any right-angled triangle. The simplicity of Pythagoras’ theorem favors the use of right-angled Cartesian coordinate systems in which the distance between two points in space is the squareroot of the sum of the squares of their coordinate differences. In Cartesian coordinates, vector algebra also finds its simplest form.

This chapter serves in most respects to define the mathematical notation and present the efficient modern methods of Cartesian vector and tensor algebra. In appendix C vector notation is extended to include differentiation and integration, and in appendix D a couple of useful non-Cartesian coordinate systems are introduced and related to Cartesian coordinates. One should, however, be aware that the emphasis on strictly algebraic treatment of geometric concepts differs from what is usually seen in texts at this level.

### B.1 Cartesian vectors

In a Cartesian coordinate system the distance between any two points, \( \mathbf{x} = (x, y, z) \) and \( \mathbf{x}' = (x', y', z') \), is given by the expression\(^1\)

\[
d(x', x) = \sqrt{(x' - x)^2 + (y' - y)^2 + (z' - z)^2}.
\]

(B.1)

This distance function implies that space is Euclidean, and therefore has all the properties one learns about in elementary geometry. Although we could prove this claim right here, it becomes nearly trivial after vector algebra has been established.

\(^1\)In this book we have chosen to follow the physics tradition in which the Cartesian coordinates are labeled \( x, y \) and \( z \). Mathematicians would prefer instead to label the coordinates \( x_1, x_2, \) and \( x_3 \), which is definitely a more systematic notation. Boldface symbols, \( \mathbf{x} = (x, y, z) \), are used to denote Cartesian coordinate triplets and vectors (see below). For calculations with pencil on paper several different notations can be used to distinguish a triplet from other symbols, for example a bar (\( \overline{x} \)), an arrow (\( \vec{x} \)) or underlining (\( \underline{x} \)).
Definition of a vector

The Cartesian distance function (1.14) depends only on the coordinate differences between two points, not on their individual values. Since all geometry is contained in the distance function, coordinate differences will be of such importance in the analytic description of space in Cartesian coordinate systems that a special notation and a special type of algebra is necessary to deal efficiently with them. The mathematical concept of a vector is usually attributed to R. W. Hamilton (1845).

A vector is a triplet of Cartesian coordinate differences between two points, say \( x \) and \( x' \),

\[
\mathbf{a} = (a_x, a_y, a_z) = (x' - x, y' - y, z' - z),
\]

This particular vector is naturally visualized by a straight arrow pointing from \( x \) to \( x' \). Since a vector only depends on the coordinate differences of its endpoints, the same symbol \( \mathbf{a} \) can be used for the vector connecting any other pair of points, say \( \mathbf{u} = (u, v, w) \) and \( \mathbf{u}' = (u', v', w') \), as long as they have the same coordinate differences. A vector has like a real arrow both direction and length, but no fixed origin. In accordance with the etymology of the word\(^2\), the same vector \( \mathbf{a} \) will carry you from \( x \) to \( x' \), and from \( \mathbf{u} \) to \( \mathbf{u}' \).

Notice how the three components of the vector, \( a_x, a_y \), and \( a_z \), are labeled by the coordinate symbols \( x, y, \) and \( z \). This is a general rule which will also be used for curvilinear coordinates (see appendix D).

Position vectors

Conceptually there is a great difference between the triplet of real numbers making up the coordinates of a point, also called its position, and the triplet of coordinate differences making up a vector. Whereas it a priori makes no geometric sense to add the coordinates of two points, the sum of the components of two vectors is just another vector. The distinction is, however, not so clear in Cartesian coordinate systems, because the three real numbers in the triplet \( x \) can be formally viewed as the difference between the true coordinates of a point and the coordinates, \( \mathbf{0} = (0, 0, 0) \), of the origin of the coordinate system. We shall for this reason permit ourselves the ambiguity of also calling \( x \) the position vector, thereby allowing the rules of vector algebra defined below also to be applied to coordinate triplets.

It must be stressed that the identification of positions and vectors is absolutely not possible in curvilinear coordinates, for example cylindrical or spherical, and even less so in non-Euclidean spaces. In those cases, true vectors can only be defined from infinitesimal coordinate differences in the local Cartesian coordinate systems that always exist in the neighborhood of any point.

B.2 Vector algebra

In keeping with the algebraization of geometry initiated four hundred years ago by Descartes, we shall focus on the algebraic rather than the geometric properties of vectors. As discussed in section 1.4 on page 11 the coordinate system is viewed as the physical reference frame for the determination of the coordinates of any point in space according to well-defined operational procedures. In this way we avoid all undefined geometrical primitives, such as the points, lines and circles of Euclidean geometry.

The following definitions endow vectors with the usual properties of the familiar geometric vectors. Geometric visualization is of course as useful as ever, and we shall whenever possible use simple sketches to illustrate what is meant.

\(^2\)The word “vector” is Latin for “one who carries”, derived from the verb “vehere”, meaning to carry, and also known from “vehicle”. In epidemiology a “vector” denotes the carrier of disease.
B. CARTESIAN COORDINATES

Linear operations

Linear operations lie at the core of vector algebra,

\[ k \mathbf{a} = (k a_x, k a_y, k a_z) \]  \hspace{1cm} \text{(scaling by a factor)} \hspace{1cm} \text{(B.3)}
\[ \mathbf{a} + \mathbf{b} = (a_x + b_x, a_y + b_y, a_z + b_z) \]  \hspace{1cm} \text{(addition)} \hspace{1cm} \text{(B.4)}
\[ \mathbf{a} - \mathbf{b} = (a_x - b_x, a_y - b_y, a_z - b_z) \]  \hspace{1cm} \text{(subtraction)} \hspace{1cm} \text{(B.5)}

Mathematically, these rules show that the set of all vectors constitute a three-dimensional vector space.

A set of vectors \( \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N \) are said to be linearly dependent if there exists a vanishing linear combination with non-zero coefficients, \( k_1 \mathbf{a}_1 + k_2 \mathbf{a}_2 + \cdots + k_N \mathbf{a}_N = \mathbf{0} \). More than three vectors are always linearly dependent because space is three-dimensional.

A straight line with origin \( \mathbf{a} \) and direction vector \( \mathbf{b} \neq \mathbf{0} \) is described by the linear vector function \( \mathbf{x}(s) = \mathbf{a} + s \mathbf{b} \) with \(-\infty < s < \infty\). Note that the origin \( \mathbf{a} = x(0) \) is a position vector whereas the direction vector \( \mathbf{b} = x(1) - x(0) \) is the difference between the coordinates of two points, and thus a true vector. The straight line is in fact the shortest path between any two points, and its length is equal to the distance between them, as it must in Euclidean geometry (see problem B.3).

Dot product

The dot product or scalar product of two vectors is familiar from geometry,

\[ \mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y + a_z b_z \]  \hspace{1cm} \text{(dot product)} \hspace{1cm} \text{(B.6)}

Two vectors are said to be orthogonal when their dot product vanishes, \( \mathbf{a} \cdot \mathbf{b} = 0 \). The square of a vector equals its dot product with itself, \( |\mathbf{a}|^2 = \mathbf{a} \cdot \mathbf{a} = a_x^2 + a_y^2 + a_z^2 \). The length of a vector is \( |\mathbf{a}| = \sqrt{a_x^2 + a_y^2 + a_z^2} \).

Cross product

The cross product or vector product is also familiar from geometry,

\[ \mathbf{a} \times \mathbf{b} = (a_y b_z - a_z b_y, a_z b_x - a_x b_z, a_x b_y - a_y b_x) \]  \hspace{1cm} \text{(cross product)} \hspace{1cm} \text{(B.7)}

It is defined entirely in terms of the coordinates and we do not in the rule itself distinguish between left-handed and right-handed coordinate systems. Whether you use your right or left hand when you draw a cross product on paper does not matter for the vector product rule, as long as you consistently use the same hand for all such drawings. The cross product \( \mathbf{a} \times \mathbf{b} \) is also called the area vector of the parallelogram spanned by the vectors \( \mathbf{a} \) and \( \mathbf{b} \), because the length \( |\mathbf{a} \times \mathbf{b}| \) is equal to its area.

Tensor product

The tensor product is not familiar from geometry,

\[ \mathbf{a} \otimes \mathbf{b} = \begin{pmatrix} a_x b_x & a_x b_y & a_x b_z \\ a_y b_x & a_y b_y & a_y b_z \\ a_z b_x & a_z b_y & a_z b_z \end{pmatrix} \]  \hspace{1cm} \text{(tensor product)} \hspace{1cm} \text{(B.8)}

It is unusual in that it produces a 3 \( \times \) 3 matrix from two vectors, but otherwise it is perfectly well defined and quite useful to have around. It is nothing but an ordinary matrix product of a column-matrix and a row-matrix, also called the direct product and sometimes in the older literature the dyadic product. Later we shall introduce more general geometric objects, called tensors, of which the simplest are represented by matrices of this kind. Tensor products — and tensors in general — cannot be given a simple visualization.
Volume product

The trilinear product of three vectors obtained by combining the cross product and the dot product is called the volume product,

\[ \mathbf{a} \times \mathbf{b} \cdot \mathbf{c} = a_x b_y c_z + a_y b_z c_x + a_z b_x c_y - a_x b_z c_y - a_y b_x c_z - a_z b_y c_x. \] (B.9)

The right-hand side shows that the volume product equals the determinant of the matrix constructed from the three vectors,

\[ \mathbf{a} \times \mathbf{b} \cdot \mathbf{c} = \det \begin{pmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{pmatrix} \] (volume product). (B.10)

The volume product is like the determinant antisymmetric under exchange of any pair of vectors (here columns). The volume product equals the signed volume of the parallelepiped spanned by the vectors \( \mathbf{a}, \mathbf{b}, \) and \( \mathbf{c} \).

Norms

The norm or length of a vector is defined by

\[ |\mathbf{a}| = \sqrt{a^2_x + a^2_y + a^2_z} \] (norm or length). (B.11)

The Cartesian distance (1.14) can now be written \( d(\mathbf{a}, \mathbf{b}) = |\mathbf{a} - \mathbf{b}| \), and this is of course the real reason why the norm is defined as it is.

The norm of an arbitrary \( 3 \times 3 \) matrix \( \mathbf{A} = \{a_{ij}\} \) is defined as\(^{3}\)

\[ |\mathbf{A}| = \sqrt{\sum_{ij} a_{ij}^2} \] (matrix norm). (B.12)

where both sums run over the coordinate labels \( x, y, \) and \( z \). It follows that the norm of a tensor product equals the product of the norms of the vector factors, \( |\mathbf{a} \mathbf{b}| = |\mathbf{a}| |\mathbf{b}| \), so this definition makes good sense.

### B.3 Basis vectors

The coordinate axes of a Cartesian coordinate system are straight lines with a common origin \( \mathbf{0} = (0, 0, 0) \) and direction vectors of unit length, called basis vectors\(^{4}\),

\[ \hat{\mathbf{e}}_x = (1, 0, 0), \quad \hat{\mathbf{e}}_y = (0, 1, 0), \quad \hat{\mathbf{e}}_z = (0, 0, 1). \] (B.13)

Every position \( \mathbf{x} \) may trivially be written as a linear combination of the basis vectors with the coordinates as coefficients,

\[ \mathbf{x} = x \hat{\mathbf{e}}_x + y \hat{\mathbf{e}}_y + z \hat{\mathbf{e}}_z, \] (B.14)

and similarly for any vector: \( \mathbf{a} = a_x \hat{\mathbf{e}}_x + a_y \hat{\mathbf{e}}_y + a_z \hat{\mathbf{e}}_z \).

---

\(^{3}\)To distinguish a matrix from a vector, the matrix symbol will be written in heavy boldface font. The distinction is unfortunately not particularly visible in print. With pencil on paper, \( 3 \times 3 \) matrices are sometimes marked with a double bar, \( \overline{1} \), or a double arrow, \( \overrightarrow{1} \).

\(^{4}\)We shall often use a “hat” to indicate that a vector has unit length. In some textbooks basis vectors are symbolized by “hatted” coordinate labels: \( \hat{x}, \hat{y}, \) and \( \hat{z} \). Although an efficient notation, it can be quite unreadable and will not be used here.
B. CARTESIAN COORDINATES

The basis vectors are normalized and mutually orthogonal,

\[ |\hat{e}_x| = |\hat{e}_y| = |\hat{e}_z| = 1, \quad \hat{e}_x \cdot \hat{e}_y = \hat{e}_y \cdot \hat{e}_z = \hat{e}_z \cdot \hat{e}_x = 0. \]  

(B.15)

Using these relations and (B.14) we find

\[ x = \hat{e}_x \cdot x, \quad y = \hat{e}_y \cdot x, \quad z = \hat{e}_z \cdot x, \]  

(B.16)

showing that the coordinates of a point may be understood as the normal projections of the point on the axes of the coordinate system.

Completeness

Combining (B.14) with (B.16) we obtain the identity

\[ \hat{e}_x (\hat{e}_x \cdot x) + \hat{e}_y (\hat{e}_y \cdot x) + \hat{e}_z (\hat{e}_z \cdot x) = x, \]

valid for all \( x \). Since this is a linear identity, we may remove \( x \) and express this completeness relation in a compact form by means of the tensor product (B.8),

\[ \hat{e}_x \hat{e}_x + \hat{e}_y \hat{e}_y + \hat{e}_z \hat{e}_z = 1, \]

(B.17)

where on the right-hand side the symbol 1 stands for the \( 3 \times 3 \) unit matrix.

Handedness

It must be emphasized that the handedness of the coordinate system has not entered the formalism. Correspondingly, the volume of the unit cube,

\[ \hat{e}_x \times \hat{e}_y \cdot \hat{e}_z = 1, \]

(B.18)

is always unity, independent of whether you call your coordinate system right-handed or left-handed. Handedness first shows up when you try to understand the world through a looking glass (see section B.6).

B.4 Index notation

Vector notation is sufficient for most areas of physics because physical quantities are mostly scalars like mass and charge, or vectors such as velocity and force. Sometimes it is, however, necessary to use a more powerful and transparent notation which generalizes better to complex expressions. It is called index notation or tensor notation, and consists in all simplicity of writing out the coordinate indices explicitly wherever they occur. Instead of thinking of a Cartesian position as a triplet \( x \) endowed with algebraic rules, we think of it as the set of coordinates \( x_i \) with the index \( i \) implicitly running over the coordinate labels, for example \( i = x, y, z \) or \( i = 1, 2, 3 \) or whatever, without having to state it explicitly every time.

Algebraic operations

Vector and index notations coexist quite peacefully as witnessed by the linear operations

\[ (k a)_i = k a_i, \quad (a + b)_i = a_i + b_i, \quad (a - b)_i = a_i - b_i. \]  

(B.19) \hspace{1cm} (B.20) \hspace{1cm} (B.21)

In each of these equations it is tacitly understood that the index runs over the coordinate labels, for example \( i = x, y, z \).
For the scalar product we also let the sum range implicitly over the coordinate labels,

\[ a \cdot b = \sum_i a_i b_i. \]  (B.22)

In full-fledged tensor calculus even the summation symbol is left out and understood as implicitly present for all indices that occur precisely twice in a term. In spite of this efficient notation was introduced by Einstein himself in 1916, we shall nevertheless refrain from using it here.

**Kronecker delta**

The nine dot products of the three basis vectors with themselves forms a \( 3 \times 3 \) matrix with two indices that each run implicitly over the three coordinate labels,

\[ \delta_{ij} \equiv \hat{e}_i \cdot \hat{e}_j = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{otherwise} \end{cases} \]  (B.23)

It is called the *Kronecker delta*, and the corresponding matrix \( \mathbf{1} = \{ \delta_{ij} \} \) is the well-known unit matrix.

The Kronecker delta is the first example of a true tensor of rank 2. Another is the tensor product (B.8) of two vectors, which in index notation takes the form,

\[ (ab)_{ij} = a_i b_j. \]  (B.24)

Having two (or more) indices is, however, not enough for a collection of values to earn the right to be called a tensor. In the following sections we shall see that what really characterizes such collections is the way they behave under coordinate transformations.

**Levi-Civita epsilon**

The volume products of the three basis vectors with themselves constitute a collection of 27 values, called the *Levi-Civita symbol*,

\[ \varepsilon_{ijk} \equiv \hat{e}_i \times \hat{e}_j \cdot \hat{e}_k = \begin{cases} +1 & \text{for } ijk = xyz yzx zxy, \\ -1 & \text{for } ijk = xzy yxz zyx, \\ 0 & \text{otherwise} \end{cases} \]  (B.25)

It is +1 for even permutations of \( xyz \) and −1 for odd permutations. It vanishes if any two indices coincide.

The bilinear cross product (B.7) can now be written as a double sum over two indices,

\[ (a \times b)_i = \sum_{jk} \varepsilon_{ijk} a_j b_k, \]  (B.26)

while the trilinear volume product (B.9) becomes a triple sum,

\[ a \times b \cdot c = \sum_{ijk} \varepsilon_{ijk} a_i b_j c_k. \]  (B.27)

The Levi-Civita symbol \( \varepsilon_{ijk} \) is in fact a tensor of third rank, completely antisymmetric in the three indices. We shall rarely use this notation, although it does come in handy when everything else fails.
B. CARTESIAN COORDINATES

B.5 Cartesian coordinate transformations

The same Euclidean world may be described geometrically by different observers with different Cartesian reference frames. Each observer constructs his own preferred Cartesian coordinate system and determines all positions relative to that. Every observer thinks that his basis vectors have the simple form (B.13) and satisfy the same orthogonality and completeness relations. Every observer believes he is right-handed. How can they ever agree on anything with such a self-centered view of the world?

The answer is that the two descriptions are related by a coordinate transformation. Since the length of a curve connecting any two points can be determined by laying out agreed-upon unit rulers along its path, it follows that both observers will agree on what is the shortest path, and thus on what is the distance between the points. Seen from one Cartesian coordinate system, which we shall call the “old”, the axes of another Cartesian coordinate system, called the “new”, will therefore also appear to be straight lines with a common origin. Furthermore, since the scalar product of two vectors can be expressed in terms of the norm (problem B.6), it must—like distance—be independent of the specific coordinate system, such that the new axes will also appear to be orthogonal in the geometry of the old coordinate system. Different observers will thus agree that their respective coordinate systems are indeed Cartesian.

Simple transformations

We begin the analysis of coordinate transformations with the familiar elementary ones: translation, rotation and reflection. These transformations express the coordinates of a point in the new system as a function of the coordinates of the same point in the old. The simple transformations defined below refer each to a special situation of the two coordinate systems. The most general transformation can be constructed from a combination of simple transformations (problem B.20).

**Simple translation**: A simple translation of the origin of coordinates along the *x*-axis by a constant amount *c* is given by

\[
\begin{align*}
x' &= x - c, \\
y' &= y, \\
z' &= z. 
\end{align*}
\]

(B.28a) (B.28b) (B.28c)

The axes of the new coordinate system are in this case parallel with the axes of the old.

**Simple rotation**: A simple rotation of the coordinate system through an angle *φ* around the *z*-axis is described by the transformation

\[
\begin{align*}
x' &= x \cos \phi + y \sin \phi, \\
y' &= -x \sin \phi + y \cos \phi, \\
z' &= z. 
\end{align*}
\]

(B.29a) (B.29b) (B.29c)

In this case the *z*-axes are parallel in the old and the new systems.

**Simple reflection**: A simple reflection of the *x*-axis in the *yz*-plane is described by

\[
\begin{align*}
x' &= -x, \\
y' &= y, \\
z' &= z. 
\end{align*}
\]

(B.30)

A simple reflection always transforms a right-handed coordinate system into a left-handed one, and vice-versa, independent of which hand you may claim to be the right one.
Arrangement of the old and new coordinate systems. The origin of the new system is e in the old, and the new basis vectors are \( \hat{a}_x, \hat{a}_y, \) and \( \hat{a}_z \) in the old. A point with coordinates \( x \) in the old, has coordinates \( x' \) in the new.

**General transformations**

In the coordinates of the old system, the new Cartesian coordinate system can be characterized by its origin \( e \) and its three orthogonal and normalized basis vectors \( \hat{a}_x, \hat{a}_y, \) and \( \hat{a}_z, \) satisfying the usual relations,

\[
|\hat{a}_x| = |\hat{a}_y| = |\hat{a}_z| = 1 \quad \hat{a}_x \cdot \hat{a}_x = \hat{a}_y \cdot \hat{a}_y = \hat{a}_z \cdot \hat{a}_z = 0
\]  

(B.31)

The position \( x' = (x', y', z') \) of a point in the new coordinate system must then correspond to following position in the old,

\[
x = e + x' \hat{a}_x + y' \hat{a}_y + z' \hat{a}_z.
\]  

(B.32)

The new coordinates are determined by multiplying with the new basis vectors and using orthonormality (B.31)

\[
x' = \hat{a}_x \cdot (x - e) = a_{xx}(x - c_x) + a_{xy}(y - c_y) + a_{xz}(z - c_z),
\]

\[
y' = \hat{a}_y \cdot (x - e) = a_{yx}(x - c_x) + a_{yy}(y - c_y) + a_{yz}(z - c_z),
\]

\[
z' = \hat{a}_z \cdot (x - e) = a_{zx}(x - c_x) + a_{zy}(y - c_y) + a_{zz}(z - c_z),
\]

where \( a_{ij} = \hat{a}_i \cdot \hat{e}_j = (\hat{a}_i)_j \) are the coordinates of the new basis vectors in the old system. This is the most general coordinate transformation that can connect any two Cartesian coordinate systems.

Using index notation, the general coordinate transformation may also be written (it is perhaps better here to think of integer indices, \( i = 1, 2, 3 \)),

\[
x'_i = \sum_j a_{ij}(x_j - c_j).
\]  

(B.33)

In matrix notation the transformation becomes even more compact\(^5\),

\[
x' = A \cdot (x - e),
\]  

(B.34)

where \( A = \{a_{ij}\} \) is called the transformation matrix.

For a true vector \( u = x_1 - x_2 \), defined as the difference between the coordinates of two positions, the transformation does not depend on the origin \( e \) of the new system, so that the true vector transformation rule becomes,

\[
u' = A \cdot u.
\]  

(B.35)

Since the basis vectors of the old system are true vectors they become \( \hat{e}'_i = A \cdot \hat{e}_i \) in the new, with coordinates \( (\hat{e}'_i)_j = \hat{e}'_i \cdot \hat{e}_j = a_{ij} \). This is quite different from the coordinates of the new basis in the old, \( a_{ij} \cdot \hat{e}_j = a_{ij} \). Keeping track of the various coordinates in the two systems can in fact be quite subtle!

**Example B.1:** The transformation matrix for a simple translation along the \( x \)-axis (B.28) is just the unit matrix, \( A = I \), whereas for a simple rotation around the \( z \)-axis (B.29) we obtain,

\[
A = \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}.
\]  

(B.36)

A simple reflection in the \( yz \)-plane (B.30) is characterized by a diagonal transformation matrix with \((-1, 1, 1)\) along the diagonal.

---

\(^5\)In ordinary mathematical vector and matrix calculus one would not use the dot to indicate matrix multiplication (nor to indicate a scalar product), but this notation is quite natural for the three-dimensional vectors and matrices that we encounter so often in physics.
Orthogonality and completeness of the new basis

The orthogonality and completeness of the new basis vectors imply that

\[ \hat{a}_i \cdot \hat{a}_j = \delta_{ij}, \quad \sum_i \hat{a}_i \hat{a}_i = 1. \]  

(B.37)

In index notation these two relations take the form,

\[ \sum_k a_{ik} a_{jk} = \sum_k a_{kj} a_{ki} = \delta_{ij}, \]  

(B.38)

which in matrix notation become the usual conditions for the matrix to be orthogonal,

\[ \mathbf{A} \cdot \mathbf{A}^T = \mathbf{A}^T \cdot \mathbf{A} = \mathbf{1}. \]  

(B.39)

Here \((\mathbf{A}^T)_{ij} = a_{ji}\) is the transposed matrix having the new basis vectors as columns.

The transposed matrix has the same determinant as the original matrix. Using that the determinant of a product of matrices is the product of the determinants and that the transposed matrix \(\mathbf{A}^T\) has the same determinant as \(\mathbf{A}\), it follows from eq. (B.39) that \((\det \mathbf{A})^2 = 1\), or

\[ \det \mathbf{A} = \pm 1. \]  

(B.40)

The transformation matrices are thus divided into two completely separate classes, those with determinant \(+1\), generically called rotations, and those with determinant \(-1\), generically called reflections. Since the simple reflection (B.30) has determinant \(-1\), all reflections may be composed from a simple reflection followed by a rotation.

* Infinitesimal rotations

Suppose the new basis lies very close to the old, such that we may write

\[ a_{ij} = \delta_{ij} + b_{ij}, \quad \text{with } |b_{ij}| \ll 1 \]  

(B.41)

Inserting this into the orthogonality relation (B.38) we find to first order,

\[ \delta_{ij} = \sum_k a_{ik} a_{jk} = \delta_{ij} + b_{ij} + b_{ji}, \]  

(B.42)

which shows that the infinitesimal transformation matrix is always antisymmetric,

\[ b_{ij} = -b_{ji}. \]  

(B.43)

Using the Levi-Civita symbol to write it as, \(b_{ij} = \sum_k \epsilon_{ijk} \phi_k\), it follows that

\[ \hat{a}_i = \hat{e}_i + \phi \times \hat{e}_i, \]  

(B.44)

showing that the new basis vectors are simply obtained by a common (solid) rotation of the old basis vectors through an infinitesimal angle \(|\phi|\) around the axis \(\hat{\phi} = \phi / |\phi|\).
B.6 Scalars, vectors and tensors

When you change the coordinate system the world stays the same; it is only the way you describe it that changes. Some geometrical quantities, for example the distance between two points, are unaffected by any coordinate transformation; others, like your current position or velocity, will change in a characteristic way. In physics we shall only use quantities that transform in a regular way under coordinate transformations and therefore may be viewed as representations of geometrical objects. This guarantees that physical laws take the same form in any coordinate system and therefore places all observers on an equal footing. The geometric objects are generically called tensors, although the most common and important ones have been given separate names.

Classification by pure rotations

Geometric quantities are primarily classified according to their behavior under pure rotations of the form,

\[ x' = A \cdot x \] (B.45)

with the condition that \( \det A = +1 \). The classes are

Scalars: A single quantity \( S \) is called a scalar, if it is invariant under pure rotations

\[ S' = S. \] (B.46)

The distance, the norm and the dot product are examples of scalars. In physics the natural constants, material constants, as well as mass and charge are scalars.

Vectors: Any triplet of quantities \( U \) is called a vector, if it transforms in the same way as the coordinates (B.33) under a pure rotation,

\[ U'_i = \sum_j a_{ij} U_j, \] (B.47)

or equivalently in matrix form,

\[ U' = A \cdot U. \] (B.48)

In physics, velocity, acceleration, momentum, angular momentum, force, and many other quantities are vectors. This definition of a vector requires triplets to have special transformation properties to qualify as vectors. A triplet containing your weight, your height, and your age, is not a vector but a collection of three scalars.

Tensors: Using the vector transformation (B.47) the tensor product of two vectors \( V \) \( W \) is seen to transform according to the rule,

\[ (V'W')_{ij} \equiv V'_i W'_j = \left( \sum_k a_{ik} V_k \right) \left( \sum_l a_{jl} W_l \right) = \sum_{kl} a_{ik} a_{jl} (VW)_{kl}, \]

where in the last step we have reordered the two sums into a convenient form.
More generally, any set of nine quantities arranged in a matrix

\[
\mathbf{T} = \{T_{ij}\} = \begin{pmatrix}
T_{xx} & T_{xy} & T_{xz} \\
T_{yx} & T_{yy} & T_{yz} \\
T_{zx} & T_{zy} & T_{zz}
\end{pmatrix}
\]  

is called a tensor of rank 2, provided it obeys the transformation law,

\[
T'_{ij} = \sum_{kl} a_{ik} a_{jl} T_{kl}.
\]  

In matrix form this may be written,

\[
\mathbf{T}' = A \cdot \mathbf{T} \cdot A^\top.
\]

In physics, the moment of inertia of an extended body and the quadrupole moment of an electric charge distribution are well-known tensors of second rank.

Tensors of higher rank may be constructed in a similar way. A tensor of rank \(r\) has \(r\) indices and is a collection of \(3^r\) quantities that transform as the direct product of \(r\) vectors. A scalar is a tensor of rank 0, and a vector a tensor of rank 1. We have so far only met one third rank tensor, the Levi-Civita symbol (B.25) (see problem B.22). When nothing else is said, a tensor is always assumed to be of rank 2.

**Subclassification by pure reflections**

Geometric quantities may be further subclassified according to their behavior under a pure reflection, defined as the transformation

\[
x' = -x.
\]

Instead of a simple reflection in the \(yz\)-plane, we have chosen a complete reflection of the coordinates through the origin. Geometrically, the reflection in the origin may be viewed as a composite of three simple reflections along each of the three coordinate axes, or as a simple reflection of a coordinate axis followed by a simple rotation through 180° around the same axis.

**Polar vectors:** A vector which obeys the transformation equation (B.47) under pure rotations as well as under pure reflections is called a polar vector. Under a pure reflection in the origin, the coordinates of a polar vector change sign just like the coordinates of a point,

\[
U' = -U.
\]

Since the coordinate axes all reverse direction, the geometrical position in space of a polar vector is unchanged by a reflection of the coordinate system and the vector may faithfully be represented by an arrow, also under reflection. In physics, acceleration, force, velocity, momentum, and electric dipole moments are all polar vectors.

Evidently, the basis vectors of the old coordinate system have the coordinates, \(\hat{e}'_i = -\hat{e}_i\), in the new (reflected) coordinate system. Basis vectors always transform as polar vectors under reflection. One should not get confused by the fact that the new basis vectors in the new system have (by definition) the same coordinates \(\hat{e}\) as the old basis vectors in the old system, because they refer to different geometric objects in the two coordinate systems.
**Axial vectors:** The cross product of two polar vectors, $U = V \times W$, behaves differently than a polar vector under a reflection. According to our rules for calculating the cross product, which are the same in all coordinate systems, we find

$$U' = V' \times W' = (-V) \times (-W) = V \times W = U,$$

without the expected change of sign. Since $U$ behaves as a polar vector under a proper rotation with determinant $+1$, we conclude that the missing minus sign is associated with any transformation with determinant $-1$, in other words with any reflection. Generalizing, we define an axial vector $U$ as a set of three quantities that transform according to the rule

$$U_i' = \det \mathbf{A} \sum_j a_{ij} U_j,$$

under an arbitrary Cartesian coordinate transformation without translation. The extra determinant factor counteracts the overall sign change otherwise associated with reflections of polar vectors. In physics, angular momentum, moment of force and magnetic dipole moments are all axial vectors. Axial vectors are also called *pseudovectors.*

The geometric direction of an axial vector depends on what we choose to be right and left. It is for this reason wrong to think of an axial vector as an arrow in space. It has magnitude and direction like a cylinder axis, but it is not oriented like a polar vector. The positive direction of an axial vector is not a geometric property, but a property fixed by convention. For consistency all humans (even the British) have agreed that one particular coordinate system and all coordinate systems that are obtained from it by proper rotation are right-handed, whereas coordinate systems that are related to this class by reflection are left-handed. We do not know whether non-human aliens would have adopted the same convention, but should we ever meet such beings we would be able to find the correct transformation between our reference frames and theirs.

**Pseudo-scalars:** The volume product of three polar vectors $P = U \cdot V \times W$ is a scalar quantity which changes sign under a pure reflection of the coordinate system. More generally a *pseudo-scalar* transforms like

$$P' = \det \mathbf{A} P,$$

under a Cartesian coordinate transformation without translation.

The sign of a pseudo-scalar is not absolute, but depends on the handedness of the coordinate system, and thus on convention. One might think that physics had no use for such quantities, because after all physics itself does not depend on coordinate systems, only its mathematical description does. Nevertheless, magnetic charge, if it is ever found, would be pseudo-scalar, and more importantly some of the familiar elementary particles, for example the pi-mesons, are described by pseudo-scalar fields.

**Pseudo-tensors:** The transformation properties of pseudo-tensors of higher rank also include an extra factor of the determinant of the transformation matrix. The Levi-Civita symbol is a pseudo-tensor of third rank (problem B.22).
Subclassification by pure translations

A pure translation takes the form

\[ x' = x - c. \]  \hspace{1cm} (B.57)

Translationally invariant vectors (and more generally, tensors) are called proper, whereas tensors that change under translations are called improper. In physics the center of mass and the electric dipole moment are improper polar vectors, whereas angular momentum, moment of force, and magnetic dipole moment are improper axial vectors. The moment of inertia and all kinds of higher multipole moments are improper tensors.

Problems

Distance

**B.1** Any distance function must satisfy the axioms

\[ d(a, a) = 0, \]  \hspace{1cm} (B.58a)

\[ d(a, b) = d(b, a), \]  \hspace{1cm} (symmetry) (B.58b)

\[ d(a, b) \leq d(a, c) + d(c, b), \]  \hspace{1cm} (triangle inequality) (B.58c)

Show that the Cartesian distance function (1.14) satisfies these axioms.

**B.2** Calculate the distance between two points on the spherical Earth in terms of longitude \( \alpha \), latitude \( \delta \) and height \( h \) over the surface with radius \( a \).

**B.3** Show using Cartesian coordinates that the shortest path (geodesic) between two arbitrary points \( a \) and \( b \) is a described by the linear function \( (1 - \xi) a + \xi b \) with \( 0 \leq \xi \leq 1 \).

Vector algebra

**B.4** Let \( a = (2, 3, -6) \) and \( b = (3, -4, 0) \). Calculate (a) the lengths of the vectors, (b) the dot product, (c) the cross product, (d) and the tensor product.

**B.5** Are the vectors \( a = (3, 1, -2) \), \( b = (4, -1, -1) \) and \( c = (1, -2, 1) \) linearly dependent (meaning that there exists a non-trivial set of coefficients such that \( a \alpha + b \beta + c \gamma = 0 \))?

**B.6** Show that

\[ |a \cdot b| \leq |a| |b|. \]  \hspace{1cm} (B.59a)

\[ |a + b| \leq |a| + |b|. \]  \hspace{1cm} (B.59b)

**B.7** Show that

\[ (a \times b \cdot c) (d \times e \cdot f) = \begin{vmatrix} a \cdot d & a \cdot e & a \cdot f \\ b \cdot d & b \cdot e & b \cdot f \\ c \cdot d & c \cdot e & c \cdot f \end{vmatrix}. \]  \hspace{1cm} (B.60)

**B.8** Show that

\[ a \times b \cdot c = a \cdot b \times c = b \cdot c \times a, \]  \hspace{1cm} (B.61)

\[ (a \times b) \times c = (a \cdot c) b - (b \cdot c) a. \]  \hspace{1cm} (B.62)

\[ (a \times b) \cdot (c \times d) = (a \cdot c)(b \cdot d) - (a \cdot d)(b \cdot c). \]  \hspace{1cm} (B.63)

\[ |a \times b|^2 = |a|^2 |b|^2 - (a \cdot b)^2. \]  \hspace{1cm} (B.64)
B.9 Show that (with the normal definition of the matrix product) the following relations make sense for the tensor product

\[(ab) \cdot c = a(b \cdot c)\]  \hspace{1cm} (B.65)
\[a \cdot (bc) = (a \cdot b)c\]  \hspace{1cm} (B.66)

This is sometimes quite useful.

B.10 Show that

\[(a \times b \cdot c)d = a \times b(c \cdot d) + b \times c(a \cdot d) + c \times a(b \cdot d)\]  \hspace{1cm} (B.67)

for arbitrary vectors \(a\), \(b\), \(c\), and \(d\).

Index notation

B.11 Show that

\[\sum_i \delta_{ii} = 3\]  \hspace{1cm} (B.68)
\[\sum_j \delta_{ij} \delta_{jk} = \delta_{ik}\]  \hspace{1cm} (B.69)

B.12 Show that

\[\sum_i \nabla_i x_j = 3,\]  \hspace{1cm} (B.70)
\[\nabla_i x_j = \delta_{ij},\]  \hspace{1cm} (B.71)
\[\nabla_i \nabla_j (x_k x_l) = \delta_{ik} \delta_{jl} + \delta_{ij} \delta_{jk}\]  \hspace{1cm} (B.72)

B.13 Show that the Levi-Civita symbol is completely antisymmetric in all three indices,

\[\epsilon_{ijk} = -\epsilon_{ikj} = -\epsilon_{kji}\]  \hspace{1cm} (B.73)

B.14 Show that the product of two Levi-Civita symbols is (see problem B.7)

\[\epsilon_{ijk} \epsilon_{lmn} = \begin{vmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{vmatrix}\]  \hspace{1cm} (B.74)

and derive from this

\[\sum_k \epsilon_{ijk} \epsilon_{lmk} = \begin{vmatrix} \delta_{il} & \delta_{im} \\ \delta_{jl} & \delta_{jm} \end{vmatrix} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}\]  \hspace{1cm} (B.75)
\[\sum_{jk} \epsilon_{ijk} \epsilon_{ljk} = 2\delta_{lj},\]  \hspace{1cm} (B.76)
\[\sum_{ijk} \epsilon_{ijk} \epsilon_{ijk} = 6\]  \hspace{1cm} (B.77)

Transformations

B.15 Show that the trace \(\sum_i T_{ii}\) of a tensor is invariant under a rotation.

B.16 Show that the Kronecker delta transforms as a tensor.
B.17 Show that the distance \(|x - y|\) is invariant under any transformation between Cartesian coordinate systems.

* B.18 Show that if \(W_1 = \sum_i T_{ij} V_j\) and if it is known that \(W\) is a vector for all vectors \(V\), then \(T_{ij}\) must be a tensor. This is called the quotient rule.

* B.19 Consider two Cartesian coordinate systems and make no assumptions about the transformation \(x_0 = f(x)\) between them. Show that the invariance of the distance, 
\[
|x' - y'| = |x - y|, \tag{B.78}
\]
implies that the transformation is of the form 
\[
x_0 = Ax + b \tag{B.79}
\]
where \(A\) is an orthogonal matrix.

* B.20 Show that the general Cartesian coordinate transformation may be built up from a combination of simple translations, rotations and reflections.

* B.21 Show that under a simple rotation, a tensor \(T_{ij}\) transforms into 
\[
T_{ij}' = \cos \phi T_{xx} \cos \phi + \sin \phi T_{xy} \sin \phi, \tag{B.79a}
\]
\[
T_{ij}' = \cos \phi T_{xy} \cos \phi + \sin \phi T_{yy} \sin \phi, \tag{B.79b}
\]
\[
T_{ij}' = \cos \phi T_{xz} \sin \phi, \tag{B.79c}
\]
\[
T_{ij}' = -\sin \phi T_{xx} \cos \phi + T_{xy} \sin \phi + \cos \phi T_{yx} \cos \phi, \tag{B.79d}
\]
\[
T_{ij}' = \cos \phi T_{yx} \cos \phi + T_{yy} \sin \phi + \sin \phi T_{zy} \cos \phi, \tag{B.79e}
\]
\[
T_{ij}' = -\sin \phi T_{xz} + \cos \phi T_{yz}, \tag{B.79f}
\]
\[
T_{ij}' = T_{xx} \cos \phi + T_{xy} \sin \phi, \tag{B.79g}
\]
\[
T_{ij}' = -T_{xz} \sin \phi + T_{yz} \cos \phi, \tag{B.79h}
\]
\[
T_{ij}' = T_{zz}. \tag{B.79i}
\]

* B.22 Show that

(a) the Levi-Civita symbol satisfies
\[
\sum_{lmn} a_{ijl} a_{jmn} \varepsilon_{lmn} = \det \Lambda \varepsilon_{ijk} \tag{B.80}
\]
where \(\Lambda\) is an arbitrary matrix.

(b) the Levi-Civita symbol (which by the definition of the cross product must be invariant, \(\epsilon_{ijk}' = \varepsilon_{ijk}\)) obeys the rule
\[
\epsilon_{ijk}' = \varepsilon_{ijk} = \det \Lambda \sum_{lmn} a_{ijl} a_{jmn} \varepsilon_{lmn} \tag{B.81}
\]
for an arbitrary coordinate transformation (which has \(\det \Lambda = \pm 1\)).

(c) the cross product of two vectors \(W = U \times V\) must transform like
\[
W'_{ij} = \det \Lambda \sum_j a_{ij} W_j \tag{B.82}
\]

B.23 Show that
\[
\frac{\partial (a \cdot b)}{\partial a} = b \tag{B.83}
\]
and that
\[
\frac{\partial |a|}{\partial a} = \frac{a}{|a|}. \tag{B.84}
\]