Chapter 13
Introduction to covariance: rotations and vectors

December 3, 2019

Covariance is introduced through the transformation between different cartesian coordinates in 3D space. The concepts of scalars, vectors and tensors, as well as scalar, vector and tensor fields, are introduced.

Contents

1 Introduction

1.1 Electrodynamics and relativity

The next few lectures focus on a central question:[1]

> How can the theory of electrodynamics be written in a way that is manifestly consistent with relativity?

Electrodynamics and relativity intersect at many points. Relativity is rooted in the observation that the speed of light,[2] c, is the same for all observers and in all directions. In this (historical) sense, relativity depends on EM (in particular light). But a more modern view is that relativity is more fundamental than EM; c appears so prominently in relativity not because it is the speed of *light* or *photons*, but because it is the limiting speed for *any* object or signal. Indeed, c is a property of the stage (spacetime itself) and not a property of the players on the stage (the particles and fields); at a deeper level, c appears only because of the (stupid) units that we use.

Behind all this is the concept of *covariance* — more important than EM itself:

> Physics should appear the same to all observers, and should be described by the same laws according to all observers.

We start by introducing covariance through simpler contexts.

1.2 One phenomenon, different observers

Consider the classic experiment of Galileo **Figure 1**: two stones dropped from a tower strike the

[1] In these lectures, the term "relativity" will mean special relativity, unless otherwise specified.

[2] In these lectures, unless otherwise specified, we shall always mean the speed of light *in vacuum*.

ground at the same time.[3] This fact of *physics* is the same whether we discuss it in (a) rectangular coordinates referred to axes that are horizontal and vertical; (b) rectangular coordinates referred to axes that are tilted; (c) polar coordinates; or (d) generalized coordinates of any type. Therefore we say

Physics is absolute. Coordinates are arbitrary. Therefore physics should be independent of coordinates.

To make sure that physics is independent of coordinates, we must first ask: What happens under coordinate transformations? Special relativity is concerned with transformation between coordinate systems in relative motion; general relativity is concerned with transformation between arbitrary (in general curvilinear) coordinates.[4] To introduce the idea of covariance, we first deal with transformation between two coordinate systems without relative motion, both described by rectangular coordinates and differing only in orientation.

1.3 Rectangular coordinates

Rectangular coordinates are often denoted by the components of a vector, so we first introduce the notation for vectors.

Notation for vectors
Vectors[5] in 3D are denoted as $\mathbf{x}$, $\mathbf{y}$, $\mathbf{p}$ etc. In cartesian coordinates $\mathbf{x} = (x^1, x^2, x^3)$. For reasons that will be apparent later, we use upper indices for coordinates (at least to start with). The components are collectively denoted as x^i. An index such as i, j runs from 1 to 3 for the three cartesian directions. The axes are called the 1-axis, 2-axis etc. The length of $\mathbf{x}$ is denoted as $|\mathbf{x}| = x$. Unit basis vectors are $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Their components are $e^i_j = i$-th component of $\mathbf{e}_j$, with $e^i_j = \delta^{ij}$.[6]

It is also convenient to express $\mathbf{x}$ as a column vector:

$$[x] \quad = \quad \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix}$$

[3]To be very precise, this experiment refers to the ideal case where there is no air resistance.

[4]Made necessary because spacetime is curved, and rectangular coordinates do not exist, except in infinitesimal neighborhoods.

[5]The word "vector" at the moment is used heuristically; later we shall specify conditions for three numbers to form a vector.

[6]Sometimes the components of $\mathbf{x}$ are also denoted as x, y, z and the axes as the x, y, z axes. Sometimes the unit vectors are also denoted as $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, $\hat{\mathbf{k}}$. The notation will be clear from the context.

Example 1
A rod of length L is inclined at an angle ϕ to the 1-axis (**Figure 2a**). Let its tip be $\mathbf{x}$. Then in this coordinate system S,

$$x^1 = L\cos\phi \qquad , \qquad x^2 = L\sin\phi$$

A new coordinate system S' has the $1'$-axis aligned along the rod (**Figure 2b**); thus.

$$x'^1 = L \qquad , \qquad x'^2 = 0$$

So the coordinate components are different; they are *relative.*

Notation
Some books denote the components in S' as $x^{1'}, x^{2'}$, to emphasize that "x" is the same (i.e., the same vector), and only "1","2" are different (i.e., different components). We do not adopt the more cumbersome (though more logical) notation.

1.4 Rotation in 2D

A more formal and general discussion of rotations is given in the next Section. Here the idea is first illustrated in 2D. Let

$$\begin{aligned} \alpha &= \text{angle between the two frames} \\ \phi &= \text{angle of } \mathbf{x} \text{ with respect to 1-axis} \\ \phi' &= \text{angle of } \mathbf{x} \text{ with respect to } 1'\text{-axis} \end{aligned}$$

See **Figure 3**. Then

$$x^1 = x\cos\phi \qquad , \qquad x^2 = x\sin\phi$$

But in the new coordinate system

$$\begin{aligned} x'^1 &= x\cos\phi' = x\cos(\phi-\alpha) \\ &= x\cos\phi\cos\alpha + x\sin\phi\sin\alpha \\ &= x^1\cos\alpha + x^2\ \sin\alpha \\ x'^2 &= x\sin\phi' = x\sin(\phi-\alpha) \\ &= x\sin\phi\cos\alpha - x\cos\phi\sin\alpha \\ &= x^2\cos\alpha - x^1\sin\alpha \end{aligned}$$

This can be written in matrix form

$$\begin{bmatrix} x'^1 \\ x'^2 \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} \tag{1}$$

or

$$\boxed{[x'] = [R]\,[x]} \tag{2}$$

The matrix $[R] = [R(\alpha)]$ depends only on the relation between the two sets of axes. It is the same for all vectors $\mathbf{x}$.

The relation (1) or (2) can also be expressed in terms of components.

$$x'^i = \sum_{j=1}^{2} R^{ij} x^j \quad , \quad i = 1, 2 \tag{3}$$

Index notation
An index such as i in (3) is called a *free* index. It appears once on each side, and can take on any value (1, 2 in 2D; 1, 2, 3 in 3D). Unless otherwise specified, it is allowed to take on each of these values successively; thus (3) is really n equations in n dimensions. On the other hand, an index such as j in (3) is called a *dummy* index. It appears twice in the same term, and is summed over all allowed values.

Summation convention
To save writing, every index $(i, j, \ldots)$ that appears twice in the same term is understood to be summed over. Thus (3) can be written simply as

$$\boxed{x'^i = R^{ij} x^j}$$

Problem 1
Write out the following equations explicitly in terms of components. Choose any one value for each free index. (a) $S = a^i b^i$; (b) $a^i = A^{ij} b^j$; (c) $C^{ik} = A^{ij} B^{jk}$; (d) $C^{ik} = A^{ij} B^{kj}$; (e) $S = A^{ii}$; (f) $S = A^{ij} B^{ji}$. §

Problem 2
Write the above expressions in terms of the matrices $[A], [B], [C]$ and the column vectors $[a], [b]$. §

Problem 3
Denote the matrix in (1) as $[R(\alpha)]$. Verify that

$$[R(-\alpha)]\,[R(\alpha)] \quad = \quad [I]$$

where $[I]$ is the identity matrix. What is the physical meaning of this mathematical relationship? §

Problem 4
Show that $[R(-\alpha)] = \left[R(\alpha)^T\right]$. Hence show that

$$\left[R(\alpha)^T\right] [R(\alpha)] \quad = \quad [I]$$

Matrices satisfying this condition are said to be *orthogonal*. §

Problem 5
Verify from (1) that

$$[R(\alpha_2)]\,[R(\alpha_1)] \quad = \quad [R(\alpha_1{+}\alpha_2)]$$

What is the physical meaning of this equation? §

1.5 How do physical laws appear to different observers

Consider a trivial "law". Let there be two rods, of length L and $2L$, along the same direction. Each rod has one end at the origin O (**Figure 4**). Let the other end-points be $\mathbf{x}$ and $\mathbf{y}$. Then the "law" is $\mathbf{y} = 2\mathbf{x}$. According to one observer S, the components satisfy

$$y^i \quad = \quad 2x^i \tag{4}$$

or in column vector form

$$[y] \quad = \quad 2\,[x] \tag{5}$$

Multiply by the rotation matrix $[R(\alpha)]$:

$$\begin{aligned} [R(\alpha)]\,[y] &= 2\,[R(\alpha)]\,[x] \\ [y'] &= 2\,[x'] \end{aligned} \tag{6}$$

or back in component form

$$y'^i \quad = \quad 2x'^i \tag{7}$$

Compare (4) and (7), or (5) and (6). Although the variables change (e.g., $y'^1 \neq y^1$), a valid law of physics takes exactly the same form in the two coordinate systems. We say the variables are *covariant*: they transform in the same way. We say the laws are *invariant*: they stay in the same form.

Although $\mathbf{y} = 2\mathbf{x}$ is a very trivial "law", this concept generalizes to all laws of physics. For example $\mathbf{F} = m\mathbf{a}$ is a more relevant law of exactly the same type — it too has the same form in all coordinate systems.

Principle of relativity
The above idea is elevated to a principle.

All valid laws of physics should take the same form in different coordinate systems S and S'.

Thus there are two types of questions.

- How are the variables in different coordinate systems related to each other?
- What laws of physics are compatible with the principle of relativity? Only these laws are allowed.

Key idea
This Chapter is rather lengthy, but the key idea is simple, and may be summarized into two points.

The transformation property is determined by the free indices. If all terms in an equation have the same free indices, then they all transform in the same way. Then if the equation holds in one frame, it will hold in every frame.

The rest of this Chapter just establishes these statements systematically.

2 Rotation of coordinates

2.1 General formulation

We now give a general formulation of rotations valid in any number of dimensions.

Basic object
The basic object is a point P, whose coordinates are

$$\begin{aligned}\mathbf{x} &= \left(x^1, x^2, x^3\right) \text{ in } S \\ \mathbf{x} &= \left(x'^1, x'^2, x'^3\right) \text{ in } S'\end{aligned}$$

The components are different, but the vector is the same (see **Figure 5**); so as a vector, they are both denoted by the same symbol $\mathbf{x}$.

The question is: How are the components related?

Defining a rotation
A rotation is a transformation of the coordinates that satisfies two properties: (a) it is linear; (b) it preserves lengths. The latter means that $|\mathbf{x}|^2$ is preserved. But for three vectors $\mathbf{z} = \mathbf{x} + \mathbf{y}$, we get

$$|\mathbf{z}|^2 = |\mathbf{x}|^2 + 2\mathbf{x}\cdot\mathbf{y} + |\mathbf{y}|^2$$

So it is obvious that dot products $\mathbf{x}\cdot\mathbf{y}$ are also preserved.

Linear property
First assume the origins of S and S' coincide. (If they do not, simply add a trivial shift.) Next, assume that the relationship between the components is linear. Then the most general transformation is

$$\begin{aligned}x'^i &= R^{ij}x^j \\ [x'] &= [R]\,[x] \qquad (8)\end{aligned}$$

which has been written first in component form and then in matrix form; here $[R]$ is an unknown matrix, called the rotation matrix. Note that the summation convention is employed.

Notation
The ij component of a matrix $[A]$ is written as $[A]^{ij} = A^{ij}$.

Identify an invariant
Although the components change (e.g., $x'^1 \neq x^1$), there is an *invariant*, i.e., a quantity which is the same in both frames. By Pythagoras' theorem

$$\begin{aligned}\sigma^2 &= \sum_i (x^i)^2 = x^i x^i \\ \sigma'^2 &= \sum_i (x'^i)^2 = x'^i x'^i\end{aligned}$$

are the same. The invariant condition can be written as follows:

$$\begin{aligned}x'^i x'^i &= x^i x^i \\ \left[x'^T\right][x'] &= \left[x^T\right][x] \qquad (9)\end{aligned}$$

Note that, for example in 3D

$$[x] = \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix} \quad , \quad \left[x^T\right] = \left[x^1, x^2, x^3\right]$$

$$\begin{aligned}\left[x^T\right][x] &= \left[x^1, x^2, x^3\right]\cdot \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix} \\ &= x^1x^1 + x^2x^2 + x^3x^3\end{aligned}$$

2.2 Condition on rotation matrix

If we choose an arbitrary matrix $[R]$, the result of the transformation (8) would not satisfy the invariant condition (9). The existence of an invariant places conditions on $[R]$. These conditions are derived in three equivalent ways.

(1) Explicitly in components for 2D
Let

$$[R] = \begin{bmatrix} p & s \\ q & r \end{bmatrix} \qquad (10)$$

Then (8) gives

$$\begin{aligned}x'^1 &= px^1 + sx^2 \\ x'^2 &= qx^1 + rx^2\end{aligned}$$

Hence

$$\begin{aligned}\sigma'^2 &= (x'^1)^2 + (x'^2)^2 \\ &= (p^2+q^2)(x^1)^2 + 2(ps+qr)x^1x^2 \\ &\quad + (s^2+r^2)(x^2)^2\end{aligned}$$

But this must equal

$$\sigma^2 = 1\cdot(x^1)^2 + 0\cdot x^1x^2 + 1\cdot(x^2)^2$$

as an identity. Hence we get three conditions

$$\begin{aligned}p^2+q^2 &= 1 \\ ps+qr &= 0 \\ s^2+r^2 &= 1 \qquad (11)\end{aligned}$$

Problem 6
Derive (11) and show that the solution is (in terms of s) $p = r = \sqrt{1-s^2}$, $q = -s$. Another solution is discarded; explain. §

(2) In general using index notation

$$\begin{aligned}x'^i &= R^{ij}x^j \\ x'^i &= R^{ik}x^k\end{aligned}$$

Note we use different dummy indices. Multiplying

$$\begin{aligned}\sigma'^2 &= x'^i x'^i = (R^{ij}x^j)(R^{ik}x^k) \\ &= (R^{ij}R^{ik})\, x^j x^k \\ \sigma^2 &= x^i x^i = \delta^{jk}\, x^j x^k\end{aligned}$$

Since they must be equal as an identity, we get

$$\boxed{R^{ij}R^{ik} = \delta^{jk}} \tag{12}$$

Here, jk are free indices, while i is a dummy index. Also, both sides are symmetric under $j \leftrightarrow k$, so there are 3 conditions in 2D ($jk = 11, 22, 12$) and 6 conditions in 3D ($jk = 11, 22, 33, 12, 23, 31$).

Let us check that in the case of 2D, (12) agrees with (11). We put the summation sign back explicitly. As an example, for $j = k = 1$

$$\begin{aligned}\sum_i R^{i1}R^{i1} &= 1 \\ R^{11}R^{11} + R^{21}R^{21} &= 1 \\ p^2 + q^2 &= 1\end{aligned}$$

Problem 7
Complete the calculation and check the other two conditions in (11). §

Problem 8
Generalize to n dimensions. How many conditions are there in (12)? §

(3) Using matrix notation

$$\begin{aligned}[x'] &= [R][x] \\ [\, x'^T\,] &= [x^T][R^T] \\ \sigma'^2 &= [x'^T][x'] = [x^T][R^T][R][x] \\ \sigma^2 &= [x^T][x] = [x^T][I][x]\end{aligned}$$

Thus

$$\boxed{[R^T][R] = [I]} \tag{13}$$

Problem 9
Check that (13) is the same as (12). §

Orthogonal matrices
A matrix satisfying (13) is said to be *orthogonal*. Rotations are represented by orthogonal matrices.

Determinant
From (13) we find

$$\begin{aligned}|\det[R]|^2 &= 1 \\ \det[R] &= \pm 1\end{aligned}$$

For all rotations that are continuously linked to the identity, we must have

$$\det[R] = +1 \tag{14}$$

These are called *proper rotations*. We shall implicitly restrict to these only. The ones that have determinant -1 are continuously linked to a reflection, say

$$[R] = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

2.3 Number of parameters

Consider the number of free parameters in the matrix $[R]$ for space of n dimensions. There are $p = n^2$ parameters in an arbitrary matrix $[R]$, but the condition that the length is preserved leads to c constraints, so there are only $f = p - c$ free parameters left.

n	p	c	f
3	9	6	3
2	4	3	1

Table 1. Number of parameters

Thus in 2D, we should be able to express the most general rotation in terms of 1 parameter only, obviously the angle of rotation. In 3D, we should be able to express the most general rotation in terms of 3 parameters (e.g., the 3 Euler angles). We shall not go into the 3D case in detail.

Problem 10
Complete the above table with one more row for arbitrary n. §

Problem 11
Go back to 2D. Of the 4 parameters p, q, r, s in (10), regard s as the free parameter and define $s = \sin\alpha$. (Why is this allowed?) Find p, q, r in terms of α by using the three equations in (11). Your answer should be as in (1). You will need to choose the sign of a square root. Explain the physical meaning of your choice. §

2.4 Combining two transformations

Consider successive rotations about a fixed axis. We know that rotational angles add (**Figure 6**). Mathematically, this can be stated as follows

$$[R(\alpha_2)][R(\alpha_1)] = [R(\alpha_1 + \alpha_2)] \tag{15}$$

If the LHS is applied to a vector $\mathbf{x}$, we get

$$[R(\alpha_2)][R(\alpha_1)][x] = [R(\alpha_2)]\,\{[R(\alpha_1)][x]\}$$

The rotation $[R(\alpha_1)]$ is done first. So matrix products should be read from right to left. The order of operations does not matter in 2D, but is important in 3D.

2.5 Specifying the rotation

A rotation can be specified in different ways.

By the angle
A rotation can be specified by the angle α. Then

$$[R] = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \quad (16)$$

Note that the parameter α is additive.

Problem 12
Using (15) and (16), derive the addition laws for $\cos(\alpha_1+\alpha_2)$, $\sin(\alpha_1+\alpha_2)$. §

By the sine
Or we can specify a rotation by the parameter $s = \sin\alpha$. Then

$$[R] = \begin{bmatrix} \sqrt{1-s^2} & s \\ -s & \sqrt{1-s^2} \end{bmatrix}$$

This is less convenient for two reasons. First, there is a square root. Secondly, s is not additive.

By the tangent
Or we can specify the rotation by the the parameter $\beta = \tan\alpha$ (at least for $|\alpha| < \pi/2$).

Problem 13
(a) Write out $[R]$ in terms of β.
(b) Let $\beta_1 = \tan\alpha_1$, $\beta_2 = \tan\alpha_2$, $\beta = \tan(\alpha_1+\alpha_2)$. Show that the law of addition for β's is

$$\beta = \frac{\beta_1 + \beta_2}{1 - \beta_1\beta_2}$$

This will look familiar (except for a sign) when we come to the addition of velocities in relativity. §

3 Scalars, vectors and tensors

3.1 Scalars and vectors

In elementary courses, probably you were told that a scalar is a quantity "without direction", and a vector is a quantity "with magnitude and direction". This is nonsense; it begs the question: What is "direction"?

Scalars
A scalar (or 3-scalar, to distinguish from concepts in Lorentz transformation to be introduced in the next Chapter) is any quantity that remains unchanged under a rotation of axes. Examples: mass, temperature, electric potential, mass density, time. A scalar is also called an *invariant* under rotations.

Basic vector
The basic vector (or 3-vector, again to distinguish from concepts in Lorentz transformation) is a displacement $\Delta\mathbf{x}$, i.e., a line segment joining two neighboring points. Its cartesian coordinates are[7]

$$\Delta x^i = (\Delta x^1, \Delta x^2, \Delta x^3)$$

Transformation laws
Under a rotation of axes, the components change according to

$$\begin{aligned} \Delta x'^i &= R^{ij}\Delta x^j \\ [\Delta x'] &= [R][\Delta x] \end{aligned}$$

where the matrix $[R]$ is independent of the vector.

Other vectors
Any three quantities $v^i = (v^1, v^2, v^3)$ that transform in the same way, i.e.,

$$\begin{aligned} v'^i &= R^{ij}v^j \\ [v'] &= [R][v] \end{aligned} \quad (17)$$

is a vector by definition. Obviously vectors may be obtained by adding or subtracting vectors, or multiplying or dividing a vector by a scalar (or invariant).

The time elapsed, Δt, is a scalar; so is the mass m. So the velocity and the momentum

$$\mathbf{v} = \frac{\Delta\mathbf{x}}{\Delta t} \ , \ \ \mathbf{p} = m\mathbf{v}$$

are also vectors. And so is the force $\mathbf{F} = \Delta\mathbf{p}/\Delta t$.

3.2 Rank-2 tensor

Basic rank-2 tensor
Let $\mathbf{x}, \mathbf{y}$ be two vectors and define the following 9 quantities as the basic rank-2 tensor.

$$t^{ij} = x^i y^j$$

Its transformation law is

$$\begin{aligned} t'^{ij} &= x'^i y'^j = (R^{il}x^l)(R^{jk}y^k) \\ &= R^{il}R^{jk}x^l y^k \\ t'^{ij} &= R^{il}R^{jk}t^{lk} \end{aligned} \quad (18)$$

General rank-2 tensor
Any 9 quantities t^{ij} which transform in this way is called a rank-2 tensor.

[7] We discuss a short displacement rather than the coordinate $\mathbf{x}$ itself (which would be a long displacement starting from the origin) because eventually (though not in this course) we wish to generalize to curved space. In curved space (e.g., the surface of the earth), a *short*, indeed infinitesimal, displacement is a straight arrow, and is a vector. A *long* displacement is not a straight arrow, and cannot be thought of as a vector.

Higher rank tensors
A rank 3 tensor is 3^3 quantities which transform as

$$t'^{ijk} \quad = \quad R^{il} R^{jm} R^{kn}\, t^{lmn}$$

The definitions for higher ranks are similar.

3.3 Examples of tensors

Several special tensors need to be mentioned.

Kronecker delta
We cannot simply write down any 3^2 quantities t^{ij} and say it is a rank-2 tensor. We must check the transformation laws. Consider the Kronecker δ. In *every* frame

$$\delta^{ij} \quad = \quad \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

In other words $\delta'^{ij} = \delta^{ij}$. Is it a rank 2 tensor? According to (18), we have to check

$$\delta^{ij} = \delta'^{ij} \quad \stackrel{?}{=} \quad R^{il} R^{jk} \delta^{lk}$$

But this condition is exactly the same as (12). So δ^{ij} is indeed a tensor.

Antisymmetric symbol
The Levi-Civita or antisymmetric symbol ϵ^{ijk} is defined as follows in *every* frame.

- $+1$ if $(ijk) = (123), (231), (312)$;
- -1 if $(ijk) = (321), (213), (132)$;
- 0 if two or more indices are the same.

In fact, ϵ^{ijk} can be defined by the following formula for the determinant of any 3×3 matrix $[A]$:

$$\det{[A]} \quad = \quad \epsilon^{ijk}\, A^{1i}\, A^{2j}\, A^{3k}$$

We claim ϵ^{ijk} is a rank-3 tensor. As before, we need to check

$$\epsilon^{ijk} = \epsilon'^{ijk} \quad \stackrel{?}{=} \quad R^{i\ell} R^{jm} R^{kn} \epsilon^{\ell mn}$$

The first equal sign is because ϵ^{ijk} is defined to have the same values in every frame.

Let us take the case of $(ijk) = (123)$. Then this amounts to checking

$$R^{1\ell} R^{2m} R^{3n} \epsilon^{\ell mn} \quad \stackrel{?}{=} \quad +1$$

But the LHS is just $\det[R]$, which is $+1$ for a proper rotation. The other components are similar.

Moment of inertia
Consider a mass m at a position $\mathbf{x}$. Its contribution to the moment of inertia is

$$I^{ij} \quad = \quad m(x^2 \delta^{ij} - x^i x^j)$$

Because (a) $x^i x^j$ is the fundamental rank 2 tensor; (b) δ^{ij} is a tensor as just proved; and (c) $x^2 = |\mathbf{x}|^2$ is a scalar, so I^{ij} is a rank-2 tensor.

More generally, if there are masses m_α at positions $\mathbf{x}_\alpha, \alpha = 1, \cdots, N$, then the moment of inertia

$$I^{ij} \quad = \quad \sum_\alpha m_\alpha (x_\alpha^2 \delta^{ij} - x_\alpha^i x_\alpha^j)$$

is also a tensor.

Problem 14
Check that this definition of moment of inertia agrees with the elementary definition in terms of perpendicular distance to an axis. §

3.4 Contraction theorem

Forward theorem
Instead of stating this theorem generally, we consider an example.

Let x^i and y^i be vectors. Then

$$\varphi \quad = \quad x^i y^i$$

$\phi' = x'^j y'^j$
$\sim (Rx)(Ry)$
$\sim RRxy$

$\sim (R^T R) xy$
$\sim xy$.
$= \phi$.

is a scalar; in other words, the dummy indices are "eaten up", no free indices remain, so the result can be "read off" as a scalar. The invariance of dot products has already been shown, but we here give another proof which can be generalized to situations with more indices.

$$\begin{aligned} x'^k y'^k \quad &= \quad (R^{ki} x^i)\,(R^{kj} y^j) \\ &= \quad (R^{ki} R^{kj})\, x^i y^j = \delta^{ij}\, x^i y^j \\ &= \quad x^i y^i \end{aligned}$$

The key step is the orthogonal property of $[R]$, which (schematically and when indices are suitably placed) allows us to cancel two powers of $[R]$: $[R][R] \mapsto [I]$. (We say schematically because as a matrix equation, one of the $[R]$ factors should be transposed.)

Problem 15
Let t^{ij} be a tensor and x^j be a vector. Then prove that

$$y^i \quad = \quad t^{ij} x^j$$

is a vector. In other words, the dummy index j is contracted away, leaving a free index i, which transforms like the component of a vector. §

Reverse theorem
The reverse theorem is also true. Again we deal with an example. Let x^i be any three variables such that for any vector y^i

$$\varphi \quad = \quad x^i y^i$$

$\phi' = \phi$

$x'y' = xy$

$Rx\, y' = xy$

$\cancel{R} R x\, y' = x R y$

$x y' = x R y$.

is a scalar. Then x^i is a vector.

To prove this, start from

$$x^i y^i = x'^k y'^k = x'^k \left(R^{ki} y^i\right)$$

Since this holds for all y^i (in fact, it is sufficient if this holds for three independent sets of y^i) we can peel off that factor and get

$$x^i = R^{ki} x'^k$$

Then multiply by R^{mi} and sum over i:

$$\begin{aligned} R^{mi} x^i &= \left(R^{mi} R^{ki}\right) x'^k \\ &= \delta^{mk}\, x'^k = x'^m \\ x'^m &= R^{mi} x^i \end{aligned}$$

Problem 16
Let x^j be any three numbers. Further suppose that it is known that for any tensor t^{ij}

$$y^i = t^{ij} x^j$$

transforms like a vector. Show that x^j is also a vector. §

Cross product
Given the forward contraction theorem and the tensor property of ϵ^{ijk}, it follows that given two vectors $\mathbf{x}$ and $\mathbf{y}$, the cross product $\mathbf{z} = \mathbf{x} \times \mathbf{y}$ is a vector, because

$$z^i = \epsilon^{ijk}\, x^j y^k$$

has one free index left.

As another example consider the definition of the magnetic field $\mathbf{B}$ in terms of the force on a moving charge:

$$\begin{aligned} \mathbf{F} &= q\mathbf{v} \times \mathbf{B} \\ F^i &= q\, \epsilon^{ijk} v^j B^k \end{aligned}$$

By the reverse contraction theorem, the quantities B^k form a vector.

Summary
In summary, the essence of the contraction theorem is this: Transformation properties can be read off from free indices. If an equation relates terms with the same free indices, then these terms will transform in the same way ($y^i = 2x^i$ being one trivial example), and thereby guarantee covariance.

Example
The upshot of all this can be illustrated with an example.

Recall that the linear velocity of a point undergoing rotations is given by

$$\begin{aligned} \mathbf{v} &= \boldsymbol{\omega} \times \mathbf{r} \\ v^i &= \epsilon^{ijk} \omega^j r^k \end{aligned}$$

where $\boldsymbol{\omega}$ is the angular velocity and $\mathbf{r}$ is the distance from the origin. Since $\mathbf{v}$ and $\mathbf{r}$ are vectors, and ϵ^{ijk} is a rank-3 tensor, by the contraction theorem, $\boldsymbol{\omega}$ is a vector.

Next, since we have shown that the moment of inertia I^{ij} is a rank-2 tensor, then the angular momentum

$$L^i = I^{ij} \omega^j$$

is also a vector. To appreciate the power of these statement, contrast the following formulas which you would have learnt in elementary physics — and which are *not* generally correct:

$$\begin{aligned} L^1 &= I^{11} \omega^1 \\ L^2 &= I^{22} \omega^2 \\ L^3 &= I^{33} \omega^3 \end{aligned}$$

The RHS does *not* define components of a vector, and these equations are *not* valid in every reference frame.

4 Scalar, vector and tensor fields

A *field* is a quantity that depends on the position, i.e., a function of x^i ($i = 1, 2, 3$) in 3D.

4.1 Scalar field

We say φ is a scalar field if it has these two properties.

- *Field*: To every point P there is a $\varphi(P)$.
- *Scalar*: $\varphi(P)$ has the same value in all frames.

$$\varphi(P) = \varphi'(P)$$

Different functional forms
Suppose the point P has coordinates (x^1, x^2, x^3) in frame S, and coordinates (x'^1, x'^2, x'^3) in S'. Then more explicitly

$$\varphi(x^1, x^2, x^3) = \varphi'(x'^1, x'^2, x'^3)$$

Example 1
Suppose the electrostatic potential is (up to some multiplicative constant)

$$\varphi(P) = \frac{x}{r^3}$$

Under a 45° rotation of axes

$$x = \frac{1}{\sqrt{2}}(-x' + y') \quad , \quad y = \frac{1}{\sqrt{2}}(x' + y')$$
$$z = z' \quad , \quad r = r'$$

Then the function φ' is given by

$$\begin{aligned}\varphi'(x', y', z') &= \varphi(x, y, z) = \frac{x}{r^3} \\ &= \frac{1}{\sqrt{2}} \frac{-x' + y'}{(r')^3}\end{aligned}$$

So φ and φ' have the same value, but different functional forms.

4.2 Vector fields

Consider for example the electric field $\mathbf{E}$. It is a vector field. By this we mean:

- *Field*: To every point P there is a set of three numbers:[8]

$$E^i = (E^1(P), E^2(P), E^3(P))$$

- *Vector*: $\mathbf{E}(P)$ transforms like a displacement:

$$E'^i(P) = R^{ij} E^j(P)$$

  where $[R]$ is the transformation matrix for coordinate displacements.

Other examples are the magnetic field $\mathbf{B}$ and the vector potential $\mathbf{A}$.

4.3 Tensor fields

In much the same way, one can define tensor fields. An example is the stress tensor of the field (see earlier Chapter).

5 Differentiation

Start with a scalar field $\varphi(\mathbf{x})$ and compare its value at two neighboring points. Then

$$\begin{aligned}\Delta\varphi &= \varphi(\mathbf{x} + \Delta\mathbf{x}) - \varphi(\mathbf{x}) \\ &= \frac{\partial\varphi}{\partial x^1}\Delta x^1 + \frac{\partial\varphi}{\partial x^2}\Delta x^2 + \ldots \\ &= \frac{\partial\varphi}{\partial x^i}\Delta x^i\end{aligned}$$

Now $\Delta\varphi$ is a scalar, and Δx^i is a vector. Hence, by the contraction theorem,

$$\partial_i\varphi = \frac{\partial\varphi}{\partial x^i}$$

is a vector field.

Gradient
We define the gradient operator $\boldsymbol{\nabla}$ as the vector operator with components ∂_i:

$$\boldsymbol{\nabla} = \hat{\mathbf{e}}_i\, \partial_i$$

[8]It does not matter whether we use upper or lower indices.

Thus we have equations such as

$$\mathbf{E} = -\boldsymbol{\nabla}\Phi$$

in electrostatics, and such equations are covariant — the two sides transform in the same way.

Curl
Combining the property of ϵ^{ijk} with the property of $\boldsymbol{\nabla}$, we see that in an expression such as

$$\begin{aligned}\mathbf{B} &= \boldsymbol{\nabla} \times \mathbf{A} \\ B^i &= \epsilon^{ijk}\partial^j A^k\end{aligned}$$

the RHS is a vector if $\mathbf{A}$ is a vector.

6 Covariance under rotations

So with all this apparatus developed, let us come back and ask: How do we know that electrodynamics is covariant under rotations? In other words, if S and S' are two observers whose coordinate systems differ by a rotation, would they see the same physics? The problem can be approached at four levels.

(1) One phenomenon at a time
We can take one phenomenon or experiment at a time, and examine it in the two frames. This is possible, but the task would never be finished, since there are infinitely many phenomena to be examined.

(2) Examining the laws of physics
All phenomena are governed by a finite number of laws; in EM they are just the Lorentz force law and Maxwell's equations. Here are some of them in component form:

$$\begin{aligned}F^1 &= q\left(E^1 + v^2 B^3 - v^3 B^2\right) \\ \partial_1 B^2 - \partial_2 B^1 &= \mu_0 J^3 + \mu_0\epsilon_0\, \partial_t E^3\end{aligned} \tag{19}$$

We can check that these equations regain the same form under a rotation; for example, a rotation by an angle α about the z-axis, under which, for example,

$$\begin{aligned}x^1 &\mapsto x^1 C + x^2 S \\ E^1 &\mapsto E^1 C + E^2 S\end{aligned}$$

where $C = \cos\alpha$, $S = \sin\alpha$.

This can be checked, but it is messy. At least the task can be finished, since there are only a finite number of equations to check.

(3) Explicitly covariant formalism
All this is made easier by writing the laws in a form that is *explicitly* covariant under rotations.

For example, the equations in (19) are written as (components of)

$$\begin{aligned}\mathbf{F} &= q\,(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \\ \boldsymbol{\nabla} \times \mathbf{B} &= \mu_0 \mathbf{J} + \mu_0 \epsilon_0\, \partial_t \mathbf{E}\end{aligned}$$

Now the covariance becomes *manifest* — it is obvious to the eye because of the properties of vectors we have established.

Normally, in undergraduate EM courses we go at least this far, though sometimes leaving the mathematical properties of vectors not explicitly proved.

(4) Invariant formalism
Even better, if we can formulate all the laws of EM in terms of an action formalism, i.e., to have everything emerge from a single statement

$$\delta \mathcal{S} = 0$$

and if we can guarantee that $\mathcal{S}$ is a rotational scalar — the same thing according to S and S' — then of course the end results must be the same in the two frames. (Or equivalently, we formulate everything in terms of a scalar Lagrangian L or scalar Hamiltonian H.) This is in fact the best approach, in that it even accommodates curvilinear coordinates with hardly any trouble. But we shall defer this discussion until we have learnt Lorentz transformations.[9]

Going to relativity
Although this Chapter is relevant in itself in establishing the rotational covariance of EM, the more important purpose is to serve as an introduction to the analogous discussion of relativistic covariance — transformation between two observers in relative motion. For this problem, we have seen some examples at the very beginning of the set of lectures, i.e., only as far as level (1). It turns out that level (2) is so complicated that we would not even bother; we shall go directly to level (3), and later also level (4).

[9] In this case, the action is much better than the Lagrangian or the Hamiltonian.

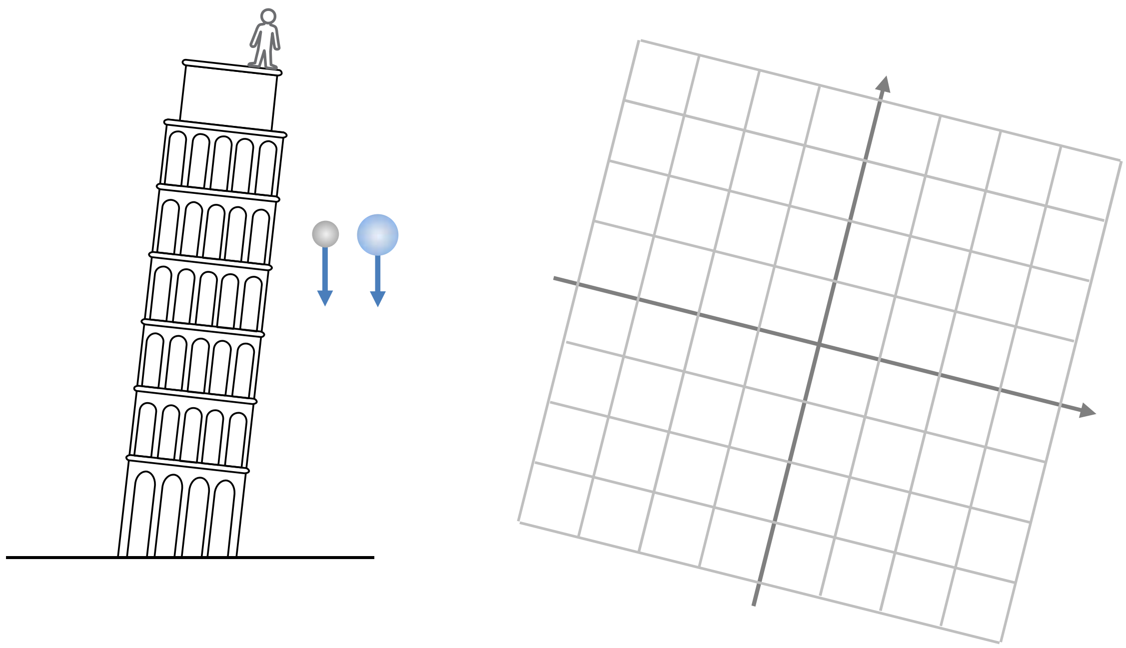

Figure 1 Physics is absolute. Coordinates are arbitrary. Therefore physics should be independent of coordinates

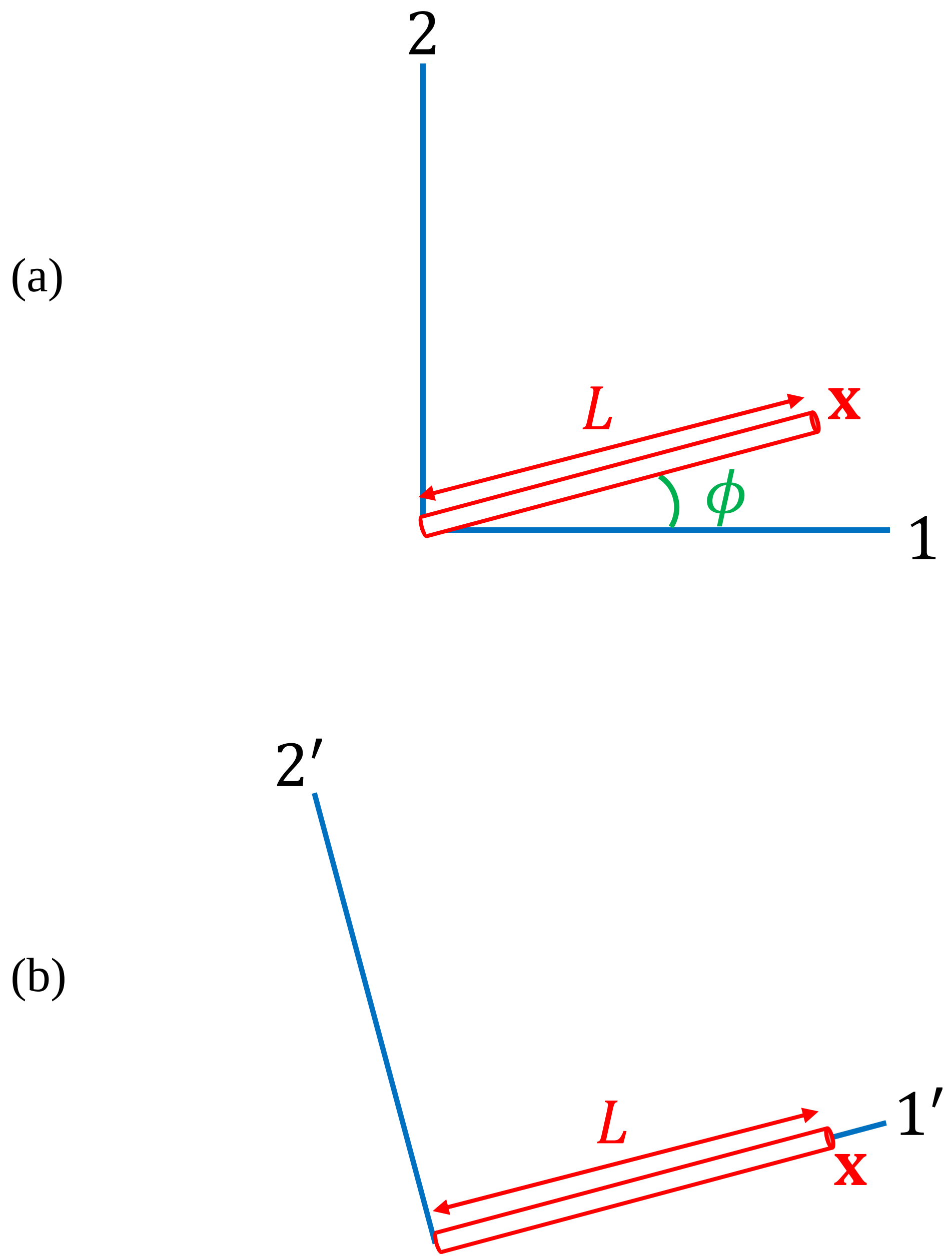

Figure 2 A rod with length L described in two coordinate systems

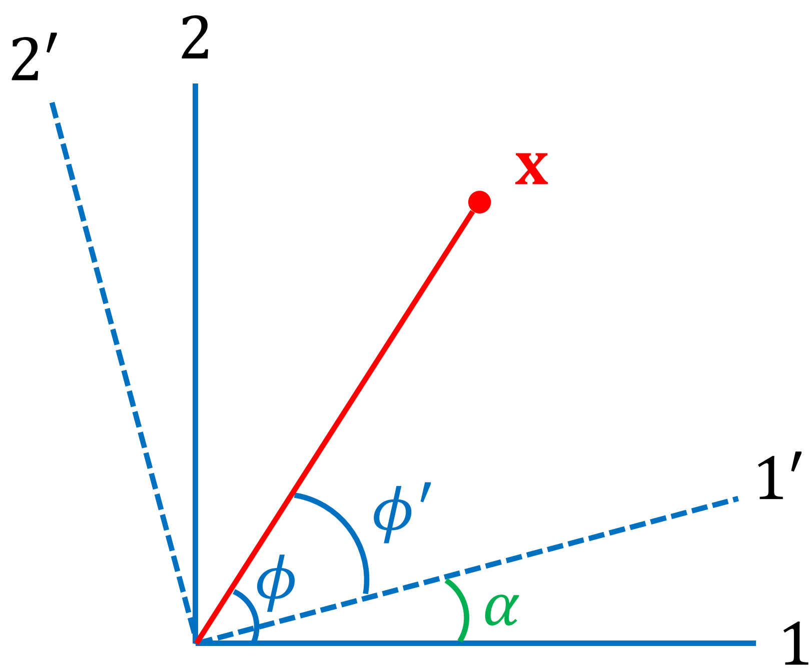

Figure 3 The components of a vector **x** described in two coordinate systems

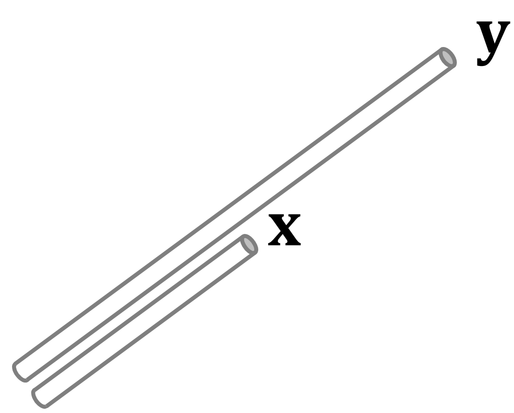

Figure 4   A simple law relating two vectors: $\mathbf{y} = 2\mathbf{x}$

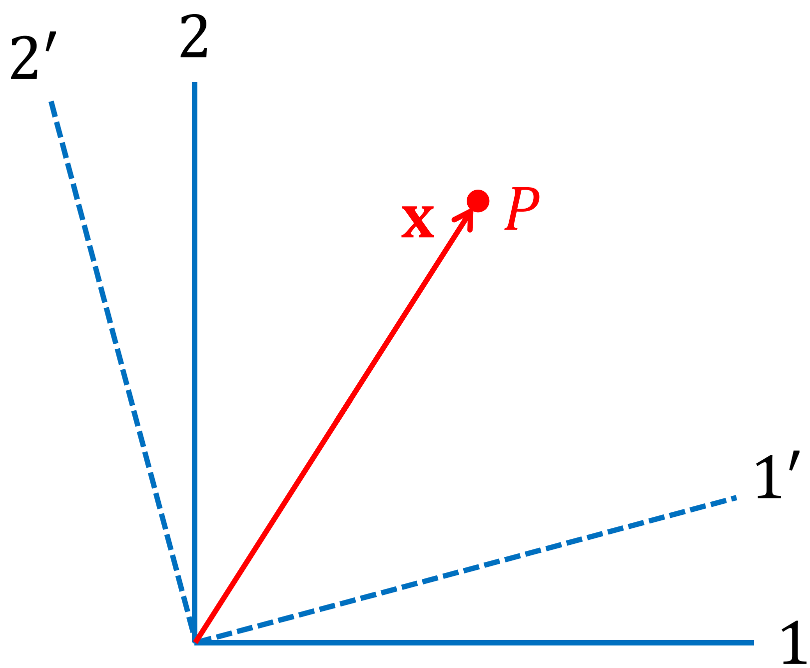

Figure 5   The same vector in two coordinate systems

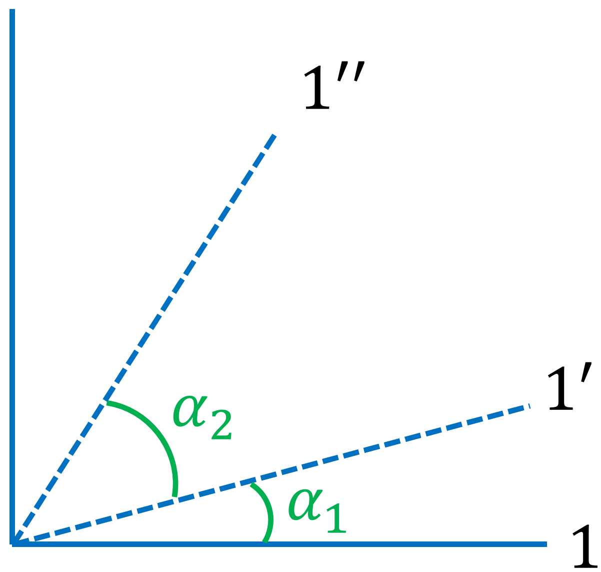

Figure 6 Combining two rotations