
Tensors

Diez B. Borja
May 25, 2022

Preface

*This notes has been written by **B. Diez** mainly from *Introducing Einstein's Relativity* by Ray D'Inverno and own knowledge.*

Contents

Preface	i
1 Tensor Algebra	2
1.1 Introduction	2
1.2 Manifolds and coordinates	2
1.3 Curves and surfaces	3
1.4 Transformation of coordinates	3
1.5 Contravariant tensors	4
1.6 Covariant and mixed tensors	5
1.7 Tensor Fields	6
1.8 Elementary Operations with Tensors	7
1.9 Index-free Interpretation of Contrvariant Vector Fields	8

Chapter 1

Tensor Algebra

1.1 Introduction

¹ To work effectively in Newtonian theory, one really needs the language of vectors. This language, first of all, is more succinct, since it summarizes a set of three equations in one. Moreover, the formalism of vectors helps to solve certain problems more readily, and, most important of all, the language reveals structure and thereby offers insight. In exactly the same way, in relativity theory, one needs the language of tensors. Again, the language helps to summarize sets of equations succinctly and to solve problems more readily, and it reveals structure in the equations. This part is devoted to learning the formalism of tensors which is a pre-condition for the rest.

The approach we adopt is to concentrate on the technique of tensors without taking into account the deeper geometrical significance behind the theory. We shall be concerned more with what you do with tensors rather than what tensors actually are. There are two distinct approaches to the teaching of tensors: the abstract or index-free (coordinate-free) approach and the conventional approach based on indices. There has been a move in recent years in some quarters to introduce tensors from the stars using the more modern abstract approach (although some have subsequently changed their mind and reverted to the conventional approach). The main advantage of this approach is that it offers deeper geometrical insight. However, it has two disadvantages. First of all, it requires much more of a mathematical background, which in turn takes time to develop. The other disadvantage is that, for all its elegance, when one wants to do real calculation with tensors, as one frequently needs to, then recourse has to be made to indices. We shall adopt the more conventional index approach, because it will prove faster and more practical. However, we advise those who wish to take their study of the subject further to look at the index-free approach at the first opportunity.

1.2 Manifolds and coordinates

*** Quizás se podría complementar algo más ***

We shall start by working with tensors defined in n dimensions since, and it is part of the power of the formalism, there is little extra effort involved. A tensor is an object defined on a geometric entity called a (differential) **manifold**. We shall now define a manifold precisely

¹Notes taken from Introducing Einstein's relativity by Ray D'Inverno

because it would involve us too much of a digression. But, in simple terms, a manifold is something which 'locally' looks like a bit of n -dimensional Euclidean space \mathbb{R}^n

We shall simply take an n -dimensional manifold M to be a set of points such that each point possesses a set of n **coordinates** x^1, x^2, \dots, x^n , where each coordinate ranges over a subset of the reals, which may, in particular, range from $-\infty$ to $+\infty$. To start off with, we can think of these coordinates as corresponding to distances or angles in Euclidean space.

1.3 Curves and surfaces

We shall frequently define these curves and surfaces parametrically

$$x^a = x^a(u), \quad a = 1, \dots, n \quad (1.1)$$

$$f(x^1, x^2, \dots, x^n) = 0 \quad (1.2)$$

Points in an m -dimensional subspace ($m < n$) must satisfy $n - m$ constraints

$$\begin{aligned} f^1(x^1, \dots, x^n) &= 0 \\ &\vdots \\ f^{n-m}(x^1, \dots, x^n) &= 0 \end{aligned} \quad (1.3)$$

1.4 Transformation of coordinates

We need to find out how quantities behave when we go from one coordinate system to another one. We therefore consider the change of coordinates $x^a \rightarrow x'^a$ given by the n equations

$$x'^a = f^a(x^1, \dots, x^n), \quad a = 1, \dots, n \quad (1.4)$$

we can write (1.4) more succinctly as $x'^a = f^a(x)$, or more simply

$$\boxed{x'^a = x^a(x)} \quad (1.5)$$

We next contemplate differentiating (1.5) with respect to each coordinate x^b

$$\left[\frac{\partial x'^a}{\partial x^b} \right]$$

the determinant J' of this matrix is called the **Jacobian** of the transformation

$$J' = \left| \frac{\partial x'^a}{\partial x^b} \right| \quad (1.6)$$

Assume that this is non-zero. Then we can solve (1.5) for the old coordinates x^a and obtain the **inverse** transformation

$$\begin{aligned} x^a &= x^a(x') \\ J &= \left| \frac{\partial x^a}{\partial x'^b} \right| \quad (\text{Jacobian of the inverse transformation}) \\ J &= \frac{1}{J'} \end{aligned}$$

In 3 dimensions, the equation of a surface is given by $z = f(x, y)$, then its total differential is defined to be

$$dz = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$$

Then, in an analogous manner, starting from (1.5) we define the total differential

$$\begin{aligned} dx'^a &= \frac{\partial x'^a}{\partial x^1} dx^1 + \cdots + \frac{\partial x'^a}{\partial x^n} dx^n \\ dx'^a &= \sum_{b=1}^n \frac{\partial x'^a}{\partial x^b} dx^b \end{aligned} \quad (1.7)$$

introducing the **Einstein summation convention**

$$\boxed{dx'^a = \frac{\partial x'^a}{\partial x^b} dx^b} \quad (1.8)$$

It defines the Kronecker delta as

$$\delta_b^a = \begin{cases} 1, & a = b \\ 0, & a \neq b \end{cases} \quad (1.9)$$

It therefore follow directly from the definition of partial differentiaton that

$$\frac{\partial x'^a}{\partial x'^b} = \frac{\partial x^a}{\partial x^b} = \delta_b^a \quad (1.10)$$

1.5 Contravariant tensors

We shall start with a prototype and then give general definition.

Consider two neighboring points in the manifold P and Q with coordinates x^a and $x^a + dx^a$ respectively. The two points define an **infinitesimal displacement** or **infinitesimal vector** \overrightarrow{PQ} . The components of this vector in the x^a -coordinate system are dx^a . The components in another coordinate system, say the x'^a -coordinate system, are dx'^a which are conected to dx^a by (1.8)

$$dx'^a = \frac{\partial x'^a}{\partial x^b} dx^b \quad (1.11)$$

The transformation matrix appearing in this equation is to be regarded as being evaluated at the point P , i.e, strictly speaking we should write

$$dx'^a = \left[\frac{\partial x'^a}{\partial x^b} \right]_P dx^b \quad (1.12)$$

A **contravariant vector** or **contravariant tensor of rank (order) 1** is a set of quantities, written X^a in the x^a -coordinates system, associated with a point P , which trasform under a change of coordinates according to

$$\boxed{X'^a = \frac{\partial x'^a}{\partial x^b} X^b} \quad (1.13)$$

where the transform matrix is evaluated at P . The infinitesimal vector dx^a is a special case of (1.13) where the components X^a are infinitesimal.

A **contravariant tensor of rank 2** is a set of n^2 quantities associated with a point P , denoted by X^{ab} in the x^a -coordinate system, which transform according to

$$X'^{ab} = \frac{\partial x'^a}{\partial x^c} \frac{\partial x'^b}{\partial x^d} X^{cd} \quad (1.14)$$

An important case is a tensor of zero rank, called a **scalar** or **scalar invariant** ϕ , which transform according to

$$\boxed{\phi' = \phi} \quad (1.15)$$

at P .

1.6 Covariant and mixed tensors

Let

$$\phi = \phi(x^a) \quad (1.16)$$

be a real-valued function on the manifold (at every point P in the manifold, $\phi(P)$ produces a real number). Also assume that ϕ is continuous and differentiable.

Remembering from (??), x^a can be thought of as a function of x'^b , (1.16) can be written equivalently as

$$\phi = \phi(x^a(x'))$$

Remembering Differentiating this with respect to x'^b , we obtain

$$\frac{\partial \phi}{\partial x'^b} = \frac{\partial \phi}{\partial x^a} \frac{\partial x^a}{\partial x'^b}$$

Then changing the order of the terms, the dummy index, and the free index (from b to a) gives

$$\frac{\partial \phi}{\partial x'^a} = \frac{\partial x^b}{\partial x'^a} \frac{\partial \phi}{\partial x^b} \quad (1.17)$$

This is the prototype equation we are looking for. Notice that it involves the inverse transformation matrix $\partial x^b / \partial x'^a$. Thus, a **covariant vector** or **covariant tensor of rank (order) 1** is a set of quantities, which transform according to

$$\boxed{X'_a = \frac{\partial x^b}{\partial x'^a} X_b} \quad (1.18)$$

Again, the transform matrix occurring is assumed to be evaluated at P .

Similarly, we define a covariant tensor of rank 2 by the transform law

$$X'_{ab} = \frac{\partial x^c}{\partial x'^a} \frac{\partial x^d}{\partial x'^b} X_{cd} \quad (1.19)$$

and so on for higher-rank tensors.

Note the convention that contravariant tensors have raised indices whereas covariant tensors have lowered indices. The way to remember this is the **co** goes **below**. The fact that the

differentials dx^a transform as a contravariant vector explains the convention that the coordinate themselves are written as x^a rather than x_a , although note that it is only the differentials and not the coordinates which have tensorial character.

We can go on to define **mixed** tensors in the obvious way. For example, a mixed tensor of rank 3- one contravariant rank and two covariant rank- satisfies

$$X'^a_{bc} = \frac{\partial x'^a}{\partial x^d} \frac{\partial x^e}{\partial x'^b} \frac{\partial x^f}{\partial x'^c} X^d_{ef} \quad (1.20)$$

If a mixed tensor has contravariant rank p and covariant rank q , then it is said to have **type** or **valence** (p, q) .

Suppose we find in one coordinate system that two tensors, X_{ab} and Y_{ab} say, are equal

$$X_{ab} = Y_{ab} \quad (1.21)$$

Let us multiply both sides by the matrices $\partial x^a / \partial x'^c$ and $\partial x^b / \partial x'^d$ and take the implied summations to get

$$\frac{\partial x^a}{\partial x'^c} \frac{\partial x^b}{\partial x'^d} X_{ab} = \frac{\partial x^a}{\partial x'^c} \frac{\partial x^b}{\partial x'^d} Y_{ab}$$

Since X_{ab} and Y_{ab} are both covariant tensors of rank 2 it follows that $X'_{ab} = Y'_{ab}$. In other words, (1.21) holds in **any** other coordinate system. In short, a tensor equation which holds in one coordinate system necessarily holds in **all** coordinate systems. Thus, although we introduce coordinate systems for convenience in tackling particular problems, if we work with tensorial equations then they hold in all coordinate systems. Put another way, tensorial equations are coordinate-independent. This is something that the index-free or coordinate-free approach makes clear from the outset.

1.7 Tensor Fields

In a vector analysis, a fixed vector is associated with a point, whereas a vector **field** defined over a region is an association of a vector to every point in that region. In exactly the same way, a tensor is a set of quantities defined at one point in the manifold. A **tensor field** defined over some region of the manifold is an association of a tensor of the same valence to every point of the region, i.e:

$$P \rightarrow T^{a\dots}_{b\dots}(P)$$

where $T^{a\dots}_{b\dots}(P)$ is the value of the tensor at P . The tensor field is called continuous or differentiable if its components in all coordinate systems are continuous or differentiable functions of the coordinates. The tensor field is called smooth if its components are differentiable to all orders, which is denoted mathematically by saying that all the components are C^∞ . Thus, for example, a contravariant vector field defined over a region is a set of n **functions** defined over that region, and the vector field is **smooth** if the functions are all C^∞ . The transformation law for contravariant vector field then becomes

$$X'^a(x') = \left[\frac{\partial x'^a}{\partial x^b} \right]_P X^b(x) \quad (1.22)$$

at each point P in the region, since the old components X^a are functions of the old x^a -coordinates and the new components X'^a are functions of the new x'^a -coordinates.

As in the case of vectors and vector fields in vector analysis, the distinction between a tensor and a tensor field is not always made completely clear. We shall for the most part be dealing with tensor field from now on, but to confo, with general usage we shall often refer to tensor fields simply as tensors. We will again shorten the transformation law such as (1.22) to the form (1.18) with everything else being implied. If we wish to emphasize that a tensor is a field, we shall write it in functional form, namely, as $T_{b,\dots}^{a,\dots}(x)$.

1.8 Elementary Operations with Tensors

Tensor calculus is concerned with **tensorial operations**, that is, operations on tensors which result in quantities which are still tensors. A simple way of establishing whether or not a quantity is a tensor is to see how it transform under a coordinate transformation. For example, we can deduce directly from the transformation law that two tensors of the same type can be added together to give a tensor of the same type, e.g.

$$X^a_{bc} = Y^a_{bc} + Z^a_{bc} \quad (1.23)$$

The same holds true for subtraction and scalar multiplication.

A covariant tensor of rank 2 is said to be **symmetric** if $X_{ab} = X_{ba}$, in which case it has only $\frac{1}{2}n(n+1)$ independent components (check this by establishing how many independent components there are of a symmetric matrix of order n). Symmetry is a tensorial property. A similar definition holds for a contravariant tensor X^{ab} . The tensor X_{ab} is said to be **anti-symmetric** or **skew symmetric** if $X_{ab} = -X_{ba}$, which has only $\frac{1}{2}n(n-1)$ independent components; this is again a tensorial property. A notation frequently used to denote the symmetric part of a tensor is

$$X_{(ab)} = \frac{1}{2}(X_{ab} + X_{ba}) \quad (1.24) \quad \text{Symmetric part of a tensor}$$

and the anti-symmetric part is

$$X_{[ab]} = \frac{1}{2}(X_{ab} - X_{ba}) \quad (1.25) \quad \text{Anti-symmetric part of a tensor}$$

In general,

$$X_{(a_1 a_2 \dots a_r)} = \frac{1}{r!} \quad (\text{sum over all permutations of the indices } a_1 \text{ to } a_r) \quad (1.26)$$

and

$$X_{[a_1 a_2 \dots a_r]} = \frac{1}{r!} \quad \text{alternating sum over all permutations of the indices } a_1 \text{ to } a_r \quad (1.27)$$

For example, we shall need to make use of the result

$$X_{[abc]} = \frac{1}{6}(X_{abc} - X_{acb} + X_{cab} - X_{cba} + X_{bca} - X_{bac}) \quad (1.28)$$

(A way to remember the above expression is to note that the positive terms are obtained by cycling the indices to the right and the corresponding negative terms by flipping the last two indices.) A **totally symmetric tensor** is defined to be one equal to its symmetric part, and a **totally anti-symmetric tensor** is one equal to its anti-symmetric part.

We can multiply two tensors of type (p_1, q_1) and (p_2, q_2) together and obtain a tensor of type $(p_1 + p_2, q_1 + q_2)$, e.g.

$$X^a_{bcd} = Y^a_b Z_{cd} \quad (1.29)$$

In particular, a tensor of type (p, q) when multiplied by a scalar field ψ is again a tensor of type (p, q) . Given a tensor of mixed type (p, q) , we can form a tensor of type $(p - 1, q - 1)$ by the process of **contraction**, which simply involves setting a raised and lowering index equal. For example,

$$X^a_{bcs} \xrightarrow{\text{contraction on } a \text{ and } b} X^a_{acd} = Y_{cd}$$

1.9 Index-free Interpretation of Contravariant Vector Fields

As we pointed out in 1.5, we must distinguish between the actual geometric object itself and its components on a particular coordinate system. The important point about tensors is that we want to make statements which are independent of any particular coordinate system being used. This is abundantly clear in the index-free approach to tensors. We shall get a feel for this approach in this section by considering the special case of a contravariant vector field, although similar index-free interpretations can be given for any tensor field. The key idea is to interpret the vector field as an **operator** which maps real-valued functions into real-valued functions. Thus, if X represents a contravariant vector field, then X operates on any real-valued function f to produce another function g , i.e. $Xf = g$. We shall show how actually to compute Xf by introducing a coordinate system. However, as we shall see, we could equally well introduce any other coordinate system, and the computation would lead to the same result.

In the x^a -coordinate system, we introduce the notation

$$\partial_a \equiv \frac{\partial}{\partial x^a}$$

and the X is defined as the operator

$$\boxed{X = X^a \partial_a} \quad (1.30)$$

$$Xf = (X^a \partial_a)f = X^a (\partial_a f) \quad (1.31)$$

for any real-valued function f . Let us compute X in some other x'^a -coordinate system. We need to use the result (1.10) expressed in the following form: we may take x^a to be a function of x'^b by (??) and x'^b to be a function of x^c by (1.5), and so, using the function of a function rule, we find

$$\delta^a_b = \frac{\partial x^a}{\partial x'^b} = \frac{\partial}{\partial x'^b} x^a(x'^c(x^d)) = \frac{\partial x^a}{\partial x'^c} \frac{\partial x'^c}{\partial x'^b} \quad (1.32)$$

Then, using the tranformation law (1.13) and (1.17) together with the above trick, we get

$$\begin{aligned}
X'^a \partial'_a &= X'^a \frac{\partial}{\partial x'^a} \\
&= \frac{\partial x'^a}{\partial x^b} X^b \frac{\partial x^c}{\partial x'^a} \frac{\partial}{\partial x^c} \\
&= \frac{\partial x^c}{\partial x'^a} \frac{\partial x'^a}{\partial x^b} X^b \frac{\partial}{\partial x^c} \\
&= \delta_b^a X^b \frac{\partial}{\partial x^c} \\
&= X^b \frac{\partial}{\partial x^b} \\
&= X^a \frac{\partial}{\partial x^a} \\
&= X^a \partial_a
\end{aligned}$$

Thus the result of operating on f by X will be the same **irrespective** of the coodinate system employed in (1.30).

In any coordinate system, we may think of the quantities $[\partial/\partial x_a]_P$ as forming a basis for all the vectors at P , since any vector at P i, by (1.30), given by

$$X_P = [X^a]_P \left[\frac{\partial}{\partial x^a} \right]_P$$

that is, a linear combination of the $[\partial/\partial x^a]_P$. the vector space of all the contravariant vectors at P is known as the **tangent space** of P and is written $T_P(M)$. (Fig. 1.1). In general, the tangent space at any point in a manifold is different from the underlying manifold. For this reason, we need to be careful in represnting a finite contravariant vector by an arrow in our figures since, strictly speaking, the arrow lies in the tangent space not the manifold. Two exceptions to this are Euclidean space and Minkowski space-time, where the tangent space at each point coincides with the manifold.

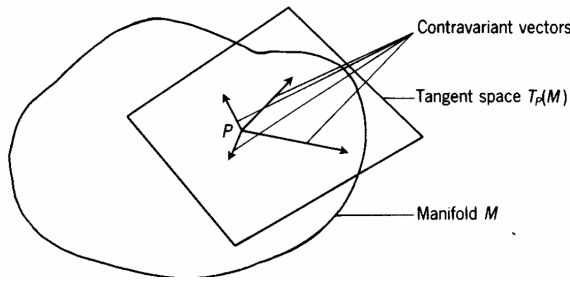


Figure 1.1: The tangent space at P .

Given two vector fields X and Y we can define a new vector field called the **commutator** or **Lie bracker** of X and Y by

$$\boxed{[X, Y] = XY - YX} \quad (1.33) \quad \text{Lie bracket}$$

Letting $[X, Y] = Z$ and operating with it on some arbitrary function f

$$\begin{aligned}
Zf &= [X, Y]f \\
&= (XY - YX)f \\
&= X(Yf) - Y(Xf) \\
&= X(Y^a \partial_a f) - Y(X^a \partial_a f) \\
&= X^b \partial_b (Y^a \partial_a f) - Y^b \partial_b (X^a \partial_a f) \\
&= (X^b \partial_b Y^a - Y^b \partial_b X^a) \partial_a f - X^a Y^b (\partial_b \partial_a f - \partial_a \partial_b f)
\end{aligned}$$

The last term vanishes since we assume commutativity of second mixed partial derivatives, i.e.

$$\partial_a \partial_b = \frac{\partial^2}{\partial x^a \partial x^b} = \frac{\partial^2}{\partial x^b \partial x^a} = \partial_b \partial_a$$

Since f is arbitrary, we obtain the result

$$[X, Y]^a = Z^a Z^b \partial_b Y^a - Y^b \partial_b X^a \quad (1.34)$$

from which it clearly follows that the commutator of two vector fields is itself a vector field. It also follows directly from the definition (1.33), that

$$[X, X] \equiv 0 \quad (1.35)$$

$$[X, Y] \equiv [Y, X] \quad (1.36)$$

$$[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] \equiv 0 \quad (1.37)$$

(1.36) shows that the Lie bracket is ant-commutative. The result (1.37) is known as **Jacobi's identity**. Notice it states that the left-hand side is not just equal to zero but is **identically** zero. What does it mean? The equation $x^2 - 4 = 0$ is only satisfied by particular values of x , namely, $+2$ and -2 . The identity $x^2 - x^2 \equiv 0$ is satisfied for all values of x . But, you may argue, the x^2 terms cancel out, and this is precisely the point. An expression is identically zero if, when all the terms are written out fully, they all cancel in pairs.