PY501 - Mathematical Physics

Hongwan Liu

Table of Contents

1 Linear Algebra **1**
1.1 A Review of Linear Algebra . . . 1
1.1.1 Vector spaces and inner products . . . 1
1.1.2 Bases and components . . . 2
1.2 Change of Bases, Covariant and Contravariant Transformations . . . 3
1.3 The Dual Space . . . 4
1.4 The Metric . . . 5
1.4.1 Raising and lowering indices . . . 5
1.4.2 Example: Some Common Metrics . . . 6
1.5 Tensors . . . 8
1.5.1 Tensor algebra . . . 9
1.5.2 Example: Rotations and Lorentz transformations . . . 10
1.5.3 Example: Linear transformations . . . 11
1.5.4 Invariants of 2D tensors . . . 12
1.5.5 Symmetric and antisymmetric tensors . . . 12
1.5.6 Kronecker and Levi-Civita tensors . . . 13
1.5.7 Isotropic Cartesian tensors . . . 14
1.5.8 Irreducible 3D Tensors Under Rotations . . . 15

1 Linear Algebra

References: Stone & Goldbart (SG) Chapter 10, Appendix A; Carroll Chapter 2

At this point in your physics career, you've gained a lot of intuition about $\mathbb{R}^n$, objects like vectors which live in the space, and their transformations. You probably also have at least some intuition about Minkowski space. We're now going to review these concepts more formally. This will better equip us to understand more general spaces in physics, such as Minkowski space and curved spacetime, as well as more general structures on such spaces, such as tensors.

1.1 A Review of Linear Algebra

The study of spaces like $\mathbb{R}^n$ falls under the subject of **linear algebra**. While a course in mathematical physics might feel a little incomplete without covering this topic in detail, linear algebra is generally well-covered in undergraduate curricula. We'll content ourselves with a lightning review of some key facts.

1.1.1 Vector spaces and inner products

So what are the structures of $\mathbb{R}^n$ that we take for granted? First, the fundamental quantities that we deal with in real space are **vectors**. A collection of objects that live in real space is simply a set, but the interesting thing about real space is that there are relations between the objects in the space. In fact, real space is an example of a **vector space**, a structure which is defined as follows:

A **vector space** V over a field $\mathbb{F}$ is a set equipped with two operations: a binary operation called **vector addition** which assigns to each pair of elements $\vec{x}, \vec{y} \in V$ a third element denoted $\vec{x} + \vec{y}$, and **scalar multiplication** which assigns to an element $\vec{x} \in V$ and $\lambda \in \mathbb{F}$ a new element $\lambda\vec{x} \in V$. There is also a distinguished element $\vec{0} \in V$ such that the follow axioms are obeyed:

1. Vector addition is commutative: $\vec{x} + \vec{y} = \vec{y} + \vec{x}$;
2. Vector addition is associative: $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$;
3. Additive identity: $\vec{0} + \vec{x} = \vec{x}$;
4. Existence of an additive inverse: for any $\vec{x} \in V$, there is an element $-\vec{x} \in V$ such that $\vec{x} + (-\vec{x}) = \vec{0}$;
5. Scalar distributive law: $\lambda(\vec{x}+\vec{y}) = \lambda\vec{x}+\lambda\vec{y}$, as well as $(\lambda+\mu)\vec{x} = \lambda\vec{x} + \mu\vec{x}$;
6. Scalar multiplication is associative: $(\lambda\mu)\vec{x} = \lambda(\mu\vec{x})$, and
7. Multiplicative identity: $1\vec{x} = \vec{x}$.

A lot of that just seems very natural, and so it might seem like a lot of useless abstraction. But the point is to be clear about what a vector in the abstract actually is, so that when we're in much less familiar settings, these formal structures are going to help us cut through the confusion. Furthermore, we can study properties of all vector spaces that would apply equally well to $\mathbb{R}^n$ as it does to any other vector space.

Here, you can see that $\mathbb{R}^n$ is a vector space over the field $\mathbb{R}$. However, you're also familiar with vector spaces over the complex numbers $\mathbb{C}$: one example is the *Hilbert space*, which underpins *quantum mechanics*. The states of a system are described as vectors $|\psi\rangle$ in a Hilbert space. The results of linear algebra apply equally well to both $\mathbb{R}^n$ and Hilbert spaces.

In addition to being able to add vectors, or multiply vectors by scalars, another important thing you can do in $\mathbb{R}^n$ is talk about distances: you can take the dot product or **inner product** of a vector with itself to talk about length, or take the inner product of two different vectors and talk about angles. Formally, vector spaces with this additional structure are called **inner product spaces**. Inner product spaces are defined as follows:[1]

> An **inner product space** is a vector space V over a field $\mathbb{F}$, together with an **inner product**, which is a map
>
> $$\langle\cdot,\cdot\rangle : V \times V \to \mathbb{F}\,, \tag{1.1}$$
>
> that satisfies the following properties for all $\vec{x},\vec{y},\vec{z} \in V$ and $\lambda,\mu \in \mathbb{F}$:
>
> 1. Conjugate symmetry: $\langle\vec{x},\vec{y}\rangle = \langle\vec{y},\vec{x}\rangle^*$, where $*$ denotes complex conjugation. If $\mathbb{F}$ is real, then this just means that the inner product should be symmetric;
> 2. Linearity in the second argument, i.e. $\langle\vec{x},\lambda\vec{y}+\mu\vec{z}\rangle = \lambda\langle\vec{x},\vec{y}\rangle + \mu\langle\vec{x},\vec{z}\rangle$. Note that this together with conjugate symmetry implies that $\langle\lambda\vec{x}+\mu\vec{y},z\rangle = \lambda^*\langle\vec{x},\vec{z}\rangle + \mu^*\langle\vec{y},\vec{z}\rangle$. The inner product is only linear in both arguments when $\mathbb{F}=\mathbb{R}$, and
> 3. Nondegenerate, i.e. if $\langle\vec{x},\vec{y}\rangle = 0$ for all $\vec{y}$, then $\vec{x}=0$.

[1] There are a few differences here compared to the usual definition in mathematics. First, in mathematics, it is common to have linearity apply to the first argument. This is of course is entirely equivalent. We use this definition to conform with our usual intuition in braket notation. Second, the inner product space is usually defined as having a *positive definite* inner product; but this unfortunately excludes Minkowski space, which is more properly classified as a pseudo-inner product. We don't really care about these differences in physics though.

The inner product on $\mathbb{R}^n$ is the dot product; in Hilbert space, it is denoted $\langle\psi'|\psi\rangle$; in Minkowski space, we have the metric tensor. We'll go into a lot more detail on this in just a bit.

The last thing that we'll talk about are **linear transformations** (also known as linear operators or linear maps), which are functions that take us between vector spaces. Let V and W be vector spaces with dimensions n and m respectively; $A : V \to W$ is a linear transformation if

$$A(\lambda\vec{x}+\mu\vec{y}) = \lambda A(\vec{x}) + \mu A(\vec{y})\,. \tag{1.2}$$

1.1.2 Bases and components

At this point, the vectors on a vector space are still abstract objects. In order to make contact with our usual representation of vectors as a column of numbers, we need to define a **basis** for the vector space. This is something that you've probably seen in linear algebra, but we'll state some facts and definitions that all are somewhat intuitive:

1. A set of vectors $\{\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n\}$ is **linearly dependent** if there exist $\lambda^1,\cdots,\lambda^n \in \mathbb{F}$, written as λ^μ for $\mu = 1,\cdots,n$, not all zero, such that

$$\lambda^1\vec{e}_1 + \lambda^2\vec{e}_2 + \cdots\lambda^n\vec{e}_n = \vec{0}\,. \tag{1.3}$$

2. If it is not linearly dependent, a set of vectors $\{\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n\}$ is **linearly independent**. For a linearly independent set, the relation

$$\lambda^1\vec{e}_1 + \lambda^2\vec{e}_2 + \cdots\lambda^n\vec{e}_n = \vec{0} \tag{1.4}$$

holds only if $\lambda^1 = \cdots = \lambda^n = 0$.

3. A set of vectors $\{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\}$ is said to **span** V if for any $\vec{x} \in V$, there are numbers x^μ such that $\vec{x}$ can be written (not necessarily uniquely) as

$$\vec{x} = x^1\vec{e}_1 + x^2\vec{e}_2 + \cdots + x^n\vec{e}_n \,. \tag{1.5}$$

   A vector space is **finite dimensional** if a finite spanning set exists.

4. A set of vectors $\{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\}$ is a **basis** if it is a *maximally linearly independent set*, i.e. introducing any additional vector makes the set linearly dependent. Equivalently, a basis is a *minimal spanning set*, i.e. deleting any of the $\vec{e}_i$ destroys the spanning property.

5. If $\{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\}$ is a basis, then any $\vec{x} \in V$ can be written

$$\vec{x} = x^1\vec{e}_1 + x^2\vec{e}_2 + \cdots + x^n\vec{e}_n \,, \tag{1.6}$$

   where the x^μ, known as the **components** of the vector with respect to this basis, are unique in that two vectors coincide if and only if they have the same components.

6. If the sets $\{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\}$ and $\{\vec{f}_1, \vec{f}_2, \cdots, \vec{f}_m\}$ are both bases for the space V, then $m = n$. This invariant integer is the **dimension**, $\dim(V)$, of the space.

At this point, you may be looking at the notation above and wondering about the placement of indices: when are indices placed above, and when are they placed below? This will made clear in the next part of our discussion.

1.2 Change of Bases, Covariant and Contravariant Transformations

Having defined the concepts of a basis, and the components of a vector with respect to a basis, we now want to understand how these components change as we choose different bases, since bases are not unique.

Suppose a vector space V has two different bases given by $\{\vec{e}_1, \cdots, \vec{e}_n\}$ and $\{\vec{e}'_1, \cdots, \vec{e}'_n\}$. Since both sets span V, every vector in $\{\vec{e}_1, \cdots, \vec{e}_n\}$ can be written as a sum of $\{\vec{e}'_1, \cdots, \vec{e}'_n\}$, and we can define a set of n^2 numbers $a^\mu{}_\nu$ which maps

$$\vec{e}_\nu = \sum_{\mu=1}^{n} a^\mu{}_\nu \vec{e}'_\mu \equiv a^\mu{}_\nu \vec{e}'_\mu = \vec{e}'_\mu a^\mu{}_\nu \,. \tag{1.7}$$

At this point, we've introduced the famous **Einstein notation**, which just says that every repeated index should be regarded as being summed over all possible values. Again, you may be worried about the placement of the indices, but all be clear as we go along. The final expression is helpful in helping you visualize the object $\vec{e}'_\mu$ as a row vector, multiplied by the **matrix** $a^\mu{}_\nu$, where μ indexes its rows, and ν indexes its columns.

$a^\mu{}_\nu$ is clearly **invertible**: every vector has a unique representation in each basis, and the map takes the coordinates of any vector in one basis to another, and so it is certainly a bijective map. We can therefore define $(a^{-1})^\mu{}_\nu$ as the inverse map,

$$\vec{e}'_\nu = (a^{-1})^\mu{}_\nu \vec{e}_\mu \,, \tag{1.8}$$

with

$$(a^{-1})^\mu{}_\nu a^\nu{}_\sigma = \delta^\mu_\sigma \,, \tag{1.9}$$

where δ^μ_σ is the Kronecker delta or the identity matrix.

So far, we have dealt with the transformation of the basis. But how does a general vector transform? Given the transformation between bases above, we see that for any arbitrary vector $\vec{x}$, which can be written as $x^\nu \vec{e}_\nu$ in one basis and $x'^\mu \vec{e}'_\mu$ in the other, are related by

$$\vec{x} = x'^\mu \vec{e}'_\mu = x^\nu \vec{e}_\nu = x^\nu (a^\mu{}_\nu \vec{e}'_\mu) = (a^\mu{}_\nu x^\nu)\vec{e}'_\mu \,, \tag{1.10}$$

or in other words,[2]

$$x'^\mu = a^\mu{}_\nu x^\nu \,. \tag{1.11}$$

One thing you should notice immediately is that the basis and the coordinates transform in the opposite way:

$$x'^\mu = a^\mu{}_\nu x^\nu \,, \qquad \vec{e}'_\mu = (a^{-1})^\sigma{}_\mu \vec{e}_\sigma \,, \tag{1.12}$$

because of course the vector itself, $x'^\mu \vec{e}'_\mu = x^\mu \vec{e}_\mu$, doesn't transform under a coordinate change at all! Any quantity that transforms under a change of basis like the basis itself is said to transform **covariantly**, while any quantity that transforms like the coordinates, i.e. in the opposite manner as the basis, is said to transform **contravariantly**. *We will always use indices on the top to indicate a quantity that transforms contravariantly, and indices on the bottom to indicate a quantity that transform covariantly.*

The best intuition for this comes from imagining a change of basis via rotation in $\mathbb{R}^2$, as shown in Fig. 1. Either we can imagine the basis vectors actually rotating counterclockwise and defining a new set of axes, as we would do in taking $x'^\nu \vec{e}'_\nu = x'^\nu (a^\mu{}_\nu \vec{e}_\mu)$, or equivalently, we can think of the components of the vector themselves rotating clockwise, with the axes just being relabeled, which corresponds to $x'^\nu \vec{e}'_\nu = (a^\mu{}_\nu x'^\nu)\vec{e}_\mu$.

[2] Very often, you will see the notation $a^\mu{}_\nu \equiv \partial x'^\mu / \partial x^\nu$, which makes total sense if you look at Eq. (1.11). In fact, the advantage of writing it this way tells you how to obtain the matrix $a^\mu{}_\nu$.

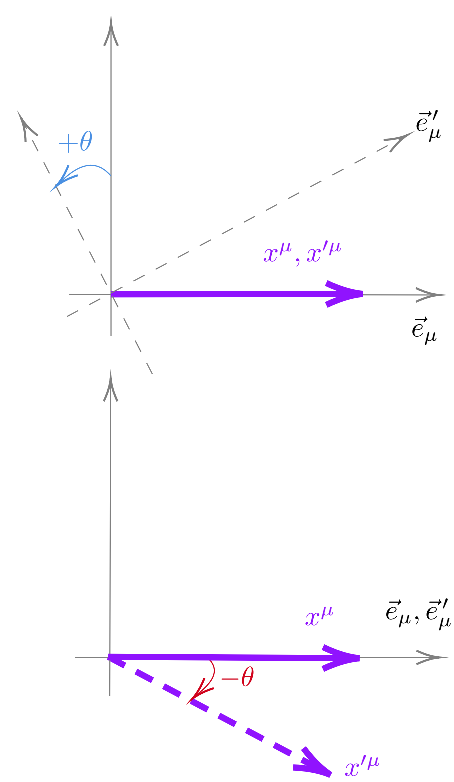

Figure 1: Two equivalent ways to think of a change of basis. (Top) the basis vectors themselves are transformed, or (bottom) the coordinates are transformed in the opposite direction. These pictures are equivalent.

1.3 The Dual Space

For every vector space V, we can define a **dual space** V^*, which is a set of linear transformations $f : V \to \mathbb{F}$, each of which takes in a vector and returns a number. The functions f are called **covectors** or **one-forms**, and you can convince yourself that V^* is also a vector space. Since these functions are linear, we have

$$f(\vec{x}) = f(x^\mu \vec{e}_\mu) = x^\mu f(\vec{e}_\mu) \equiv x^\mu f_\mu \,, \tag{1.13}$$

where in the last equality I have defined the set of numbers $f_\mu \equiv f(\vec{e}_\mu)$, which I can construct given the basis $\{\vec{e}_\mu\}$ in V. Under a change of basis in V,

$$f_\mu = f(\vec{e}_\mu) = f(a^\nu{}_\mu \vec{e}'_\nu) = a^\nu{}_\mu f(\vec{e}'_\nu) \equiv a^\nu{}_\mu f'_\nu \,, \tag{1.14}$$

where $f'_\nu \equiv f(\vec{e}'_\nu)$ are again a set of numbers that we can construct given the basis $\{\vec{e}'_\nu\}$ in V. Notice that under a change of basis in V, f_μ transforms *covariantly*, i.e. in the same manner as the change of basis in V.

Given a basis $\vec{e}_\mu$ of V, we can define a **dual basis** for V^*, which is the set of covectors $\vec{e}^{*\mu} \in V^*$ such that

$$\vec{e}^{*\mu}(\vec{e}_\nu) = \delta^\mu_\nu \,. \tag{1.15}$$

This is clearly a basis, since for any $f \in V^*$,

$$f(\vec{x}) = x^\mu f_\mu = x^\mu f_\nu \delta^\nu_\mu = x^\mu f_\nu \vec{e}^{*\nu}(\vec{e}_\mu) = f_\nu \vec{e}^{*\nu}(x^\mu \vec{e}_\mu) \,, \tag{1.16}$$

or in other words,

$$f = f_\mu \vec{e}^{*\mu} \,. \tag{1.17}$$

We should therefore view $f_\mu = f(\vec{e}_\mu)$ as the *components* of f under the induced dual basis $\vec{e}^{*\mu}$.

You should already have a sense that V and V^* are very closely related; in fact, the map $\vec{e}_\mu \mapsto \vec{e}^{*\mu}$ is an **isomorphism**, i.e. a map of every element in V to another in V^* that preserves their respective relation to each other under addition and scalar multiplication.

(**End of Lecture**: Wednesday 3 Sep 2025)

1.4 The Metric

So far, everything we have discussed has been about vector spaces. We are now going to turn our attention to inner product spaces over $\mathbb{R}$, where the additional inner product structure is defined, giving us a way of talking about distances and angles.

As a reminder, the inner product is a map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ that takes two vectors in V, and spits out a number. Having chosen a basis $\{\vec{e}_\mu\}$ for our vector space V, we can now define a quantity known as the **metric** or **metric tensor** $g_{\mu\nu}$,

$$g_{\mu\nu} \equiv \langle \vec{e}_\mu, \vec{e}_\nu \rangle . \tag{1.18}$$

For $\mathbb{R}^n$, for example, with the inner product given by the dot product, we have simply $g_{\mu\nu} = \delta_{\mu\nu}$, while for Minkowski space, the metric is given by[3]

$$\eta_{\mu\nu} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \tag{1.19}$$

[3] I will generally stick with the mostly minus convention for the Minkowski metric. My apologies to mostly plus aficionados, but I'm just slightly more comfortable with the mostly minus convention at this point.

Knowing the metric fully defines the inner product, since

$$\langle \vec{x}, \vec{y} \rangle = \langle x^\mu \vec{e}_\mu, y^\nu \vec{e}_\nu \rangle = g_{\mu\nu} x^\mu y^\nu . \tag{1.20}$$

Moreover, the structure of the inner product also guarantees that $g_{\mu\nu} = g_{\nu\mu}$, i.e. $g_{\mu\nu}$ is **symmetric**. Another thing we can note is that as a matrix, $g_{\mu\nu} x^\nu = 0$ only if $x^\nu = 0$ by the definition of the inner product; this means that $g_{\mu\nu}$ is invertible. We can therefore define the **inverse of the metric**, which we denote $g^{\mu\nu}$, with

$$g_{\mu\nu} g^{\nu\sigma} = g^{\mu\nu} g_{\nu\sigma} = \delta^\mu_\sigma . \tag{1.21}$$

For now, this is just a relationship between matrices; we'll come back and revisit the metric when we have discussed tensors later on.

1.4.1 Raising and lowering indices

Let's look at the expression in Eq. (1.20) more closely. We can reinterpret $\langle \vec{x}, \vec{y} \rangle = g_{\mu\nu} x^\mu y^\nu$ as $(g_{\mu\nu} x^\mu) y^\nu$. With this rewriting, can think of $(g_{\mu\nu} x^\mu)$ as being the components of the object $\langle \vec{x}, \cdot \rangle$, which takes in a vector $\vec{y}$ and returns $\langle \vec{x}, \vec{y} \rangle$. In fact, $\langle \vec{x}, \cdot \rangle$ is an object in V^*, mapping vectors in V to real numbers, and can be written as $\langle \vec{x}, \cdot \rangle = g_{\mu\nu} x^\mu \vec{e}^{*\nu}$, so that

$$g_{\mu\nu} x^\mu \vec{e}^{*\nu} (y^\sigma \vec{e}_\sigma) = g_{\mu\nu} x^\mu y^\sigma \delta^\nu_\sigma = g_{\mu\nu} x^\mu y^\nu = \langle \vec{x}, \vec{y} \rangle . \tag{1.22}$$

Clearly then, $g_{\mu\nu} x^\mu$ are indeed the components of $\langle \vec{x}, \cdot \rangle$ in the basis of V^* induced by our chosen basis of V, and therefore transforms *covariantly*.

If all that was a bit dense, the upshot is that, starting from a contravariant quantity x^μ, we can **lower its index** by defining

$$x_\nu \equiv g_{\mu\nu} x^\mu \,, \tag{1.23}$$

which 1) is a quantity that transforms covariantly (and so has a lower index), and 2) can be **contracted** with a contravariant quantity to form a real number, or a **scalar** or an **invariant**. Intuitively, it is one half of the inner product: you need to put together a covariant and contravariant piece to obtain a scalar.

Multiplying Eq. (1.23) by $g^{\sigma\nu}$ on both sides, we also find

$$g^{\sigma\nu} x_\nu \equiv g^{\sigma\nu} g_{\mu\nu} x^\mu = g^{\sigma\nu} g_{\nu\mu} x^\mu = \delta^\sigma_\mu x^\mu = x^\sigma \,, \tag{1.24}$$

which shows that I can also **raise an index** by multiplying by the inverse tensor. Ultimately, all I'm doing is switching between the components of the two objects

$$\langle \vec{x}, \cdot \rangle \leftrightarrow \vec{x} \,, \tag{1.25}$$

which are in 1-to-1 correspondence with each other between the isomorphic vector spaces V and V^*.

Finally, notice that every time we perform a contraction, we sum over one upper and one lower index. This is because every contraction represents the pairing of a function in V^*, with a vector in V, and results in a scalar. Another way of understanding this is that you want to pair up a contravariant with a covariant quantity, so that you end up with a quantity that doesn't transform, i.e. a scalar. I have never encountered a situation where you want to sum over the components of two objects which transform in the same way.

1.4.2 Example: Some Common Metrics

Let's pause for a moment and take a look at some important examples.

To digest all of this information, let's revisit $\mathbb{R}^2$ with all of this technology. $\mathbb{R}^2$ is a 2D vector space, with vectors $x^i \vec{e}_i$ that look like, for example, $3\vec{e}_x + 2\vec{e}_y$, where 3 and 2 are the components of the vector, and $\{\vec{e}_x, \vec{e}_y\}$ is a chosen basis for the space. $\mathbb{R}^2$ also comes with an inner product, which is the usual dot product. In $\mathbb{R}^2$, we can choose a basis that is **orthonormal**, i.e. with a metric given by

$$g_{ij} = \langle \vec{e}_i, \vec{e}_j \rangle = \vec{e}_i \cdot \vec{e}_j = \delta_{ij} \,. \tag{1.26}$$

I can use this metric to raise and lower indices of covariant or contravariant quantities, so for example

$$g_{ij} x^j = \delta_{ij} x^j = x_i \,, \tag{1.27}$$

but if you explicitly plug in the indices, you can see that $x^0 = x_0$ and $x^1 = x_1$, which shows that *the position of indices doesn't matter in* $\mathbb{R}^n$. Inner products between two vectors can be written in component form as

$$g_{ij} x^i y^j = x_j y^j \,, \tag{1.28}$$

i.e. the sum of the product of individual components, as in the usual dot product. You can think of x_j as the components of $\langle \vec{x}, \cdot \rangle$. In $\mathbb{R}^2$, you can also think of

x_j as a row matrix, which maps column vectors (which contain components of a vector) to numbers by matrix multiplication.

Now let's consider a basis change, given by $\vec{e}'_i = a^j{}_i \vec{e}_j$, where

$$a^j{}_i = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} , \tag{1.29}$$

where j indexes the row and i indexes the column. This is a clockwise rotation of the basis vectors by some constant angle θ. Explicitly,

$$\begin{aligned} \vec{e}'_1 &= \cos\theta\,\vec{e}_1 - \sin\theta\,\vec{e}_2 \,, \\ \vec{e}'_2 &= \sin\theta\,\vec{e}_1 + \cos\theta\,\vec{e}_2 \,. \end{aligned} \tag{1.30}$$

At the same time, for any vector $\vec{x} = x^i \vec{e}_i$, the coordinates x^i transforms in the opposite sense, i.e. $x'^i = (a^{-1})^i{}_j x^j$, where

$$(a^{-1})^i{}_j = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} , \tag{1.31}$$

so that

$$\begin{aligned} x'^1 &= \cos\theta\, x^1 - \sin\theta\, x^2 \\ x'^2 &= \sin\theta\, x^1 + \cos\theta\, x^2 \,, \end{aligned} \tag{1.32}$$

which is instead a *counterclockwise* rotation of the components.[4] You can check that the vector itself, $x^i \vec{e}_i$, remains unchanged. This is the same intuition we had from Fig. 1.

We now graduate to something hopefully still familiar, but a little more nontrivial: 4D Minkowski space, where the 0-dimension is time, and dimensions 1,2,3 are spatial dimensions. The vectors that live in this space are called **4-vectors**, and they are of the form $x^\mu \vec{e}_\mu$, where $\{\vec{e}_0, \vec{e}_1, \vec{e}_2, \vec{e}_3\}$ forms a basis. Once again, we have an inner product and an associated metric; we can choose a basis such that the metric is the **Minkowski metric**

$$\eta_{\mu\nu} = \langle \vec{e}_i, \vec{e}_j \rangle \,, \tag{1.33}$$

where

$$\eta_{\mu\nu} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \tag{1.34}$$

You can verify for yourself that $\eta^{\mu\nu}$, the inverse of the matrix, has the same entries as $\eta_{\mu\nu}$.

Once again, I can lower indices by hitting a contravariant quantity with the metric, e.g. $x_\mu = \eta_{\mu\nu} x^\nu$, but this time, you can see that $x_0 = x^0$, and $x_i = -x^i$. Therefore, *the position of the indices does matter in Minkowski space*, and we need to be a little more careful.[5] You can still think of x_μ as the components of $\langle \vec{x}, \cdot \rangle$, but x_μ is now no longer just a simple transposition (i.e. a row matrix) relative to x^μ (which we can view as a column matrix); you also need to change the sign of the spatial components. Inner products can be written, as before, as

$$g_{\mu\nu} x^\mu y^\nu = x_\nu y^\nu \,, \tag{1.35}$$

but note that because of the negative signs in the metric, *you are no longer guaranteed that the inner product is positive.*

[4] What may be confusing is that Eq. (1.30) and Eq. (1.32) look identical! But how we interpret what's happening is different. In the first, each line tells us how each basis vector is separately rotated, so you're looking out for $\vec{e}_1$ transforming into something else. But in the second, it is the transformation of the arrow denoted by (x^1, x^2) going into (x'^1, x'^2) that we are interested in.

[5] In the wild though, some people freely mix upper and lower indices even in Minkowski space, which is possible with a little extra care. I usually prefer not to do this.

We can again consider basis changes such as the **Lorentz boost**, given by $\vec{e}'_\nu = \Lambda^\mu{}_\nu \vec{e}_\mu$, where for example

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{1.36}$$

where $|\beta| < 1$ and $\gamma = (1-\beta^2)^{-1/2}$. Under this transformation,

$$\begin{aligned} \vec{e}'_0 &= \gamma \vec{e}_0 + \beta\gamma \vec{e}_1 \\ \vec{e}'_1 &= \beta\gamma \vec{e}_0 + \gamma \vec{e}_1 \\ \vec{e}'_2 &= \vec{e}_2 \\ \vec{e}'_3 &= \vec{e}_3 \,. \end{aligned} \tag{1.37}$$

On the other hand, the coordinates transform as $x'^\nu = (\Lambda^{-1})^\nu{}_\mu x^\mu$, where

$$(\Lambda^{-1})^\nu{}_\mu = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{1.38}$$

i.e.

$$\begin{aligned} x'^0 &= \gamma x^0 - \beta\gamma x^1 \\ x'^1 &= -\beta\gamma x^0 + \gamma x^1 \\ x'^2 &= x^2 \\ x'^3 &= x^3 \,. \end{aligned} \tag{1.39}$$

(**End of Lecture**: Monday 8 Sep 2025)

1.5 Tensors

We have now seen vector spaces and their dual spaces. We can now start defining even more general objects by putting vector spaces and dual spaces together!

Consider three vector spaces U, V and W over $\mathbb{F}$. We can define the **tensor product** of spaces such as $V \otimes W$, or even $U \otimes V \otimes W$.

1. It is distributive, i.e. for $\vec{a} \in U$ and $\vec{x} \in V$,

$$\begin{aligned} \vec{a} \otimes (\vec{x} + \vec{y}) &= \vec{a} \otimes \vec{x} + \vec{a} \otimes \vec{y} \,, \\ (\vec{a} + \vec{b}) \otimes \vec{x} &= \vec{a} \otimes \vec{x} + \vec{b} \otimes \vec{x} \,; \end{aligned} \tag{1.40}$$

2. It is associative, so that we can chain together three vector spaces like $U \otimes V \otimes W$ without worrying about whether it's $(U \otimes V) \otimes W$ or $U \otimes (V \otimes W)$;

3. It commutes with $\mathbb{F}$, i.e.

$$\lambda(\vec{a} \otimes \vec{x}) = (\lambda \vec{a}) \otimes \vec{x} = \vec{a} \otimes (\lambda \vec{x}) \,, \tag{1.41}$$

but it is *not* commutative over the vectors, i.e. $\vec{a} \otimes \vec{b} \neq \vec{b} \otimes \vec{a}$ in general.

Consider a vector space V with a basis $\{\vec{e}_\mu\}$. This basis induces a basis in tensor products of V and V^* in the natural way; for example, in $V^* \otimes V^*$, this induces the basis $\vec{e}^{*\mu} \otimes \vec{e}^{*\nu}$. These basis vectors act on pairs of basis vectors, $(\vec{e}_\alpha, \vec{e}_\beta)$, in the expected way, so for our $V^* \otimes V^*$ example, we get

$$\vec{e}^{*\mu} \otimes \vec{e}^{*\nu}(\vec{e}_\alpha, \vec{e}_\beta) = \delta^\mu_\alpha \delta^\nu_\beta \,. \tag{1.42}$$

It can also act on elements of $V \otimes V$ in a manner that you might also expect:

$$\vec{e}^{*\mu} \otimes \vec{e}^{*\nu}(\vec{e}_\alpha \otimes \vec{e}_\beta) = \delta^\mu_\alpha \delta^\nu_\beta \,. \tag{1.43}$$

A good example of what a tensor is and what it does is the metric itself, which we often refer to as **metric tensor**. It is a tensor in $V^* \otimes V^*$,

$$\mathbf{g} = g_{\mu\nu} \vec{e}^{*\mu} \otimes \vec{e}^{*\nu} \,. \tag{1.44}$$

It acts on pairs of vectors $\vec{x}$ and $\vec{y}$ and returns a number in $\mathbb{F}$:

$$\mathbf{g}(\vec{x}, \vec{y}) = g_{\mu\nu} \vec{e}^{*\mu} \otimes \vec{e}^{*\nu}(x^\alpha \vec{e}_\alpha, y^\beta \vec{e}_\beta) = g_{\mu\nu} x^\alpha y^\beta \delta^\mu_\alpha \delta^\nu_\beta = g_{\mu\nu} x^\mu y^\nu = x^\mu y_\mu \,. \tag{1.45}$$

As with vectors and covectors, once we've picked a basis, we will only need to worry about the components, with the understanding that the object itself is specified by both the components and the basis, and that the spaces act in the natural way that you expect.

Under a change of basis $\{\vec{e}_\mu\} \mapsto \{\vec{e}'_\mu\}$, the metric itself undergoes a transformation:

$$g_{\mu\nu} \mapsto g'_{\mu\nu} = \langle \vec{e}'_\mu, \vec{e}'_\nu \rangle = \langle a^\sigma{}_\mu \vec{e}_\sigma, a^\lambda{}_\nu \vec{e}_\lambda \rangle = a^\sigma{}_\mu a^\lambda{}_\nu g_{\sigma\lambda} \,. \tag{1.46}$$

Each lower index is acted on by the change-of-basis transformation, with each transformation given by a covariant transformation. We say therefore that $g_{\mu\nu}$ a doubly covariant **tensor**, which explains why we often also refer to $g_{\mu\nu}$ as the **metric tensor**. You can see now that we can look into objects with more general number of indices, say $Q^{\alpha\beta}{}_{\gamma\delta\epsilon}$, which transforms as

$$Q'^{\alpha\beta}{}_{\gamma\delta\epsilon} = (a^{-1})^\alpha{}_{\alpha'} (a^{-1})^\beta{}_{\beta'} a^{\gamma'}{}_\gamma a^{\delta'}{}_\delta a^{\epsilon'}{}_\epsilon Q^{\alpha'\beta'}{}_{\gamma'\delta'\epsilon'} \,, \tag{1.47}$$

which is a doubly contravariant, triply covariant tensor, or a type $(2,3)$ tensor. The total number of indices is what we call the **rank** of the tensor. Notice how when we were writing down the transformation of $Q^{\alpha\beta}{}_{\gamma\delta\epsilon}$, the indices lined up: we contracted upper indices with lower indices, so that under each application of the change-of-basis transformation, upper indices remain upper indices.

Another thing you will notice is that I have been very careful with the relative positions of the tensors. This is good practice, but very often you'll find people get sloppy and collapse all the indices when they think the notation is obvious. The one tensor where this is always okay is the Kronecker delta δ^μ_ν, since we always have this tensor returning 1 if $\mu = \nu$ and 0 if $\mu \neq \nu$ regardless of the position of the indices.

1.5.1 Tensor algebra

So what can we do with tensors? Well, we can add tensors together, but we have to ensure that you're adding things that are transforming with a change of bases in the same way. So for example

$$A^\mu{}_{\nu\lambda} = B^{\mu\tau}{}_{\nu\lambda\tau} + C^\mu{}_{\nu\lambda} \tag{1.48}$$

is legal, but

$$A^{\mu}{}_{\nu\lambda} \overset{\text{wrong}}{=} B^{\nu}{}_{\mu\lambda} + C^{\mu}{}_{\nu\lambda\sigma\sigma} + D^{\mu}{}_{\nu\lambda\tau} \tag{1.49}$$

makes no sense.

We can multiply tensors together. Suppose we have tensors $A^{\mu}{}_{\nu\lambda}$ and $B^{\mu}{}_{\nu\lambda\tau}$, which are tensors of type $(1,2)$ and $(1,3)$ respectively. Then, we can multiply them together to get

$$C^{\alpha\beta}{}_{\nu\lambda\rho\sigma\tau} = A^{\alpha}{}_{\nu\lambda} B^{\beta}{}_{\rho\sigma\tau} \,, \tag{1.50}$$

which is a tensor of type $(2,5)$.

You can also contract indices, or equivalently, multiply tensors by the metric, so that for example

$$C^{\alpha\beta}{}_{\alpha\beta\rho\sigma\tau} = g^{\alpha\lambda} g^{\beta\mu} C_{\lambda\mu\alpha\beta\rho\sigma\tau} \tag{1.51}$$

is now a tensor of type $(0,3)$. Contracting two vectors is a special case that leads to a real number, also called a scalar or sometimes an **invariant**, which does not transform under a change of basis. Let me stress again: *you must contract one upper and one lower index!* You can write down objects like $B_{\alpha\beta\beta}$, but these objects are *not* tensors. Notice also that free indices (indices that are not contracted) on the left and right sides of an equation must match, to tell you which components of the tensor you're talking about consistently.

At this point, you've seen a lot of index notation—how to raise and lower indices, how to contract indices and so on, and what these operations mean. Let me summarize by giving you some golden rules of index notation:

1. Free indices (indices that are not summed over) must agree on both the left-hand side and the right-hand side of any equation.
2. Indices that are not free should be contracted in pairs, with one upper and one lower index (unless in $\mathbb{R}^n$, where the position doesn't matter).
3. Contracted indices can always be relabeled, since they are dummy indices.
4. There should never be more than two of the same index appearing in a term formed by the product of a bunch of tensors, vectors etc. If this happens to you, it is likely because sum of the indices are supposed to be contracted in a sum. Relabel your contracted indices!

1.5.2 Example: Rotations and Lorentz transformations

Let's study the properties of basis transformations $a^{\mu}{}_{\nu}$ that leave the metric invariant. In $\mathbb{R}^n$, these are the transformations such that angles and lengths are all preserved. Under the transformation $a^{\mu}{}_{\nu}$, the metric tensor transforms as

$$g_{\mu\nu} \mapsto a^{\sigma}{}_{\mu} a^{\lambda}{}_{\nu} g_{\sigma\lambda} \,, \tag{1.52}$$

Transformations O that leave the metric invariant are therefore of the form

$$O^{\sigma}{}_{\mu} g_{\sigma\lambda} O^{\lambda}{}_{\nu} = g_{\mu\nu} \,. \tag{1.53}$$

In $\mathbb{R}^n$, if we start with the canonical metric $g_{\sigma\lambda} = \delta_{\sigma\lambda}$, then these transformations must satisfy

$$O^{\sigma}{}_{\mu} \delta_{\sigma\lambda} O^{\lambda}{}_{\nu} = \delta_{\mu\nu} \implies O^{\lambda}{}_{\mu} O_{\lambda\nu} = \delta_{\mu\nu} \implies (O^{-1})_{\mu}{}^{\lambda} = O^{\lambda}{}_{\mu} \,. \tag{1.54}$$

Therefore, as matrices, we must have $O^{-1} = O^{\mathsf{T}}$. The set of all such matrices is the **orthogonal matrices**, which can be thought of as a **group** called $O(n)$, called the **orthogonal group**.[6] This group is made up of matrices corresponding to *rotations* and *reflections*. The group containing matrices corresponding only to rotations is called the **special orthogonal group**, $SO(n)$.

In Minkowski space, we have instead

$$\Lambda^{\sigma}{}_{\mu}\eta_{\sigma\lambda}\Lambda^{\lambda}{}_{\nu} = \eta_{\mu\nu}\,. \tag{1.55}$$

The Λ matrices also form a group that we call the **Lorentz group**, $O(1,3)$ for spacetime. The set of all Lorentz transformations (both boosts and spatial rotations), as well as time reversal and reflections, make up this group. Again, the group with just boosts and rotations is called $SO(1,3)$.

[6] A group G is a set of elements with an operation $\cdot$, that obeys the following: 1) associativity, i.e. $(a\cdot b)\cdot c = a\cdot(b\cdot c)$ for all $a,b,c\in G$; 2) it contains an identity element e in G such that $a\cdot e = e\cdot a = a$, and 3) every element a in G has an inverse b, such that $a\cdot b = b\cdot a = e$.

1.5.3 Example: Linear transformations

In the crash course in linear algebra, I mentioned the concept of linear transformations briefly. Let's linear transformations $M : V \to V$, mapping vectors in a vector space V over a field $\mathbb{F}$ to other vectors in V. The linear transformation must satisfy the following property:

$$M(\lambda\vec{x} + \mu\vec{y}) = \lambda M(\vec{x}) + \mu M(\vec{y}) \tag{1.56}$$

for all $\vec{x},\vec{y}\in V$, and $\lambda,\mu\in\mathbb{F}$. This object exists independently of any basis, but given a basis, it can be represented by a matrix $M^{\mu}{}_{\nu}$, obtained by examining the action of the transformation on the basis vectors:

$$M(\vec{e}_\mu) = M^{\nu}{}_{\mu}\vec{e}_\nu\,. \tag{1.57}$$

Suppose M acting on some arbitrary vector $\vec{x} = x^\mu\vec{e}_\mu$ gives the result $\vec{y} = y^\nu\vec{e}_\nu$. We can see however that

$$\vec{y} = y^\nu\vec{e}_\nu = M(x^\mu\vec{e}_\mu) = x^\mu M(\vec{e}_\mu) = M^{\nu}{}_{\mu}x^\mu\vec{e}_\nu\,, \tag{1.58}$$

or in other words, the components transform as

$$y^\nu = M^{\nu}{}_{\mu}x^\mu\,. \tag{1.59}$$

We can therefore see that given a basis, M behaves just like matrix multiplication, as we already knew to be true from linear algebra.

In another basis related to the old one via $\vec{e}_\nu = a^{\sigma}{}_{\nu}\vec{e}'_\sigma$, we see that

$$M(\vec{e}_\mu) = M^{\nu}{}_{\mu}\vec{e}_\nu = M^{\nu}{}_{\mu}a^{\sigma}{}_{\nu}\vec{e}'_\sigma\,, \tag{1.60}$$

but also

$$M(\vec{e}_\mu) = M(a^{\lambda}{}_{\mu}\vec{e}'_\lambda) = a^{\lambda}{}_{\mu}M(\vec{e}'_\lambda) \implies M(\vec{e}'_\lambda) = (a^{-1})^{\mu}{}_{\lambda}M(\vec{e}_\mu)\,. \tag{1.61}$$

Therefore, in the new basis,

$$M(\vec{e}'_\lambda) = a^{\sigma}{}_{\nu}M^{\nu}{}_{\mu}(a^{-1})^{\mu}{}_{\lambda}\vec{e}'_\sigma\,, \tag{1.62}$$

i.e. under the change of basis,

$$M^{\sigma}{}_{\lambda} \mapsto a^{\sigma}{}_{\nu}M^{\nu}{}_{\mu}(a^{-1})^{\mu}{}_{\lambda}\,, \tag{1.63}$$

which is the transformation rule for type (1,1) tensors, and also what you may be familiar with in linear algebra about the change of basis of linear transformations.[7]

[7] If you are confused as to why a^{-1} appears to be acting on the *covariant* index of $M^{\nu}{}_{\mu}$, note that we defined a as the transformation mapping the $'$ basis to the unprimed, and so a^{-1} should map the unprimed covariant basis vectors to $'$ basis vectors. It's all consistent!

Compare this with the transformation rule for the metric tensor,

$$g_{\mu\nu} \mapsto (a^{-1})^{\sigma}{}_{\mu}(a^{-1})^{\lambda}{}_{\nu} g_{\sigma\lambda}\,, \tag{1.64}$$

and you'll notice a big difference! And that's because the metric tensor is a type (0,2) tensor. The key lesson here is that *2D tensors are much more than just matrices*; while it can be useful to write out tensor components as matrices, one needs to keep in mind that tensors also come equipped with a transformation rule!

1.5.4 Invariants of 2D tensors

Invariants are extremely useful: they don't transform under a basis change, and so they're very easy to deal with. They're also giving you information about the tensor that is independent of the basis. The first commonly discussed invariant of 2D tensors are the **determinant**, which you can compute from the matrix representation of the tensor; we will discuss this in much greater detail after we've built up some machinery to discuss it. The second invariant is called the **trace**. In matrix language, this is simply the sum of the terms along the diagonal. It's hard to understand how this could be an invariant from the matrix perspective, but in terms of indices, it is simply given by

$$\text{tr}(M) = g^{\mu\nu}M_{\mu\nu} = g_{\mu\nu}M^{\mu\nu} = \delta^{\nu}_{\mu}M^{\mu}{}_{\nu} = M^{\mu}{}_{\mu}\,, \tag{1.65}$$

where I'm showing you how to obtain the trace for 2D tensors of all types.

(**End of Lecture**: Wednesday Sep 10 2024)

1.5.5 Symmetric and antisymmetric tensors

A tensor is said to be **symmetric** in some indices if the tensor values are the same when the two indices are swapped. For example, we say that $S^{\mu\nu}$ is a symmetric tensor if $S^{\mu\nu} = S^{\nu\mu}$. Note that we have also

$$S^{\mu}{}_{\nu} = g_{\nu\alpha}S^{\mu\alpha} = g_{\nu\alpha}S^{\alpha\mu} = S_{\nu}{}^{\mu}\,, \tag{1.66}$$

and so the position of the indices doesn't matter in determining whether a tensor is symmetric or not. Moreover, under a change of basis,

$$S'^{\alpha\beta} = a^{\alpha}{}_{\mu}a^{\beta}{}_{\nu}S^{\mu\nu} = a^{\alpha}{}_{\mu}a^{\beta}{}_{\nu}S^{\nu\mu} = S'^{\beta\alpha}\,, \tag{1.67}$$

and so a symmetric matrix stays symmetric under a change of basis.

We can similarly define a tensor to be **antisymmetric** in some indices if, under a swap of the two indices, the tensor picks up a minus sign, e.g. $A^{\mu\nu} = -A^{\nu\mu}$. Again, a tensor that is antisymmetric when indices are in one position are still antisymmetric if the indices are raised or lowered; they also remain antisymmetric under an arbitrary change of basis.

The contraction of a symmetric and an antisymmetric tensor is always zero, since e.g.

$$S_{\mu\nu}A^{\mu\nu} = S_{\nu\mu}A^{\mu\nu} = -S_{\nu\mu}A^{\nu\mu} = -S_{\mu\nu}A^{\mu\nu} = 0\,, \tag{1.68}$$

where in the second last step I have simply relabeled the contracted indices (since they are dummy indices), and noted that $S_{\mu\nu}A^{\mu\nu}$ is equal to its negative, and therefore has to be zero.

Every tensor can always be decomposed into a symmetric piece and an antisymmetric piece. To see this, take an arbitrary tensor $B^{\mu\nu}$. We can always rewrite this as

$$B^{\mu\nu} = \frac{1}{2}(B^{\mu\nu} + B^{\nu\mu}) + \frac{1}{2}(B^{\mu\nu} - B^{\nu\mu}) \equiv S^{\mu\nu} + A^{\mu\nu}\,, \tag{1.69}$$

where $S^{\mu\nu} = (B^{\mu\nu} + B^{\nu\mu})/2$ a symmetric tensor, and $A^{\mu\nu} = (B^{\mu\nu} - B^{\nu\mu})/2$ is antisymmetric.

1.5.6 Kronecker and Levi-Civita tensors

We'll now discuss two special tensors that arise very commonly in tensor algebra. The first of them, the **Kronecker delta** δ^μ_ν, is defined as a (1,1) tensor that, in some basis, is unity if $\mu = \nu$ and zero otherwise. Let's check what happens under an arbitrary change of basis:

$$\delta^\mu_\nu \mapsto (a^{-1})^\lambda{}_\nu a^\mu{}_\sigma \delta^\sigma_\lambda = \delta^\mu_\nu\,. \tag{1.70}$$

In other words, *the Kronecker delta always has the same numerical components in all coordinate systems*.[8] [9]

The **Levi-Civita symbol** $\epsilon_{\mu_1\mu_2\cdots\mu_n}$ is defined as an object with n indices such that $\epsilon_{12\cdots n} = 1$, and $\epsilon_{\cdots i_p \cdots i_q \cdots} = -\epsilon_{\cdots i_q \cdots i_p \cdots}$, i.e. every time two indices are exchanged, the result differs by a minus sign. This definition guarantees that when two indices are equal, the Levi-Civita symbol is zero. One particularly important use-case of the Levi-Civita symbol is in expressing antisymmetric quantities in component form. For example, the **cross product** $\vec{a}\times\vec{b}$ can be written as

$$(\vec{a}\times\vec{b})_k = \epsilon_{ijk} a^i b^j\,. \tag{1.71}$$

The **determinant** of an $n\times n$ matrix M can likewise be written as (assuming summation over repeated indices)

$$\epsilon_{\mu_1\cdots\mu_n}\mathrm{det}(M) = \epsilon_{\nu_1\cdots\nu_n} M^{\nu_1}{}_{\mu_1}\cdots M^{\nu_n}{}_{\mu_n}\,, \tag{1.72}$$

or

$$\mathrm{det}(M) = \frac{1}{n!}\epsilon^{\mu_1\cdots\mu_n}\epsilon_{\nu_1\cdots\nu_n} M^{\nu_1}{}_{\mu_1}\cdots M^{\nu_n}{}_{\mu_n}\,. \tag{1.73}$$

Let's suppose there is an n-dimensional tensor $\eta_{\mu_1\mu_2\cdots\mu_n}$ whose components coincide with $\epsilon_{\mu_1\mu_2\cdots\mu_n}$ in one particular basis. Then under a change of basis,

$$\begin{aligned}\eta_{\mu_1\cdots\mu_n} &\mapsto a^{\nu_1}{}_{\mu_1} a^{\nu_2}{}_{\mu_2}\cdots a^{\nu_n}{}_{\mu_n}\epsilon_{\nu_1\cdots\nu_n}\\ &= \epsilon_{\mu_1\cdots\mu_n}\mathrm{det}(a) = \eta_{\mu_1\cdots\mu_n}\mathrm{det}(a)\,.\end{aligned} \tag{1.74}$$

We see that the Levi-Civita symbol is *almost* a tensor whose components do not transform, up to a pesky determinant.

At this point, let's examine the determinant of the metric tensor itself, $g \equiv \mathrm{det}(g_{\mu\nu})$. We see that under the same transformation, the determinant after the change of basis is

$$g' \equiv \mathrm{det}(g'_{\mu\nu}) = \mathrm{det}(a^\lambda{}_\mu a^\sigma{}_\nu g_{\lambda\sigma}) = (\mathrm{det}(a))^2 g\,, \tag{1.75}$$

and therefore the quantity $\sqrt{|g'|} = |\mathrm{det}(a)|\sqrt{|g|}$; note that the absolute value is important, since the metric tensor can have negative determinant (e.g. in Minkowski space!). Now let's consider the object

$$\varepsilon_{\mu_1\mu_2\cdots\mu_n} = \sqrt{|g|}\epsilon_{\mu_1\mu_2\cdots\mu_n} \tag{1.76}$$

[8] This explains why the indices are almost never distinguished in the literature: both $\delta^\mu{}_\nu$ and $\delta_\nu{}^\mu$ act the same way on any object: just replace μ with ν or vice-versa.

[9] Sometimes, people talk about the (0,2) tensor $\delta_{\mu\nu} = g_{\mu\lambda}\delta^\lambda_\nu = g_{\mu\nu}$, which as you can see is really the metric tensor (whose coordinates *can* change with a change of basis). I will usually avoid using $\delta_{\mu\nu}$, unless we are talking about $\mathbb{R}^n$, where the index position doesn't matter.

[10] An easy way to see that $\varepsilon_{\mu_1\cdots\mu_n}$ needs to pick up a sign under non-orientation-preserving transformations is by considering what happens to a cross-product under reflection, where $\vec{x}\mapsto -\vec{x}$. With the sign flip, $\varepsilon^{ijk}x_j y_k \mapsto -\varepsilon^{ijk}x_j y_k$ as needed, since the final vector as a result of the cross product also needs to pick up a minus sign under reflection.

Under the same transformation, we now see that provided $\det(a) > 0$, which are referred to as *orientation preserving changes of basis*,[10]

$$\varepsilon_{\mu_1\cdots\mu_n} \mapsto \sqrt{|g|}\det(a)\epsilon_{\mu_1\cdots\mu_n} = \sqrt{|g'|}\epsilon_{\mu_1\cdots\mu_n}\,, \tag{1.77}$$

and so once again, we find a tensor known as the **Levi-Civita tensor** $\varepsilon_{\mu_1\cdots\mu_n}$ that always has the same form in any basis (although we must evaluate the determinant of the transformed metric, which is in general different for each basis).

If we limit ourselves to rotations in $\mathbb{R}^n$ and boosts (and rotations) in Minkowski space, then $\varepsilon_{\mu_1\cdots\mu_n} = \epsilon_{\mu_1\cdots\mu_n}$ is a good old tensor. However, if we include transformations like reflections, for example, then we need a lot more care.[11]

[11] You might have heard a joke about how 'a tensor is something that transforms like a tensor'. We can always exclude some class of basis changes from consideration to make a class of objects well-behaved and tensor-like.

Let's circle back and show that the determinant is indeed an invariant under a change of basis. We have

$$\begin{aligned}
\det(M') =& \frac{1}{n!}\epsilon^{\mu_1\cdots\mu_n}\epsilon_{\nu_1\cdots\nu_n}M'^{\nu_1}{}_{\mu_1}\cdots M'^{\nu_n}{}_{\mu_n}\\
=& \frac{1}{n!}\epsilon^{\mu_1\cdots\mu_n}\epsilon_{\nu_1\cdots\nu_n}\left(a^{\sigma_1}{}_{\mu_1}(a^{-1})^{\nu_1}{}_{\rho_1}M^{\rho_1}{}_{\sigma_1}\cdots a^{\sigma_n}{}_{\mu_n}(a^{-1})^{\nu_n}{}_{\rho_n}M^{\rho_n}{}_{\sigma_n}\right)\\
=& \frac{1}{n!}a^{\sigma_1}{}_{\mu_1}\cdots a^{\sigma_n}{}_{\mu_n}\epsilon^{\mu_1\cdots\mu_n}(a^{-1})^{\nu_1}{}_{\rho_1}\cdots(a^{-1})^{\nu_n}{}_{\rho_n}\epsilon_{\nu_1\cdots\nu_n}\\
&\times M^{\rho_1}{}_{\sigma_1}\cdots M^{\rho_n}{}_{\sigma_n}\\
=& \frac{1}{n!}\epsilon^{\sigma_1\cdots\sigma_n}\det(a)\epsilon_{\rho_1\cdots\rho_n}\det(a^{-1})M^{\rho_1}{}_{\sigma_1}\cdots M^{\rho_n}{}_{\sigma_n}\\
=& \det(M)\,,
\end{aligned} \tag{1.78}$$

since $\det(a)\det(a^{-1}) = 1$.

1.5.7 Isotropic Cartesian tensors

Another special tensor that we'll now discuss is the isotropic, Cartesian tensor. Consider a Cartesian coordinate system with orthonormal basis vectors, so that $g_{ij} = \delta_{ij}$, the Kronecker delta function. We looked at the set of orthogonal matrices form what we called $O(n)$, which are matrices with the property that $O^{-1} = O^{\mathsf{T}}$. When we perform a change of basis under these matrices, we found that this leaves the metric invariant, since

$$g'_{kl} = O^i{}_k O^j{}_l \delta_{ij} = O_{jk}O^j{}_l = O^{\mathsf{T}}_{kj}O^j{}_l = \delta_{kl}\,. \tag{1.79}$$

You can check that the same thing is true for products of δ_{ij}, for example $T_{ijklmn} = \delta_{ij}\delta_{kl}\delta_{mn}$.

What is the most general form a tensor of rank m that is invariant under an $O(n)$ transformation? These are important questions that have tedious answers to them, and so we won't try to prove these results (see for example Ref. [1], but rather just state them. The most general $O(n)$ invariant tensor of rank 4 I_{ijkl} is

$$I_{ijkl} = \alpha\delta_{ij}\delta_{kl} + \beta\delta_{ik}\delta_{lj} + \gamma\delta_{il}\delta_{jk} \tag{1.80}$$

for some numbers α, β and γ.

How about for $SO(n)$, which is simply all the orthogonal matrices with determinant 1, corresponding to just rotations, and not reflections? We can check, for example, that ϵ_{ijk} is invariant under an $SO(3)$ transformation O, since

$$\begin{aligned}\epsilon'_{lmn} &= O^i{}_l O^j{}_m O^k{}_n \epsilon_{ijk} \\ &= \epsilon_{lmn} \text{det}(O) \\ &= \epsilon_{lmn}\,, \end{aligned} \tag{1.81}$$

since by definition, an $SO(n)$ matrix has determinant 1. With this, it's not surprising that, for example, the most general $SO(4)$-invariant, rank-4 tensor J_{ijkl} is

$$J_{ijkl} = \alpha\delta_{ij}\delta_{kl} + \beta\delta_{il}\delta_{jk} + \gamma\delta_{il}\delta_{jk} + \lambda\epsilon_{ijkl}\,.$$

1.5.8 Irreducible 3D Tensors Under Rotations

Let's start from the vector space $V = \mathbb{R}^3$ over the field $\mathbb{R}$, corresponding to our usual intuition of 3-dimensional vectors, with an inner product given by the usual dot product. Consider T_{ij}, the components of a rank-2 tensor in $V^* \otimes V^*$. Let's consider how T_{ij} transforms under rotations, i.e. under $SO(3)$ transformations. We can write down the transformation law

$$T_{ij} \mapsto T'_{ij} = O^k{}_i O^l{}_j T_{kl}\,. \tag{1.82}$$

We saw, however, that we can always split T_{ij} into symmetric S_{ij} and antisymmetric parts A_{ij}, i.e. $T_{ij} = S_{ij} + A_{ij}$, where

$$\begin{aligned} S_{ij} &= \frac{1}{2}(T_{ij} + T_{ji})\,, \\ A_{ij} &= \frac{1}{2}(T_{ij} - T_{ji})\,. \end{aligned} \tag{1.83}$$

We also showed that under any change of basis, the symmetric part remains symmetric, and the antisymmetric part remains antisymmetric. Intuitively, we've split the vector space $V^* \otimes V^*$ into two **subspaces**, one containing symmetric tensors and the other containing antisymmetric tensors, and every element in $V^* \otimes V^*$ can be written as a sum of elements from these two subspaces. Elements in each subspace don't mix under $SO(3)$ transformations: symmetric tensors stay symmetric, and antisymmetric tensors stay antisymmetric.

Let's take a closer look at A_{ij}. The diagonal components are all zero, i.e. $A_{11} = A_{22} = A_{33} = 0$, by the definition of antisymmetry. Since $A_{12} = -A_{21}$, $A_{13} = -A_{31}$, and $A_{23} = -A_{32}$, we really only have 3 independent components. This is suggestive, because a vector in $\mathbb{R}^3$ also has three independent components; is A_{ij} secretly similar to a vector? The answer is yes! Let's define $B^k \equiv \epsilon^{ijk} A_{ij}$, or equivalently

$$A_{ij} = \frac{1}{2}\epsilon_{ijk} B^k\,. \tag{1.84}$$

You can think of this as stacking up (A_{23}, A_{31}, A_{12}) into a vector. We can check that B_k transforms as a vector, since we know that ϵ^{ijk} is invariant under an SO(3) transformation, i.e.

$$\epsilon^{ijk} A_{ij} \mapsto \epsilon^{ijk} O^m{}_i O^n{}_j A_{mn}\,. \tag{1.85}$$

Recalling that

$$\epsilon^{ijk} O^m{}_i O^n{}_j O^p{}_k = \epsilon^{mnp} \text{det}(O) = \epsilon^{mnp}\,, \tag{1.86}$$

where we've used $\text{det}(O) = 1$ for special orthogonal matrices, and that in matrix form $O^\intercal O = \mathbb{I}$, we also have

$$\epsilon^{ijk} O^m{}_i O^n{}_j O^p{}_k (O^\intercal)^q{}_p = \epsilon^{ijq} O^m{}_i O^n{}_j = (O^\intercal)^q{}_p \epsilon^{mnp}\,. \tag{1.87}$$

Putting everything together, we find

$$B^k = \epsilon^{ijk} A_j \mapsto (O^\intercal)^k{}_p \epsilon^{mnp} A_{mn} = (O^\intercal)^k{}_p B^p \,, \tag{1.88}$$

which is precisely how a contravariant vector component should transform.

Let's take stock. We've shown that any element in $V^* \otimes V^*$ can be written as a sum of an symmetric tensor, and an antisymmetric tensor, which is secretly a vector. Under a change of basis, each part—the symmetric tensor and the vector—transform into another symmetric tensor and a vector respectively, with no mixing of the subspaces.

The symmetric tensor, it turns out, can be broken down even more. Let's write

$$S_{ij} = \frac{s}{3}\delta_{ij} + \left(S_{ij} - \frac{s}{3}\delta_{ij}\right) \,, \tag{1.89}$$

where $s \equiv S^k{}_k$ is the trace of S_{ij} (remember that metric here is simply δ_{ij}). Under an SO(3) rotation, the first term remains invariant, since orthogonal matrices leaves the Kronecker delta Euclidean metric invariant, and the trace s is a scalar and also invariant. The second term transforms nontrivially, but evaluating its trace, we find

$$\delta^{ij}\left(S_{ij} - \frac{s}{3}\delta_{ij}\right) = s - \frac{s}{3}\cdot \delta^{ij}\delta_{ij} = 0 \,.$$

Since any trace is a scalar and an invariant, under any change of basis, $S_{ij} - (s/3)\delta_{ij}$ always remains a **symmetric, traceless tensor**. This object has **five** degrees of freedom—a symmetric tensor has 6 degrees of freedom (S_{11}, S_{22}, S_{33}, S_{12}, S_{13}, S_{23}), and imposing the traceless condition removes 1 degree of freedom.

We have just shown that any rank-2 tensor T_{ij} in Euclidean $\mathbb{R}^3$ can be decomposed as

$$T_{ij} = \frac{s}{3}\delta_{ij} + \frac{1}{2}\epsilon^{ijk} B_k + \left(S_{ij} - \frac{s}{3}\delta_{ij}\right) \,, \tag{1.90}$$

into a scalar, vector and symmetric traceless tensor respectively, with each piece transforming separately under an SO(3) rotation, so that

$$\begin{aligned} T_{ij} \mapsto T'_{ij} &= \frac{s}{3}\delta_{ij} + \frac{1}{2}\epsilon^{ijk} B'_k + \left(S'_{ij} - \frac{s}{3}\delta_{ij}\right) \\ &= \frac{s}{3}\delta_{ij} + \frac{1}{2}\epsilon^{ijk} O^l{}_k B_l + O^k{}_i O^l{}_j \left(S_{kl} - \frac{s}{3}\delta_{kl}\right) \,. \end{aligned} \tag{1.91}$$

The scalar, vector and symmetric traceless tensor are all examples of **irreducible tensors**, i.e. they cannot be decomposed any further into smaller subspaces that transform independently under SO(3), something which I didn't prove, but is true. Very often, noting that T_{ij} lives in the tensor product of two vector spaces, people write this as

$$\mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{3} \oplus \mathbf{5} \,, \tag{1.92}$$

where $\mathbf{1}$ standing for the irreducible scalar, $\mathbf{3}$ for the irreducible vector, and $\mathbf{5}$ for the irreducible symmetric traceless tensor. This decomposition is very useful: very often, you can analyze each part separately, simplifying the problem significantly.

There's a lot more to be said about tensor decomposition, irreducible tensors and their applications in physics, but a thorough exploration of this will touch on group theory and representation theory, which we will not discuss in this course.

(**End of Lecture**: Monday Sep 15 2024)

References

[1] July 2016. URL: https://mathematica.stackexchange.com/questions/77855/finding-basis-of-isotropic-tensors-of-rank-n.