

# What is a tensor ?

Johann Colombano-Rut

February 2016

v3.2

01/05/2020

This is the print version of the post [www.naturelovesmath.com](http://www.naturelovesmath.com)  
[Cliquer ici pour le pdf en francais](#)

The (foolish) purpose of this post is to tackle the concept of tensor, while trying to keep it accessible to the widest audience possible. For the layman (or specialist from another field of study) who just wants to know what this is about without the rigidity and complexity of reading an indecipherable math handbook, for the student who already had a course on tensor calculus but can't figure out some of the "obvious" stuff the teacher didn't linger on, and finally for any math teacher or expert looking for a fresh eye on the subject (meaning not the handbook type) maybe to help describe it and teach it to his own audience.

It doesn't mean there won't be any formalism though, it will just be stripped down to the bare minimum while trying not to take too big of a shortcut when needed.

So, you will find a lot of visual representations along this post, that I hope will assist and guide you to understand each step, but keep in mind that visual aids are just that, they should not be the basis for understanding, only afterthoughts and guides.

A few sections are a little heavy on technical terms and symbols, that's because they are very important (so, to the layman : just get over it, tensors are not for the faint of heart.).

Note : some paragraphs [in red square brackets] have been added or edited in 2016 while some [in blue brackets] in 2020. The sections containing these changes are indicated by a (\*) in the table of content.

Buckle up, you're in for a ride!

A tensor is based on two concepts :

- There are several ways of looking at the components of a vector. We are first going to define them before trying to explain their subtleties and what this really means.
- There is a "mirror world" of vector spaces, and it's the back-and-forth between these that will bring most of the defining properties of tensors.

These two seemingly simple points will help define what is a tensor of order one and help generalize the concept. It is necessary in order to really understand what they are made of and why they are so useful, particularly in physics.

---

## CONTENTS

---

I	COVARIANT AND CONTRAVARIANT COMPONENTS OF A VECTOR . . . . .	4
1	The case of an orthonormal basis . . . . .	4
2	The case of an orthogonal basis not normalized . . . . .	5
II	LINEAR FORMS AND THE DUAL SPACE (*) . . . . .	7
III	LINK BETWEEN VECTOR AND LINEAR FORM (*) . . . . .	9
IV	LINKS BETWEEN CONTRAVARIANT AND COVARIANT COMPONENTS (*) . .	10
1	The case of a non orthogonal basis . . . . .	13
V	LET'S TAKE A LOOK AT MATRICES (*) . . . . .	17
1	The metric tensor's dirty secrets (*) . . . . .	19
2	Multilinearity in the spotlight . . . . .	21
VI	WHAT IS A TENSOR? . . . . .	23
VII	THE CONCEPT OF COVARIANCE IN PHYSICS . . . . .	24
VIII	ZOOLOGY : EXAMPLES OF TENSORS . . . . .	26
IX	A WORD OF PHILOSOPHY . . . . .	27

---

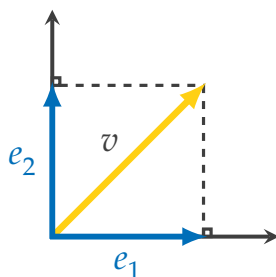
 COVARIANT AND CONTRAVARIANT COMPONENTS OF A VECTOR
 

---

## 1 THE CASE OF AN ORTHONORMAL BASIS

To fix the position of a point in a vector space, one needs as many numbers (real numbers, but not necessarily) as there are dimensions in this space. These are a point's coordinates.

$(1, 1)$  for example gives us the position of a point distant of 1 from the origin. We will here (and through the whole post) use a real Euclidean vector space of two dimensions. So we have a scalar product, and an orthonormal basis  $e = \{e_1, e_2\}$ , that is to say a reference frame made of two orthogonal vectors of norm 1.



A vector is defined by its unique linear combination of the basis vectors (The unicity being of course due to the fact that we chose a basis of reference...).

Here,  $v = v^1 \times e_1 + v^2 \times e_2$

$$\text{So } v = 1 \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \times \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

$$v \quad = \quad 1 \times \begin{array}{c} \longrightarrow \\ e_1 \end{array} \quad + \quad 1 \times \begin{array}{c} \uparrow \\ e_2 \end{array} \quad \text{Soit} \quad \begin{array}{c} \nearrow v \\ \longleftarrow \\ \longrightarrow \end{array}$$

The  $v^i$  are the contravariant components of the vector  $v$  in the basis  $e$  (those are not exponents, they are just superscript numbers or letters). The main thing to remember here is that the contravariant components of a vector are the "number of times we take each vector of the basis". They are the "building blocks" of the vector.

Note that from here, we'll always write down these vector components in a column matrix, and we'll

see that there is a good reason for that convention.

We can find again these vector components by projecting it onto each axis (meaning onto each vector of the basis), which is exactly what the scalar product was born to do :

$$v \cdot e_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1$$

$$v \cdot e_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1$$

These are the covariant components of the vector  $v$  (noted via subscripts) :

$$\begin{aligned} v_1 &= 1 \\ v_2 &= 1 \end{aligned}$$

The covariant components are the result of a projection, so they are images of a linear form applied to the vector. Even if that sentence does not make sense right now, it will soon, so bear with me.

This covariant / contravariant distinction doesn't seem to mean much right now because they are the same for this vector  $v$ . We will see that here it is because of the orthonormal basis (and the fact that we are using the usual euclidean scalar product), and that in most other cases those components really are different.

## 2 THE CASE OF AN ORTHOGONAL BASIS NOT NORMALIZED

Let's take a situation where we can see this difference. Suppose we need to change the scale of our vectors by taking an orthogonal basis that's "smaller" (so it is not of norm 1, it is not normalized)

$$b = \left\{ \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix} \right\}$$

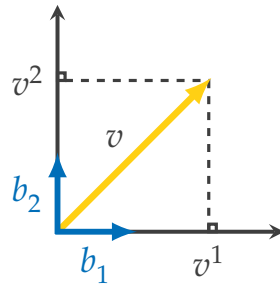
Remark : By smaller, we mean that its vectors have a smaller norm than the first basis, when we express every vector in the standard basis.

We can easily show that the contravariant components of the vector  $v$  become  $\begin{pmatrix} v^1 \\ v^2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$

because  $\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2 \times \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix} + 2 \times \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix}$ .

$$\begin{array}{c} \text{yellow arrow } v \\ \nearrow \end{array} = 2 \times \begin{array}{c} \text{blue arrow } b_1 \\ \rightarrow \end{array} + 2 \times \begin{array}{c} \text{blue arrow } b_2 \\ \uparrow \end{array} \quad \text{Soit} \quad \begin{array}{c} \text{yellow arrow } v \\ \nearrow \\ \text{blue arrows } b_1, b_2 \\ \rightarrow \uparrow \end{array}$$

While the basis "shrunk", the contravariant components "got bigger" : that's why they're called contravariant.



$$v^i = 2 \quad \text{in the basis } (b_i)$$

Note that the length (meaning Euclidean norm) of the vector hasn't changed, only its relation to the basis has changed.

Its covariant components are then  $(v_1 \ v_2) = (\frac{1}{2} \ \frac{1}{2})$

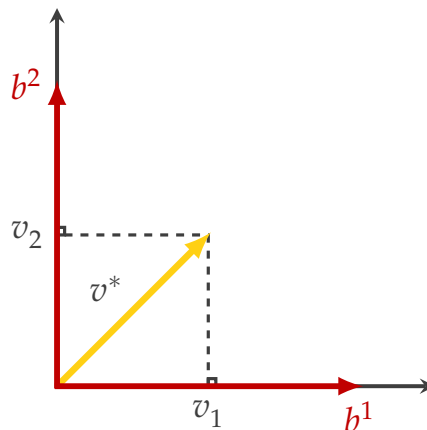
$$\text{because } \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix} = \frac{1}{2}.$$

So here it's the contrary : the covariant component "shrunk", just like the basis.

But, since the vector hasn't changed, the fact that its components have "shrunk" means that they must be expressed in a "bigger" basis, so a different one...

Why did we change the basis ?

This is because to obtain our covariant components, we applied a scalar product, so we went into the dual space... We will come back to that in a moment. For now, let's just say that the covariant components are in fact the dual vector's components, which we note  $v^*$  expressed in a new basis  $(b^i)$  :



$$v_i = \frac{1}{2} \quad \text{in the basis } (b^i)$$

What's a dual space ? That's a very good question, thanks for asking.

# II

---

## LINEAR FORMS AND THE DUAL SPACE (\*)

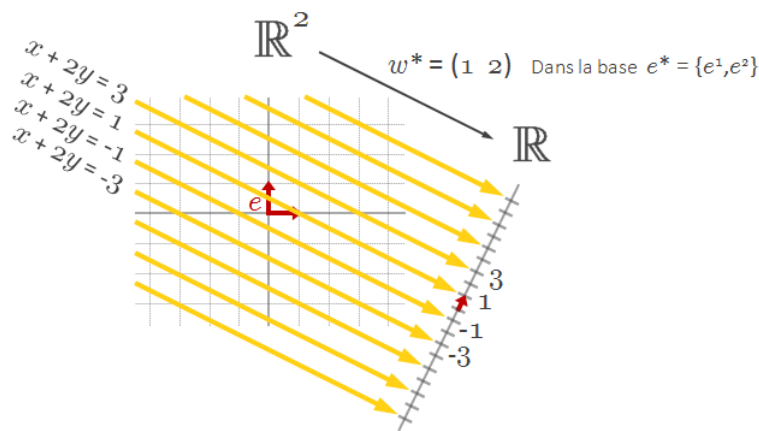
---

Linear forms are the "second heart" of tensors. Now we're stepping on the accelerator it's going to become abstract. We're not in Kansas anymore, Toto.

A linear form is a sort of linear function on a vector space. We are still in the "nice" Euclidean space  $\mathbb{R}^2$ . Here, a linear form is a function that takes a vector and transforms it into a linear combination of its components (giving us a real number).

Let's take a vector  $w$  of components  $\begin{pmatrix} x \\ y \end{pmatrix}$  and apply the linear combination  $x + 2y$ .

We note that linear form  $w^* : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto x + 2y$  and we get for example  $w^*\left(\begin{pmatrix} 4 \\ 5 \end{pmatrix}\right) = 4 + 2 \times 5 = 14$ .



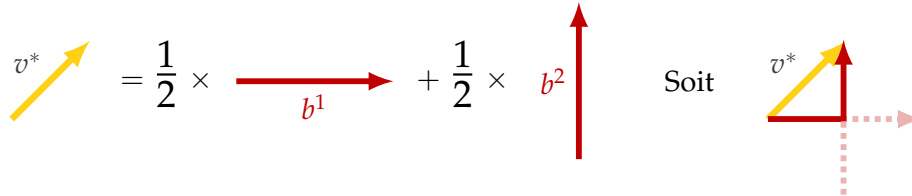
[ A linear form is a projection : all the points situated on a particular arrow are sent to the same real number. A linear form is then a "direction of projection" in a way. Note that we could have represented the real number line in a different position, that one is just convenient for visualization. ]

Since a linear form is defined by a (unique) linear combination, we can represent it with two numbers, just like a vector. Here we have  $w^*$  with components  $(w_1 \ w_2) = (1 \ 2)$  that we'll write in a row matrix. Just like vectors, linear forms "live" in a vector space of their own (that we are gonna call "dual space" of  $\mathbb{R}^2$ ), and their components will depend on a basis (which will be, you guessed it, the dual basis).

Up until now, everything is fine. We've defined a new world, the dual space, but it is very similar to our well-known vector space, so nothing to fret about. By the way, linear forms are also called

covectors. Figures.

BUT (yeah, it couldn't last, could it?), everything is reversed. The building blocks of a linear form are called its covariant components :



[Let's recap :

- A vector is defined by the linear combination of its components with the vectors of the basis, so for  $v$  with contravariant components  $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ , we have  $v = 1 \times b_1 + 2 \times b_2$ , so :

$$v = v^i b_i$$

↑  
Invariant

When we change the basis, the vector has to "adapt" its components to compensate for it. This is so  $v$  does not change, and it's why its (contravariant) components get "bigger" when the basis gets "smaller". The vector  $v$  has an intrinsic existence : its existence does not depend on the basis, contrary to its components. Another way of seeing this is to consider that a vector is only defined in the standard basis, sort of "its only natural basis" in a way.

- A linear form is defined by the product of its components by the components of a given vector, so for the linear form  $v^*$  of covariant components  $(1 \ 2)$ , we have  $v^* \left( \begin{pmatrix} x \\ y \end{pmatrix} \right) = 1 \times x + 2 \times y$ . But since the vector is expressed in a given basis, we only need to know how the linear form acts on the basis' vectors, so for each  $i$  we have  $v^*(b_i) = (1 \ 2) b_i$  which we can sum up in :

$$v_i = v \cdot b_i$$

↑  
Invariant

Here when we change the basis, the linear form doesn't have to compensate anything and their components vary in the same way as the basis.

**Remark :** We are here in a euclidean space where  $\cdot$  represents the **usual** scalar product. In a more general quadratic space (or even in a Euclidean space where the scalar product is different), we would have to use the corresponding symmetric bilinear form.

When we use the scalar product to project a vector onto a basis vector to get its covariant component, we are in fact "switching" to the dual space : the projections of  $v$  onto the basis vectors are linear forms, and the results of these projections are the covariant components in the dual basis. These are by definition the components of the dual vector in the dual basis. That's why we changed from the blue basis to the red one.]



## LINK BETWEEN VECTOR AND LINEAR FORM (\*)

Let  $E = \mathbb{R}^2$  be our Euclidian vector space, and  $E^*$  its dual.

[ Here, it is the scalar product that is going to link the elements of  $E$  and those of  $E^*$ . More precisely, it induces an isomorphism :

$$\begin{array}{lcl} E & \longrightarrow & E^* = \mathcal{L}(E, \mathbb{R}) \\ v & \longmapsto & v^* : E \rightarrow \mathbb{R} \\ & & x \mapsto v \cdot x \end{array}$$

This mapping gives us a one-to-one correspondance between  $E$  and  $E^*$ . For each vector  $v \in E$ , there is a unique linear form  $v^* \in E^*$ . The two spaces are said to be isomorphic. Not surprising that they look so much alike!

*Important notes :*

- [ This correspondance is natural and does not depend on the choice of basis, because a scalar product has been added. Without scalar product, there would not be any isomorphism more natural than another.
- Covariant components, on the other hand do vary and depend on the choice of basis (just as contravariant components do). ]
- Since we are in that particular Euclidian space where  $v \cdot v' = xx' + yy'$  (the **usual** scalar product), it follows that  $v$  and  $v^*$  have the same components in the standard basis [ (in fact in any orthonormal basis for this scalar product) ], they are "the same", or "superposed" if we draw them in the same graph.  
If we had a different scalar product, like  $v \cdot v' = 2xx' + 3yy'$  for example, the two would then have different components in the standard basis (since it would not be orthonormal for this new scalar product) and then different covariant and contravariant components in general. This would be a different correspondance between  $E$  and  $E^*$ .
- In the more general case of a quadratic space  $(E, \phi)$  where  $\phi$  is a non-degenerate bilinear symmetric form, not only would  $v$  and  $v^*$  have different components, but every use of the scalar product  $\cdot$  would have to be replaced by  $\phi$ . It would then again be a completely different correspondance between  $E$  and  $E^*$ . The Minkowskian space of special relativity would be an example of that case. ]

# IV

---

## LINKS BETWEEN CONTRAVARIANT AND COVARIANT COMPONENTS (\*)

---

Let's get back to business. We have a vector  $v = v^i b_i$  (definition of a vector as a linear combination of its basis)

Note that we are now using Einstein's notation :  $v^i b_i = \sum_{i=1}^2 v^i b_i = v^1 b_1 + v^2 b_2$ .

This notation only consists in removing the summation symbol, but beware : only opposite indices are summed over (one superscript and one subscript). Never sum over identical indices if they are at the same level.

[ In our example,

$$v = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (\text{in the standard basis}),$$
$$(v^i) = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \quad \text{in the basis } \{b_i\} = \left\{ \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix} \right\}$$

So we indeed have  $v = v^i b_i$ .]

By taking the scalar product by the  $b_j$  of the vector, we get

$$v \cdot b_j = (v^i b_i) \cdot b_j$$

$$v_j = (b_i \cdot b_j) v^i$$

We introduce a special notation for the scalar products of the basis vectors :  $g_{ij} = b_i \cdot b_j$

Just to be clear :  $g$  is a matrix, where the component at the intersection of line  $i$  and column  $j$  is  $b_i \cdot b_j$ .

[ Back to our example,  $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix} = g \begin{pmatrix} 2 \\ 2 \end{pmatrix}$

We have  $g_{11} = b_1 \cdot b_1 = \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix} = \frac{1}{4}$  and  $g_{22} = b_2 \cdot b_2 = \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix} = \frac{1}{4}$  and  $g_{12} = g_{21} = 0$

So  $g = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{pmatrix}$ .

We can also show that  $v^j = g^{ij} v_i$  with  $g^{ij} = b^i \cdot b^j$  making the components of the inverse matrix  $g^{-1}$ .

Let's calculate the components of the dual basis' vectors which we'll note  $b^* = \{b^1, b^2\}$ .

We then have  $b^i b_j = \delta_j^i$  so then  $b^1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$  and  $b^2 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$

And while we're there  $g^{-1} = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$ .

We just verified that  $b^*$  is in fact "bigger" than the original basis. Mystery solved.

Let's remember that  $v^* = v_i b^i$

$$v^* = (1 \ 1) \quad (\text{in the standard dual basis}),$$

$$(v_i) = \left(\frac{1}{2} \ \frac{1}{2}\right) \quad \text{in the dual basis } \{b^i\} = \{(2 \ 0), (0 \ 2)\}$$

Note : this can be generalized. Had we "shrunk" only one vector of the basis in  $E$ , say  $b_1$ , the dual basis vector  $b^2$  would have "grown", but  $b^1$  would have stayed the same. Had we rotated one vector, it would have been the other one's dual that would have reflected that rotation (in the same direction). There is a symmetry, a sort of "mirror effect" between a basis and its dual, so that we always have  $b_i b^j = \delta_i^j$  and that's what generates this reversal.

Beware! Here the dual space is defined via the scalar product, so a vector and its dual are always superposed (except for the basis). With another scalar product, or a pseudo scalar product like in the case of Minkowski's space, the correspondance would be different and the vectors would not be superposed. The metric tensor and the scalar product's associated bilinear form are often confused with one another, but it is very wrong... The two are usually different (though we are particularly fond of the case where they are the same, hence the confusion).

To sum-up, the dual space depends on the scalar product used (or quadratic form), but the one-to-one mapping between the two spaces (the metric tensor) depends not only on the scalar product but also on the chosen basis!

So we now have the means to go directly from one type of components to the other (that is to say from the space to its dual). Let's see what this  $g$  is exactly.

First, note that in the case of an orthonormal basis (norms=1 and orthogonal), the scalar products of each couple of basis vectors would be zero if they're different ( $i \neq j$ ) and equal to 1 otherwise (when  $i = j$ ). Here  $g$  is simply the identity matrix. We then have  $v^i = v_i$  which means equality between the covariant and contravariant components.

(Remark : the notation  $v^i = v_i$  is an abuse, we do not sum indices here!)

Now, it's not going to be as easy in our "shrunken" basis. To understand what the matrix  $g$  is exactly, we will need to take a look at changes of basis. When we go from a basis  $e$  to a basis  $e'$ , we can define a new basis matrix  $P$  so that  $e'_i = P e_j$

Say what? Don't we usually express the old basis in function of the new one? Yes, but it's confusing. Let's do thing in their natural order instead of writing convenience.

With that in mind, the contravariant components of a vector  $v$  will obey  $v^{ti} = P^{-1}v$ , they "vary contrary to the basis" because the matrix is inverted.

The covariant componants will then verify  $v'_i = {}^tPv$ , they "vary like the basis".

- here we are using the **usual** scalar product, so we should really write  $v'_i = {}^tPIv$  in order to not forget that (or  $v_i = {}^tPv^*$  which is the exact same thing).
- [also note that here the products between a matrix and a vector suppose that we wrote the latter as a column matrix. If we chose the row matrix notation, we would need to transpose :  $v_i = v^*P$ ]

From here, we can say in a way that the  $g_{ij}$  and  $g^{ij}$  matrices are the "new basis matrices" between the basis of our vector space and the dual basis of its dual space (it would be the identity matrix in the case of an orthonormal basis in a Euclidean space).

$$\begin{array}{ccc}
 v = \begin{pmatrix} x \\ y \end{pmatrix} & \xrightarrow{I} & v^* = \begin{pmatrix} x \\ y \end{pmatrix} \\
 \downarrow P^{-1} & & \downarrow {}^tP \\
 \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} & \xrightarrow{g_{ij}} & \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}
 \end{array}$$

[Notice that in this graph, all components are written in columns.]

In a way, these matrices transform a vector into a linear form (and vice-versa) by lowering or raising their indices. They transform contravariant components into covariant components (and vice-versa).

If we use the matrix notation, we have  $(v_1 \ v_2) = g \begin{pmatrix} v^1 \\ v^2 \end{pmatrix}$  et  $\begin{pmatrix} v^1 \\ v^2 \end{pmatrix} = g^{-1} (v_1 \ v_2)$ .

This  $g$  is a weird matrix (even when it's the identity). Usually, a matrix represents a linear transformation (be it invertible or not) that applies to a vector (column matrix in our notation) and produces another vector [also written in a column). Here we see that  $g$  screws with our notation, it "lowers the column components into rows" and "raises the row components into columns"!

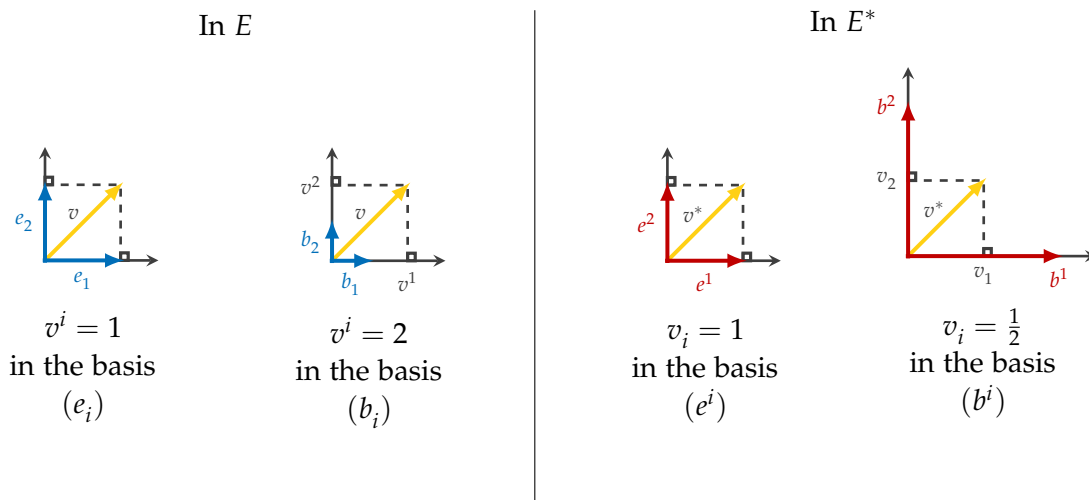
The reason is simple :  $g$  is a **tensor**. So it looks a lot like a usual linear transformation (and it partly is), but works a bit differently. The matrices of linear transformations are tensors with two indices, but one is covariant, the other contravariant.  $g$  is different : it is two times covariant in the form  $g_{ij}$  and two times contravariant when in the form  $g^{ij}$ .

So it's not so simple as to identify  $g$  and its matrix. We should rather say its components table, or its associated matrix that applies not to linear forms but to their associated vectors in the dual space (then it's also an abuse to call the elements of the dual space linear forms, but let's not dwell on that technicality that would complicate things unnecessarily, with my apologies to those who would take offense)...

When  $g$  is two times covariant it transforms a vector into a linear form, and vice-versa when it is two times contravariant. That's a different beast!

$g$  is called the **metric tensor**.

Ok now, it is time to simplify all that "components mess" thanks to the dual space :



In conclusion, by passing to a "smaller basis", the contravariant components of the vector  $v$  got "bigger", while its covariant components (which are also the contravariant components of its associated linear form or dual vector) got "smaller".

Let's do it again in a more complicated case (just barely). It will also be more interesting and more visual too!

### 1 THE CASE OF A NON ORTHOGONAL BASIS

This time, we'll take a basis system  $f$  in which the vectors make a  $\frac{\pi}{3}$  angle. For example we can take  $f = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix} \right\}$  Let's use the same vector  $v$  of components  $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$  in the (orthonormal)

standard basis. Its contravariant components are  $v^i = \begin{pmatrix} \frac{3-\sqrt{3}}{3} \\ \frac{2\sqrt{3}}{3} \end{pmatrix}$

(01/05/2019 update)

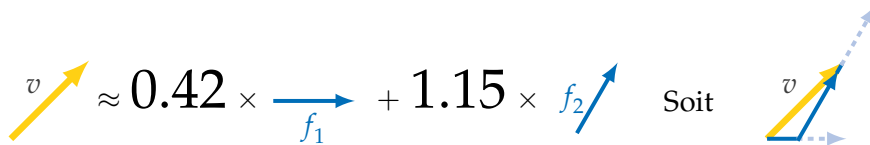
Contravariant components are precisely the vector's components in a different basis. Hence we usually calculate them using a change of basis matrix :  $v^i = P^{-1} \begin{pmatrix} x \\ y \end{pmatrix}$

$P = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{pmatrix}$  because the matrix is made of the basis' vectors. We need to invert it :

$$P^{-1} = \frac{1}{\frac{\sqrt{3}}{2}} \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ -0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{3}} \frac{\sqrt{3}}{2} & -\frac{2}{\sqrt{3}} \frac{1}{2} \\ 0 & \frac{2}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} 1 & -\frac{\sqrt{3}}{3} \\ 0 & \frac{2\sqrt{3}}{3} \end{pmatrix}$$

So  $v^i = \begin{pmatrix} 1 & -\frac{\sqrt{3}}{3} \\ 0 & \frac{2\sqrt{3}}{3} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{3-\sqrt{3}}{3} \\ \frac{2\sqrt{3}}{3} \end{pmatrix}$

We can then check that  $\begin{pmatrix} 1 \\ 1 \end{pmatrix} = v^1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + v^2 \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}$

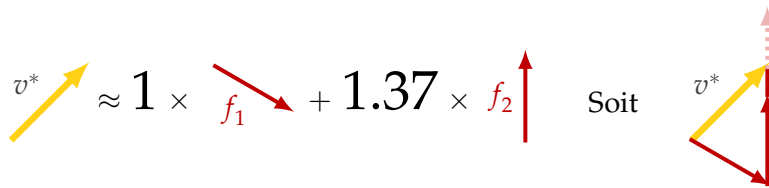


Its covariant components are  $v_i = \begin{pmatrix} 1 & \frac{1+\sqrt{3}}{2} \end{pmatrix}$  via the scalar products  $v \cdot f_i$

Now, to represent them in a graphic, we need the dual basis' vectors.

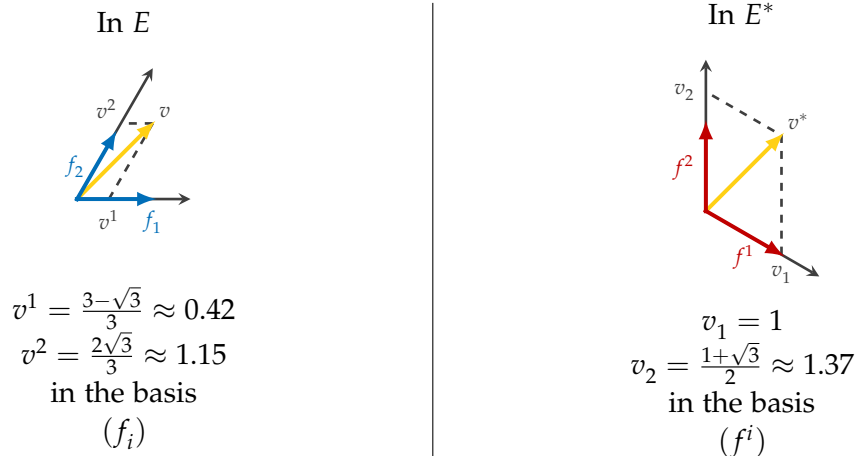
[Whatever the choice of scalar product], they are defined by the following property :  $f^i(f_j) = \delta_j^i$  with  $\delta_j^i = 1$  if  $i = j$  and 0 otherwise. Visually, it means that the dual basis  $f^*$  is orthogonal [(in the sense of the usual euclidean scalar product)] to the  $f$  basis (that is part of the reason why the components are linked the way they are, and the reason why we should never represent the covariant and contravariant components in a common basis, contrary to widely used erroneous representations).

That gives us the dual basis system  $f^* = \{ \begin{pmatrix} 1 & -\frac{\sqrt{3}}{3} \end{pmatrix}, \begin{pmatrix} 0 & \frac{2\sqrt{3}}{3} \end{pmatrix} \}$ .



(01/05/2019 update) Let's review that last calculation :

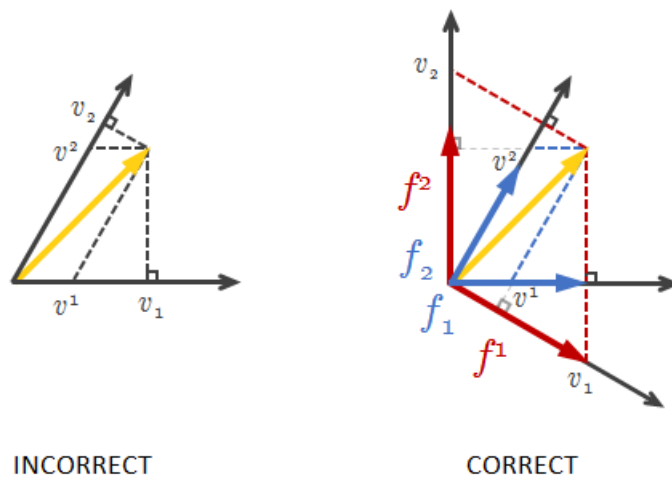
$$v^* = v_i f^i = v_1 f^1 + v_2 f^2 = 1 \times \begin{pmatrix} 1 & -\frac{\sqrt{3}}{3} \end{pmatrix} + \frac{1+\sqrt{3}}{2} \times \begin{pmatrix} 0 & \frac{2\sqrt{3}}{3} \end{pmatrix} = \begin{pmatrix} 1 & 1 \end{pmatrix}$$



Here we can clearly see that the contravariant components  $v^i$  are expressed in the  $f_j$  basis, while the covariant components  $v_i$  are expressed in the dual basis  $f^j$ . Also, one [ALWAYS] reads components by parallel projection since we express the vectors by linear combination of the basis vectors : remember the Chasles relation of vector addition ?

Of course, since the dual basis is orthogonal to the first base, each dashed line is perpendicular to the axis of the other base. So one needs to be careful when representing them in a common graph, and add both basis systems, even if it's a little bit heavy on the eyes.

The orthogonal projection in a space is always equivalent to the parallel projection in its dual space...



We then have a correct visual representation of these two types of components (in a Euclidean space may I remind you). It is rather intuitive that in a "squeezed" basis, we can read coordinates and components in two different ways : by orthogonal projection (covariant components) or by parallel projection (contravariant components). But beware ! There is a trap in the left-hand picture : the orthogonal projection is a parallel projection on the other base... The right way to do it is depicted on the right-hand side.

[Important reminder : What about a non-Euclidean space or one with a non usual scalar product? Well, then  $v$  and  $v^*$  would not be superposed anymore... So beware, in that representation we are here in a very special case! In particular, that representation does not apply to a Minkowski space. So any such representation of the components of vectors in special relativity would be erroneous...]



## LET'S TAKE A LOOK AT MATRICES (\*)

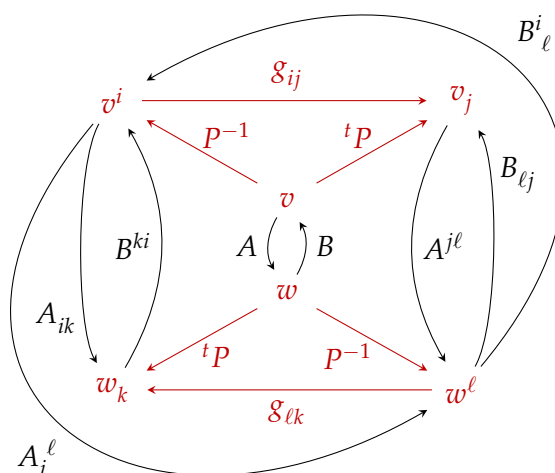
Fasten your seat-belt. No, really do, it's going to get bumpy.

Suppose we let  $v$  be a vector in the standard basis, and  $A$  a linear map that sends this vector unto another vector  $w$  (still in the standard basis).

[Now let's say we decide to change the basis. We write  $P$  the change of basis matrix.

What will the vector become? They don't change of course.

What will become their components? ] The tensors of rank two map their respective components before and after the change of basis :



[Important notes :

- In this graph, all products assume column matrix notation (even for dual components).
- Covariant components are those of "classical" duality through the usual euclidean scalar product (hence  $v = v^*$  and  $w = w^*$ ).

In the convenient case where  $A$  is invertible, we then have a linear map  $B = A^{-1}$  and then  $B^{ki}$  is the inverse tensor of  $A_{ik}$  and  $B_{lj}$  is the inverse of  $A^{j^l}$  and again  $B^i_\ell$  that of  $A_i^\ell$

Remark : The  $-1$  exponent must be reserved only for matrices or linear maps and never be used with tensors, because the inverse of a bilinear form (associated with the matrix  $A$  in the standard basis) is not well defined, one must instead consider the bivector associated with the inverse matrix  $A^{-1}$  in the standard basis.

The inverse tensor of  $A_{ik}$  only exists if  $A$  is invertible and it is then  $B^{ki}$ . The letter  $B$  is used to indicate that this tensor is associated with the matrix  $B = A^{-1}$  after a change of basis, and not with the matrix  $A$ .

As proof,  $(A_{ik}) = {}^tPAP$  and  $(B^{kj}) = P^{-1}B^tP^{-1}$  so it follows that  $(A_{ik})(B^{kj}) = I$ .

The notation  $B^{ki}$  for the inverse tensor is the only correct notation, since the redundancy of  $(A^{-1})^{ki}$  or  $(A^{ki})^{-1}$  would certainly confuse things. For practical use however, we note  $A^{ki}$  even if it technically is an abuse.

Here they are :

- $A_{ik}$  is a two times covariant tensor. It is a bilinear form that takes a vector and returns a linear form, or that transforms two vectors in a scalar  $(x, y) \mapsto {}^txAy$ . [In fact, it is precisely the bilinear form of matrix  $A$  in the standard canonical basis, described after a change of basis  $P$ . The matrices  $(A_{ik})$  and  $A$  are called congruent, they represent the same bilinear form].  
Congruence invariant : the rank.

In our example, we have  $(A_{ik}) = {}^tPAP$  and applied to the contravariant components, we get

$$A_{ik}v^i = w_k$$

- $A_i^\ell$  is a one time covariant and one time contravariant tensor (in that order). It is a linear transformation, that takes one vector and gives back another vector. [In fact, it is precisely the linear transformation of matrix  $A$  in the canonical basis, described after a change of basis  $P$ . The matrices  $(A_i^\ell)$  and  $A$  are called similar, they represent the same linear transformation.]  
Similarity invariants : the rank, the determinant, the trace, the eigen values, the characteristic polynomial, the minimal polynomial, the Jordan form (a complete set of invariants), Young's tables...

In our example, we have  $(A_i^\ell) = P^{-1}AP$  and applied to the contravariant vector, we get

$$A_i^\ell v^i = w^\ell$$

- $A^j_k$  is also a one time contravariant and one time covariant tensor (in that order). It also is a linear transformation, that takes one linear form and gives back another linear form. [In fact, it is precisely the linear transformation of matrix  ${}^tA$  in the dual canonical basis, described after a change of basis  ${}^tP$  in the dual space. The transposed matrices of  $(A^j_k)$  and  $A$  are similar and represent the same linear transformation.]

In our example, we have  $(A^j_k) = {}^tPA{}^tP^{-1}$  and applied to the covariant components, we get

$$A^j_k v_j = w_k$$

Note : If  $A$  is symmetric, then  $A^j_k$  is nothing but the transpose of the previous tensor  $A_i^\ell \dots$  ]

- $A^{j\ell}$  is a two times contravariant tensor. It is a transformation that takes a linear form and gives back a vector, or takes two linear forms and gives back a scalar. In a change of basis, we have  $(A^{j\ell}) = P^{-1}A^tP^{-1}$ . I suggest to call these matrices  $A^{j\ell}$  and  $A$  conbluar. No? similuent then? In fact their transposed matrices are congruent via  $P^{-1}$ , [ which means that they (the transposed ones) represent the same bilinear form. ]

In our example, we have  $(A^{j\ell}) = P^{-1}A^tP^{-1}$  and applied to the covariant components, we get

$$A^{j\ell}v_j = w^\ell$$

Here we can clearly see the bivalent role of  $A$ . It can be seen either as a linear map, a bilinear form or a bivector, because in the canonical basis, they are all described by the same matrix (the dual canonical basis is itself).

However, as soon as we change basis it takes one particular form depending on the type of change it is.  $A$  "acquires" its covariant/contravariant properties depending on the basis.

And for dessert, what links them all : the metric tensor of course since  $g_{kj}g^{ik} = \delta_j^i$  gives us  $A_{ij} = g_{ki}A^k_j = g_{kj}g_{li}A^{lk}$ .

## 1 THE METRIC TENSOR'S DIRTY SECRETS (\*)

Let's rotate our favorite vector  $\vec{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$  by an angle of  $\frac{\pi}{2}$ . To do so, we just have to multiply it by the rotation matrix  $R = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$  to get our new rotated vector  $w = Rv = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ .

The new basis will be  $f = \{f_1, f_2\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix} \right\}$ .

We then have all we need to determine our change of basis matrix  $P$  so that  $f_j = Pe_j$ .

The contravariant components of  $v$  will be  $v^i = P^{-1}v$  and its covariant components  $v_i = {}^tPIv$  (The  $I$  is the identity matrix, here only to remind us that we used the usual scalar product). The dual basis components will then be  $f^j = {}^tP^{-1}e_j$ . (let's not forget that  $e_j = e^j$  since  $e$  is the standard basis)

Bonus :

$$g_{ij} = \underset{\substack{\uparrow \\ \text{Matrix of the bilinear form associated with the scalar product}}}{{}^tPIP}$$

and

$$g^{ij} = \underset{\substack{\uparrow \\ \text{Matrix of the bilinear form associated with the scalar product}}}{P^{-1}I^tP^{-1}}$$

We can easily check that  $g_{ij} = (g^{ij})^{-1}$  and  $v_i = g_{ik}v^k$ .

[What use is the identity matrix here? Well, it's just a reminder that in a more general case of a quadratic space, one just has to replace  $I$  by the matrix of the symmetric bilinear form...]

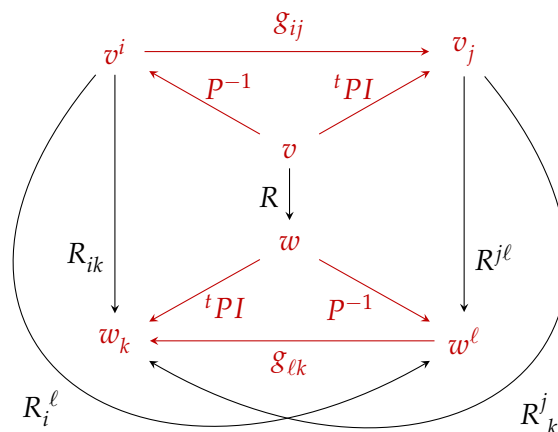
Now, let's see the components of  $w$  in the new basis without using the metric tensor (no real reason except to see what becomes of the rotation matrix) :

For the contravariant components, it's easy, we changed basis, so we are going to do the same for the rotation matrix which will then become  $R' = P^{-1}RP$  and  $w^i = R'v^i$ . This change of basis is due to the fact that the rotation matrix is a mixed tensor, one time covariant and one time contravariant. [Another notation would be  $R' = R_i^\ell$  (careful of the indices' order!).]

Now for the covariant components, it's a little bit different. We would think that using the metric tensor would be enough to raise or lower one index of the tensor  $R'$ , but that's not the case...

In fact, the rotation matrix applied to  $v_i$  will be  $R'' = QRQ^{-1}$  with  $Q = {}^tPI$

We have to make a change of basis "in the other direction" and by transposing. [Or again  $R'' = R_k^j$ .]



We have then a change of basis matrix  $P$  going from  $e$  to  $f$ , and  ${}^tP^{-1}I$  going from  $e$  to  $f^*$ . In both cases we use the matrix  $P$ , but to get to the dual space, we need to invert rows and columns and go in the opposite direction. In a way, the first space is generated by the matrix's columns, and the dual space by the inverse matrix's rows.

But let's not stop just there. Let's go a step further to understand what that means.

[In a given basis, making up the columns of the change of basis matrix  $P$ , we have, by definition of the metric tensor :

Matrix of the bilinear form associated with the scalar product

$${}^tPIP = I$$

Matrix of the metric tensor

As a result, some types of change of basis will give us the same metric tensor. Since orthogonal matrices  $Q$  in a Euclidean space all verify  ${}^tQIQ = I$  so each time the change of basis matrix is a

rotation or a permutation of the axes for example, the metric tensor will then be the identity. It will be the case in every orthonormal basis. Of course, because a rotation does not change the orthonormality of a basis in a Euclidean space!]

Now say we are in a Minkowski space. That is to say we replace the scalar product with a pseudo scalar product represented by  $I_{3,1}$  the diagonal matrix which has three 1 and one  $-1$  or vice-versa depending on the convention (now we are obviously in a four-dimensional space but it doesn't change any formula). For the physicists, this form's little name is  $\eta_{\mu\nu}$ , called the Minkowski form. Then, when the change of basis matrix is a general Lorentz transformation  $\Lambda \in O(3,1)$ , we will get Minkowski's tensor metric because

Matrix of the bilinear form associated with the pseudo scalar product

$${}^t\Lambda \overset{\downarrow}{\eta_{\mu\nu}} \Lambda = \overset{\uparrow}{\eta_{\mu\nu}}$$

Matrix of Minkowski's tensor metric

This rotation would then become  $R' = \Lambda^{-1}R\Lambda$  to transform the contravariant components of vectors, and  $R'' = QRQ^{-1}$  with  $Q = {}^t\Lambda\eta_{\mu\nu}$  to transform covariant components.

Note that  $I$  has here been replaced by  $\eta_{\mu\nu}$ .

Conclusion : as long as we are making changes of basis that belong to the same class of transformations (i.e. generate the same metric tensor), we have unique formulas to effect these changes without even knowing the basis vectors!

That is THE big advantage of tensors : no need to take into account each change of basis individually, we only have to consider a set or class of such changes thanks to the metric tensor!

Boom.

Before our conclusion, let's generalize (just a little bit).

We have here two goals : the first is to describe one way to look at the multilinearity of tensors, and the second is to shed light on a very abstract definition of tensors we often see in old tensorial calculus books. It goes : tensors are objects with indices that transform in a particular way in a change of basis. It's the tensoriality criteria.

## 2 MULTILINEARITY IN THE SPOTLIGHT

A multilinear form is a map that takes in several vectors and combines them to give a number. Their main property is linearity : if a vector is expressed in a new basis, then this change affects the form.

Let's note  $T$  a bilinear form that takes in two vectors in two **different** vector spaces :

$$T : E \times F \rightarrow \mathbb{R}$$

$$(u, v) \mapsto T(u, v) = T_{ij}u^i v^j$$

Suppose  $u$  changes basis, via a change of basis matrix  $P$  in its vector space, and that, independantly,  $v$  also changes basis via a change of basis matrix  $S$  in its own vector space. Then multilinearity implies that  $T(P^{-1}u, S^{-1}v) = T_{ij}(P^i_\ell u^\ell)(S^j_n v^n) = P^i_\ell S^j_n T_{ij}u^\ell v^n$

Now if instead of taking in vectors from different vector spaces we take in vectors in several copies of the same vector space :

$$T : E \times E \rightarrow \mathbb{R}$$

$$(u, v) \mapsto T(u, v) = T_{ij}u^i v^j$$

Then, a change of basis to express one vector implies the same change to any other :  $T(P^{-1}u, P^{-1}v) = T_{ij}(P^i_\ell u^\ell)(P^j_n v^n) = P^i_\ell P^j_n T_{ij}u^\ell v^n$

A tensor is a multilinear form that in a way combines these two situations. It takes in vectors from two different spaces ( $E$  and its dual  $E^*$ ) but these are not independent... A change of basis in one affects the other. For example :

$$T : E \times E^* \rightarrow \mathbb{R}$$

$$(u, v^*) \mapsto T(u, v^*) = T^i_j u^j v^*_i$$

If a change of basis matrix  $P$  acts on the original space's basis vectors, then  $P^{-1}$  will act on its dual basis, and we will have  $T(P^{-1}u, P v^*) = T^i_j (P^j_\ell u^\ell)(P^o_i v^*_o) = P^j_\ell P^o_i T^i_j u^\ell v^*_o$

Note the subtle difference : the tensor  $P^o_i$  is the inverse of  $P^j_\ell$ .

We now have a concrete criteria : if an object does not change in this way during a change of basis, then it is not a tensor...

# VI

---

## WHAT IS A TENSOR ?

---

Now we finally have all the tools we need to answer this question.

Tensors are not just the generalization of the concepts of vectors, matrices and multidimensional tables of numbers. They are first and foremost a generalization of the concept of linear form (and multilinear forms). They combine the advantages of both objects : the contravariant components conserve linear combinations (just as vectors do), and the covariant components conserve their proportions relative to the basis (just as linear forms do).

Tensors are dual objects (even when they're of order one), as in a person and his image in the mirror.

This duality gives tensors a relative independence from the basis system : we define a set of changes of basis that correspond to a unique metric tensor, and we will be able to easily effect these changes on tensors.

This brings joy to the physicists, because it allows them to describe the laws of nature so that they do not change for a different spectator, a change that is reduced to a change of basis linked to a particular metric tensor, and that is fundamental.

# VII

---

## THE CONCEPT OF COVARIANCE IN PHYSICS

---

Physicists have two goals :

1. Defining quantities that do not change with a change of frame (inertial frames in special relativity for example - in this case they are said "Lorentz invariant")
2. Modeling the laws of physics in the form of relations (equations really) that do not change with a change of frame. These will be relations between tensors of the same type, or between invariant quantities. We say that these equations are "covariant".

Note : A common misuse is to label physical quantities as covariant (or not covariant). This generates a lot of confusion... Because if this quantity is a scalar or a tensor in Minkowski space (and whatever type of tensor, contravariant or covariant), then it will necessarily be Lorentz covariant! On the other hand, if it is not a tensor, or if it is a tensor from a different space (for example in three dimensions) then it would not be Lorentz covariant. It's as simple as that and doesn't need extra complications. This invariance or covariance is relative to a particular group of transformations that depicts the types of change of basis "authorized" (a set chose precisely so they would all have the same metric tensor and be in that way equivalent) :

- In special relativity, the equations have to stay the same with any change of inertial frames (frames in a straight and uniform movement), so they have to be covariant under transformations of the [Lorentz group](#). Same thing for quantum electrodynamics. For example, the

relation  $\vec{p} = m_0 \vec{v}$  describing momentum as the product of rest mass by the speed  $\vec{v} = \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \\ \frac{dz}{dt} \end{pmatrix}$  in

3 dimensions does not stay the same under Lorentz transformations ( $\vec{p}$  is not Lorentz covariant). Adding a time component is not enough to correct this (That's why special relativity is so much more than just the union of space and time...), one also needs to play with the proper time  $\tau$  introduced in special relativity, and we then have a new definition of the momentum, called

the four-momentum vector  $\mathbf{p} = m_0 \mathbf{u}$  where  $\mathbf{u} = \begin{pmatrix} c \frac{dt}{d\tau} \\ \frac{dx}{d\tau} \\ \frac{dy}{d\tau} \\ \frac{dz}{d\tau} \end{pmatrix}$  is the four-velocity vector. By the way,



it gives us  $\mathbf{p} = \begin{pmatrix} E/c \\ \vec{p} \end{pmatrix}$  linking momentum and energy. The relation  $\mathbf{p} = m_0\mathbf{u}$  is then Lorentz covariant.

Note that even if the relation stays the same, the tensors components will usually change with a change of frame (but always in the same way, keeping the relation intact). A "real" invariant would be its pseudo-norm which will have the same value in every frame. For example with the four-velocity  $\mathbf{u}^\mu\mathbf{u}_\mu = c^2$ , or for the four-momentum  $\mathbf{p}^\mu\mathbf{p}_\mu = \frac{E^2}{c^2} - \vec{p}^2 = m_0^2c^2$  and these equations are not only Lorentz covariant, they are also completely invariant (their values do not vary).

By the way, at rest ( $\vec{p} = 0$ ) we obtain the famous equation  $E = m_0c^2$ , and, that's precisely the point, in every reference frame.

- In general relativity, the equations have to stay the same under any change of frame (and not just inertial ones). These changes are described this time by the group of diffeomorphisms of space-time, and the best way to get generally covariant relations is to use tensor fields (But it is quite a bit more complex to generalize the covariance of equations in general relativity, because the derivation does not transform a tensor in another tensor, so one has to define a covariant derivation...).

# VIII

---

## ZOOLOGY : EXAMPLES OF TENSORS

---

A tensor is characterized by its order and valence :

The order corresponds to the number of indices. For each index, it needs as many numbers as there are dimensions, so a total of  $n^{order}$  numbers. When indices are not used, the tensor is underlined as many times as its order, for example  $T_i^j = \underline{T}$ .

The valence  $(h, k)$  gives us the number of contravariant indices  $h$  and the number of covariant indices  $k$ .

Note that in a change of basis, the valence gives us the number of times  $k$  we have to multiply by the change of basis matrix, and the number of times  $h$  we have to multiply by its inverse.

Tensors of order 0 : these are the scalars. Real numbers in a real space, or complex numbers or others depending on the field... They are "usual" numbers without indices, that do not depend on any basis system and then do not need any notion of covariance or contravariance.

Tensors of order 1 : these are the vectors ( $T^i$  with contravariant components) and covectors or linear forms ( $T_i$  with covariant components).

Tensors of order 2 : these are the matrices  $T_j^i$  (linear transformations) or tensors of valence  $(1, 1)$ , tensors of valence  $(2, 0)$  (bivectors, like the inverse metric tensor  $g^{ij}$ , the electromagnetic tensor, etc.), and finally the tensors of valence  $(0, 2)$  (bilinear forms or bi-covectors, like the scalar product via its associated bilinear form, the trace, the metric tensor  $g_{ij}$ , the Ricci tensor, symplectic forms, etc.).

We also find in this category numerous physical quantities : the magnetic field, magnetic flow, angular momentum, torque, angular velocity, angular acceleration, areal velocity, etc. Also called spinors, girators, axial vectors, pseudovectors... there is a vast terminology. Note : the term "quadrivector" or "four-vector" used by physicists does not correspond to a tensor of order two, but only to a vector in four-dimensional space.

Tensors of order 3 : these are tensors with three indices, written  $T^{ijk}, T_{ijk}, T_{jk}^i, T_k^{ij}$ .

# IX

---

## A WORD OF PHILOSOPHY

---

If you are surprised that this post is part of the "About numbers" series, then you probably make a clear distinction between the concept of number (natural, integer, rational, real, complex, etc.) and that of more abstract mathematical objects. I do not. I won't discuss it in length, but let me say a few words.

First, every concept of number can be described as a vector in a particular space... Complex numbers are an example (more on that coming in a future post), but real numbers are another (in a trivial one-dimensional space), integers and rational numbers too (in a lattice), etc.

Also, vectors, matrices and tensors' components are numbers, so they can be seen as multi-dimensional numbers in a way (even if it does not tell the whole story).

So the concept of number takes a very, very broader sense in mathematics than intuition usually tells us... But, a table of number is not a number! I hear you say.

To that I'll just answer :  $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$  ,  $\begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$  ,  $\begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}$  , etc...

;-)

Feedback greatly appreciated, thanks!

Take 10 seconds to answer [3 questions online](#)