

Giancarlo Bernacchi

TENSORS

made easy

An informal introduction
to Maths of General Relativity

To Sophie, Alexandre and Elena

2017

4th edition: March 2017
© Giancarlo Bernacchi

All rights reserved

Giancarlo Bernacchi
Rho (Milano) - Italia
giancarlobernacchi@libero.it

*Text edited by means of
Sun Microsystem OpenOffice Writer 3
(except some figures)*

*Edition December 2016
Revision March 2017*

Contents

Introduction	3
Notations and conventions	5
1 Vectors and Covectors	7
1.1 Vectors	7
1.2 Basis-vectors and components	7
1.3 Covectors	9
1.4 Scalar product (heterogeneous)	11
1.5 The Kronecker δ	12
1.6 Metaphors and models	13
1.7 The “T-mosaic” model	14
2 Tensors	19
2.1 Outer product between vectors and covectors	21
2.2 Matrix representation of tensors	23
2.3 Sum of tensors and product by a number	25
2.4 Symmetry and skew-symmetry	25
2.5 Representing tensors in T-mosaic	26
2.6 Tensors in T-mosaic model: definitions	27
2.7 Tensor inner product	29
2.8 Outer product in T-mosaic	37
2.9 Contraction	38
2.10 Inner product as outer product + contraction	40
2.11 Multiple connection of tensors	40
2.12 “Scalar product” or “identity” tensor	41
2.13 Inverse tensor	42
2.14 Vector-covector “dual switch” tensor	45
2.15 Vectors / covectors homogeneous scalar product	46
2.16 \mathbf{G} applied to basis-vectors	49
2.17 \mathbf{G} applied to a tensor	49
2.18 Relations between \mathbf{I} , \mathbf{G} , δ	51
3 Change of basis	53
3.1 Basis change in T-mosaic	56
3.2 Invariance of the null tensor	59
3.3 Invariance of tensor equations	60
4 Tensors in manifolds	61
4.1 Coordinate systems	63
4.2 Coordinate lines and surfaces	63
4.3 Coordinate bases	64
4.4 Coordinate bases and non-coordinate bases	67
4.5 Change of the coordinate system	69
6 Contravariant and covariant tensors	71

4.7	Affine tensors	72
4.8	Cartesian tensors	73
4.9	Magnitude of vectors	73
4.10	Distance and metric tensor	73
4.11	Euclidean distance	74
4.12	Generalized distances	74
4.13	Tensors and not –	76
4.14	Covariant derivative	76
4.15	The gradient $\hat{\nabla}$ at work	80
4.16	Gradient of some fundamental tensors	86
4.17	Covariant derivative and index raising / lowering	86
4.18	Christoffel symbols	87
4.19	Covariant derivative and invariance of tensor equations	89
4.20	T-mosaic representation of gradient, divergence and covariant derivative	90
5	Curved manifolds	92
5.1	Symptoms of curvature	92
5.2	Derivative of a scalar or vector along a line	94
5.3	T-mosaic representation of derivatives along a line	98
5.4	Parallel transport along a line of a vector	99
5.5	Geodesics	100
5.6	Positive and negative curvature	102
5.7	Flat and curved manifold	104
5.8	Flat local system	106
5.9	Local flatness theorem	106
5.10	Riemann tensor	111
5.11	Symmetries of tensor \mathbf{R}	116
5.12	Bianchi identity	117
5.13	Ricci tensor and Ricci scalar	117
5.14	Einstein tensor	118
	Appendix	121
	Bibliographic references	125

Introduction

Tensor Analysis has a particular, though not exclusive, interest for the theory of General Relativity, in which curved spaces have a central role; so we cannot restrict ourselves to Cartesian tensors and to usual 3D space. In fact, all texts of General Relativity include some treatment of Tensor Analysis at various levels, but often it reduces to a schematic summary that is almost incomprehensible to the first approach, or else the topic is distributed along the text to be introduced piecemeal when needed, inevitably disorganized and fragmented. On the other hand, works devoted to the tensors at elementary level are virtually missing, neither is worth embarking on a preliminary study of specialized texts: even if one might overcome the mathematical difficulties, he would not be on the shortest way leading to relativity (but perhaps to differential geometry or surroundings). What we ultimately need is a simple introduction, not too formal – maybe heuristic – suitable to be used as a first approach to the subject, before dealing with the study of specific texts of General Relativity. We will cite as references those articles or handouts that we found closer to this approach and to which we have referred in writing these notes.

Our goal is not to do something better, but more easily affordable; we payed attention to the fundamentals and tried to make clear the line of reasoning, avoiding, as it often happens in such circumstances, too many implicit assumptions (which usually are not so obvious to the first reading).

As a general approach, we fully agree (and we do it with the faith of converts) on the appropriateness of the “geometric approach” to Tensor Analysis: really, it does not require more effort, but gives a strength to the concept of tensor that the old traditional approach “by components”, still adopted by many texts, cannot give.

Addressing the various topics we have adopted a “top-down” strategy. We are convinced that it is always better, even didactically, address the problems in their generality (not Cartesian tensors as introduction to general tensors, not flat spaces before curved ones); to shrink, to go into details, etc. there will always be time, but it would be better a later time. After all, there is a strategy better than others to open Chinese boxes: to begin with the larger one. This should somehow contain the irresistible temptation to unduly transfer in a general context results that apply only in particular instances.

We also tried to apply a little-known corollary of “Occam's razor”: do not introduce restrictive assumptions before necessary (that is why, for example, we introduce a “dual switch” before the metric tensor). The desire not to appear pedestrian is the factor that increases by an order of magnitude the illegibility of mathematics books: we felt that it is better to let it go and accept the risk of being judged somewhat dull by those who presume to know the story in depth.

This book does not claim to originality, but aims to didactics. With an exception: the model (or metaphor) that represents the tensor as a piece or “tessera” of a mosaic (we'll call it “T-mosaic”). We are not aware that this metaphor, useful to write without embarrassment tensor equations and even a bit obvious, has ever been presented in a text. That might be even surprising. But the magicians never reveal their tricks, and mathematicians sometimes resemble them. If it will be appreciated, we will feel honored to have been just us to uncover the trick.

To read these notes should be enough some differential calculus (until partial

derivatives); for what concerns matrices, it suffices to know that they are tables with rows and columns (and that swapping them you can create a great confusion), or slightly more. Instead, we will assiduously use the Einstein sum convention for repeated indexes as a great simplification in writing tensor equations.

As Euclid said to the Pharaoh, there is no royal road to mathematics; this does not mean that it is always compulsory to follow the most impervious pathway.

The hope of the author of these notes is they could be useful to someone as a conceptual introduction to the subject.

G.B.

Notations and conventions

In these notes we'll use the standard notation for components of tensors, namely upper (apices) and lower (subscript) indexes in Greek letters, for example T_{β}^{α} , as usual in Relativity. The coordinates will be marked by upper indexes, such as x^{μ} , basis-vectors will be represented by \vec{e}_{α} , basis-covectors by \tilde{e}^{α} .

In a tensor formula such as $P_{\alpha} = g_{\alpha\beta} V^{\beta}$ the index α that appears once at left side member and once at right side member is a “free” index, while β that occurs twice in the right member only is a “dummy” index.

We will constantly use the Einstein sum convention, whereby a *repeated index means an implicit summation over that index*. For example, $P_{\alpha} = g_{\alpha\beta} V^{\beta}$ is for

$$P_{\alpha} = g_{\alpha 1} V^1 + g_{\alpha 2} V^2 + \dots + g_{\alpha n} V^n. \text{ Further examples are: } A^{\alpha} B_{\alpha} = \sum_{\alpha=1}^n A^{\alpha} B_{\alpha},$$

$$A_y^{\alpha\mu\nu} B_{\alpha\nu}^{\kappa} = \sum_{\alpha=1}^n \sum_{\nu=1}^n A_y^{\alpha\mu\nu} B_{\alpha\nu}^{\kappa} \text{ and } A_{\alpha}^{\alpha} = A_1^1 + A_2^2 + \dots + A_n^n.$$

We note that, due to the sum convention, the “chain rule” for partial derivatives can

be simply written $\frac{\partial f}{\partial x^{\alpha}} = \frac{\partial f}{\partial x^{\mu}} \frac{\partial x^{\mu}}{\partial x^{\alpha}}$ and the sum over μ comes automatically.

A dummy index, unlike a free index, does not survive the summation and thus it does not appear in the result. Its name can be freely changed as long as it does not collide with other homonymous indexes in the same term.

In all equations *the dummy indexes must always be balanced up-down* and they cannot occur more than twice in each term. The free indexes appear only once in each term. In an equation both left and right members must have the same free indexes.

These conventions make it much easier writing correct relationships.

The little needed of matrix algebra will be said when necessary. We only note that the multiplication of two matrices requires us to “devise” a dummy index. Thus, for instance, the product of matrices $[A_{\alpha\beta}]$ and $[B^{\alpha\beta}]$ becomes $[A_{\alpha\mu}] \cdot [B^{\mu\beta}]$ (also for matrices we locate indexes up or down depending on the tensors they represent, without a particular meaning for the matrix itself).

The mark \bullet will be used indiscriminately for scalar products between vectors and covectors, both heterogeneous and homogeneous, as well as for inner tensor products.

Other notations will be explained when introduced: there is some redundancy in notations and we will use them with easiness according to the convenience. In fact, we had better getting familiar with all various alternative notations that can be found in the literature.

The indented paragraphs marked \square are “inserts” in the thread of the speech and they can, if desired, be skipped at first reading (they are usually justifications or proofs).

“Mnemo” boxes suggest simple rules to remind complex formulas.

1 Vectors and covectors

1.1 Vectors

A set where we can define operations of addition between any two elements and multiplication of an element by a number (such that the result is still an element of the set) is called a *vector space* and its elements *vectors*. *

The usual vectors of physics are oriented quantities that fall under this definition; we will denote them generically by \vec{V} .

In particular, we will consider the vector space formed by the set of vectors defined *at a certain point P*.

1.2 Basis-vectors and components

The maximum number n of vectors independent to one another that we can put together is the *dimension* of a vector space. These n vectors, chosen arbitrarily, form a ***basis of vectors*** (it doesn't matter they are unit or equal length vectors).

We denote the basis-vectors by $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ (the generic basis-vector by \vec{e}_α , $\alpha = 1, 2, 3, \dots, n$ and the basis as a whole by $\{\vec{e}_\alpha\}$).

Any further vector can be expressed as a linear combination of the n basis-vectors. In other words, any vector can be written as an expansion (or decomposition) on the basis:

$$\vec{V} = \vec{e}_1 V^1 + \vec{e}_2 V^2 + \dots + \vec{e}_n V^n \quad 1.1$$

that, using Einstein's sum convention, we write:

$$\vec{V} = \vec{e}_\alpha V^\alpha \quad 1.2$$

The n numbers V^α ($\alpha = 1, 2, \dots, n$) are the ***components of the vector*** \vec{V} on the basis we have chosen ($\forall V^\alpha \in \mathbb{R}$).

This way of representing the vector \vec{V} is simply its “recipe”: $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ (i.e. the \vec{e}_α) is the list of the ingredients, and V^1, V^2, \dots, V^n (i.e. the V^α) are the quantities of each one.

The choice of the n basis-vectors is arbitrary. Obviously, the components V^α of a given vector \vec{V} change with the chosen basis (unlike the vector itself!)

* In fact, it must be defined $\vec{C} = a\vec{A} + b\vec{B}$, where $\vec{A}, \vec{B}, \vec{C}$ are elements of the same set and a, b numbers $\in \mathbb{R}$. We gloss over other requests more obvious.

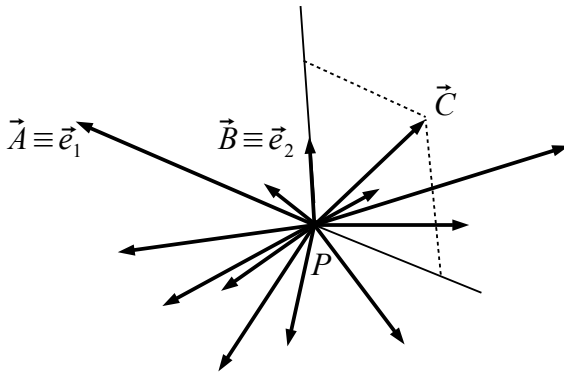
The expansion *eq.1.1* allows us to represent a vector as an n -tuple:

$$(V^1, V^2, \dots V^n)$$

provided there is a basis $\vec{e}_1, \vec{e}_2, \dots \vec{e}_n$ fixed in advance.

As an alternative to *eq.1.2* we will also write $\vec{V} \xrightarrow{\text{comp}} V^\alpha$ with the same meaning.

▪ Let's represent graphically the vector-space of plane vectors ($n = 2$) emanating from P , drawing only some vectors among the infinite ones:



Once selected $\vec{A} \equiv \vec{e}_1$, $\vec{B} \equiv \vec{e}_2$ as basis-vectors, a third vector \vec{C} is no more independent, but it can be expressed as a linear combination of the first two; in this case:

$$\vec{C} = -\frac{1}{2}\vec{e}_1 + 2\vec{e}_2$$

The components of \vec{C} on the established basis are thus $(-\frac{1}{2}, 2)$.

▪ We observe that the expansion (or “recipe”) of a basis-vector on its own basis (i.e. the basis made by themselves) cannot be other than:

$$\begin{aligned} \vec{e}_1 &\rightarrow (1, 0, 0, \dots 0) \\ \vec{e}_2 &\rightarrow (0, 1, 0, \dots 0) \\ &\dots \dots \dots \\ \vec{e}_n &\rightarrow (0, 0, 0, \dots 1) \end{aligned}$$

regardless whether they are unit basis-vector or not and their length.

1.3 Covectors

We define covector \tilde{P} (or dual vector or “one-form”) a linear scalar function of the vector \vec{V} . In other words: \tilde{P} applied to a vector results in a number:

$$\tilde{P}(\vec{V}) = \text{number} \in \mathbb{R} \quad 1.3$$

\tilde{P} can thus be seen as an “operator” that for any vector taken in input gives a number as output.

By what rule? One of the possible rules has a particular interest because it establishes a sort of reciprocity or *duality* between the vectors \vec{V} and covectors \tilde{P} , and it is what we aim to formalize.

▪ To begin with, we apply the covector \tilde{P} to a basis-vector (instead of a generic vector). By definition, let's call the result the α -th component of \tilde{P} :

$$\tilde{P}(\vec{e}_\alpha) = P_\alpha \quad 1.4$$

We will then say that the n numbers P_α ($\alpha = 1, 2, \dots, n$) are the components of the covector \tilde{P} (as well as the V^α were the components of the vector \vec{V}).

In operational terms we can state a rule (which will be generalized later on):

Apply the covector \tilde{P} to the basis-vectors \vec{e}_α to get the components of the covector \tilde{P} itself.

Note that this definition of component sets up a precise relationship between vectors and covectors.

Due to that definition, also a covector can be represented as an n -tuple

$$(P_1, P_2, \dots, P_n)$$

associated with a basis $\vec{e}^1, \vec{e}^2, \dots, \vec{e}^n$ of covectors.

Even a covector can then be expressed as an expansion on its own basis:

$$\tilde{P} = \vec{e}^\alpha P_\alpha \quad (\alpha = 1, 2, \dots, n) \quad 1.5$$

using the components P_α as coefficients of the expansion.

By itself, the choice of the basis of covectors is arbitrary, but at this point, having already fixed both the covector \tilde{P} and (by means of eq.1.4) its components, the basis of covectors $\{\vec{e}^\alpha\}$ follows, so that

the last equation (eq.1.5) can be true. In short, once used the vector basis $\{\vec{e}_\alpha\}$ to define the components of the covector, the choice of the covector basis $\{\tilde{e}^\alpha\}$ is forced.

▪ Before giving a mathematical form to this link between the two bases, we observe that, by using the definition eq.1.4 given above, the rule according to which \tilde{P} acts on the generic vector \vec{V} can be specified:

$$\tilde{P}(\vec{V}) = \tilde{P}(\vec{e}_\alpha V^\alpha) = V^\alpha \tilde{P}(\vec{e}_\alpha) = V^\alpha P_\alpha \quad 1.6$$

Applying \tilde{P} to \vec{V} means multiplying orderly by couples the components of both and sum up.

In this way the meaning of \tilde{P} is completely defined.

Also on the set of covectors \tilde{P} it is now possible to define operations of sum of covectors and multiplication of a covector by a number. This confers also to the set of covectors the status of a vector space.*

We have thus defined two separate vector spaces, one for the vectors, the other for the covectors (both related to a point P), equal in dimension and in dual relationship between each other.

▪ We can now clarify how the duality relationship between the two vector spaces links the bases $\{\vec{e}_\alpha\}$ and $\{\tilde{e}^\alpha\}$.

Provided $\vec{V} = \vec{e}_\alpha V^\alpha$ and $\tilde{P} = \tilde{e}^\beta P_\beta$ we can write:

$$\tilde{P}(\vec{V}) = \tilde{e}^\beta P_\beta (V^\alpha \vec{e}_\alpha) = P_\beta V^\alpha \tilde{e}^\beta(\vec{e}_\alpha)$$

but, on the other hand, (eq.1.6):

$$\tilde{P}(\vec{V}) = P_\alpha V^\alpha$$

Both expansions are identical only if:

$$\tilde{e}^\beta(\vec{e}_\alpha) = \begin{cases} 1 & \text{for } \beta=\alpha \\ 0 & \text{otherwise} \end{cases}$$

By defining the *Kronecker symbol* as: $\delta_\alpha^\beta = \begin{cases} 1 & \text{for } \beta=\alpha \\ 0 & \text{otherwise} \end{cases} \quad 1.7$

* From the definition given for $\tilde{P}(\vec{V})$ it follows that: i) the application \tilde{P} is linear: $\tilde{P}(a\vec{A}+b\vec{B}) = P_\alpha(aA^\alpha+bB^\alpha) = aP_\alpha A^\alpha+bP_\alpha B^\alpha = a\tilde{P}(\vec{A})+b\tilde{P}(\vec{B})$;
 ii) the set of the covectors \tilde{P} is a vector space, too: $(a\tilde{P}+b\tilde{Q})(\vec{V}) = a\tilde{P}(\vec{V})+b\tilde{Q}(\vec{V}) = aP_\alpha V^\alpha+bQ_\alpha V^\alpha = (aP_\alpha+bQ_\alpha)V^\alpha = R_\alpha V^\alpha = \tilde{R}(\vec{V})$

we can write the **duality condition**:

$$\tilde{e}^\beta(\tilde{e}_\alpha) = \delta_\alpha^\beta \quad 1.8$$

A vector space of vectors and a vector space of covectors are dual if and only if this relation between their bases holds good.

▪ We observe now that $\tilde{P}(\tilde{V}) = V^\alpha P_\alpha$ (eq.16) lends itself to an alternative interpretation.

Because of its symmetry, the product $V^\alpha P_\alpha$ can be interpreted not only as $\tilde{P}(\tilde{V})$ but also as $\tilde{V}(\tilde{P})$, interchanging operator and operand:

$$\tilde{V}(\tilde{P}) = V^\alpha P_\alpha = \text{number} \in \mathbb{R} \quad 1.9$$

In this “reversed” interpretation a vector can be regarded as a linear scalar function of a covector \tilde{P} . *

As a final step, if we want the duality to be complete, it would be possible to express the components V^α of a vector as the result of the application of \tilde{V} to the basis-covectors \tilde{e}^α (by symmetry with the definition of component of covector eq.1.4):

$$V^\alpha = \tilde{V}(\tilde{e}^\alpha) \quad 1.10$$

▫ It's easy to observe that this is the case, because:

$$\tilde{V}(\tilde{P}) = \tilde{V}(\tilde{e}^\alpha P_\alpha) = P_\alpha \tilde{V}(\tilde{e}^\alpha)$$

but since $\tilde{V}(\tilde{P}) = P_\alpha V^\alpha$ (eq.1.9), equating the right members the assumption follows.

Eq.1.10 is the dual of eq.1.4 and together they express a general rule:

▮ To get the components of a vector or covector apply the vector or covector to its *dual* basis.

1.4 Scalar product (heterogeneous)

The application of a covector to a vector or vice versa (the result does not change) so far indicated with notations like $\tilde{P}(\tilde{V})$ or $\tilde{V}(\tilde{P})$ has the meaning of a *scalar product* between the two ones. An alternative notation makes use of parentheses $\langle \dots \rangle$ emphasizing the symmetry of the two operands.

* It's easy to verify that the application \tilde{V} is linear:

$$\tilde{V}(a\tilde{P}+b\tilde{Q}) = V^\alpha(aP_\alpha+bQ_\alpha) = aV^\alpha P_\alpha + bV^\alpha Q_\alpha = a\tilde{V}(\tilde{P}) + b\tilde{V}(\tilde{Q})$$

The linearity of the operation is already known.

All writings:

$$\tilde{P}(\vec{V}) = \vec{V}(\tilde{P}) = \langle \tilde{P}, \vec{V} \rangle = \langle \vec{V}, \tilde{P} \rangle = V^\alpha P_\alpha \quad 1.11$$

are equivalent and represent the heterogeneous scalar product between a vector and a covector.

By the new introduced notation the duality condition between bases *eq.1.8* is currently expressed as:

$$\langle \tilde{e}^\beta, \vec{e}_\alpha \rangle = \delta_\alpha^\beta \quad 1.12$$

The homogeneous scalar product between vectors or covectors of the same kind requires a different definition that will be introduced later.

1.5 The Kronecker δ

Eq.1.7 defines the Kronecker δ : $\delta_\alpha^\beta = \begin{cases} 1 & \text{for } \beta=\alpha \\ 0 & \text{otherwise} \end{cases}$

A remarkable property of the Kronecker δ often used in the calculations is that δ_α^β acts as an operator that identifies the two indexes α, β turning one into the other.

For example: $\delta_\alpha^\beta V^\alpha = V^\beta$; $\delta_\alpha^\beta P_\beta = P_\alpha$ 1.13

Note that the summation that is implicit in the first member collapses to the single value of the second member. This happens because δ_α^β removes from the sum all terms whose indexes α, β are different, making them equal to zero.

Roughly speaking: δ_α^β “changes” one of the two indexes of its operand into the other one; what survives is the free index (note the “balance” of the indexes in both *eq.1.13*).

◦ We prove the first of *eq.1.13*:

Multiplying $\delta_\alpha^\beta = \tilde{e}^\beta(\vec{e}_\alpha)$ (*eq.1.8*) by V^α gives:

$$\delta_\alpha^\beta V^\alpha = \tilde{e}^\beta(\vec{e}_\alpha) V^\alpha = \tilde{e}^\beta(\vec{V}) = V^\beta$$

(the last equality of the chain is the rule *eq.1.10*).

Similarly for the other equation *eq.1.13*.

In practice, this property of Kronecker δ turns useful while calculating. Each time we can make a product like $\tilde{e}^\beta(\vec{e}_\alpha)$ to appear in an expression, we can replace it by δ_α^β , that soon produces a change of the index in one of the factors, as shown in *eq.1.13*.

So far we have considered the Kronecker symbol δ_α^β as a number; we will see later that δ_α^β is to be considered as a component of a tensor.

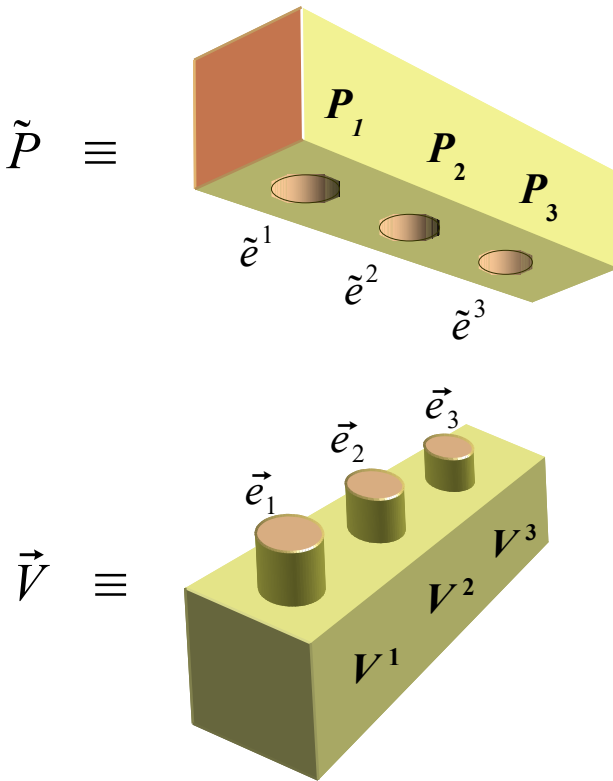
1.6 Metaphors and models

The fact that vectors and covectors can both be represented as a double array of basis elements and components

$$\begin{matrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 & \dots & \vec{e}_n \\ V^1 & V^2 & V^3 & \dots & V^n \end{matrix} \quad \text{and} \quad \begin{matrix} \tilde{e}^1 & \tilde{e}^2 & \tilde{e}^3 & \dots & \tilde{e}^n \\ P_1 & P_2 & P_3 & \dots & P_n \end{matrix}$$

suggests an useful metaphor to display graphically the formulas.

- Vectors and covectors are represented as interlocking cubes or “building blocks” bearing pins and holes that allow to hookup each other.



Vectors are cubes with pins upward; covectors are cubes with holes downward. We'll refer to pins and holes as “connectors”.

Pins and holes are in number of n , the dimension of space.

Each pin represents a \vec{e}_α ($\alpha = 1, 2, \dots, n$); each hole is for a \tilde{e}^α .

In correspondence with the various pins or holes we have to imagine the respective components written sideways on the body of the cube, as shown in the picture.

The example above refers to a 3D space (for $n = 4$, for instance, the representation would provide cubes with 4 pins or holes).

The n pins and the n holes must connect together simultaneously. Their connection emulates the heterogeneous scalar product between vectors and covectors and creates an object with no exposed connectors (a scalar). The heterogeneous scalar product

$$\left. \begin{array}{l} \tilde{e}^\alpha P_\alpha \\ \vec{e}_\alpha V^\alpha \end{array} \right\} \xrightarrow{\text{scalar product}} P_\alpha V^\alpha$$

may be indeed put in a pattern that may be interpreted as the metaphor of interlocking cubes:

$$\left. \begin{array}{l} \tilde{e}_1^\times + \tilde{e}_2^\times + \tilde{e}_3^\times \dots \tilde{e}_n^\times \\ P_1^\times \quad P_2^\times \quad P_3^\times \quad \dots \quad P_n^\times \\ \vec{e}_1^\times + \vec{e}_2^\times + \vec{e}_3^\times \dots \vec{e}_n^\times \\ V^1^\times + V^2^\times + V^3^\times \dots V^n^\times \end{array} \right\} \xrightarrow{\text{scalar product}} \begin{array}{l} P_1^\times + P_2^\times + P_3^\times \dots P_n^\times \\ V^1^\times \quad V^2^\times \quad V^3^\times \dots V^n^\times \end{array}$$

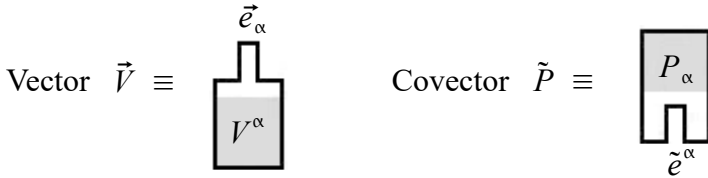
as shown in the previous picture, once we imagine to have made the fit between the two cubes.

1.7 The “T-mosaic” model

In the following we will adopt a more simple drawing without perspective, representing the cube as seen from the short side, like a piece of a mosaic (a “tessera”) in two dimensions with no depth, so that all the pins and holes will be aligned along the line of sight and merge into one which represents them collectively. The generic name of the array (\vec{e}_α for pins or \tilde{e}^α for holes) is thought as written upon

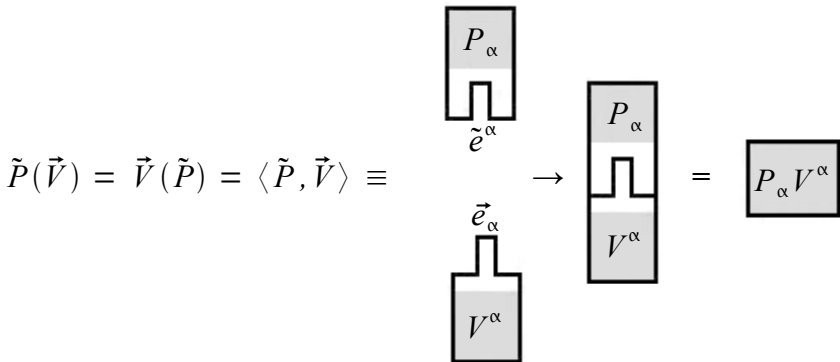
it, while the generic name of the component array (V^α for vectors or P_α for covectors) is written in the piece. This representation is the basis of the model that we call “**T-mosaic**” for its ability to be generalized to the case of tensors.

- So, from now on we will use the two-dimensional representation:

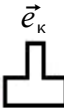



Blocks like those in the figure above give a synthetic representation of the expansion by components on the given basis (the “recipe” *eq.1.2*, *eq.1.5*).

- The connection of the blocks, i.e. the application of a block to another, represents the *heterogeneous scalar product*:

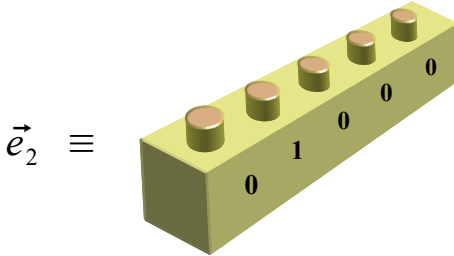


The connection is made between homonymous connectors (= with the same index: in this example \tilde{e}^α with \vec{e}_α). When the blocks fit together, the connectors disappear and bodies merge *multiplying to each other*.

- Basis-vector will be drawn as: 
- and basis-covector as: 

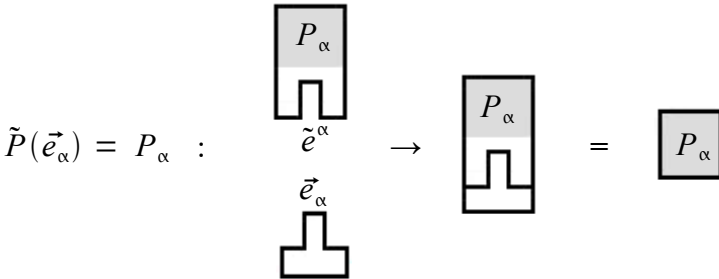
For simplicity, nothing will be written in the body of the tessera which represents a basis-vector or basis-covector, but we have to remember that, according to the perspective representation, a series of 0 together with a single 1 are inscribed in the side of the block.

Example in 5D:

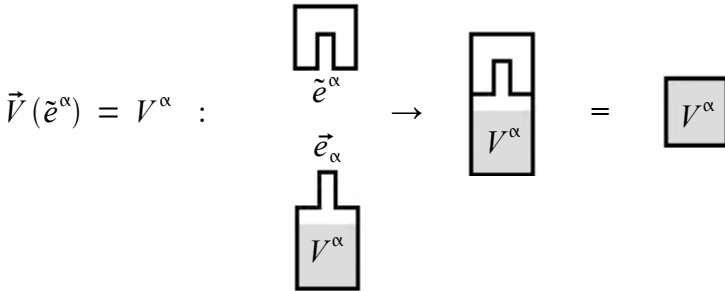


This means that, in the scalar product, all the products that enter in the sum go to zero, except one.

- The application of a covector to a basis-vector to get the covector components (eq.1.4) is represented by the connection:



Similarly, the application of a vector to a basis-covector to get the components of the vector (eq.1.10) corresponds to the connection:

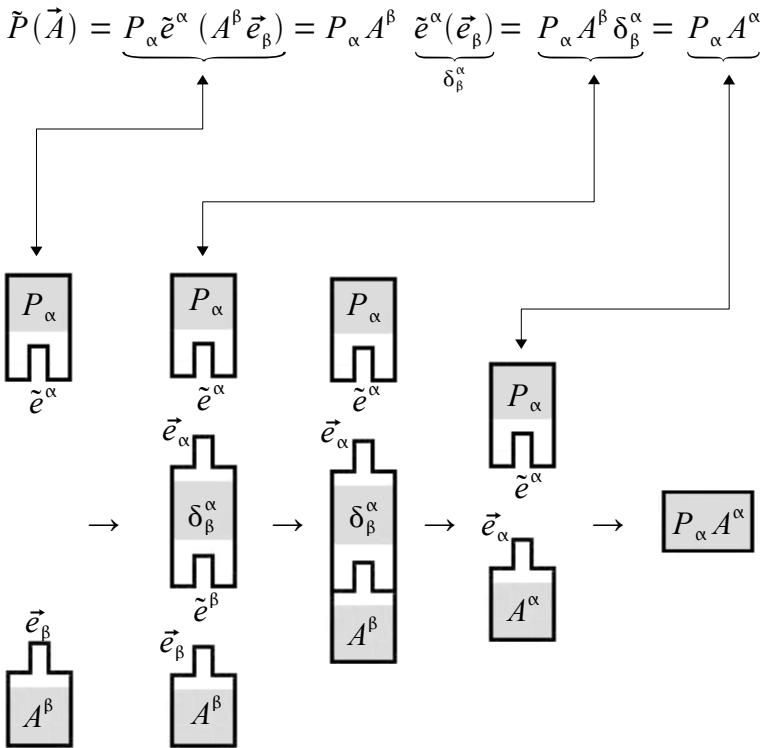


A “smooth” block, i.e. a block without free connectors, is a scalar (= a number).

▪ Notice that:

In T-mosaic representation the connection always occurs between connectors with the same index (same name), in contrast to what happens in algebraic formulas (where it is necessary to diversify them in order to avoid the summations interfere with each other).

▪ However, it is still possible to perform blockwise the connection between different indexes, in which case it is necessary to insert a block of Kronecker δ as a “plug adapter”:

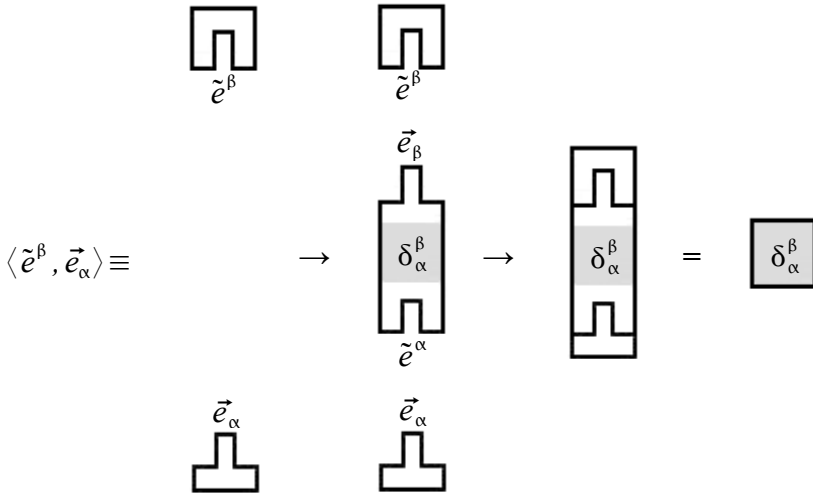


The chain of equalities above is the usual sequence of steps when using the conventional algebraic notation, which makes use of different indexes and the symbol Kronecker δ .

Note the correspondence between successive steps in algebraic formulas and blocks.

It is worth noting that in the usual T-mosaic representation the connection occurs directly between $P_\alpha \tilde{e}^\alpha$ and $A^\alpha \vec{e}_\alpha$ by means of the homonymic α -connectors, skipping the first 3 steps.

- The T-mosaic representation of the duality relation (eq.1.8 or eq.1.12) is a significant example of the use of Kronecker δ as a “plug adapter”:



In practice, it does not matter plugging directly connectors of the same name or connectors with different index via an interposed plug adapter; the first way is easier and faster, the second allows a better correspondence with algebraic formulas.

We will see later that using the Kronecker δ as a “plug adapter” block is justified by its tensor character.

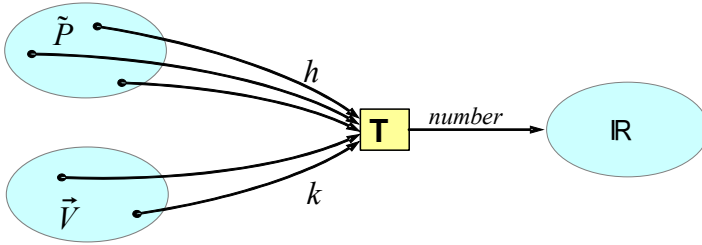
2 Tensors

The concept of tensor \mathbf{T} is an extension of those of vector and covector.

A tensor is a linear scalar function of h covectors and k vectors ($h, k = 0, 1, 2, \dots$).

We may see \mathbf{T} as an operator that takes h covectors and k vectors in input to give a number as a result:

$$\mathbf{T}(\vec{A}, \vec{B}, \dots, \vec{P}, \vec{Q}, \dots) = \text{number} \in \mathbb{R} \quad 2.1$$



By $\binom{h}{k}$ or even $r = h + k$ we denote the rank or order of the tensor.

A tensor $\binom{0}{0}$ is a scalar; a tensor $\binom{1}{0}$ is a vector; a tensor $\binom{0}{1}$ is a covector.

Warning: in the notation $\mathbf{T}(\ , \dots)$ the parenthesis following the symbol of tensor contains the list of arguments that the tensor takes as input (*input list*). It does not simply “qualify” the tensor, but it represents an operation already performed. In fact, $\mathbf{T}(\ , \dots)$ is the *result* of applying \mathbf{T} to the list within parentheses. The “naked tensor” is simply written \mathbf{T} .

▪ The *components of a tensor* \mathbf{T} are defined in a way similar to vectors and covectors by applying the tensor to basis-vectors and covectors. In simple words: we input into the tensor \mathbf{T} the required amount of basis-vectors and basis-covectors (h basis-covectors and k basis-vectors); the number that comes out is the component of the tensor on the given bases. For example, the components of a $\binom{0}{2}$ tensor result by giving to the tensor the various couples $\vec{e}_\alpha, \vec{e}_\beta$:

$$\mathbf{T}(\vec{e}_\alpha, \vec{e}_\beta) = T_{\alpha\beta} \quad 2.2$$

in order to obtain n^2 numbers marked by the double index $\alpha\beta$ (for instance, from \vec{e}_1, \vec{e}_2 we get T_{12} , and so on).

In general:

$$\mathbf{T}(\vec{e}_\alpha, \vec{e}_\beta, \dots, \tilde{e}^\mu, \tilde{e}^\nu, \dots) = T_{\alpha\beta\dots}^{\mu\nu\dots} \quad 2.3$$

▪ This equation allows us to specify the calculation rule left undetermined by *eq.2.1* (which number does come out by applying the tensor to its input list?). For example, given a tensor \mathbf{S} of rank $\binom{1}{1}$ for which *eq.2.3* becomes $\mathbf{S}(\vec{e}_\alpha, \tilde{e}^\gamma) = S_\alpha^\gamma$, we could get:

$$\mathbf{S}(\vec{V}, \tilde{P}) = \mathbf{S}(\vec{e}_\alpha V^\alpha, \tilde{e}^\gamma P_\gamma) = V^\alpha P_\gamma \mathbf{S}(\vec{e}_\alpha, \tilde{e}^\gamma) = V^\alpha P_\gamma S_\alpha^\gamma \quad 2.4$$

In general *eq.2.1* works as follows:

$$\mathbf{T}(\vec{A}, \vec{B}, \dots, \tilde{P}, \tilde{Q}, \dots) = A^\alpha B^\beta \dots P_\mu Q_\nu \dots T_{\alpha\beta\dots}^{\mu\nu\dots} \quad 2.5$$

and its result is a number (one for each set of values μ, ν, α, β).

An expression like $A^\alpha B^\beta P_\mu Q_\nu T_{\alpha\beta}^{\mu\nu}$ that contains only balanced dummy indexes is a scalar because no index survives in the result of the implicit multiple summation.

Eq.2.4, eq.2.5 appear as an extension to tensors of *eq.1.6, eq.1.9* valid for vectors and covectors. Applying the tensor \mathbf{T} to its input list $(, , \dots)$ is like the application $\vec{V}(\tilde{P})$ or $\tilde{P}(\vec{V})$ of a vector or covector to its respective argument, meaning heterogeneous scalar product. It may be seen as a sort of multiple scalar product, i.e. a sequence of scalar products between the tensor and the vectors / covectors of the list (indeed, it is a tensor “inner product” similar to the scalar product of vectors and covectors as we’ll explain later on).

▪ Speaking of tensor components we have so far referenced to bases of vectors and covectors. However, it is possible to express the tensor as an expansion by components on *its own basis*. That is to say, in the case of a $\binom{0}{2}$ tensor:

$$\mathbf{T} = \tilde{e}^{\alpha\beta} T_{\alpha\beta} \quad 2.6$$

(it’s the “recipe” of the tensor, similar to that of vectors). This time the basis is a tensorial one and consists of basis-tensors $\tilde{e}^{\alpha\beta}$ with double index, in number of n^2 .

This expression has only a formal meaning until we specify what sort of basis is that and how it is related to the basis of vectors / covectors defined before. To do this we need to define first the outer product between vectors and /or covectors.

2.1 Outer product between vectors and covectors

Given the vectors \vec{A}, \vec{B} and the covectors \vec{P}, \vec{Q} we define *vector outer product* between the vectors \vec{A} and \vec{B} :

$$\vec{A} \otimes \vec{B} \text{ such that: } \vec{A} \otimes \vec{B} (\vec{P}, \vec{Q}) = \vec{A}(\vec{P}) \vec{B}(\vec{Q}) \quad 2.7$$

Namely: $\vec{A} \otimes \vec{B}$ is an operator acting on a *couple of covectors* (i.e. vectors of the other kind) in terms of a scalar products, as stated in the right member. The result is here again a number $\in \mathbb{R}$.

It can be immediately seen that the outer product \otimes is *non-commutative*: $\vec{A} \otimes \vec{B} \neq \vec{B} \otimes \vec{A}$ (indeed, note that $\vec{B} \otimes \vec{A}$ against the same operand would give as a different result $\vec{B}(\vec{P}) \vec{A}(\vec{Q})$).

Also note that $\vec{A} \otimes \vec{B}$ is a rank $\binom{2}{0}$ tensor because it matches the given definition of tensor: it takes 2 covectors as input and gives a number as result.

Similarly we can define the outer products between covectors or between vectors and covectors, ranked $\binom{0}{2}$ and $\binom{1}{1}$ respectively:

$$\begin{aligned} \vec{P} \otimes \vec{Q} \text{ such that: } \vec{P} \otimes \vec{Q} (\vec{A}, \vec{B}) &= \vec{P}(\vec{A}) \vec{Q}(\vec{B}) \quad \text{and also} \\ \vec{P} \otimes \vec{A} \text{ such that: } \vec{P} \otimes \vec{A} (\vec{B}, \vec{Q}) &= \vec{P}(\vec{B}) \vec{A}(\vec{Q}) \quad \text{etc.} \end{aligned}$$

▪ Starting from vectors and covectors and making outer products between them we can build tensors of gradually increasing rank.

In general, the inverse is *not* true: not every tensor can be expressed as the outer product of tensors of lower rank.

▪ We can now characterize the tensor-basis in terms of outer product of basis-vectors and / or covectors. To fix on the case $\binom{0}{2}$, it is:

$$\tilde{e}^{\mu\nu} = \tilde{e}^\mu \otimes \tilde{e}^\nu \quad 2.8$$

▫ In fact, * from the definition of component $T_{\alpha\beta} = \mathbf{T}(\vec{e}_\alpha, \vec{e}_\beta)$

(eq.2.2), using the expansion $\mathbf{T} = T_{\mu\nu} \tilde{e}^{\mu\nu}$ (eq.2.6) we get:

$$T_{\alpha\beta} = T_{\mu\nu} \tilde{e}^{\mu\nu}(\vec{e}_\alpha, \vec{e}_\beta) \text{ , which is true only if:}$$

* The most direct demonstration, based on a comparison between the two forms

$\mathbf{T} = \vec{P} \otimes \vec{Q} = P_\alpha \tilde{e}^\alpha \otimes Q_\beta \tilde{e}^\beta = P_\alpha Q_\beta \tilde{e}^\alpha \otimes \tilde{e}^\beta = T_{\alpha\beta} \tilde{e}^\alpha \otimes \tilde{e}^\beta$ and $\mathbf{T} = T_{\alpha\beta} \tilde{e}^{\alpha\beta}$
holds only in the case of tensors decomposable as tensor outer product.

$\tilde{e}^{\mu\nu}(\vec{e}_\alpha, \vec{e}_\beta) = \delta_\alpha^\mu \delta_\beta^\nu$ because in that case:

$$T_{\alpha\beta} = T_{\mu\nu} \delta_\alpha^\mu \delta_\beta^\nu .$$

Since $\delta_\alpha^\mu = \tilde{e}^\mu(\vec{e}_\alpha)$ and $\delta_\beta^\nu = \tilde{e}^\nu(\vec{e}_\beta)$ the second-last equation becomes:

$$\tilde{e}^{\mu\nu}(\vec{e}_\alpha, \vec{e}_\beta) = \tilde{e}^\mu(\vec{e}_\alpha) \tilde{e}^\nu(\vec{e}_\beta)$$

which, by definition of \otimes (eq.2.7), is equivalent to:

$$\tilde{e}^{\mu\nu} = \tilde{e}^\mu \otimes \tilde{e}^\nu , \text{ q.e.d.}$$

The basis of tensors has been so reduced to the basis of (co)vectors. The $\binom{0}{2}$ tensor under consideration can then be expanded on the basis of covectors:

$$\mathbf{T} = T_{\alpha\beta} \tilde{e}^\alpha \otimes \tilde{e}^\beta \tag{2.9}$$

It's again the “recipe” of the tensor, as well as eq.2.6, but this time it uses basis-vectors and basis-covectors as “ingredients”.

▪ In general, a tensor can be expressed as a linear combination of (or as an expansion over the basis of) elementary outer products $\vec{e}_\alpha \otimes \vec{e}_\beta \otimes \dots \tilde{e}^\mu \otimes \tilde{e}^\nu \otimes \dots$ whose coefficients are the components.

For instance, a $\binom{3}{1}$ tensor can be expanded as:

$$\mathbf{T} = T_{\mu}^{\alpha\beta\gamma} \vec{e}_\alpha \otimes \vec{e}_\beta \otimes \vec{e}_\gamma \otimes \tilde{e}^\mu \tag{2.10}$$

which is usually simply written:

$$\mathbf{T} = T_{\mu}^{\alpha\beta\gamma} \vec{e}_\alpha \vec{e}_\beta \vec{e}_\gamma \tilde{e}^\mu \tag{2.11}$$

Note the “balance” of upper / lower indexes.

The symbol \otimes can be normally omitted without ambiguity.

▫ In fact, products $\vec{A}\vec{B}$ or $\vec{V}\vec{P}$ can be unambiguously interpreted as $\vec{A} \otimes \vec{B}$ or $\vec{V} \otimes \vec{P}$ because for other products other explicit symbols are used, such as $\vec{V}(\vec{P})$, $\langle \vec{V}, \vec{P} \rangle$ or $\vec{V} \cdot \vec{W}$ for scalar products, $\mathbf{T}(\dots)$ or even \bullet for inner tensor products.

▪ **The order of the indexes** is important and is stated by eq.2.10 or eq.2.11 which represent the tensor in terms of an outer product: it is understood that changing the order of the indexes means to change the order of factors in the outer product, in general non-commutative.

Thus, in general $T^{\alpha\beta\gamma} \neq T^{\beta\alpha\gamma} \neq T^{\gamma\alpha\beta} \dots$

To preserve the order of the indexes is often enough a notation with a double sequence for upper and lower indexes like $Y_{\beta\nu}^{\alpha\gamma\mu}$. However, this notation is ambiguous and turns out to be improper when indexes are raised / lowered. Actually, it would be convenient using a scanning with reserved columns like $Y_{\cdot|\beta|\cdot}^{\alpha|\cdot|\gamma|\cdot}$; it aligns in a single sequence upper and lower indexes and assigns to each index a specific place wherein it may move up and down without colliding with other indexes. To avoid any ambiguity we ought to use a notation such as $Y_{\cdot\beta\cdot\nu}^{\alpha\cdot\gamma\mu\cdot}$, where the dot \cdot is used to keep busy the column, or simply $Y_{\beta\nu}^{\alpha\gamma\mu}$ replacing the dot with a space.

2.2 Matrix representation of tensors

- A tensor of rank 0 (a scalar) is a number
 - a tensor of rank 1, that is $\binom{0}{1}$ or $\binom{1}{0}$, is an n -tuple of numbers
 - a tensor of rank 2, that is $\binom{2}{0}, \binom{1}{1}, \binom{0}{2}$, is a square matrix $n \times n$
 - a tensor of rank 3 is a “cubic lattice” of $n \times n \times n$ numbers, etc...
- (In each case the numbers are the components of the tensor).

A tensor of rank r can be thought of as an r -dimensional grid of numbers, or components (a single number or n -array or $n \times n$ matrix, etc.); in all cases we must think the component grid as associated with an underlying “bases grid” of the same dimension.

On the contrary, not every n -tuple of numbers or $n \times n$ matrix, and so on, is a vectors, tensor, etc.

Only tensors of rank 1 and 2, since represented as vectors or matrices, can be treated with the usual rules of matrix calculus (the interest to do that may be due to the fact that the inner product between tensors of these ranks is reduced to the product of matrices).

In particular, tensors $\binom{2}{0}$, $\binom{1}{1}$ or $\binom{0}{2}$ can all be represented by matrices, although built on different bases grids. For instance, for tensors of rank $\binom{2}{0}$ the basis grid is:

$$\begin{array}{cccc}
\vec{e}_1 \otimes \vec{e}_1 & \vec{e}_1 \otimes \vec{e}_2 & \cdots & \vec{e}_1 \otimes \vec{e}_n \\
\vec{e}_2 \otimes \vec{e}_1 & \vec{e}_2 \otimes \vec{e}_2 & & \vec{e}_2 \otimes \vec{e}_n \\
\vdots & & & \vdots \\
\vec{e}_n \otimes \vec{e}_1 & \vec{e}_n \otimes \vec{e}_2 & \cdots & \vec{e}_n \otimes \vec{e}_n
\end{array}$$

On this basis the matrix of the tensor is:

$$\mathbf{T} \equiv \begin{bmatrix} T^{11} & T^{12} & \cdots & T^{1n} \\ T^{21} & T^{22} & & T^{2n} \\ \vdots & & & \vdots \\ T^{n1} & T^{n2} & \cdots & T^{nn} \end{bmatrix} = [T^{\alpha\beta}] \quad 2.12$$

Similarly:

$$\text{for } \binom{1}{1} \text{ tensor: } \mathbf{T} \equiv [T_{\beta}^{\alpha}] \text{ on grid } \vec{e}_{\alpha} \otimes \vec{e}_{\beta}$$

$$\text{for } \binom{0}{2} \text{ tensor: } \mathbf{T} \equiv [T_{\alpha\beta}] \text{ on grid } \tilde{e}^{\alpha} \otimes \tilde{e}^{\beta}$$

▪ The outer product between vectors and tensors has similarities with the Cartesian product (= set of pairs).

For example $\vec{V} \otimes \vec{W}$ generates a basis grid $\vec{e}_{\alpha} \otimes \vec{e}_{\beta}$ that includes all the couples that can be formed by \vec{e}_{α} and \vec{e}_{β} and a corresponding matrix of components with all possible couples such as $V^{\alpha} W^{\beta}$ ($= T^{\alpha\beta}$) ordered by row and column.

If \mathbf{X} is a rank $\binom{2}{0}$ tensor, the outer product $\mathbf{X} \otimes \vec{V}$ produces a 3-dimensional cubic grid $\vec{e}_{\alpha} \otimes \vec{e}_{\beta} \otimes \vec{e}_{\gamma}$ of all the triplets built by \vec{e}_{α} , \vec{e}_{β} , \vec{e}_{γ} and a similar cubic structure of all possible couples $X^{\alpha\beta} T^{\gamma}$.

▪ It should be stressed that the dimension of the space on which the tensor spreads itself is the rank r of the tensor and has nothing to do with n , the dimension of geometric space. For example, a rank 2 tensor is a matrix (2-dimensional) in a space of any dimension; what varies is the number n of rows and columns.

Also the T-mosaic blockwise representation is invariant with the dimension of space: the number of connectors does not change since they are not individually represented, but as arrays.

2.3 Sum of tensors and product by a number

As in the case of vectors and covectors, the set of all tensors of a certain rank $\binom{h}{k}$ defined at a point P has the structure of a vector space, once defined operations of sum between tensors and multiplication of tensors by numbers.

The sum of tensors (of the same rank!) gives a tensor whose components are the sums of the components:

$$\mathbf{A} + \mathbf{B} = \mathbf{C} \Leftrightarrow C_{\mu\nu\dots}^{\alpha\beta\dots} = A_{\mu\nu\dots}^{\alpha\beta\dots} + B_{\mu\nu\dots}^{\alpha\beta\dots} \quad 2.13$$

The product of a number a by a tensor has the effect of multiplying by a all the components:

$$a \mathbf{A} \xrightarrow{\text{comp}} a A_{\mu\nu\dots}^{\alpha\beta\dots} \quad 2.14$$

2.4 Symmetry and skew-symmetry

A tensor is told *symmetric* with respect to a pair of indexes if their exchange leaves it unchanged; *skew-symmetric* if the exchange of the two indexes involves a change of sign (the two indexes must be *both upper or both lower indexes*).

For example: $T_{\alpha\beta\gamma\dots}$ is

- symmetric with respect to indexes α, γ if $T_{\gamma\beta\alpha\dots} = T_{\alpha\beta\gamma\dots}$
- skew-symmetric with respect to indexes β, γ if $T_{\alpha\gamma\beta\dots} = -T_{\alpha\beta\gamma\dots}$

Note that the symmetry has to do with the order of the arguments in the input list of the tensor. For example, taken a $\binom{0}{2}$ tensor:

$$\mathbf{T}(\vec{A}, \vec{B}) = A^\alpha B^\beta \mathbf{T}(\vec{e}_\alpha, \vec{e}_\beta) = A^\alpha B^\beta T_{\alpha\beta}$$

$$\mathbf{T}(\vec{B}, \vec{A}) = B^\beta A^\alpha \mathbf{T}(\vec{e}_\beta, \vec{e}_\alpha) = A^\alpha B^\beta T_{\beta\alpha}$$

\Rightarrow the order of the arguments in the list is not relevant if and only if the tensor is symmetric, since:

$$T_{\alpha\beta} = T_{\beta\alpha} \Leftrightarrow \mathbf{T}(\vec{A}, \vec{B}) = \mathbf{T}(\vec{B}, \vec{A}) \quad 2.15$$

• For tensors of rank $\binom{0}{2}$ or $\binom{2}{0}$, represented by a matrix, symmetry / skew-symmetry reflect into their matrices:

- symmetric tensor \Rightarrow symmetric matrix: $[T_{\alpha\beta}] = [T_{\beta\alpha}]$
- skew-symmetric tensor \Rightarrow skew-symmetric matrix $[T_{\alpha\beta}] = -[T_{\beta\alpha}]$

- Any tensor \mathbf{T} ranked $\binom{0}{2}$ or $\binom{2}{0}$ can always be decomposed into a symmetric part \mathbf{S} and skew-symmetric part \mathbf{A} :

$$\mathbf{T} = \mathbf{S} + \mathbf{A}$$

that is

$$T_{\alpha\beta} = S_{\alpha\beta} + A_{\alpha\beta} \quad \text{where:} \quad 2.16$$

$$S_{\alpha\beta} = \frac{1}{2}(T_{\alpha\beta} + T_{\beta\alpha}) \quad \text{and} \quad A_{\alpha\beta} = \frac{1}{2}(T_{\alpha\beta} - T_{\beta\alpha}) \quad 2.17$$

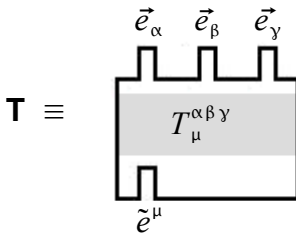
Tensors of rank greater than 2 can have more complex symmetries with respect to exchanges of 3 or more indexes, or even groups of indexes.

- The symmetries are intrinsic properties of the tensor and do not depend on the choice of bases: if a tensor has a certain symmetry in a basis, it keeps the same in other bases (as it will become clear later).

2.5 Representing tensors in T-mosaic

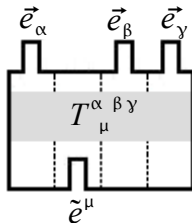
- Tensor \mathbf{T} "naked"**: it consists of a body that bears the indication of the components (in the shaded band); at edges the connectors:

- pins \vec{e}_α (in the top edge)
- holes \tilde{e}^μ (in the bottom edge)



$$\mathbf{T} = \underbrace{T_{\mu}^{\alpha\beta\gamma} \tilde{e}^\mu \vec{e}_\alpha \vec{e}_\beta \vec{e}_\gamma}_{\text{component \& connectors}}$$

- When necessary to make explicit the order of indexes by means of reserved columns a representation by compartments will be used, as:



$$\text{for } \mathbf{T} = T_{\mu}^{\alpha\beta\gamma} \vec{e}_\alpha \tilde{e}^\mu \vec{e}_\beta \vec{e}_\gamma$$

The exposed connectors correspond to free indexes; they determine the rank, given by $\binom{h}{k} = \binom{\text{number of pins } \vec{e}}{\text{number of holes } \tilde{e}}$ or even by $r = h + k$.

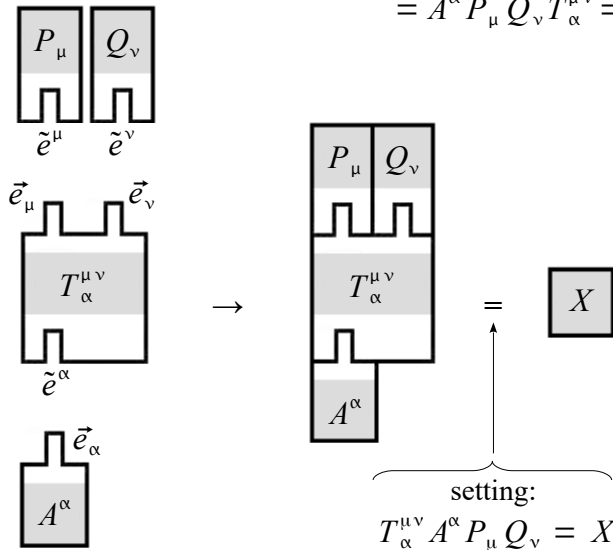
- Blocks representing tensors can connect to each other with the usual rule: pins or holes of a tensor can connect with holes or pins (with equal generic index) of vectors / covectors, or other tensors. The meaning of each single connection is similar to that of the heterogeneous scalar product of vectors and covectors: connectors disappear and bodies merge *multiplying to each other*.
- When (at least) one of the factors is a tensor ($r > 1$) we properly refer to it as an *inner tensor product*.
- The shapes of the blocks can be deformed for graphic reasons without changing the order of the connectors. It is convenient to keep fixed the orientation of the connectors (pins “on”, holes “down”).

2.6 Tensors in T-mosaic model: definitions

Previously stated definitions for tensors and tensor components have simple graphical representation in T-mosaic:

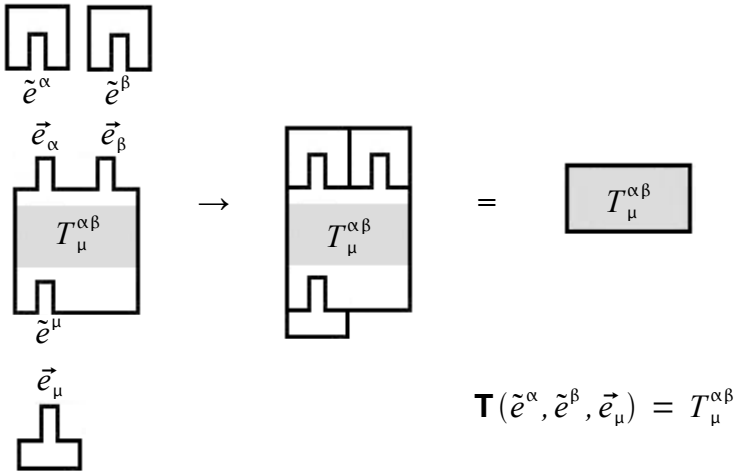
- **Definition of tensor:** plug all r connectors of the tensor with vectors or covectors:

$$\mathbf{T}(\vec{A}, \vec{P}, \vec{Q}) = \mathbf{T}(A^\alpha \vec{e}_\alpha, P_\mu \tilde{e}^\mu, Q_\nu \tilde{e}^\nu) = A^\alpha P_\mu Q_\nu \mathbf{T}(\vec{e}_\alpha, \tilde{e}^\mu, \tilde{e}^\nu) = A^\alpha P_\mu Q_\nu T_\alpha^{\mu\nu} = X$$



The result is a *single* number X (all the indexes are dummy).

- **Components of tensor \mathbf{T}** : obtained by saturating with basis-vectors and basis-covectors all connectors of the “naked” tensor. For example:

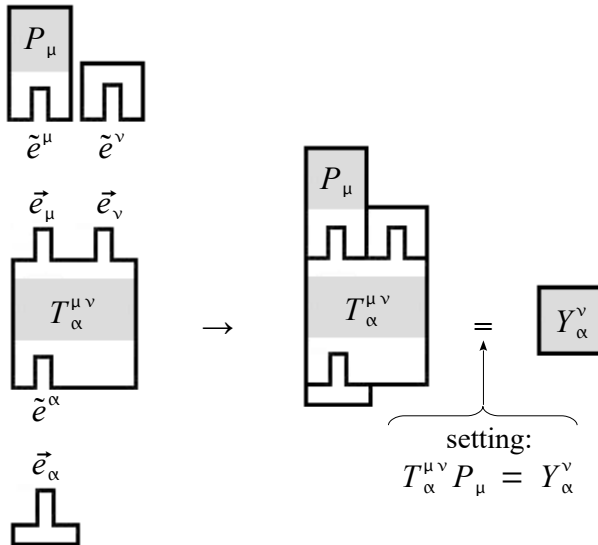


The result $T_{\mu}^{\alpha\beta}$ are n^3 numbers, the components (3 are the indexes).

Intermediate cases between the previous two can occur, too:

- **Input list with basis- and other vectors / covectors:**

$$\mathbf{T}(\vec{e}_{\alpha}, \tilde{P}, \tilde{e}^{\nu}) = \mathbf{T}(\vec{e}_{\alpha}, P_{\mu} \tilde{e}^{\mu}, \tilde{e}^{\nu}) = P_{\mu} \mathbf{T}(\vec{e}_{\alpha}, \tilde{e}^{\mu}, \tilde{e}^{\nu}) = P_{\mu} T_{\alpha}^{\mu\nu} = Y_{\alpha}^{\nu}$$



The result Y_α^ν are n^2 numbers (survived indexes are 2): they are the components of a double tensor $\mathbf{Y} = Y_\alpha^\nu \vec{e}_\nu \tilde{e}^\alpha$; but it is improper to say that the result is a tensor!

Remarks:

- The order of connection, stated by the input list, is important. Note that a different result $\mathbf{T}(\vec{e}_\alpha, \tilde{e}^\mu, \tilde{P}) = Z_\alpha^\mu \neq \mathbf{T}(\vec{e}_\alpha, \tilde{P}, \tilde{e}^\nu) = Y_\alpha^\nu$ would be obtained by connecting \tilde{P} with \tilde{e}^ν instead of \tilde{e}^μ .
- In general a “coated” or “saturated tensor”, that is a tensor with all connectors plugged (by vector, covectors or other) so as to be a “smooth object” is a number or a multiplicity of numbers.
- It is worth emphasizing the different meaning of notations that are similar only in appearance:

$$\left| \begin{array}{ll}
 \mathbf{T} = T_\mu^{\alpha\beta} \vec{e}_\alpha \vec{e}_\beta \tilde{e}^\mu & \text{is the tensor “naked”} \\
 \mathbf{T}(\underbrace{\tilde{e}^\alpha, \tilde{e}^\beta, \vec{e}_\mu}_{\substack{\uparrow \\ \text{Input list:} \\ \text{“T applied to ...”}}}) = T_\mu^{\alpha\beta} & \text{is the result of applying the tensor to the} \\
 & \text{list of basis-vectors (i.e. a component).}
 \end{array} \right.$$

- The input list may also be *incomplete* (i.e. it may contain a number of arguments $< r$, rank of the tensor), so that some connectors remains unplugged. The application of a tensor to an incomplete list cannot result in a number or a set of number, but is a tensor lowered in rank.

2.7 Tensor inner product

The *tensor inner product* consists in the connection of two tensors \mathbf{T} and \mathbf{Y} realized by means of a pair of connectors (indexes) of different kind belonging one to \mathbf{T} , the other to \mathbf{Y} . The connection is a *single* one and involves only one \vec{e} and only one \tilde{e} , but if the tensors are of high rank can be realized in more than one way.

To denote the tensor inner product we will use the symbol \bullet and we'll write $\mathbf{T} \bullet \mathbf{Y}$ to denote the product of the tensor \mathbf{T} on \mathbf{Y} , even if this writing can be ambiguous and requires further clarifications.

The two tensors that connect can be vectors, covectors or tensors of rank $r > 1$.

A particular case of tensor inner product is the familiar heterogeneous dot product between a vector and a covector.

Let's define the total rank R of the inner product as the sum of ranks of the two tensors involved: $R = r_1 + r_2$ (R is the total number of connectors of the tensors involved in the product). For the heterogeneous scalar (or dot) product is $R = 1 + 1 = 2$.

In a less trivial or strict sense the tensor inner product occurs between tensors of which at least one of rank $r > 1$, that is, $R > 2$. In any case, the tensor inner product lowers by 2 the rank R of the result.

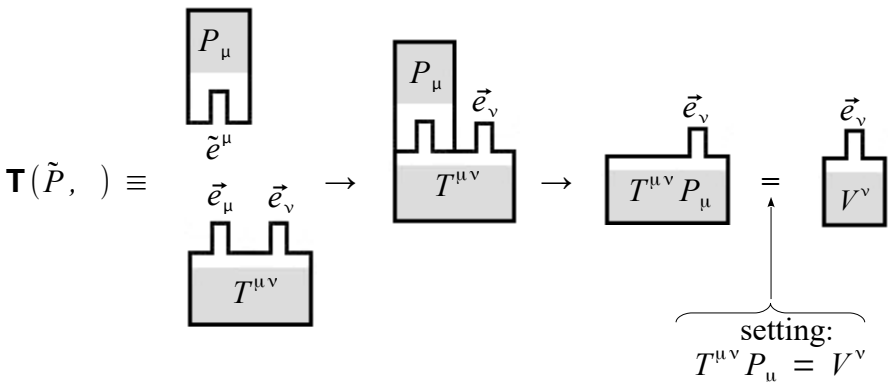
We will examine examples of tensor inner products of total rank R gradually increasing, detecting their properties and peculiarities.

▪ **Tensor inner product tensor • vector / covector**

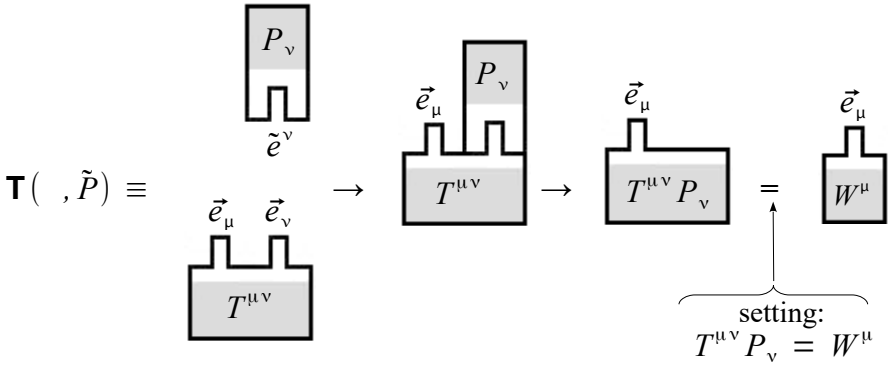
For the moment we limit ourselves to the case $\mathbf{T} \cdot \vec{V}$ or $\mathbf{T} \cdot \tilde{P}$ for which $R = 2 + 1 = 3$.

Let's exemplify the case $\mathbf{T} \cdot \tilde{P}$. We observe that making the tensor inner product of a rank $\binom{2}{0}$ tensor $\mathbf{T} = T^{\mu\nu} \tilde{e}_\mu \tilde{e}_\nu$ with a covector P means applying the former, instead of a complete list like $\mathbf{T}(\tilde{P}, \tilde{Q})$, to an incomplete list formed by the single element \tilde{P} .

We see immediately that this product $\mathbf{T} \cdot \tilde{P}$ can be made *in two ways*, depending on which one of the two connectors \tilde{e} of \mathbf{T} is involved, and the results are in general different; we will distinguish them by writing $\mathbf{T}(\tilde{P}, \quad)$ or $\mathbf{T}(\quad, \tilde{P})$ where the space is for the missing argument. The respective schemes are:



and:



Only one connector of the tensor is plugged and then the rank r of the tensor decrease by 1 (in the examples the result is a vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$).

Algebraic expressions corresponding to the two cases are:

$$\begin{aligned}
 \mathbf{T} \cdot \tilde{P} &= T^{\mu\nu} \vec{e}_\mu \vec{e}_\nu (P_\alpha \tilde{e}^\alpha) = T^{\mu\nu} P_\alpha \vec{e}_\mu \vec{e}_\nu (\tilde{e}^\alpha) = \\
 &= \begin{cases} T^{\mu\nu} P_\alpha \vec{e}_\mu \vec{e}_\nu (\tilde{e}^\alpha) = T^{\mu\nu} P_\alpha \vec{e}_\nu \delta_\mu^\alpha = T^{\mu\nu} P_\mu \vec{e}_\nu = V^\nu \vec{e}_\nu \\
 T^{\mu\nu} P_\alpha \vec{e}_\mu \vec{e}_\nu (\tilde{e}^\alpha) = T^{\mu\nu} P_\alpha \vec{e}_\mu \delta_\nu^\alpha = T^{\mu\nu} P_\nu \vec{e}_\mu = W^\mu \vec{e}_\mu \end{cases}
 \end{aligned}$$

The notation $\mathbf{T}(\tilde{P})$ or $\mathbf{T} \cdot \tilde{P}$ may designate both cases, but turns to be ambiguous because it does not specify the indexes involved in the inner product. Likewise ambiguous is the writing $\vec{e}_\mu \vec{e}_\nu (\tilde{e}^\alpha)$ in equations written above.

In fact, in algebra as well as in T-mosaic block connection, we need to know which indexes connect (via connectors' plugging or indirectly by Kronecker δ), i.e. on which indexes the inner product is performed. Only if \mathbf{T} is symmetrical with respect to μ, ν the result is the same and there is no ambiguity.

- In general, a writing like $\tilde{e}^\alpha \tilde{e}^\beta \vec{e}_\mu \vec{e}_\nu (\tilde{e}^\kappa)$ (which stands for $\tilde{e}^\alpha \otimes \tilde{e}^\beta \otimes \vec{e}_\mu \otimes \vec{e}_\nu (\tilde{e}^\kappa)$ with \otimes implied) has the meaning of an inner product of the covector \tilde{e}^κ by one element among those of the different kind aligned in the chain of outer products, for example:

$$\tilde{e}^\alpha \tilde{e}^\beta \vec{e}_\mu \vec{e}_\nu (\tilde{e}^\kappa) = \tilde{e}^\alpha \tilde{e}^\beta \vec{e}_\nu \delta_\mu^\kappa \quad \text{or} \quad \tilde{e}^\alpha \tilde{e}^\beta \vec{e}_\mu \vec{e}_\nu (\tilde{e}^\kappa) = \tilde{e}^\alpha \tilde{e}^\beta \vec{e}_\mu \delta_\nu^\kappa$$

The results are different depending on the element (the index) “hooked”, that must be known from the context (no matter the position of \tilde{e}^κ or \vec{e}_κ in the chain).

Likewise, when the internal product is made with a vector, as in $\tilde{e}^\alpha \tilde{e}^\beta \vec{e}_\mu \vec{e}_\nu (\vec{e}_\kappa)$ we have two chances:

$$\underbrace{\tilde{e}^\alpha \tilde{e}^\beta \vec{e}_\mu \vec{e}_\nu}_{\delta_\kappa^\alpha} (\vec{e}_\kappa) = \tilde{e}^\beta \vec{e}_\mu \vec{e}_\nu \delta_\kappa^\alpha \quad \text{or} \quad \tilde{e}^\alpha \tilde{e}^\beta \underbrace{\vec{e}_\mu \vec{e}_\nu}_{\delta_\kappa^\beta} (\vec{e}_\kappa) = \tilde{e}^\alpha \vec{e}_\mu \vec{e}_\nu \delta_\kappa^\beta$$

depending on whether the dot product goes to hook \tilde{e}^α or \tilde{e}^β .

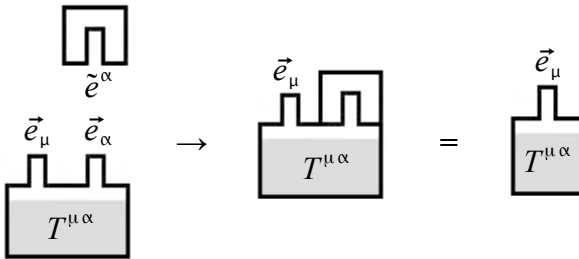
Formally, the inner product of a vector or covector \vec{e}_κ or \tilde{e}^κ addressed to a certain element (different in kind) inside a tensorial chain $\tilde{e}^\alpha \tilde{e}^\beta \vec{e}_\mu \vec{e}_\nu \dots$ removes the “hooked” element from the chain and insert a δ with the vanished indexes. The chain welds with a ring less, without changing the order of the remaining ones.

Note that product $\tilde{P} \bullet \mathbf{T}$ on fixed indexes cannot give a result different from $\mathbf{T} \bullet \tilde{P}$. Hence $\tilde{P} \bullet \mathbf{T} = \mathbf{T} \bullet \tilde{P}$ and the tensor inner product is *commutative* in this case.

• Same as above is the case of tensor inner product tensor • basis-covector. Here as well, the product can be realized in two ways: $\mathbf{T}(\tilde{e}^\mu)$ or $\mathbf{T}(\tilde{e}^\alpha)$. We carry out the latter case only, using the same \mathbf{T} of previous examples:

$$\mathbf{T}(\tilde{e}^\alpha) = \mathbf{T} \bullet \tilde{e}^\alpha = T^{\mu\nu} \vec{e}_\mu \vec{e}_\nu (\tilde{e}^\alpha) = T^{\mu\nu} \vec{e}_\mu \delta_\nu^\alpha = T^{\mu\alpha} \vec{e}_\mu$$

which we represent as:



$T^{\mu\alpha}$, the component of \mathbf{T} , would be the result if the input list had been complete. Having omitted a basis-covector in the input list has the effect of leaving uncovered a connector \vec{e} , so that the result is here a vector.

Similar considerations apply to inner products tensor \bullet vector where \mathbf{T} is ranked $\binom{0}{2}$.

- If \mathbf{T} is ranked $\binom{1}{1}$ the tensor inner products $\mathbf{T} \bullet \vec{V}$ and $\mathbf{T} \bullet \vec{P}$ have an unique (not ambiguous) meaning because in both cases there is only an obliged way to perform them (it is clear in T-mosaic).

Even here the products commute: for example $\mathbf{T} \bullet \vec{V} = \vec{V} \bullet \mathbf{T}$.

- In conclusion: inner tensor products with rank $R = 3$ may or may not be twofold, but they are anyways commutative.

▪ **Inner tensor product tensor \bullet tensor** ($R \geq 4$)

When the total rank R of the inner tensor product is $R \geq 4$ the multiplicity of possible products increases and further complications arise, due to the non-commutativity of the products and for what concerns the order of the indexes that survive (when they come from both tensors, how order them?). Let's specify better.

- Given two tensors \mathbf{A}, \mathbf{B} we main by:

- *multiplicity* the number of inner product possible upon various indexes (connectors) *after fixed the order of the factors* ($\mathbf{A} \bullet \mathbf{B}$ or $\mathbf{B} \bullet \mathbf{A}$).
- *commutativity* the eventuality that is $\mathbf{A} \bullet \mathbf{B} = \mathbf{B} \bullet \mathbf{A}$ *after fixed the pair of indexes* (connectors) *on which the products are performed*.

- Example of a case $R=2+2=4$ is the following, with $\mathbf{A} \binom{0}{2}$ and $\mathbf{B} \binom{2}{0}$.

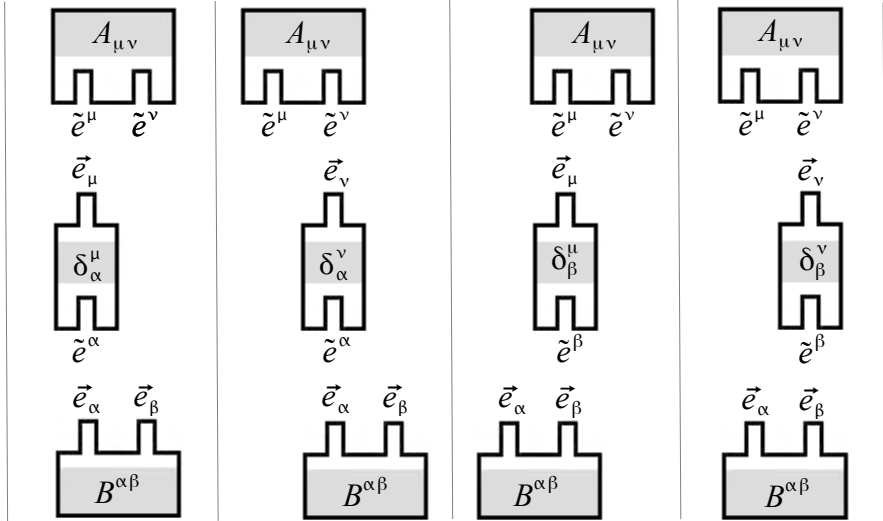
① **Multiplicity**

4 different tensor inner products $\mathbf{A} \bullet \mathbf{B}$ are possible, depending on which connectors plug (that is, which indexes are involved):

$$\mathbf{A} \bullet \mathbf{B} = (A_{\mu\nu} \tilde{e}^\mu \tilde{e}^\nu) \bullet (B^{\alpha\beta} \vec{e}_\alpha \vec{e}_\beta) =$$

$$= \begin{cases} \xrightarrow{\mu, \alpha} & = A_{\mu\nu} B^{\alpha\beta} \tilde{e}^\nu \vec{e}_\beta \delta_\alpha^\mu = A_{\alpha\nu} B^{\alpha\beta} \tilde{e}^\nu \vec{e}_\beta = C_\nu^\beta \tilde{e}^\nu \vec{e}_\beta \\ \xrightarrow{\nu, \alpha} & = A_{\mu\nu} B^{\alpha\beta} \tilde{e}^\mu \vec{e}_\beta \delta_\alpha^\nu = A^{\alpha\beta} B_{\mu\alpha} \tilde{e}^\mu \vec{e}_\beta = D_\mu^\beta \tilde{e}^\mu \vec{e}_\beta \\ \xrightarrow{\mu, \beta} & = A_{\mu\nu} B^{\alpha\beta} \tilde{e}^\nu \vec{e}_\alpha \delta_\beta^\mu = A^{\alpha\beta} B_{\nu\beta} \tilde{e}^\nu \vec{e}_\alpha = E_\nu^\alpha \tilde{e}^\nu \vec{e}_\alpha \\ \xrightarrow{\nu, \beta} & = A_{\mu\nu} B^{\alpha\beta} \tilde{e}^\mu \vec{e}_\alpha \delta_\beta^\nu = A^{\alpha\beta} B_{\mu\beta} \tilde{e}^\mu \vec{e}_\alpha = F_\mu^\alpha \tilde{e}^\mu \vec{e}_\alpha \end{cases}$$

which correspond respectively to the following T-mosaic patterns: *



• The 4 results of $\mathbf{A} \cdot \mathbf{B}$ are in general different. But it may happen that:

$$\mathbf{A} \text{ symmetrical } (\Rightarrow \tilde{e}^\mu, \tilde{e}^\nu \text{ interchangeable}) \Rightarrow \begin{cases} 1^\circ \text{ product} = 2^\circ \\ 3^\circ \text{ product} = 4^\circ \end{cases}$$

$$\mathbf{B} \text{ symmetrical } (\Rightarrow \vec{e}_\alpha, \vec{e}_\beta \text{ interchangeable}) \Rightarrow \begin{cases} 1^\circ \text{ product} = 3^\circ \\ 2^\circ \text{ product} = 4^\circ \end{cases}$$

The 4 products coincide only if both \mathbf{A} and \mathbf{B} are symmetrical.

② Non-commutativity and indexes' arrangement

$\mathbf{B} \cdot \mathbf{A}$ represent 4 more inner products, distinct from one another and different from the previous **; e.g. the first one is (product on α, μ):

$$\mathbf{B} \cdot \mathbf{A} = (B^{\alpha\beta} \vec{e}_\alpha \vec{e}_\beta) \cdot (A_{\mu\nu} \tilde{e}^\mu \tilde{e}^\nu) = C_\nu^\beta \vec{e}_\beta \tilde{e}^\nu .$$

Comparing with the first among the products above we see that the inner tensorial product on a given pair of indexes between tensors of rank $r > 1$ is in general non-commutative: $\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$.

* We use here the artifice of δ as “plug adapter” to keep the correspondence of indexes between equations and blocks.

** Of course, the 4 possible pairs of indexes (connectors) are the same as the previous case, but their order is inverted. Note that the multiplicities of $\mathbf{A} \cdot \mathbf{B}$ are counted separately from those of $\mathbf{B} \cdot \mathbf{A}$.

▪ How to account for the different results if T-puzzle graphical representation for $\mathbf{A} \cdot \mathbf{B}$ and $\mathbf{B} \cdot \mathbf{A}$ is unique? The distinction lies in a different reading of the result that takes into account the order of the items and give them the correct priority. In fact, the result must be constructed by the following rule: *

list first connectors (indices) of the first operand that survive free, then free connectors (indices) survived of the second one.
Reading of the connectors (indexes) is always from left to right.

Obviously indexes that make the connection don't appear in the result.

▪ If both \mathbf{A} and \mathbf{B} are symmetrical, then $\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}$, and in addition the 4+4 inner products achievable with \mathbf{A} and \mathbf{B} converge in one, as is easily seen in T-mosaic.

▫ As can be seen by T-mosaic, the complete variety of products for $R=4$ includes, in addition to the 4 cases just seen, the following:

• 7 cases of inner product $\vec{V} \cdot \mathbf{T}$ with $\mathbf{T} \binom{0}{3}, \binom{1}{2}, \binom{2}{1}$

• 7 cases of inner product $\tilde{P} \cdot \mathbf{T}$ with $\mathbf{T} \binom{3}{0}, \binom{2}{1}, \binom{1}{2}$

with multiplicity 3, 2 or 1, all commutative;

• 6 cases of tensor product $\mathbf{A} \cdot \mathbf{B}$ with $\mathbf{A} \binom{1}{1}$ and $\mathbf{B} \binom{2}{0}, \binom{1}{1}, \binom{0}{2}$

all with multiplicity 2, *non-commutative*,

giving a total of 24 distinct inner products (+ 24 commutated) for $R=4$ (including 4 cases considered before). The number of products grow very rapidly with R ; only the presence of symmetries in tensors reduces the number of different products.

▪ ***Incomplete input lists and tensor inner product***

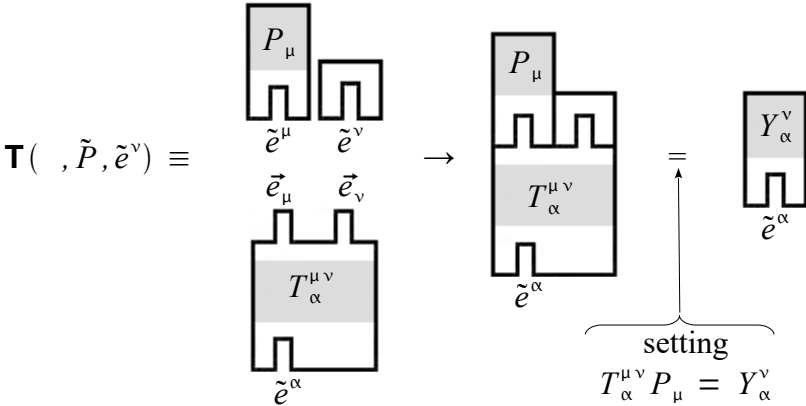
There is a partial overlap of the notions of inner product and the application of a tensor to an “input list” ** The complete input list is used to saturate all the connectors (indexes) of the tensor in order to give a number as a result; however, the list may be incomplete, and

* This rule is fully consistent with the usual interpretation of the inner tensor product as "external tensor product + contraction" which will be given later on.

**At least for what concerns the product “tensor • vector or covector” (input lists contain only vectors or covectors, not tensors).

then one or more connectors (indexes) residues of the tensor are found in the result, which turns out to be a tensor of lowered rank. The extreme case of “incomplete input list” is the tensor inner product tensor • vector or covector, where all the arguments of the list are missing, except one. Conversely the application of a tensor to a list (complete or not) can be seen as a succession of more than one single tensor inner products with vectors and covectors.

The higher the rank r of the tensor, the larger the number of possible incomplete input lists, i.e. of situations intermediate between the complete list and the single tensor inner product with vector or covector. For example, for $r = 3$ there are 6 possible cases of incomplete list. * In the case of a $\binom{2}{1}$ tensor one among them is:



The result is n covectors, one for each value of ν .

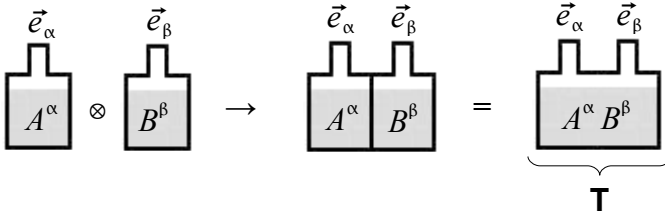
This scheme can be interpreted as a tensor inner product between the tensor T and the covector \tilde{P} , followed by a second tensor inner product between the result and the basis-covector \tilde{e}^ν , treated as such in algebraic terms:

$$\begin{aligned}
 i) \quad \mathbf{T}(\tilde{P}) &= T_\alpha^{\mu\nu} \tilde{e}_\mu^\alpha \tilde{e}_\nu^\beta \underbrace{(P_\kappa \tilde{e}^\kappa)}_{\delta_\mu^\kappa} = T_\alpha^{\mu\nu} P_\kappa \tilde{e}^\alpha \tilde{e}_\nu^\beta \delta_\mu^\kappa = T_\alpha^{\mu\nu} P_\mu \tilde{e}^\alpha \tilde{e}_\nu^\beta = Y_\alpha^\nu \tilde{e}^\alpha \tilde{e}_\nu^\beta \\
 ii) \quad Y_\alpha^\nu \tilde{e}^\alpha \tilde{e}_\nu^\beta (\tilde{e}^\gamma) &= Y_\alpha^\nu \tilde{e}^\alpha \delta_\nu^\gamma = Y_\alpha^\gamma \tilde{e}^\alpha
 \end{aligned}$$

* 3 lists lacking of one argument + 3 lists lacking of 2 arguments.

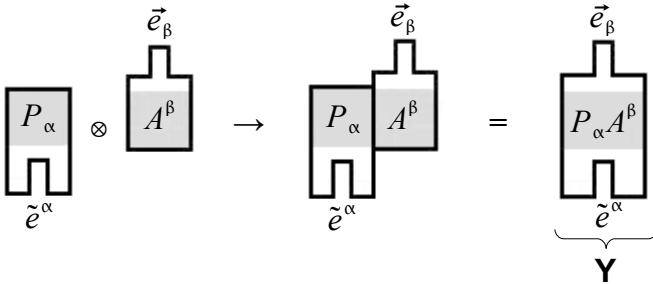
2.8 Outer product in T-mosaic

According to T-mosaic metaphor the outer product \otimes means “union by side gluing”. It may involve vectors, covectors or tensors. Of the more general case is part the outer product between vectors and/or covectors; for example:



$$\vec{A} \otimes \vec{B} = A^\alpha B^\beta \vec{e}_\alpha \otimes \vec{e}_\beta = A^\alpha B^\beta \vec{e}_\alpha \vec{e}_\beta = T^{\alpha\beta} \vec{e}_\alpha \vec{e}_\beta = \mathbf{T}$$

or else:



$$\vec{P} \otimes \vec{A} = P_\alpha A^\beta \vec{e}^\alpha \otimes \vec{e}_\beta = P_\alpha A^\beta \vec{e}^\alpha \vec{e}_\beta = Y_\alpha^\beta \vec{e}^\alpha \vec{e}_\beta = \mathbf{Y}$$

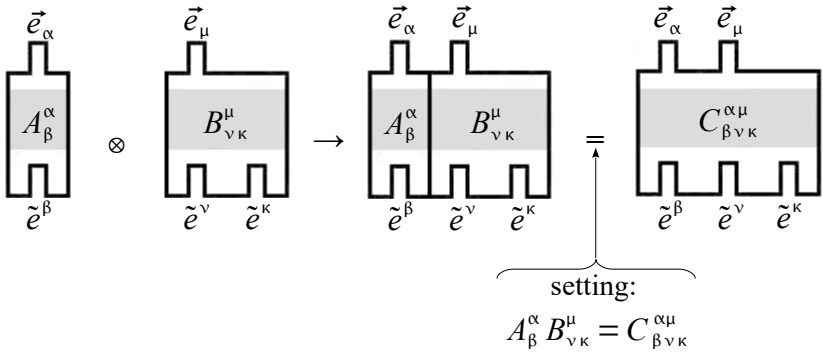
The outer product makes the bodies to blend together and components multiply (without the sum convention comes in operation).

It is clear from the T-mosaic representation that the symbol \otimes has a meaning similar to that of the conjunction “and” in a list of items and can usually be omitted without ambiguity.

- The same logic applies to tensors of rank $r > 1$, and one speaks about outer tensor product.

The tensor outer product operates on tensors of any rank, and merge them into one “for side gluing” The result is a composite tensor of rank equal to the sum of the ranks.

For example, in a case $\mathbf{A} \binom{1}{1} \otimes \mathbf{B} \binom{1}{2} = \mathbf{C} \binom{2}{3}$:



$$(A_{\beta}^{\alpha} \vec{e}_{\alpha} \otimes \tilde{e}^{\beta}) \otimes (B_{\nu\kappa}^{\mu} \vec{e}_{\mu} \otimes \tilde{e}^{\nu} \otimes \tilde{e}^{\kappa}) = A_{\beta}^{\alpha} B_{\nu\kappa}^{\mu} \vec{e}_{\alpha} \otimes \tilde{e}^{\beta} \otimes \vec{e}_{\mu} \otimes \tilde{e}^{\nu} \otimes \tilde{e}^{\kappa} = C_{\beta\nu\kappa}^{\alpha\mu} \vec{e}_{\alpha} \otimes \tilde{e}^{\beta} \otimes \vec{e}_{\mu} \otimes \tilde{e}^{\nu} \otimes \tilde{e}^{\kappa}$$

usually written $C_{\beta\nu\kappa}^{\alpha\mu} \vec{e}_{\alpha} \tilde{e}^{\beta} \vec{e}_{\mu} \tilde{e}^{\nu} \tilde{e}^{\kappa}$.

- Note the non-commutativity: $\mathbf{A} \otimes \mathbf{B} \neq \mathbf{B} \otimes \mathbf{A}$.

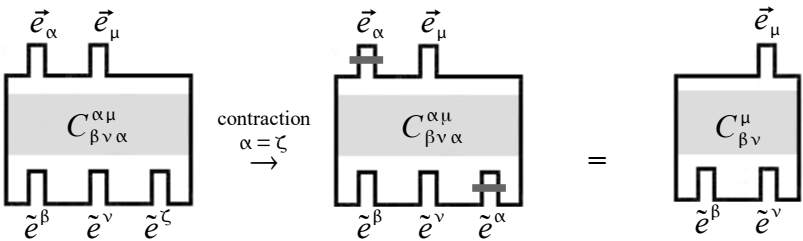
2.9 Contraction

Identifying two indexes of different kind (*one upper and one lower*), a tensor undergoes a *contraction* (the repeated index becomes dummy and appears no longer in the result). Contraction is an “unary” operation.

The tensor rank lowers from $\binom{h}{k}$ to $\binom{h-1}{k-1}$.

In T-mosaic metaphor both a pin and a hole belonging to the same tensor disappear.

Example: contraction for $\alpha = \zeta$:



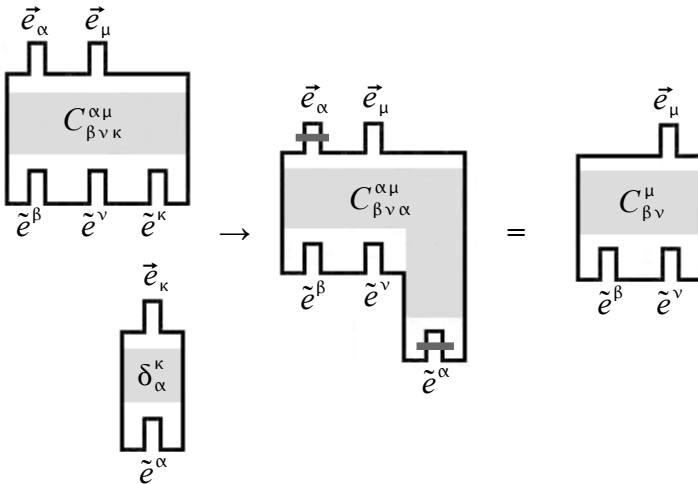
It is worth noting that the contraction is not a simple canceling or “simplification” of equal upper and lower indexes, but a sum which is triggered by repeated indexes. For example:

$$C_{\nu\mu}^{\mu} = C_{\nu 1}^1 + C_{\nu 2}^2 + \dots + C_{\nu n}^n = C_{\nu}$$

- In a sense, the contraction consists of an inner product that operates inside the tensor itself, plugging two connectors of different kind.
- The contraction may be seen as the result of an inner product by a tensor carrying a connector (an index) which is already present, with opposite kind, in the operand. Typically this tensor is the Kronecker δ . For example:

$$C_{\beta\nu\kappa}^{\alpha\mu} \delta_{\alpha}^{\kappa} = C_{\beta\nu\alpha}^{\alpha\mu} = C_{\beta\nu}^{\mu} \text{ where contraction is on index } \alpha.$$

The T-mosaic representation of the latter is the following:



- For subsequent repeated contractions a tensor $\binom{h}{k}$ can be reduced until $h = 0$ or $k = 0$. A tensor of rank $\binom{h}{h}$ can be reduced to a scalar.
- The contraction of a $\binom{1}{1}$ tensor gives a scalar, the sum of the main diagonal elements of its matrix, and is called the **trace**:

$$A_{\alpha}^{\alpha} = A_1^1 + A_2^2 + \dots + A_n^n \tag{2.18}$$

The trace of **I** is $\delta_{\alpha}^{\alpha} = 1 + 1 + \dots = n$, dimension of the manifold. *

* **I** is the tensor “identity” whose components are Kronecker- δ : see later.

2.10 Inner product as outer product + contraction

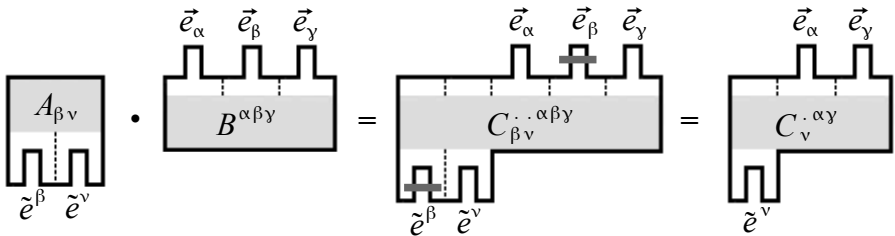
The inner tensor product $\mathbf{A} \cdot \mathbf{B}$ can be interpreted as *tensor outer product* $\mathbf{A} \otimes \mathbf{B}$ followed by a *contraction of indexes*:

$$\mathbf{A} \cdot \mathbf{B} = \text{contraction of } (\mathbf{A} \otimes \mathbf{B}).$$

The multiplicity of possible contractions of indexes gives an account of the multiplicity of inner tensor product represented by $\mathbf{A} \cdot \mathbf{B}$.

For example, in a case $\mathbf{A} \cdot \mathbf{B} = \mathbf{C} \binom{0}{2} \binom{3}{0} = \mathbf{C} \binom{2}{1}$, with total rank $R=5$:

$$\mathbf{A} \cdot \mathbf{B} = A_{\beta v} \cdot B^{\alpha \beta \gamma} = C_{\beta v}^{\dots \alpha \beta \gamma} = C_v^{\dots \alpha \gamma} \quad (\text{product on } \beta):$$



The 1st step (outer product) is univocal; the 2nd step (contraction) implies the choice of indexes (connectors) on which contraction must take place. Gradually choosing a connector after the other exhausts the choice of possibilities (multiplicity) of the inner product $\mathbf{A} \cdot \mathbf{B}$. The same considerations apply to the switched product $\mathbf{B} \cdot \mathbf{A}$.

▪ This 2-steps modality is equivalent to the rules stated for the writing of the result of the tensor inner product. Its utility stands out especially in complex cases such as:

$$A_{\dots \beta \cdot}^{\mu \nu \cdot \zeta} \cdot B_{\alpha \cdot \gamma}^{\cdot \beta \cdot} = C_{\dots \beta \cdot \alpha \cdot \gamma}^{\mu \nu \cdot \zeta \cdot \beta \cdot} = C_{\dots \alpha \gamma}^{\mu \nu \zeta \cdot \cdot} \quad 2.19$$

2.11 Multiple connection of tensors

In T-mosaic the tensor inner product is always performed by means of a simple connection between a pin \vec{e} of a tensor and a hole \tilde{e} of the other.

Multiple connections between two tensors, easy to fulfill in T-mosaic, can be algebraically described as an inner product followed by contractions in number of $m-1$, where m is the number of plugged connectors).

2.12 “Scalar product” or “identity” tensor

The operation of (heterogeneous) scalar product takes in input a vector and a covector to give a number \Rightarrow it is a tensor of rank $\binom{1}{1}$. We denote it by \mathbf{I} :

$$\mathbf{I}(\tilde{P}, \vec{V}) \equiv \langle \tilde{P}, \vec{V} \rangle = P_\alpha V^\alpha \quad 2.20$$

and can expand it as: $\mathbf{I} = I_\beta^\alpha \tilde{e}_\alpha \otimes \tilde{e}^\beta$, that is $\mathbf{I} = I_\beta^\alpha \tilde{e}_\alpha \tilde{e}^\beta$.

How is it \mathbf{I} ? Let us calculate its components giving to it the dual basis-vectors as input list:

$$I_\beta^\alpha = \mathbf{I}(\tilde{e}^\alpha, \vec{e}_\beta) = \langle \tilde{e}^\alpha, \vec{e}_\beta \rangle = \delta_\beta^\alpha \quad 2.21$$

$$\text{hence:} \quad \mathbf{I} = \delta_\beta^\alpha \tilde{e}^\beta \vec{e}_\alpha \quad 2.22$$

\Rightarrow the Kronecker δ symbol is a tensor:

$$\delta \equiv \mathbf{I} \equiv [\delta_\beta^\alpha] = \begin{bmatrix} \langle \tilde{e}^1, \vec{e}_1 \rangle & \langle \tilde{e}^1, \vec{e}_2 \rangle & \cdots & \langle \tilde{e}^1, \vec{e}_n \rangle \\ \langle \tilde{e}^2, \vec{e}_1 \rangle & \langle \tilde{e}^2, \vec{e}_2 \rangle & & \langle \tilde{e}^2, \vec{e}_n \rangle \\ \vdots & & & \vdots \\ \langle \tilde{e}^n, \vec{e}_1 \rangle & \langle \tilde{e}^n, \vec{e}_2 \rangle & \cdots & \langle \tilde{e}^n, \vec{e}_n \rangle \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots \\ 0 & 1 & \cdots \\ \cdots & \cdots & 1 \end{bmatrix} \quad 2.23$$

$$\text{namely:} \quad \delta \equiv \mathbf{I} \xrightarrow{\text{comp}} \delta_\beta^\alpha \quad 2.24$$

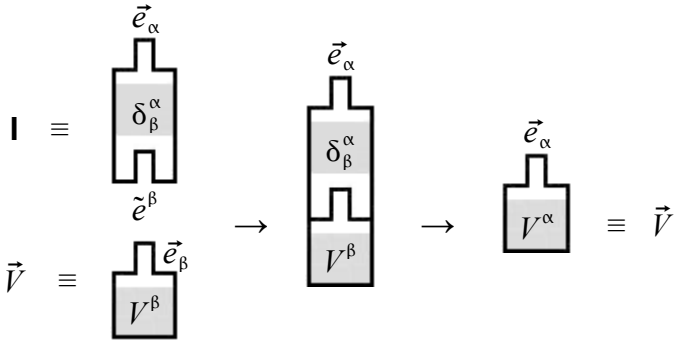
The heterogeneous scalar product is the tensor \mathbf{I} , identity tensor its components are the Kronecker's δ_β^α
its related matrix is the unit matrix $I = \text{diag}(+1)$

How does \mathbf{I} act? Its complete input list contains two arguments: a vector and a covector. Let's apply it to a partial list that contains only one of the two; there are 2 possibilities:

- ① $\mathbf{I}(\vec{V}) = \delta_\beta^\alpha \tilde{e}^\beta \vec{e}_\alpha (V^\gamma \vec{e}_\gamma) = \delta_\beta^\alpha V^\gamma \tilde{e}^\beta \vec{e}_\alpha (\vec{e}_\gamma) = \delta_\beta^\alpha \delta_\gamma^\beta V^\gamma \vec{e}_\alpha = \delta_\gamma^\alpha V^\gamma \vec{e}_\alpha = V^\alpha \vec{e}_\alpha = \vec{V}$
- ② $\mathbf{I}(\tilde{P}) = \delta_\beta^\alpha \tilde{e}^\beta \vec{e}_\alpha (P_\gamma \tilde{e}^\gamma) = \delta_\beta^\alpha P_\gamma \tilde{e}^\beta \vec{e}_\alpha (\tilde{e}^\gamma) = \delta_\beta^\alpha \delta_\alpha^\gamma P_\gamma \tilde{e}^\beta = \delta_\beta^\gamma P_\gamma \tilde{e}^\beta = P_\beta \tilde{e}^\beta = \tilde{P}$

(note: $\delta_\beta^\alpha \delta_\gamma^\beta = \delta_\gamma^\alpha$)

The T-mosaic blockwise representation of $\mathbf{I}(\vec{V})$ is:



Actually, \mathbf{I} transforms a vector into itself and a covector into itself:

$$\mathbf{I}(\vec{V}) = \vec{V} \quad ; \quad \mathbf{I}(\tilde{P}) = \tilde{P} \tag{2.25}$$

hence its name “identity tensor”.

The same tensor is also called “fundamental mixed tensor”.*

We also observe that, for $\forall \mathbf{T}$:

$$\mathbf{T}(\mathbf{I}) = \mathbf{I}(\mathbf{T}) = \mathbf{T} \tag{2.26}$$

2.10 Inverse tensor

Given a tensor \mathbf{T} , if there exists a *unique* tensor \mathbf{Y} such that $\mathbf{T} \cdot \mathbf{Y} = \mathbf{I}$, we say that \mathbf{Y} is the inverse of \mathbf{T} , i.e. $\mathbf{Y} = \mathbf{T}^{-1}$ and

$$\mathbf{T} \cdot \mathbf{T}^{-1} = \mathbf{I} \tag{2.27}$$

Only tensors \mathbf{T} of rank 2, and among them only $\binom{1}{1}$ type or *symmetric* $\binom{0}{2}$ or $\binom{2}{0}$ type tensors, can satisfy these conditions ** and have an inverse \mathbf{T}^{-1} .

* Hereafter the notation δ will be abandoned in favor of \mathbf{I} .

** That eq.2.27 is satisfied by tensors \mathbf{T} and \mathbf{T}^{-1} both ranked $r = 2$ can easily be shown in terms of T-mosaic blocks. If \mathbf{T} e \mathbf{T}^{-1} are tensors ranked $\binom{0}{2}$ or $\binom{2}{0}$, eq.2.27 stands for 4 different inner products and, for a given \mathbf{T} , can be satisfied for more than one \mathbf{T}^{-1} ; only if we ask that \mathbf{T} is symmetric ($\Rightarrow \mathbf{T}^{-1}$ simmetric, too) the tensor \mathbf{T}^{-1} that satisfies it is unique and the uniqueness of the inverse is guaranteed. For this reason one restricts the field to symmetric tensors only.

For tensors of rank $r \neq 2$ the inverse is not defined.

▪ The inverse \mathbf{T}^{-1} of a symmetric tensor \mathbf{T} of rank $\binom{0}{2}$ or $\binom{2}{0}$ has the following properties:

- the indexes position interchanges upper \leftrightarrow lower *
- it is represented by the inverse matrix
- it is symmetric

For example: given the $\binom{2}{0}$ symmetric tensor $\mathbf{T} \xrightarrow{comp} T^{\alpha\beta}$, its inverse is the $\binom{0}{2}$ tensor $\mathbf{T}^{-1} \xrightarrow{comp} T_{\mu\nu}$ such that:

$$T^{\alpha\gamma} T_{\gamma\nu} = \delta_{\nu}^{\alpha} \quad 2.28$$

Furthermore, denoting T and \bar{T} the matrices associated to \mathbf{T} and \mathbf{T}^{-1} , we have:

$$T \cdot \bar{T} = I \quad 2.29$$

where I is the unit matrix; namely, the matrices T and \bar{T} are inverse to each other.

▫ Indeed, for $\binom{2}{0}$ tensors the definition *eq.2.27* turns into *eq.2.28*.

But the latter, transcribed in matrix terms, takes the form $[T^{\alpha\gamma}] \cdot [\bar{T}_{\gamma\nu}] = I$, which is equivalent to *eq.2.29*.

Since the inverse of a symmetric matrix is symmetric as well, the symmetry of \mathbf{T}^{-1} follows: if a tensor is symmetric, then its inverse is symmetric, too.

- The correspondence between the inverse matrix and the inverse tensor forwards to the tensor other properties of the inverse matrix:
- the commutativity $T \cdot \bar{T} = \bar{T} \cdot T = I$, which applies to inverse matrices, is also true for inverse tensors: $\mathbf{T} \cdot \mathbf{T}^{-1} = \mathbf{T}^{-1} \cdot \mathbf{T} = \mathbf{I}$
- in order that \mathbf{T}^{-1} exists, the inverse matrix must also exist and for that it is required to be $\det T \neq 0$. **

Currently we say that tensors $T^{\alpha\beta}$ and $T_{\alpha\beta}$ are inverse to each

* This fact is often expressed by saying that the inverse of a double-contravariant tensor is a double-covariant tensor, and vice versa.

** A matrix can have an inverse only if its determinant is not zero.

other: usually we call the components of both tensors with the same symbol T and distinguish them only by the position of the indexes. However, it is worth realizing that we are dealing with different tensors, and they cannot run both under the same symbol \mathbf{T} (if we call \mathbf{T} one of the two, we must use another name for the other: in this instance \mathbf{T}^{-1}).

- It should also be reminded that (only) if $T^{\alpha\beta}$ is diagonal (i.e. $T^{\alpha\beta} \neq 0$ only for $\alpha = \beta$) its inverse will be diagonal as well, with components $T_{\alpha\beta} = 1/T^{\alpha\beta}$, and viceversa.
- The property of a tensor to have an inverse is intrinsic to the tensor itself and does not depend on the choice of bases: if a tensor has inverse in a basis, it has inverse in any other basis, too (as will become clear later on).
- An obvious property belongs to the mixed tensor T_{β}^{α} of rank $\binom{1}{1}$ “related” with both $T_{\alpha\beta}$ and $T^{\alpha\beta}$, defined by their inner product:

$$T_{\beta}^{\alpha} = T^{\alpha\gamma} T_{\gamma\beta}$$

Comparing this relation with *eq.2.28* (written as $T^{\alpha\gamma} T_{\gamma\beta} = \delta_{\beta}^{\alpha}$ with β in place of ν) we see that:

$$T_{\beta}^{\alpha} = \delta_{\beta}^{\alpha} \tag{2.30}$$

Indeed, the mixed fundamental tensor $\binom{1}{1} \delta_{\beta}^{\alpha}$, or tensor \mathbf{I} , is the “common relative” of all couples of inverse tensors $T_{\alpha\beta}, T^{\alpha\beta}$.)*

- The mixed fundamental tensor $\mathbf{I} \xrightarrow{comp} \delta_{\beta}^{\alpha}$ has (with few others **) the property to be the inverse of itself.

▫ Indeed, an already noticed *** property of Kronecker's δ :

$$\delta_{\beta}^{\alpha} \delta_{\gamma}^{\beta} = \delta_{\gamma}^{\alpha}$$

is the condition of inverse (similar to *eq.2.28*) for δ_{β}^{α} .
(T-mosaic icastically shows the meaning of this relation).

* That does not mean, of course, that it is the only existing mixed double tensor (think to $C_{\alpha}^{\beta} = A_{\alpha\gamma} B^{\gamma\beta}$ when $\mathbf{A} \in \mathbf{B}$ are not related to each other).

** Also auto-inverse are the tensors $\binom{1}{1}$, whose matrices are mirror images of \mathbf{I} .

***Already noticed about the calculation of $\mathbf{I}(\tilde{P})$.

2.14 Vector-covector “dual switch” tensor

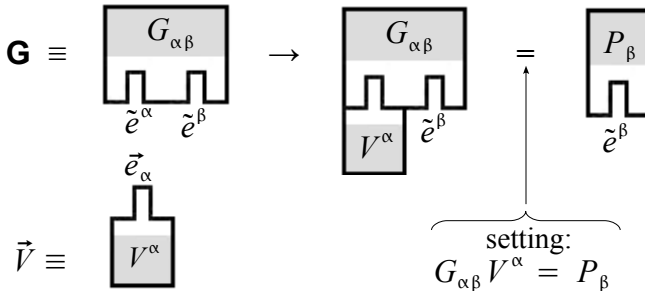
A tensor of rank $\binom{0}{2}$ needs two vectors as input to give a scalar. If the input list is incomplete and consists in one vector only, the result is a covector:

$$\mathbf{G}(\vec{V}) = G_{\alpha\beta} \tilde{e}^\alpha \tilde{e}^\beta (V^\gamma \vec{e}_\gamma) = G_{\alpha\beta} V^\gamma \tilde{e}^\alpha \tilde{e}^\beta \cdot \vec{e}_\gamma = G_{\alpha\beta} V^\alpha \tilde{e}^\beta = P_\beta \tilde{e}^\beta = \tilde{P}$$

$\xleftarrow{\delta_y^\alpha} \xrightarrow{\delta_y^\beta}$

having set $G_{\alpha\beta} V^\alpha = P_\beta$.

The representation in T-mosaic is:



to be read: $\mathbf{G}(\vec{V}) = G_{\alpha\beta} V^\alpha \tilde{e}^\beta = P_\beta \tilde{e}^\beta = \tilde{P}$

By means of a $\binom{0}{2}$ tensor we can then transform a vector \vec{V} into a covector \tilde{P} belonging to the dual space.

Let us pick out a $\binom{0}{2}$ tensor \mathbf{G} as a “dual switch” to be used from now on to transform any vector \vec{V} into its “dual” covector \tilde{V} :

$$\mathbf{G}(\vec{V}) = \tilde{V} \tag{2.29}$$

\mathbf{G} establishes a correspondence $\mathbf{G}: \vec{V} \rightarrow \tilde{V}$ between the two dual vector spaces (we use here the same name V with different marks above in order to emphasize the relationship and the term “dual” as “related through G in the dual space”).

The choice of \mathbf{G} is arbitrary, nevertheless we must choose a tensor which has inverse \mathbf{G}^{-1} in order to perform the “switching” in the opposite sense as well, from \tilde{V} to \vec{V} . In addition, if we want the inverse \mathbf{G}^{-1} to be unique, we must pick out a \mathbf{G} which is *symmetric*, as we know. Note that, in this manner, using one or the other of the two \tilde{e} connectors of \mathbf{G} becomes indifferent.

Applying \mathbf{G}^{-1} to the switch definition *eq.2.31*:

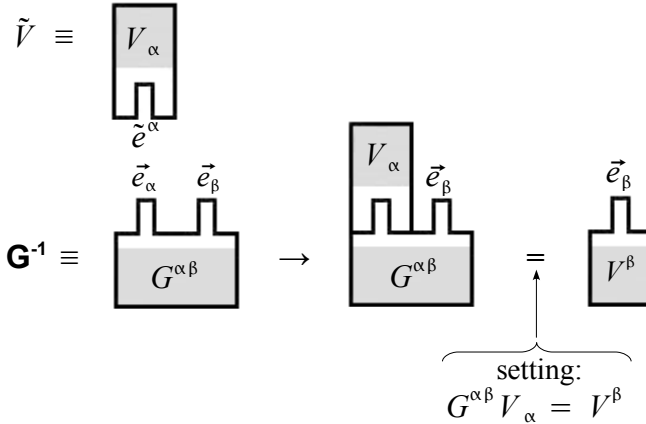
$$\mathbf{G}^{-1}(\mathbf{G}(\vec{V})) = \mathbf{G}^{-1}(\vec{V})$$

but $\mathbf{G}^{-1}(\mathbf{G}) = \mathbf{I}$ and $\mathbf{I}(\vec{V}) = \vec{V}$, then:

$$\vec{V} = \mathbf{G}^{-1}(\vec{V}) \tag{2.32}$$

\mathbf{G}^{-1} thus establishes an inverse correspondence $\mathbf{G}^{-1}: \vec{V} \rightarrow \vec{V}$.

In T-mosaic terms:



to be read $\mathbf{G}^{-1}(\vec{V}) = G^{\alpha\beta} V_\alpha \vec{e}_\beta = V^\beta \vec{e}_\beta = \vec{V}$

• The vector \leftrightarrow covector switching can be expressed componentwise. In terms of components $\mathbf{G}(\vec{V}) = \vec{V}$ is written as:

$$G_{\alpha\beta} V^\alpha = V_\beta \tag{2.33}$$

which is roughly interpreted: $G_{\alpha\beta}$ “*lowers the index*”.

Conversely, $\mathbf{G}^{-1}(\vec{V}) = \vec{V}$ can be written:

$$G^{\alpha\beta} V_\alpha = V^\beta \tag{2.34}$$

and interpreted as: $G^{\alpha\beta}$ “*raises the index*”.

2.15 Vectors / covectors homogeneous scalar product

It presupposes the notion of switch tensor \mathbf{G} .

We define the *homogeneous scalar product between two vectors* as the scalar product between one vector and the dual covector of the other:

$$\vec{A} \cdot \vec{B} \stackrel{\text{def}}{=} \langle \vec{A}, \vec{B} \rangle \quad \text{or, likewise,} \quad \stackrel{\text{def}}{=} \langle \vec{A}, \vec{B} \rangle \quad 2.35$$

From the first equality, expanding on the bases and using the switch \mathbf{G} we get:

$$\begin{aligned} \vec{A} \cdot \vec{B} &= \langle \vec{A}, \vec{B} \rangle = \langle A^\alpha \vec{e}_\alpha, B_\gamma \vec{e}^\gamma \rangle = \langle A^\alpha \vec{e}_\alpha, \underbrace{G_{\beta\gamma} B^\beta}_{B_\gamma} \vec{e}^\gamma \rangle = \\ &= G_{\beta\gamma} A^\alpha B^\beta \langle \vec{e}_\alpha, \vec{e}^\gamma \rangle = G_{\beta\gamma} A^\alpha B^\beta \delta_\alpha^\gamma = G_{\beta\gamma} A^\gamma B^\beta = \mathbf{G}(\vec{A}, \vec{B}) \end{aligned}$$

In short:
$$\vec{A} \cdot \vec{B} = \mathbf{G}(\vec{A}, \vec{B}) \quad 2.36$$

Hence, the dual switch tensor \mathbf{G} is also the “*scalar product between two vectors*” tensor.

The same result follows from the second equality eq.2.35.

The symmetry of \mathbf{G} guarantees the commutative property of the scalar product:

$$\vec{A} \cdot \vec{B} = \vec{B} \cdot \vec{A}$$

or
$$\mathbf{G}(\vec{A}, \vec{B}) = \mathbf{G}(\vec{B}, \vec{A}) \quad 2.37$$

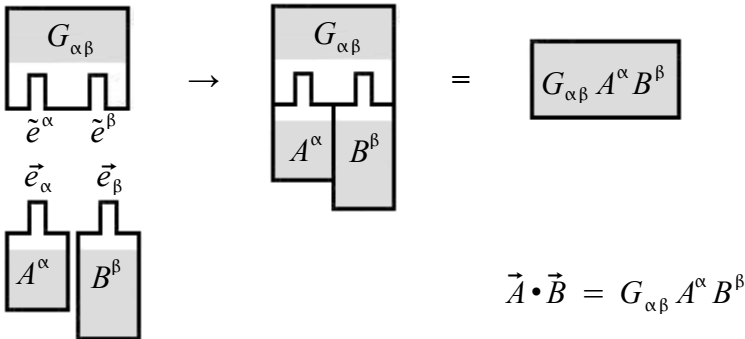
(note this is just the condition of symmetry for the tensor \mathbf{G} eq.2.15).

▪ The *homogeneous scalar product between two covectors* is then defined by means of \mathbf{G}^{-1} :

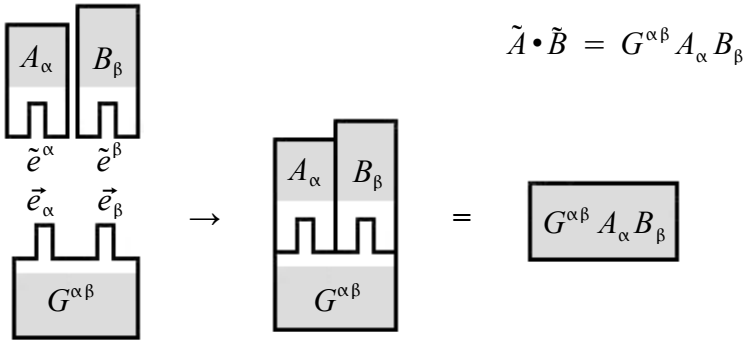
$$\tilde{A} \cdot \tilde{B} = \mathbf{G}^{-1}(\tilde{A}, \tilde{B}) \quad 2.38$$

\mathbf{G}^{-1} , the inverse dual switch, is thus the “*scalar product between two covector*” tensor.

▪ In the T-mosaic metaphor the scalar (or inner) product between two vectors takes the form:



while the scalar (or inner) product between two covectors is:



The result is a number in both cases.

▪ How is it **G**? Let's compute its components using as arguments the basis-vectors instead of \tilde{A}, \tilde{B} :

$$G_{\alpha\beta} = \mathbf{G}(\vec{e}_\alpha, \vec{e}_\beta) = \vec{e}_\alpha \cdot \vec{e}_\beta \quad \Rightarrow$$

$$\Rightarrow \mathbf{G} \equiv [G_{\alpha\beta}] = \begin{bmatrix} \vec{e}_1 \cdot \vec{e}_1 & \vec{e}_1 \cdot \vec{e}_2 & \dots & \vec{e}_1 \cdot \vec{e}_n \\ \vec{e}_2 \cdot \vec{e}_1 & \vec{e}_2 \cdot \vec{e}_2 & & \vec{e}_2 \cdot \vec{e}_n \\ \vdots & & & \vdots \\ \vec{e}_n \cdot \vec{e}_1 & \vec{e}_n \cdot \vec{e}_2 & \dots & \vec{e}_n \cdot \vec{e}_n \end{bmatrix} \quad 2.39$$

It's clear that the symmetry of **G** is related to the commutativity of the scalar product:

$$\vec{e}_\alpha \cdot \vec{e}_\beta = \vec{e}_\beta \cdot \vec{e}_\alpha \quad \Leftrightarrow \quad G_{\alpha\beta} = G_{\beta\alpha} \quad \Leftrightarrow \quad \mathbf{G} \text{ symmetric}$$

and that its associated matrix is also symmetric.

In a similar way for the inverse switch:

$$G^{\alpha\beta} = \mathbf{G}^{-1}(\tilde{e}^\alpha, \tilde{e}^\beta) = \tilde{e}^\alpha \cdot \tilde{e}^\beta \quad \Rightarrow$$

$$\Rightarrow \mathbf{G}^{-1} \equiv [G^{\alpha\beta}] = \begin{bmatrix} \tilde{e}^1 \cdot \tilde{e}^1 & \tilde{e}^1 \cdot \tilde{e}^2 & \dots & \tilde{e}^1 \cdot \tilde{e}^n \\ \tilde{e}^2 \cdot \tilde{e}^1 & \tilde{e}^2 \cdot \tilde{e}^2 & & \tilde{e}^2 \cdot \tilde{e}^n \\ \vdots & & & \vdots \\ \tilde{e}^n \cdot \tilde{e}^1 & \tilde{e}^n \cdot \tilde{e}^2 & \dots & \tilde{e}^n \cdot \tilde{e}^n \end{bmatrix} \quad 2.40$$

Notation

Various equivalent writings for the homogeneous scalar product are:

▪ between vectors:

$$\vec{A} \cdot \vec{B} = \vec{B} \cdot \vec{A} = \langle \vec{A}, \vec{B} \rangle = \langle \vec{A}, \vec{B} \rangle = \mathbf{G}(\vec{A}, \vec{B}) = \mathbf{G}(\vec{B}, \vec{A})$$

▪ between covectors:

$$\vec{A} \cdot \vec{B} = \vec{B} \cdot \vec{A} = \langle \vec{A}, \vec{B} \rangle = \langle \vec{A}, \vec{B} \rangle = \mathbf{G}^{-1}(\vec{A}, \vec{B}) = \mathbf{G}^{-1}(\vec{B}, \vec{A})$$

The notation $\langle \ , \ \rangle$ is reserved to heterogeneous scalar product vector-covector

2.16 **G** applied to basis-vectors

G transforms a basis-vector into a covector, but in general *not* into a basis-covector:

$$\mathbf{G}(\vec{e}_\alpha) = G_{\mu\nu} \tilde{e}^\mu \tilde{e}^\nu (\vec{e}_\alpha) = G_{\mu\nu} \tilde{e}^\mu \delta_\alpha^\nu = G_{\mu\alpha} \tilde{e}^\mu \quad 2.41$$

and this is not $= \tilde{e}^\alpha$ because $G_{\mu\alpha} \tilde{e}^\mu$ is a sum over all \tilde{e}^μ and can't collapse to a single value \tilde{e}^α (except particular cases). Likewise:

$$\mathbf{G}^{-1}(\tilde{e}^\alpha) = G^{\mu\nu} \vec{e}_\mu \vec{e}_\nu (\tilde{e}^\alpha) = G^{\mu\nu} \vec{e}_\mu \delta_\nu^\alpha = G^{\mu\alpha} \vec{e}_\mu \neq \vec{e}_\alpha \quad 2.42$$

- Notice that in *eq.2.41* we have made use of the duality condition (*eq.1.8*) that link the bases of vectors and covectors; it is a relation different from that stated by the “dual switching”.

2.17 **G** applied to a tensor

Can we assume that the converter **G** acts on the single index of the tensor (i.e. on the single connector \vec{e}_α) as if it were isolated, in the same way it would act on a vector, that is, changing it into a \tilde{e}^ν without changing other indices and their sequence? Not quite. (The same question concerns its inverse **G**⁻¹).

We see that the application of **G** to the tensor $\binom{3}{0} \mathbf{X} = X^{\alpha\beta\gamma} \vec{e}_\alpha \vec{e}_\beta \vec{e}_\gamma$:

$$G_{\alpha\nu} X^{\alpha\beta\gamma} = X^{\cdot\beta\gamma}_\nu$$

has, in this example, really the effect of lowering the index α involved in the product without altering the order of the remaining others. However, this is not the case for any index. What we need is to examine a series of cases more extensive than a single example, having in mind that the application of **G** to a tensor according to usual rules of the tensor inner product $\mathbf{G}(\mathbf{T}) = \mathbf{G} \cdot \mathbf{T}$ gives rise to a

multiplicity of different products (the same applies to \mathbf{G}^{-1}). To do so, let's think of the inner product as of the various possible contractions of the external product: in this interpretation it is up to the different contractions to produce multiplicity.

▫ Referring to the case, let's examine the other inner products you can do on the various indices, passing through the outer product $G_{\mu\nu} X^{\alpha\beta\gamma} = X_{\mu\nu}^{\dots\alpha\beta\gamma}$ and then performing the contractions. The μ - α contraction leads to the known result $X_{\alpha\nu}^{\dots\alpha\beta\gamma} = X_{\nu}^{\dots\beta\gamma}$; in addition:

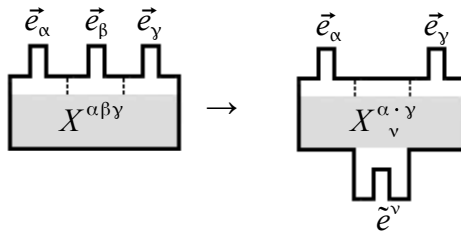
- contraction μ - β gives $X_{\beta\nu}^{\dots\alpha\beta\gamma} = X_{\nu}^{\dots\alpha\gamma}$
- contraction μ - γ gives $X_{\gamma\nu}^{\dots\alpha\beta\gamma} = X_{\nu}^{\dots\alpha\beta}$

(since \mathbf{G} is symmetric, contractions of ν with α , β or γ give the same results as the contractions of μ).

Other results rise by the application of \mathbf{G} in post-multiplication:

from $X^{\alpha\beta\gamma} G_{\mu\nu} = X^{\alpha\beta\gamma}_{\dots\mu\nu}$ we get $X^{\beta\gamma}_{\dots\nu}$, $X^{\alpha\gamma}_{\dots\nu}$, $X^{\alpha\beta}_{\dots\nu}$ and other similar contractions for ν .

We observe that, among the results, there are lowerings of indices (e.g. $X_{\nu}^{\dots\beta\gamma}$, lowering $\alpha \searrow \nu$, and $X^{\alpha\beta}_{\dots\nu}$, lowering $\gamma \searrow \nu$), together with other results that are *not* lowerings of indices (for example $X_{\nu}^{\dots\alpha\gamma}$: β is lowered $\beta \searrow \nu$ but also shifted). $X^{\alpha\gamma}_{\nu}$ does not appear among the results: it is therefore *not possible* to get by means of \mathbf{G} the transformation:



as if \vec{e}_{β} were isolated.

It's easy to see that, given the usual rules of the inner tensor product, *only indices placed at the beginning or at the end of the string can be raised / lowered by \mathbf{G}* (a similar conclusion applies to \mathbf{G}^{-1}).

It is clear, however, that this limitation does not manifest itself until the indexes' string includes only two, that is, for tensors of rank $r=2$.

▪ For tensors of rank $r=2$ and, restricted to extreme indexes for tensors

of higher rank, we can enunciate the rules:

G applied to a tensor *lowers an index* (the one by which it connects). Only the writing by components, e.g. $G_{\mu\nu} T^{\alpha\mu}_{\cdot\cdot\gamma} = T^{\alpha}_{\cdot\nu\gamma}$, clarifies which indexes are involved.

G⁻¹ applied to a tensor *raises an index* (the one by which it connects). For example $T^{\alpha\beta}_{\cdot\cdot\gamma\mu} G^{\mu\nu} = T^{\alpha\beta\cdot\nu}_{\cdot\cdot\gamma}$.

Mnemo

Examples: $G_{\kappa\nu} T^{\kappa\alpha}_{\cdot\cdot\beta} = T^{\alpha}_{\nu\cdot\beta}$: hook κ , raise it, rename it ν
 $T^{\alpha\beta}_{\cdot\cdot\lambda\mu} G^{\mu\nu} = T^{\alpha\beta\cdot\nu}_{\cdot\cdot\lambda}$: hook μ , raise it, rename it ν

▪ A special case is when the tensor **G** applies to its inverse **G⁻¹**:

$$G_{\alpha\beta} G^{\beta\gamma} = G^{\gamma}_{\alpha}$$

which can be thought to as a raising as well as a lowering of an index. But since the two tensors **G** and **G⁻¹** are inverse to each other:

$$G_{\alpha\beta} G^{\beta\gamma} = \delta^{\gamma}_{\alpha}$$

and from the last two it follows (not surprisingly, given the *eq.2.30!*) that:

$$G^{\gamma}_{\alpha} = \delta^{\gamma}_{\alpha} \tag{2.43}$$

2.18 Relations between **I**, **G**, δ

All equivalent to $\langle \tilde{A}, \tilde{B} \rangle = \langle \vec{A}, \vec{B} \rangle = \vec{A} \cdot \vec{B} = \tilde{A} \cdot \tilde{B}$ and to one another are the following expressions:

$$\mathbf{I}(\tilde{A}, \tilde{B}) = \mathbf{I}(\vec{A}, \vec{B}) = \mathbf{G}(\vec{A}, \vec{B}) = \mathbf{G}^{-1}(\tilde{A}, \tilde{B}) \tag{2.44}$$

(easy to see using T-mosaic).

▪ The componentwise writing for $\mathbf{G}(\mathbf{G}^{-1}) = \mathbf{I}$ gives:

$$G^{\alpha}_{\beta} = I^{\alpha}_{\beta},$$

that, together with *eq.2.43*, leads to:

$$G^{\alpha}_{\beta} = I^{\alpha}_{\beta} = \delta^{\alpha}_{\beta} \tag{2.45}$$

Likewise:

$$G^{\alpha\beta} = I^{\alpha\beta} = \delta^{\alpha\beta} ; G_{\alpha\beta} = I_{\alpha\beta} = \delta_{\alpha\beta} \tag{2.46}$$

- That does *not* mean that **G** and **I** are the same tensor. We observe that the name **G** is reserved to the tensor whose components are $G_{\alpha\beta}$; the other two tensors, whose components are G_{β}^{α} and $G^{\alpha\beta}$, are different tensors and cannot be labeled by the same name. In fact, the tensor with components $G^{\alpha\beta}$ is **G**⁻¹, while that one whose components are G_{β}^{α} coincides with **I**.

It is thus matter of three distinct tensors: *

$$\left| \begin{array}{l} \mathbf{I} \xrightarrow{comp} G_{\beta}^{\alpha} = I_{\beta}^{\alpha} = \delta_{\beta}^{\alpha} \\ \mathbf{G} \xrightarrow{comp} G_{\alpha\beta} = I_{\alpha\beta} = \delta_{\alpha\beta} \\ \mathbf{G}^{-1} \xrightarrow{comp} G^{\alpha\beta} = I^{\alpha\beta} = \delta^{\alpha\beta} \end{array} \right.$$

even if the components' names are somewhat misleading.

Note that only δ_{β}^{α} is the Kronecker delta, represented by the matrix *diag*(+1).

Besides, neither $I_{\alpha\beta}$ nor $I^{\alpha\beta}$ (and not even $\delta_{\alpha\beta}$ and $\delta^{\alpha\beta}$) deal with the identity tensor **I** (but rather with **G**; their matrix may be *diag*(+1) or not).

The conclusion is that we can label by the same name tensors with indexes moved up / down when using a componentwise notation, but we must be careful, when switching to the tensor notation, to avoid identifying as one different tensors.

Just to avoid any confusion, in practice the notations $I_{\alpha\beta}$, $I^{\alpha\beta}$, $\delta_{\alpha\beta}$, $\delta^{\alpha\beta}$ are almost never used.

* The rank is also different: $\binom{1}{1}, \binom{0}{2}, \binom{2}{0}$ respectively. Their matrix representation can be formally the same if **G** \equiv *diag* (+1), but on a *different* “basis grid”!

3 Change of basis

The vector \vec{V} has expansion $V^\alpha \vec{e}_\alpha$ on the basis of vectors $\{\vec{e}_\alpha\}$.

Its components will vary as the basis changes. In which way?

It's worth noting that the components change, but it does not the vector \vec{V} !

Let us denote by $V^{\beta'}$ the components on the new basis $\{\vec{e}_{\beta'}\}$; it is now $\vec{V} = V^{\beta'} \vec{e}_{\beta'}$, hence:

$$\vec{V} = V^{\beta'} \vec{e}_{\beta'} = V^\alpha \vec{e}_\alpha \quad 3.1$$

In other words, the same vector can be expanded on the new basis as well as it was expanded on the old one (from now on the upper ' will denote the new basis).

- Like all vectors, each basis-vector of the new basis $\vec{e}_{\beta'}$ can be expanded on the old basis-vectors \vec{e}_α :

$$\vec{e}_{\beta'} = \Lambda_{\beta'}^\alpha \vec{e}_\alpha \quad (\beta' = 1, 2, \dots, n) \quad 3.2$$

$\Lambda_{\beta'}^\alpha$ are the coefficients of the expansion that describes the “recipe” of the new $\vec{e}_{\beta'}$ on the old basis $\{\vec{e}_\alpha\}$ (i.e. on the basis of the old “ingredients”).

Taken as a whole, these coefficients express the new basis in terms of the old one; they can be arranged in a matrix $n \times n$ $[\Lambda_{\beta'}^\alpha]$

- *Conversely* we can express the old basis-vectors on the basis of new ones as:

$$\vec{e}_\alpha = \Lambda_\alpha^{\beta'} \vec{e}_{\beta'}, \quad (\alpha = 1, 2, \dots, n) \quad 3.3$$

The two matrices that appear in *eq.3.2* and *eq.3.3* are *inverse* to each other: exchanging the indexes the matrix is inverted because the sense of the transformation reverses; applying both them one after the other we return to the starting point.

- We now aim to find the relation between the new components of a given vector \vec{V} and the old ones. Using *eq.3.3*:

$$\left. \begin{aligned} \vec{V} &= V^\alpha \vec{e}_\alpha = V^\alpha \Lambda_\alpha^{\beta'} \vec{e}_{\beta'} \\ \text{and since: } \vec{V} &= V^{\beta'} \vec{e}_{\beta'} \end{aligned} \right\} \Rightarrow V^{\beta'} = \Lambda_\alpha^{\beta'} V^\alpha \quad 3.4$$

▪ What about covectors? First let us deduce the transformation law for components. From $P_\alpha = \tilde{P}(\tilde{e}_\alpha)$ *eq.1.4*, which also holds good in the new basis, by means of the transformation of basis-vectors *eq.3.2* that we already know, we get:

$$P_{\beta'} = \tilde{P}(\tilde{e}_{\beta'}) = \tilde{P}(\Lambda_{\beta'}^\alpha \tilde{e}_\alpha) = \Lambda_{\beta'}^\alpha \tilde{P}(\tilde{e}_\alpha) = \Lambda_{\beta'}^\alpha P_\alpha \quad 3.5$$

From the definition of component of \tilde{P} and using the *eq.3.5* above, we deduce the inverse transformation for basis-covectors (from new to old ones):

$$\left. \begin{aligned} \tilde{P} &= P_\alpha \tilde{e}^\alpha \\ \tilde{P} &= P_{\beta'} \tilde{e}^{\beta'} = \Lambda_{\beta'}^\alpha P_\alpha \tilde{e}^{\beta'} \end{aligned} \right\} \Rightarrow \tilde{e}^\alpha = \Lambda_{\beta'}^\alpha \tilde{e}^{\beta'}$$

and hence the direct from old to new:

$$\tilde{e}^{\beta'} = \Lambda_\alpha^{\beta'} \tilde{e}^\alpha \quad 3.6$$

▪ To summarize, all direct transformations (from old to new basis) are ruled by two matrices:

- $[\Lambda_{\alpha}^{\beta'}]$ expresses the transformation of the components of vectors (*eq.3.4*) and of basis-covectors (*eq.3.6*)
- $[\Lambda_{\beta'}^{\alpha}]$ expresses the transformation of the components of covectors (*eq.3.5*) and of basis-vectors (*eq.3.2*)

The two matrices are one the transposed inverse * of the other.

* The transpose M^T of a matrix M is obtained by interchanging rows and columns; the transposed inverse $(M^{-1})^T$ coincides with the inverse transpose $(M^T)^{-1}$. The notation we use cannot distinguish a matrix from its transpose (and the transposed inverse from the inverse, too) because the upper / lower indexes do not qualify rows and columns in a fixed way. Hence, the two matrices $\Lambda_{\alpha}^{\beta'}$ in *eq.3.3* and *eq.3.6* are not the same: they are the inverse in the first case and the transposed inverse in the second. It can be seen by expanding the two equations:

$$\begin{aligned} \text{eq.3.3: } \tilde{e}_\alpha = \Lambda_{\alpha}^{\beta'} \tilde{e}_{\beta'} &\Rightarrow \begin{cases} \tilde{e}_1 = \Lambda_1^{1'} \tilde{e}_{1'} + \Lambda_1^{2'} \tilde{e}_{2'} + \dots \\ \tilde{e}_2 = \Lambda_2^{1'} \tilde{e}_{1'} + \Lambda_2^{2'} \tilde{e}_{2'} + \dots \\ \dots \end{cases} \Rightarrow \Lambda_{\alpha}^{\beta'} = \begin{bmatrix} \Lambda_1^{1'} & \Lambda_1^{2'} & \dots \\ \Lambda_2^{1'} & \Lambda_2^{2'} & \dots \\ \dots & \dots & \dots \end{bmatrix} \\ \text{eq.3.6: } \tilde{e}^{\beta'} = \Lambda_{\alpha}^{\beta'} \tilde{e}^\alpha &\Rightarrow \begin{cases} \tilde{e}^{1'} = \Lambda_1^{1'} \tilde{e}^1 + \Lambda_1^{2'} \tilde{e}^2 + \dots \\ \tilde{e}^{2'} = \Lambda_1^{2'} \tilde{e}^1 + \Lambda_2^{2'} \tilde{e}^2 + \dots \\ \dots \end{cases} \Rightarrow \Lambda_{\alpha}^{\beta'} = \begin{bmatrix} \Lambda_1^{1'} & \Lambda_2^{1'} & \dots \\ \Lambda_1^{2'} & \Lambda_2^{2'} & \dots \\ \dots & \dots & \dots \end{bmatrix} \end{aligned}$$

On the contrary, the inverse transformation (from old to new basis) is ruled by two matrices that are the inverse of the previous two.

So, just one matrix (with its inverse and transposed inverse) is enough to describe all the transformations of (components of) vectors and covectors, basis-vectors and basis-covectors under change of basis.

- Let us agree to call Λ (without further specification) *the matrix that transforms vector components and basis-covectors* from the old bases system with index α to the new system indexed β' :

$$\Lambda \stackrel{\text{def}}{=} \left[\Lambda_{\alpha}^{\beta'} \right] \quad 3.7$$

Given this position, the transformations of bases and components are summarized in the following table:

$$\begin{array}{cc} \vec{e}_{\alpha} \xrightarrow{(\Lambda^{-1})^T} \vec{e}_{\beta'} & \tilde{e}^{\alpha} \xrightarrow{\Lambda} \tilde{e}^{\beta'} \\ V^{\alpha} \xrightarrow{\Lambda} V^{\beta'} & P_{\alpha} \xrightarrow{(\Lambda^{-1})^T} P_{\beta'} \end{array} \quad 3.8$$

Roughly speaking: if the bases vary in a manner, the components vary in the opposite one in order to leave the vector unchanged (as well as the number that expresses the measure of a quantity increases by making smaller the unit we use, and vice versa).

The transformations in the opposite sense (\leftarrow) require to invert all the matrices of the table (remind that the inverse of $(\Lambda^{-1})^T$ is Λ^T).

- For a transformation to be reversible, it is required that its matrix Λ is invertible, that means $\det \Lambda \neq 0$.
- The duality relation between bases of vectors and covectors holds good in the new basis, too:

$$\begin{aligned} \langle \tilde{e}^{\nu'}, \vec{e}_{\beta'} \rangle &= \langle \Lambda_{\mu}^{\nu'} \tilde{e}^{\mu}, \Lambda_{\beta}^{\alpha} \vec{e}_{\alpha} \rangle = \Lambda_{\mu}^{\nu'} \Lambda_{\beta}^{\alpha} \langle \tilde{e}^{\mu}, \vec{e}_{\alpha} \rangle = \\ &= \Lambda_{\mu}^{\nu'} \Lambda_{\beta}^{\alpha} \delta_{\alpha}^{\mu} = \Lambda_{\alpha}^{\nu'} \Lambda_{\beta}^{\alpha} = \delta_{\beta}^{\nu'} \end{aligned}$$

(the two matrices are inverse to each other, $[\Lambda_{\mu}^{\nu'}] = [\Lambda_{\beta}^{\alpha}]^{-1}$ and hence related by $\Lambda_{\alpha}^{\nu'} \Lambda_{\beta}^{\alpha} = \delta_{\beta}^{\nu'}$).

- The matrix Λ that rules the transformations can be expressed in terms of both old and new basis:

$$\delta_\alpha^\mu = \langle \vec{e}_\alpha, \tilde{e}^\mu \rangle = \langle \vec{e}_\alpha, \Lambda_{\nu'}^\mu \tilde{e}^{\nu'} \rangle = \Lambda_{\nu'}^\mu \langle \vec{e}_\alpha, \tilde{e}^{\nu'} \rangle ,$$

only possible if $\langle \vec{e}_\alpha, \tilde{e}^{\nu'} \rangle = \Lambda_{\alpha'}^{\nu'}$ since then $\delta_\alpha^\mu = \Lambda_{\nu'}^\mu \Lambda_{\alpha'}^{\nu'}$.

Hence, the element of the matrix Λ is:

$$\Lambda_{\alpha'}^{\nu'} = \langle \vec{e}_\alpha, \tilde{e}^{\nu'} \rangle \tag{3.9}$$

and the transformation matrix is thus built up by crossing old and new bases.

- In practice, it is not convenient trying to remind the transformation matrices to use in each different case, nor reasoning in terms of matrix calculus: the balance of the indexes inherent in the sum convention is an automatic mechanism that leads to write the right formulas in every case.

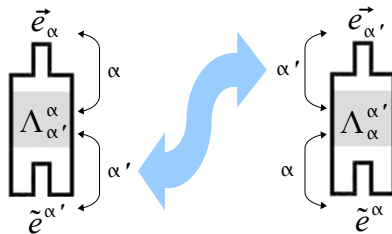
Mnemo

The transformation of a given object is correctly written by simply taking care to balance the indexes. For example, the transformation of components of covector from new to old basis, provisionally written $P_\mu = \Lambda P_{\nu'}$, can only be completed as $P_\mu = \Lambda_\mu^{\nu'} P_{\nu'}$ \Rightarrow the matrix element to use is $\Lambda_\mu^{\nu'}$

3.1 Basis change in T-mosaic

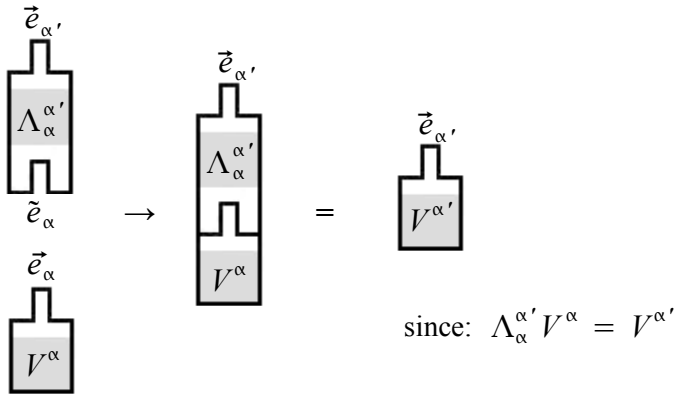
In the T-mosaic metaphor, the change of basis requires *each connector to be converted*. The conversion is performed by applying a “basis converter” extra-block to the connector; for convenience we represent it as a standard block with a pin and a hole (but beware: it is *not* a tensor!).

The basis-converter block is one and only, in two variants distinguished only by the different position of the indices with apex ' :

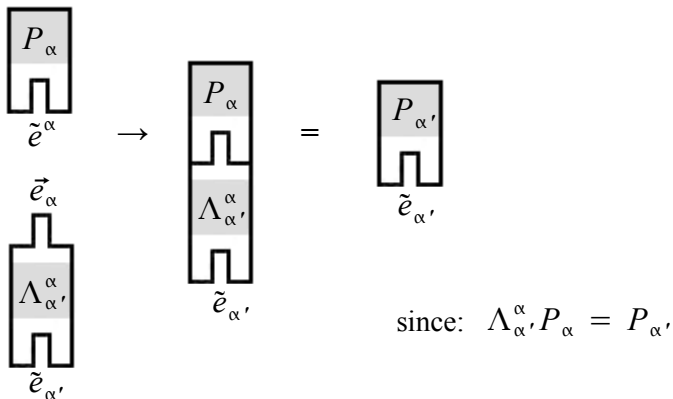


The connection is done “wearing” the basis-converters blocks as “shoes” on the connectors of the tensor, docking them on the pin-side or the hole-side as needed, "apex on apex" or "non-apex on non-apex". The body of the converter blocks will be marked with the element of the transformation matrix $\Lambda_{\alpha}^{\alpha'}$ or $\Lambda_{\alpha'}^{\alpha}$, with the apices ' up or down oriented in the same way with those of connectors (this implicitly leads to the correct choice of $\Lambda_{\alpha}^{\alpha'}$ or $\Lambda_{\alpha'}^{\alpha}$).

For example, to basis-transform the components of the vector $V^{\alpha} \vec{e}_{\alpha}$ we must apply on the \vec{e}_{α} connector a converter block that hooks it and replaces with $\vec{e}_{\alpha'}$:

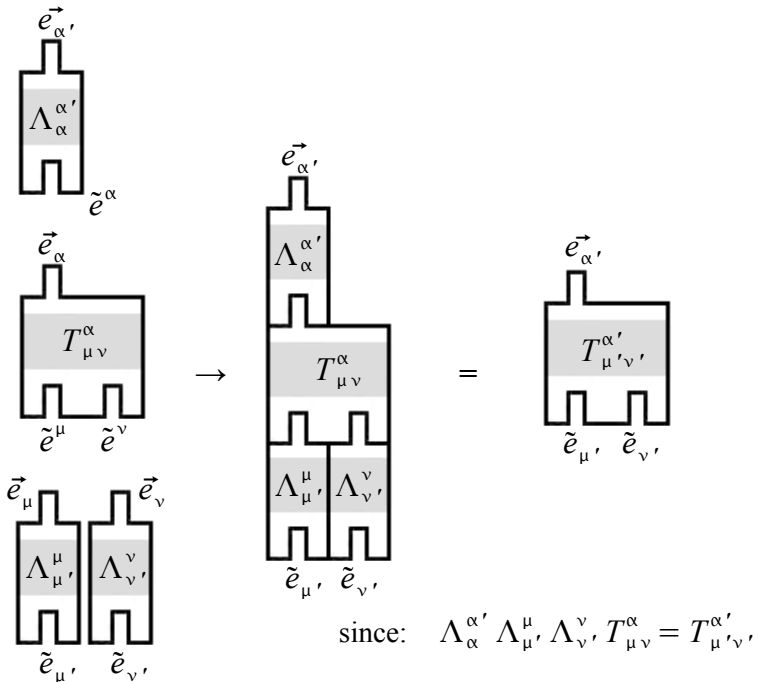


Similarly, to basis-transform the components of the covector $P_{\alpha} \tilde{e}^{\alpha}$, the appropriate converter block applies to the old connector \tilde{e}^{α} :



- This rule doesn't apply to basis-vectors $\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array}$ (the converter block would contain in that case the inverse matrix, if one); however the block representation of these instances has no practical interest.
- Subject to a basis transformation, a tensor needs to convert **all** its connectors by means of the matrices $\Lambda_{\alpha}^{\alpha'}$ or $\Lambda_{\alpha'}^{\alpha}$; an appropriate conversion block must be applied to each connector.

For example:



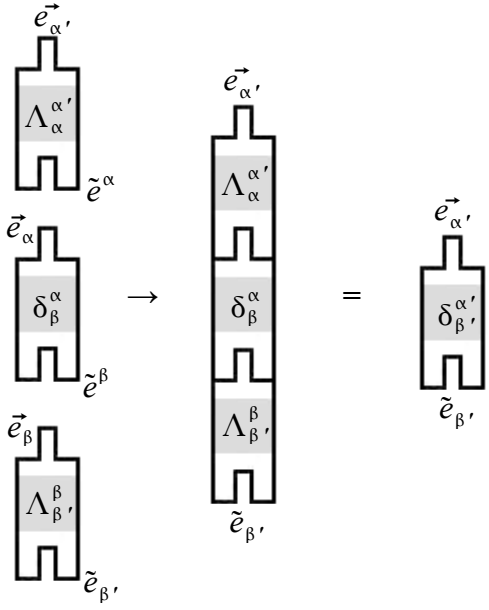
The position just made:

$$T_{\mu'\nu'}^{\alpha'} = \Lambda_{\alpha}^{\alpha'} \Lambda_{\mu}^{\mu'} \Lambda_{\nu}^{\nu'} T_{\mu\nu}^{\alpha} \quad 3.10$$

exemplifies the transformation rule for tensors:

to old \rightarrow new basis transform the components of a tensor we must apply as many $\Lambda_{\alpha}^{\alpha'}$ as upper indexes and as many $\Lambda_{\alpha'}^{\alpha}$ as lower indexes.

▪ A very special case is that of tensor $\mathbf{I} = \delta_{\beta}^{\alpha} \vec{e}_{\alpha} \tilde{e}^{\beta}$ whose components never change: $\delta_{\beta}^{\alpha} = \text{diag}(+1)$ is true *in any basis*:



since: $\Lambda_{\alpha}^{\alpha'} \Lambda_{\beta}^{\beta'} \delta_{\beta}^{\alpha} = \Lambda_{\beta}^{\alpha'} \Lambda_{\beta}^{\beta'} = \delta_{\beta}^{\alpha'}$

3.2 Invariance of the null tensor

A tensor is the null tensor when all its components are zero:

$$\mathbf{C} = 0 \iff \forall C_{\mu\dots}^{\alpha\dots} = 0 \quad * \quad 3.11$$

If a tensor is null in a certain basis, it is also null in any other basis.

▫ In fact:

$$\mathbf{C} = 0 \implies C_{\mu\dots}^{\alpha\dots} = 0 \implies C_{\mu'}^{\alpha'} \dots = \Lambda_{\alpha}^{\alpha'} \Lambda_{\mu}^{\mu'} \dots C_{\mu\dots}^{\alpha\dots} = 0 \quad 3.12$$

for all $\Lambda_{\alpha}^{\alpha'}, \Lambda_{\mu}^{\mu'} \dots$, that is for any transformation of bases.

From that the invariance of tensor equations follows.

* It makes no sense equating a tensor to a number; writing $\mathbf{C} = 0$ is only a conventional notation for $\forall C_{\mu\dots}^{\alpha\dots} = 0$. Otherwise written $\mathbf{0}$.

3.3 Invariance of tensor equations

Vectors, covectors and tensors are the invariants of a given “landscape”: changing the basis, only the way by which they are represented by their components varies. Then, also the relationships between them, or tensor equations, are invariant. For tensor equations we mean equations in which only tensors (vectors, covectors and scalars included) are involved, no matter whether written in tensor notation (\mathbf{T} , \vec{V} , ecc.) or componentwise.

Reduced to its essential terms, a tensor equation is an equality between two tensor like:

$$\mathbf{A} = \mathbf{B} \quad \text{or} \quad A_{\mu\dots}^{\alpha\dots} = B_{\mu\dots}^{\alpha\dots} \quad 3.13$$

Now, it's enough putting $\mathbf{A} - \mathbf{B} = \mathbf{C}$ to reduce to the equivalent form *eq.3.12* $\mathbf{C} = 0$ or $C_{\mu\dots}^{\alpha\dots} = 0$ that, as is valid in a basis, is valid in all bases. It follows that *eq.3.13*, too, as is valid in a basis is valid in all bases. So the equations in tensor form are not affected by the particular basis chosen, but they apply in all bases without changes.

In practice

|| an equation derived in a system of bases, once expressed in tensor form, applies as well in any system of bases.

Also some properties of tensors, if expressed by tensor relations, are valid regardless of the basis. It is the case, inter alia, of the properties of symmetry / skew-symmetry and invertibility. For example, *eq.2.15* that expresses the symmetry properties of a double tensor is a tensor equation, and that ensures that the symmetries of a tensor are the same in any basis.

The condition of invertibility *eq.2.27*, *eq.2.28*, is a tensor equation too, and therefore a tensor which has inverse in a basis has inverse in any other basis.

As will be seen in the following paragraphs, the change of bases can be induced by a transformation of the coordinates of space in which tensors are set. Tensor equations, inasmuch invariant under change of basis regardless of the reason that it is due, will be *invariant under coordinate transformation* (that is, valid in any reference accessible via coordinate transformation).

In these features reside the strength and the most interest of the tensor formulation.

4 Tensors in manifolds

We have so far considered vectors and tensors defined at a single point P . This does not mean that P should be an isolated point: vectors and tensors are usually given as vector fields or tensor fields defined in some domain or continuum of points.

Henceforth we will not restrict to \mathbb{R}^n space, but we shall consider a wider class of spaces retaining some basic analytical properties of \mathbb{R}^n , such as the differentiability (of functions therein defined). Belongs to this larger class any n -dimensional space \mathcal{M} whose points may be put in a one-to-one (= bijective) and continuous correspondence with the points of \mathbb{R}^n (or its subset). Continuity of correspondence means that points close in space \mathcal{M} have as image points close in \mathbb{R}^n , that is a requisite for the differentiability in \mathcal{M} .

Under these conditions we refer to \mathcal{M} as a *differentiable manifold*.

(Sometime we'll also use, instead of the term "manifold", the less technical "space" with the same meaning).

Roughly speaking, a differentiable manifold of dimension n is a space that can be continuously "mapped" in \mathbb{R}^n (with the possible exception of some points).

- In more precise terms:

It is required that every infinitesimal neighborhood of the generic point $P \in \mathcal{M}$ has as image an infinitesimal neighborhood of the corresponding point P' in \mathbb{R}^n or, turning to finite terms, that for every open set $U \subset \mathcal{M}$ there is a continuous correspondence $\varphi: U \rightarrow \mathbb{R}^n$. A couple of elements (U, φ) is called a "chart".

- One complication comes from the fact that, in general, there is not a correspondence that transports the whole space \mathcal{M} into \mathbb{R}^n , i.e. there is not a single chart (\mathcal{M}, φ) . To map \mathcal{M} completely we need a collection of charts (U_i, φ_i) , called "atlas".

- The charts of the atlas must have overlapping areas at the edges so as to ensure the transition between one chart to the other. Some points (or rather, some neighborhoods) will appear on more than one charts, each of which being the result of a different correspondence rule. For example, a point Q will be mapped as point $\varphi(Q)$ on a chart and as $\psi(Q)$ on another one: it

is required that the correspondence $\varphi(Q) \leftrightarrow \psi(Q)$ is itself continuous and differentiable.

An example of a differentiable manifold is a two-dimensional spherical surface like the Earth's surface, that can be covered by two-dimensional plane charts (even though no single chart can cover the entire globe and some points, different depending on the type of correspondence used, are left out; for example, the Mercator projection excludes the poles).

The fact that there is no correspondence able to transport by itself the whole manifold into \mathbb{R}^n is not an heavy limitation when, as often occurs, only single points are left out. One correspondence can suffice in these cases to describe *almost* completely the space. In practical terms, we can say that there is a correspondence $\varphi: \mathcal{M} \rightarrow \mathbb{R}^n$, meaning now with \mathcal{M} the space deprived of some points.*

In any case the limitation that until now has brought us to consider vectors emerging from a single point is still valid: each vector or tensor is related to the point where it is defined and cannot be “operated” with vectors or tensors defined at different points, except they are infinitely close. This limitation comes from the fact that we cannot presume to freely transport vectors (and tensors) in any space as it is usually done in the plane or in spaces \mathbb{R}^n .

As already mentioned, the set of tensors of rank $\binom{h}{k}$ defined at a point P is itself a vector space of dimension n^{h+k} . In particular, the already known vector space of dimension n of the vectors defined in a point P is called the *tangent space*; the similar for covectors is named the *cotangent space*.

Due to the regularity of the correspondence φ , an enough small neighborhood of any point P of the differentiable manifold will behave as a neighborhood of \mathbb{R}^n : a differentiable manifold appears *locally* like a flat one, even if its overall structure is different.

In practice, the term differentiable manifold means a space to which we can apply a coordinate system: this is indeed the meaning of the

* In the following, when we talk about a coordinate system defined on a manifold, we will tacitly assume that some single points can be left outside. To be more precise, may be that “null measure sets” (as is a line in a 2D space) are excluded.

correspondence φ .

4.1 Coordinate systems

Given a n -dimensional differentiable manifold \mathcal{M} , defining a certain correspondence $\varphi: \mathcal{M} \rightarrow \mathbb{R}^n$ means “tagging” each point P of \mathcal{M} with n numbers x^μ ($\mu = 1, 2, \dots, n$) that identify it uniquely. The n -tuple expresses the *coordinates* of P (another way, punctual, to designate the correspondence φ is to write $P \leftrightarrow \{x^\mu\}$).

Since the manifold is differentiable, the correspondence is continuous and transforms each neighborhood of P in a neighborhood of $\{x^\mu\}$.

The numbers that express the coordinates may be lengths, angles or otherwise. We note that a coordinate system does not presuppose any definition of distance between two points of the manifold.

4.2 Coordinate lines and surfaces

In the neighborhood of a point P whose coordinates are (x^1, x^2, \dots, x^n) no other points can have the same coordinates, but there are points that have *some* coordinates equal to P (provided that at least one is different).

The points that share with P only a single coordinate forms a (hyper) surface in $n-1$ dimensions. These *coordinate surfaces* passing through P are n in number, one for each coordinate which remains fixed.

The points that share with P $n-1$ coordinates and differ only in one, let's call it x^ν , align on a line along which only x^ν changes.

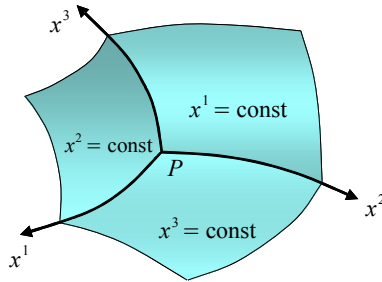
n *coordinate lines* of this kind emerge from P .

A coordinate line originates from the intersection of $n-1$ coordinate (hyper)surfaces (the intersection leaves free one coordinate only). Thus, at each point P of the manifold n coordinate surfaces and n coordinate lines intersect.

- The representation comes easy in 3D:

At any point P three coordinate surfaces intersect; on each of them the value of a coordinate remains constant. These surfaces, intersecting two by two, generate 3 coordinate lines; on each of them two coordinates are constant and only the third varies. The coordinate surfaces are labeled by the coordinate that remains constant, the coordinate lines by the coordinate that varies, as

shown in the figure.

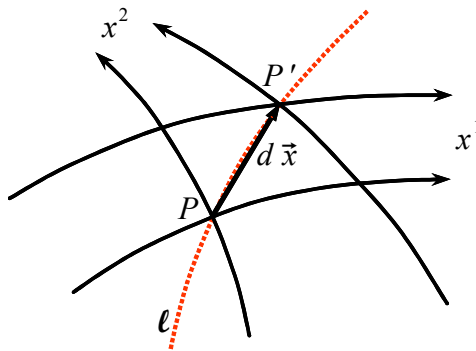


Roughly speaking: defining a coordinate system in a manifold means to draw an infinitely dense n -dimensional grid of coordinate lines within the space (in general not straight nor intersecting at right angle).

4.3 Coordinate bases

An opportunity that arises after defined a coordinate system is to link the bases of vectors and covectors to the coordinates themselves.

The idea is to associate to the coordinates a basis of vectors consisting of n basis-vectors, each of which is tangent to one of the n coordinate lines that intersect in the point.



Given a generic point P of coordinate x^μ , infinite parametric lines* pass through it; let l be one of them. Moving along l by an infinitesimal the point moves from P to P' while the parameter

* A parametric line in a n -dimensional manifold is defined by a system of n equations $x^\mu = x^\mu(s)$ with parameter s ($\mu = 1, 2, \dots, n$). To each value of s corresponds a point of the line.

increases from s to $s + ds$ and the coordinates of the point move from x^1, x^2, \dots, x^n to $x^1 + dx^1, x^2 + dx^2, \dots, x^n + dx^n$.

The displacement from P to P' is then $d\vec{x} \xrightarrow{comp} dx^\mu$, or:

$$d\vec{x} = dx^\mu \vec{e}_\mu. \tag{4.1}$$

This equation works as a definition of a vector basis $\{\vec{e}_\mu\}$ in P , in such a manner that each basis-vector is tangent to a coordinate line.

Note that $d\vec{x}$ is a vector (inasmuch it is independent of the coordinate system which we refer to).

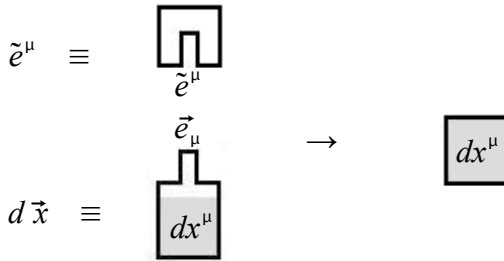
A basis of vectors defined in this way, related to the coordinates, is called a **coordinate vector basis**. Of course it is a basis among other ones, but the easiest to use because only here $d\vec{x} \xrightarrow{comp} dx^\mu$! *

It is worth noting that in general it depends upon the point P .

▪ A further relation between $d\vec{x}$ and its components dx^μ is:

$$dx^\mu = \langle \tilde{e}^\mu, d\vec{x} \rangle \tag{4.2}$$

(it is nothing but the rule *eq.1.10*, according to which we get the components by applying the vector to the dual basis); its blockwise representation in T-mosaic metaphor looks like:



The basis $\{\tilde{e}^\mu\}$ is the **coordinate covector basis**, dual to the previously introduced coordinate vector basis.

▪ We aim now to link even this covector basis to the coordinates.

Let us consider a scalar function of the point f defined on the manifold (at least along ℓ) as a function of the coordinates. Its variation from the initial point P along a path $d\vec{x}$ is given by its total differential:

$$df = \frac{\partial f}{\partial x^\mu} dx^\mu \tag{4.3}$$

* If the basis is not that defined by *eq.4.1*, we can still decompose $d\vec{x}$ along the coordinate lines x^μ , but dx^μ is no longer a component along a basis-vector.

We note that the derivatives $\frac{\partial f}{\partial x^\mu}$ are the components of a covector because their product by the components dx^μ of the vector $d\vec{x}$ gives the scalar df (in other words, eq.4.3 can be interpreted as an heterogeneous scalar product).*

Let us denote by $\tilde{\nabla} f \xrightarrow{comp} \frac{\partial f}{\partial x^\mu}$ this covector, whose components *in a coordinate basis* are the partial derivatives of the function f and therefore it can be identified with the gradient of f itself.

In symbols, *in a coordinate basis*:

$$\tilde{grad} \equiv \tilde{\nabla} \xrightarrow{comp} \frac{\partial}{\partial x^\mu} \quad 4.4$$

or, in short notation $\partial_\mu \equiv \frac{\partial}{\partial x^\mu}$ for partial derivatives:

$$\tilde{\nabla} \xrightarrow{comp} \partial_\mu \quad 4.5$$

Note that the gradient is a covector, not a vector.

The total differential (eq.4.3) can now be written in vector form:

$$df = \tilde{\nabla} f (d\vec{x}) = \langle \tilde{\nabla} f, d\vec{x} \rangle$$

If as scalar function f we take one of the coordinates x^ν ** we get:

$$dx^\nu = \langle \tilde{\nabla} x^\nu, d\vec{x} \rangle$$

By comparison with the already known $dx^\nu = \langle \tilde{e}^\nu, d\vec{x} \rangle$ (eq.4.2) \Rightarrow

$$\Rightarrow \tilde{e}^\nu = \tilde{\nabla} x^\nu \quad \text{in a coordinate basis} \quad 4.6$$

This means that in coordinate bases the basis-covector \tilde{e}^ν coincides with the gradient of the coordinate x^ν and is therefore oriented in the direction of the faster variation of the coordinate itself, which is that of the normal to the coordinate surface $x^\nu = \text{const}$.

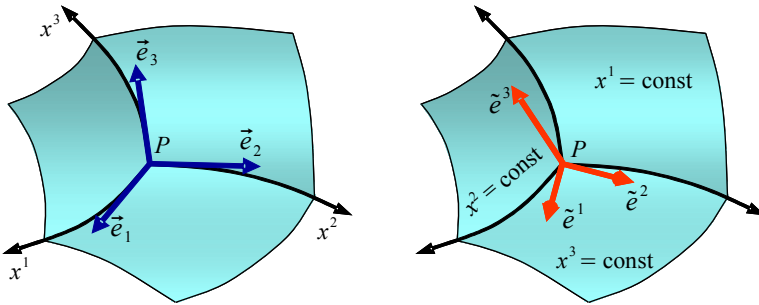
In summary, provided we use *coordinate bases*:

- the basis-vectors \vec{e}_μ are tangent to coordinate lines
- the basis-covectors \tilde{e}^ν are normal to coordinate surfaces

* Note that, as we ask dx^μ to be a component of $d\vec{x}$, we assume to work *in a coordinate basis*.

** It means taking as line ℓ the coordinate line x^ν .

as shown in the following figures that illustrate the 3D case:



When the coordinate system is orthogonal (\Rightarrow the coordinate lines, in general not straight, intersect at each point at right angle) basis-vectors and basis-covectors have the same direction.* If the coordinates are not orthogonal the two sets of vectors and covectors are differently oriented.

Covectors, as well as vectors, can be represented by arrows, although we must keep in mind that they are entities of a different kind (for instance, vectors and covectors are not directly summable).**

4.4 Coordinate bases and non-coordinate bases

In 3D Cartesian coordinates it is usual to write the infinitesimal displacement as:

$$d\vec{x} = \vec{i} dx + \vec{j} dy + \vec{k} dz$$

or, labeling by \hat{e} the *unit* vectors:

$$d\vec{x} = \hat{e}_1 dx^1 + \hat{e}_2 dx^2 + \hat{e}_3 dx^3 = \hat{e}_\mu dx^\mu \quad 4.7$$

which is precisely the condition of coordinate basis $d\vec{x} = \vec{e}_\mu dx^\mu$ (eq.4.1) with $\vec{e}_\mu = \hat{e}_\mu$.

This means that unit vectors \hat{e}_μ , or $\vec{i}, \vec{j}, \vec{k}$, are a coordinate basis. Vectors $\vec{i}, \vec{j}, \vec{k}$, are in fact everywhere oriented as coordinate lines, and this is true not only punctually, but in the whole space \mathbb{R}^3 .

It is not always the case. For example, in plane polar coordinates (ρ, θ)

* In general, however, they will not match in length or magnitude.

** Some authors shy away from this representation, that is entirely lawful with the warning of the diversity of kind.

$d\vec{r}$ is expressed as:

$$d\vec{x} = \hat{e}_\rho d\rho + \hat{e}_\theta \rho d\theta \quad 4.8$$

and this can be identified with the coordinate basis condition $d\vec{x} = \vec{e}_\mu dx^\mu$ provided we take as a basis:

$$\vec{e}_\rho = \hat{e}_\rho, \quad \vec{e}_\theta = \rho \hat{e}_\theta \quad 4.9$$

Here the basis-vector \vec{e}_θ "includes" ρ , it's no longer a unit vector and also varies from point to point. Conversely, the basis of unit vectors $\{\hat{e}_\rho, \hat{e}_\theta\}$ which is currently used in Vector Analysis for polar coordinates is *not* a coordinate basis.

- An example of a non-coordinate basis in 3D Cartesian coordinates is obtained by applying a 45° rotation to unit vectors \vec{i}, \vec{j} on the horizontal plane, leaving \vec{k} unchanged. Expressing \vec{i}, \vec{j} in terms of the new rotated vectors we \vec{i}', \vec{j}' get for $d\vec{x}$:

$$d\vec{x} = \frac{\sqrt{2}}{2}(\vec{i}' - \vec{j}')dx + \frac{\sqrt{2}}{2}(\vec{i}' + \vec{j}')dy + \vec{k}dz, \quad 4.10$$

and this relation matches the coordinate bases condition (eq.4.1) only by taking as basis vectors:

$$\frac{\sqrt{2}}{2}(\vec{i}' - \vec{j}'), \quad \frac{\sqrt{2}}{2}(\vec{i}' + \vec{j}'), \quad \vec{k}$$

that is nothing but the old coordinate basis $\{\vec{i}, \vec{j}, \vec{k}\}$. It goes without saying that the rotated basis $\{\vec{i}', \vec{j}', \vec{k}\}$ is not a coordinate basis (as we clearly see from eq.4.10).

This example shows that, as already said, *only in a coordinate basis* $d\vec{x} \xrightarrow{comp} dx^\mu$ holds good. Not otherwise!

- As an example we can deduce the dual coordinate basis of covectors from the duality relation $\langle \vec{e}^\mu, \vec{e}_\nu \rangle = \delta_\mu^\nu$ in both the following cases:

① In 3D Cartesian coordinates it is easy to see that the vector coordinate basis $\{\vec{e}_\nu\} \equiv \{\vec{i}, \vec{j}, \vec{k}\}$ coincides with the covector coordinate basis $\{\vec{e}^\mu\} \equiv \{\vec{i}, \vec{j}, \vec{k}\}$ obtained by duality *

* \vec{i} and \tilde{i} (in general \vec{e}_ν and \vec{e}^μ) remain distinct entities, although superimposable. In Cartesian they have the same expansion by components and their "arrows" coincide, but under change of coordinates they return to differ.

- Indeed, it is already known that in Cartesian the coordinate basis is made of the unit vectors $\vec{i}, \vec{j}, \vec{k}$. They can be written in terms of components:

$$\vec{i} = (1, 0, 0) \quad , \quad \vec{j} = (0, 1, 0) \quad , \quad \vec{k} = (0, 0, 1)$$

The coordinate covector basis will necessarily be:

$$\tilde{i} = (1, 0, 0) \quad , \quad \tilde{j} = (0, 1, 0) \quad , \quad \tilde{k} = (0, 0, 1)$$

because only this way you can get:

$$\begin{aligned} \langle \tilde{i}, \vec{i} \rangle &= 1 \quad , \quad \langle \tilde{i}, \vec{j} \rangle = 0 \quad , \quad \langle \tilde{i}, \vec{k} \rangle = 0 \quad , \\ \langle \tilde{j}, \vec{i} \rangle &= 0 \quad , \quad \langle \tilde{j}, \vec{j} \rangle = 1 \quad , \quad \langle \tilde{j}, \vec{k} \rangle = 0 \quad , \\ \langle \tilde{k}, \vec{i} \rangle &= 0 \quad , \quad \langle \tilde{k}, \vec{j} \rangle = 0 \quad , \quad \langle \tilde{k}, \vec{k} \rangle = 1 \quad , \end{aligned}$$

as required by the condition of duality.

- ⊛ In polar coordinates, on the contrary, the vector coordinate basis $\{\vec{e}_\nu\} \equiv \{\vec{e}_\rho, \vec{e}_\theta\}$ does not match with the covector basis deduced from duality condition $\{\tilde{e}^\nu\} \equiv \{\tilde{e}^\rho, \tilde{e}^\theta\}$.

- Indeed: unit vectors in polar coordinates are by definition:

$$\hat{e}_\rho = (1, 0) \quad , \quad \hat{e}_\theta = (0, 1) \quad . \quad \text{Thus, from eq.4.9:}$$

$$\vec{e}_\rho = (1, 0) \quad , \quad \vec{e}_\theta = (0, \rho)$$

Let be $\tilde{e}^\rho = (a, b)$, $\tilde{e}^\theta = (c, d)$. From $\langle \tilde{e}^\mu, \vec{e}_\nu \rangle = \delta_\nu^\mu \Rightarrow$

$$\Rightarrow \left\{ \begin{array}{l} 1 = \langle \tilde{e}^\rho, \vec{e}_\rho \rangle = (a, b) \cdot (1, 0) = a \\ 0 = \langle \tilde{e}^\rho, \vec{e}_\theta \rangle = (a, b) \cdot (0, \rho) = b\rho \Rightarrow b=0 \\ 0 = \langle \tilde{e}^\theta, \vec{e}_\rho \rangle = (c, d) \cdot (1, 0) = c \\ 1 = \langle \tilde{e}^\theta, \vec{e}_\theta \rangle = (c, d) \cdot (0, \rho) = d\rho \Rightarrow d = \frac{1}{\rho} \end{array} \right\} \Rightarrow \tilde{e}^\rho = (1, 0) \\ \tilde{e}^\theta = (0, \frac{1}{\rho})$$

Note that the dual switch tensor \mathbf{G} , even if already defined, has no role in the calculation of a basis from its dual.

4.5 Change of the coordinate system

The change of coordinates, by itself, does not request the change of the bases, but we must do it if we want to continue working in coordinate basis in the new system as well. As usual, the change of basis will be done by means of a transformation matrix $\Lambda = [\Lambda_\mu^{\nu'}]$ which depends upon both the old $\{x^\mu\}$ and new $\{x^{\nu'}\}$ coordinate

systems (an apex ' marks the new ones). But which Λ must we use for a given transformation of coordinates?

To find the actual form of Λ we impose to vector basis a twofold condition: that the starting basis is a coordinate basis (first equality) and that the arrival basis is a coordinated basis, too (second equality):

$$d\vec{x} = \vec{e}_\mu dx^\mu = \vec{e}_{\nu'} dx^{\nu'} \quad 4.11$$

The generic new coordinate $x^{\nu'}$ will be related to all the old ones x^1, x^2, \dots, x^n by a function $x^{\nu'} = x^{\nu'}(x^1, x^2, \dots, x^n)$, i.e.

$$x^{\nu'} = x^{\nu'}(\dots, x^\mu, \dots)$$

whose total differential is:

$$dx^{\nu'} = \frac{\partial x^{\nu'}}{\partial x^\mu} dx^\mu$$

Substituting into the double equality *eq.4.11*:

$$d\vec{x} = \vec{e}_\mu dx^\mu = \vec{e}_{\nu'} \frac{\partial x^{\nu'}}{\partial x^\mu} dx^\mu \quad \Rightarrow \quad \vec{e}_\mu = \vec{e}_{\nu'} \frac{\partial x^{\nu'}}{\partial x^\mu} \quad 4.12$$

A comparison with the change of basis-vectors *eq.3.3* (inverse, *from new to old*):

$$\vec{e}_\mu = \Lambda_\mu^{\nu'} \vec{e}_{\nu'}$$

leads to the conclusion that:

$$\Lambda_\mu^{\nu'} = \frac{\partial x^{\nu'}}{\partial x^\mu} \quad 4.13$$

$\Rightarrow \Lambda_\mu^{\nu'}$ is the matrix element of Λ complying with the definition *eq.3.7* that we were looking for.*

Since the *Jacobian matrix* of the transformation is defined as:

$$J \stackrel{\text{def}}{=} \begin{bmatrix} \frac{\partial x^{1'}}{\partial x^1} & \frac{\partial x^{1'}}{\partial x^2} & \dots & \frac{\partial x^{1'}}{\partial x^n} \\ \frac{\partial x^{2'}}{\partial x^1} & \frac{\partial x^{2'}}{\partial x^2} & & \frac{\partial x^{2'}}{\partial x^n} \\ \vdots & & & \vdots \\ \frac{\partial x^{n'}}{\partial x^1} & \frac{\partial x^{n'}}{\partial x^2} & \dots & \frac{\partial x^{n'}}{\partial x^n} \end{bmatrix} = \left[\frac{\partial x^{\nu'}}{\partial x^\mu} \right] \quad 4.14$$

* Or equally, as in the case, element of its transpose.

we can identify:

$$\Lambda \equiv J \qquad 4.15$$

▪ The coordinate transformation has thus as *related matrix* the Jacobian matrix $\Lambda \equiv J$ (whose elements are the partial derivatives of the new coordinates with respect to the old ones). It states, together with its inverse and transpose, how the old coordinate bases transform into the new coordinate bases, and consequently how the components of vectors and tensors transform.

After a coordinate transformation we have to transform bases and components by means of the related matrix $\Lambda \equiv J$ (and its inverse / transpose) to continue working in a coordinate basis.

▪ As always, to avoid confusion, it is more practical to start from the transformation formulas and take care to balance the indexes; the indexes of Λ will suggest the proper partial derivatives to be taken.

For example, $P_{v'} = \Lambda_{v'}^{\mu} P_{\mu} \Rightarrow$ use the matrix $\Lambda_{v'}^{\mu} = \frac{\partial x^{\mu}}{\partial x^{v'}}$.

Note that the upper / lower position of the index marked ' in partial derivatives agrees with the upper / lower position on the symbol Λ .

4.6 Contravariant and covariant tensors

We summarize the transformation laws for *components* of vectors and covectors *in a coordinate basis*:

- vector components: $V^{v'} = \frac{\partial x^{v'}}{\partial x^{\mu}} V^{\mu}$ (*contravariant* scheme)
- covector components: $P_{v'} = \frac{\partial x^{\mu}}{\partial x^{v'}} P_{\mu}$ (*covariant* scheme)

In the traditional formulation of Tensor Calculus “by components”, the two transformation schemes above are assumed as definitions of *contravariant vector* and *covariant vector* respectively.

The equivalence with the terminology used so far is:

- vector \leftrightarrow contravariant vector
- covector (or “1-form”) \leftrightarrow covariant vector

For a tensor of higher rank, for example $\binom{1}{2}$, the following law of transformation of the components applies:

$$T_{\mu' \nu'}^{\alpha'} = \frac{\partial x^{\alpha'}}{\partial x^{\alpha}} \frac{\partial x^{\mu}}{\partial x^{\mu'}} \frac{\partial x^{\nu}}{\partial x^{\nu'}} T_{\mu \nu}^{\alpha} \quad 4.16$$

which is just a different way of writing *eq.3.10 in case of coordinate basis* (in that case *eq.4.13* holds).

The generalization of *eq.3.10* to tensors of any rank is obvious.

In general, a tensor of rank $\binom{h}{k}$ is told contravariant of order (= rank) h and covariant of order k .

Traditional texts begin here, using the transformation laws under coordinate transformation of components* as a definition: a tensor is defined as a multidimensional entity whose components transform according to example *eq.4.16*.

4.7 Affine tensors

In conformity with the traditional definition, tensors are those quantities that transform according to the contravariant or covariant schemes under *any* admissible coordinate transformation. If we restrict the class of the allowed coordinate transformations, the number of quantities that underlie the schemes of “tensorial” (contravariant / covariant) transformation increases. In particular, if we restrict to consider *linear transformations*, we identify the class of *affine* tensors, wider than that of tensors without further qualification.

The coordinate transformations to which we restrict in this case have the form $x^{\nu'} = a_{\mu}^{\nu'} x^{\mu}$ where $a_{\mu}^{\nu'}$ are purely numerical coefficients.

These coefficients can be arranged in a matrix $[a_{\mu}^{\nu'}]$ that expresses the law of coordinate transformation. But also the transformation matrix for vector components, i.e. the Jacobian matrix J given by

$$\Lambda_{\mu}^{\nu'} = \frac{\partial x^{\nu'}}{\partial x^{\mu}}$$

because here is $\frac{\partial x^{\nu'}}{\partial x^{\mu}} = a_{\mu}^{\nu'}$.

So, for affine tensors, the law of coordinate transformation and the transformation law for vector components coincide and are both expressed by the same matrix Λ (the components of covectors transform as usual by the transposed inverse of Λ).

* Although the traditional “by components” approach avoids speaking about bases, it is implied that there *we always work in a coordinate basis* (both old and new).

4.8 Cartesian tensors

If we further restrict to linear *orthogonal* transformations, we identify the class of Cartesian tensors (wider than affine tensors).

A linear orthogonal transformation is represented by an orthogonal matrix.* Also for Cartesian tensors the same matrix $[a_\mu^{\nu'}]$ rules both the coordinate transformation and the transformation of vector components. Moreover, since an orthogonal matrix coincides with its transposed inverse, the same matrix $[a_\mu^{\nu'}]$ transforms as well covector components. So, as vectors and covectors transform in the same way, they are the same thing: the distinction between vectors and covectors (or contravariant tensors and covariant tensors) falls in the case of Cartesian tensors.

4.9 Magnitude of vectors

We define *magnitude* (or *norm*) of a vector \vec{V} the scalar $|\vec{V}|$ such that:

$$|\vec{V}|^2 = \vec{V} \cdot \vec{V} = \mathbf{G}(\vec{V}, \vec{V}) \quad 4.17$$

Note that this definition, given by means of a scalar product, depends on the switch \mathbf{G} that has been chosen.

4.10 Distance and metric tensor

In a manifold with a coordinate system x^μ a *distance* ds from a point P to a point P' shifted by $d\vec{x}$ can be defined as:

$$ds^2 = |d\vec{x}|^2 = \mathbf{G}(d\vec{x}, d\vec{x}) \quad 4.18$$

In fact, what we define is an infinitesimal distance, also called “arc element”.

After defined a distance the manifold becomes a *metric space*.

Since its definition is given by the scalar product, the distance depends on the \mathbf{G} we use. If we want the distance to be defined in a certain way, it follows that \mathbf{G} must be chosen appropriately.

We will denote \mathbf{g} the tensor $\binom{0}{2}$ we pick out among the possible switches \mathbf{G} to get the desired definition of distance.

* A matrix is orthogonal when its rows are orthogonal vectors (i.e. scalar product in pairs = 0), and so its columns.

The distance ds given by the tensor equation:

$$ds^2 = \mathbf{g}(d\vec{x}, d\vec{x}) \quad 4.19$$

is an invariant scalar and does not depend on the coordinate system.

At this point, the tensor \mathbf{g} defines the geometry of the manifold, i.e. its metric properties, based on the definition of distance between two points. For this reason \mathbf{g} is called **metric tensor**, or “metric”.

▪ If (and only if) we use *coordinate bases*, the distance can be written:

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta \quad 4.20$$

▫ Indeed, (only) in a coordinate basis is $d\vec{x} = \vec{e}_\mu dx^\mu$, hence:

$$\begin{aligned} ds^2 = \mathbf{g}(d\vec{x}, d\vec{x}) &= \mathbf{g}(\vec{e}_\alpha dx^\alpha, \vec{e}_\beta dx^\beta) = \mathbf{g}(\vec{e}_\alpha, \vec{e}_\beta) dx^\alpha dx^\beta = \\ &= g_{\alpha\beta} dx^\alpha dx^\beta \end{aligned}$$

4.11 Euclidean distance

If the manifold is the usual Euclidean space \mathbb{R}^3 , using a Cartesian coordinate system, the distance is classically defined:

$$ds = \sqrt{dx^2 + dy^2 + dz^2} \quad 4.21$$

or:

$$ds^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 \quad 4.22$$

which express the Pythagoras theorem and coincides with the form (eq.4.20) that applies in coordinate bases

$$ds^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 = g_{\alpha\beta} dx^\alpha dx^\beta \quad 4.23$$

provided we take:

$$g_{\alpha\beta} = \begin{cases} 1 & \text{for } \alpha = \beta \\ 0 & \text{for } \alpha \neq \beta \end{cases} \quad \text{i.e.} \quad \mathbf{g} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \text{diag}(+1) \quad 4.24$$

Notice that in this case $\mathbf{g} = \mathbf{g}^{-1}$, namely $g_{\alpha\beta} = g^{\alpha\beta}$.

4.12 Generalized distances

The definition of distance given by eq.4.20 is more general than the Euclidean case and includes any bilinear form in dx^1, dx^2, \dots, dx^n such as:

$$g_{11}dx^1 dx^1 + g_{12}dx^1 dx^2 + g_{13}dx^1 dx^3 + \dots + g_{21}dx^2 dx^1 + g_{22}dx^2 dx^2 + \dots + g_{nn}dx^n dx^n \quad 4.25$$

where $g_{\alpha\beta}$ are subject to few restrictions. Inasmuch they build up the matrix that represents the metric tensor \mathbf{g} , we must at least require that the matrix is symmetric and $\det[g_{\alpha\beta}] \neq 0$ such that \mathbf{g} is invertible and \mathbf{g}^{-1} exists.

- Note, however, that \mathbf{g} is symmetrizable in any case: for example, if it were $g_{12} \neq g_{21}$ in the bilinear form eq.4.25, just taking $g'_{12} = g'_{21} = (g_{12} + g_{21})/2$ nothing would change in the definition of distance.

Metric spaces where the distance is defined by eq.4.19 with any \mathbf{g} , provided symmetric and invertible, are called **Riemann spaces**.

- Given a manifold, its metric properties are fully described by the metric tensor \mathbf{g} associated to it. As a tensor, \mathbf{g} does not depend on the coordinate system imposed onto the manifold: \mathbf{g} does not change if the coordinate system changes. However, we do not know how to represent \mathbf{g} by itself: its representation is only possible in terms of components $g_{\alpha\beta}$ and they do depend upon the particular coordinate system (and therefore upon coordinate basis). For a given \mathbf{g} we then have different matrices $[g_{\alpha\beta}]$, one for each coordinate system, which mean the same \mathbf{g} , i.e. the same metric properties for the manifold.

For instance, for the 2D Euclidean plane, the metric tensor \mathbf{g} , which expresses the usual Euclidean metric properties is represented by

$$[g_{\alpha\beta}] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ in Cartesian coordinates and by } [g_{\alpha'\beta'}] = \begin{bmatrix} 1 & 0 \\ 0 & \rho^2 \end{bmatrix} \text{ in}$$

polar coordinates. Of course, there are countless other $[g_{\mu'\nu'}]$ which can be obtained from the previous ones by coordinate transformation and subsequent change of basis by means of the related matrix Λ

(remind that $[g_{\alpha\beta}] \xrightarrow{\Lambda_{\mu'}^{\alpha} \Lambda_{\nu'}^{\beta}} [g_{\mu'\nu'}]$, i.e. $g_{\mu'\nu'} = \Lambda_{\mu'}^{\alpha} \Lambda_{\nu'}^{\beta} g_{\alpha\beta}$).

In other words, the same Euclidean metric can be expressed by all the $[g_{\mu'\nu'}]$ attainable by coordinate transformation from the Cartesian

$$[g_{\alpha\beta}] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} .$$

- Summary: $\left. \begin{array}{l} \text{geometry of manifold} \\ \text{coordinate system} \end{array} \right\} \Leftrightarrow \mathbf{g} \Rightarrow \left. \begin{array}{l} \\ \\ \end{array} \right\} g_{\mu\nu}$

4.13 Tensors and not -

The contra / covariant transformation laws (exemplified by eq.4.16) provide a criterion for determining whether a quantity is a tensor. If the quantity changes (in fact its components change) according to those schemes, the quantity is a tensor.

For example, the transformation for (components of) $d\vec{x}$ (eq.4.11):

$$dx^{v'} = \frac{\partial x^{v'}}{\partial x^{\mu}} dx^{\mu}$$

confirms that the infinitesimal displacement vector is a contravariant tensor of the first order.

Instead, the “position” or “radius-vector” \vec{x} is not a tensor because its components do not transform according to the law of coordinate transformation (which in general has nothing to do with the contra / covariant schemes).

▪ Neither is a tensor the derivative of (components of) a vector. In fact, given a vector that transforms its components:

$$V^{\mu} = V^{v'} \frac{\partial x^{\mu}}{\partial x^{v'}} \quad 4.26$$

its derivatives transform (according to Leibniz rule * and using the “chain rule”) as:

$$\frac{\partial V^{\mu}}{\partial x^{\alpha}} = \frac{\partial V^{v'}}{\partial x^{\alpha}} \frac{\partial x^{\mu}}{\partial x^{v'}} + V^{v'} \frac{\partial}{\partial x^{\alpha}} \frac{\partial x^{\mu}}{\partial x^{v'}} = \frac{\partial V^{v'}}{\partial x^{k'}} \frac{\partial x^{k'}}{\partial x^{\alpha}} \frac{\partial x^{\mu}}{\partial x^{v'}} + V^{v'} \frac{\partial^2 x^{\mu}}{\partial x^{\alpha} \partial x^{v'}} \quad 4.27$$

which is *not* the tensor transformation scheme (it would be for a tensor $\binom{1}{1}$ without the additional term in ∂^2). This conclusion is not surprising because partial derivatives are not independent of the coordinates.

4.14 Covariant derivative

▪ In general, the derivative of a vector (which is itself a vector) is not simply the derivative of the components, because even the basis-vectors vary from point to point in the manifold:

$$\frac{\partial \vec{V}}{\partial x^{\mu}} \equiv \partial_{\mu} \vec{V} = \partial_{\mu} (\vec{e}_v V^v) = \vec{e}_v \partial_{\mu} V^v + V^v \partial_{\mu} \vec{e}_v \quad 4.28$$

* Leibniz rule for derivative of a product: $(f \cdot g)' = f' \cdot g + f \cdot g'$

Of the two terms, the first describes the variability of the components of the vector, the second the variability of the basis-vectors along the coordinate lines x^μ .

We observe that $\partial_\mu \vec{e}_\nu$, the partial derivative of a basis-vector, although *not* a tensor (see Appendix) is a vector in a geometric sense (because it represents the change of the basis vector \vec{e}_ν along the infinitesimal arc dx^μ on the coordinate line x^μ) and therefore can be expanded on the basis of vectors.

Hence, setting $\partial_\mu \vec{e}_\nu = \vec{\Gamma}$, we can write:

$$\vec{\Gamma} = \Gamma^\lambda \vec{e}_\lambda \quad 4.29$$

But in reality $\vec{\Gamma}$ is already labeled with lower indexes μ, ν , hence $\vec{\Gamma}_{\mu\nu} = \Gamma^\lambda_{\mu\nu} \vec{e}_\lambda$, so that:

$$\partial_\mu \vec{e}_\nu = \Gamma^\lambda_{\mu\nu} \vec{e}_\lambda \quad 4.30$$

The vectors $\vec{\Gamma}_{\mu\nu}$ are n^2 in number, one for each pair μ, ν , with n components $\Gamma^\lambda_{\mu\nu}$ each (for $\lambda = 1, 2, \dots, n$).

The coefficients $\Gamma^\lambda_{\mu\nu}$ are called **Christoffel symbols** or *connection coefficients*; they are in total n^3 coefficients.

Eq.4.30 is equivalent to:

$$\partial_\mu \vec{e}_\nu \xrightarrow{\text{comp}} \Gamma^\lambda_{\mu\nu} \quad 4.31$$

The Christoffel symbols $\Gamma^\nu_{\mu\lambda}$ are therefore the components of the vectors “partial derivative of a basis vector” on the basis of the vectors themselves

namely:

The “recipe” of the partial derivative of a basis vector uses as “ingredients” the same basis vectors while Christoffel symbols represent their amounts.

Inserting eq.4.30, eq.4.28 becomes:

$$\partial_\mu \vec{V} = \vec{e}_\nu \partial_\mu V^\nu + \Gamma^\lambda_{\mu\nu} \vec{e}_\lambda V^\nu =$$

in the last right term ν, λ are dummy summation indexes: they can therefore be freely changed (and interchanged). Interchanging ν, λ :

$$= \vec{e}_\nu \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} \vec{e}_\nu V^\lambda = \vec{e}_\nu (\partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda)$$

In short:
$$\partial_\mu \vec{V} = \vec{e}_\nu \underbrace{(\partial_\mu V^\nu + \Gamma_{\mu\lambda}^\nu V^\lambda)}_{V^\nu_{;\mu}} \quad 4.32$$

We define **covariant derivative of a vector** with respect to x^μ :

$$V^\nu_{;\mu} \stackrel{\text{def}}{=} \partial_\mu V^\nu + \Gamma_{\mu\lambda}^\nu V^\lambda \quad 4.33$$

(we introduce here the subscript ; to denote the *covariant derivative*). The covariant derivative $V^\nu_{;\mu}$ thus differs from the respective partial derivative for a corrective term in Γ . Unlike the ordinary derivative the covariant derivative is in nature a tensor: the n^2 covariant derivatives of eq.4.33 are the components of a tensor, as we will see (in Appendix a check calculation *in extenso*).

- The derivative of a vector is written in terms of covariant derivative (from eq.4.32) as:

$$\partial_\mu \vec{V} = \vec{e}_\nu V^\nu_{;\mu} \quad 4.34$$

The (partial) derivative of a vector is thus a vector whose scalar components are the n covariant derivatives of the vector.

- In the same way we treat the case of the derivative of a covector:

$$\frac{\partial \tilde{A}}{\partial x^\mu} \equiv \partial_\mu \tilde{A} = \partial_\mu (\tilde{e}^\nu A_\nu) = \tilde{e}^\nu \partial_\mu A_\nu + A_\nu \partial_\mu \tilde{e}^\nu \quad 4.35$$

As before, the term $\partial_\mu \tilde{e}^\nu$ describes the variability of the basis-covector \tilde{e}^ν along the coordinate line x^μ and, though not a covector in tensorial sense, can be expanded on the basis of covectors.

Setting $\partial_\mu \tilde{e}^\nu = \tilde{L}$, we write:

$$\tilde{L} = L_\lambda \tilde{e}^\lambda$$

But in reality \tilde{L} is already labeled with an upper ν and a lower μ index, then:

$$\tilde{L}^\nu_\mu = L^\nu_{\mu\lambda} \tilde{e}^\lambda$$

hence:

$$\partial_\mu \tilde{e}^\nu = L^\nu_{\mu\lambda} \tilde{e}^\lambda \quad 4.36$$

which, substituted into eq.4.35, gives:

$$\partial_\mu \tilde{A} = \tilde{e}^\nu \partial_\mu A_\nu + L^\nu_{\mu\lambda} \tilde{e}^\lambda A_\nu =$$

interchanging the dummy indexes ν, λ in the last term:

$$= \tilde{e}^\nu \partial_\mu A_\nu + L_{\mu\nu}^\lambda \tilde{e}^\nu A_\lambda = \tilde{e}^\nu (\partial_\mu A_\nu + L_{\mu\nu}^\lambda A_\lambda) \quad 4.37$$

- What is the relationship between Γ and L ? They are simply opposite in sign:

$$L_{\mu\nu}^\lambda = -\Gamma_{\mu\nu}^\lambda \quad 4.38$$

- It can be shown by calculating separately the derivatives of both right and left members of the duality condition $\langle \tilde{e}^\mu, \vec{e}_\nu \rangle = \delta_\nu^\mu$. Since $\delta_\nu^\mu = 0$ or 1 (however a constant), then $\partial_\kappa(\delta_\nu^\mu) = 0$.

On the other hand:

$$\begin{aligned} \partial_\kappa \langle \tilde{e}^\mu, \vec{e}_\nu \rangle &= \langle \partial_\kappa \tilde{e}^\mu, \vec{e}_\nu \rangle + \langle \tilde{e}^\mu, \partial_\kappa \vec{e}_\nu \rangle \\ &= \langle L_{\kappa\lambda}^\mu \tilde{e}^\lambda, \vec{e}_\nu \rangle + \langle \tilde{e}^\mu, \Gamma_{\kappa\nu}^\lambda \vec{e}_\lambda \rangle \\ &= L_{\kappa\lambda}^\mu \langle \tilde{e}^\lambda, \vec{e}_\nu \rangle + \Gamma_{\kappa\nu}^\lambda \langle \tilde{e}^\mu, \vec{e}_\lambda \rangle \\ &= L_{\kappa\lambda}^\mu \delta_\nu^\lambda + \Gamma_{\kappa\nu}^\lambda \delta_\lambda^\mu \\ &= L_{\kappa\nu}^\mu + \Gamma_{\kappa\nu}^\mu \end{aligned}$$

Hence $0 = L_{\kappa\nu}^\mu + \Gamma_{\kappa\nu}^\mu \Rightarrow eq.4.38$, q.e.d.

Therefore: $\partial_\mu \tilde{e}^\nu = -\Gamma_{\mu\lambda}^\nu \tilde{e}^\lambda \quad 4.39$

and from eq.4.37: $\partial_\mu \tilde{A} = \tilde{e}^\nu \underbrace{(\partial_\mu A_\nu - \Gamma_{\mu\nu}^\lambda A_\lambda)}_{A_{\nu;\mu}} \quad 4.40$

We define **covariant derivative of a covector** with respect to x^μ :

$$A_{\nu;\mu} \stackrel{\text{def}}{=} \partial_\mu A_\nu - \Gamma_{\mu\nu}^\lambda A_\lambda \quad 4.41$$

which is analogous to eq.4.33.

- After this position, the derivative of a covector is written (from eq.4.40):

$$\partial_\mu \tilde{A} = \tilde{e}^\nu A_{\nu;\mu} \quad 4.42$$

▮ The (partial) derivative of a covector has as scalar components n covariant derivatives of the covector.

4.15 The gradient $\tilde{\nabla}$ at work

We have already stated that the gradient $\tilde{\nabla}$ is a covector whose components *in coordinate bases* are the partial derivatives with respect to the coordinates (eq.4.4, eq.4.5):

$$grad \equiv \tilde{\nabla} \xrightarrow{comp} \partial_\mu$$

namely

$$\tilde{\nabla} = \tilde{e}^\mu \partial_\mu \quad 4.43$$

- That $\tilde{\nabla}$ is a tensor has been deduced from eq.4.3 in case of $\tilde{\nabla}$ applied to f , but it is in general true, confirmed by the fact that the “chain rule” applied to its components $\partial_{v'} = \partial_\mu \frac{\partial x^\mu}{\partial x^{v'}}$ coincides with the covariant transformation scheme eq.4.16.

As a tensor, various tensor operation can be performed by $\tilde{\nabla}$ applying it to other tensors. We will examine the cases:

- outer product with vectors, covectors or tensors \Rightarrow **gradient**
- inner product with vector or tensor \Rightarrow **divergence**

① Tensor outer product $\tilde{\nabla} \otimes$ vector

The result is the tensor $\binom{1}{1}$ **gradient of vector**. In coordinate bases:

$$\tilde{\nabla} \otimes \vec{V} = \tilde{e}^\mu \partial_\mu \otimes \vec{V}$$

that is to say *

$$\tilde{\nabla} \vec{V} = \tilde{e}^\mu \partial_\mu \vec{V} \quad 4.44$$

$\partial_\mu \vec{V}$ comes from eq.4.32. Substituting:

$$\tilde{\nabla} \vec{V} = \tilde{e}^\mu \tilde{e}_\nu (\partial_\mu V^\nu + \Gamma_{\mu\lambda}^\nu V^\lambda) \quad 4.45$$

On the other hand, $\tilde{\nabla} \vec{V}$ may be written as expansion:

$$\tilde{\nabla} \vec{V} = \nabla_\mu V^\nu \tilde{e}^\mu \tilde{e}_\nu \quad 4.46$$

and by comparison between the last two we see that

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma_{\mu\lambda}^\nu V^\lambda \quad 4.47$$

which is equivalent to eq.4.33 in an alternative notation:

* It is convenient to think of the symbol \otimes as a connective between tensors. In T-mosaic it means to juxtapose and past the two blocks $\tilde{e}^\mu \partial_\mu$ and \vec{V}

$$\nabla_{\mu} V^{\nu} \equiv V^{\nu}_{;\mu} \quad 4.48$$

Eq.4.47 represents the *covariant derivative of vector*, qualifying it as component of the tensor gradient of vector $\tilde{\nabla} \vec{V}$ ($\equiv \tilde{\nabla} \otimes \vec{V}$).

The covariant derivatives of a vector are the n^2 components of the tensor “gradient of vector” $\tilde{\nabla} \otimes \vec{V}$ (more often written $\tilde{\nabla} \vec{V}$).

- In Appendix we show that the covariant derivatives $V^{\nu}_{;\mu}$ transform according to the tensor contra / covariant scheme exemplified by eq.4.16 and this confirms that the gradient $\tilde{\nabla} \vec{V}$ of vector of which they are components is indeed a tensor.

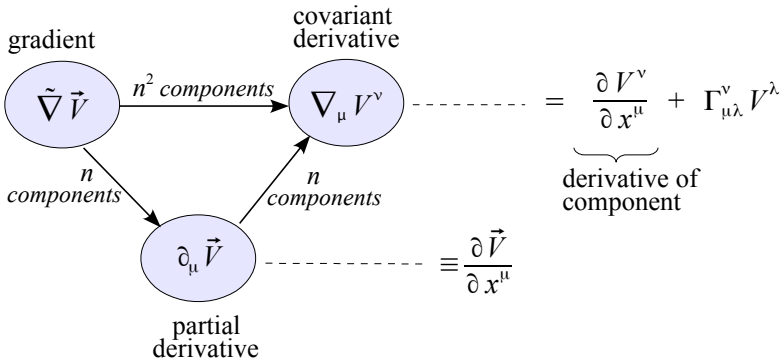
- The partial derivative of a vector (eq.4.34) may be written in the new notation in terms of $\tilde{\nabla}$ as:

$$\partial_{\mu} \vec{V} = \vec{e}_{\nu} \nabla_{\mu} V^{\nu} \quad 4.49$$

namely:

$$\partial_{\mu} \vec{V} \xrightarrow{\text{comp}} \nabla_{\mu} V^{\nu} \quad 4.50$$

- What we have seen about gradient tensor and derivative of a vector can be summarized in the sketch:



- When in particular the vector is a basis-vector the sketch above reduces to:

$$\tilde{\nabla} \vec{e}_{\nu} \xrightarrow{n^2 \text{ components}} \Gamma_{\mu\nu}^{\lambda}$$

$$\partial_{\mu} \vec{e}_{\nu} \xrightarrow{n} \Gamma_{\mu\nu}^{\lambda}$$

according to eq.4.44, eq.4.31:

$$\tilde{\nabla} \tilde{e}_\nu = \tilde{e}^\mu \partial_\mu \tilde{e}_\nu = \tilde{e}^\mu \tilde{e}_\lambda \Gamma_{\mu\nu}^\lambda . \quad 4.51$$

Terminology

We reserve the term “covariant derivative” to the components which together form the “gradient tensor”:

$$\underbrace{\tilde{\nabla} \tilde{V}}_{\text{gradient tensor}} \xrightarrow{\text{comp}} \underbrace{\nabla_\mu V^\nu}_{\text{covariant derivative}} = \underbrace{\partial_\mu V^\nu}_{\text{partial derivative}} + \underbrace{\Gamma_{\mu\lambda}^\nu}_{\text{Christoffel symbol}} V^\lambda$$

Notation

The covariant derivative of V^ν with respect to x^μ is symbolized

$$\nabla_\mu V^\nu \quad \text{or} \quad V^\nu_{;\mu}$$

Not to be confused with the partial derivative $\frac{\partial V^\nu}{\partial x^\alpha}$, indicated as

$$\partial_\mu V^\nu \quad \text{or} \quad V^\nu_{,\mu}$$

The lower index after the , means partial derivative

The lower index after the ; means covariant derivative

② **Tensor outer product $\tilde{\nabla} \otimes$ covector**

The result is the $\binom{0}{2}$ tensor **gradient of covector**. In coordinate bases:

$$\tilde{\nabla} \otimes \tilde{A} = \tilde{e}^\mu \partial_\mu \otimes \tilde{A}$$

or

$$\tilde{\nabla} \tilde{A} = \tilde{e}^\mu \partial_\mu \tilde{A} \quad 4.52$$

We go on as in the case of gradient of a vector:

$\partial_\mu \tilde{A}$ comes from eq.4.40; replacing:

$$\tilde{\nabla} \tilde{A} = \tilde{e}^\mu \tilde{e}^\nu (\partial_\mu A_\nu - \Gamma_{\mu\nu}^\lambda A_\lambda) \quad 4.53$$

and by comparison with the generic definition

$$\tilde{\nabla} \tilde{A} = \nabla_\mu A_\nu \tilde{e}^\mu \tilde{e}^\nu \quad 4.54$$

we get:

$$\nabla_\mu A_\nu = \partial_\mu A_\nu - \Gamma_{\mu\nu}^\lambda A_\lambda \quad 4.55$$

which is equivalent to eq.4.41 with the alternative notation:

$$\nabla_\mu A_\nu \equiv A_{\nu;\mu} \quad 4.56$$

Eq.4.55 represents the *covariant derivative of covector* and qualifies it as a component of the tensor gradient of covector $\tilde{\nabla} \tilde{A} (\equiv \tilde{\nabla} \otimes \tilde{A})$. Eq.4.55 is the analogous of eq.4.47, which was stated for vectors.

▪ The partial derivative of a covector (eq.4.42) may be written in the new notation in terms of $\tilde{\nabla}$ as:

$$\partial_\mu \tilde{A} = \tilde{e}_\nu \nabla_\mu A_\nu \tag{4.57}$$

namely:

$$\partial_\mu \tilde{A} \xrightarrow{comp} \nabla_\mu A_\nu \tag{4.58}$$

▪ For the gradient and the derivative of a covector a sketch like that shown for vectors applies, with few obvious changes.

Mnemo

To write the covariant derivative of vectors and covectors:

- we had better writing the Γ before the component, not vice versa
- 1st step:

$$\nabla_\mu A^\nu = \partial_\mu A^\nu + \Gamma_\mu^\nu \dots \quad \text{or} \quad \nabla_\mu A_\nu = \partial_\mu A_\nu - \Gamma_{\mu\nu}^\dots$$
 (adjust the indexes of Γ according to those of the 1st member)
- 2nd step: paste a dummy index to Γ and to the component A that follows so as to balance the upper / lower indexes:

$$\dots + \Gamma_{\mu\kappa}^\nu A^\kappa \quad \text{or} \quad \dots - \Gamma_{\mu\nu}^\kappa A_\kappa$$

• **Tensor outer product $\tilde{\nabla} \otimes$ tensor $\binom{h}{k}$**

$\tilde{\nabla} \otimes \mathbf{T}$ (or $\tilde{\nabla} \mathbf{T}$) is a $\binom{h}{k+1}$ tensor, the *gradient of tensor*.

By the same procedure of previous cases we can find that, for example for a rank $\binom{2}{1}$ tensor $\mathbf{T} = T_y^{\alpha\beta} \tilde{e}_\alpha \tilde{e}_\beta \tilde{e}^\gamma$ is:

$$\tilde{\nabla} \mathbf{T} = \nabla_\mu T_y^{\alpha\beta} \tilde{e}_\mu \tilde{e}_\alpha \tilde{e}_\beta \tilde{e}^\gamma \quad \text{namely:}$$

$$\tilde{\nabla} \mathbf{T} \xrightarrow{comp} \nabla_\mu T_y^{\alpha\beta} = \partial_\mu T_y^{\alpha\beta} + \Gamma_{\mu\kappa}^\alpha T_y^{\kappa\beta} + \Gamma_{\mu\kappa}^\beta T_y^{\alpha\kappa} - \Gamma_{\mu\gamma}^\kappa T_y^{\alpha\beta} \tag{4.59}$$

In general, the *covariant derivative of a tensor* $\binom{h}{k}$ is obtained by attaching to the partial derivative h terms in Γ with positive sign and k terms in Γ with negative sign. Treat the indexes individually, one after the other in succession.

Mnemo

To write the covariant derivative of tensors, for example $\nabla_\mu T_\gamma^{\alpha\beta}$:

- after the term in ∂_μ consider the indexes of T one by one:
- write the first term in Γ as it were $\nabla_\mu T^{\alpha}$, then complete T with its remaining indexes β : $\rightarrow +\Gamma_{\mu\kappa}^\alpha \rightarrow +\Gamma_{\mu\kappa}^\alpha T_\gamma^{\kappa\beta}$
- write the second term in Γ as it were $\nabla_\mu T^\beta$, then complete T with its remaining indexes α : $\rightarrow +\Gamma_{\mu\kappa}^\beta \rightarrow +\Gamma_{\mu\kappa}^\beta T_\gamma^{\alpha\kappa}$
- write the third Γ as it were $\nabla_\mu T_\gamma$, then complete T with its remaining indexes $\alpha\beta$: $\rightarrow -\Gamma_{\mu\gamma}^\kappa \rightarrow -\Gamma_{\mu\gamma}^\kappa T_\kappa^{\alpha\beta}$

▪ **Considerations on the form of components of tensor $\tilde{\nabla}$**

Once stated that the gradient $\tilde{\nabla}$ is a tensor and ∂_μ are its n components, it follows that $\tilde{\nabla}\mathbf{T}$ is a tensor ranked $r+1$ if \mathbf{T} has rank r and that $\partial_\mu\mathbf{T}$ are its component at first level, but *not* in general its *scalar* components: they are scalar for $r=0$, i.e. $\mathbf{T} \equiv f$ scalar, but even for $r=1$, $\partial_\mu\mathbf{T}$ are vectors and assume a form more complex than the simple partial derivative (eq.4.32), while for $r > 1$ they are tensors. Notice that in any case $\partial_\mu\mathbf{T}$ are n in number while the scalar components of $\tilde{\nabla}\mathbf{T}$ ought to be n^{r+1} when \mathbf{T} has rank r . The *scalar* components of the gradient $\tilde{\nabla}\mathbf{T}$ are in fact the n^{r+1} covariant derivatives. They have a form that depends upon the rank r of the tensor, since ∂_μ applied to \mathbf{T} operates (see eq.4.28) the derivatives of each basis-vector or covector of \mathbf{T} as well as those of the components.

Denoting ∇_μ the covariant derivative we may write in general:

$$\tilde{\nabla} \xrightarrow{comp} \nabla_\mu \tag{4.60}$$

where ∇_μ are the *scalar* components which consist of:

- ∂_μ if applied to a scalar
- $\partial_\mu + \Gamma \dots$ if applied to (a component of) a vector \vec{V}
- $\partial_\mu - \Gamma \dots$ if applied to (a component of) a covector \tilde{P}
- $\partial_\mu \pm$ various terms in $\Gamma \dots$ according to the rank of tensor \mathbf{T}

The components covariant derivative ∇_μ then take a “variable arrangement” depending on the object to which $\tilde{\nabla}$ applies.

④ Inner product $\tilde{\nabla} \cdot$ vector

$$\tilde{\nabla}(\vec{V}) = \nabla_{\mu} V^{\mu} = V^{\mu}_{;\mu} = \text{div } \vec{V} \quad 4.61$$

It is the scalar *divergence of a vector*.

Note that $\tilde{\nabla}$ applied as inner product to a vector looks similar to a covariant derivative:

$$\nabla_{\mu} V^{\mu} = \partial_{\mu} V^{\mu} + \Gamma^{\mu}_{\mu\kappa} V^{\kappa} \quad 4.62$$

but with one (repeated) index only, and is actually a number. Roughly:

The divergence of a vector $\text{div } \vec{V}$ looks like a covariant derivative, except for a single repeated index: $\nabla_{\mu} V^{\mu}$ or $V^{\mu}_{;\mu}$

To justify the presence of additional terms in Γ it is appropriate to think of the divergence $\nabla_{\mu} V^{\mu}$ as a covariant derivative $\nabla_{\nu} V^{\mu}$ followed by an index contraction $\nu = \mu$ (more precisely it is a contraction of the gradient):

$$\tilde{\nabla} \otimes \vec{V} = \nabla_{\nu} V^{\mu} \tilde{e}^{\nu} \otimes \tilde{e}_{\mu} \xrightarrow{\text{contraction } \nu=\mu} \nabla_{\mu} V^{\mu} \quad 4.63$$

which results in a scalar product, that is a scalar.

This definition of divergence is a direct generalization of the divergence as known from Vector Analysis, where it is defined in Cartesian coordinates as the sum of the partial derivatives of the components of the vector. In fact, in a Cartesian frame, it is $\text{div } \vec{V} = \nabla_{\mu} V^{\mu} = \partial_{\mu} V^{\mu}$ provided the coefficients Γ are null.

⑤ Inner product $\tilde{\nabla} \cdot$ tensor $\binom{h}{k}$

Divergence(s) of tensors of higher order can similarly be defined provided at least one upper index exists.

It results in a tensor of rank $\binom{h-1}{k}$, tensor *divergence of a tensor*.

If there are more than one upper index, various divergences can be defined, one for each upper index. For each index of the tensor, both upper or lower, including the one involved, an adequate term in Γ must be added. For example, the divergence with respect to β index of the rank $\binom{2}{1}$ tensor $\mathbf{T} = T^{\alpha\beta}_{\gamma} \tilde{e}_{\alpha} \tilde{e}_{\beta} \tilde{e}^{\gamma}$ is the $\binom{1}{1}$ tensor:

$$\begin{aligned} \tilde{\nabla}(\mathbf{T}) &= \nabla_{\beta} T^{\alpha\beta}_{\gamma} \tilde{e}_{\alpha} \tilde{e}^{\gamma} && \text{namely:} \\ \tilde{\nabla}(\mathbf{T}) &\xrightarrow{\text{comp}} \nabla_{\beta} T^{\alpha\beta}_{\gamma} = \partial_{\beta} T^{\alpha\beta}_{\gamma} + \Gamma^{\alpha}_{\beta\kappa} T^{\kappa\beta}_{\gamma} + \Gamma^{\beta}_{\beta\kappa} T^{\alpha\kappa}_{\gamma} - \Gamma^{\kappa}_{\beta\gamma} T^{\alpha\beta}_{\kappa} \end{aligned} \quad 4.64$$

It comes to the contraction of $\tilde{\nabla} \mathbf{T}$ (eq.4.59) for $\mu = \beta$.

4.16 Gradient of some fundamental tensors

It is interesting to calculate the gradient of some fundamental tensors: the identity tensor and the metric tensor.

$$\begin{aligned} \blacksquare \tilde{\nabla} \mathbf{I} \xrightarrow{comp} \nabla_{\mu} \delta_{\beta}^{\alpha} &= \partial_{\mu} \delta_{\beta}^{\alpha} + \Gamma_{\mu\kappa}^{\alpha} \delta_{\beta}^{\kappa} - \Gamma_{\mu\beta}^{\kappa} \delta_{\kappa}^{\alpha} = \\ &= 0 + \Gamma_{\mu\beta}^{\alpha} - \Gamma_{\mu\beta}^{\alpha} = 0 \end{aligned} \quad 4.65$$

as might be expected.

$$\blacksquare \tilde{\nabla} \mathbf{g} \xrightarrow{comp} \nabla_{\mu} g_{\alpha\beta} = \partial_{\mu} g_{\alpha\beta} - \Gamma_{\mu\alpha}^{\kappa} g_{\kappa\beta} - \Gamma_{\mu\beta}^{\kappa} g_{\alpha\kappa} = 0 \quad 4.66$$

▫ Indeed (eq.2.37, eq.4.30):

$$\begin{aligned} \partial_{\mu} g_{\alpha\beta} &= \partial_{\mu} (\vec{e}_{\alpha} \cdot \vec{e}_{\beta}) = \underbrace{\partial_{\mu} \vec{e}_{\alpha}}_{=\Gamma_{\mu\alpha}^{\kappa} \vec{e}_{\kappa}} \cdot \vec{e}_{\beta} + \vec{e}_{\alpha} \cdot \underbrace{\partial_{\mu} \vec{e}_{\beta}}_{=\Gamma_{\mu\beta}^{\kappa} \vec{e}_{\kappa}} = \\ &= \Gamma_{\mu\alpha}^{\kappa} \vec{e}_{\kappa} \cdot \vec{e}_{\beta} + \vec{e}_{\alpha} \cdot \Gamma_{\mu\beta}^{\kappa} \vec{e}_{\kappa} = \Gamma_{\mu\alpha}^{\kappa} g_{\kappa\beta} + \Gamma_{\mu\beta}^{\kappa} g_{\alpha\kappa} \quad * \end{aligned}$$

This important result, less obvious than the previous one, states that

in each point of any manifold with Riemannian metric it is:

$$\tilde{\nabla} \mathbf{g} = 0 \quad 4.67$$

Likewise:

$$\tilde{\nabla} \mathbf{g}^{-1} = 0 \quad 4.68$$

▫ It can be obtained like eq.4.66, passing through $\nabla_{\mu} g^{\alpha\beta}$ and $\partial_{\mu} g^{\alpha\beta}$ and then considering that $\partial_{\mu} \tilde{e}^{\alpha} = -\Gamma_{\mu\kappa}^{\alpha} \tilde{e}^{\kappa}$ (eq.4.39).

The last eq.4.67, eq.4.68 are often respectively written as:

$$\nabla_{\kappa} g_{\alpha\beta} = 0 \quad \text{or} \quad g_{\alpha\beta;\kappa} = 0 \quad 4.69$$

and:
$$\nabla_{\kappa} g^{\mu\nu} = 0 \quad \text{or} \quad g^{\mu\nu}{}_{;\kappa} = 0$$

4.17 Covariant derivative and index raising / lowering

The covariant derivative and the raising or lowering of the indexes are operations that commute (= you can reverse the order in which they take place). For example:

$$g^{\alpha\mu} T_{\alpha\beta;\kappa} = T^{\mu}_{\beta;\kappa} \quad 4.70$$

* $\tilde{\nabla} \mathbf{g} = 0$ all over the space does *not* mean that \mathbf{g} is unvarying, but that it is punctually *stationary* (like a function in its maximum or minimum) in *each* point.

(on the left, covariant derivative followed by raising the index; on the right, raised index followed by covariant derivative).

▫ Indeed:

$$T_{\beta;\kappa}^{\mu} = \left(g^{\alpha\mu} T_{\alpha\beta} \right)_{;\kappa} = \underbrace{g^{\alpha\mu}}_{=0}{}_{;\kappa} T_{\alpha\beta} + g^{\alpha\mu} T_{\alpha\beta;\kappa} = g^{\alpha\mu} T_{\alpha\beta;\kappa}$$

having used the *eq.4.69*.

If is thus allowed “to raise/lower indexes under covariant derivative”.

4.18 Christoffel symbols

Christoffel symbols $\Gamma_{\mu\nu}^{\lambda}$ give account of the variability of basis-vectors from one point to the other of the manifold. In addition, they are what make the difference between covariant derivatives and ordinary derivatives: $\forall \Gamma = 0 \Leftrightarrow \partial_{\mu} = \bar{\nabla}_{\mu}$

▮ If (and only if) all Christoffel symbols Γ are null, ordinary derivatives and covariant derivatives coincide.

▪ In the usual 3D Euclidean space described by a Cartesian coordinate system, basis-vectors do not vary from point to point and consequently $\forall \Gamma = 0$ everywhere (so that covariant derivatives coincide with partial derivatives).

That is no more true when the same space is described by a spherical coordinate system ρ, ϕ, θ because in this case the basis-vectors $\vec{e}_{\rho}, \vec{e}_{\phi}, \vec{e}_{\theta}$ are functions of the point (as we saw for the plane polar coordinates) and this implies that at least some Γ are $\neq 0$.

Just this fact makes sure that the $\Gamma_{\mu\nu}^{\lambda}$ cannot be the components of a hypothetical tensor $\binom{1}{2}$ as might seem: the hypothetical tensor $\mathbf{\Gamma}$ would be null in Cartesian (since all its components are zero) and not null in spherical coordinates, against the invariance of a tensor under change of coordinates.

The Christoffel symbols are rather a set of n^3 coefficients depending on 3 indexes but do not form any tensor. However, observing the way in which it was introduced ($\vec{\Gamma} = \Gamma^{\lambda} \vec{e}_{\lambda}$, *eq.4.29*), each $\vec{\Gamma}$ related to a fixed pair of lower indexes is a vector whose components are marked by the upper index λ . It is therefore lawful (only for this index) the raising / lowering by means of \mathbf{g} :

$$g_{\kappa\lambda} \Gamma_{\mu\nu}^\lambda = \Gamma_{\mu\nu\kappa} \quad 4.71$$

▪ An important property of the Christoffel symbols is to be symmetric in a coordinate basis with respect to the exchange of lower indexes:

$$\Gamma_{\mu\nu}^\lambda = \Gamma_{\nu\mu}^\lambda \quad 4.72$$

▫ To prove this we operate a change of coordinates and a consequent transformation of coordinate basis by the appropriate related matrix. Let us express the old $\Gamma_{\mu\nu}^\lambda$ in terms of the new basis (using for this purpose the matrix $\Lambda_v^{\kappa'} = \frac{\partial x^{\kappa'}}{\partial x^v}$) :

$$\begin{aligned} \Gamma_{\mu\nu}^\lambda \vec{e}_\lambda &\stackrel{\text{def}}{=} \partial_\mu \vec{e}_\nu = \partial_\mu (\Lambda_v^{\kappa'} \vec{e}_{\kappa'}) = (\partial_\mu \Lambda_v^{\kappa'}) \vec{e}_{\kappa'} + \Lambda_v^{\kappa'} (\partial_\mu \vec{e}_{\kappa'}) = \\ &= \partial_\mu \frac{\partial x^{\kappa'}}{\partial x^v} \vec{e}_{\kappa'} + \frac{\partial x^{\kappa'}}{\partial x^v} \partial_\mu \vec{e}_{\kappa'} = \frac{\partial^2 x^{\kappa'}}{\partial x^\mu \partial x^v} \vec{e}_{\kappa'} + \frac{\partial x^{\kappa'}}{\partial x^v} \frac{\partial x^{\lambda'}}{\partial x^\mu} \frac{\partial}{\partial x^{\lambda'}} \vec{e}_{\kappa'}, \end{aligned}$$

where we have used the chain rule to insert $x^{\lambda'}$.

Since this expression is symmetric with respect to the indexes μ, ν , the same result is obtained by calculating $\Gamma_{\nu\mu}^\lambda \vec{e}_\lambda \stackrel{\text{def}}{=} \partial_\nu \vec{e}_\mu$ on the new basis, then $\Gamma_{\mu\nu}^\lambda = \Gamma_{\nu\mu}^\lambda$.

▪ An important relationship links Γ to \mathbf{g} and shows that in coordinate bases the coefficients Γ are functions of \mathbf{g} and its first derivatives only :

$$\Gamma_{\mu\beta}^\kappa = \frac{1}{2} g^{\kappa\alpha} (-g_{\mu\beta,\alpha} + g_{\beta\alpha,\mu} + g_{\alpha\mu,\beta}) \quad 4.73$$

(note the cyclicity of indexes μ, β, α within the parenthesis).

▫ That comes out from:

$$\begin{aligned} \tilde{\nabla} \mathbf{g} = 0 &\Rightarrow \nabla_\mu g_{\alpha\beta} = \partial_\mu g_{\alpha\beta} - \Gamma_{\mu\alpha}^\lambda g_{\lambda\beta} - \Gamma_{\mu\beta}^\lambda g_{\alpha\lambda} = 0 \\ &\Rightarrow (\text{using the comma instead of } \partial \text{ as derivative symbol}): \end{aligned}$$

$$g_{\alpha\beta,\mu} = \Gamma_{\mu\alpha}^\lambda g_{\lambda\beta} + \Gamma_{\mu\beta}^\lambda g_{\alpha\lambda} \quad 4.74$$

Cyclically rotating the three indexes α, β, μ we can write two other similar equations for $g_{\beta\mu,\alpha} = \dots$ and $g_{\mu\alpha,\beta} = \dots$

By summing member to member: 1st – 2nd + 3rd equation and using the symmetry properties of the lower indexes eq.4.72 (which applies in a

coordinate basis) we get:

$$g_{\alpha\beta,\mu} - g_{\beta\mu,\alpha} + g_{\mu\alpha,\beta} = 2g_{\lambda\alpha}\Gamma_{\mu\beta}^{\lambda}$$

Multiplying both members by $g^{\kappa\alpha}$ and since $g^{\kappa\alpha}g_{\lambda\alpha} = \delta_{\lambda}^{\kappa}$: *

$$\Gamma_{\mu\beta}^{\kappa} = \frac{1}{2}g^{\kappa\alpha}(g_{\alpha\beta,\mu} - g_{\beta\mu,\alpha} + g_{\mu\alpha,\beta}) \quad , \text{ c.v.}$$

This result has as important consequence:**

$$\forall g_{\alpha\beta} = \text{constant} \quad \Leftrightarrow \quad \forall \Gamma = 0 \quad 4.75$$

If the components $g_{\alpha\beta}$ of the metric tensor \mathbf{g} are constant, all the coefficients Γ are zero, and vice versa.

This is especially true in Cartesian coordinate systems.

In general we will qualify (incorrectly) “flat” a coordinate system where $\forall \Gamma = 0$. ***

4.19 Covariant derivative and invariance of tensor equations

A tensor equation contains only tensors, thus it may contain covariant derivatives but not ordinary derivatives (which are not tensors). For instance, $\tilde{\nabla}\mathbf{g} = 0$ or its componentwise equivalent $\nabla_{\mu}g_{\alpha\beta} = 0$ (which actually stands for n^3 equations) are tensor equations; it is *not* the case of eq.4.73.

A strategy often used to obtain tensor equations is called “*comma goes to semi-colon*” and exploits the fact that in a “flat” coordinate system, since all Γ are zero, the ordinary derivative and the covariant derivative coincide. In practice:

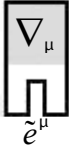
Working in a “flat” coordinate system (e.g., Cartesian), a partial derivatives equation inferred in this system can be turned into a tensor equation simply by replacing partial derivatives with covariant derivatives, namely by replacing the commas with semi-colons. The tensor equation thus obtained is invariant and applies in any reference accessible via coordinate transformation.

* From eq.2.28, since \mathbf{g} and \mathbf{g}^{-1} are inverse tensors.

** Which is already implicit in eq.2.39 and eq.4.31.

***In fact, it is the manifold to be (or not) flat. But: flat manifold $\Leftrightarrow \exists$ a coordinate system where $\forall \Gamma = 0$, as we shall see later.

4.20 T-mosaic representation of gradient, divergence and covariant derivative

$\tilde{\nabla}$ is the covector  that applies to scalars, vectors or tensors.

▪ Gradient of scalar: $\tilde{\nabla} f = \begin{array}{|c|c|} \hline \nabla_\mu & f \\ \hline \tilde{e}^\mu & \\ \hline \end{array} = \begin{array}{|c|} \hline \nabla_\mu f \\ \hline \tilde{e}^\mu \\ \hline \end{array}$

▪ Gradient of vector: $\tilde{\nabla} \otimes \vec{V} = \begin{array}{|c|c|} \hline \nabla_\mu & V^\nu \\ \hline \tilde{e}^\mu & \vec{e}_\nu \\ \hline \end{array} = \begin{array}{|c|} \hline \nabla_\mu V^\nu \\ \hline \tilde{e}^\mu \\ \hline \end{array} = \underbrace{\nabla_\mu V^\nu}_{\text{covariant derivative}} \tilde{e}^\mu \vec{e}_\nu$

▪ Gradient of covector: $\tilde{\nabla} \otimes \tilde{P} = \begin{array}{|c|c|} \hline \nabla_\mu & P_\nu \\ \hline \tilde{e}^\mu & \tilde{e}^\nu \\ \hline \end{array} = \begin{array}{|c|} \hline \nabla_\mu P_\nu \\ \hline \tilde{e}^\mu \tilde{e}^\nu \\ \hline \end{array} = \underbrace{\nabla_\mu P_\nu}_{\text{covariant derivative}} \tilde{e}^\mu \tilde{e}^\nu$

▪ Gradient of tensor,

e.g. \mathbf{g} : $\tilde{\nabla} \otimes \mathbf{g} = \begin{array}{|c|c|} \hline \nabla_\lambda & g_{\mu\nu} \\ \hline \tilde{e}^\lambda & \tilde{e}^\mu \tilde{e}^\nu \\ \hline \end{array} = \begin{array}{|c|} \hline \nabla_\lambda g_{\mu\nu} \\ \hline \tilde{e}^\lambda \tilde{e}^\mu \tilde{e}^\nu \\ \hline \end{array}$

- Divergence of vector:

$$\text{div } \vec{A} = \tilde{\nabla}(\vec{A}) = \begin{array}{c} \boxed{\nabla_\mu} \\ \text{---} \\ \tilde{e}^\mu \\ \text{---} \\ \boxed{A^\mu} \end{array} \rightarrow \boxed{\begin{array}{c} \nabla_\mu \\ \text{---} \\ A^\mu \end{array}} = \boxed{\nabla_\mu A^\mu}$$

- Divergence of tensor $\mathbf{T} = T_\lambda^{\nu\mu} \vec{e}_\nu \vec{e}_\mu \tilde{e}^\lambda$ (with respect to index μ):

$$\tilde{\nabla}(\mathbf{T}) = \begin{array}{c} \boxed{\nabla_\mu} \\ \text{---} \\ \tilde{e}^\mu \\ \text{---} \\ \begin{array}{c} \vec{e}_\nu \quad \vec{e}_\mu \\ \text{---} \\ \boxed{T_\lambda^{\nu\mu}} \\ \text{---} \\ \tilde{e}^\lambda \end{array} \end{array} \rightarrow \begin{array}{c} \vec{e}_\nu \quad \boxed{\nabla_\mu} \\ \text{---} \\ \boxed{T_\lambda^{\nu\mu}} \\ \text{---} \\ \tilde{e}^\lambda \end{array} = \begin{array}{c} \vec{e}_\nu \\ \text{---} \\ \boxed{\nabla_\mu T_\lambda^{\nu\mu}} \\ \text{---} \\ \tilde{e}^\lambda \end{array}$$

5 Curved manifolds

5.1 Symptoms of curvature

Consider a spherical surface within the usual 3D Euclidean space. The spherical surface is a manifold 2D.

A 3D observer, or \mathcal{O}_{3D} , which has an outside view, sees the curvature of the spherical surface in the third dimension. A 2D observer \mathcal{O}_{2D} , who lives on the spherical surface has no outside view. \mathcal{O}_{3D} sees the light rays propagate on the spherical surface along maximum circle arcs; for \mathcal{O}_{2D} they will be straight lines, by definition (they are the only possible inertial trajectories, run in absence of forces). How can \mathcal{O}_{2D} realize that his space is curved?

By studying the geometry in its local context \mathcal{O}_{2D} will discover the common properties of the 2D Euclidean geometry, the usual plane geometry. For example, he finds that the ratio of the circumference to diameter is a constant $= \pi$ for all circles and that the sum of the inner angles of any triangle is 180° . In fact, in its environment, i.e. in the small portion of the spherical surface on which he lives, his two-dimensional space is locally flat, as for us is flat the surface of a water pool on the Earth's surface (the observer \mathcal{O}_{3D} sees this neighborhood practically lying on the plane tangent to the spherical surface in the point where \mathcal{O}_{2D} is located).

However, when \mathcal{O}_{2D} comes to consider very large circles, he realizes that the circumference / diameter ratio is no more a constant and it is smaller and smaller than π as the circle enlarges, and that the sum of the inner angles of a triangle is variable but always $> 180^\circ$.

\mathcal{O}_{2D} can deduce from these facts that his space is curved, though not capable of understanding the third dimension.

The curvature of a space is thus an intrinsic property of the space itself and there is no need of an outside view to describe it. In other words, it is not necessary for a n -dimensional space to be considered embedded in another $n + 1$ dimensional space to reveal the curvature.

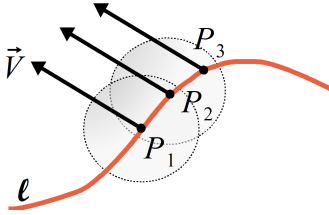
▪ Another circumstance that allows \mathcal{O}_{2D} to discover the curvature of his space is that, carrying a vector parallel to itself on a large enough closed loop, the returned vector is no longer parallel (= does not overlap) to the initial vector.

In a flat space such as the Euclidean one vectors can be transported

parallel to itself without difficulty, so that they can be often considered delocalized. For example, to measure the relative velocity of two particles far apart, we may imagine to carry the vector \vec{v}_2 parallel to itself from the second particle to the first so as to make their origins coincide in order to carry out the subtraction $\vec{v}_2 - \vec{v}_1$.

But what does it mean a parallel transport in a curved space?

In a curved space, this expression has no precise meaning. However, \mathcal{O}_{2D} can think to perform a parallel transport of a vector by a step-by-step strategy, taking advantage from the fact that in the neighborhood of each point the space looks substantially flat.



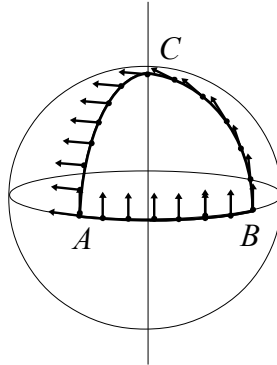
Let l be the curve along which we want to parallel-transport the (foot of the) vector \vec{V} . At first the observer \mathcal{O}_{2D} and the vector are in P_1 ; the vector is transported parallel to itself to P_2 , a point of the (infinitesimal) neighborhood of P_1 . Then \mathcal{O}_{2D} goes to P_2 and parallel transports the vector to P_3 that belongs to the neighborhood of P_2 , and so on. In this way the transport is always within a neighborhood in which the space is flat and the notion of parallel transport is unambiguous: vector in $P_1 \parallel$ vector in P_2 ; vector in $P_2 \parallel$ vector in P_3 , etc.

But the question is: is vector in $P_1 \parallel$ vector in P_n ? That is: does the parallelism which works locally step-by-step by infinitesimal amounts hold globally as well?

To answer this question we must transport the vector along a closed path: only if the vector that returns from parallel transport is superimposable onto the initial vector we can conclude for global parallelism.

In fact it is not so, at least for large enough circuits: just take for example the transport $A-B-C-A$ along the meridians and the equator in

the picture:



The mismatch between the initial vector and the one that returns after the parallel transport can give a measure of the degree of curvature of the space (this measure is fully accessible to the two-dimensional observer \mathcal{O}_{2D}).

To do so we must give first a more precise mathematical definition of parallel transport of a vector along a line.

5.2 Derivative of a scalar or vector along a line

Given a scalar or a vector defined (at least at the points of a parametric line) within the manifold, let us express the derivatives of the scalar or vector along the line as the incremental ratio of the scalar or vector with respect to the parameter.

The line ℓ is defined by n parametric equations $x^\mu(\tau)$. The parameter τ can be (or not) the arc length. At any point of the line let be defined a tangent vector (a unit vector only if τ is the arc):

$$\vec{U} = \frac{d\vec{x}}{d\tau} \quad 5.1$$

$$\Rightarrow U^\mu = \frac{dx^\mu}{d\tau} \quad \text{i.e.} \quad \vec{U} \xrightarrow{\text{comp}} \frac{dx^\mu}{d\tau} \quad \text{in a coordinate basis.}$$

① derivative of a scalar f along a line

From the total differential of f : $df = \frac{\partial f}{\partial x^\mu} dx^\mu$ we immediately deduce the definition of derivative along a line:

$$\frac{df}{d\tau} = \frac{\partial f}{\partial x^\mu} \frac{dx^\mu}{d\tau} \quad 5.2$$

but, since for a scalar f partial and covariant derivatives identify, *in a coordinate base* is:

$$\frac{df}{d\tau} = U^\mu \partial_\mu f \equiv U^\mu \nabla_\mu f = \langle \vec{\nabla} f, \vec{U} \rangle \quad 5.3$$

Roughly speaking: the derivative of f along the line can be seen as a “*projection*” or “*component*” of the gradient in direction of the tangent vector. It is a scalar.

- The derivative of a scalar along a line is a generalization of the partial ∂_μ (or covariant ∇_μ) derivative and reduces to it along coordinated lines.

- ◻ Indeed: along a coordinate line $x^{\bar{\mu}}$ it is $\vec{U} // \vec{e}_{\bar{\mu}} \Rightarrow U^{\bar{\mu}}$ is the only nonzero component; * the summation on μ implicit in $U^\mu \nabla_\mu f$ (eq.5.3) thus collapses to the single $\bar{\mu}$ -th term.

Moreover, if τ is the arc, then $d\tau \equiv dx^{\bar{\mu}} \Rightarrow U^{\bar{\mu}} = \frac{dx^{\bar{\mu}}}{d\tau} = 1$.

For these reasons, eq.5.3 reduces to:

$$\frac{df}{d\tau} = U^{\bar{\mu}} \nabla_{\bar{\mu}} f [\text{no sum}] = \nabla_{\bar{\mu}} f \quad ** \quad 5.4$$

q.e.d.

- Because of its relationship with the covariant derivative and the gradient, the derivative of a scalar along a line is also denoted with:

$$\frac{df}{d\tau} \equiv \nabla_{\vec{U}} f \quad 5.5$$

- Eq.5.3 and eq.5.5 enable us to write symbolically:

$$\frac{d}{d\tau} = \nabla_{\vec{U}} = U^\mu \partial_\mu = U^\mu \nabla_\mu \quad 5.6$$

* Note that along the coordinate line $x^{\bar{\mu}}$, provided τ is the arc:

$$\vec{U} \stackrel{\text{def}}{=} \frac{d\vec{x}}{d\tau} = \frac{\vec{e}_\mu dx^\mu}{d\tau} = \frac{\vec{e}_{\bar{\mu}} dx^{\bar{\mu}}}{dx^{\bar{\mu}}} [\text{no sum}] = \vec{e}_{\bar{\mu}}$$

** The indication [no sum] offs the summation convention for repeated indexes in the case.

② derivative of a vector \vec{V} along a line

In analogy with the result just found for the scalar f (eq.5.3) we define the derivative of a vector \vec{V} along a line ℓ with tangent vector \vec{U} as the “projection” or “component” of the gradient on the tangent vector:

$$\frac{d\vec{V}}{d\tau} \stackrel{\text{def}}{=} \langle \tilde{\nabla} \vec{V}, \vec{U} \rangle \quad 5.7$$

The derivative of a vector along a line is a vector, since $\tilde{\nabla} \vec{V}$ is a $\binom{1}{1}$ tensor. Expanding on the bases:

$$\begin{aligned} \frac{d\vec{V}}{d\tau} &= \langle \tilde{\nabla} \vec{V}, \vec{U} \rangle = \langle \nabla_\mu V^\nu \tilde{e}^\mu \tilde{e}_\nu, U^\kappa \tilde{e}_\kappa \rangle = \nabla_\mu V^\nu U^\kappa \underbrace{\langle \tilde{e}^\mu, \tilde{e}_\kappa \rangle}_{\delta_\kappa^\mu} \tilde{e}_\nu = \\ &= \underbrace{U^\mu \nabla_\mu V^\nu}_{\stackrel{\text{def}}{=} \frac{DV^\nu}{d\tau}} \tilde{e}_\nu = \frac{DV^\nu}{d\tau} \tilde{e}_\nu \quad 5.8 \end{aligned}$$

Eq.5.8 can be written:
$$\frac{d\vec{V}}{d\tau} \xrightarrow{\text{comp}} \frac{DV^\nu}{d\tau} \quad 5.9$$

after set:
$$\frac{DV^\nu}{d\tau} \stackrel{\text{def}}{=} U^\mu \nabla_\mu V^\nu \quad 5.10$$

$\frac{DV^\nu}{d\tau}$ is the ν -component of the derivative of vector $\frac{d\vec{V}}{d\tau}$ and is referred to as *covariant derivative along line*, expanded:

$$\begin{aligned} \frac{DV^\nu}{d\tau} &\stackrel{\text{def}}{=} U^\mu \nabla_\mu V^\nu = \underbrace{U^\mu}_{\equiv \frac{dx^\mu}{d\tau}} \left(\frac{\partial V^\nu}{\partial x^\mu} + \Gamma_{\mu\lambda}^\nu V^\lambda \right) = \frac{\partial V^\nu}{\partial x^\mu} \frac{dx^\mu}{d\tau} + \Gamma_{\mu\lambda}^\nu V^\lambda U^\mu = \\ &= \frac{dV^\nu}{d\tau} + \Gamma_{\mu\lambda}^\nu V^\lambda U^\mu \quad 5.11 \end{aligned}$$

A new derivative symbol has been introduced since the components of $\frac{d\vec{V}}{d\tau}$ are *not* simply $\frac{dV^\nu}{d\tau}$, but have the more complex form eq.5.11.

▪ Another notation, the most widely used in practice, for the derivative of a vector along a line is the following:

$$\frac{d\vec{V}}{d\tau} \equiv \nabla_{\vec{U}} \vec{V} \quad 5.12$$

(that recalls the definition “projection of the gradient $\nabla \vec{V}$ on \vec{U} ”).

▪ Let us recapitulate the various notations for the derivative of a vector along a line. Beside the definition *eq.5.7* $\frac{d\vec{V}}{d\tau} \stackrel{\text{def}}{=} \langle \tilde{\nabla} \vec{V}, \vec{U} \rangle$ were appended *eq.5.12*, *eq.5.9*, *eq.5.10*, summarized in:

$$\frac{d\vec{V}}{d\tau} \equiv \nabla_{\vec{U}} \vec{V} \xrightarrow{\text{comp}} \frac{DV^\nu}{d\tau} \equiv U^\mu \nabla_\mu V^\nu \quad 5.13$$

or, symbolically:

$$\frac{d}{d\tau} \equiv \nabla_{\vec{U}} \xrightarrow{\text{comp}} \frac{D}{d\tau} \equiv U^\mu \nabla_\mu \quad 5.14$$

▪ The relationship between the derivative of a vector along a line and the covariant derivative arises when the line ℓ is a coordinate line $x^{\bar{\mu}}$.

▫ In this case $U^{\bar{\mu}} = 1$ provided τ is the arc and the summation on $\bar{\mu}$ collapses, as already seen (*eq.5.4*) for the derivative of a scalar; then the chain *eq.5.8* changes into:

$$\frac{d\vec{V}}{d\tau} = U^{\bar{\mu}} \nabla_{\bar{\mu}} V^\nu \vec{e}_\nu = \nabla_{\bar{\mu}} V^\nu \vec{e}_\nu \quad 5.15$$

namely:
$$\frac{d\vec{V}}{d\tau} \xrightarrow{\text{comp}} \nabla_{\bar{\mu}} V^\nu$$

The derivative of a vector along a coordinate line $x^{\bar{\mu}}$ has as components the n covariant derivatives corresponding to the blocked index $\bar{\mu}$.

▪ Comparing the two cases of derivative along a line for a scalar and a vector we see (*eq.5.6*, *eq.5.14*) that in both cases the derivative is designated by $\frac{d}{d\tau} = \nabla_{\vec{U}}$, but the operator $\nabla_{\vec{U}}$ has a different meaning in the two cases:

- $\nabla_{\vec{U}} = U^\mu \nabla_\mu$ for a scalar
- $\nabla_{\vec{U}} \xrightarrow{\text{comp}} U^\mu \nabla_\mu$ for a vector.

▪ The derivative along a line of scalars and vectors are sometimes called “absolute derivatives” or (more properly) “*directional covariant derivatives*”.

5.3 T-mosaic representation of derivatives along a line

① Derivative along a line of a scalar:

$$\begin{aligned}
 \frac{df}{d\tau} = \langle \tilde{\nabla} f, \vec{U} \rangle &= \begin{array}{c} \boxed{\nabla_\mu f} \\ \text{---} \\ \tilde{e}^\mu \\ \text{---} \\ \tilde{e}_\mu \\ \text{---} \\ \boxed{U^\mu} \end{array} \rightarrow \begin{array}{c} \boxed{\nabla_\mu f} \\ \text{---} \\ \boxed{U^\mu} \end{array} = \boxed{U^\mu \nabla_\mu f} = \\
 &= \underbrace{\boxed{\nabla_{\vec{U}} f}}_{\text{(scalar)}}
 \end{aligned}$$

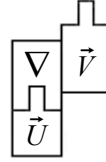
② Derivative along a line of a vector:

$$\begin{aligned}
 \frac{d\vec{V}}{d\tau} = \langle \tilde{\nabla} \vec{V}, \vec{U} \rangle &= \begin{array}{c} \tilde{e}_\nu \\ \text{---} \\ \boxed{\nabla_\mu} \quad \boxed{V^\nu} \\ \text{---} \\ \tilde{e}^\mu \\ \text{---} \\ \tilde{e}_\mu \\ \text{---} \\ \boxed{U^\mu} \end{array} \rightarrow \begin{array}{c} \tilde{e}_\nu \\ \text{---} \\ \boxed{\nabla_\mu V^\nu} \\ \text{---} \\ \tilde{e}^\mu \\ \text{---} \\ \tilde{e}_\mu \\ \text{---} \\ \boxed{U^\mu} \end{array} = \begin{array}{c} \tilde{e}_\nu \\ \text{---} \\ \boxed{U^\mu \nabla_\mu V^\nu} \end{array} = \\
 &= \begin{array}{c} \tilde{e}_\nu \\ \text{---} \\ \boxed{\frac{DV^\nu}{d\tau}} \end{array} \equiv \underbrace{\nabla_{\vec{U}} \vec{V}}_{\text{(vector)}}
 \end{aligned}$$

Mnemo

Derivative along a line of a vector: the previous block diagram can help recalling the right formulas.

In particular, writing $\nabla_{\vec{v}} \vec{V}$ recalls the disposition of the symbols in the block:



Also note that:

- $\frac{d\vec{V}}{d\tau} \equiv \langle \vec{\nabla} \vec{V}, \vec{U} \rangle \equiv \nabla_{\vec{v}} \vec{V}$ contain the sign \rightarrow and are vectors
- $\frac{DV^\nu}{d\tau} \equiv U^\mu \nabla_\mu V^\nu$ are scalar components (without sign \rightarrow)

If a single thing is to keep in mind choose this:

- directional covariant derivative $\equiv U^\mu \nabla_\mu$: applies to both a scalar and vector components

5.4 Parallel transport along a line of a vector

Let \vec{V} a vector defined in each point of a parametric line $x^\mu(\tau)$.

Moving from the point $P(\tau)$ to the point $P(\tau + \Delta\tau)$ the vector undergoes an increase:

$$\vec{V}(\tau + \Delta\tau) = \vec{V}(\tau) + \frac{d\vec{V}}{d\tau}(\tau) \Delta\tau + O(\Delta\tau)^2 \quad * \quad 5.16$$

In case the first degree term is missing, namely:

$$\vec{V}(\tau + \Delta\tau) = \vec{V}(\tau) + O(\Delta\tau)^2 \quad \Leftrightarrow \quad \frac{d\vec{V}}{d\tau} = 0 \quad 5.17$$

we say that the vector \vec{V} is **parallel transported**.

Roughly: transporting its origin along the line for an amount $\Delta\tau$ the vector undergoes only a variation of the 2nd order (and thus keeps itself almost unchanged).

The parallel transport of a vector \vec{V} is defined locally, in its own flat local system: over a finite length the parallelism between initial and transported vector is no longer maintained .

- In each point of the line let a tangent vector $\vec{U} = \frac{d\vec{x}}{d\tau}$ be defined.

* We denote by $O(x)$ a quantity of the same order of x . In this case it is matter of a quantity of the same order of $(\Delta\tau)^2$, i.e. of the second order with respect to $\Delta\tau$.

The *parallel transport condition of \vec{V} along \vec{U}* can be expressed by one of the following forms:

$$\nabla_{\vec{U}} \vec{V} = 0 \quad , \quad \langle \tilde{\nabla} \vec{V}, \vec{U} \rangle = 0 \quad , \quad U^\mu \nabla_\mu V^\nu = 0 \quad 5.18$$

all equivalent to the early definition because (eq.5.7, eq.5.13):

$$\nabla_{\vec{U}} \vec{V} = \frac{d\vec{V}}{d\tau} = \langle \tilde{\nabla} \vec{V}, \vec{U} \rangle \xrightarrow{comp} U^\mu \nabla_\mu V^\nu .$$

The vector \vec{V} to be transported can have any orientation with respect to \vec{U} . Given a regular line $x^\mu(\tau)$ and a vector \vec{V} defined in one of its points, it is always possible to parallel transport the vector along the line.

5.5 Geodesics

When it happens that the tangent vector \vec{U} is transported parallel to itself along a line, the line is a *geodesic*.

Geodesic condition for a line is then:

$$\frac{d\vec{U}}{d\tau} = 0 \quad \text{along the line} \quad 5.19$$

Equivalent to eq.5.19 are the following statements:

$$\nabla_{\vec{U}} \vec{U} = 0 \quad , \quad \langle \tilde{\nabla} \vec{U}, \vec{U} \rangle = 0 \quad 5.20$$

or componentwise:

$$5.21$$

$$U^\mu \nabla_\mu U^\nu = 0 \quad , \quad \frac{dU^\nu}{d\tau} + \Gamma_{\mu\lambda}^\nu U^\lambda U^\mu = 0 \quad , \quad \frac{d^2 x^\nu}{d\tau^2} + \Gamma_{\mu\lambda}^\nu \frac{dx^\lambda}{d\tau} \frac{dx^\mu}{d\tau} = 0$$

which come respectively from eq.5.12, eq.5.7, eq.5.10, eq.5.11 (and eq.5.1). All the eq.5.19, eq.5.20, eq.5.21 are equivalent and represent the *equations of the geodesic*.

Some properties of geodesics are the following:

- Along a geodesic the tangent vector is constant in magnitude.*

$$\begin{aligned} \square \text{ Indeed: } \frac{d}{d\tau} (g_{\alpha\beta} U^\alpha U^\beta) &= U^\mu \nabla_\mu (g_{\alpha\beta} U^\alpha U^\beta) = \\ &= U^\mu (g_{\alpha\beta} U^\alpha \nabla_\mu U^\beta + g_{\alpha\beta} U^\beta \nabla_\mu U^\alpha + U^\alpha U^\beta \nabla_\mu g_{\alpha\beta}) = 0 \\ \text{because } U^\mu \nabla_\mu U^\beta &= U^\mu \nabla_\mu U^\alpha = 0 \text{ is the parallel transport} \end{aligned}$$

* The converse is not true: the constancy in magnitude of the tangent vector is not a sufficient condition for a line to be a geodesic.

condition of \vec{U} along the geodesic and $\nabla_{\mu} g_{\alpha\beta} = 0$ (eq.4.67).

Hence: $g_{\alpha\beta} U^{\alpha} U^{\beta} = const \Rightarrow \mathbf{g}(\vec{U}, \vec{U}) \stackrel{\text{def}}{=} |\vec{U}|^2 = const$, q.e.d.

- A geodesic parameterized by a parameter τ is still a geodesic when re-parameterized * with an “affine” parameter s like $s = a + b\tau$ (two parameters are called *affine* if linked by a linear relationship).

- The geodesics are the “straightest possible lines” in a curved space, and can be thought of as inertial trajectories of particles not subject to external forces. Under these conditions a particle transports its velocity vector parallel to itself (constant in magnitude and tangent to the path). A curve must have precise characteristics to be a geodesic: it must be a path that requires no acceleration. For example on a 2D spherical surface the geodesics are the greatest circles, like the meridians and the equator; not so the parallels: they can be traveled only if a constant acceleration directed along the meridian is in action.

- We add some comments to clarify a situation that is not immediately intuitive.

An aircraft flying along a parallel with speed \vec{v} tangent to the parallel itself does *not* parallel-transport this vector along its path. A question arises about how would evolve a vector \vec{E} at first coincident with \vec{v} , but for which we require to be parallel transported all along the trip.

For this purpose let's imagine to cut from the spherical surface a thin ribbon containing the path, in this case the whole Earth's parallel. This ribbon can be flattened and layed down on a plane; on this plane the path is an arc of a circumference (whose radius is the generatrix of the cone whose axis coincides with the Earth's axis and is tangent to the spherical surface on the parallel); the length of the arc will be that of the Earth's parallel. Let's now parallel-transport on the plane the vector \vec{E}_0 initially tangent to the arc of circumference until the end of the flattened arc (working on a plan, there is no ambiguity about the meaning of parallel transport). Let \vec{E}_{fin} be the vector at the end of the transport. Now we imagine bringing back to its original place on

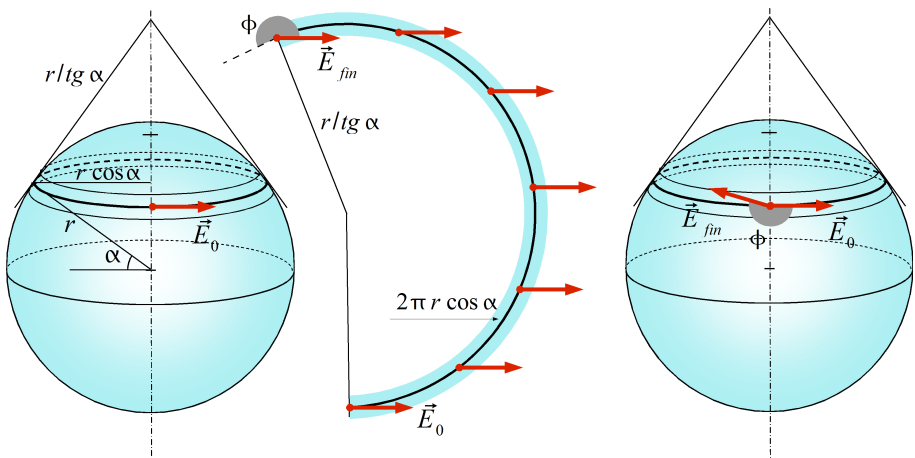
* A line may be parametrized in more than one way. For example, a parabola can be parametrized by $x = t; y = t^2$ or $x = t^3; y = t^6$.

the spherical surface the ribbon with the various subsequent positions taken by the vector \vec{E} drawn on. In that way we achieve a graphical representation of the evolution of the vector \vec{E} parallel transported along the Earth's parallel.

The steps of the procedure are explained in the figure below. After having traveled a full circle of parallel transport on the Earth's parallel of latitude α the vector \vec{E} has undergone a clockwise rotation by an angle $\phi = \frac{2\pi r \cos \alpha}{r / \tan \alpha} = 2\pi \sin \alpha$, from

\vec{E}_0 to \vec{E}_{fin} .

Only for $\alpha = 0$ (along the equator) we get $\phi = 0$ after a turn.



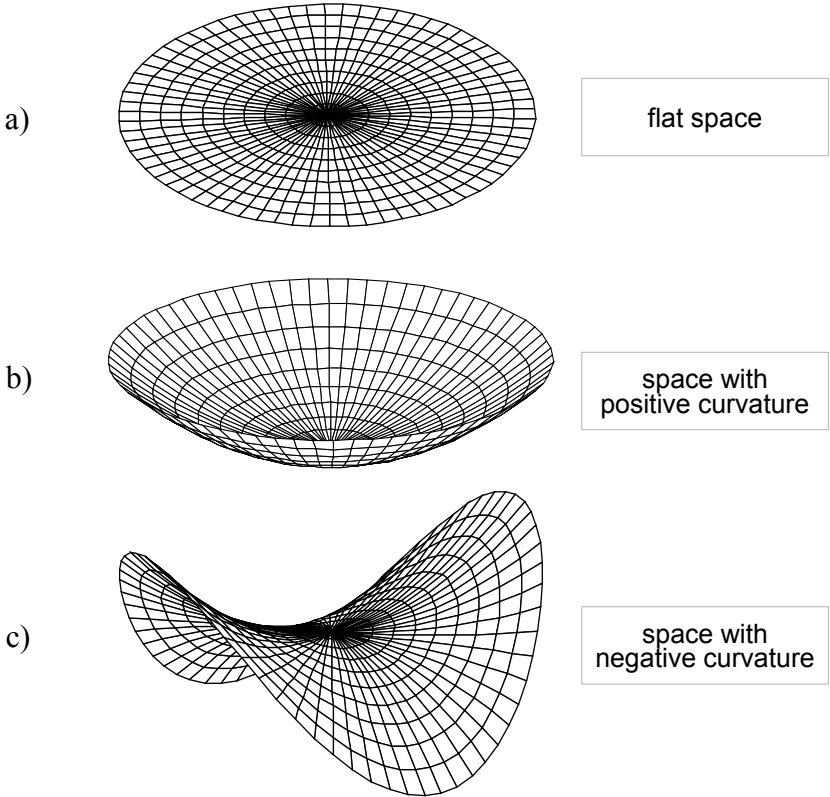
The trick to flatten on the plane the tape containing the trajectory works whatever is the initial orientation of the vector you want to parallel transport (along the “flattened” path the vector may be moved parallel to itself even if it is out of the plane).

5.6 Positive and negative curvature

We have so far considered the case of a 2D space “spherical surface” with constant curvature only. There are, however, other possible types of curved spaces and other types of curvature. In fact, the curvature

can be *positive* (e.g. a spherical space) or *negative* (hyperbolic space). In a space with positive curvature the sum of the internal angles of a triangle is $>180^\circ$, the ratio of circumference to diameter is $< \pi$ and two parallel lines end up with meeting each other; in a space with negative curvature the sum of the internal angles of a triangle is $<180^\circ$, the ratio of circumference to diameter is $> \pi$ and two parallel lines end up to diverge.

▪ A space with negative curvature or hyperbolic curvature is less easily representable than a space with positive curvature or spherical. An intuitive approach is to imagine a thin flat metallic plate (a) that is unevenly heated. The thermal dilatation will be greater in the points where the higher is the temperature.



If the plate is heated more in the central part the thermal dilatation will be greater at the center and the foil will sag dome-shaped (b).

If the plate is more heated in the periphery, the thermal dilation will be greater in the outer band and the plate will “embark” assuming a wavy shape (c).

The simplest case of negative curvature (c) is shown, in which the curved surface has the shape of a hyperbolic paraboloid (the form often taken by fried potato chips!).

5.7 Flat and curved manifold

The metric tensor \mathbf{g} completely characterizes the geometry of the manifold, in particular its being flat or curved.* Now, either \mathbf{g} is constant all over the manifold (i.e. unvarying) or varies with the point P . By definition:

$$\mathbf{g} = \mathbf{g}(P) \quad \Leftrightarrow \quad \text{curved manifold} \quad 5.22$$

and conversely:

$$\mathbf{g} \text{ constant} \quad \Leftrightarrow \quad \text{flat manifold} \quad 5.23$$

Since \mathbf{g} is a tensor, it is independent of the coordinate system, even if its representation by components $g_{\mu\nu}$ depends upon. Then \mathbf{g} may be constant all over the manifold despite in *some* system its components $g_{\mu\nu}$ are P -dependent,** only if $g_{\mu\nu}$ are P -dependent in *all* systems we are sure that \mathbf{g} is not constant. Only in that case we can say that \mathbf{g} is P -dependent, or $\mathbf{g} = \mathbf{g}(P)$.

According to what said above, componentwise translations of eq.5.22, eq.5.23, are respectively:

- if *some* $g_{\mu\nu}$ is P -dependent in *all* coordinate systems, then the manifold is curved (and viceversa)
- if *all* $g_{\mu\nu}$ are constant in *some* coordinate system, then the manifold is flat (and viceversa)

What is discriminating is the existence of a reference system where all $g_{\mu\nu}$ are independent of the point P , that is to say that all elements of the matrix $[g_{\mu\nu}]$ are numerical constants. If that system exists, the manifold is flat, otherwise it is curved.

* It will be seen later on apropos of Riemann's tensor.

Here and below we consider only *positive definite* metrics (for indefinite metrics eq.5.22 applies only to the left \Leftarrow and eq.5.23 only to the right \Rightarrow).

** It is useful to keep in mind as a good example the polar coordinates (ρ, θ) , where $g_{\theta\theta} = 1/\rho^2$ despite the flatness of the plane.

In practice, given a $g_{\mu\nu}$ P -dependent, it is not easy to determine if the manifold is flat or curved: you should find a coordinate system where all $g_{\mu\nu}$ = numerical constants, or exclude the possibility to find one.

- Note that in a flat manifold $\forall \Gamma = 0$ due to eq.4.73.
- Due to a theorem of matrix calculus, any symmetric matrix * (such as $[g_{\mu\nu}]$ with constant coefficients) can be put into the canonical form $diag(\pm 1)$ by means of an other appropriate matrix Λ .
Remind that the matrix Λ is related to a certain coordinate transformation (it is *not* the transformation, but it comes down!). In fact, to reduce the metric to canonical form, we have to apply a coordinate transformation to the manifold so that (the transposed inverse of) its Jacobian matrix Λ_{α}^{μ} , makes the metric canonical:

$$\Lambda_{\alpha}^{\mu} \Lambda_{\beta}^{\nu} g_{\mu\nu} = g_{\alpha'\beta'} = diag(\pm 1) \tag{5.24}$$

For a flat space, being the metric independent of the point, the matrix Λ holds globally, for the manifold as a whole. Hence, a flat manifold admits a coordinate system in which the metric tensor at each point has the canonical form

$$[g_{\mu\nu}] = diag(\pm 1) \tag{5.25}$$

which can be achieved by a proper matrix Λ valid for the whole manifold.

This property can be taken as a definition of flatness:

Is flat a manifold that admits a coordinate system where $g_{\mu\nu}$ can be everywhere expressed in canonical form $diag(\pm 1)$

In particular, if $[g_{\mu\nu}] = diag(+1)$ the manifold is the Euclidean space; if $[g_{\mu\nu}] = diag(\pm 1)$ with a single -1 the manifold is a Minkowski space. The occurrences of $+1$ and -1 in the main diagonal, or their sum called “signature”, is a characteristic of the given space: transforming the coordinates the various $+1$ and -1 can interchange their places on the diagonal but do not change in number (*Sylvester's theorem*).

Hereafter by $\eta_{\mu\nu}$ we will refer to the canonical metric $diag(\pm 1)$, regardless of the number of $+1$ and -1 in the main diagonal.

Conversely, for curved manifolds there is no coordinate system in

* We mean here numeric matrices.

which the metric tensor has the canonical form $diag(\pm 1)$ on the whole manifold. Only *punctually* we can do this: given any point P , we can turn the metric to the canonical form in this point by means of a certain matrix Λ , but that works only for that point (as well as the coordinate transformation from which it comes). The coordinate transformation that has led to the matrix Λ good for P does not work for all other points of the manifold: to put the metric into canonical form here too, you have to use other transformations and other matrices.

5.8 Flat local system

Indeed, for each point of a curved manifold, we can do better: not only putting the metric punctually in canonical form, but also making its first derivatives to vanish at that point.

In that way the (canonical) metric is stationary * in the point; that is, it is valid, except for infinitesimals of higher order, also *locally* in the neighborhood of the point. This ensures that the metric is flat in the neighborhood of the point (a metric that is the canonical $\eta_{\mu\nu}$ in a single point is not enough to guarantee the flatness: it is necessary that this metric remains stationary in the neighborhood. In other words, the flatness is a concept that cannot be defined punctually, but locally in a neighborhood).

All that is formalized in the following important theorem.

5.9 Local flatness theorem

Given a differentiable manifold and any point P in it, it is possible to find a coordinate system that reduces the metric to the form:

$$g_{\mu\nu}(P') = \eta_{\mu\nu}(P) + O[(x^\kappa - x_0^\kappa)^2] \quad 5.26$$

where $P'(x^\kappa)$ is any point of the neighborhood of $P(x_0^\kappa)$. ** The term $O[\dots]$ is for an infinitesimal of the second order with respect to the displacement from $P(x_0^\kappa)$ to $P'(x^\kappa)$, that we have improperly denoted for simplicity by $(x^\kappa - x_0^\kappa)$.

* The metric is in any case stationary in each point: $\tilde{\nabla} \mathbf{g} = 0$, eq. 4.67. Analogy with a function $f(x)$ that is stationary in a maximum or minimum. In those points $f' = 0$ and for small shifts to the right or left $f(x)$ varies only for infinitesimals of the second order with respect to the shift.

** See note on eq. 5.16.

The above equation is merely a Taylor series expansion around the point P , missing of the first order term $(x^k - x_0^k) \frac{\partial g_{\mu\nu}}{\partial x^k}(P)$ because the first derivative is supposed to be null in P .

We explicitly observe that the neighborhood of P containing P' belongs to the manifold, *not* to the tangent space. On the other hand $\eta_{\mu\nu}$ is the (global) flat metric of the tangent space in P .

Roughly, the theorem means that the space is locally flat in the neighborhood of each point and there it can be confused with the tangent space to less than a second-order infinitesimal.

The analogy is with the Earth's surface, that can be considered locally flat in the neighborhood of every point.

- The proof of the theorem is a semi-constructive one: it can be shown that there is a coordinate transformation whose related matrix makes the metric canonical in P with all its first derivatives null; then we'll try to rebuild the coordinate transformation and its related matrix (or at least the initial terms of the series expansions of its inverse).

We report here a scheme of proof.

- Given a manifold with a coordinate system x^α and a point P in it we transform to new coordinates $x^{\mu'}$ to take the metric into canonical form in P with zero first derivatives.

At any point P' of the P -neighborhood it is:

$$g_{\mu'\nu'}(P') = \Lambda_{\mu'}^\alpha \Lambda_{\nu'}^\beta g_{\alpha\beta}(P') \tag{5.27}$$

(denoted $\Lambda_{\alpha'}^{\mu'}$ the transformation matrix Λ (see eq.3.7, eq.3.8), to transform $g_{\alpha\beta}$ we have to use twice its transposed inverse).

In the latter both g and Λ are meant to be calculated in P' .

In a differentiable manifold, what happens in P' can be approximated by what happens in P and a Taylor series. We expand in Taylor series around P both members of eq.5.27, first the left: *

$$g_{\mu'\nu'}(P') = g_{\mu'\nu'}(P) + (x^{y'} - x_0^{y'}) g_{\mu'\nu',y'}(P) + \frac{1}{2} (x^{y'} - x_0^{y'}) (x^{\lambda'} - x_0^{\lambda'}) g_{\mu'\nu',y'\lambda'}(P) + \dots \tag{5.28}$$

and then the right one:

* Henceforth we will often use the comma , to denote partial derivatives.

$$\begin{aligned} \Lambda_{\mu}^{\alpha}, \Lambda_{\nu}^{\beta}, g_{\alpha\beta}(P') &= \Lambda_{\mu}^{\alpha}, \Lambda_{\nu}^{\beta}, g_{\alpha\beta}(P) + (x^{y'} - x_0^{y'}) \frac{\partial}{\partial x^{y'}} (\Lambda_{\mu}^{\alpha}, \Lambda_{\nu}^{\beta}, g_{\alpha\beta})(P) + \\ &+ \frac{1}{2} (x^{y'} - x_0^{y'}) (x^{\lambda'} - x_0^{\lambda'}) \frac{\partial^2}{\partial x^{y'} \partial x^{\lambda'}} (\Lambda_{\mu}^{\alpha}, \Lambda_{\nu}^{\beta}, g_{\alpha\beta})(P) + \dots \end{aligned} \quad 5.29$$

Now let's equal order by order the right terms of *eq.5.28* and *eq.5.29*:

$$I) \quad g_{\mu',\nu'}(P) = \Lambda_{\mu}^{\alpha}, \Lambda_{\nu}^{\beta}, g_{\alpha\beta}(P) \quad 5.30$$

$$II) \quad g_{\mu',\nu',\gamma'}(P) = \frac{\partial}{\partial x^{y'}} (\Lambda_{\mu}^{\alpha}, \Lambda_{\nu}^{\beta}, g_{\alpha\beta})(P) \quad 5.31$$

$$III) \quad g_{\mu',\nu',\gamma'\lambda'}(P) = \frac{\partial^2}{\partial x^{y'} \partial x^{\lambda'}} (\Lambda_{\mu}^{\alpha}, \Lambda_{\nu}^{\beta}, g_{\alpha\beta})(P) \quad 5.32$$

Hereafter all terms are calculated in P and for that we omit this specification below.

Recalling that $\Lambda_{\mu}^{\alpha} = \frac{\partial x^{\alpha}}{\partial x^{\mu'}}$ and carrying out the derivatives of product we get:

$$I) \quad g_{\mu',\nu'} = \frac{\partial x^{\alpha}}{\partial x^{\mu'}} \frac{\partial x^{\beta}}{\partial x^{\nu'}} g_{\alpha\beta} \quad 5.33$$

$$II) \quad g_{\mu',\nu',\gamma'} = \frac{\partial^2 x^{\alpha}}{\partial x^{y'} \partial x^{\mu'}} \frac{\partial x^{\beta}}{\partial x^{\nu'}} g_{\alpha\beta} + \frac{\partial x^{\alpha}}{\partial x^{\mu'}} \frac{\partial^2 x^{\beta}}{\partial x^{y'} \partial x^{\nu'}} g_{\alpha\beta} + \frac{\partial x^{\alpha}}{\partial x^{\mu'}} \frac{\partial x^{\beta}}{\partial x^{\nu'}} g_{\alpha\beta,\gamma'} \quad 5.34$$

$$III) \quad g_{\mu',\nu',\gamma'\lambda'} = \frac{\partial^3 x^{\alpha}}{\partial x^{y'} \partial x^{\lambda'} \partial x^{\mu'}} \frac{\partial x^{\beta}}{\partial x^{\nu'}} g_{\alpha\beta} + \dots \quad 5.35$$

The right side members of these equations contain terms such as $g_{\alpha\beta,\gamma'}$ that can be rewritten, using the chain rule, as $g_{\alpha\beta,\gamma'} = g_{\alpha\beta,\sigma} \frac{\partial x^{\sigma}}{\partial x^{\gamma'}}$.

The typology of the right side member terms thus reduces to:

- known terms: the $g_{\alpha\beta}$ and their derivatives with respect to the old coordinates, such as $g_{\alpha\beta,\sigma}$
- unknowns to be determined: the derivatives of the old coordinates with respect to the new ones.

(We'll see that just the derivatives of various order of old coordinates with respect to new ones allow us to rebuild the matrix Λ related to the transformation and the transformation itself.)

▪ Now let's perform the counting of equations and unknowns for the case $n = 4$, the most interesting for General Relativity.

The equations are counted by the left side members of I, II, III ; the unknowns by the right side members of them.

I) There are 10 equations, as many as the independent elements of the 4×4 symmetric tensor $g_{\mu'\nu'}$. The $g_{\alpha\beta}$ are known.

To get $g_{\mu'\nu'} = \eta_{\mu'\nu'} \Rightarrow$ we have to put $= 0$ or ± 1 each of the 10 independent elements of $g_{\mu'\nu'}$ (in the right side member).

Consequently, among the 16 unknown elements of matrix Λ , 6 can be arbitrarily assigned, while the other 10 are determined by equations I (eq.5.33). (The presence of six degrees of freedom means that there is a multiplicity of transformations and hence of matrices Λ able to make canonical the metric). So have been assigned or calculated values for all 16 first derivatives.

II) There are 40 equations ($g_{\mu'\nu'}$ has 10 independent elements; γ' can take 4 values). The independent second derivatives like $\frac{\partial^2 x^\alpha}{\partial x^{\gamma'} \partial x^{\mu'}}$ are 40 (4 values for α at numerator; 10 different pairs * of γ', μ' at denominator).

All the $g_{\alpha\beta}$ and first derivatives are already known. Hence, we can set $= 0$ all the 40 $g_{\mu'\nu',\gamma'}$ at first member and determine all the 40 second derivatives as a consequence.

III) There are 100 equations (10 independent elements $g_{\mu'\nu'}$ to derive with respect to the 10 different pairs * generated by 4 numbers).

Known the other factors, 80 third derivatives like $\frac{\partial^3 x^\alpha}{\partial x^{\lambda'} \partial x^{\gamma'} \partial x^{\mu'}}$ are to be determined (the 3 indexes at denominator give 20 combinations, * the index at numerator 4 other). Now, we cannot find a set of values for the 80 third derivatives such that all the 100 $g_{\mu'\nu',\gamma'\lambda'}$ vanish; 20 among them remain nonzero.

* It is a matter of combinations (in the case of 4 items) with repeats, different for at least one element.

We have thus shown that, for a generic point P , by assigning appropriate values to the derivatives of various order, it is possible to get (more than) a metric $g_{\mu',\nu'}$ such that in this point:

- $g_{\mu',\nu'} = \eta_{\mu',\nu'}$, the metric of flat manifold
- all its first derivatives are null: $\forall g_{\mu',\nu',\gamma'} = 0$
- some second derivatives are nonzero: \exists some $g_{\mu',\nu',\gamma'\lambda'} \neq 0$

▪ This metric $\eta_{\mu',\nu'}$ with zero first derivatives and second derivatives not all zero characterizes the **flat local system** in P (but the manifold itself is curved because of the nonzero second derivatives).

The metric $\eta_{\mu',\nu'}$ with zero first and second derivatives is instead that of the tangent space in P (which is a space everywhere flat). The flat local metric of the manifold and that of the tangent space coincide in P and (the metric being stationary) even in its neighborhood except for a difference of infinitesimals of higher order.

▪ Finally, we note that, having calculated or assigned appropriate values to the derivatives of various order of old coordinates with respect to the new ones, we can reconstruct the series expansions of the (inverse) transformation of coordinates $x^\alpha(x^{\mu'})$ and the elements of the matrix $\Lambda_{\mu'}^\alpha$, (transposed inverse of Λ). Indeed, in their series expansions in terms of new coordinates $x^{\mu'}$ around P

$$x^\alpha(x^{\mu'}) \rightarrow x^\alpha(P') = x^\alpha(P) + (x^{y'} - x_0^{y'}) \frac{\partial x^\alpha}{\partial x^{y'}}(P) + \\ + \frac{1}{2}(x^{y'} - x_0^{y'})(x^{\lambda'} - x_0^{\lambda'}) \frac{\partial^2 x^\alpha}{\partial x^{y'} \partial x^{\lambda'}}(P) + \dots$$

$$\Lambda_{\mu'}^\alpha(x^{\mu'}) \rightarrow \Lambda_{\mu'}^\alpha(P') = \Lambda_{\mu'}^\alpha(P) + (x^{y'} - x_0^{y'}) \frac{\partial \Lambda_{\mu'}^\alpha}{\partial x^{y'}}(P) + \\ + \frac{1}{2}(x^{y'} - x_0^{y'})(x^{\lambda'} - x_0^{\lambda'}) \frac{\partial^2 \Lambda_{\mu'}^\alpha}{\partial x^{y'} \partial x^{\lambda'}}(P) + \dots = \\ = \frac{\partial x^\alpha}{\partial x^{\mu'}}(P) + (x^{y'} - x_0^{y'}) \frac{\partial^2 x^\alpha}{\partial x^{y'} \partial x^{\mu'}}(P) + \\ + \frac{1}{2}(x^{y'} - x_0^{y'})(x^{\lambda'} - x_0^{\lambda'}) \frac{\partial^3 x^\alpha}{\partial x^{y'} \partial x^{\lambda'} \partial x^{\mu'}}(P) + \dots$$

only derivatives of old coordinates with respect to the new ones appear as coefficients.

Once known its inverse, the direct transformation $x^{\mu'}(x^\alpha)$ and the related matrix $\Lambda = [\Lambda_\alpha^{\mu'}]$ that induce the canonical metric in P are in principle implicitly determined, q.e.d.

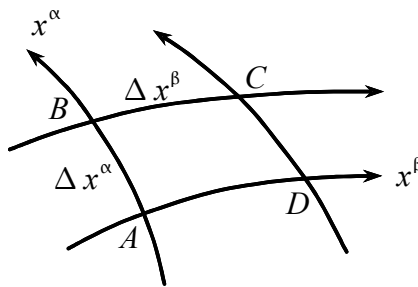
- It is worth noting that also in the flat local system the strategy “comma goes to semicolon” is applicable. Indeed, eq.4.73 ensures that even in the flat local system is $\forall \Gamma = 0$ and therefore covariant derivatives ∇_μ and ordinary derivatives ∂_μ coincide.

- That enables us to get some results in a straightforward way. For example, as in the flat local system (like in all flat manifolds) $g_{\alpha\beta,\mu} = 0$ and $g_{\alpha\beta,\mu} = g_{\alpha\beta;\mu}$, it follows $g_{\alpha\beta;\mu} \equiv \nabla_\mu g_{\alpha\beta} = 0$, which is the well-known result $\tilde{\nabla} \mathbf{g} = 0$, true in general. In the following we will use again this strategy.

5.10 Riemann tensor

Parallel transporting a vector along a closed line in a curved manifold, the final vector differs from the initial by an amount $\Delta \vec{V}$ due to the curvature. This amount $\Delta \vec{V}$ depends on the path, but it can be used as a measure of the curvature in a point if calculated along a closed infinitesimal path around the point.

Given a point A of the n -dimensional manifold we build a “parallelogram loop” $ABCD$ relying on the coordinate lines of two generic coordinates x^α, x^β picked among n



and we parallel-transport the vector \vec{V} along the circuit $ABCD A_{fin}^*$

* The loop $ABCD A_{fin}^*$ always closes: in fact it lies on the surface or “hypersurface” in which the $n-2$ not involved coordinates are constant.

(of course, the construction could be repeated using all possible couples of coordinates with $\alpha, \beta = 1, 2, \dots, n$).

Tangent to the coordinates are the basis-vectors $\vec{e}_\alpha, \vec{e}_\beta$, coordinated to x^α, x^β .

The parallel transport of \vec{V} requires that $\frac{d\vec{V}}{d\tau} = 0$ or, component-wise $U^\alpha \nabla_\alpha V^\nu = 0$ (eq.5.13), which along a segment of coordinate line x^α reduces to $\nabla_\alpha V^\nu = 0$ (eq.5.15). But:

$$\nabla_\alpha V^\nu = 0 \Rightarrow \frac{\partial V^\nu}{\partial x^\alpha} + \Gamma_{\alpha\lambda}^\nu V^\lambda = 0 \Rightarrow \frac{\partial V^\nu}{\partial x^\alpha} = -\Gamma_{\alpha\lambda}^\nu V^\lambda \quad 5.36$$

and a similar relation is obtained for the transport along x^β .

Because of the transport along the line segment AB , the components of the vector \vec{V} undergo an increase:

$$V^\nu(B) = V^\nu(A) + \left(\frac{\partial V^\nu}{\partial x^\alpha} \right)_{(AB)} \Delta x^\alpha = V^\nu(A) - \left(\Gamma_{\alpha\lambda}^\nu V^\lambda \right)_{(AB)} \Delta x^\alpha \quad 5.37$$

where (AB) means “calculated in an intermediate point between A and B ”. Along each of the four segments of the path there the changes are:

$$AB: \quad V^\nu(B) = V^\nu(A) - \left(\Gamma_{\alpha\lambda}^\nu V^\lambda \right)_{(AB)} \Delta x^\alpha$$

$$BC: \quad V^\nu(C) = V^\nu(B) - \left(\Gamma_{\beta\lambda}^\nu V^\lambda \right)_{(BC)} \Delta x^\beta$$

$$CD: \quad V^\nu(D) = V^\nu(C) + \left(\Gamma_{\alpha\lambda}^\nu V^\lambda \right)_{(CD)} \Delta x^\alpha$$

$$DA_{fin}: \quad V^\nu(A_{fin}) = V^\nu(D) + \left(\Gamma_{\beta\lambda}^\nu V^\lambda \right)_{(DA)} \Delta x^\beta$$

(in the last two segments $\Delta x^\alpha, \Delta x^\beta$ are negative in sign). Summing:

$$\begin{aligned} V^\nu(A_{fin}) - V^\nu(A) &= \left(\Gamma_{\beta\lambda}^\nu V^\lambda \right)_{(DA)} \Delta x^\beta - \left(\Gamma_{\beta\lambda}^\nu V^\lambda \right)_{(BC)} \Delta x^\beta + \\ &\quad + \left(\Gamma_{\alpha\lambda}^\nu V^\lambda \right)_{(CD)} \Delta x^\alpha - \left(\Gamma_{\alpha\lambda}^\nu V^\lambda \right)_{(AB)} \Delta x^\alpha \\ &= \left[\left(\Gamma_{\beta\lambda}^\nu V^\lambda \right)_{(DA)} - \left(\Gamma_{\beta\lambda}^\nu V^\lambda \right)_{(BC)} \right] \Delta x^\beta + \\ &\quad + \left[\left(\Gamma_{\alpha\lambda}^\nu V^\lambda \right)_{(CD)} - \left(\Gamma_{\alpha\lambda}^\nu V^\lambda \right)_{(AB)} \right] \Delta x^\alpha \end{aligned}$$

$$= -\frac{\partial}{\partial x^\alpha} (\Gamma_{\beta\lambda}^\nu V^\lambda) \Delta x^\alpha \Delta x^\beta + \frac{\partial}{\partial x^\beta} (\Gamma_{\alpha\lambda}^\nu V^\lambda) \Delta x^\beta \Delta x^\alpha$$

(in the last step we have used once again the theorem of the finite increase; now we mean the two terms in (..) be calculated in an intermediate point between (BC) and (DA) and in an intermediate point between (AB) and (CD) respectively, dropping to indicate it)

$$= -\left(\Gamma_{\beta\lambda,\alpha}^\nu V^\lambda + \Gamma_{\beta\lambda}^\nu \frac{\partial V^\lambda}{\partial x^\alpha} \right) \Delta x^\alpha \Delta x^\beta + \left(\Gamma_{\alpha\lambda,\beta}^\nu V^\lambda + \Gamma_{\alpha\lambda}^\nu \frac{\partial V^\lambda}{\partial x^\beta} \right) \Delta x^\beta \Delta x^\alpha$$

and reusing the result $\frac{\partial V^\lambda}{\partial x^\alpha} = -\Gamma_{\alpha\mu}^\lambda V^\mu$ (eq.5.36):

$$= -\left(\Gamma_{\beta\lambda,\alpha}^\nu V^\lambda - \Gamma_{\beta\lambda}^\nu \Gamma_{\alpha\mu}^\lambda V^\mu \right) \Delta x^\alpha \Delta x^\beta + \left(\Gamma_{\alpha\lambda,\beta}^\nu V^\lambda - \Gamma_{\alpha\lambda}^\nu \Gamma_{\beta\mu}^\lambda V^\mu \right) \Delta x^\beta \Delta x^\alpha$$

Using μ as dummy index instead of λ in the 1st and 3rd term:

$$\begin{aligned} &= -\left(\Gamma_{\beta\mu,\alpha}^\nu V^\mu - \Gamma_{\beta\lambda}^\nu \Gamma_{\alpha\mu}^\lambda V^\mu \right) \Delta x^\alpha \Delta x^\beta + \left(\Gamma_{\alpha\mu,\beta}^\nu V^\mu - \Gamma_{\alpha\lambda}^\nu \Gamma_{\beta\mu}^\lambda V^\mu \right) \Delta x^\beta \Delta x^\alpha \\ &= V^\mu \Delta x^\alpha \Delta x^\beta \left(-\Gamma_{\beta\mu,\alpha}^\nu + \Gamma_{\beta\lambda}^\nu \Gamma_{\alpha\mu}^\lambda + \Gamma_{\alpha\mu,\beta}^\nu - \Gamma_{\alpha\lambda}^\nu \Gamma_{\beta\mu}^\lambda \right) \end{aligned}$$

We conclude:

$$\Delta V^\nu = V^\mu \Delta x^\alpha \Delta x^\beta \underbrace{\left(-\Gamma_{\beta\mu,\alpha}^\nu + \Gamma_{\alpha\mu,\beta}^\nu + \Gamma_{\beta\lambda}^\nu \Gamma_{\alpha\mu}^\lambda - \Gamma_{\alpha\lambda}^\nu \Gamma_{\beta\mu}^\lambda \right)}_{R_{\mu\beta\alpha}^\nu} \quad 5.38$$

$R_{\mu\beta\alpha}^\nu$ must be a tensor because the other factors are tensors as well as ΔV^ν . Its rank is $\binom{1}{3}$ for the rank balancing: $\binom{1}{0} = \binom{1}{0} \binom{1}{0} \binom{1}{0} \cdot \binom{1}{3}$.

Set:

$$R_{\mu\beta\alpha}^\nu \stackrel{\text{def}}{=} \left(-\Gamma_{\beta\mu,\alpha}^\nu + \Gamma_{\alpha\mu,\beta}^\nu + \Gamma_{\beta\lambda}^\nu \Gamma_{\alpha\mu}^\lambda - \Gamma_{\alpha\lambda}^\nu \Gamma_{\beta\mu}^\lambda \right) \quad 5.39$$

we can write:

$$\Delta V^\nu = V^\mu \Delta x^\alpha \Delta x^\beta R_{\mu\beta\alpha}^\nu \quad 5.40$$

that is:

$$\Delta \vec{V}(\tilde{P}) = \mathbf{R}(\tilde{P}, \vec{V}, \vec{\Delta}x, \vec{\Delta}x) \quad 5.41$$

being \tilde{P} = any covector, or even (multiplying eq.5.40 by \vec{e}_ν):

$$\Delta \vec{V} = \vec{e}_\nu V^\mu \Delta x^\alpha \Delta x^\beta R_{\mu\alpha\beta}^\nu \quad *$$
5.42

It is tacitly supposed a passage to the limit for $\Delta x^\alpha, \Delta x^\beta \rightarrow 0$: \mathbf{R} and its components are related to the point P at which the infinitesimal loop degenerates.

$\mathbf{R} \equiv R_{\mu\beta\alpha}^\nu \vec{e}_\nu \tilde{e}^\mu \tilde{e}^\beta \tilde{e}^\alpha$, the *Riemann tensor*, expresses the “proportionality constant” between the change undergone by a parallel transported vector along an (infinitesimal) loop and the area of the loop, and contains the whole information about the curvature. In general it depends upon the point: $\mathbf{R} = \mathbf{R}(P)$.

If $\mathbf{R} = 0$ from eq.5.41 it follows $\Delta \vec{V} = 0 \Rightarrow$ the vector \vec{V} overlaps itself after the parallel transport along the infinitesimal loop around the point $P \Rightarrow$ the manifold is flat at that point. Hence:

$$\mathbf{R}(P) = 0 \Rightarrow \text{manifold flat in } P$$

If that applies globally the manifold is everywhere flat.

If $\mathbf{R} \neq 0$ in P ,the manifold is curved in that point, although it is possible to define a local flat system in P .

▪ \mathbf{R} is a function of Γ and its first derivatives (eq.5.39), namely of \mathbf{g} and its first and second derivatives (see eq.4.73).

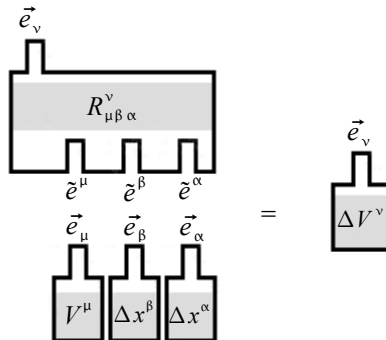
▪ In the flat local coordinate system where \mathbf{g} assumes the canonical form with zero first derivatives, \mathbf{R} depends only on the second derivatives of \mathbf{g} . Let us clarify now this dependence.

Because in each point of the manifold in its flat local system is $\forall \Gamma = 0$, the definition eq.5.39 given for $R_{\mu\beta\alpha}^\nu$ reduces to:

$$R_{\mu\beta\alpha}^\nu = -\Gamma_{\beta\mu,\alpha}^\nu + \Gamma_{\alpha\mu,\beta}^\nu$$
5.43

* That is equivalent to \mathbf{R} applied to an incomplete list. Completing the list gives eq.5.41.

Eq.5.40 is eq.5.42 componentwise.



From eq.4.73 that gives the Γ as functions of g and their derivatives

$$\Gamma_{\beta\mu}^{\nu} = \frac{1}{2} g^{\nu\rho} (-g_{\beta\mu,\rho} + g_{\mu\rho,\beta} + g_{\rho\beta,\mu}) \quad \text{and since } \frac{\partial}{\partial x^{\alpha}} g^{\nu\rho} = 0 :$$

$$\begin{aligned} R_{\mu\beta\alpha}^{\nu} &= -\frac{1}{2} g^{\nu\rho} (-g_{\beta\mu,\rho\alpha} + g_{\mu\rho,\beta\alpha} + g_{\rho\beta,\mu\alpha}) + \\ &\quad + \frac{1}{2} g^{\nu\rho} (-g_{\alpha\mu,\rho\beta} + g_{\mu\rho,\alpha\beta} + g_{\rho\alpha,\mu\beta}) \\ &= \frac{1}{2} g^{\nu\rho} (+g_{\beta\mu,\rho\alpha} - g_{\rho\beta,\mu\alpha} - g_{\alpha\mu,\rho\beta} + g_{\rho\alpha,\mu\beta}) \end{aligned}$$

and multiplying both members by $g_{\nu\kappa}$ we get (recall that $g_{\nu\kappa} g^{\nu\rho} = \delta_{\kappa}^{\rho}$ and hence $\rho \rightarrow \kappa$):

$$R_{\kappa\mu\beta\alpha} = \frac{1}{2} (g_{\beta\mu,\kappa\alpha} - g_{\kappa\beta,\mu\alpha} - g_{\alpha\mu,\kappa\beta} + g_{\kappa\alpha,\mu\beta})$$

To write, as usual, $\alpha\beta$ as first pair and $\mu\nu$ as second we have to make the index exchanges $\kappa \rightarrow \alpha, \mu \rightarrow \beta, \beta \rightarrow \mu, \alpha \rightarrow \nu$; rearranging the terms and by the symmetry of $g_{\alpha\beta}$ we get:

$$R_{\alpha\beta\mu\nu} = \frac{1}{2} (g_{\alpha\nu,\beta\mu} + g_{\beta\mu,\alpha\nu} - g_{\alpha\mu,\beta\nu} - g_{\beta\nu,\alpha\mu}) \quad 5.44$$

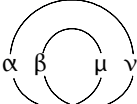
that is true in each point of the manifold in its respective flat local system (but *not* in general).

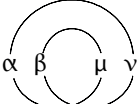
- It is worth noting that the eq.5.44 is no longer valid in a generic system because it is not a tensor equation: outside the flat local system we need to retrieve the last two terms in $\Gamma\Gamma$ from eq.5.39. Therefore in a generic coordinate system it is:

$$R_{\alpha\beta\mu\nu} = \frac{1}{2} (g_{\alpha\nu,\beta\mu} + g_{\beta\mu,\alpha\nu} - g_{\alpha\mu,\beta\nu} - g_{\beta\nu,\alpha\mu}) + g_{\eta\alpha} (\Gamma_{\mu\lambda}^{\eta} \Gamma_{\nu\beta}^{\lambda} - \Gamma_{\nu\lambda}^{\eta} \Gamma_{\mu\beta}^{\lambda}) \quad 5.45$$

Mnemo

To write down $R_{\alpha\beta\mu\nu}$ in the local flat system (eq.5.44) the “Pascal snail” can be used as a mnemonic aid to suggest the pairs of the indexes before the comma:

$+$  $+ g_{\alpha\nu, \dots} + g_{\beta\mu, \dots}$

$-$  $- g_{\alpha\mu, \dots} - g_{\beta\nu, \dots}$

Use the remaining indexes for pairs after the comma.

- In $R_{\alpha\beta\mu\nu}$ it is usual to identify two pairs of indexes: $R_{\begin{matrix} \alpha\beta \\ 1^{st} \end{matrix} \begin{matrix} \mu\nu \\ 2^{nd} \\ \text{pair pair} \end{matrix}}$

- Bringing order also in the indexes of *eq.5.39* we rewrite it as:

$$R_{\beta\mu\nu}^{\alpha} = \Gamma_{\beta\nu,\mu}^{\alpha} - \Gamma_{\beta\mu,\nu}^{\alpha} + \Gamma_{\lambda\mu}^{\alpha} \Gamma_{\beta\nu}^{\lambda} - \Gamma_{\lambda\nu}^{\alpha} \Gamma_{\beta\mu}^{\lambda} \quad 5.46$$

- The Riemann tensor \mathbf{R} provides a flatness *criterion* more general than that one stated by *eq.5.22* e *eq.5.23* and also valid for *indefinite* metrics (that can be flat even if P -dependent).

5.11 Symmetries of tensor \mathbf{R}

The symmetries of $R_{\alpha\beta\mu\nu}$ can be studied in the flat local system by interchanging indexes in *eq.5.44*.

For example, exchanging the indexes $\alpha \leftrightarrow \beta$ within the first pair gives:

$$R_{\beta\alpha\mu\nu} = \frac{1}{2} (\mathcal{g}_{\beta\nu,\alpha\mu} + \mathcal{g}_{\alpha\mu,\beta\nu} - \mathcal{g}_{\beta\mu,\alpha\nu} - \mathcal{g}_{\alpha\nu,\beta\mu})$$

whose right side member is the same as in the starting equation changed in sign and thus $\Rightarrow R_{\beta\alpha\mu\nu} = -R_{\alpha\beta\mu\nu}$.

It turns out from *eq.5.44* that \mathbf{R} is: *i)* skew-symmetric with respect to the exchange of indexes within a pair; *ii)* symmetric with respect to the exchange of a pair with the other; it also enjoys a property such that: *iii)* the sum with cyclicity in the last 3 indexes is null:

$$i) \quad R_{\beta\alpha\mu\nu} = -R_{\alpha\beta\mu\nu} \quad R_{\alpha\beta\nu\mu} = -R_{\alpha\beta\mu\nu} \quad 5.47$$

$$ii) \quad R_{\mu\nu\alpha\beta} = R_{\alpha\beta\mu\nu} \quad 5.48$$

$$iii) \quad R_{\alpha\beta\mu\nu} + R_{\alpha\mu\nu\beta} + R_{\alpha\nu\beta\mu} = 0 \quad 5.49$$

The four previous relations, although deduced in the flat local system (*eq.5.44* applies there only) are tensor equations and thus valid in any reference frame.

From *i)* it follows that the components with repeated indexes within the pair are null (for example $R_{11\mu\nu} = R_{\beta\alpha 33} = R_{221\nu} = R_{1111} = 0$). Indeed, from *i)*: $R_{\alpha\alpha\mu\nu} = -R_{\alpha\alpha\mu\nu} \Rightarrow R_{\alpha\alpha\mu\nu} = 0$; and so on.

- In the end, due to symmetries, among the n^4 components of $R_{\alpha\beta\mu\nu}$ only $n^2(n^2-1)/12$ are independent and $\neq 0$ (namely 1 for $n=2$; 6 for $n=3$, 20 for $n=4$; ...).

5.12 Bianchi identity

It's another relation linking the covariant first derivatives of $R_{\alpha\beta\mu\nu}$:

$$R_{\alpha\beta\mu\nu;\lambda} + R_{\alpha\beta\nu\lambda;\mu} + R_{\alpha\beta\lambda\mu;\nu} = 0 \quad 5.50$$

(the first two indexes $\alpha\beta$ are fixed; the others $\mu\nu\lambda$ rotate).

- To get this result we place again in the flat local system of a point P and calculate the derivatives of $R_{\alpha\beta\mu\nu}$ from eq.5.44:

$$\begin{aligned} R_{\alpha\beta\mu\nu;\lambda} &= \frac{\partial}{\partial x^\lambda} \frac{1}{2} (g_{\alpha\nu,\beta\mu} + g_{\beta\mu,\alpha\nu} - g_{\alpha\mu,\beta\nu} - g_{\beta\nu,\alpha\mu}) \\ &= \frac{1}{2} (g_{\alpha\nu,\beta\mu\lambda} + g_{\beta\mu,\alpha\nu\lambda} - g_{\alpha\mu,\beta\nu\lambda} - g_{\beta\nu,\alpha\mu\lambda}) \end{aligned}$$

and similarly for $R_{\alpha\beta\nu\lambda,\mu}$ e $R_{\alpha\beta\lambda\mu,\nu}$.

Summing member to member these three equations and taking into account that in $g_{\alpha\beta,\gamma\delta\kappa}$ the indexes can be exchanged at will within the two groups $\alpha\beta, \gamma\delta$ *, the right member vanishes so that:

$$R_{\alpha\beta\mu\nu;\lambda} + R_{\alpha\beta\nu\lambda;\mu} + R_{\alpha\beta\lambda\mu;\nu} = 0$$

Since in the flat local system there is no difference between derivative and covariant derivative, namely:

$$R_{\alpha\beta\mu\nu;\lambda} = R_{\alpha\beta\mu\nu,\lambda} ; R_{\alpha\beta\nu\lambda;\mu} = R_{\alpha\beta\nu\lambda,\mu} ; R_{\alpha\beta\lambda\mu;\nu} = R_{\alpha\beta\lambda\mu,\nu}$$

we can write $R_{\alpha\beta\mu\nu;\lambda} + R_{\alpha\beta\nu\lambda;\mu} + R_{\alpha\beta\lambda\mu;\nu} = 0$ which is a tensor relationship and applies in any coordinate system, q.e.d.

5.12 Ricci tensor and Ricci scalar

$R_{\alpha\beta\mu\nu}$ may be contracted ** on indexes of the same pair or on indexes of different pairs. In the first case the result is 0, in the second it is significant.

The following schemes illustrate some possible contractions and their results:

* $g_{\alpha\beta}$ is symmetric and the derivation order doesn't matter.

** In fact, it is $R_{\beta\mu\nu}^\alpha$ to undergo a contraction.

$$\left. \begin{array}{l}
 R_{\alpha\beta\mu\nu} \xrightarrow{\cdot g^{\alpha\beta}} R_{\beta\mu\nu}^{\beta} \rightarrow R_{\mu\nu} \\
 \downarrow = \\
 -R_{\beta\alpha\mu\nu} \xrightarrow{\cdot g^{\beta\alpha}} -R_{\alpha\mu\nu}^{\alpha} \rightarrow -R_{\mu\nu}
 \end{array} \right\} \Rightarrow R_{\mu\nu} = 0$$

5.51

$$\left. \begin{array}{l}
 R_{\alpha\beta\mu\nu} \xrightarrow{\cdot g^{\alpha\mu}} R_{\beta\mu\nu}^{\mu} \rightarrow R_{\beta\nu} \\
 \downarrow = \\
 R_{\mu\nu\alpha\beta} \xrightarrow{\cdot g^{\mu\alpha}} R_{\nu\alpha\beta}^{\alpha} \rightarrow R_{\nu\beta}
 \end{array} \right\} \Rightarrow R_{\beta\nu} = R_{\nu\beta}$$

Other similar cases still give as result 0 or $\pm R_{\beta\nu}$.

We define **Ricci tensor** the $\binom{0}{2}$ tensor with components:

$$R_{\beta\nu} \stackrel{\text{def}}{=} \text{contraction of } R_{\alpha\beta\mu\nu} \text{ with respect to } 1^{\text{st}} \text{ and } 3^{\text{rd}} \text{ index} \quad 5.52$$

▪ Caution should be paid when contracting Riemann tensor on any two indexes: before contracting we must move the two indexes to contract in 1st and 3rd position using the symmetries of $R_{\alpha\beta\mu\nu}$, in order to perform anyway the contraction 1-3 whose result is known and positive by definition.

In such manner, due to the symmetries of $R_{\alpha\beta\mu\nu}$ and using schemes like eq.5.51 we see that:

- have sign + the results of contraction on indexes 1-3 or 1-4
- have sign – the results of contraction on indexes 1-4 or 2-3

(while contractions on indexes 1-2 or 3-4 have 0 as result)

- The tensor $R_{\beta\nu}$ is symmetric (see eq.5.51).
- By further contraction of the Ricci tensor we get the **Ricci scalar** R :

$$R_{\beta\nu} \xrightarrow{\cdot g^{\beta\nu}} R_{\nu}^{\nu} \rightarrow R \quad (= \text{trace of } R_{\beta\nu}). \quad * \quad 5.53$$

5.13 Einstein tensor

From Bianchi identity eq.5.49, raising and contracting twice the indexes to reduce Riemann to Ricci tensors and using the commutation property between index raising and covariant derivative:

*Remind that the trace is defined only for a mixed tensor T_{ν}^{μ} as the sum T_{μ}^{μ} of elements of the main diagonal of its matrix. Instead, the trace of $T_{\mu\nu}$ or $T^{\mu\nu}$ is by definition the trace of $g^{\kappa\mu} T_{\mu\nu}$ or, respectively, of $g_{\kappa\mu} T^{\mu\nu}$.

$$g^{\beta\nu} g^{\alpha\mu} (R_{\alpha\beta\mu\nu;\lambda} + R_{\alpha\beta\nu\lambda;\mu} + R_{\alpha\beta\lambda\mu;\nu}) = 0$$

$$g^{\beta\nu} (R^{\mu}_{\beta\mu\nu;\lambda} + R^{\mu}_{\beta\nu\lambda;\mu} + R^{\mu}_{\beta\lambda\mu;\nu}) = 0$$

contracting $R^{\mu}_{\beta\mu\nu;\lambda}$ on indexes 1-3 (sign +)
and $R^{\mu}_{\beta\lambda\mu;\nu}$ on indexes 1-4 (sign -) :

$$g^{\beta\nu} (R_{\beta\nu;\lambda} + R^{\mu}_{\beta\nu\lambda;\mu} - R_{\beta\lambda;\nu}) = 0$$

using symmetry $R^{\mu}_{\beta\nu\lambda;\mu} = -R_{\beta\nu\lambda;\mu}$ to allow
the hooking and rising of index β :

$$g^{\beta\nu} (R_{\beta\nu;\lambda} - R^{\mu}_{\beta\nu\lambda;\mu} - R_{\beta\lambda;\nu}) = 0$$

$$R^{\nu}_{\nu;\lambda} - R^{\nu\mu}_{\nu\lambda;\mu} - R^{\nu}_{\lambda;\nu} = 0$$

contracting $R^{\nu\mu}_{\nu\lambda;\mu}$ on indexes 1-3 (sign +)

$$R^{\nu}_{\nu;\lambda} - R^{\mu}_{\lambda;\mu} - R^{\nu}_{\lambda;\nu} = 0$$

The first term contracts again; the 2nd and 3rd term differ for a dummy index and are in fact the same:

$$R_{;\lambda} - 2R^{\mu}_{\lambda;\mu} = 0 \quad 5.54$$

that, thanks to the identity $(2R^{\mu}_{\lambda} - \delta^{\mu}_{\lambda} R)_{;\mu} \equiv 2R^{\mu}_{\lambda;\mu} - R_{;\lambda}$, may be written as:

$$(2R^{\mu}_{\lambda} - \delta^{\mu}_{\lambda} R)_{;\mu} = 0 \quad 5.55$$

This expression or its equivalent *eq.5.54* are sometimes called *twice contracted Bianchi identities*.

Operating a further rising of the indexes:

$$g^{\nu\lambda} (2R^{\mu}_{\lambda} - \delta^{\mu}_{\lambda} R)_{;\mu} = 0$$

$$(2R^{\nu\mu} - \delta^{\nu\mu} R)_{;\mu} = 0$$

that, since $\delta^{\nu\mu} \equiv g^{\nu\mu}$ (*eq.2.46*), gives:

$$(2R^{\nu\mu} - g^{\nu\mu} R)_{;\mu} = 0$$

$$\left(R^{\nu\mu} - \frac{1}{2} g^{\nu\mu} R \right)_{;\mu} = 0 \quad 5.56$$

We define **Einstein tensor** the $\binom{0}{2}$ tensor **G** whose components are: *

* Nothing to do, of course, with the “dual switch” earlier denoted by the same symbol!

$$G^{\nu\mu} \stackrel{\text{def}}{=} R^{\nu\mu} - \frac{1}{2} g^{\nu\mu} R$$

The eq.5.56 can then be written:

$$G^{\nu\mu}{}_{;\mu} = 0 \tag{5.57}$$

which shows that the tensor **G** has null divergence.

Since both $R^{\nu\mu}$ and $g^{\nu\mu}$ are symmetric under indexes exchange, **G** is symmetric too: $G^{\nu\mu} = G^{\mu\nu}$.

The tensor **G** has an important role in General Relativity because it is the only divergence-free double-tensor that can be derived from **R** and as such contains the information about the curvature of the manifold.

G is our last stop. What's all that for? Blending physics and mathematics with an overdose of intuition, one day of a hundred years ago Einstein wrote:

$$G^{\mu\nu} = \kappa T^{\mu\nu}$$

anticipating, according to some, a theory of the Third Millennium to the twentieth century. This equation expresses a direct proportionality between the content of matter-energy represented by the "stress-energy-momentum" tensor **T** – a generalization in the Minkowski space-time of the stress tensor – and the curvature of space expressed by **G** (it was to ensure the conservation of energy and momentum that Einstein needed a zero divergence tensor like **G**). In a sense, **T** belongs to the realm of physics, while **G** is matter of mathematics: that maths we have toyed with so far. Compatibility with the Newton's gravitation law allows to give a value to the constant κ and write down the fundamental equation of General Relativity in the usual form:

$$G^{\mu\nu} = \frac{8\pi G}{c^4} T^{\mu\nu}$$

May be curious to note in the margin as, in this equation defining the fate of the Universe, beside fundamental constants such as the gravitational constant G and the light speed c , peeps out the ineffable π , which seems to claim in that manner its key role in the architecture of the world, although no one seems to make a great account of that (those who think of this observation as trivial try to imagine a Universe in which π is, yes, a constant, but different from 3.14).

Appendix

1 - Transformation of Γ under coordinate change

We explicit $\Gamma_{\mu\nu}^\lambda$ from eq.4.30 $\Gamma_{\mu\nu}^\lambda \vec{e}_\lambda = \partial_\mu \vec{e}_\nu$ by (scalar) multiplying both members by \tilde{e}^k :

$$\Gamma_{\mu\nu}^\lambda \underbrace{\langle e_\lambda, \tilde{e}^k \rangle}_{\delta_\lambda^k} = \langle \partial_\mu \vec{e}_\nu, \tilde{e}^k \rangle$$

$$\Gamma_{\mu\nu}^k = \langle \partial_\mu \vec{e}_\nu, \tilde{e}^k \rangle$$

Now let us operate a coordinate change $x^\xi \rightarrow x^{\zeta'}$ (and coordinate bases, too) and express the right member in the new frame:

$$\begin{aligned} \Gamma_{\mu\nu}^k &= \langle \partial_\mu \vec{e}_\nu, \tilde{e}^k \rangle \\ &= \left\langle \frac{\partial x^{\alpha'}}{\partial x^\mu} \frac{\partial}{\partial x^{\alpha'}} (\Lambda_v^{\beta'} \vec{e}_{\beta'}), \Lambda_y^k \tilde{e}^{y'} \right\rangle \quad (\text{chain rule on } x^{\alpha'}) \\ &= \Lambda_\mu^{\alpha'} \left\langle \frac{\partial}{\partial x^{\alpha'}} (\Lambda_v^{\beta'} \vec{e}_{\beta'}), \Lambda_y^k \tilde{e}^{y'} \right\rangle \\ &= \Lambda_\mu^{\alpha'} \Lambda_y^k \left\langle \frac{\partial}{\partial x^{\alpha'}} (\Lambda_v^{\beta'} \vec{e}_{\beta'}), \tilde{e}^{y'} \right\rangle \\ &= \Lambda_\mu^{\alpha'} \Lambda_y^k \left\langle \frac{\partial \vec{e}_{\beta'}}{\partial x^{\alpha'}} \Lambda_v^{\beta'} + \vec{e}_{\beta'} \frac{\partial}{\partial x^{\alpha'}} \Lambda_v^{\beta'}, \tilde{e}^{y'} \right\rangle \\ &= \Lambda_\mu^{\alpha'} \Lambda_y^k \Lambda_v^{\beta'} \left\langle \frac{\partial \vec{e}_{\beta'}}{\partial x^{\alpha'}}, \tilde{e}^{y'} \right\rangle + \Lambda_\mu^{\alpha'} \Lambda_y^k \frac{\partial}{\partial x^{\alpha'}} \Lambda_v^{\beta'} \langle \vec{e}_{\beta'}, \tilde{e}^{y'} \rangle \\ &= \Lambda_\mu^{\alpha'} \Lambda_y^k \Lambda_v^{\beta'} \underbrace{\left\langle \frac{\partial \vec{e}_{\beta'}}{\partial x^{\alpha'}}, \tilde{e}^{y'} \right\rangle}_{\Gamma_{\alpha'\beta'}^{y'}} + \Lambda_\mu^{\alpha'} \Lambda_y^k \frac{\partial}{\partial x^{\alpha'}} \Lambda_v^{\beta'} \delta_\beta^{y'} \\ &= \Lambda_\mu^{\alpha'} \Lambda_y^k \Lambda_v^{\beta'} \Gamma_{\alpha'\beta'}^{y'} + \Lambda_\mu^{\alpha'} \Lambda_y^k \frac{\partial}{\partial x^{\alpha'}} \Lambda_v^{\beta'} \\ &= \Lambda_\mu^{\alpha'} \Lambda_y^k \Lambda_v^{\beta'} \Gamma_{\alpha'\beta'}^{y'} + \frac{\partial x^{\alpha'}}{\partial x^\mu} \frac{\partial x^k}{\partial x^{\beta'}} \frac{\partial^2 x^{\beta'}}{\partial x^{\alpha'} \partial x^\nu} \\ &= \Lambda_\mu^{\alpha'} \Lambda_y^k \Lambda_v^{\beta'} \Gamma_{\alpha'\beta'}^{y'} + \frac{\partial x^k}{\partial x^{\beta'}} \frac{\partial^2 x^{\beta'}}{\partial x^\mu \partial x^\nu} \quad (\text{chain rule on } x^{\alpha'}) \end{aligned}$$

The first term would describe a tensor transform, but the additional term leads to a different law and confirms that $\Gamma_{\mu\nu}^k$ is not a tensor.

2 - Transformation of covariant derivative under coordinate change

To $V^{\kappa}_{;\mu} = V^{\kappa}_{,\mu} + \Gamma^{\kappa}_{\mu\nu} V^{\nu}$

let's apply a coordinate change $x^{\xi} \rightarrow x^{\zeta'}$ and express the right member in the new coordinate frame:

$$\begin{aligned}
 V^{\kappa}_{;\mu} &= V^{\kappa}_{,\mu} + \Gamma^{\kappa}_{\mu\nu} V^{\nu} \\
 &= \frac{\partial}{\partial x^{\mu}} (\Lambda^{\kappa}_{\gamma'} V^{\gamma'}) + (\text{transformation of } \Gamma) \Lambda^{\nu}_{\sigma'} V^{\sigma'} \\
 &= \Lambda^{\alpha'}_{\mu} \frac{\partial}{\partial x^{\alpha'}} (\Lambda^{\kappa}_{\gamma'} V^{\gamma'}) + (\text{transformation of } \Gamma) \Lambda^{\nu}_{\sigma'} V^{\sigma'} \\
 &= \Lambda^{\alpha'}_{\mu} \Lambda^{\kappa}_{\gamma'} \frac{\partial V^{\gamma'}}{\partial x^{\alpha'}} + V^{\gamma'} \Lambda^{\alpha'}_{\mu} \frac{\partial}{\partial x^{\alpha'}} (\Lambda^{\kappa}_{\gamma'}) + \\
 &\quad + \left(\Lambda^{\alpha'}_{\mu} \Lambda^{\kappa}_{\gamma'} \Lambda^{\beta'}_{\nu} \Gamma^{\gamma' \nu}_{\alpha' \beta'} + \frac{\partial x^{\kappa}}{\partial x^{\beta'}} \frac{\partial^2 x^{\beta'}}{\partial x^{\mu} \partial x^{\nu}} \right) \Lambda^{\nu}_{\sigma'} V^{\sigma'} \\
 &= \Lambda^{\alpha'}_{\mu} \Lambda^{\kappa}_{\gamma'} \frac{\partial V^{\gamma'}}{\partial x^{\alpha'}} + V^{\gamma'} \Lambda^{\alpha'}_{\mu} \frac{\partial}{\partial x^{\alpha'}} (\Lambda^{\kappa}_{\gamma'}) + \\
 &\quad + \Lambda^{\alpha'}_{\mu} \Lambda^{\kappa}_{\gamma'} \Lambda^{\beta'}_{\nu} \Gamma^{\gamma' \nu}_{\alpha' \beta'} \Lambda^{\nu}_{\sigma'} V^{\sigma'} + \frac{\partial x^{\kappa}}{\partial x^{\beta'}} \frac{\partial^2 x^{\beta'}}{\partial x^{\mu} \partial x^{\nu}} \Lambda^{\nu}_{\sigma'} V^{\sigma'} \\
 &= \Lambda^{\alpha'}_{\mu} \Lambda^{\kappa}_{\gamma'} \frac{\partial V^{\beta'}}{\partial x^{\alpha'}} + V^{\gamma'} \Lambda^{\alpha'}_{\mu} \frac{\partial}{\partial x^{\alpha'}} \frac{\partial x^{\kappa}}{\partial x^{\gamma'}} + \\
 &\quad + \Lambda^{\alpha'}_{\mu} \Lambda^{\kappa}_{\gamma'} \Gamma^{\gamma' \nu}_{\alpha' \beta'} V^{\beta'} + \frac{\partial x^{\kappa}}{\partial x^{\beta'}} \frac{\partial^2 x^{\beta'}}{\partial x^{\mu} \partial x^{\sigma'}} V^{\sigma'} \quad * \\
 &= \Lambda^{\alpha'}_{\mu} \Lambda^{\kappa}_{\gamma'} \frac{\partial V^{\gamma'}}{\partial x^{\alpha'}} + V^{\gamma'} \frac{\partial x^{\alpha'}}{\partial x^{\mu}} \frac{\partial^2 x^{\kappa}}{\partial x^{\alpha'} \partial x^{\gamma'}} + \\
 &\quad + \Lambda^{\alpha'}_{\mu} \Lambda^{\kappa}_{\gamma'} \Gamma^{\gamma' \nu}_{\alpha' \beta'} V^{\beta'} + \frac{\partial x^{\kappa}}{\partial x^{\beta'}} \frac{\partial^2 x^{\beta'}}{\partial x^{\mu} \partial x^{\sigma'}} V^{\sigma'}
 \end{aligned}$$

or, changing some dummy indexes:

$$\begin{aligned}
 &= \Lambda^{\alpha'}_{\mu} \Lambda^{\kappa}_{\gamma'} \frac{\partial V^{\gamma'}}{\partial x^{\alpha'}} + V^{\sigma'} \frac{\partial x^{\alpha'}}{\partial x^{\mu}} \frac{\partial^2 x^{\kappa}}{\partial x^{\alpha'} \partial x^{\sigma'}} + \\
 &\quad + \Lambda^{\alpha'}_{\mu} \Lambda^{\kappa}_{\gamma'} \Gamma^{\gamma' \nu}_{\alpha' \beta'} V^{\beta'} + \frac{\partial x^{\kappa}}{\partial x^{\alpha'}} \frac{\partial^2 x^{\alpha'}}{\partial x^{\mu} \partial x^{\sigma'}} V^{\sigma'}
 \end{aligned}$$

* Recall that $\Lambda^{\beta'}_{\nu} \Lambda^{\nu}_{\sigma'} = \frac{\partial x^{\beta'}}{\partial x^{\nu}} \frac{\partial x^{\nu}}{\partial x^{\sigma'}} = \frac{\partial x^{\beta'}}{\partial x^{\sigma'}} = \delta^{\beta'}_{\sigma'}$

$$\begin{aligned}
&= \Lambda_{\mu}^{\alpha'} \Lambda_{\gamma'}^{\kappa} V_{,\alpha'}^{\gamma'} + \Lambda_{\mu}^{\alpha'} \Lambda_{\gamma'}^{\kappa} \Gamma_{\alpha'\beta'}^{\gamma'} V^{\beta'} + \text{terms in } \partial^2 \\
&= \Lambda_{\mu}^{\alpha'} \Lambda_{\gamma'}^{\kappa} (V_{,\alpha'}^{\gamma'} + \Gamma_{\alpha'\beta'}^{\gamma'} V^{\beta'}) + \text{terms in } \partial^2
\end{aligned}$$

The two terms in ∂^2 cancel each other because opposite in sign. To show that let's compute:

$$\begin{aligned}
\frac{\partial}{\partial x^{\sigma'}} \frac{\partial x^{\kappa}}{\partial x^{\mu}} &= \frac{\partial}{\partial x^{\sigma'}} \left(\frac{\partial x^{\alpha'}}{\partial x^{\mu}} \frac{\partial x^{\kappa}}{\partial x^{\alpha'}} \right) = \frac{\partial x^{\alpha'}}{\partial x^{\mu}} \frac{\partial}{\partial x^{\sigma'}} \frac{\partial x^{\kappa}}{\partial x^{\alpha'}} + \frac{\partial x^{\kappa}}{\partial x^{\alpha'}} \frac{\partial}{\partial x^{\sigma'}} \frac{\partial x^{\alpha'}}{\partial x^{\mu}} = \\
&= \frac{\partial x^{\alpha'}}{\partial x^{\mu}} \frac{\partial^2 x^{\kappa}}{\partial x^{\sigma'} \partial x^{\alpha'}} + \frac{\partial x^{\kappa}}{\partial x^{\alpha'}} \frac{\partial^2 x^{\alpha'}}{\partial x^{\sigma'} \partial x^{\mu}}
\end{aligned}$$

On the other hand: $\frac{\partial}{\partial x^{\sigma'}} \frac{\partial x^{\kappa}}{\partial x^{\mu}} = \frac{\partial}{\partial x^{\sigma'}} \delta_{\mu}^{\kappa} = 0$, hence:

$$\frac{\partial x^{\alpha'}}{\partial x^{\mu}} \frac{\partial^2 x^{\kappa}}{\partial x^{\sigma'} \partial x^{\alpha'}} = - \frac{\partial x^{\kappa}}{\partial x^{\alpha'}} \frac{\partial^2 x^{\alpha'}}{\partial x^{\sigma'} \partial x^{\mu}} , \text{ q.e.d.}$$

The transformation for $V_{;\mu}^{\kappa}$ is then:

$$\begin{aligned}
V_{;\mu}^{\kappa} &= V_{,\mu}^{\kappa} + \Gamma_{\mu\nu}^{\kappa} V^{\nu} = \Lambda_{\mu}^{\alpha'} \Lambda_{\gamma'}^{\kappa} (V_{,\alpha'}^{\gamma'} + \Gamma_{\alpha'\beta'}^{\gamma'} V^{\beta'}) = \Lambda_{\mu}^{\alpha'} \Lambda_{\gamma'}^{\kappa} V_{;\alpha'}^{\gamma'} \Rightarrow \\
&\Rightarrow V_{;\mu}^{\kappa} \text{ transforms as a (component of a) } \binom{1}{1} \text{ tensor.}
\end{aligned}$$

That allows us to define a tensor $\tilde{\nabla} \vec{V} \xrightarrow{\text{comp}} V_{;\mu}^{\kappa}$.

3 - Non-tensoriality of basis-vectors, their derivatives and gradients

Are not tensors: a) basis-vectors \vec{e}_{ν} ; b) their derivatives $\partial_{\mu} \vec{e}_{\nu}$; c) their gradients $\tilde{\nabla} \vec{e}_{\nu}$

a) "Basis-vector" is a role that is given to certain vectors: within a vector space n vectors are chosen to wear the "jacket" of basis-vectors. Under change of coordinates these vectors remain unchanged, tensorially transforming their components (according to the usual contravariant scheme *eq.3.4*), while their role of basis-vectors is transferred to other vectors of the vector space. More than a law of transformation of the basis-vectors, *eq.3.2* is the law ruling the role or jacket transfer.

For instance, under transformation from Cartesian coordinates to spherical, the vector $\vec{V} \equiv (1, 0, 0)$ that in Cartesian plays the role of basis-vector \vec{i} transforms its components (following the contravariant

scheme *eq.3.4*) and remains unchanged, but loses the role of basis-vector which is transferred to new vectors according to *eq.3.2* (which *looks like* a covariant scheme).

Vectors underlying basis-vectors have then a tensorial character that *does not belong to basis-vectors as such*.

b) Scalar components of $\partial_\mu \vec{e}_\nu$ are the Christoffel symbols (*eq.4.30*, *eq.4.31*) which, as shown in Appendix 1, do not behave as tensors: that excludes that $\partial_\mu \vec{e}_\nu$ is a tensor.

The same conclusion is reached whereas the derivatives of basic-vectors $\partial_\mu \vec{e}_\nu$ are null in Cartesian coordinates but not in spherical ones. Since a tensor which is null in a reference must be zero in all, the *derivatives of basis-vectors* $\partial_\mu \vec{e}_\nu$ *are not tensors*.

c) Also the gradients of basis-vectors $\tilde{\nabla} \vec{e}_\nu$ have as scalar components the Christoffel symbols (*eq.4.51*); since they have not a tensorial character (see Appendix 1), it follows that *gradients of basis-vectors* $\tilde{\nabla} \vec{e}_\nu$ *are not tensors*.

As above, the same conclusion is reached whereas the $\tilde{\nabla} \vec{e}_\nu$ are zero in Cartesian but not zero in spherical coordinates.

We note that this does not conflict with the fact that $\partial \vec{V}$, $\tilde{\nabla} \vec{V}$ own tensor character: the scalar components of both are the covariant derivatives which transform as tensors.

Bibliographic references

- Among texts specifically devoted to Tensor Analysis the following retain a relatively soft profile:

Bertschinger, E. 1999, *Introduction to Tensor Calculus for General Relativity*, Massachusetts Institute of Technology – Physics 8.962, pag. 34

Download: <http://web.mit.edu/edbert/GR/gr1.pdf>

In just over 30 pages a systematic, well structured and clear though concise treatment of Tensor Analysis aimed to General Relativity. The approach to tensors is of the geometric type, the best suited to acquire the basic concepts in a systematic way. The notation conforms to standard texts of Relativity. No use is made of matrix calculus. These notes are modeled in part on this text, possibly the best among those known to the author. Recommended without reserve; it can be at times challenging at first approach.

Kay, D. C. 1988, *Tensor Calculus*, McGraw-Hill, pag. 228

This book belongs to the mythical “Schaum Outline Series”, but without having the qualities. The approach to Tensor Analysis is the most traditional “by components” (except a final chapter where the broad lines of the geometric approach are piled up in a very formal and almost incomprehensible summary). The traditional separation between statements and problems characteristic of Schaum's Outline is here interpreted in a not very happy way: the theoretical parts are somewhat abstract and formal, the problems are mostly examples of mechanisms not suitable to clarify the concepts. In fact, the conceptual part is the great absent in this book (which you can read almost to the end without having clear the very fundamental invariance properties of vectors and tensors subject to coordinate transformations!). Extensive use of matrix calculus (to which a far too insufficient review chapter is devoted). The second half of the book is dedicated to topics of interest for Differential Geometry but marginal for General Relativity. The notations used are not always standard. This text is not recommended as an introduction to the mathematics of General Relativity, to which it is rather misleading. Self-study risky. On the contrary it is useful for consultation of certain subjects, when the same are somewhat known.

Lovelock, D., Rund, H. 1989, *Tensors, Differential Forms and Variational Principles*, Dover Publication, pag. 366

Text with a somewhat mathematical setting, but still accessible and attentive to the conceptual propositions. The approach is that traditional “by components” ; are first introduced affine tensors and then general tensors according to a (questionable) policy of gradualism. The topics that are preliminary to General Relativity do not exceed one third of the book. It may be useful as a summary and reference for some single subjects.

Fleisch, D.A. 2012, *A Student's Guide to Vectors and Tensors*, Cambridge University Press, pag. 133

A very "friendly" introduction to Vector and Tensor Analysis, understandable even without special prerequisites, nevertheless with a good completeness (until the introduction of the Riemann tensor, but without going into curved spaces). The approach to tensor is traditional. Beautiful illustrations, many examples from physics, many calculations carried out in full, together with a speech that gives the impression of proceeding methodically and safely, without jumps and without leaving behind dark spots, make this book an excellent autodidactic tool, also accessible to a good high school student.

- Among the texts of General Relativity (GR), the following contains a sufficiently practicable introduction to Tensor Analysis carried on by a geometrical approach:

Schutz, B. F. 1985, *A First Course in General Relativity*, Cambridge University Press, pag. 376

Classic introductory text to General Relativity, it contains a quite systematic and complete discussion, although fragmented, of tensors. At least in the first part of the book (the one that concerns us here), the discourse is carefully argued, even meticulous, with all the explanations of the case and only few occasional disconnection (even if the impression is sometimes that of a bit cumbersome "machinery"). This is a good introduction to the topic which has the only defect of "flattening" significantly concepts and results without giving a different emphasis depending on their importance. For that it would be better dealing with it with some landmark already acquired. It has been used in various circumstances as a reference while writing these notes.

Dunsby, P. K. S. 2000 *An Introduction to Tensors and Relativity*, University of Cape Town, South Africa, pag. 130

Download: <http://www.mth.uct.ac.za/omei/gr/> (PS format)

This is a schematic but well-made summary that follows the text of Schutz. With few variations, some simplifications and much more incisiveness.

Carroll, S.M. 2004 *Spacetime and geometry*, Addison-Wesley, pag. 513

Text that has established itself as a leader among those recently published on the subject GR. It's a fascinating example of scientific prose of conversational style: reading it sequentially you may have the impression of participating in a course of lectures. Tensors are introduced and developed closely integrated with the GR. This text places itself at a level of difficulty somewhat higher than that of Schutz.

Carroll, S. M. 1997 *Lecture Notes on General Relativity*, University of California Santa Barbara, pag. 231

Download: http://xxx.lanl.gov/PS_cache/gr-qc/pdf/9712/9712019v1.pdf

These are the original readings, in a even more conversational tone, from which the text quoted above has been developed. All that is important is located here, too,

except for some advanced topics of GR.

Dettmann, C. P. 2007 *General Relativity*, University of Bristol, pag.36

Download: http://www.maths.bris.ac.uk/~macpd/gen_rel/bnotes.pdf

A thorough and original synthesis of the entire GR in few pages, including tensors. Readable, despite the compression. The structure of the text makes it easy to consult, even with regard to tensors.

McMahon, D. 2006 *Relativity Demystified*, Mc Graw-Hill, pag. 344

One of the “Self-teaching Guide” of the “Demystified” Series. It aims to achieve a “working knowledge” of the subject, without dwelling too much upon theory and subtilizing concepts. Pragmatic approach, sometimes a little rough, but interesting for a number of exercises developed in full. Useful, but not sufficient for those not satisfied by the approximative.

Ta-Pei Cheng 2010 *Relativity, Gravitation and Cosmology – A Basic introduction*, Oxford University Press, pag.435

A beautiful volume from the Oxford Master Series that deals exhaustively both GR that cosmology. The setting is teaching, and various topics are covered thoroughly, widely, and sometimes in a somewhat insisted way. In short, it is difficult not to understand, but first there is a lot to read. Tensors are first introduced in outline in the early chapters, but then, prior to Part IV "Relativity: full tensor formulation", two chapters are dedicated to them, albeit shrinking to what is interesting for GR. Appreciable discussion and examples of curved spaces. The book is also worthy of consideration for the considerable number of problems proposed, and in many cases solved.