# General Relativity Fall 2017
# Lecture 1: Linear algebra, vectors and tensors

Yacine Ali-Haïmoud
September 5th 2017

The goal of this pure mathematics lecture is to provide a refresher for linear algebra concepts, (re)-introduce tensors, and define our notation.

## A. Basic definitions

Definition: A **real vector space** is a set $\mathcal{V}$ with an internal operation $+$, such that $(\mathcal{V}, +)$ is a commutative group, and an external operation $\cdot$ such that $\forall\ (\lambda, X) \in (\mathbb{R}, \mathcal{V}), \lambda \cdot X \in \mathcal{V}$. This operation is associative: $\lambda \cdot (\mu \cdot X) = (\lambda\mu) \cdot X$, distributive: $(\lambda + \mu) \cdot X = \lambda \cdot X + \mu \cdot X$ and $\lambda \cdot (X + Y) = \lambda \cdot X + \lambda \cdot Y$, and such that $1 \cdot X = X$.

Definition: A **basis** $(e_{(1)}, ..., e_{(n)})$ is a set of linearly independent vectors of $\mathcal{V}$ that spans $\mathcal{V}$, i.e. such that any vector of $\mathcal{V}$ can be written as a linear combination

$$X = X^\mu e_{(\mu)} \equiv \sum_{\mu=1}^{n} X^\mu e_{(\mu)}. \tag{1}$$

The real numbers $X^\mu$ are the **components** of the vector $X$ on the basis $(e_{(1)}, ..., e_{(n)})$. For any finite-dimensional vector space, the number $n$ of basis vectors is independent of the chosen basis and is the **dimension** of the space.

Eq. (1) introduces the **summation convention**: repeated indices are to be summed over, *unless explicitly specified*. We will make sure that pairs of repeated indices always appear as one up and one down.

## B. Examples

- $\mathbb{R}^n$ is a real vector spaces of dimension $n$. $\mathbb{C}^n$ is a real vector space of dimension $2n$.
- Let $\mathcal{F}$ be the set of infinitely differentiable functions $\mathbb{R}^n \to \mathbb{R}$, and define the vector space $\mathcal{V}$ as the set of linear maps $\mathcal{F} \to \mathbb{R}$. Define the sum and external product as they would be naturally. **Exercise:** show that this is a vector space of infinite dimension.
- Another example, which will be the starting point of differential geometry: *derivative operators*. Let $p \in \mathbb{R}^n$. Consider the subset $\mathcal{T}_p$ of the vector space $\mathcal{V}$ of linear operators defined above, that satisfy *Leibniz's rule*: i.e. for all $T \in \mathcal{T}_p$, for all $f, g$ in $\mathcal{F}$, $T(fg) = f(p)T(g) + g(p)T(f)$. Clearly, $\mathcal{T}_p$ contains the usual derivative operators at $p$ $\partial/\partial x^1\big|_p, .., \partial/\partial x^n\big|_p$. **Exercise:** show that this is indeed a vector space. Prove that it has dimension $n$, i.e. that it is the space of directional derivative operators. $\mathcal{T}_p$ is called the **tangent space at** $p$.

## C. Change of basis

Suppose we are given two bases $(e_{(1)}, ..., e_{(n)})$ and $(e'_{(1)}, ..., e'_{(n)})$. Let us denote by ${\Lambda'}^{\nu}{}_{\mu'}$ the $\nu$-th component of $e'_{(\mu')}$ in the unprimed basis and by $\Lambda^{\mu'}{}_{\nu}$ the $\mu'$-th component of $e_{(\nu)}$ in the primed basis:

$$e'_{(\mu')} = {\Lambda'}^{\nu}{}_{\mu'}\ e_{(\nu)}, \tag{2}$$

$$e_{(\nu)} = \Lambda^{\mu'}{}_{\nu}\ e'_{(\mu')}. \tag{3}$$

Substituting Eq. (2) into Eq. (3) and vice-versa we obtain

$$\Lambda^{\nu'}{}_{\nu}{\Lambda'}^{\nu}{}_{\mu'} = \delta^{\nu'}_{\mu'}, \tag{4}$$

$$ {\Lambda'}^{\nu}{}_{\mu'}\Lambda^{\mu'}{}_{\mu} = \delta^{\nu}_{\mu}, \tag{5}$$

which is just stating that $\Lambda$ and $\Lambda'$, seen as matrices, are the inverse of one another. Keeping track of all the primes can be annoying, so unless explicitly specified, we shall keep primes *only on indices* from now on, but one should keep in mind the full underlying meaning. So we write the above equations as

$$e_{(\mu')} = \Lambda^{\nu}{}_{\mu'}\, e_{(\nu)}, \tag{6}$$

$$e_{(\nu)} = \Lambda^{\mu'}{}_{\nu}\, e_{(\mu')}, \tag{7}$$

$$\Lambda^{\nu'}{}_{\nu}\Lambda^{\nu}{}_{\mu'} = \delta^{\nu'}_{\mu'}, \tag{8}$$

$$\Lambda^{\nu}{}_{\mu'}\Lambda^{\mu'}{}_{\mu} = \delta^{\nu}_{\mu}. \tag{9}$$

We denote by $X^\mu$ the components of a vector in the unprimed basis and by $X^{\mu'}$ its components in the primed basis (again, think of this as $X'^{\mu'}$, but we'll only keep the prime on the index for short):

$$X = X^\mu e_{(\mu)} = X^{\mu'} e_{(\mu')}. \tag{10}$$

Using the equations given above, we arrive at (**Exercise:** show this explicitly):

$$X^{\mu'} = \Lambda^{\mu'}{}_{\mu} X^\mu, \qquad X^\mu = \Lambda^{\mu}{}_{\mu'} X^{\mu'}. \tag{11}$$

### D. Dual vectors

DEFINITION: A **dual vector** (or **covector** or **one-form**) is a linear map from $\mathcal{V} \to \mathbb{R}$. The space of dual vectors $\mathcal{V}^*$ is called the **dual space**. It has the same dimension as the vector space $\mathcal{V}$.

Given a basis $(e_{(1)}, ..., e_{(n)})$ we define the the dual basis $(e^{*(1)}, ..., e^{*(n)})$ such that $e^{*(\mu)}(e_{(\nu)}) = \delta^\mu_\nu$. **Exercise:** show that it is a basis of the dual space; conclude that the dual and vector spaces have the same dimensions.

Given a dual vector $X^*$, we can write it as a linear combination of the dual basis vectors:

$$X^* = X^*_\mu\, e^{*(\mu)}, \tag{12}$$

where $X^*_\mu$ are again the components of $X^*$.

**Exercise:** Show that the dual bases of two bases $(e_{(\mu)}), (e_{(\mu')})$ are related through

$$e^{*(\mu')} = \Lambda^{\mu'}{}_{\mu} e^{*(\mu)}, \qquad e^{*(\mu)} = \Lambda^{\mu}{}_{\mu'} e^{*(\mu')}, \tag{13}$$

and that the components of dual vectors transform as

$$X^*_{\mu'} = \Lambda^{\nu}{}_{\mu'} X^*_{\mu}, \qquad X^*_{\mu} = \Lambda^{\mu'}{}_{\mu} X^*_{\mu'}. \tag{14}$$

Hence we see that the transformation law of the components of a vector is identical to the change-of-basis law for dual basis vectors, and conversely, the transformation law of the components of a dual vector is identical to the change-of-basis law for basis vectors. This is why vectors are sometimes referred to as **contravariant** vectors and dual vectors as **covariant** vectors.

We can also define the dual of the dual space $\mathcal{V}_{**}$, and the dual-dual basis $e_{**(\mu)}$ such that $e_{**(\nu)}(e^{*(\mu)}) = \delta^\mu_\nu = e^{*(\mu)}(e_{(\nu)})$. This space has the same dimension as $\mathcal{V}^*$ hence $\mathcal{V}$. For a given vector $X$ we can define the dual-dual vector $X_{**}$ such that for any dual vector $Y^*$, $X_{**}(Y^*) \equiv Y^*(X)$. This mapping is linear, injective (if two vectors have the same image they must be identical), hence bijective since the two spaces have the same dimension. So **we can identify the dual-dual and initial vector space**. From now on we will drop the $*$ when referring to a dual basis: $e^{(\mu)} \equiv e^{*(\mu)}$, and $e_{**(\mu)} \equiv\equiv e_{(\mu)}$. For a vector $X$, we define $X(Y^*) \equiv Y^*(X)$.

### E. Tensors, abstract index notation

DEFINITION: a **tensor** of rank $(k, l)$ is a multilinear map from $\underbrace{\mathcal{V}^* \times ... \times \mathcal{V}^*}_{k \text{ times}} \times \underbrace{\mathcal{V} \times ... \times \mathcal{V}}_{l \text{ times}} \to \mathbb{R}$. So dual vectors are tensors of rank $(0, 1)$ and vectors (seen as dual-dual vectors) are tensors or rank (1, 0).

Given a basis for the initial vector space, a basis for tensors of rank $(k, l)$ is given by the **outer products** $e_{(\mu_1)} \otimes ...e_{(\mu_k)} \otimes e^{(\nu_1)} \otimes ...e^{(\nu_l)}$, where $\mu_i, \nu_i \in [1, ..., n]$. Here we used the identification of the dual-dual with the initial vector space. So we may write a rank $(k, l)$ tensor $\mathbf{T}$ as

$$\mathbf{T} = T^{\mu_1...\mu_k}{}_{\nu_1...\nu_l} e_{(\mu_1)} \otimes ...e_{(\mu_k)} \otimes e^{(\nu_1)} \otimes ...e^{(\nu_l)}, \tag{15}$$

where $T^{\mu_1...\mu_k}{}_{\nu_1...\nu_l}$ are the components of $\mathbf{T}$. **Exercise:** Show that the components of $\mathbf{T}$ are given by

$$T^{\mu_1...\mu_k}{}_{\nu_1...\nu_l} = \mathbf{T}(e^{(\mu_1)}, ..., e^{(\mu_k)}, e_{(\nu_1)}, ..., e_{(\nu_l)}). \tag{16}$$

**Exercise:** Show that under a change of basis (and corresponding change in dual basis), the components of a tensor change as

$$T^{\mu_1'...\mu_k'}{}_{\nu_1'...\nu_l'} = \Lambda^{\mu_1'}{}_{\mu_1}...\Lambda^{\mu_k'}{}_{\mu_k}\Lambda^{\nu_1}{}_{\nu_1'}...\Lambda^{\nu_l}{}_{\nu_l'}T^{\mu_1...\mu_k}{}_{\nu_1...\nu_l}. \tag{17}$$

**Note:** some textbooks may use this transformation law to *define* a tensor. I prefer the geometric, basis-indpendent definition of a multi-linear map on vectors and dual vectors.

An example of tensor of rank $(1, 1)$ is the **identity tensor** $\mathbf{I} = \delta^\alpha{}_\beta$. **Exercise:** Show using the tensor-component transformation law that this indeed transforms as a tensor.

By now it should be clear that vectors, dual vectors and tensors are **geometric objects** that have neaning independently of any basis, and that their components are specific to a given basis. The notation $\mathbf{T}$ clearly does not convey the rank of a tensor, and it would be bulky to keep adding asterisks, arrows, bars, or any other symbols. So we adopt the following **abstract index notation**, and denote by $T^{\alpha_1...\alpha_k}{}_{\beta_1...\beta_l}$ a tensor of rank $(k, l)$, seen as a geometric, basis-independent quantity. For instance, we denote a vector by $X^\alpha$ and a dual vector by $Y_\alpha$, dropping asterisks. $T^{\alpha\beta}{}_{\gamma\delta}$ is a tensor of rank (2, 2), i.e. a linear map from $\mathcal{V}^* \times \mathcal{V}^* \times \mathcal{V} \times \mathcal{V}$.

The obvious disadvantage of this notation is that it looks similar to the components. In order to circumvent this, we will (try to) stick to using the first half of the alphabet to denote a geometric, basis-independent object, and the second half to refer to components in a specific basis. For instance, we may write something like

$$X^\alpha = X^\mu e^\alpha_{(\mu)}. \tag{18}$$

In words, this equation says that the components of the vector $X^\alpha$ on the basis $(e^\alpha_{(1)}, ..., e^\alpha_{(n)})$ are $X^1, ..., X^n$. We will also (try to) stick to having parenthesis on the labeling of basis vectors, to distinguish them from components of co-vectors.

### F. Tensor operations

Consider a rank $(k, l)$ tensor $\mathbf{T}$ and a rank $(p, q)$ tensor $\mathbf{S}$. We define the **outer product** tensor $\mathbf{O}$ as the $(k+p, l+q)$ tensor such that,

$$\mathbf{O}(X^{(1)}, ..., X^{(k)}, ...X^{(k+p)}; Y_{(1)}, ..., Y_{(l)}, ...Y_{(l+q)}) \equiv \mathbf{T}(X^{(1)}, ..., X^{(k)}; Y_{(1)}, ..., Y_{(l)})\mathbf{S}(X^{(k+1)}, ..., X^{(k+p)}; Y_{(l+1)}, ..., Y_{(l+q)}. \tag{19}$$

In abstract index notation, this has the more compact expression

$$O^{\alpha_1...\alpha_k...\alpha_{k+p}}{}_{\beta_1...\beta_l...\beta_{l+q}} \equiv T^{\alpha_1...\alpha_k}{}_{\beta_1...\beta_l}S^{\alpha_{k+1}...\alpha_{k+p}}{}_{\beta_{l+1}...\beta_{l+q}}. \tag{20}$$

Given a tensor $\mathbf{T}$ of rank $(k, l)$ we define the **contraction** on the $p$-th upper index and $q$-th lower index as follows: choose a basis $e_{(\mu)}$ with corresponding dual basis $e^{(\mu)}$, we define the tensor $\mathbf{CT}$ of rank $(k-1, l-1)$ as

$$\mathbf{CT}(X^{(1)}, ..., X^{(k-1)}; Y_{(1)}, ..., Y_{(l-1)}) \equiv \mathbf{T}(X^{(1)}, ... \underset{(p\text{-th slot})}{e^{(\mu)}}, ..., X^{(k-1)}; Y_{(1)}, ... \underset{(q\text{-th slot})}{e_{(\mu)}}, ..., Y_{(l-1)}). \tag{21}$$

**Exercise:** Show that the resulting tensor is independent of the basis chosen to define the contraction. In abstract index notation, this tensor is

$$T^{\alpha_1,...\gamma,...\alpha_{k-1}}{}_{\beta_1,...\gamma,...\beta_{l-1}}, \tag{22}$$

where the repeated index $\gamma$ is at the $p$-th upper slot and the $q$-th lower slot.

We say that a tensor is **symmetric** in the slots $p$ and $q$ if

$$\mathbf{T}(X^{(1)}, ...X^{(p)}, ..., X^{(q)}, ..) = \mathbf{T}(X^{(1)}, ...X^{(q)}, ..., X^{(p)}, ..), \tag{23}$$

or, in abstract index notation,

$$T^{\alpha_1,...,\alpha_p,...,\alpha_q,...} = T^{\alpha_1,...,\alpha_q,...,\alpha_p,...}. \tag{24}$$

For example, a tensor $\mathbf{g}$ of rank $(0, 2)$ is symmetric if $g_{\alpha\beta} = g_{\beta\alpha}$.

The property of **antisymmetry** is similarly defined. For instance, a tensor $\mathbf{F}$ of rank $(0, 3)$ is antisymmetric in its first 2 indices if $F_{\alpha\beta\gamma} = -F_{\beta\alpha\gamma}$.

A tensor of rank $(k, 0)$ or rank $(0, l)$ is **completely symmetric** if it is symmetric in all slot pairs. A **completely antisymmetric** tensor is defined similarly.

**Exercise:** What is the dimension of the space spanned by tensors of rank $(k, 0)$ on a vector space of dimension $n$? What is the dimension of the space spanned by completely symmetric $(0, k)$ tensors? How about completely antisymmetric $(k, 0)$ tensors?

We now define the operations of **symmetrization** and **antisymmetrization**. We first define $\mathcal{P}_n$ as the set of permutations of $[1, ...n]$. For a permutation $\sigma \in \mathcal{P}_n$, we define its **signature** $s(\sigma)$ as $+1$ if $\sigma$ contains an even number of pair exchanges and $-1$ if it contains an odd number. For instance, $[1, 2, 3] \Rightarrow [2, 3, 1]$ has signature $+1$, and $[1, 2, 3] \Rightarrow [2, 1, 3]$ has signature $-1$.

Given a tensor $\mathbf{T}$, we define its symmetric and antisymmetric parts as

$$T^{(\alpha_1...\alpha_n)} \equiv \frac{1}{n!} \sum_{\sigma\in\mathcal{P}_n} T^{\alpha_{\sigma(1)}...\alpha_{\sigma(2)}}, \tag{25}$$

$$T^{[\alpha_1...\alpha_n]} \equiv \frac{1}{n!} \sum_{\sigma\in\mathcal{P}_n} s(\sigma) T^{\alpha_{\sigma(1)}...\alpha_{\sigma(2)}} \tag{26}$$

These tensors are, respectively, completely symmetric and antisymmetric. For instance

$$T_{(\alpha\beta)} \equiv \frac{1}{2}\left(T_{\alpha\beta} + T_{\beta\alpha}\right), \tag{27}$$

$$T_{[\alpha\beta]} \equiv \frac{1}{2}\left(T_{\alpha\beta} - T_{\beta\alpha}\right). \tag{28}$$

**Exercise:** Write explicitly the completely symmetric and antisymmetric parts of a tensor of rank $(0, 3)$. Note that only tensors of rank $(2, 0)$ [or $(0, 2)$] can be written as the sum of their symmetric part and antisymmetric parts.

### G. Metric

In order to do physics, we require more structure: that of a *metric*. The metric is a central object in general relativity, which conveys the notions of distances and angles (among other things). For now we will be content with a mathematical definition. When we say "the" metric we assume that there is a preferred metric on the vector space of interest.

Definition. A **metric** is a symmetric, non-degenerate tensor of rank $(0, 2)$. Non-degenerate means that $g_{\alpha\beta}A^\beta = 0 \Leftrightarrow A^\alpha = 0$, i.e. there cannot be any non-zero vector $A^\alpha$ such that the dual vector $g_{\alpha\beta}A^\beta$ is the zero dual vector.

Two vectors $X, Y$ are said to be **orthogonal** if $\mathbf{g}(X, Y) = g_{\alpha\beta}X^\alpha Y^\beta = 0$. A vector $X$ is said to have unit norm is $\mathbf{g}(X, X) = \pm 1$, and is said to be **null** if $\mathbf{g}(X, X) = 0$, i.e. $X$ is orthogonal to itself.

**Exercise:** show that we can always find a basis in which $g_{\mu\nu} = \pm\delta_{\mu\nu}$, i.e. in which only the diagonal components are non-vanishing, and they are all equal to plus or minus one. Such a basis is an **orthonormal basis**.

**Exercise:** (Sylvester's law of inertia) show that the number of pluses and minuses (the **signature** of the metric) is independent of the basis in which the metric is diagonalized.

Definition: The **inverse metric** $g^{\alpha\beta}$ is the unique symmetric, non-degenerate tensor of rank $(2, 0)$ such that $g^{\alpha\beta}g_{\beta\gamma} = \delta^\alpha_\gamma$. The fact that this tensor must be non-degenerate is easy to prove: if $g^{\alpha\beta}A_\beta = 0$, then, multiplying by $g_{\gamma\alpha}$, we get $A_\gamma = 0$. The fact that it is unique follows from the non-degeneracy of $\mathbf{g}$ (**Exercise:** prove it). Finally, the existence can be proven by explicitly defining $g^{\mu\nu} = g_{\mu\nu}$ in an orthonormal basis.

The metric tensor can be used to map vectors to one-forms: given a vector $X^\alpha$, we may define the one-form $X^*_\alpha$ such that, for any vector $Y$, $X^*(Y) \equiv \mathbf{g}(X, Y)$. In abstract index notation, we have $X^*_\beta Y^\beta = g_{\alpha\beta}X^\alpha Y^\beta$ for all $Y$, implying $X^*_\alpha \equiv g_{\alpha\beta}X^\beta$. So the **metric can be used to lower indices**. We will no longer write the asteriks explicitly. Note that without a metric, we did not have a basis-independent way to define a dual vector given a vector. Similarly, we may **use the inverse metric to raise indices**: given a one-form $X_\alpha$ we define the vector $X^\alpha \equiv g^{\alpha\beta} X_\beta$.

More generally, **the metric and inverse metric can be used to raise and lower indices of a tensor of arbitrary rank**, for instance

$$T_\alpha{}^\beta{}_\gamma{}^\delta \equiv g_{\alpha\sigma}\, g^{\delta\rho}\, T^{\sigma\beta}{}_{\gamma\rho}. \tag{29}$$

Finally, note that $g^\alpha_{\ \gamma} \equiv g^{\alpha\beta} g_{\beta\gamma} = \delta^\alpha_\gamma$.