TENSOR CALCULUS

Part 1: tensor algebra

Marko Vojinović

March 2010

Contents

Chapter 1

Tensor algebra

1.1 Vectors

Pick a D-dimensional manifold $\mathcal{M}$ (we shall be mainly interested with the most relevant case for physics, $D = 4$), and some point $\mathcal{P} \in \mathcal{M}$ on it.

Draw an arbitrary curve $\mathcal{C}$ in $\mathcal{M}$ passing through $\mathcal{P}$, and parametrize it using some parameter λ. Now focus on the **directional derivative operator** of $\mathcal{C}$ with respect to λ, evaluated at $\mathcal{P}$:

$$\left.\frac{d}{d\lambda}\right|_{\lambda(\mathcal{P})}$$

Though it may sound a little strange at this point, call this quantity the **tangent vector** of the curve $\mathcal{C}$ at point $\mathcal{P}$, and denote it with a boldface letter $\boldsymbol{A}$:

$$\boldsymbol{A} \equiv \left.\frac{d}{d\lambda}\right|_{\lambda(\mathcal{P})}.$$

Now construct a local coordinate system in the neighborhood of $\mathcal{P}$. It is consisted of D curves, which are parametrized by parameters called **coordinates**, and denoted x^μ. Assign a directional derivative at $\mathcal{P}$ to all these curves, thus obtaining a set of **coordinate tangent vectors** for coordinate lines. Denote these as:

$$\mathbf{e}_\mu \equiv \left.\frac{\partial}{\partial x^\mu}\right|_{x^\mu(\mathcal{P})},$$

ie. explicitly[1]:

$$\mathbf{e}_0 \equiv \left.\frac{\partial}{\partial x^0}\right|_{x^0(\mathcal{P})}, \qquad \mathbf{e}_1 \equiv \left.\frac{\partial}{\partial x^1}\right|_{x^1(\mathcal{P})}, \qquad \dots \qquad \mathbf{e}_{D-1} \equiv \left.\frac{\partial}{\partial x^{D-1}}\right|_{x^{D-1}(\mathcal{P})}.$$

On one hand, any curve $\mathcal{C}$ (passing through $\mathcal{P}$ and parametrized with λ) can be written as a map from a real line to a manifold,

$$\mathcal{C}: \qquad \mathbb{R} \to \mathcal{M}, \qquad \lambda \mapsto \mathcal{P}.$$

On the other hand, it can be written in a coordinate frame as a set of functions $x^\mu(\lambda)$,

$$\bar{\mathcal{C}}: \qquad \mathbb{R} \to \mathbb{R}^D, \qquad \lambda \mapsto x^\mu(\lambda),$$

while these functions are mapped to the manifold via coordinate curves,

$$\phi: \qquad \mathbb{R}^D \to \mathcal{M}, \qquad x^\mu \mapsto \mathcal{P},$$

[1]In general relativity physics one typically counts coordinates from 0 to $D-1$, rather than from 1 to D. The x^0 coordinate is usually interpreted as *coordinate time*, although it need not have any relation to the "real" physical time.

such that $\mathcal{C} = \bar{\mathcal{C}} \circ \phi$. Knowing this, we can always write (according to the usual rules of calculus in $\mathbb{R}^D$):

$$\frac{d}{d\lambda}\Big|_{\lambda(\mathcal{P})} = \frac{dx^\mu}{d\lambda}\frac{\partial}{\partial x^\mu}\Big|_{\lambda(\mathcal{P})},$$

which can be rewritten in our new notation as

$$\boldsymbol{A} = A^\mu \mathbf{e}_\mu \equiv A^0\mathbf{e}_0 + A^1\mathbf{e}_1 + \cdots + A^{D-1}\mathbf{e}_{D-1}.$$

Here we have introduced the notation $A^\mu \equiv \dfrac{dx^\mu}{d\lambda}$ for the coefficients.

Now use this to define the concept of **tangent space**:

> Consider all possible curves $\mathcal{C}$ parametrized in all possible ways and passing through a single point $\mathcal{P} \in \mathcal{M}$. The **set of all their directional derivatives evaluated at** $\mathcal{P}$ satisfies the axioms of a **vector space**, isomorphic to $\mathbb{R}^D$. Call this space **tangent space of** $\mathcal{M}$ **at** $\mathcal{P}$, and denote it as $T\mathcal{M}_\mathcal{P}$ (or just $T\mathcal{M}$, when the choice of the point is unambiguous).

The elements of this space are called **tangent vectors at** $\mathcal{P}$. The tangent vectors of D coordinate lines provide a natural **basis** in this vector space, and is usually denoted $\mathbf{e}_\mu$. As we have seen above, any vector $\boldsymbol{A} \in T\mathcal{M}$ can be represented as a linear combination of basis vectors,

$$\boldsymbol{A} = A^\mu \mathbf{e}_\mu, \qquad \forall \boldsymbol{A} \in T\mathcal{M}.$$

The coefficients in the expansion are called **components of vector** $\boldsymbol{A}$ **in basis** $\mathbf{e}_\mu$.

Note that the choice of basis is quite arbitrary, so the components A^μ are *not unique* for a single unique vector $\boldsymbol{A}$. Of course, once the basis vectors are fixed, the components also become unique. It is important to emphasize that each vector $\boldsymbol{A}$ is a **geometric object**, and exists regardless of any coordinates, basis vectors and components one may or may not assign to it. It is defined as a directional derivative of some curve passing through $\mathcal{P}$, and need no coordinates for its definition.

1.2 1-forms

Consider *linear functionals* of the tangent space $T\mathcal{M}_\mathcal{P}$, ie. functions that map vectors into numbers,

$$\boldsymbol{f}: \qquad T\mathcal{M} \to \mathbb{R}, \qquad \boldsymbol{A} \mapsto \boldsymbol{f}[\boldsymbol{A}] \in \mathbb{R},$$

and are linear:

$$\boldsymbol{f}[a\boldsymbol{A} + b\boldsymbol{B}] = a\boldsymbol{f}[\boldsymbol{A}] + b\boldsymbol{f}[\boldsymbol{B}], \qquad \forall a, b \in \mathbb{R}, \quad \forall \boldsymbol{A}, \boldsymbol{B} \in T\mathcal{M}_\mathcal{P}.$$

A set of all these linear functionals over $T\mathcal{M}_\mathcal{P}$ also has an algebraic structure of a **vector space**, and is called a **dual tangent space** of $T\mathcal{M}_\mathcal{P}$, and denoted $T\mathcal{M}^*_\mathcal{P}$ (or just $T\mathcal{M}^*$ when the choice of the point $\mathcal{P} \in \mathcal{M}$ is unambiguous). Linear functionals $\boldsymbol{f}$, elements of this space, are called **1-forms**.

It can be shown that the space $T\mathcal{M}^*$ is also D-dimensional, and that its dual space (the dual of a dual) is isomorphic to the original tangent space $T\mathcal{M}$.

Given a set of basis vectors $\mathbf{e}_\mu$ in $T\mathcal{M}$, one can naturally construct a set of basis functionals in $T\mathcal{M}^*$, denoted $\mathbf{e}^\mu$, via the *biorthogonality relation*:

$$\mathbf{e}^\mu[\mathbf{e}_\nu] = \delta^\mu_\nu \equiv \begin{cases} 1 & \text{for } \mu = \nu, \\ 0 & \text{for } \mu \neq \nu. \end{cases}$$

This set of basis 1-forms $\mathbf{e}^\mu$ is said to be **biorthogonal** to the basis vectors $\mathbf{e}_\mu$.

Once the basis 1-forms have been chosen, one can expand any 1-form $\boldsymbol{f}$ using this basis as:

$$\boldsymbol{f} = f_\mu \mathbf{e}^\mu \equiv f_1\mathbf{e}^1 + f_2\mathbf{e}^2 + \cdots + f_{D-1}\mathbf{e}^{D-1},$$

where the coefficients f_μ are called **components of the 1-form $\boldsymbol{f}$**. Also, given an expansion of some vector $\boldsymbol{A} \in T\mathcal{M}$ with respect to basis $\mathbf{e}_\mu$, $\boldsymbol{A} = A^\mu \mathbf{e}_\mu$, one can calculate the action of $\boldsymbol{f}$ on $\boldsymbol{A}$ using the biorthogonality relation and linearity of $\boldsymbol{f}$ as:

$$\boldsymbol{f}[\mathbf{A}] = f_\mu \mathbf{e}^\mu[\mathbf{A}] = f_\mu \mathbf{e}^\mu[A^\nu \mathbf{e}_\nu] = f_\mu A^\nu \mathbf{e}^\mu[\mathbf{e}_\nu] = f_\mu A^\nu \delta^\mu_\nu = f_\mu A^\mu \in \mathbb{R}.$$

This process is called **contraction**.

If the tangent space basis $\mathbf{e}_\mu$ was chosen to be a set of tangent vectors to coordinate curves in $\mathcal{M}$ (it need not be chosen this way!!!), ie. if

$$\mathbf{e}_\mu = \frac{\partial}{\partial x^\mu}\Big|_{x^\mu(\mathcal{P})},$$

it is called **coordinate basis of tangent vectors**, and the corresponding biorthogonal basis of 1-forms is also called **coordinate basis of 1-forms**. It also deserves special notation:

$$e^\mu = \boldsymbol{d}x^\mu.$$

We shall explain later the distinction between coordinate and noncoordinate bases, and the meaning of the symbol $\boldsymbol{d}$.

1.3 The $\times$, $\otimes$ and $\oplus$

Consider two arbitrary nonempty sets, A and B. One can construct their **Cartesian product** as a set of all ordered pairs of all their elements:

$$A \times B = \{(a, b) \,|\, a \in A, b \in B\}.$$

If the two sets are not just any sets, but have some algebraic structure, one may wish to supplement the $\times$ with additional axioms which will make the product "behave well".

For example, if U and V are vector spaces, they already carry two operations, *addition of vectors* and *multiplication of a vector with a scalar*. So construct a Cartesian product of two vector spaces, and provide appropriate axioms such that it is linear with respect to addition and multiplication with a scalar (denote it as $\otimes$ in order to distinguish it from general $\times$):

$$(\boldsymbol{u}_1 + \boldsymbol{u}_2) \otimes \mathbf{v} = \boldsymbol{u}_1 \otimes \mathbf{v} + \boldsymbol{u}_2 \otimes \mathbf{v}, \qquad \boldsymbol{u} \otimes (\mathbf{v}_1 + \mathbf{v}_2) = \boldsymbol{u} \otimes \mathbf{v}_1 + \boldsymbol{u} \otimes \mathbf{v}_2,$$

$$(c\boldsymbol{u}) \otimes \mathbf{v} = c(\boldsymbol{u} \otimes \mathbf{v}), \qquad \boldsymbol{u} \otimes (c\mathbf{v}) = c(\boldsymbol{u} \otimes \mathbf{v}),$$

$$\forall \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u} \in U, \qquad \forall \mathbf{v}_1, \mathbf{v}_2, \mathbf{v} \in V, \qquad \forall c \in \mathbb{R}.$$

Call it the **tensor product** of vector spaces U and V. The tensor product is constructed in such a way that it represents the most general bilinear operation. The resulting set $U \otimes V$ is also vector space, which dimension is the product of dimensions of U and V. It is the **largest possible space** that contains U and V as subspaces (once). Note:

- The two vector spaces need not be the same, so the tensor product is **not commutative**:

  $$\boldsymbol{u} \otimes \mathbf{v} \neq \mathbf{v} \otimes \boldsymbol{u}.$$

  It is not commutative even if the two spaces U and V are the same.

- The two vector spaces need to be constructed over the same field of scalars (in this case $\mathbb{R}$).

- An arbitrary vector from $U \otimes V$ in general **cannot be written as a product $\boldsymbol{u} \otimes \mathbf{v}$**, but only as a linear combination of such products.

Given two arbitrary matrices, one can construct their tensor product by multiplying all their elements in all possible combinations. For example,

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \otimes \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax & bx & cx \\ ay & by & cy \\ az & bz & cz \\ dx & ex & fx \\ dy & ey & fy \\ dz & ez & fz \end{bmatrix}.$$

In general, if one matrix is of type $m \times n$ and the other of type $p \times q$, then their tensor product is a matrix of type $mp \times nq$.

The tensor product is also called **Kronecker product** and **outer multiplication**.

As another example, assume again that U and V are vector spaces. Construct again their Cartesian product, but this time provide somewhat different axioms with respect to addition and multiplication with a scalar (denote it as $\oplus$ in order to distinguish it from $\otimes$ and $\times$):

$$(\boldsymbol{u}_1 + \boldsymbol{u}_2) \oplus (\boldsymbol{v}_1 + \boldsymbol{v}_2) = \boldsymbol{u}_1 \oplus \boldsymbol{v}_1 + \boldsymbol{u}_2 \oplus \boldsymbol{v}_2,$$

$$c(\boldsymbol{u} \oplus \boldsymbol{v}) = (c\boldsymbol{u}) \oplus (c\boldsymbol{v}),$$

$$\forall \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u} \in U, \qquad \forall \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v} \in V, \qquad \forall c \in \mathbb{R}.$$

Call it the **direct sum** of vector spaces U and V. The direct sum is constructed such that the resulting set $U \oplus V$ is also a vector space, which dimension is the sum of dimensions of U and V. It is the **smallest possible space** that contains U and V as subspaces (independently). Note:

- The two vector spaces need not be the same, so the direct sum is **not commutative**:

  $$\boldsymbol{u} \oplus \boldsymbol{v} \neq \boldsymbol{v} \oplus \boldsymbol{u}.$$

  It is not commutative even if U and V are the same.

- The two vector spaces need to be constructed over the same field of scalars (in this case $\mathbb{R}$).

- An arbitrary vector from $U \oplus V$ **can always be written as a sum $\boldsymbol{u} \oplus \boldsymbol{v}$**, where $\boldsymbol{u}$ and $\boldsymbol{v}$ are some vectors from U and V.

Given two arbitrary matrices, one can construct their direct sum by combining them in a bigger block-diagonal matrix. For example,

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \oplus \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \left[\begin{array}{ccc|c} a & b & c & 0 \\ d & e & f & 0 \\ \hline 0 & 0 & 0 & x \\ 0 & 0 & 0 & y \\ 0 & 0 & 0 & z \end{array}\right].$$

In general, if one matrix is of type $m \times n$ and the other of type $p \times q$, then their direct sum is a matrix of type $(m + p) \times (n + q)$.

1.4 Tensor algebra

Return again to the tangent space $T\mathcal{M}_{\mathcal{P}}$ and its dual, $T\mathcal{M}^*_{\mathcal{P}}$. Use the tensor product and direct sum to construct a vector space of **tensors of type** (p, q):

$$T_{p,q} \equiv \underbrace{T\mathcal{M} \otimes \cdots \otimes T\mathcal{M}}_{p \text{ times}} \otimes \underbrace{T\mathcal{M}^* \otimes \cdots \otimes T\mathcal{M}^*}_{q \text{ times}} \oplus \text{ all possible permutations.}$$

Tensors of type (p,q) are said to be p **times contravariant and** q **times covariant**. The reason for this terminology will be explained later.

Take an example for an illustration. A space of tensors of type $(2,1)$ is

$$T_{2,1} = T\mathcal{M} \otimes T\mathcal{M} \otimes T\mathcal{M}^* \quad \oplus \quad T\mathcal{M} \otimes T\mathcal{M}^* \otimes T\mathcal{M} \quad \oplus \quad T\mathcal{M}^* \otimes T\mathcal{M} \otimes T\mathcal{M}.$$

Given bases $\mathbf{e}_\mu$ and $\mathbf{e}^\mu$ in $T\mathcal{M}$ and $T\mathcal{M}^*$, we can represent any tensor from $T_{2,1}$ as a linear combination in appropriate basis. For example, any tensor $\boldsymbol{A}$ from the space $T\mathcal{M} \otimes T\mathcal{M} \otimes T\mathcal{M}^*$ can be written as

$$\boldsymbol{A} = A^{\mu\nu}{}_{\lambda} \mathbf{e}_\mu \otimes \mathbf{e}_\nu \otimes \mathbf{e}^\lambda,$$

and we can write a corresponding tensor from $T_{2,1}$ as:

$$\tilde{\boldsymbol{A}} = \boldsymbol{A} \oplus \boldsymbol{0} \oplus \boldsymbol{0} = (A^{\mu\nu}{}_{\lambda} \mathbf{e}_\mu \otimes \mathbf{e}_\nu \otimes \mathbf{e}^\lambda) \oplus \boldsymbol{0} \oplus \boldsymbol{0}.$$

Here $\boldsymbol{0}$ is the zero vector from spaces $T\mathcal{M} \otimes T\mathcal{M}^* \otimes T\mathcal{M}$ and $T\mathcal{M}^* \otimes T\mathcal{M} \otimes T\mathcal{M}$. Since there is a natural correspondence between $\tilde{\boldsymbol{A}}$ and $\boldsymbol{A}$, one usually omits these extra zeroes, and writes $\boldsymbol{A} \equiv \tilde{\boldsymbol{A}} \in T_{2,1}$.

Using this convention, there are two more types of tensors which belong to $T_{2,1}$:

$$\boldsymbol{B} = B^{\mu}{}_{\nu}{}^{\lambda} \mathbf{e}_\mu \otimes \mathbf{e}^\nu \otimes \mathbf{e}_\lambda, \qquad \boldsymbol{C} = C_{\mu}{}^{\nu\lambda} \mathbf{e}^\mu \otimes \mathbf{e}_\nu \otimes \mathbf{e}_\lambda.$$

Note that there is natural *morphism* between vector spaces $T\mathcal{M} \otimes T\mathcal{M} \otimes T\mathcal{M}^*$, $T\mathcal{M} \otimes T\mathcal{M}^* \otimes T\mathcal{M}$ and $T\mathcal{M}^* \otimes T\mathcal{M} \otimes T\mathcal{M}$. Namely, for every tensor $\boldsymbol{B} \in T\mathcal{M} \otimes T\mathcal{M}^* \otimes T\mathcal{M}$ there is a corresponding (non-unique!) tensor $\tilde{\boldsymbol{B}} \in T\mathcal{M} \otimes T\mathcal{M} \otimes T\mathcal{M}^*$ such that both of them have the same components (in the same basis):

$$\boldsymbol{B} = B^{\mu}{}_{\nu}{}^{\lambda} \mathbf{e}_\mu \otimes \mathbf{e}^\nu \otimes \mathbf{e}_\lambda \qquad \Rightarrow \qquad \tilde{\boldsymbol{B}} = B^{\mu}{}_{\nu}{}^{\lambda} \mathbf{e}_\mu \otimes \mathbf{e}_\lambda \otimes \mathbf{e}^\nu,$$

and similarly for $\boldsymbol{C} \in T\mathcal{M}^* \otimes T\mathcal{M} \otimes T\mathcal{M}$. It is obvious that these tensors, although different, contain the same information. One can consider appropriate equivalence classes across these spaces, and choose a representative from the appropriate space as one sees fit. Therefore it is not necessary to make a sharp distinction between the three types of tensors, and we can consider only the complete big space $T_{2,1}$.

In a similar fashion, one can write down tensors and their components and bases for all other spaces $T_{p,q}$. For example, a tensor from the space $T_{4,3}$ could be written as

$$\boldsymbol{A} = A^{\mu\nu}{}_{\rho\sigma\lambda}{}^{\alpha\beta} \mathbf{e}_\mu \otimes \mathbf{e}_\nu \otimes \mathbf{e}^\rho \otimes \mathbf{e}^\sigma \otimes \mathbf{e}^\lambda \otimes \mathbf{e}_\alpha \otimes \mathbf{e}_\beta.$$

Now **define the space of tensors of type** $(0,0)$, $T_{0,0}$, to be just the field of numbers, $\mathbb{R}$,

$$T_{0,0} \equiv \mathbb{R}.$$

Tensors of this type are called **scalars**. Using this, one can make the following final definition:

The direct sum of all possible tensor spaces of type (p,q) over all possible values for p and q,

$$T(\mathcal{P}) \equiv \bigoplus_{p=0}^{\infty} \bigoplus_{q=0}^{\infty} T_{p,q}(\mathcal{P}),$$

is called the **tensor algebra** at point $\mathcal{P}$. It is one of the most general algebraic structures one can construct from a given vector space. Its basic operations are

- multiplication with a scalar (carried over from the vector space $T\mathcal{M}$),
- addition (also carried over from the vector space $T\mathcal{M}$), and
- multiplication (the tensor product, $\otimes$).

Explicitly, the structure of the tensor algebra is the following:

$$\begin{array}{rll} T & = \mathbb{R} & \text{scalars} \\ & \oplus\ T\mathcal{M} & \text{vectors} \\ & \oplus\ T\mathcal{M}^* & \text{1-forms} \\ & \oplus\ T\mathcal{M}\otimes T\mathcal{M} & \text{tensors of type } (2,0) \\ & \oplus\ T\mathcal{M}\otimes T\mathcal{M}^* \quad \oplus \quad T\mathcal{M}^*\otimes T\mathcal{M} & \text{tensors of type } (1,1) \\ & \oplus\ T\mathcal{M}^*\otimes T\mathcal{M}^* & \text{tensors of type } (0,2) \\ & \oplus\ T\mathcal{M}\otimes T\mathcal{M}\otimes T\mathcal{M} & \text{tensors of type } (3,0) \\ & \oplus\ \ldots & \text{and so on.} \end{array}$$

1.5 Elementary operations on tensors

There are several various operations on tensors one can construct within the tensor algebra. We shall introduce them all one by one, typically via examples.

Addition. While addition is defined overall in the tensor algebra, it is nontrivial only when one considers tensors of the same type (p, q). For example, given $\boldsymbol{A}, \boldsymbol{B} \in T_{2,1}$, we have

$$\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B} = (A^{\mu\nu}{}_{\lambda} + B^{\mu\nu}{}_{\lambda})\, \mathbf{e}_\mu \otimes \mathbf{e}_\nu \otimes \mathbf{e}^\lambda, \qquad \text{or in components:} \qquad C^{\mu\nu}{}_{\lambda} = A^{\mu\nu}{}_{\lambda} + B^{\mu\nu}{}_{\lambda}.$$

Multiplication with a scalar. Any tensor can be multiplied with a number. For example, given a tensor $\boldsymbol{A} \in T_{2,1}$ and a scalar $c \in \mathbb{R}$, we have

$$\boldsymbol{B} = c\,\boldsymbol{A} = cA^{\mu\nu}{}_{\lambda}\mathbf{e}_\mu \otimes \mathbf{e}_\nu \otimes \mathbf{e}^\lambda, \qquad \text{or in components:} \qquad B^{\mu\nu}{}_{\lambda} = cA^{\mu\nu}{}_{\lambda}.$$

Tensor product. Also called *outer multiplication* and *Kronecker product*, it has already been introduced. It is defined for all tensors. For example, given $\boldsymbol{A} \in T_{1,0}$ and $\boldsymbol{B} \in T_{2,1}$, we have

$$\boldsymbol{C} = \boldsymbol{A} \otimes \boldsymbol{B} = A^\rho B^{\mu\nu}{}_{\lambda}\mathbf{e}_\rho \otimes \mathbf{e}_\mu \otimes \mathbf{e}_\nu \otimes \mathbf{e}^\lambda, \qquad \text{or in components:} \qquad C^{\rho\mu\nu}{}_{\lambda} = A^\rho B^{\mu\nu}{}_{\lambda}.$$

Note:

- The resulting tensor has components which are just ordinary multiplication of components of $\boldsymbol{A}$ and $\boldsymbol{B}$, in all possible combinations.
- $\boldsymbol{C} \in T_{3,1}$, which means that the result is **outside** of both spaces $T_{1,0}$ and $T_{2,1}$ (this is the reason for the name "outer multiplication").
- The product $\boldsymbol{A} \otimes \boldsymbol{B} \neq \boldsymbol{B} \otimes \boldsymbol{A}$ because $\otimes$ is not commutative. However, both products "carry the same information" since the components are multiplied using ordinary real number multiplication, which *is* commutative.
- Multiplication with a scalar can be considered as a special case of tensor product between a tensor of type $T_{0,0}$ with some other arbitrary tensor.

Contraction. Contraction is the procedure of using some piece of a tensor coming from $T\mathcal{M}^*$ and letting it act on some other piece coming from $T\mathcal{M}$ as a functional. For example, given $\boldsymbol{A} \in T_{2,1}$, we can write it down in a basis,

$$\boldsymbol{A} = A^{\mu\nu}{}_{\lambda}\mathbf{e}_\mu \otimes \mathbf{e}_\nu \otimes \mathbf{e}^\lambda \qquad \in \qquad T\mathcal{M} \otimes T\mathcal{M} \otimes T\mathcal{M}^*.$$

Now let the $T\mathcal{M}^*$ part act on the first $T\mathcal{M}$ part, as follows:

$$\begin{array}{rcll} \boldsymbol{B} = C_{1,3}(\boldsymbol{A}) & = & A^{\mu\nu}{}_{\lambda} C_{1,3}(\mathbf{e}_\mu \otimes \mathbf{e}_\nu \otimes \mathbf{e}^\lambda) & \text{linearity of functionals} \\ & = & A^{\mu\nu}{}_{\lambda}\mathbf{e}_\nu \otimes \mathbf{e}^\lambda[\mathbf{e}_\mu] & \text{definition: action of “3 on 1”} \\ & = & A^{\mu\nu}{}_{\lambda}\mathbf{e}_\nu \otimes \delta^\lambda_\mu & \text{biorthogonality relation} \\ & = & A^{\mu\nu}{}_{\lambda}\delta^\lambda_\mu \mathbf{e}_\nu & \text{tensor multiplication with a scalar} \\ & = & A^{\mu\nu}{}_{\mu}\mathbf{e}_\nu & \text{Einstein summation convention} \end{array}$$

In component language,

$$B^{\nu} \equiv C_{1,3}(A^{\mu\nu}{}_{\lambda}) = A^{\mu\nu}{}_{\mu}.$$

Note:

- For the example above, one can define contractions $C_{1,3}$ and $C_{2,3}$ which are different operations in general (in components, $A^{\mu\nu}{}_{\mu}$ and $A^{\mu\nu}{}_{\nu}$). For a tensor of type (p,q) a total of pq different contractions are possible.
- Given a tensor of type (p,q), the result of any of its contractions is a tensor of type $(p-1,q-1)$. One basis 1-form $\mathbf{e}^{\lambda}$ always "eats itself out" with one basis vector $\mathbf{e}_{\mu}$.
- Tensors of type $(p,0)$ and $(0,q)$ cannot be contracted.
- Tensors of type $(1,1)$ can be represented as matrices. The (single definable) contraction of these tensors is equivalent to **taking the trace** of the corresponding matrix, ie. summing the diagonal elements. Therefore, contraction is the **generalization of the idea of trace** from $(1,1)$ tensors to arbitrary tensors. However, tensors of type $(2,0)$ and $(0,2)$, although representable in matrix form, *cannot be contracted*, regardless of any matrix trace operation.
- The contraction is independent of the choice of the basis vectors. This is not obvious, and will be proved later when we discuss transformations between bases.

Inner product. Also known as **matrix multiplication**, it is the sequence of taking the tensor product of two tensors and then contracting them in some way. For example, given two tensors, $\boldsymbol{A} \in T_{2,1}$ and $\boldsymbol{B} \in T_{1,1}$, we can write:

$$\boldsymbol{C} = \boldsymbol{A} \cdot \boldsymbol{B} = C_{3,4}(\boldsymbol{A} \otimes \boldsymbol{B}) = A^{\mu\nu}{}_{\lambda} B^{\lambda}{}_{\rho} \mathbf{e}_{\mu} \otimes \mathbf{e}_{\nu} \otimes \mathbf{e}^{\rho} \qquad \text{or in components:} \qquad C^{\mu\nu}{}_{\rho} = A^{\mu\nu}{}_{\lambda} B^{\lambda}{}_{\rho}.$$

As another example, we consider matrices, ie. tensors of type $(1,1)$. The inner product of two such tensors,

$$\boldsymbol{A} \cdot \boldsymbol{B} = C_{2,3}(\boldsymbol{A} \otimes \boldsymbol{B}) = A^{\mu}{}_{\lambda} B^{\lambda}{}_{\rho} \mathbf{e}_{\mu} \otimes \mathbf{e}^{\rho},$$

is again a tensor of type $(1,1)$, ie. a matrix (so that the product of two elements in $T_{1,1}$ is again an element of $T_{1,1}$ — hence the name "inner" product). From the component language it is obvious that this is exactly the usual matrix multiplication.

Note that matrix multiplication is not commutative because tensor product $\otimes$ is not commutative.

Transpose. The transpose of a tensor is defined by interchanging the positions of two basis vectors in a tensor product:

$$(\mathbf{e}_{\mu} \otimes \mathbf{e}_{\nu})^T = \mathbf{e}_{\nu} \otimes \mathbf{e}_{\mu}, \qquad (\mathbf{e}_{\mu} \otimes \mathbf{e}^{\nu})^T = \mathbf{e}^{\nu} \otimes \mathbf{e}_{\mu}, \qquad (\mathbf{e}^{\mu} \otimes \mathbf{e}^{\nu})^T = \mathbf{e}^{\nu} \otimes \mathbf{e}^{\mu}.$$

For example, given a tensor $\boldsymbol{A} \in T_{2,1}$, one can define three different transposes, $\boldsymbol{A}^{T_{12}}$, $\boldsymbol{A}^{T_{13}}$ and $\boldsymbol{A}^{T_{23}}$, as follows:

$$\boldsymbol{B}_1 = \boldsymbol{A}^{T_{12}} = A^{\mu\nu}{}_{\lambda}(\mathbf{e}_{\mu} \otimes \mathbf{e}_{\nu} \otimes \mathbf{e}^{\lambda})^{T_{12}} = A^{\mu\nu}{}_{\lambda} \mathbf{e}_{\nu} \otimes \mathbf{e}_{\mu} \otimes \mathbf{e}^{\lambda}, \qquad \text{or in components:} \qquad B^{\nu\mu}{}_{\lambda} = A^{\mu\nu}{}_{\lambda},$$

$$\boldsymbol{B}_2 = \boldsymbol{A}^{T_{13}} = A^{\mu\nu}{}_{\lambda}(\mathbf{e}_{\mu} \otimes \mathbf{e}_{\nu} \otimes \mathbf{e}^{\lambda})^{T_{13}} = A^{\mu\nu}{}_{\lambda} \mathbf{e}^{\lambda} \otimes \mathbf{e}_{\nu} \otimes \mathbf{e}_{\mu}, \qquad \text{or in components:} \qquad B_{\lambda}{}^{\nu\mu} = A^{\mu\nu}{}_{\lambda},$$

$$\boldsymbol{B}_3 = \boldsymbol{A}^{T_{23}} = A^{\mu\nu}{}_{\lambda}(\mathbf{e}_{\mu} \otimes \mathbf{e}_{\nu} \otimes \mathbf{e}^{\lambda})^{T_{23}} = A^{\mu\nu}{}_{\lambda} \mathbf{e}_{\mu} \otimes \mathbf{e}^{\lambda} \otimes \mathbf{e}_{\nu}, \qquad \text{or in components:} \qquad B^{\mu}{}_{\lambda}{}^{\nu} = A^{\mu\nu}{}_{\lambda}.$$

Essentially, we are switching positions between a pair of indices. The transpose operation is a straightforward generalization of a transpose of a matrix.

Exterior product. Also called **wedge product**, it is a completely antisymmetrized tensor product, and is denoted as $\wedge$. For example, given two vectors, $\boldsymbol{A}, \boldsymbol{B} \in T_{1,0}$, we have

$$\boldsymbol{C} = \boldsymbol{A} \wedge \boldsymbol{B} = \boldsymbol{A} \otimes \boldsymbol{B} - \boldsymbol{B} \otimes \boldsymbol{A} = (A^{\mu} B^{\nu} - A^{\nu} B^{\mu}) \mathbf{e}_{\mu} \otimes \mathbf{e}_{\nu},$$

or in component language,

$$C^{\mu\nu} = A^\mu B^\nu - A^\nu B^\mu.$$

Note:

- The resulting tensor $\boldsymbol{C}$ in the above example is called a **bivector**. The result of taking the wedge of three vectors, $\boldsymbol{A} \wedge \boldsymbol{B} \wedge \boldsymbol{C}$, is called a **trivector**, and so on to the general case of a p-vector, a completely antisymmetric tensor of type $(p, 0)$.
- The result of wedging two 1-forms, $\boldsymbol{f} \wedge \boldsymbol{g}$ is called a 2-form, and so on to the general case of a p-form, a completely antisymmetric tensor of type $(0, p)$.
- The wedge between a vector and a 1-form is not defined. Only wedges inside spaces $T_{p,0}$ exist (p-vectors), and similarly for wedges inside spaces $T_{0,p}$ (p-forms).
- The above formula for the wedge of two vectors **cannot be used recursively!** Namely, we could try to calculate $\boldsymbol{A} \wedge \boldsymbol{B} \wedge \boldsymbol{C}$ as:

$$\begin{aligned} \boldsymbol{A} \wedge \boldsymbol{B} \wedge \boldsymbol{C} &= \boldsymbol{A} \wedge (\boldsymbol{B} \otimes \boldsymbol{C} - \boldsymbol{C} \otimes \boldsymbol{B}) \\ &= \boldsymbol{A} \otimes (\boldsymbol{B} \otimes \boldsymbol{C} - \boldsymbol{C} \otimes \boldsymbol{B}) - (\boldsymbol{B} \otimes \boldsymbol{C} - \boldsymbol{C} \otimes \boldsymbol{B}) \otimes \boldsymbol{A} \\ &= \boldsymbol{A} \otimes \boldsymbol{B} \otimes \boldsymbol{C} - \boldsymbol{A} \otimes \boldsymbol{C} \otimes \boldsymbol{B} - \boldsymbol{B} \otimes \boldsymbol{C} \otimes \boldsymbol{A} + \boldsymbol{C} \otimes \boldsymbol{B} \otimes \boldsymbol{A}, \end{aligned}$$

  which is *wrong*, because the result is not totally antisymmetric! Last two terms have the wrong sign, and two more terms are missing! The correct result is

$$\boldsymbol{A} \wedge \boldsymbol{B} \wedge \boldsymbol{C} = \boldsymbol{A} \otimes \boldsymbol{B} \otimes \boldsymbol{C} - \boldsymbol{A} \otimes \boldsymbol{C} \otimes \boldsymbol{B} + \boldsymbol{B} \otimes \boldsymbol{C} \otimes \boldsymbol{A} - \boldsymbol{C} \otimes \boldsymbol{B} \otimes \boldsymbol{A} - \boldsymbol{B} \otimes \boldsymbol{A} \otimes \boldsymbol{C} + \boldsymbol{C} \otimes \boldsymbol{A} \otimes \boldsymbol{B},$$

  and is always constructed in such a way that the interchange of any two vectors changes the overall sign of the product.
- If $\boldsymbol{A}$ and $\boldsymbol{B}$ are a p-vector and a q-vector respectively (or a p-form and a q-form), then the following commutation rule is valid:

$$\boldsymbol{A} \wedge \boldsymbol{B} = (-1)^{pq} \boldsymbol{B} \wedge \boldsymbol{A}.$$

- The exterior product is the basic ingredient in the so-called **exterior calculus** or **algebra of differential forms**, which will be introduced in the next section.

1.6 Differential forms, exterior calculus

Start from the tensor algebra $T(\mathcal{P})$, and consider its quotient $T/\boldsymbol{A} \otimes \boldsymbol{A}$ for $\boldsymbol{A} \in T\mathcal{M}$, ie. the subset of all completely antisymmetric tensors. Furthermore, consider only the tensors of type $(p, 0)$. This quotient is a *subalgebra* of the tensor algebra, and is called **exterior algebra** $\Lambda(\mathcal{P})$. Its typical element, say $\boldsymbol{A} \in T_{3,0}$ can be written in the usual form

$$\boldsymbol{A} = A^{\mu\nu\lambda} \mathbf{e}_\mu \otimes \mathbf{e}_\nu \otimes \mathbf{e}_\lambda.$$

However, since it is completely antisymmetric, it can also be rewritten in the form

$$\boldsymbol{A} = \frac{1}{3!} A^{\mu\nu\lambda} \mathbf{e}_\mu \wedge \mathbf{e}_\nu \wedge \mathbf{e}_\lambda$$

(the factor $1/3!$ appears due to the overcounting of components in the sum). In general, any p-vector from Λ can be written in the form

$$\boldsymbol{A} = \frac{1}{p!} A^{\mu_1 \cdots \mu_p} \mathbf{e}_{\mu_1} \wedge \cdots \wedge \mathbf{e}_{\mu_p}.$$

Thus, we see that $\Lambda(\mathcal{P})$ is an algebra with operations of addition and scalar multiplication inherited from $T(\mathcal{P})$, while the product of elements is the exterior product, $\wedge$.

Completely analogously, instead using tensors of type $(p, 0)$, one can use tensors of type $(0, p)$, and construct the **algebra of differential forms**, $\Lambda^*(\mathcal{P})$, whose elements are

$$\boldsymbol{f} = \frac{1}{p!} f_{\mu_1 \dots \mu_p} \mathbf{e}^{\mu_1} \wedge \cdots \wedge \mathbf{e}^{\mu_p}.$$

Note:

- Both algebras Λ and Λ^* are finite-dimensional. If the tangent space $T\mathcal{M}$ has dimension D, there are at most D linearly independent basis vectors $\mathbf{e}_\mu$ which can be used to construct a basis element in the exterior algebra,

  $$\mathbf{e}_0 \wedge \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_{D-1}.$$

  If we try to wedge another vector, the result will be zero, since the antisymmetry of $\wedge$ implies $\mathbf{e}_\mu \wedge \mathbf{e}_\mu = 0$. The total dimension of Λ is 2^D. Similarly for Λ^*.

- Due to the natural factorization of tensor algebra T, exterior algebra Λ inherits this structure:

  $$\begin{array}{rcll} \Lambda & = & \Lambda^0 \oplus \Lambda^1 \oplus \cdots \oplus \Lambda^D & \\ & = & \mathbb{R} & \text{scalars} \\ & \oplus & T\mathcal{M} & \text{vectors} \\ & \oplus & T\mathcal{M} \wedge T\mathcal{M} & \text{2-vectors} \\ & \oplus & \dots & \\ & \oplus & \underbrace{T\mathcal{M} \wedge \cdots \wedge T\mathcal{M}}_{D \text{ times}} & D\text{-vectors.} \end{array}$$

  Analogous factorization can be written for Λ^*, with slightly different terminology:

  $$\begin{array}{rcll} \Lambda^* & = & \Lambda^{*0} \oplus \Lambda^{*1} \oplus \cdots \oplus \Lambda^{*D} & \\ & = & \mathbb{R} & \text{scalars or 0-forms} \\ & \oplus & T\mathcal{M}^* & \text{1-forms} \\ & \oplus & T\mathcal{M}^* \wedge T\mathcal{M}^* & \text{2-forms} \\ & \oplus & \dots & \\ & \oplus & \underbrace{T\mathcal{M}^* \wedge \cdots \wedge T\mathcal{M}^*}_{D \text{ times}} & D\text{-forms.} \end{array}$$

- If a p-form $\boldsymbol{f}$ can be written as an exterior product of p 1-forms $\boldsymbol{g}_1, \dots, \boldsymbol{g}_p$,

  $$\boldsymbol{f} = \boldsymbol{g}_1 \wedge \cdots \wedge \boldsymbol{g}_p,$$

  it is called a **simple** p-form. Otherwise it can be written only as a linear combination of such products, and is not simple. Similar terminology is used also for p-vectors.

- Exterior algebra is also known by the name **Grassmann algebra**, and its elements called **Grassmann numbers** or **anticommutative numbers** (as opposed to real numbers), due to the anticommutativity of exterior product.

- If we consider the case $D = 2$, and choose two vectors $\boldsymbol{A} = A^\mu \mathbf{e}_\mu$ and $\boldsymbol{B} = B^\mu \mathbf{e}_\mu$, their exterior product can be calculated explicitly as

  $$\begin{array}{rcl} \boldsymbol{A} \wedge \boldsymbol{B} & = & (A^0 \mathbf{e}_0 + A^1 \mathbf{e}_1) \wedge (B^0 \mathbf{e}_0 + B^1 \mathbf{e}_1) \\ & = & A^0 B^1 \mathbf{e}_0 \wedge \mathbf{e}_1 + A^1 B^0 \mathbf{e}_1 \wedge \mathbf{e}_0 \\ & = & (A^0 B^1 - A^1 B^0) \mathbf{e}_0 \wedge \mathbf{e}_1 \\ & = & \det \left[\begin{array}{cc} A^0 & A^1 \\ B^0 & B^1 \end{array} \right] \mathbf{e}_0 \wedge \mathbf{e}_1 \end{array}$$

Thus, wedge product and exterior algebra represent the generalization of the concepts of **determinant** (D-forms) and **minors** (p-forms), and provide natural formalism for their systematic description.

- Exterior algebra was first introduced and explored by Hermann Grassmann in his work on geometry called "Theory of Extension" (from 1844). Hence the name "exterior".

1.7 Transformations of basis, principle of relativity

So far we have discussed tensor algebra using one particular choice of the basis vectors $\mathbf{e}_\mu$ in $T\mathcal{M}_\mathcal{P}$. This is however far from unique, and now we shall consider what happens when we switch from one basis to another.

Start from the basis $\mathbf{e}_\mu$, and construct a new basis, $\mathbf{e}_{\mu'}$, as a linear combination of the old one:

$$\mathbf{e}_{\mu'} = M^\mu{}_{\mu'}\mathbf{e}_\mu.$$

Note:

- Since the new basis vectors must be linearly independent, the **transformation matrix** $M \equiv [M^\mu{}_{\mu'}]$ must be nonsingular, $\det M \neq 0$. Other than that, it is **completely arbitrary**.
- The "prime" is used to denote the new basis. However, it is more instructive to put a prime on the **index** rather than on the symbol **e**. Such notation might seem unusual, but it has several advantages over the more common one, as we shall see below.
- Indices μ and μ' in $M^\mu{}_{\mu'}$ are to be understood as completely independent of each other, despite the same symbol μ. Due to the prime, there can be no confusion.
- Since $\det M \neq 0$, one can introduce the **inverse transformation matrix**, denoted as $M^{-1} = [M^{\mu'}{}_\mu]$. It cannot be confused with $M^\mu{}_{\mu'}$ due to the different position of the prime. Since the two matrices are inverse to each other, following identities hold:

  $$M^\mu{}_{\mu'}M^{\mu'}{}_\nu = \delta^\mu_\nu, \qquad M^{\mu'}{}_\mu M^\mu{}_{\nu'} = \delta^{\mu'}_{\nu'},$$

  where δ^μ_ν is the familiar Kronecker delta symbol.

Given a vector $\boldsymbol{A} \in T\mathcal{M}$, one can expand it as a linear combination in both bases:

$$\boldsymbol{A} = A^\mu\mathbf{e}_\mu = A^{\mu'}\mathbf{e}_{\mu'}.$$

Using the transformation rule between two bases, one can calculate the transformation rule between components A^μ and $A^{\mu'}$:

$$A^{\mu'} = M^{\mu'}{}_\nu A^\nu.$$

Components of a vector transform using the inverse transformation matrix. It is very important to stress that $\boldsymbol{A}$ is a **geometric object**, an **arrow** (tangent to some curve in the manifold $\mathcal{M}$), which is **independent** of any choice of basis. The components of $\boldsymbol{A}$ transform precisely in such a way to cancel the transformation of the basis vectors, so that $\boldsymbol{A}$ **does not transform at all**:

$$\boldsymbol{A}' = A^{\mu'}\mathbf{e}_{\mu'} = (M^{\mu'}{}_\nu A^\nu)(M^\lambda{}_{\mu'}\mathbf{e}_\lambda) = \underbrace{M^{\mu'}{}_\nu M^\lambda{}_{\mu'}}_{\delta^\lambda_\nu} A^\nu\mathbf{e}_\lambda = A^\lambda\mathbf{e}_\lambda = \boldsymbol{A}.$$

We see that **only quantities that carry an index actually change**, so we drop the prime from $\boldsymbol{A}'$, and keep primes only on indices.

Now consider the biorthogonal basis, $\mathbf{e}^{\mu}$. Starting from bases $\mathbf{e}_{\mu}$ and $\mathbf{e}_{\mu'}$ in $T\mathcal{M}$, one can construct biorthogonal bases $\mathbf{e}^{\mu}$ and $\mathbf{e}^{\mu'}$ in $T\mathcal{M}^*$, and ask what is the relation between them. The answer is easy to obtain from the biorthogonality relation:

$$\mathbf{e}^{\mu'} = M^{\mu'}{}_{\mu}\mathbf{e}^{\mu}.$$

Basis 1-forms transform using the inverse transformation matrix. Consequently, given that any 1-form $\boldsymbol{f} \in T\mathcal{M}^*$ is also a geometric object independent of any choice of basis, one can easily deduce the transformation rule for the components of 1-forms:

$$f_{\mu'} = M^{\mu}{}_{\mu'} f_{\mu}.$$

Components of 1-forms transform using the original transformation matrix.

Now introduce some terminology:

- Any quantity which transforms the same way as a basis in $T\mathcal{M}$ is said to be transformed **covariantly** (ie. "same as" the basis). Since the basis is transformed using the matrix $M^{\mu}{}_{\mu'}$, from the position of its indices it is easy to see that all quantities that carry a subscript ("down") index transform covariantly.
- Any quantity which transforms in the opposite way from the basis in $T\mathcal{M}$ is said to be transformed **contravariantly** (ie. "opposite of" the basis). Since in this case the inverse transformation matrix $M^{\mu'}{}_{\mu}$ is used, from the position of its indices it is easy to see that all quantities that carry a superscript ("up") index transform contravariantly.
- The covariance/contravariance is determined solely by the position of the index, and is independent of the nature of the object transformed (components, bases, etc.)

Finally, we can consider the general case of any tensor. Every tensor is (like a vector and a 1-form) a **geometric object**, independent of any basis chosen, so it **does not transform** when the basis is being changed. The components of a tensor are always transformed in precisely such a way as to cancel the transformation of the basis, so that the total tensor remains unchanged. As an example, consider the tensor $\boldsymbol{A} \in T_{2,1}$. It can be written as a linear combination in both bases:

$$\boldsymbol{A} = A^{\mu\nu}{}_{\lambda}\mathbf{e}_{\mu} \otimes \mathbf{e}_{\nu} \otimes \mathbf{e}^{\lambda} = A^{\mu'\nu'}{}_{\lambda'}\mathbf{e}_{\mu'} \otimes \mathbf{e}_{\nu'} \otimes \mathbf{e}^{\lambda'}.$$

Knowing the transformation rules for both basis vectors and basis 1-forms, we can deduce that components of $\boldsymbol{A}$ transform in accord with the appropriate positioning of the indices:

$$A^{\mu'\nu'}{}_{\lambda'} = A^{\mu\nu}{}_{\lambda} M^{\mu'}{}_{\mu} M^{\nu'}{}_{\nu} M^{\lambda}{}_{\lambda'}.$$

Thus, components $A^{\mu\nu}{}_{\lambda}$ transform twice contravariantly and once covariantly, the basis $\mathbf{e}_{\mu} \otimes \mathbf{e}_{\nu} \otimes \mathbf{e}^{\lambda}$ transforms twice covariantly and once contravariantly, and **$\boldsymbol{A}$ does not transform at all**, since the former two cancel each other out.

A general tensor of type (p,q) has components with p indices "up" and q indices "down", and is thus usually said to be "p times contravariant and q times covariant". But it is extremely important to stress that **only the components of a tensor have this property**, while the **basis of the tensor has the opposite property**, in such a way that the **tensor itself is invariant** with respect to the change of basis.

We stress that the invariance property of tensors is so important because it offers itself as a perfect tool for the embodiment of the **principle of general relativity**:

The **principle of general relativity** states:

All laws of physics must be expressed in such a way as to not depend on the choice of any particular coordinate system.

If we want to express some physical theory in such a way as to fulfill this axiom, it is necessary and sufficient to **express all the laws as tensor equations**, since these are automatically independent of the choice of the basis used. Note:

- This is the "why" for the whole story of tensor calculus in theory of general relativity.
- The above principle is the reason why the theory of general relativity is named the way it is named.

Next we turn attention to one technical issue — contraction. We have defined contraction of tensors as action of one basis 1-form onto one basis vector. Now we shall prove that this procedure is in fact independent on the choice of the basis itself. To this end, consider an example, $\boldsymbol{A} \in T_{2,1}$, and its contraction $C_{2,3}$ in some basis:

$$C_{2,3}(\boldsymbol{A}) = A^{\mu\nu}{}_{\nu}\mathbf{e}_{\mu}.$$

We can perform the same procedure in the primed basis, to obtain

$$C'_{2,3}(\boldsymbol{A}) = A^{\mu'\nu'}{}_{\nu'}\mathbf{e}_{\mu'}.$$

Now transform this back to the old basis, noting that $A^{\mu'\nu'}{}_{\nu'}$ has only one index (the other two are summed over):

$$C'_{2,3}(\boldsymbol{A}) = A^{\mu'\nu'}{}_{\nu'}\mathbf{e}_{\mu'} = (A^{\mu\nu'}{}_{\nu'}M^{\mu'}{}_{\mu})(M^{\lambda}{}_{\mu'}\mathbf{e}_{\lambda}) = A^{\lambda\nu'}{}_{\nu'}\mathbf{e}_{\lambda} = A^{\lambda\nu}{}_{\rho}M^{\nu'}{}_{\nu}M^{\rho}{}_{\nu'}\mathbf{e}_{\lambda} = A^{\lambda\rho}{}_{\rho}\mathbf{e}_{\lambda} = C_{2,3}(\boldsymbol{A}).$$

As we can see, the contraction is invariant with the change of basis. This is due to the fact that when contracting, we always sum over one contravariant and one covariant index in components of the tensor being contracted. While the components themselves change with respect to all indices, the changes of summed indices cancel each other, since they transform in the opposite way.

Finally, there is one more very important topic to be addressed. Namely, the original basis $\mathbf{e}_{\mu}$ was introduced as a set of **tangent vectors of the coordinate curves** at point $\mathcal{P}$, ie. the set of differential operators,

$$\mathbf{e}_{\mu} \equiv \frac{\partial}{\partial x^{\mu}},$$

evaluated at $\mathcal{P}$. After a change to a new basis $\mathbf{e}_{\mu'}$ using the arbitrary nonsingular transformation matrix M, one can ask is it possible to construct a new set of coordinate curves, such that the new basis can be represented as a set of differential operators acting on these new coordinate curves. The general answer to this question is **no, this is not always possible**. In order to see this, consider a set of alternative coordinate curves which pass through $\mathcal{P}$ and are parametrized with $x^{\mu'}$. Then construct the new basis relative to them in the same way as the old one was constructed:

$$\mathbf{e}_{\mu'} \equiv \frac{\partial}{\partial x^{\mu'}},$$

evaluated at $\mathcal{P}$. In order to deduce the explicit form of the transformation matrix between the new and old basis, one can employ the chain rule for differentials familiar from ordinary calculus:

$$\frac{\partial}{\partial x^{\mu'}} = \frac{\partial x^{\mu}}{\partial x^{\mu'}}\frac{\partial}{\partial x^{\mu}},$$

and just rewrite it in tensor notation:

$$\mathbf{e}_{\mu'} = M^{\mu}{}_{\mu'}\mathbf{e}_{\mu}, \qquad \text{where} \qquad M^{\mu}{}_{\mu'} \equiv \frac{\partial x^{\mu}}{\partial x^{\mu'}}.$$

However, knowing that the partial derivatives commute, the following identity for M holds:

$$\frac{\partial}{\partial x^{\lambda'}} M^{\mu}{}_{\mu'} = \frac{\partial}{\partial x^{\mu'}} M^{\mu}{}_{\lambda'}.$$

Now remember that in general case **matrix M is arbitrary**, and thus **may not satisfy the above identity!** This means that the **new basis may not be constructible from any set of coordinate curves**.

Consequently, starting from some set of coordinate curves, we may construct a basis (called **coordinate basis**), and then use some arbitrary matrix M to switch to a new basis. Depending on whether M was chosen in such a way as to satisfy the above identity, the new basis is called **coordinate (holonomic) basis** or **noncoordinate (anholonomic) basis**. Therefore, there are two distinct classes of basis vectors, and later we shall give a convenient method to determine whether a given basis is coordinate or noncoordinate one.

In older literature on tensor calculus, one can find a definition saying that a tensor (of type (p, q)) is a **set of quantities** $A^{\mu_1\cdots\mu_p}{}_{\nu_1\ldots\nu_q}$ that satisfies the following **transformation rule**:

$$A^{\mu'_1\cdots\mu'_p}{}_{\nu'_1\ldots\nu'_q} = A^{\mu_1\cdots\mu_p}{}_{\nu_1\ldots\nu_q} \frac{\partial x^{\mu'_1}}{\partial x^{\mu_1}} \cdots \frac{\partial x^{\mu'_p}}{\partial x^{\mu_p}} \frac{\partial x^{\nu_1}}{\partial x^{\nu'_1}} \cdots \frac{\partial x^{\nu_1}}{\partial x^{\nu'_1}}.$$

This "component" kind of approach to tensors is flawed on several grounds:

- A tensor is not just a set of components, but rather a linear combination of these components with some set of basis vectors. If basis is ignored, the geometric nature of tensors and their essential invariance with respect to the choice of this basis is not obvious, which makes the principle of relativity hard to understand.
- Due to the possible arbitrary transformation of basis, a perfectly valid tensor can be expressed via perfectly valid components which do not satisfy the above definition.
- It is **not** sufficient to consider just coordinate bases (as is usually done is some books), because (a) one misses a whole lot of powerful geometry and insight from the formalism and (b) the general transformation matrices $M^{\mu'}{}_{\mu}$ and $M^{\mu}{}_{\mu'}$ are **absent from the theory**, which is a **major handicap**.

The final point is actually the most severe one. In order to appreciate and understand why, let us just say that matrices $M^{\mu'}{}_{\mu}$ and $M^{\mu}{}_{\mu'}$ **have physical interpretation** of **gravitational field potentials**. In four dimensions, they are usually called **tetrads**, and are denoted $\mathbf{e}^{\mu'}{}_{\mu}$ and $\mathbf{e}^{\mu}{}_{\mu'}$. They play a very fundamental role in theory of general relativity, as we shall see later on. For example, if we omit them from the formalism, we have **no way to incorporate fermion fields** and couple them to gravity. We shall revisit tetrads in third chapter.

Contents of "Tensor Calculus Part 2" (in preparation):

For further reading

[1] Milutin Blagojević, “Gravitation and Gauge Symmetries”, Institute of Physics Publishing, London (2002), ISBN 0-7503-0767-6

[2] Charles W. Misner, Kip S. Thorne, John Archibald Wheeler, “Gravitation”, W. H. Freeman and Co. (1973), ISBN 978-0-7167-0344-0.