

# Tensor Analysis

Emre Özer  
Autumn 2019

Notes based on the book *“Introduction to Tensor Analysis and the Calculus of Moving Surfaces”* by PAVEL GRINFELD.

## Contents

<b>1</b>	<b>Basics</b>	<b>2</b>
1.1	Position vector as function of coordinates . . . . .	2
1.2	Covariant basis . . . . .	3
1.3	Covariant metric tensor . . . . .	3
1.4	Contravariant metric tensor . . . . .	3
1.5	Contravariant basis . . . . .	4
1.6	Line element . . . . .	4
1.7	The Christoffel symbol . . . . .	5
<b>2</b>	<b>The Tensor Property</b>	<b>7</b>
2.1	Variants and coordinate transformations . . . . .	7
2.2	Definitions and examples of tensors . . . . .	7
2.3	Properties of tensors . . . . .	9
2.4	Index juggling . . . . .	10
<b>3</b>	<b>Differentiation</b>	<b>11</b>
3.1	Gradient of a scalar field . . . . .	11
3.2	Covariant derivative . . . . .	11
3.2.1	Motivation . . . . .	11
3.2.2	Definitions of derivatives, divergence and Laplacian . . . . .	12
3.2.3	Properties of the covariant derivative . . . . .	13

# 1 Basics

Let's start with a description of objects in Euclidean spaces in tensor terms.

**Definition** (Vector). A *vector* is a directed line segment - a *length with a direction*.

**Definition** (Dot product). We define the dot product *geometrically*, meaning independent of any coordinates. Given two vectors  $\mathbf{u}$  and  $\mathbf{v}$ , their dot product is

$$\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}| |\mathbf{v}| \cos \alpha, \quad (1.1)$$

where  $|\mathbf{u}|$  and  $|\mathbf{v}|$  are the *lengths* of the vectors and  $\alpha$  is the *angle* between them. Note that although this definition looks like it assumes the lengths of the vectors and angles between them are defined, only the length is the *primary* concept. The dot product can be defined in terms of lengths alone, to see this, consider

$$\begin{aligned} (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) &= |\mathbf{u} + \mathbf{v}|^2 = u^2 + v^2 + 2(\mathbf{u} \cdot \mathbf{v}), \\ \implies \mathbf{u} \cdot \mathbf{v} &= \frac{1}{2} (|\mathbf{u} + \mathbf{v}|^2 - u^2 - v^2), \end{aligned}$$

where  $u^2$  and  $v^2$  are the squared lengths of  $\mathbf{u}$  and  $\mathbf{v}$ . We have assumed commutativity and distributivity of the dot product which can be easily proved from definition (1.1). The expression above doesn't involve angles, so we may define the angle from the dot product.

**Definition** (Variant). Objects whose values *change* depending on the choice of coordinates.

**Example** (Variants). The *components* of vectors such as velocity and force, the basis vectors, the metric tensor...

**Definition** (Invariant). Objects that do not depend on the choice of coordinates.

**Example** (Invariants). Physical quantities such as force and displacement vectors, scalars, physical fields...

We would hope that any physical object is an invariant since the universe doesn't have a preferred coordinate system. Coordinates are just labels, invariants objects do not need labels to exist. This is a very powerful statement, as it allows us to *find physical quantities* just by looking at the invariants. An example of this is the relativistic action, constructed from the invariant proper time.

**Definition** (Position vector). The *position vector* is a vector  $\mathbf{R}$  that represents the positions of points in Euclidean space with respect to an arbitrary origin  $O$ .

Note that the position vector is introduced with no reference to a coordinate system. So, it is *invariant* (or *geometric*).

## 1.1 Position vector as function of coordinates

Let  $x^i$  denote a set of coordinates. Now, since the coordinates assign a *unique label* to each point, the position vector now becomes a function of the coordinates  $x^i$ . Denote this functions as

$$\mathbf{R} = \mathbf{R}(x), \quad (1.2)$$

where  $x$  stands for the coordinates  $x^i$ . This equation features the symbol  $\mathbf{R}$  in two different roles. On the left-hand side,  $\mathbf{R}$  represents the geometric position vector. On the right-hand side,  $\mathbf{R}$  stands for the vector-valued function that yields the position vector for every valid combination

of coordinates. As it is often the case, the same letter is used to represent two different objects.  $\mathbf{R}$  is the invariant position vector while  $\mathbf{R}(x)$  is a vector-valued function of three variables.

It is essential to treat  $\mathbf{R}$  as a primary object that is subject to its own set of rules and operations. It is counterproductive to think of  $\mathbf{R}$  as a triplet of components with respect to a some Cartesian grid. There is no basis with respect to which  $\mathbf{R}$  can be decomposed. In fact, we will define the basis from  $\mathbf{R}(x)$ .

## 1.2 Covariant basis

**Definition** (Covariant Basis). The *covariant basis* is obtained from differentiating the position vector  $\mathbf{R}(x)$  with respect to each of the coordinates:

$$\mathbf{e}_i = \frac{\partial \mathbf{R}(x)}{\partial x^i}. \quad (1.3)$$

Note that the basis vectors may vary in space. Such basis are called *local*. At all points in the Euclidean space, the covariant basis  $\mathbf{e}_i$  provides a convenient basis for decomposing vectors. The *contravariant components*  $v^i$  of a vector  $\mathbf{v}$  are the scalar values that produce  $\mathbf{v}$  when used in a linear combination with the vectors  $\mathbf{e}_i$ :

$$\mathbf{v} = v^i \mathbf{e}_i. \quad (1.4)$$

## 1.3 Covariant metric tensor

**Definition** (Covariant metric tensor). The *covariant metric tensor* is defined as the pairwise dot products of the covariant basis vectors,

$$g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j. \quad (1.5)$$

**Corollary.** Since the dot product is commutative, the metric tensor is symmetric:

$$\mathbf{e}_i \cdot \mathbf{e}_j = \mathbf{e}_j \cdot \mathbf{e}_i \implies g_{ij} = g_{ji}.$$

**Proposition** (Dot product). Suppose two vectors  $\mathbf{v}$  and  $\mathbf{u}$  are located at the same point, with components  $v^i$  and  $u^i$ . Then their dot product  $\mathbf{u} \cdot \mathbf{v}$  is given by

$$\mathbf{u} \cdot \mathbf{v} = u^i \mathbf{e}_i \cdot v^j \mathbf{e}_j = u^i v^j (\mathbf{e}_i \cdot \mathbf{e}_j) = u^i v^j g_{ij}. \quad (1.6)$$

**Corollary.** The length of a vector in terms of the metric is given by

$$|\mathbf{v}| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{g_{ij} v^i v^j}.$$

**Proposition.** The covariant metric interpreted as a matrix is positive-definite.

*Proof.* Let  $\mathbf{u}$  be a non-zero vector, and let  $g$  denote the metric tensor as a matrix. Then,

$$\mathbf{u}^T g \mathbf{u} = u^i g_{ij} u^j = \mathbf{u} \cdot \mathbf{u} = u^2 > 0. \quad \square$$

## 1.4 Contravariant metric tensor

**Definition** (Contravariant metric tensor). The *contravariant metric tensor*  $g^{ij}$  is defined as the matrix inverse of the covariant metric tensor  $g_{ij}$ :

$$g^{ik} g_{kj} = \delta_k^i \iff g^{ik} g_{jk} = \delta_k^i. \quad (1.7)$$

**Corollary.** The contravariant metric tensor is symmetric,  $g^{ij} = g^{ji}$ .

*Proof.* The inverse of any symmetric matrix is symmetric. Let  $A$  and  $B$  be inverses of each other and let  $A_{ij} = A_{ji}$ . We have  $AB = I$  so,

$$A_{ij} B^{jk} = A_{ji} B^{jk} = B^{jk} A_{ji} = \delta_i^k \implies B^{jk} = B^{kj}$$

since we know  $BA = I \iff B^{kj} A_{ji} = \delta_i^k$ .  $\square$

## 1.5 Contravariant basis

**Definition** (Contravariant basis). The *contravariant basis*  $\mathbf{e}^i$  is defined with respect to the covariant basis by the set of relations

$$\mathbf{e}^i \cdot \mathbf{e}_j = \delta_j^i, \quad (1.8)$$

so that the covariant and contravariant basis are *mutually orthonormal*.

**Corollary.** The contravariant basis can be explicitly written in terms of the contravariant metric and the covariant basis as follows:

$$\delta_j^i = g^{ik} g_{kj} = g^{ik} \mathbf{e}_k \cdot \mathbf{e}_j = \mathbf{e}^i \cdot \mathbf{e}_j \implies \mathbf{e}^i = g^{ik} \mathbf{e}_k. \quad (1.9)$$

**Corollary.** The pairwise dot product of the contravariant basis vectors yield the contravariant metric tensor,

$$\mathbf{e}^i \cdot \mathbf{e}^j = g^{ik} \mathbf{e}_k \cdot g^{jm} \mathbf{e}_m = g^{ik} g^{jm} g_{km} = g^{ij}.$$

**Corollary** (Covariant vector components). The definition for the covariant vector components follows from the contravariant basis. We have

$$\mathbf{v} = v^i \mathbf{e}_i = v^i g_{ij} \mathbf{e}^j = v_j \mathbf{e}^j$$

where we naturally defined  $v_j = g_{ij} v^i$ .

**Corollary** (Vector component decomposition). It's straightforward to show that

$$v^i = \mathbf{v} \cdot \mathbf{e}^i \quad \text{and} \quad v_i = \mathbf{v} \cdot \mathbf{e}_i.$$

## 1.6 Line element

We may express the line element  $ds^2$  in terms of the metric  $g_{ij}$ . If the coordinates  $x^i$  parameterize a surface, the line element is also called the *first fundamental form*. Consider a small displacement  $\delta \mathbf{R}$ . By the chain rule, we have

$$\delta \mathbf{R} = \frac{\partial \mathbf{R}(x)}{\partial x^i} \delta x^i + o(\delta x).$$

The length of the small displacement is given by

$$\delta s^2 = \delta \mathbf{R} \cdot \delta \mathbf{R} = \delta x^i \delta x^j \frac{\partial \mathbf{R}}{\partial x^i} \cdot \frac{\partial \mathbf{R}}{\partial x^j} + o(\delta x^2) = \delta x^i \delta x^j g_{ij} + o(\delta x^2).$$

Taking the limit as  $\delta x^i \rightarrow dx^i$  yields the *line element*:

$$ds^2 = dx^i dx^j g_{ij}. \quad (1.10)$$

The arc length of a curve parameterized by some  $x^i = x^i(t)$  is then given by

$$L = \int \sqrt{g_{ij} \frac{dx^i}{dt} \frac{dx^j}{dt}} dt. \quad (1.11)$$

To see why this is, simply reparameterize  $\delta \mathbf{R}$  in terms of  $t$ , so

$$\delta \mathbf{R} = \frac{\partial \mathbf{R}(t(x))}{\partial x^i} \frac{dx^i(t)}{dt} \delta t + o(\delta t) \implies ds^2 = \frac{dx^i}{dt} \frac{dx^j}{dt} g_{ij}.$$

## 1.7 The Christoffel symbol

In curvilinear coordinate systems, the basis varies from one point to another. The variation can be described by the partial derivatives  $\partial \mathbf{e}_i / \partial x^j$ . The expression  $\partial \mathbf{e}_i / \partial x^j$  represents  $N^2$  vectors: each of the basis elements differentiated with respect to each of the coordinates. Each of the  $N^2$  vectors  $\partial \mathbf{e}_i / \partial x^j$  can be decomposed with respect to the covariant basis  $\mathbf{e}_k$ . The resulting  $N^3$  coefficients form the *Christoffel symbol*  $\Gamma_{ij}^k$ :

$$\Gamma_{ij}^k \mathbf{e}_k = \frac{\partial \mathbf{e}_i}{\partial x^j} \iff \Gamma_{ij}^k = \mathbf{e}^k \cdot \frac{\partial \mathbf{e}_i}{\partial x^j}. \quad (1.12)$$

The Christoffel symbol provides all information about how the basis vectors change depending on position. This will become important when we define gradients later on.

**Proposition.** The Christoffel symbol is symmetric in its lower indices, so  $\Gamma_{ij}^k = \Gamma_{ji}^k$ .

*Proof.* By definition, we write  $\mathbf{e}^i = \partial \mathbf{R} / \partial x^i$ , hence

$$\frac{\partial \mathbf{e}_i}{\partial x^j} = \frac{\partial^2 \mathbf{R}}{\partial x^j \partial x^i} = \frac{\partial^2 \mathbf{R}}{\partial x^i \partial x^j} = \frac{\partial \mathbf{e}_j}{\partial x^i}.$$

The result follows.  $\square$

The Christoffel symbol also appears in the decomposition of the partial derivatives  $\partial \mathbf{e}^i / \partial x^j$  of the *contravariant* basis  $\mathbf{e}^i$ . Transform equation (1.12) by the product rule,

$$\Gamma_{ij}^k = \mathbf{e}^k \cdot \frac{\partial \mathbf{e}_i}{\partial x^j} = \frac{\partial (\mathbf{e}^k \cdot \mathbf{e}_i)}{\partial x^j} - \mathbf{e}_i \cdot \frac{\partial \mathbf{e}^k}{\partial x^j} = -\mathbf{e}_i \cdot \frac{\partial \mathbf{e}^k}{\partial x^j}.$$

where the first term vanishes since  $\mathbf{e}^k \cdot \mathbf{e}_i = \delta_i^k$  is constant. Hence, we arrive at an alternative definition for the Christoffel symbol:

$$\Gamma_{ij}^k \mathbf{e}^i = -\frac{\partial \mathbf{e}^k}{\partial x^j} \iff \Gamma_{ij}^k = -\mathbf{e}_i \cdot \frac{\partial \mathbf{e}^k}{\partial x^j}. \quad (1.13)$$

**Proposition.** The Christoffel symbol can be written in terms of the metric as:

$$\Gamma_{ij}^k = \frac{1}{2} g^{km} \left( \frac{\partial g_{mi}}{\partial x^j} + \frac{\partial g_{mj}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^m} \right).$$

*Proof.* We simply compute the right hand side. The first term is

$$\frac{\partial g_{mi}}{\partial x^j} = \frac{\partial}{\partial x^j} (\mathbf{e}_m \cdot \mathbf{e}_i) = \mathbf{e}_m \cdot \frac{\partial \mathbf{e}_i}{\partial x^j} + \mathbf{e}_i \cdot \frac{\partial \mathbf{e}_m}{\partial x^j} = g_{m\ell} \mathbf{e}^\ell \cdot \frac{\partial \mathbf{e}_i}{\partial x^j} + g_{i\ell} \mathbf{e}^\ell \cdot \frac{\partial \mathbf{e}_m}{\partial x^j}$$

Similarly, the other two terms are

$$\frac{\partial g_{mj}}{\partial x^i} = g_{m\ell} \mathbf{e}^\ell \cdot \frac{\partial \mathbf{e}_j}{\partial x^i} + g_{j\ell} \mathbf{e}^\ell \cdot \frac{\partial \mathbf{e}_m}{\partial x^i} \quad \text{and} \quad \frac{\partial g_{ij}}{\partial x^m} = g_{i\ell} \mathbf{e}^\ell \cdot \frac{\partial \mathbf{e}_j}{\partial x^m} + g_{j\ell} \mathbf{e}^\ell \cdot \frac{\partial \mathbf{e}_i}{\partial x^m}.$$

Substituting the Christoffel symbols, we have

$$\frac{\partial g_{mi}}{\partial x^j} + \frac{\partial g_{mj}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^m} = 2g_{\ell m} \Gamma_{ij}^\ell.$$

The result follows:

$$\frac{1}{2} g^{km} (2g_{\ell m} \Gamma_{ij}^\ell) = \delta_\ell^k \Gamma_{ij}^\ell = \Gamma_{ij}^k. \quad \square$$

**Example** (Velocity components). Consider a particle moving along a curve with parameterization with respect to time  $x^i \equiv x^i(t)$ . Then, the velocity components  $V^i$  are given by

$$V^i = \mathbf{e}^i \cdot \mathbf{V} = \mathbf{e}^i \cdot \frac{d\mathbf{R}}{dt} = \mathbf{e}^i \cdot \frac{\partial \mathbf{R}}{\partial x^j} \frac{dx^j}{dt} = \delta_j^i \frac{dx^j}{dt} = \frac{dx^i}{dt}.$$

**Example** (Acceleration components). The acceleration  $\mathbf{A}$  is more complicated:

$$\mathbf{A} = \frac{d\mathbf{V}}{dt} = \frac{d}{dt}(V^i \mathbf{e}_i) = \frac{dV^i}{dt} \mathbf{e}_i + V^i \frac{d\mathbf{e}_i}{dt} = \frac{dV^i}{dt} \mathbf{e}_i + V^i \frac{\partial \mathbf{e}_i}{\partial x^j} \frac{dx^j}{dt} = \frac{dV^i}{dt} \mathbf{e}_i + V^i V^j \frac{\partial \mathbf{e}_i}{\partial x^j}.$$

The components  $A^i$  are then given by

$$A^i = \mathbf{e}^i \cdot \mathbf{A} = \frac{dV^i}{dt} + \mathbf{e}^i \cdot \frac{\partial \mathbf{e}_k}{\partial x^j} V^j V^k = \frac{dV^i}{dt} + V^j V^k \Gamma_{jk}^i.$$

**Example** (Vector field along curve). For some vector field  $\mathbf{U}(t)$  defined along the curve  $x^i \equiv x^i(t)$ , the rate of change of  $\mathbf{U}$  is given by

$$\frac{d\mathbf{U}(t)}{dt} = \frac{d}{dt}(U^i \mathbf{e}_i) = \frac{dU^i}{dt} \mathbf{e}_i + U^i \frac{\partial \mathbf{e}_i}{\partial x^j} V^j = \left( \frac{dU^i}{dt} + U^k V^j \Gamma_{kj}^i \right) \mathbf{e}_i.$$

Hence, if  $\mathbf{U}(t)$  is constant along the curve, we have

$$\frac{dU^i}{dt} + U^k V^j \Gamma_{kj}^i = 0.$$

## 2 The Tensor Property

### 2.1 Variants and coordinate transformations

A variant is an object that can be constructed by a similar rule in various coordinate systems. For example, the covariant basis  $\mathbf{e}_i$  is a variant because it is obtained in any coordinate system by the same rule—partial differentiation of the position vector. The metric tensor  $g_{ij}$  is also a variant because it is formed by pairwise dot products of the elements of the covariant basis.

Naturally, when the same rule is applied in different coordinate systems, the results differ. Tensors are variants which transform in specific ways.

Now, we consider “nice” coordinate transformations. By nice, we mean the transformations are continuously differentiable. Consider two coordinate systems,  $x$  and  $x'$ , related to each other by

$$x^i = x^i(x'), \quad \text{and} \quad x^{i'} = x^{i'}(x),$$

where, as a notation we denote  $(x')^i$  as  $x^{i'}$ . We also note  $i$  and  $i'$  are different indices, so they are not summed over each other. An important identity related with such transformations is:

$$\frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^{i'}}{\partial x^j} = \frac{\partial x^i}{\partial x^j} = \delta^i_j = \frac{\partial x^{i'}}{\partial x^{j'}} = \frac{\partial x^{i'}}{\partial x^i} \frac{\partial x^i}{\partial x^{j'}}. \quad (2.1)$$

We adopt the following notation in terms of the Jacobians associated with the transformation  $x \rightarrow x'$ ,

$$J^{i'}_i \equiv \frac{\partial x^{i'}}{\partial x^i}, \quad \text{and} \quad J^i_{i'} \equiv \frac{\partial x^i}{\partial x^{i'}}. \quad (2.2)$$

We note that although we have used the same letter  $J$  to denote both objects above, they do not refer to the same object. Namely, the two are matrix inverses of each other due to the identity (2.1).

### 2.2 Definitions and examples of tensors

**Definition** (Covariant tensor). A variant  $T_i$  is called a *covariant tensor of order one* if its components in the coordinates  $x^{i'}$  and  $x^i$  are related by

$$T_{i'} = \frac{\partial x^i}{\partial x^{i'}} T_i = J^i_{i'} T_i \iff T_i = J^{i'}_i T_{i'}. \quad (2.3)$$

**Definition** (Contravariant tensor). A variant  $T^i$  is called a *contravariant tensor of order one* if its components in the coordinates  $x^{i'}$  and  $x^i$  are related by

$$T^{i'} = \frac{\partial x^{i'}}{\partial x^i} T^i = J^{i'}_i T^i \iff T^i = J^i_{i'} T^{i'}. \quad (2.4)$$

**Proposition** (Contraction invariance). Let  $S^i$  be a contravariant and  $T_i$  be a covariant tensor. Then, the contraction

$$U = S^i T_i$$

is *invariant*. To see why, we simply write the transformation rule for  $U$ ,

$$U' = S^{i'} T_{i'} = J^{i'}_j S^j J^k_{i'} T_k = \delta^k_j S^j T_k = S^j T_j = U. \quad \square \quad (2.5)$$

**Proposition** (Covariant basis). The covariant basis  $\mathbf{e}_i$  is a tensor of rank one. We express  $\mathbf{e}_{i'}$  in the  $x'$  coordinates and apply the chain rule to obtain the transformation rule,

$$\mathbf{e}_{i'} = \frac{\partial \mathbf{R}}{\partial x^{i'}} = \frac{\partial \mathbf{R}}{\partial x^i} \frac{\partial x^i}{\partial x^{i'}} = J^i_{i'} \mathbf{e}_i. \quad \square \quad (2.6)$$



**Proposition** (Contravariant vector components). If the covariant basis vectors transform as covariant tensors, the contravariant vector components must transform in the *opposite sense* in order to keep the vector itself invariant. More explicitly, for an arbitrary vector  $\mathbf{v}$ , we have

$$\mathbf{v} = v^{i'} \mathbf{e}_{i'} = v^{i'} J_{i'}^i \mathbf{e}_i = v^i \mathbf{e}_i \implies v^i = v^{i'} J_{i'}^i. \quad \square \quad (2.7)$$

Hence, the contravariant vector components transform as contravariant tensors.

**Definition** (Higher order tensors). A variant  $T^{ij}$  is a *doubly contravariant*,  $T_{ij}$  is a *doubly covariant* and  $T_j^i$  is a *mixed* tensor of order two if their components transform as

$$T^{i'j'} = J^{i'}_i J^{j'}_j T^{ij}, \quad T_{i'j'} = J^i_{i'} J^j_{j'} T_{ij}, \quad T^{i'}_{j'} = J^{i'}_i J^j_{j'} T^i_j. \quad (2.8)$$

Higher order generalizations are obvious.

**Proposition** (Kronecker delta). The Kronecker symbol  $\delta^i_j$  transforms as a *mixed tensor of order two*. To see this, we simply multiply (contract)  $\delta^i_j$  with  $J^{i'}_i J^j_{j'}$ ,

$$J^{i'}_i J^j_{j'} \delta^i_j = J^{i'}_i J^j_{j'} = \delta^{i'}_{j'}. \quad \square \quad (2.9)$$

**Proposition** (Contravariant basis). The contravariant basis  $\mathbf{e}^i$  transform as contravariant tensors. By definition, we have

$$\mathbf{e}^{i'} \cdot J^j_{j'} \mathbf{e}_j = \mathbf{e}^{i'} \cdot \mathbf{e}_{j'} = \delta^{i'}_{j'} = J^{i'}_i J^j_{j'} \delta^i_j = J^{i'}_i J^j_{j'} \mathbf{e}^i \cdot \mathbf{e}_j.$$

Comparing the first and the last term, we see

$$\mathbf{e}^{i'} = J^{i'}_i \mathbf{e}^i. \quad \square \quad (2.10)$$

**Proposition** (Covariant metric). The covariant metric is a *doubly covariant tensor*. The proof follows from (2.6):

$$g_{i'j'} = \mathbf{e}_{i'} \cdot \mathbf{e}_{j'} = J^i_{i'} J^j_{j'} \mathbf{e}_i \cdot \mathbf{e}_j = J^i_{i'} J^j_{j'} g_{ij}. \quad \square \quad (2.11)$$

**Proposition** (Contravariant metric). The contravariant metric is a *doubly contravariant tensor*. The proof follows from (2.10):

$$g^{i'j'} = \mathbf{e}^{i'} \cdot \mathbf{e}^{j'} = J^{i'}_i J^{j'}_j \mathbf{e}^i \cdot \mathbf{e}^j = J^{i'}_i J^{j'}_j g^{ij}. \quad \square \quad (2.12)$$

**Proposition** (Partial derivatives). Given a scalar field  $F(x)$ , the collection of partial derivatives  $\partial F / \partial x^i$  transform as covariant tensor components.

$$\frac{\partial F}{\partial x^{i'}} = \frac{\partial F}{\partial x^i} \frac{\partial x^i}{\partial x^{i'}} = J^i_{i'} \frac{\partial F}{\partial x^i}. \quad \square$$

**Proposition** (Second derivatives). Given a scalar field  $F$ , the collection of second derivatives  $\partial^2 F / \partial x^i \partial x^j$  do not transform as a second order tensor.

$$\begin{aligned} \frac{\partial^2 F}{\partial x^{i'} \partial x^{j'}} &= \frac{\partial}{\partial x^{i'}} \left( \frac{\partial F}{\partial x^{j'}} \right) = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial}{\partial x^i} \left( \frac{\partial F}{\partial x^{j'}} \right) = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial}{\partial x^i} \left( \frac{\partial F}{\partial x^j} \frac{\partial x^j}{\partial x^{j'}} \right) \\ &= J^i_{i'} J^j_{j'} \frac{\partial^2 F}{\partial x^i \partial x^j} + J^i_{i'} \frac{\partial F}{\partial x^j} \frac{\partial}{\partial x^i} \left( \frac{\partial x^j}{\partial x^{j'}} \right) \neq J^i_{i'} J^j_{j'} \frac{\partial^2 F}{\partial x^i \partial x^j}, \end{aligned}$$

since there exists transformations  $x \rightarrow x'$  such that

$$\frac{\partial}{\partial x^i} \left( \frac{\partial x^j}{\partial x^{j'}} \right) = \frac{\partial^2 x^i}{\partial x^{k'} \partial x^{j'}} \frac{\partial x^{k'}}{\partial x^i} \neq 0.$$

the second term is not necessarily zero.  $\square$

**Proposition** (Derivative of tensor field). Given a covariant (or contravariant) tensor field  $T_i$ , the variant  $\partial T_i / \partial x^j$  is *not* a tensor.

$$\begin{aligned} \frac{\partial T_{i'}}{\partial x^{j'}} &= \frac{\partial}{\partial x^{j'}} (J_{i'}^i T_i) = \frac{\partial}{\partial x^{j'}} \left( \frac{\partial x^i}{\partial x^{i'}} \right) T_i + J_{i'}^i \frac{\partial x^j}{\partial x^{j'}} \frac{\partial T_i}{\partial x^j} \\ &= \frac{\partial^2 x^i}{\partial x^{j'} \partial x^{i'}} T_i + J_{i'}^i J_{j'}^j \frac{\partial T_i}{\partial x^j} \neq J_{i'}^i J_{j'}^j \frac{\partial T_i}{\partial x^j}. \quad \square \end{aligned}$$

**Proposition** (Skew-symmetric part of above). The skew-symmetric part  $S_{ij}$

$$S_{ij} = \frac{\partial T_i}{\partial x^j} - \frac{\partial T_j}{\partial x^i}$$

of the variant  $\partial T_i / \partial x^j$  is a tensor.

$$\begin{aligned} S_{i'j'} &= \frac{\partial^2 x^i}{\partial x^{j'} \partial x^{i'}} T_i + J_{i'}^i J_{j'}^j \frac{\partial T_i}{\partial x^j} - \frac{\partial^2 x^j}{\partial x^{j'} \partial x^{i'}} T_j - J_{i'}^i J_{j'}^j \frac{\partial T_j}{\partial x^i} \\ &= J_{i'}^i J_{j'}^j \left( \frac{\partial T_i}{\partial x^j} - \frac{\partial T_j}{\partial x^i} \right) = J_{i'}^i J_{j'}^j S_{ij}. \quad \square \end{aligned}$$

## 2.3 Properties of tensors

**Proposition** (Linear combinations of tensors). A linear combination of tensors is a tensor, provided only tensors of the same type are summed. To illustrate, consider the linear combination of two contravariant tensors,

$$A^i = c_1 B^i + c_2 C^i \implies A^{i'} = c_1 B^{i'} + c_2 C^{i'} = c_1 J_{i'}^i B^i + c_2 J_{i'}^i C^i = J_{i'}^i (c_1 B^i + c_2 C^i) = J_{i'}^i A^i.$$

The generalization to higher order tensors is trivial. This also shows that the sum of different type of tensor components, such as  $B^i + C_i$  doesn't really make sense. The resulting object is not a tensor.

**Proposition** (Tensor product). The tensor product of tensors is a tensor. Consider the following example:

$$A^i_{jkl} = B^i_j C_{kl} \implies A^{i'}_{j'k'l'} = B^{i'}_{j'} C_{k'l'} = B^i_j C_{kl} J_{i'}^i J_{j'}^j J_{k'}^k J_{l'}^l = A^i_{jkl} J_{i'}^i J_{j'}^j J_{k'}^k J_{l'}^l.$$

Similarly, the dot product of two vector-valued tensors is a tensor, and the product of a scalar-valued tensor with a vector-valued tensor is a tensor.

**Proposition** (Contraction theorem). Contraction of a tensor is a tensor. Consider the example  $T_k = S^i_{ik}$ , where  $S^i_{jkl}$  is a tensor of order three. We have

$$S^{i'}_{j'k'} = S^i_{jk} J_{i'}^i J_{j'}^j J_{k'}^k \implies S^{i'}_{i'k'} = S^i_{jk} J_{i'}^i J_{i'}^j J_{k'}^k = S^i_{jk} \delta^j_i J_{k'}^k = S^i_{ik} J_{k'}^k,$$

hence we have

$$T_{k'} = T_k J_{k'}^k$$

and so the proposition holds.  $\square$

**Corollary** (Contraction of Kronecker delta). The contraction  $\delta^i_i$  has no free indices so it is a tensor of order zero, and an invariant. Its value equals the number of dimensions.

**Example.** Suppose  $\mathbf{V}_{ij}$  is a tensor with vector elements. The components  $V^k_{ij}$  of  $\mathbf{V}_{ij}$  with respect to the covariant basis  $\mathbf{e}_k$  is a tensor:

$$V^{k'}_{i'j'} = \mathbf{e}^{k'} \cdot \mathbf{V}_{i'j'} = J^{k'}_k J_{i'}^i J_{j'}^j \mathbf{e}^k \cdot \mathbf{V}_{ij} = J^{k'}_k J_{i'}^i J_{j'}^j V^k_{ij}.$$

**Proposition.** If a variant transforms as a tensor *from a particular coordinate system*  $x^\alpha$  *to any* coordinate system  $x^i$ , then it is a tensor. In other words,

$$T^i_j = J^i_\alpha J^\beta_j T^\alpha_\beta \implies T^{i'}_{j'} = J^{i'}_i J^j_{j'} T^i_j,$$

where we note that in the first expression  $T^\alpha_\beta$  is restricted to the particular coordinate system  $x^\alpha$ .

*Proof.* First, we write  $T^\alpha_\beta$  in terms of  $T^i_j$ ,

$$J^j_\nu J^\mu_i T^i_j = J^j_\nu J^\mu_i J^i_\alpha J^\beta_j T^\alpha_\beta = \delta^\mu_\alpha \delta^\beta_\nu T^\alpha_\beta = T^\mu_\nu.$$

From this, the proposition follows by direct substitution and applying the chain rule,

$$T^{i'}_{j'} = J^{\nu}_{j'} J^{i'}_\mu T^\mu_\nu = J^{\nu}_{j'} J^{i'}_\mu J^j_\nu J^\mu_i T^i_j = J^{i'}_i J^j_{j'} T^i_j. \quad \square$$

## 2.4 Index juggling

We've already seen an example of index juggling with the contravariant basis, see (1.9). In general, raising an index is the contraction of a lower index of a variant with the contravariant metric:

$$A^i = g^{ij} A_j. \quad (2.13)$$

Similarly, lowering an index is the contraction of an upper index of a variant with the covariant metric:

$$A_i = g_{ij} A^j. \quad (2.14)$$

More generally, indices of higher order tensors can be raised and lowered individually, for example:

$$T^i_j = g^{ik} T_{kj} = g^{ik} g_{jl} T_k^l. \quad (2.15)$$

**Proposition.** Operations of raising and lowering indices are inverses of each other. This follows directly from the definition of the contravariant metric - we have:

$$A_i = g_{ij} A^j = g_{ij} g^{jk} A_k = \delta^k_i A_k. \quad \square$$

**Proposition.** Contracted indices can be swapped invariantly, meaning

$$A^i B_i = g^{ij} A_j g_{ik} B^k = \delta^j_k A_j B^k = A_j B^j. \quad \square$$

**Proposition.** Given a symmetric tensor  $T_{ij} = T_{ji}$ , we have  $T^i_j = T_j^i$ . Proved by direct computation:

$$T^i_j = g^{ik} T_{kj} = g^{ik} T_{jk} = T_j^i. \quad \square$$

**Corollary** (Metric tensor and Kronecker delta). Objects related by index juggling are considered equivalent. From this point of view, the Kronecker symbol and the metric tensor are equivalent, since lowering upper index of  $\delta^i_j$  yields the covariant metric tensor:

$$\delta^i_j g_{ik} = g_{jk}$$

Similar relations are obtained by raising the lower index of  $\delta^i_j$  and raising (lowering) one of the indices of the covariant (contravariant) metrics.

For notational purposes, the metric tensor appearing in many tensor relations is omitted by the use of index juggling. An example is the dot product between two vectors  $\mathbf{U}$  and  $\mathbf{V}$ :

$$\mathbf{U} \cdot \mathbf{V} = g_{ij} U^i V^j = U_j V^j.$$

### 3 Differentiation

#### 3.1 Gradient of a scalar field

**Definition** (Gradient - geometric). Given a scalar field  $F$ , the gradient  $\text{grad } F$  is a vector field which, at any point, is directed in the direction of greatest increase in  $F$  with magnitude equal to the directional derivative of  $F$  in that direction.

Note that this definition is *geometric*. There is no reference to any system of coordinates, so we expect  $\text{grad } F$  to be *invariant*. We will obtain the correct definition for  $\text{grad } F$  in two steps: constructing an invariant quantity and making sure that the invariant quantity reduces to the correct form in Cartesian coordinates.

**Proposition** (Step 1: Invariance). We need a tensor expression with no free indices. So, consider

$$\text{grad } F = \frac{\partial F}{\partial x^i} \mathbf{e}^i.$$

Is this invariant? Yes, obviously:

$$(\text{grad } F)' = \frac{\partial F}{\partial x^{i'}} \mathbf{e}^{i'} = J^{i'}_{i'} \frac{\partial F}{\partial x^i} J^{i'}_j \mathbf{e}^j = \delta^i_j \frac{\partial F}{\partial x^i} \mathbf{e}^j = \frac{\partial F}{\partial x^i} \mathbf{e}^i = \text{grad } F. \quad \square$$

**Proposition** (Step 2: Cartesian expression). Now that we have an invariant expression, it is a candidate for the gradient. If the expression reduces to the gradient in Cartesian coordinates, it will represent the gradient in any coordinate system since it is invariant. In cartesian coordinates the metric reduces to the Kronecker symbol, hence we have  $\mathbf{e}^i = \mathbf{e}_i$  so

$$\text{grad } F = \frac{\partial F}{\partial x} \hat{\mathbf{x}} + \frac{\partial F}{\partial y} \hat{\mathbf{y}} + \frac{\partial F}{\partial z} \hat{\mathbf{z}}.$$

This is the expression for the gradient in Cartesian coordinates, so we're done.

**Definition** (Gradient - tensor form). The gradient of a scalar field  $f$  in coordinates  $x^i$  is

$$\text{grad } F = \frac{\partial F}{\partial x^i} \mathbf{e}^i. \quad (3.1)$$

#### 3.2 Covariant derivative

Since the derivative  $\partial/\partial x^i$  of a tensor field is not a tensor, we cannot use it to define the divergence or the Laplacian. We need a new definition, which transforms as a tensor so that we can construct invariants from tensor fields. This new operation is called the *covariant derivative*.

##### 3.2.1 Motivation

Consider a vector field  $\mathbf{V}$ . Since it is an invariant ( $\mathbf{V}' = \mathbf{V}$ ), its partial derivatives with respect to coordinates  $x^j$  transform as a covariant tensor:

$$\frac{\partial \mathbf{V}}{\partial x^{j'}} = \frac{\partial x^j}{\partial x^{j'}} \frac{\partial \mathbf{V}}{\partial x^j} = J^j_{j'} \frac{\partial \mathbf{V}}{\partial x^j}.$$

Now, we express  $\mathbf{V}$  in terms of the covariant basis  $\mathbf{V} = V^i \mathbf{e}_i$ . Differentiating with respect to  $x^j$  yields

$$\frac{\partial \mathbf{V}}{\partial x^j} = \frac{\partial V^i}{\partial x^j} \mathbf{e}_i + V^i \frac{\partial \mathbf{e}_i}{\partial x^j} = \frac{\partial V^i}{\partial x^j} \mathbf{e}_i + V^i \Gamma^k_{ij} \mathbf{e}_k.$$

Now, we factor out the  $\mathbf{e}_i$  by renaming contracted indices,

$$\frac{\partial \mathbf{V}}{\partial x^j} = \left( \frac{\partial V^i}{\partial x^j} + \Gamma_{jk}^i V^k \right) \mathbf{e}_i \equiv \nabla_j V^i \mathbf{e}_i.$$

Since we know how  $\partial \mathbf{V} / \partial x^j$  transforms, we can immediately deduce the transformation rule for the right hand side:

$$\begin{aligned} (\nabla_j V^i)' \mathbf{e}_{i'} &= J_{i'}^i (\nabla_j V^i)' \mathbf{e}_i = J_{i'}^i J_j^j \nabla_j V^i \mathbf{e}_i. \\ \implies (\nabla_j V^i)' &\equiv \nabla_{j'} V^{i'} = J_{j'}^j J_i^i \nabla_j V^i. \end{aligned}$$

Hence, we see that the quantity  $\nabla_j V^i$  transforms as a mixed tensor of order two. Now, we repeat the same calculation for the covariant components  $V_i$ ,

$$\frac{\partial \mathbf{V}}{\partial x^j} = \frac{\partial}{\partial x^j} (V_i \mathbf{e}^i) = \frac{\partial V_i}{\partial x^j} \mathbf{e}^i + V_i \frac{\partial \mathbf{e}^i}{\partial x^j} = \frac{\partial V_i}{\partial x^j} \mathbf{e}^i - V_i \Gamma_{jk}^i \mathbf{e}^k = \left[ \frac{\partial V_i}{\partial x^j} - \Gamma_{ij}^k V_k \right] \mathbf{e}^i \equiv \nabla_j V_i \mathbf{e}^i.$$

Similarly, it is obvious that  $\nabla_j V_i$  transforms as a second rank covariant tensor. Finally, since the Christoffel symbol vanishes in Cartesian coordinates,  $\nabla_j$  simply reduces to partial differentiation. Now, we're in a position to define the covariant derivative.

### 3.2.2 Definitions of derivatives, divergence and Laplacian

**Definition** (Covariant derivative). The covariant derivative of a contravariant tensor field  $T^i$  is

$$\nabla_j T^i \equiv \frac{\partial T^i}{\partial x^j} + \Gamma_{jk}^i T^k. \quad (3.2)$$

Similarly, the covariant derivative of a covariant tensor field  $T_i$  is

$$\nabla_j T_i \equiv \frac{\partial T_i}{\partial x^j} - \Gamma_{ij}^k T_k. \quad (3.3)$$

**Definition** (Contravariant derivative). The contravariant derivative is defined by raising the index of the covariant derivative:

$$\nabla^j \equiv g^{jk} \nabla_k. \quad (3.4)$$

**Proposition** (Divergence). The divergence of a vector field  $\mathbf{V}$  in terms of its components is simply

$$\text{div } \mathbf{V} = \nabla_i V^i = \nabla^i V_i. \quad (3.5)$$

This is because the divergence of a vector field must be invariant. It is defined in a coordinate free sense, at a point  $x^i$  as

$$\text{div } \mathbf{V} = \lim_{\Delta\tau \rightarrow 0} \frac{1}{\Delta\tau} \oint_{\sigma} \mathbf{V} \cdot d\boldsymbol{\sigma}.$$

where  $\Delta\tau$  is the volume enclosed by surface  $\sigma$  which surrounds the point  $x^i$ . Since the definition is coordinate free, the divergence must be an invariant. Finally, we require that it evaluates to the correct form in Cartesian coordinates, which it does since  $\Gamma_{jk}^i \equiv 0$  in Cartesian coordinates;

$$\nabla_i V^i = \frac{\partial V^x}{\partial x} + \frac{\partial V^y}{\partial y} + \frac{\partial V^z}{\partial z} = \text{div } \mathbf{V}.$$

**Proposition** (Laplacian). The Laplacian of a scalar field is given, in terms of the covariant derivative, as

$$\Delta F = g^{ij} \nabla_i \nabla_j F = \nabla^i \nabla_i F.$$

Again, seeing why this is true is very simple. It is invariant, and it reduces to the correct Laplacian in Cartesian coordinates. That's it! The Laplacian must be invariant because, by definition it is the divergence of the gradient. Since both the divergence and the gradient have coordinate independent definitions, it follows that the Laplacian is also geometric.

**Definition** (Covariant derivative for higher order tensors). We generalize the definition of the covariant derivative to higher order tensors - the rule is apparent from a mixed tensor  $T^i_j$ ,

$$\nabla_k T^i_j = \frac{\partial T^i_j}{\partial x^k} + \Gamma^i_{k\ell} T^\ell_j - \Gamma^\ell_{jk} T^\ell_i. \quad (3.6)$$

### 3.2.3 Properties of the covariant derivative

**Proposition.** The covariant derivative of a contravariant vector  $\mathbf{V}^i$  transforms as a tensor. To see why, write the invariant  $\mathbf{V}^i \mathbf{e}_i$  and consider  $\partial \mathbf{V}^i \mathbf{e}_i / \partial x^j$ .

**Proposition** (Commutativity). Covariant derivatives commute with each other, so we have

$$(\nabla_\ell \nabla_k - \nabla_k \nabla_\ell) T^i \equiv 0.$$

*Proof.* Since the resulting derivative is a tensor, if it vanishes identically in one coordinate system it must vanish in all coordinate systems. So, we show that it vanishes identically in Cartesian coordinates, where  $\nabla_k = \partial / \partial x^k$ ,

$$(\nabla_\ell \nabla_k - \nabla_k \nabla_\ell) T^i = \frac{\partial^2 T^i}{\partial x^k \partial x^\ell} - \frac{\partial^2 T^i}{\partial x^\ell \partial x^k} \equiv 0. \quad \square$$

**Proposition** (Distributivity). This property follows trivially from the fact that the covariant derivative is linear in all its arguments,

$$\nabla_k (T^i_j + S^i_j) = \nabla_k T^i_j + \nabla_k S^i_j.$$

**Proposition** (Product rule). The covariant derivative obeys the product rule for derivatives,

$$\begin{aligned} \nabla_k (T^i U_j) &= \frac{\partial (T^i U_j)}{\partial x^k} + \Gamma^i_{mk} T^m U_j - \Gamma^m_{jk} T^i U_m \\ &= T^i \left[ \frac{\partial U_j}{\partial x^k} - \Gamma^m_{jk} U_m \right] + U_j \left[ \frac{\partial T^i}{\partial x^k} + \Gamma^i_{mk} T^m \right] \\ &= T^i \nabla_k U_j + U_j \nabla_k T^i. \end{aligned}$$

**Proposition** (Product rule - vectors). The product rule applies with vector dot product, the proof is identical as above.

$$\nabla_k (\mathbf{T}^i \cdot \mathbf{U}_j) = \mathbf{T}^i \cdot \nabla_k \mathbf{U}_j + \mathbf{U}_j \cdot \nabla_k \mathbf{T}^i.$$

**Proposition** (Metric property). The bases  $\mathbf{e}_i$  and  $\mathbf{e}^i$  and the metrics  $g_{ij}$ ,  $g^{ij}$  and  $\delta^i_j$  all vanish under  $\nabla_k$ .

*Proof.* First, we show  $\nabla_k \mathbf{e}_i \equiv 0$  and  $\nabla_k \mathbf{e}^i \equiv 0$ .

$$\begin{aligned} \nabla_k \mathbf{e}_i &= \frac{\partial \mathbf{e}_i}{\partial x^k} - \Gamma^m_{ik} \mathbf{e}_m = \Gamma^j_{ik} \mathbf{e}_j - \Gamma^m_{ik} \mathbf{e}_m \equiv 0, \\ \nabla_k \mathbf{e}^i &= \frac{\partial \mathbf{e}^i}{\partial x^k} + \Gamma^i_{mk} \mathbf{e}^m = -\Gamma^i_{jk} \mathbf{e}^j + \Gamma^i_{mk} \mathbf{e}^m \equiv 0. \end{aligned}$$

Then, we simply use the product rule:

$$\begin{aligned} \nabla_k g_{ij} &= \nabla_k (\mathbf{e}_i \cdot \mathbf{e}_j) = \mathbf{e}_i \cdot \nabla_k \mathbf{e}_j + \mathbf{e}_j \cdot \nabla_k \mathbf{e}_i \equiv 0, \\ \nabla_k g^{ij} &= \nabla_k (\mathbf{e}^i \cdot \mathbf{e}^j) = \mathbf{e}^i \cdot \nabla_k \mathbf{e}^j + \mathbf{e}^j \cdot \nabla_k \mathbf{e}^i \equiv 0, \\ \nabla_k \delta^i_j &= \nabla_k (\mathbf{e}^i \cdot \mathbf{e}_j) = \mathbf{e}^i \cdot \nabla_k \mathbf{e}_j + \mathbf{e}_j \cdot \nabla_k \mathbf{e}^i \equiv 0. \quad \square \end{aligned}$$

**Proposition.** From the metrinilic property, the relation

$$\frac{\partial g_{ij}}{\partial x^k} = \Gamma_{ijk} + \Gamma_{jik}$$

follows. This is obvious:

$$0 = \nabla_k g_{ij} = \frac{\partial g_{ij}}{\partial x^k} - \Gamma_{ik}^m g_{jm} - \Gamma_{jk}^m g_{im} = \frac{\partial g_{ij}}{\partial x^k} - \Gamma_{jik} - \Gamma_{ijk}. \quad \square$$

Similarly, we have

$$\frac{\partial g^{ij}}{\partial x^k} = -\Gamma_k^{ij} - \Gamma_k^{ji}.$$

**Proposition** (Index juggling). Thanks to the metrinilic property, we can juggle indices across the covariant derivative, meaning

$$S_i = \nabla_k T_i \iff S^i = \nabla_k T^i.$$

*Proof.* We contract both sides by  $g^{im}$ , renaming the index  $i \rightarrow m$ :

$$S^i = g^{im} S_m = g^{im} \nabla_k T_m = \nabla_k (g^{im} T_m) - T_m \nabla_k g^{im} = \nabla_k (g^{im} T_m) = \nabla_k T^i.$$

**Proposition** (Index juggling). Contracted indices can exchange flavours across the covariant derivative, meaning

$$S^i \nabla_k T_i = S_i \nabla_k T^i.$$

*Proof.* We simply raise and lower the indices and use the product rule:

$$S^i \nabla_k T_i = g^{i\ell} S_\ell \nabla_k (g_{im} T^m) = g^{i\ell} S_\ell [g_{im} \nabla_k T^m + T^m \nabla_k g_{im}] = g^{i\ell} S_\ell g_{im} \nabla_k T^m = \delta_m^\ell S_\ell \nabla_k T^m = S_m \nabla_k T^m = S_m \nabla_k T^i.$$

**Proposition** (Contraction commutativity). The covariant derivative commutes with contraction. This means, in the expression

$$\nabla_k T_{ij}^i,$$

we may first apply the contraction and define  $S_j = T_{ij}^i$  and then apply  $\nabla_k$ , or we may first apply  $\nabla_k$  to the tensor  $T_{ij}^i$  and then apply the contraction of  $r$  and  $i$  indices to the result.

*Proof.* We simply write the expression with the contraction,

$$\nabla_k T_{ij}^i = \frac{\partial T_{ij}^i}{\partial x^k} + \underbrace{\Gamma_{km}^i T_{ij}^m - \Gamma_{ki}^m T_{mj}^i}_{\equiv 0} - \Gamma_{kj}^m T_{im}^i = \frac{\partial T_{ij}^i}{\partial x^k} - \Gamma_{kj}^m T_{im}^i = \nabla_k S_j. \quad \square$$