Introduction to tensors

Joakim Strandberg

July 2005

1 Contravariant and covariant vectors

1.1 The definitions of contravariant and covariant components

The purpose of this text is to be an introductory text to what tensors are and where they come from. To achieve this goal understanding of what contravariant and covariant vectors are is essential. Therefore this text begins with the definitions of these concepts.

Definition of contravariant components of a vector The contravariant components $(a^1, a^2, \ldots, a^n)$ of a n-dimensional vector $\mathbf{v}$ with respect to the linearly independent vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ (called *coordinate axes*) are defined by:

$$\mathbf{v} = a^1\mathbf{x}_1 + a^2\mathbf{x}_2 + \cdots + a^n\mathbf{x}_n \tag{1}$$

Definition of covariant components of a vector The covariant components $(b_1, b_2, \ldots, b_n)$ of a n-dimensional vector $\mathbf{v}$ with respect to the linearly independent vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ (called *coordinate axes*) are defined by:

$$\begin{array}{ccc} b_1 & = & \mathbf{v} \cdot \mathbf{x}_1 \\ b_2 & = & \mathbf{v} \cdot \mathbf{x}_2 \\ \vdots & & \vdots \\ b_n & = & \mathbf{v} \cdot \mathbf{x}_n \end{array} \tag{2}$$

Why the components $(a^1, a^2, \ldots, a^n)$ are called contravariant and $(b_1, b_2, \ldots, b_n)$ are called covariant is explained on page 2. Problem 1 illustrates how to compute the contravariant and covariant components of a given vector.

Problem 1 Given the vectors $\mathbf{x}_1 = (1, 0)$, $\mathbf{x}_2 = (1, 2)$ and $\mathbf{v} = (1, 2)$.

a) Determine the contravariant components (a^1, a^2) of $\mathbf{v}$ with respect to the vectors $\mathbf{x}_1$ and $\mathbf{x}_2$.

b) Determine the covariant components (b_1, b_2) of $\mathbf{v}$ with respect to the vectors $\mathbf{x}_1$ and $\mathbf{x}_2$.

Solution 1a The vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ are clearly independent. Using the definition of contravariant components of a vector

$$\mathbf{v} = a^1\mathbf{x}_1 + a^2\mathbf{x}_2 \tag{3}$$

$$\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix} a^1 + \begin{pmatrix}1\\2\end{pmatrix} a^2 \tag{4}$$

This is an equation system

$$\begin{pmatrix}1 & 1\\0 & 2\end{pmatrix}\begin{pmatrix}a^1\\a^2\end{pmatrix} = \begin{pmatrix}1\\2\end{pmatrix} \tag{5}$$

Using Cramer's rule

$$a^1 = \frac{\begin{vmatrix}1 & 1\\2 & 2\end{vmatrix}}{\begin{vmatrix}1 & 1\\0 & 2\end{vmatrix}} = \frac{2-2}{2} = 0 \tag{6}$$

$$a^2 = \frac{\begin{vmatrix}1 & 1\\0 & 2\end{vmatrix}}{2} = \frac{2}{2} = 1 \tag{7}$$

Thus $(a^1, a^2) = (0, 1)$.

Solution 1b Using the definition of covariant components

$$b_1 = \mathbf{v} \cdot \mathbf{x}_1 = (1, 2) \cdot (1, 0) = 1 \tag{8}$$

$$b_2 = \mathbf{v} \cdot \mathbf{x}_2 = (1, 2) \cdot (1, 2) = 5 \tag{9}$$

Thus $(b_1, b_2) = (1, 5)$.

Notice that in this example $(a^1, a^2) \neq (b_1, b_2)$.

Important! Generally the contravariant and covariant components of a vector are different. But when the "coordinate axes" $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are orthonormal (they all have unit length and and are orthogonal) there is no difference between contravariant and covariant components!

1.2 Why the components are called contravariant and covariant

Problem 2 Why are the components of a vector $\mathbf{v}$ defined by (1) and (2) called contravariant and covariant components?

Solution 2 If the coordinate axes are scaled by a factor of c, then the contravariant components become c times smaller but the covariant components become c times greater. Thus the contravariant components of a vector scale against the coordinate axes and the covariant components scale with. In Latin *contra* means *against* and *co* means *with*. Thereof the names contravariant and covariant components. Problem 3 illustrates this.

Problem 3 Given the vectors $\mathbf{x}_1 = (c, 0)$, $\mathbf{x}_2 = (c, 2c)$ and $\mathbf{v} = (1, 2)$.

a) Determine the contravariant components (a^1, a^2) of $\mathbf{v}$ with respect to the vectors $\mathbf{x}_1$ and $\mathbf{x}_2$.

b) Determine the covariant components (b_1, b_2) of $\mathbf{v}$ with respect to the vectors $\mathbf{x}_1$ and $\mathbf{x}_2$.

Solution 3a The equation system now becomes

$$\begin{pmatrix} c & c \\ 0 & 2c \end{pmatrix} \begin{pmatrix} a^1 \\ a^2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \tag{10}$$

Using Cramer's rule

$$a^1 = \frac{\begin{vmatrix} 1 & c \\ 2 & 2c \end{vmatrix}}{\begin{vmatrix} c & c \\ 0 & 2c \end{vmatrix}} = \frac{2c - 2c}{2c^2} = 0 \tag{11}$$

$$a^2 = \frac{\begin{vmatrix} c & 1 \\ 0 & 2 \end{vmatrix}}{2c^2} = \frac{2c}{2c^2} = \frac{1}{c} \tag{12}$$

Thus $(a^1, a^2) = \frac{1}{c}(0, 1)$.

Solution 3b Using the definition of covariant components

$$b_1 = \mathbf{v} \cdot \mathbf{x}_1 = (1, 2) \cdot (c, 0) = c \tag{13}$$

$$b_2 = \mathbf{v} \cdot \mathbf{x}_2 = (1, 2) \cdot (c, 2c) = 5c \tag{14}$$

Thus $(b_1, b_2) = c(1, 5)$.

1.3 Why the contravariant and covariant components are interesting

Problem 4 Why are only the definitions (1) and (2) of the components of a vector interesting? For example, it's possible to define components

$$\begin{array}{ccc} c_1 & = & \mathbf{v} \cdot \mathbf{x}_1/|\mathbf{x}_1| \\ c_2 & = & \mathbf{v} \cdot \mathbf{x}_2/|\mathbf{x}_2| \\ \vdots & & \vdots \\ c_n & = & \mathbf{v} \cdot \mathbf{x}_n/|\mathbf{x}_n| \end{array} \tag{15}$$

Why not also give the components $(c_1, c_2, \ldots, c_n)$ a name? Why are the contravariant and covariant components so "important" that they have been given names?

Solution 4 When *parametrizations of the Euclidean plane* and *tensors of first order* have been explained, it is shown on page 7 how the definition of work in physics give rise to contravariant and covariant components.

Problem 5 What is the definition of the Euclidean plane?

Solution 5 Intuitively the Euclidean plane can be imagined as a flat piece of paper. In this paper one can introduce two orthogonal coordinate axes usually denoted x- and y-axes. Two numbers x and y can then be measured to describe the position of a point. How is a point defined? Instead of saying that the two numbers *describe the position* of a point, the two numbers are defined to *be* a point.

Definition of the Euclidean plane The Euclidean plane is defined as the set of all ordered 2-tuples $\mathbf{r} = (x, y)$ called *points* where x and y are real numbers.

That the 2-tuples are ordered means that $(x, y) \neq (y, x) \Leftrightarrow x \neq y$.

Definition of a parametrization of the Euclidean plane A parametrization of the Euclidean plane is a specification of unique coordinates (u^1, u^2) (where u^1 and u^2 are real) to every point in the plane.

$$\mathbf{r}(u^1, u^2) = (x(u^1, u^2), y(u^1, u^2)) \tag{16}$$

This text does not consider all possible parametrizations but restricts itself to *allowable parametrizations*.

Important! Notice that coordinates (in a parametrization) are NOT written with subscripts (u_1, u_2) but with superscripts (u^1, u^2). The 2 reasons for this will be explained on page ???. It's not explained here in the text because the explanation requires the *Einstein summation convention* and *contravariant tensors of first order*.

Definition of an allowable parametrization of the Euclidean plane An allowable parametrization of the Euclidean plane satisfies:

i) Every point in the Euclidean plane is given unique coordinates (u^1, u^2) where $(u^1, u^2 \in \mathbb{R})$
$$\mathbf{r}(u^1, u^2) = (x(u^1, u^2), y(u^1, u^2)) \tag{17}$$

ii) The vectors $\frac{\partial \mathbf{r}}{\partial u^1}$ and $\frac{\partial \mathbf{r}}{\partial u^2}$ are defined everywhere, linearly independent and $\neq \mathbf{0}$.

Some examples of allowable parametrizations:

Problem 6 Determine which of the following parametrizations which are allowable parametrizations.

a) $\mathbf{r} = (u^1, u^2)$

b) $\mathbf{r} = (u^1 + u^2 + (u^2)^3, 2u^2)$

Solution 6a The parametrization $\mathbf{r} = (u^1, u^2)$ leads to

$$\begin{aligned} \frac{\partial \mathbf{r}}{\partial u^1} &= (1, 0) \\ \frac{\partial \mathbf{r}}{\partial u^2} &= (0, 1) \end{aligned} \tag{18}$$

It is clear that the vectors $\frac{\partial \mathbf{r}}{\partial u^1}$ and $\frac{\partial \mathbf{r}}{\partial u^2}$ are defined everywhere, linearly independent and $\neq \mathbf{0}$.

Solution 6b The parametrization $\mathbf{r} = (u^1 + u^2 + (u^2)^3, 2u^2)$ leads to

$$\begin{aligned}\frac{\partial \mathbf{r}}{\partial u^1} &= (1, 0) \\ \frac{\partial \mathbf{r}}{\partial u^2} &= (1 + 3(u^2)^2, 2)\end{aligned} \tag{19}$$

Here the vectors $\frac{\partial \mathbf{r}}{\partial u^1}$ and $\frac{\partial \mathbf{r}}{\partial u^2}$ are defined everywhere. They satisfy $\neq \mathbf{0}$ because $(u^2)^2 \geq 0$ and are linearly independent. Notice that the vector $\frac{\partial \mathbf{r}}{\partial u^2}$ is dependent on the point under consideration. This example shows that the vectors $\frac{\partial \mathbf{r}}{\partial u^1}$ and $\frac{\partial \mathbf{r}}{\partial u^2}$ may be different from point to point.

Problem 7 Why is the second requirement (ii) important in the definition of allowable parametrizations?

Solution 7 The second requirement (ii) guarantees that a vector defined at a point in the Euclidean plane always can be written as a linear combination of $\frac{\partial \mathbf{r}}{\partial u^1}$ and $\frac{\partial \mathbf{r}}{\partial u^2}$ defined at the point.

Important! From now on when coordinates are introduced in the Euclidean plane it is assumed they are allowed parametrizations.

Problem 8 Why is the letter $\mathbf{r}$ used to define a point? Why not use the notation $\mathbf{p} = (x, y)$ instead, where the letter p has been chosen since it's the first letter in the word point as is done in the book *Elementary differential geometry* by Barret O'Neill?

Solution 8 This text follows the notational convention used in the book *Vektoranalys* by Anders Ramgard. Ramgard does not explain why he chose $\mathbf{r}$ to denote a point or vector in Euclidean n-space. It may be because of the following three reasons:

(i) In Euclidean 2-space where $\mathbf{r} = (x, y)$, one can introduce a parametrization (r, ϕ) called polar coordinates $\mathbf{r} = (r \cos \phi, r \sin \phi)$. Notice that the length of the vector $\mathbf{r}$ is $|\mathbf{r}| = r$, where r originates from the word radius (here radius of a circle).

(ii) In Euclidean 3-space where $\mathbf{r} = (x, y, z)$, one can introduce a parametrization (r, ϕ, θ) called spherical coordinates $\mathbf{r} = (r \sin\theta \cos\phi, r \sin\theta \sin\phi, \cos\theta)$. The length of the vector $\mathbf{r}$ is $|\mathbf{r}| = r$, where r originates from the radius of the sphere.

(iii) Euclidean n-space is denoted $\mathbb{R}^n$, where the letter $\mathbb{R}$ here comes from the word real.

2 First order tensors defined on the Euclidean plane

The following problem leads to the definition of contravariant tensors of 1:st order.

Problem 9 Let a vector $\mathbf{v}$ be given by the contravariant components (a^1, a^2) at a fixed point P in the Euclidean plane using the coordinates u^1, u^2:

$$\mathbf{v} = a^1 \frac{\partial \mathbf{r}}{\partial u^1} + a^2 \frac{\partial \mathbf{r}}{\partial u^2} \tag{20}$$

Using another coordinate system $\overline{u}^1, \overline{u}^2$, the contravariant components of the vector $\mathbf{v}$ are $(\overline{a}^1, \overline{a}^2)$ and $\mathbf{v}$ can be written:

$$\mathbf{v} = \overline{a}^1 \frac{\partial \mathbf{r}}{\partial \overline{u}^1} + \overline{a}^2 \frac{\partial \mathbf{r}}{\partial \overline{u}^2} \tag{21}$$

What is the relation between (a^1, a^2) and $(\overline{a}^1, \overline{a}^2)$? Express $(\overline{a}^1, \overline{a}^2)$ as a function of (a^1, a^2).

Solution 9 Using the chain-rule (here $\alpha = 1, 2$):

$$\frac{\partial \mathbf{r}}{\partial u^\alpha} = \frac{\partial \mathbf{r}}{\partial \overline{u}^1}\frac{\partial \overline{u}^1}{\partial u^\alpha} + \frac{\partial \mathbf{r}}{\partial \overline{u}^2}\frac{\partial \overline{u}^2}{\partial u^\alpha} \tag{22}$$

It is allowed to use the chain-rule here since $\frac{\partial \mathbf{r}}{\partial \overline{u}^1}$, $\frac{\partial \mathbf{r}}{\partial \overline{u}^2}$, $\frac{\partial \overline{u}^1}{\partial u^\alpha}$ and $\frac{\partial \overline{u}^2}{\partial u^\alpha}$ are continuous functions. That $\frac{\partial \overline{u}^1}{\partial u^\alpha}$ and $\frac{\partial \overline{u}^2}{\partial u^\alpha}$ are continuous functions is shown in problem 11. Now using equations (20), (21) and (22) leads to expressing $(\overline{a}^1, \overline{a}^2)$ as a function of (a^1, a^2)

$$\begin{aligned}
\mathbf{v} &= a^1 \frac{\partial \mathbf{r}}{\partial u^1} + a^2 \frac{\partial \mathbf{r}}{\partial u^2} \\
&= a^1 \left(\frac{\partial \mathbf{r}}{\partial \overline{u}^1}\frac{\partial \overline{u}^1}{\partial u^1} + \frac{\partial \mathbf{r}}{\partial \overline{u}^2}\frac{\partial \overline{u}^2}{\partial u^1} \right) + a^2 \left(\frac{\partial \mathbf{r}}{\partial \overline{u}^1}\frac{\partial \overline{u}^1}{\partial u^2} + \frac{\partial \mathbf{r}}{\partial \overline{u}^2}\frac{\partial \overline{u}^2}{\partial u^2} \right) \\
&= \left(a^1 \frac{\partial \overline{u}^1}{\partial u^1} + a^2 \frac{\partial \overline{u}^1}{\partial u^2} \right) \frac{\partial \mathbf{r}}{\partial \overline{u}^1} + \left(a^1 \frac{\partial \overline{u}^2}{\partial u^1} + a^2 \frac{\partial \overline{u}^2}{\partial u^2} \right) \frac{\partial \mathbf{r}}{\partial \overline{u}^2} \\
&= \overline{a}^1 \frac{\partial \mathbf{r}}{\partial \overline{u}^1} + \overline{a}^2 \frac{\partial \mathbf{r}}{\partial \overline{u}^2}
\end{aligned} \tag{23}$$

Identifying the coefficients in front of $\frac{\partial \mathbf{r}}{\partial \overline{u}^1}$ and $\frac{\partial \mathbf{r}}{\partial \overline{u}^2}$ in the last two steps

$$\overline{a}^1 = a^1 \frac{\partial \overline{u}^1}{\partial u^1} + a^2 \frac{\partial \overline{u}^1}{\partial u^2} \tag{24}$$

$$\overline{a}^2 = a^1 \frac{\partial \overline{u}^2}{\partial u^1} + a^2 \frac{\partial \overline{u}^2}{\partial u^2} \tag{25}$$

Summarizing (here $\beta = 1, 2$):

$$\overline{a}^\beta = a^1 \frac{\partial \overline{u}^\beta}{\partial u^1} + a^2 \frac{\partial \overline{u}^\beta}{\partial u^2} \tag{26}$$

Equation (26) is the source of inspiration for tensors (1:st order contravariant tensors). There are two kinds of tensors of 1:st order:

Definition of a contravariant tensor of 1:st order (or a contravariant vector) on the Euclidean plane Let a 2-tuple of real numbers (a^1, a^2) be associated with a point P in the Euclidean plane with coordinates u^1, u^2. Associate also with the point P a 2-tuple of real numbers $\overline{a}^1, \overline{a}^2$ with respect to the coordinates $\overline{u}^1, \overline{u}^2$. If these numbers satisfy:

$$\overline{a}^\beta = a^1 \frac{\partial \overline{u}^\beta}{\partial u^1} + a^2 \frac{\partial \overline{u}^\beta}{\partial u^2} \tag{27}$$

we say that a contravariant vector at P is given. The numbers a^1, a^2 and $\overline{a}^1, \overline{a}^2$ are called the components of the contravariant vector in the respective coordinate systems u^1, u^2 or $\overline{u}^1, \overline{u}^2$. Contravariant vectors are indicated by a superscript index.

Definition of a covariant tensor of 1:st order (or a covariant vector) on the Euclidean plane Let a 2-tuple of real numbers (b_1, b_2) be associated with a point P in the Euclidean plane with coordinates u^1, u^2. Associate also with the point P a 2-tuple of real numbers $\overline{b}_1, \overline{b}_2$ with respect to the coordinates $\overline{u}^1, \overline{u}^2$. If these numbers satisfy:

$$\overline{b}_\beta = b_1 \frac{\partial u^1}{\partial \overline{u}^\beta} + b_2 \frac{\partial u^2}{\partial \overline{u}^\beta} \tag{28}$$

we say that a covariant vector at P is given. The numbers b_1, b_2 and $\overline{b}_1, \overline{b}_2$ are called the components of the covariant vector in the respective coordinate systems u^1, u^2 or $\overline{u}^1, \overline{u}^2$. Covariant vectors are indicated by a subscript index.

Problem 10 Determine if the vector (du^1, du^2) is contravariant, covariant or neither.

Solution 10 The vector (du^1, du^2) is a contravariant vector since the components du^1 and du^2 transform contravariantly under change of coordinate system from (u^1, u^2) to $(\overline{u}^1, \overline{u}^2)$ accordning to the chain-rule

$$du^1 = d\overline{u}^1 \frac{\partial u^1}{\partial \overline{u}^1} + d\overline{u}^2 \frac{\partial u^1}{\partial \overline{u}^2} \tag{29}$$

$$du^2 = d\overline{u}^1 \frac{\partial u^2}{\partial \overline{u}^1} + d\overline{u}^2 \frac{\partial u^2}{\partial \overline{u}^2} \tag{30}$$

Summarizing:

$$du^\gamma = d\overline{u}^1 \frac{\partial u^\gamma}{\partial \overline{u}^1} + d\overline{u}^2 \frac{\partial u^\gamma}{\partial \overline{u}^2} \qquad (\gamma = 1, 2) \tag{31}$$

That the vector (du^1, du^2) transforms contravariantly is part of the reason why coordinates are written with superscripts.

Returning to problem 4 posed on page 3 Why are the contravariant and covariant components so "important" that they have been given names?

Solution When calculating the work performed by a force on a particle under an infinitesimal displacement in the Euclidean plane, its natural to assume that the work performed is independent of the coordinate system used. If the work performed is independent of the coordinate system then it turns out that if the displacement is expressed by its contravariant components then the force must be expressed by its covariant components. To see this, assume that a force $\mathbf{F}$ is exerted on a particle which moves (du^1, du^2) in an ordinary Cartesian coordinate system $\mathbf{r} = (u^1, u^2)$. The work done dW is

$$dW = F_1 du^1 + F_2 du^2 \tag{32}$$

With new coordinates $(\overline{u}^1, \overline{u}^2)$, the force is $\mathbf{F} = (F_{\overline{1}}, F_{\overline{2}})$. We would like to write

$$dW = F_{\overline{1}} d\overline{u}^1 + F_{\overline{2}} d\overline{u}^2 \tag{33}$$

What then is the connection between (F_1, F_2) and $(F_{\overline{1}}, F_{\overline{2}})$? Using the fact that du^1 and du^2 transform contravariantly

$$du^\gamma = d\overline{u}^1 \frac{\partial u^\gamma}{\partial \overline{u}^1} + d\overline{u}^2 \frac{\partial u^\gamma}{\partial \overline{u}^2} \tag{34}$$

$$\begin{aligned} dW &= F_1 du^1 + F_2 du^2 \\ &= F_1 \left(d\overline{u}^1 \frac{\partial u^1}{\partial \overline{u}^1} + d\overline{u}^2 \frac{\partial u^1}{\partial \overline{u}^2} \right) + F_2 \left(d\overline{u}^1 \frac{\partial u^2}{\partial \overline{u}^1} + d\overline{u}^2 \frac{\partial u^2}{\partial \overline{u}^2} \right) \\ &= \left(F_1 \frac{\partial u^1}{\partial \overline{u}^1} + F_2 \frac{\partial u^2}{\partial \overline{u}^1} \right) d\overline{u}^1 + \left(F_1 \frac{\partial u^1}{\partial \overline{u}^2} + F_2 \frac{\partial u^2}{\partial \overline{u}^2} \right) d\overline{u}^2 \\ &= F_{\overline{1}} d\overline{u}^1 + F_{\overline{2}} d\overline{u}^2 \end{aligned} \tag{35}$$

Identifying:

$$F_{\overline{1}} = F_1 \frac{\partial u^1}{\partial \overline{u}^1} + F_2 \frac{\partial u^2}{\partial \overline{u}^1} \tag{36}$$

$$F_{\overline{2}} = F_1 \frac{\partial u^1}{\partial \overline{u}^2} + F_2 \frac{\partial u^2}{\partial \overline{u}^2} \tag{37}$$

Equations (36) and (37) can be summarized by:

$$F_{\overline{\beta}} = F_1 \frac{\partial u^1}{\partial \overline{u}^\beta} + F_2 \frac{\partial u^2}{\partial \overline{u}^\beta} \tag{38}$$

Comparing (38) and (28) shows that the components (F_1, F_2) must transform covariantly if the work performed by the force is independent of the choice of coordinates used.

Problem 11 A contravariant 1:st order tensor is a 2-tuple of real numbers (a^1, a^2) that is transformed under change from coordinates (u^1, u^2) to $(\overline{u}^1, \overline{u}^2)$ as:

$$\overline{a}^\beta = a^1 \frac{\partial \overline{u}^\beta}{\partial u^1} + a^2 \frac{\partial \overline{u}^\beta}{\partial u^2} \tag{39}$$

But is it obvious that $\frac{\partial \overline{u}^\beta}{\partial u^1}$ and $\frac{\partial \overline{u}^\beta}{\partial u^2}$ always exists? It is sufficient in this problem to investigate if $\frac{\partial \overline{u}^1}{\partial u^1}$ exists under change of allowable parametrizations (it is not difficult to generalize). If $\frac{\partial \overline{u}^1}{\partial u^1}$ doesn't always exists, then a first order tensor is defined by an expression that is not always defined. How can one define something by something that is not defined???

Solution 11 The expression $\frac{\partial \overline{u}^1}{\partial u^1}$ is always well defined under coordinate change between allowable parametrizations. To see this, use $\overline{u}^1 = \overline{u}^1(x, y)$ and the chain rule:

$$\frac{\partial \overline{u}^1}{\partial u^1} = \frac{\partial \overline{u}^1}{\partial x} \frac{\partial x}{\partial u^1} + \frac{\partial \overline{u}^1}{\partial y} \frac{\partial y}{\partial u^1} \tag{40}$$

The expression $\frac{\partial \overline{u}^1}{\partial u^1}$ is defined if $\frac{\partial \overline{u}^1}{\partial x}$, $\frac{\partial x}{\partial u^1}$, $\frac{\partial \overline{u}^1}{\partial y}$ and $\frac{\partial y}{\partial u^1}$ are defined (it is then also allowed to use the chain-rule). Since $\frac{\partial \mathbf{r}}{\partial u^1} = (\frac{\partial x(u^1,u^2)}{\partial u^1}, \frac{\partial y(u^1,u^2)}{\partial u^1})$ is well defined according to the definition of allowable parametrizations $\Rightarrow \frac{\partial x}{\partial u^1}$ and $\frac{\partial y}{\partial u^1}$ are defined. But is $\frac{\partial \overline{u}^1}{\partial x}$ and $\frac{\partial \overline{u}^1}{\partial y}$ defined? Since $\frac{\partial \mathbf{r}}{\partial \overline{u}^1}$ and $\frac{\partial \mathbf{r}}{\partial \overline{u}^2}$ are defined, the differentials dx and dy can be written:

$$dx = \frac{\partial x}{\partial \overline{u}^1} d\overline{u}^1 + \frac{\partial x}{\partial \overline{u}^2} d\overline{u}^2 \tag{41}$$

$$dy = \frac{\partial y}{\partial \overline{u}^1} d\overline{u}^1 + \frac{\partial y}{\partial \overline{u}^2} d\overline{u}^2 \tag{42}$$

$$\begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} \frac{\partial x}{\partial \overline{u}^1} & \frac{\partial x}{\partial \overline{u}^2} \\ \frac{\partial y}{\partial \overline{u}^1} & \frac{\partial y}{\partial \overline{u}^2} \end{pmatrix} \begin{pmatrix} d\overline{u}^1 \\ d\overline{u}^2 \end{pmatrix} \tag{43}$$

Using Cramer's rule:

$$d\overline{u}^1 = \frac{\begin{vmatrix} dx & \frac{\partial x}{\partial \overline{u}^2} \\ dy & \frac{\partial y}{\partial \overline{u}^2} \end{vmatrix}}{\begin{vmatrix} \frac{\partial x}{\partial \overline{u}^1} & \frac{\partial x}{\partial \overline{u}^2} \\ \frac{\partial y}{\partial \overline{u}^1} & \frac{\partial y}{\partial \overline{u}^2} \end{vmatrix}} = \frac{\frac{\partial y}{\partial \overline{u}^2}}{\frac{\partial x}{\partial \overline{u}^1}\frac{\partial y}{\partial \overline{u}^2} - \frac{\partial y}{\partial \overline{u}^1}\frac{\partial x}{\partial \overline{u}^2}} dx - \frac{\frac{\partial x}{\partial \overline{u}^2}}{\frac{\partial x}{\partial \overline{u}^1}\frac{\partial y}{\partial \overline{u}^2} - \frac{\partial y}{\partial \overline{u}^1}\frac{\partial x}{\partial \overline{u}^2}} dy \tag{44}$$

If $\frac{\partial \overline{u}^1}{\partial x}$, $\frac{\partial \overline{u}^1}{\partial y}$ are defined then $d\overline{u}^1$ can be written:

$$d\overline{u}^1 = \frac{\partial \overline{u}^1}{\partial x} dx + \frac{\partial \overline{u}^1}{\partial y} dy \tag{45}$$

Identifying:

$$\frac{\partial \overline{u}^1}{\partial x} = \frac{\frac{\partial y}{\partial \overline{u}^2}}{\frac{\partial x}{\partial \overline{u}^1}\frac{\partial y}{\partial \overline{u}^2} - \frac{\partial y}{\partial \overline{u}^1}\frac{\partial x}{\partial \overline{u}^2}} \tag{46}$$

$$\frac{\partial \overline{u}^1}{\partial y} = -\frac{\frac{\partial x}{\partial \overline{u}^2}}{\frac{\partial x}{\partial \overline{u}^1}\frac{\partial y}{\partial \overline{u}^2} - \frac{\partial y}{\partial \overline{u}^1}\frac{\partial x}{\partial \overline{u}^2}} \tag{47}$$

The denominator $\frac{\partial x}{\partial \overline{u}^1}\frac{\partial y}{\partial \overline{u}^2} - \frac{\partial y}{\partial \overline{u}^1}\frac{\partial x}{\partial \overline{u}^2}$ is equal to the determinant

$$\begin{vmatrix} \frac{\partial x}{\partial \overline{u}^1} & \frac{\partial x}{\partial \overline{u}^2} \\ \frac{\partial y}{\partial \overline{u}^1} & \frac{\partial y}{\partial \overline{u}^2} \end{vmatrix} \tag{48}$$

which is equal to $\pm$the area spanned by the vectors $\frac{\partial \mathbf{r}}{\partial \overline{u}^1}$ and $\frac{\partial \mathbf{r}}{\partial \overline{u}^2}$. And this area is $\neq 0$ since $\frac{\partial \mathbf{r}}{\partial \overline{u}^1}$ and $\frac{\partial \mathbf{r}}{\partial \overline{u}^2}$ are two linearly independent vectors $\neq \mathbf{0}$. The nominators $\frac{\partial y}{\partial \overline{u}^2}$ and $\frac{\partial x}{\partial \overline{u}^2}$ are defined since $\frac{\partial \mathbf{r}}{\partial \overline{u}^2}$ is defined. Thus the right-hand sides of equations (46) and (47) are defined, and hence the left-hand sides $\frac{\partial \overline{u}^1}{\partial x}$ and $\frac{\partial \overline{u}^1}{\partial y}$ are also defined. Thus the expression $\frac{\partial \overline{u}^1}{\partial u^1}$ is always defined under coordinate change between allowed parametrizations.

Problem 12 Determine if the vector $\frac{\partial \mathbf{r}}{\partial u^1}$ is contravariant, covariant or neither.

Solution 12 Investigating how the vector $\frac{\partial \mathbf{r}}{\partial \overline{u}^1}$ transforms when coordinates are changed from $\overline{u}^\alpha$ to u^β.

$$\frac{\partial \mathbf{r}}{\partial \overline{u}^1} = \left(\frac{\partial x}{\partial \overline{u}^1}, \frac{\partial y}{\partial \overline{u}^1} \right) \tag{49}$$

Using the chain-rule

$$\frac{\partial x}{\partial \overline{u}^1} = \frac{\partial x}{\partial u^1}\frac{\partial u^1}{\partial \overline{u}^1} + \frac{\partial x}{\partial u^2}\frac{\partial u^2}{\partial \overline{u}^1} \tag{50}$$

$$\frac{\partial y}{\partial \overline{u}^1} = \frac{\partial y}{\partial u^1}\frac{\partial u^1}{\partial \overline{u}^1} + \frac{\partial y}{\partial u^2}\frac{\partial u^2}{\partial \overline{u}^1} \tag{51}$$

Thus

$$\begin{aligned} \frac{\partial \mathbf{r}}{\partial \overline{u}^1} &= \left(\frac{\partial x}{\partial u^1}\frac{\partial u^1}{\partial \overline{u}^1} + \frac{\partial x}{\partial u^2}\frac{\partial u^2}{\partial \overline{u}^1}, \frac{\partial y}{\partial u^1}\frac{\partial u^1}{\partial \overline{u}^1} + \frac{\partial y}{\partial u^2}\frac{\partial u^2}{\partial \overline{u}^1} \right) \\ &= \left(\frac{\partial x}{\partial u^1}, \frac{\partial y}{\partial u^1} \right) \frac{\partial u^1}{\partial \overline{u}^1} + \left(\frac{\partial x}{\partial u^2}, \frac{\partial y}{\partial u^2} \right) \frac{\partial u^2}{\partial \overline{u}^1} \\ &= \frac{\partial \mathbf{r}}{\partial u^1}\frac{\partial u^1}{\partial \overline{u}^1} + \frac{\partial \mathbf{r}}{\partial u^2}\frac{\partial u^2}{\partial \overline{u}^1} \end{aligned} \tag{52}$$

$$\frac{\partial \mathbf{r}}{\partial \overline{u}^\beta} = \frac{\partial \mathbf{r}}{\partial u^1}\frac{\partial u^1}{\partial \overline{u}^\beta} + \frac{\partial \mathbf{r}}{\partial u^2}\frac{\partial u^2}{\partial \overline{u}^\beta} \tag{53}$$

This means that $\frac{\partial \mathbf{r}}{\partial u^\beta}$ does not define a contravariant nor covariant vector, but it is the "vector" $(\frac{\partial \mathbf{r}}{\partial u^1}, \frac{\partial \mathbf{r}}{\partial u^2})$ that transforms covariantly. Covarians is usually denoted by a subscript, which inspires the following notation that is frequently used

$$\mathbf{r}_\alpha \equiv \frac{\partial \mathbf{r}}{\partial u^\alpha} \tag{54}$$

$$\mathbf{r}_{\overline{\alpha}} \equiv \frac{\partial \mathbf{r}}{\partial \overline{u}^\alpha} \tag{55}$$

Using this notation, it's possible to write

$$\mathbf{r}_{\overline{\beta}} = \mathbf{r}_1 \frac{\partial u^1}{\partial \overline{u}^\beta} + \mathbf{r}_2 \frac{\partial u^2}{\partial \overline{u}^\beta} \tag{56}$$

Problem 13 Do the covariant components b_β of a vector $\mathbf{v}$ (eq. (2) defined on page 1) with respect to the coordinate axes $\mathbf{r}_1$ and $\mathbf{r}_2$ transform covariantly according to the definition of covariant 1:st order tensors?

$$b_\beta = \mathbf{v} \cdot \mathbf{r}_\beta \tag{57}$$

Solution 13 Yes, they do.

$$\begin{aligned} b_{\overline{\beta}} &= \mathbf{v} \cdot \mathbf{r}_{\overline{\beta}} \\ &= \mathbf{v} \cdot \left(\mathbf{r}_1 \frac{\partial u^1}{\partial \overline{u}^\beta} + \mathbf{r}_2 \frac{\partial u^2}{\partial \overline{u}^\beta} \right) \\ &= (\mathbf{v} \cdot \mathbf{r}_1) \frac{\partial u^1}{\partial \overline{u}^\beta} + (\mathbf{v} \cdot \mathbf{r}_2) \frac{\partial u^2}{\partial \overline{u}^\beta} \\ &= b_1 \frac{\partial u^1}{\partial \overline{u}^\beta} + b_2 \frac{\partial u^2}{\partial \overline{u}^\beta} \end{aligned} \tag{58}$$

3 First order tensors defined on Euclidean n-space

3.1 The generalization

Mathematicians like to generalize things. How can 1:st order tensors defined on the Euclidean plane be generalized? Probably the first obvious generalization is to generalize from to the Euclidean plane to Euclidean n-space.

Definition of Euclidean n-space Let n be a positive integer. Euclidean n-space is defined as the set of all ordered n-tuples $\mathbf{r} = (p_1, \ldots, p_n)$ called *points* where $p_1, \ldots, p_n$ are real numbers (the index k of p_k does not imply covarians, but is only an index).

Definition of an allowable parametrization of Euclidean n-space An allowable parametrization of Euclidean n-space satisfies:

i) Every point in Euclidean n-space is given unique coordinates $(u^1, \ldots, u^n)$ where $(u^1, \ldots, u^n \in \mathbb{R})$

$$\mathbf{r}(u^1, \ldots, u^n) = (p_1(u^1, \ldots, u^n), \ldots, p_n(u^1, \ldots, u^n)) \tag{59}$$

ii) The vectors $\frac{\partial \mathbf{r}}{\partial u^1}, \dots, \frac{\partial \mathbf{r}}{\partial u^n}$ are defined everywhere, linearly independent and $\neq \mathbf{0}$.

As before, when coordinates are introduced it is assumed they are allowed parametrizations. Generalizing problem 9:

Problem 14 Let a n-dimensional vector $\mathbf{v}$ be given by the contravariant components $(a^1, \dots, a^n)$ at a fixed point P in Euclidean n-space using the coordinates $u^1, \dots, u^n$:

$$\mathbf{v} = a^1 \frac{\partial \mathbf{r}}{\partial u^1} + \dots + a^n \frac{\partial \mathbf{r}}{\partial u^n} \tag{60}$$

Using another coordinate system $\overline{u}^1, \dots, \overline{u}^n$, the contravariant components of the vector $\mathbf{v}$ are $(\overline{a}^1, \dots, \overline{a}^n)$ and $\mathbf{v}$ can be written:

$$\mathbf{v} = \overline{a}^1 \frac{\partial \mathbf{r}}{\partial \overline{u}^1} + \dots + \overline{a}^n \frac{\partial \mathbf{r}}{\partial \overline{u}^n} \tag{61}$$

What is the relation between (a^1, a^2) and $(\overline{a}^1, \overline{a}^2)$? Express $(\overline{a}^1, \overline{a}^2)$ as a function of (a^1, a^2).

Solution 14 Using a similar argument as in problem 9

$$\overline{a}^\beta = a^1 \frac{\partial \overline{u}^\beta}{\partial u^1} + \dots + a^n \frac{\partial \overline{u}^\beta}{\partial u^n} \tag{62}$$

Therefore the generalization of 1:st order tensors to be defined on Euclidean n-space becomes:

Definition of a contravariant tensor of 1:st order (or a contravariant vector) on Euclidean n-space A contravariant tensor of 1:st order is a quantity whose n components a^α are transformed according to

$$\overline{a}^\beta = a^1 \frac{\partial \overline{u}^\beta}{\partial u^1} + \ldots + a^n \frac{\partial \overline{u}^\beta}{\partial u^n} \tag{63}$$

under change of coordinate system in Euclidean n-space.

Definition of a covariant tensor of 1:st order (or a covariant vector) on Euclidean n-space A covariant tensor of 1:st order is a quantity whose n components b_β are transformed according to

$$\overline{b}_\beta = b_1 \frac{\partial u^1}{\partial \overline{u}^\beta} + \ldots + b_n \frac{\partial u^n}{\partial \overline{u}^\beta} \tag{64}$$

under change of coordinate system in Euclidean n-space.

3.2 The Einstein summation convention

Here it is convenient to introduce the Einstein summation convention that is used extensively in tensor analysis[1] and relativity theory. The point of the Einstein summation

[1] The term *tensor analysis* was coined by Einstein in 1916.

convention is to simplify notation. Consider the definition of 1:st order contravariant tensors. The notation may first be simplified by the introduction of a summation sign:

$$\begin{aligned}\overline{a}^{\beta} &= a^1 \frac{\partial \overline{u}^{\beta}}{\partial u^1} + \ldots + a^n \frac{\partial \overline{u}^{\beta}}{\partial u^n} \\ &= \sum_{\alpha=1}^{n} a^{\alpha} \frac{\partial \overline{u}^{\beta}}{\partial u^{\alpha}}\end{aligned} \tag{65}$$

This sum is written in the Einstein notation as:

$$\sum_{\alpha=1}^{n} a^{\alpha} \frac{\partial \overline{u}^{\beta}}{\partial u^{\alpha}} = a^{\alpha} \frac{\partial \overline{u}^{\beta}}{\partial u^{\alpha}} \tag{66}$$

In the expression $\frac{\partial \overline{u}^{\beta}}{\partial u^{\alpha}}$ the index β is considered superscript and the index α is considered subscript.

SUMMATION CONVENTION If in a product a letter figures twice, once as superscript and once as subscript, summation must be carried out from 1 to n with respect to this letter. The summation sign $\sum$ will be omitted.

For example

$$a^{\alpha} b_{\alpha} = \sum_{\alpha=1}^{n} a^{\alpha} b_{\alpha} = a^1 b_1 + a^2 b_2 + \cdots + a^n b_n \tag{67}$$

Applying the Einstein summation to the definition of 1:st order covariant tensors:

$$\begin{aligned}\bar{b}_{\beta} &= b_1 \frac{\partial u^1}{\partial \overline{u}^{\beta}} + \ldots + b_n \frac{\partial u^n}{\partial \overline{u}^{\beta}} \\ &= \sum_{\gamma=1}^{n} b_{\gamma} \frac{\partial u^{\gamma}}{\partial \overline{u}^{\beta}} \\ &= b_{\gamma} \frac{\partial u^{\gamma}}{\partial \overline{u}^{\beta}}\end{aligned} \tag{68}$$

Equation (20) on page 5 can be written

$$\mathbf{v} = a^{\beta} \mathbf{r}_{\beta} \tag{69}$$

Rewriting equation (56) with the Einstein summation notation (and of course generalized to Euclidean n-space)

$$\mathbf{r}_{\overline{\beta}} = \mathbf{r}_{\alpha} \frac{\partial u^{\alpha}}{\partial \overline{u}^{\beta}} \tag{70}$$

An index with respect to which summation must be carried out is called a *summation index* or *dummy index*. The other indices are said to be *free indices*. Dummy indices may be changed during computations without warning, for example

$$a^{\alpha} b_{\alpha} = a^i b_i = a^1 b_1 + a^2 b_2 + \cdots + a^n b_n \tag{71}$$

$$a^{\alpha} b_{\alpha} + a^{\gamma} c_{\gamma} = a^{\alpha} (b_{\alpha} + c_{\alpha}) \tag{72}$$

Problem 15 Why are the coordinates $(u^1, u^2, \ldots, u^n)$ written with superscripts?

Solution 15 One could be led to believe that the vector $(u^1, u^2, \ldots, u^n)$ transforms contravariantly under coordinate changes because the indices of the coordinate variables are written with superscripts. This is not true. It is the vector $(du^1, du^2, \ldots, du^n)$ that transforms contravariantly. In expressions du^α, $\frac{\partial u^\alpha}{\partial u^\beta}$, $\frac{\partial \mathbf{r}}{\partial u^\beta}$ the index α is considered superscript and β subscript. This convention together with the Einstein summation convention simplifies the tensor notation.

4 Higher order tensors

Now that first order tensors are generalized to be defined on Euclidean n-space, can the tensor concept be generalized further? Yes, it can and the way first order tensors are generalized is inspired by the definition of work in physics. The following problems illustrate this.

Problem 16 Let a^α be the contravariant components of an arbitrary vector $\mathbf{v}$ which means $\overline{a}^\gamma = a^\alpha \frac{\partial \overline{u}^\gamma}{\partial u^\alpha}$. Consider the expression

$$I = b_\alpha a^\alpha \tag{73}$$

Show that if I is invariant under coordinate change, then b_α are the covariant components of a vector $\mathbf{w}$ which means b_α transforms covariantly.

Solution 16 Let $\overline{a}^\gamma$, $\overline{b}_\gamma$ be the components with respect to the coordinate system $\overline{u}^\alpha$. Since I is invariant

$$I = \overline{b}_\gamma \overline{a}^\gamma = b_\alpha a^\alpha \tag{74}$$

Inserting $\overline{a}^\gamma = \frac{\partial \overline{u}^\gamma}{\partial u^\alpha} a^\alpha$

$$\overline{b}_\gamma \frac{\partial \overline{u}^\gamma}{\partial u^\alpha} a^\alpha = b_\alpha a^\alpha \tag{75}$$

This expression is valid for an arbitrary a^α. Thus it's possible to identify

$$\overline{b}_\gamma \frac{\partial \overline{u}^\gamma}{\partial u^\alpha} = b_\alpha \tag{76}$$

Multiply both sides with $\frac{\partial u^\alpha}{\partial \overline{u}^\beta}$ and sum with respect to α from 1 to n

$$\overline{b}_\gamma \frac{\partial \overline{u}^\gamma}{\partial u^\alpha} \frac{\partial u^\alpha}{\partial \overline{u}^\beta} = b_\alpha \frac{\partial u^\alpha}{\partial \overline{u}^\beta} \tag{77}$$

Here is now used

$$\frac{\partial \overline{u}^\gamma}{\partial u^\alpha} \frac{\partial u^\alpha}{\partial \overline{u}^\beta} = \delta^\gamma_\beta \tag{78}$$

where δ^γ_β is the Kronecker delta which is $= 1$ when $\beta = \gamma$ and $= 0$ otherwise. Equation (78) follows from the fact that the variables $\overline{u}^1, \overline{u}^2, \ldots, \overline{u}^n$ are independent and the chain-rule

$$\begin{aligned} \delta^\gamma_\beta &= \frac{\partial \overline{u}^\beta}{\partial \overline{u}^\gamma} \\ &= \frac{\partial \overline{u}^\beta}{\partial u^1} \frac{\partial u^1}{\partial \overline{u}^\gamma} + \frac{\partial \overline{u}^\beta}{\partial u^2} \frac{\partial u^2}{\partial \overline{u}^\gamma} + \cdots + \frac{\partial \overline{u}^\beta}{\partial u^n} \frac{\partial u^n}{\partial \overline{u}^\gamma} \qquad \text{(no summation over } n) \\ &= \frac{\partial \overline{u}^\gamma}{\partial u^\alpha} \frac{\partial u^\alpha}{\partial \overline{u}^\beta} \end{aligned} \tag{79}$$

This means

$$\overline{b}_\gamma \delta^\gamma_\beta = b_\alpha \frac{\partial u^\alpha}{\partial \overline{u}^\beta} \tag{80}$$

$$\overline{b}_\beta = b_\alpha \frac{\partial u^\alpha}{\partial \overline{u}^\beta} \tag{81}$$

Thus b_α transforms covariantly. This problem can be generalized, for example to define 2:nd order covariant tensors as in problem 17 and 2:nd order contravariant tensors as in problem 18.

Problem 17 Let b^α and c^β be two arbitrary contravariant 1:st order tensors. Consider

$$I = a_{\alpha\beta} b^\alpha c^\beta \tag{82}$$

If I is invariant under coordinate change, determine the transformation behaviour of $a_{\alpha\beta}$.

Solution 17 Since I is invariant

$$I = \overline{a}_{\omega\lambda} \overline{b}^\omega \overline{c}^\lambda = a_{\alpha\beta} b^\alpha c^\beta \tag{83}$$

Inserting $\overline{b}^\omega = b^\alpha \frac{\partial \overline{u}^\omega}{\partial u^\alpha}$ and $\overline{c}^\lambda = c^\beta \frac{\partial \overline{u}^\lambda}{\partial u^\beta}$

$$I = \overline{a}_{\omega\lambda} \left(b^\alpha \frac{\partial \overline{u}^\omega}{\partial u^\alpha} \right) \left(c^\beta \frac{\partial \overline{u}^\lambda}{\partial u^\beta} \right) = a_{\alpha\beta} b^\alpha c^\beta \tag{84}$$

Since b^α and c^β are arbitrary it's possible to identify

$$a_{\alpha\beta} = \overline{a}_{\omega\lambda} \frac{\partial \overline{u}^\omega}{\partial u^\alpha} \frac{\partial \overline{u}^\lambda}{\partial u^\beta} \tag{85}$$

Multiply both sides by $\frac{\partial u^\alpha}{\partial \overline{u}^\mu} \frac{\partial u^\beta}{\partial \overline{u}^\nu}$ and sum with respect to both α and β from 1 to n

$$a_{\alpha\beta} \frac{\partial u^\alpha}{\partial \overline{u}^\mu} \frac{\partial u^\beta}{\partial \overline{u}^\nu} = \overline{a}_{\omega\lambda} \frac{\partial \overline{u}^\omega}{\partial u^\alpha} \frac{\partial u^\alpha}{\partial \overline{u}^\mu} \frac{\partial \overline{u}^\lambda}{\partial u^\beta} \frac{\partial u^\beta}{\partial \overline{u}^\nu} = \overline{a}_{\omega\lambda} \delta^\omega_\mu \delta^\lambda_\nu \tag{86}$$

$$\overline{a}_{\mu\nu} = a_{\alpha\beta} \frac{\partial u^\alpha}{\partial \overline{u}^\mu} \frac{\partial u^\beta}{\partial \overline{u}^\nu} \tag{87}$$

If I is invariant under coordinate change then $a_{\alpha\beta}$ defines a 2:nd order covariant tensor.

Definition of a covariant tensor of 2:nd order defined on Euclidean n-space A covariant tensor of order 2 is a quantity whose n^2 components $a_{\alpha\beta}$ are transformed according to

$$\overline{a}_{\mu\nu} = a_{\alpha\beta} \frac{\partial u^\alpha}{\partial \overline{u}^\mu} \frac{\partial u^\beta}{\partial \overline{u}^\nu} \tag{88}$$

under change of coordinate system in Euclidean n-space.

Problem 18 Let b_α and c_β be two arbitrary covariant 1:st order tensors. Consider

$$I = a^{\alpha\beta} b_\alpha c_\beta \tag{89}$$

If I is invariant under coordinate change, determine the transformation behaviour of $a^{\alpha\beta}$.

Solution 18 Since I is invariant

$$I = \overline{a}^{\omega\lambda}\overline{b}_\omega \overline{c}_\lambda = a^{\alpha\beta} b_\alpha c_\beta \tag{90}$$

Inserting $\overline{b}_\omega = b_\alpha \frac{\partial u^\alpha}{\partial \overline{u}^\omega}$ and $\overline{c}_\lambda = c_\beta \frac{\partial u^\beta}{\partial \overline{u}^\lambda}$

$$I = \overline{a}^{\omega\lambda} \left(b_\alpha \frac{\partial u^\alpha}{\partial \overline{u}^\omega} \right) \left(c_\beta \frac{\partial u^\beta}{\partial \overline{u}^\lambda} \right) = a^{\alpha\beta} b_\alpha c_\beta \tag{91}$$

Since b_α and c_β are arbitrary it's possible to identify

$$a^{\alpha\beta} = \overline{a}^{\omega\lambda} \frac{\partial u^\alpha}{\partial \overline{u}^\omega} \frac{\partial u^\beta}{\partial \overline{u}^\lambda} \tag{92}$$

Multiply both sides by $\frac{\partial \overline{u}^\mu}{\partial u^\alpha} \frac{\partial \overline{u}^\nu}{\partial u^\beta}$ and sum with respect to both α and β from 1 to n

$$a^{\alpha\beta} \frac{\partial \overline{u}^\mu}{\partial u^\alpha} \frac{\partial \overline{u}^\nu}{\partial u^\beta} = \overline{a}^{\omega\lambda} \frac{\partial u^\alpha}{\partial \overline{u}^\omega} \frac{\partial \overline{u}^\mu}{\partial u^\alpha} \frac{\partial u^\beta}{\partial \overline{u}^\lambda} \frac{\partial \overline{u}^\nu}{\partial u^\beta} = \overline{a}^{\omega\lambda} \delta^\mu_\omega \delta^\nu_\lambda \tag{93}$$

$$\overline{a}^{\mu\nu} = a^{\alpha\beta} \frac{\partial \overline{u}^\mu}{\partial u^\alpha} \frac{\partial \overline{u}^\nu}{\partial u^\beta} \tag{94}$$

If I is invariant under coordinate change then $a^{\alpha\beta}$ defines a 2:nd order contravariant tensor.

Definition of a contravariant tensor of 2:nd order defined on Euclidean n-space A contravariant tensor of order 2 is a quantity whose n^2 components $a^{\alpha\beta}$ are transformed according to

$$\overline{a}^{\mu\nu} = a^{\alpha\beta} \frac{\partial \overline{u}^\mu}{\partial u^\alpha} \frac{\partial \overline{u}^\nu}{\partial u^\beta} \tag{95}$$

under change of coordinate system in Euclidean n-space.

Now, there is a third kind of tensor, the *mixed* 2:nd order tensor.

Problem 19 Let b_α be an arbitrary covariant 1:st order tensor and c^β an arbitrary contravariant 1:st order tensor. Consider

$$I = {a_\alpha}^\beta b^\alpha c_\beta \tag{96}$$

If I is invariant under coordinate change, determine the transformation behaviour of ${a_\alpha}^\beta$.

Important! Notice that ${a_\alpha}^\beta$ is not written as a^β_α. This is because in tensor notation it's not only if the index is superscript or subscript that matters, but the position is also important. This has to do with the possibility of associating a new tensor with a given one through the metric tensor by raising or lowering an index and is explained later in this text.

Solution 19 Since I is invariant

$$I = {\overline{a}_\omega}^\lambda \overline{b}^\omega \overline{c}_\lambda = {a_\alpha}^\beta b^\alpha c_\beta \tag{97}$$

Inserting $\overline{b}^\omega = b^\alpha \frac{\partial \overline{u}^\omega}{\partial u^\alpha}$ and $\overline{c}_\lambda = c_\beta \frac{\partial u^\beta}{\partial \overline{u}^\lambda}$

$$I = {\overline{a}_\omega}^\lambda \left(b^\alpha \frac{\partial \overline{u}^\omega}{\partial u^\alpha} \right) \left(c_\beta \frac{\partial u^\beta}{\partial \overline{u}^\lambda} \right) = {a_\alpha}^\beta b^\alpha c_\beta \tag{98}$$

Since b^α and c_β are arbitrary it's possible to identify

$$a_\alpha{}^\beta = \overline{a}_\omega{}^\lambda \frac{\partial \overline{u}^\omega}{\partial u^\alpha} \frac{\partial u^\beta}{\partial \overline{u}^\lambda} \tag{99}$$

Multiply both sides by $\frac{\partial u^\alpha}{\partial \overline{u}^\mu} \frac{\partial \overline{u}^\nu}{\partial u^\beta}$ and sum with respect to both α and β from 1 to n

$$a_\alpha{}^\beta \frac{\partial u^\alpha}{\partial \overline{u}^\mu} \frac{\partial \overline{u}^\nu}{\partial u^\beta} = \overline{a}_\omega{}^\lambda \frac{\partial \overline{u}^\omega}{\partial u^\alpha} \frac{\partial u^\alpha}{\partial \overline{u}^\mu} \frac{\partial u^\beta}{\partial \overline{u}^\lambda} \frac{\partial \overline{u}^\nu}{\partial u^\beta} = \overline{a}_\omega{}^\lambda \delta^\omega_\mu \delta^\nu_\lambda \tag{100}$$

$$\overline{a}_\mu{}^\nu = a_\alpha{}^\beta \frac{\partial u^\alpha}{\partial \overline{u}^\mu} \frac{\partial \overline{u}^\nu}{\partial u^\beta} \tag{101}$$

If I is invariant under coordinate change then $a_\alpha{}^\beta$ defines a 2:nd order mixed tensor. It is covariant to 1:st order and contravariant to 1:st order.

Definition of a mixed tensor of 2:nd order defined on Euclidean n-space A mixed tensor of order 2, contravariant to order 1 and covariant to order 1, is a quantity whose n^2 components $a_\alpha{}^\beta$ are transformed according to

$$\overline{a}_\mu{}^\nu = a_\alpha{}^\beta \frac{\partial u^\alpha}{\partial \overline{u}^\mu} \frac{\partial \overline{u}^\nu}{\partial u^\beta} \tag{102}$$

under change of coordinate system in Euclidean n-space.

Continuing in the same way, to define a general order tensor, consider

$$I = c_{\nu_1 \nu_2 \cdots \nu_s}{}^{\alpha_1 \alpha_2 \cdots \alpha_r} a^{\nu_1} a^{\nu_2} \cdots a^{\nu_n} b_{\alpha_1} b_{\alpha_2} \cdots b_{\alpha_n} \tag{103}$$

If I is invariant under coordinate change, then $c_{\nu_1 \nu_2 \cdots \nu_s}{}^{\alpha_1 \alpha_2 \cdots \alpha_r}$ defines a mixed tensor, contravariant to order r and covariant to order s, and transforms as

$$\overline{c}_{\mu_1 \mu_2 \cdots \mu_s}{}^{\beta_1 \beta_2 \cdots \beta_r} = c_{\nu_1 \nu_2 \cdots \nu_s}{}^{\alpha_1 \alpha_2 \cdots \alpha_r} \frac{\partial u^{\nu_1}}{\partial \overline{u}^{\mu_1}} \frac{\partial u^{\nu_2}}{\partial \overline{u}^{\mu_2}} \cdots \frac{\partial u^{\nu_s}}{\partial \overline{u}^{\mu_s}} \frac{\partial \overline{u}^{\beta_1}}{\partial u^{\alpha_1}} \frac{\partial \overline{u}^{\beta_2}}{\partial u^{\alpha_2}} \cdots \frac{\partial \overline{u}^{\beta_r}}{\partial u^{\alpha_r}} \tag{104}$$

Definition of a general order mixed tensor defined on Euclidean n-space A mixed tensor, contravariant to order r and covariant to order s, is a quantity whose n^{r+s} components $c_{\nu_1 \nu_2 \cdots \nu_s}{}^{\alpha_1 \alpha_2 \cdots \alpha_r}$ are transformed according to

$$\overline{c}_{\mu_1 \mu_2 \cdots \mu_s}{}^{\beta_1 \beta_2 \cdots \beta_r} = c_{\nu_1 \nu_2 \cdots \nu_s}{}^{\alpha_1 \alpha_2 \cdots \alpha_r} \frac{\partial u^{\nu_1}}{\partial \overline{u}^{\mu_1}} \frac{\partial u^{\nu_2}}{\partial \overline{u}^{\mu_2}} \cdots \frac{\partial u^{\nu_s}}{\partial \overline{u}^{\mu_s}} \frac{\partial \overline{u}^{\beta_1}}{\partial u^{\alpha_1}} \frac{\partial \overline{u}^{\beta_2}}{\partial u^{\alpha_2}} \cdots \frac{\partial \overline{u}^{\beta_r}}{\partial u^{\alpha_r}} \tag{105}$$

Problem 20 What's the definition of a zero order tensor?

Solution 20 A zero order tensor is defined as a scalar. Why is this so? For example, a 2:nd order tensor has n^2 components, a 1:st order tensor has n^1 components, and thus a zero order tensor should have $n^0 = 1$ components.

Now that the tensor concept has been introduced, the rest of this text is devoted to explaining why tensors are written $c_{\nu_1 \nu_2 \cdots \nu_s}{}^{\alpha_1 \alpha_2 \cdots \alpha_r}$ and not $c^{\alpha_1 \alpha_2 \cdots \alpha_r}_{\nu_1 \nu_2 \cdots \nu_s}$.

4.1 The connection with vectors and matrices

Here it is convenient to give a first motivation to why a 2:nd order mixed tensor written as $a_\alpha{}^\beta$, instead of a^β_α. When 1:st order tensors are viewed as vectors and 2:nd order tensors are viewed as matrices, the convention is to let the first index α in $a_\alpha{}^\beta$ denote the row number and the second index β denote the column number in the matrix.

CONVENTION First order tensor: The index denotes the row in a column vector. Second order tensor: The first index denotes the row and the second index denotes the column in a matrix.

For example if ${a_\alpha}^\beta$ is defined on the Euclidean plane and ${a_1}^1 = 1$, ${a_1}^2 = 2$, ${a_2}^1 = 3$, ${a_2}^2 = 4$, then this can be written in matrix notation

$$({a_\alpha}^\beta) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \tag{106}$$

If b_β is defined on Euclidean 4-space and $b_1 = 1$, $b_2 = 3$, $b_3 = 5$, $b_4 = 7$ then this is written as a column matrix

$$(b_\beta) = \begin{pmatrix} 1 \\ 3 \\ 5 \\ 7 \end{pmatrix} \tag{107}$$

If the tensor notation would be a_α^β where α and β are right on top of each other, then with the above convention it would not be possible to decide which index denotes the row and column in the matrix. Maybe one could choose the convention so that the lower index always denotes the row and the upper index always denotes the column. But then, which index denotes the row in the tensor $a^{\alpha\beta}$?

4.2 The Kronecker delta and symmetric tensors

In equation (78) the Kronecker delta δ_μ^ν was only introduced as a symbol that takes the value 1 when $\mu = \nu$ and 0 otherwise. But the Kronecker delta is more than a symbol. In a given coordinate system u^α it is defined as

$$\delta_\mu^\nu = \frac{\partial u^\nu}{\partial u^\mu} \tag{108}$$

and is a mixed second order tensor, contravariant to order 1 and covariant to order 1. If the lower index μ denotes the row and ν the column then δ_μ^ν can be written

$$(\delta_\mu^\nu) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{109}$$

Notice that the indices over the Kronecker delta tensor is written on top of each other and not ${\delta_\mu}^\nu$ nor ${\delta^\nu}_\mu$. This is because one can change places of the indices $\delta_\mu^\nu = \delta_\nu^\mu$. A tensor with this property is called a symmetric tensor and it does not matter which index denotes the row and which index denotes the column. Since it is a mixed tensor it transforms as

$$\begin{aligned} \overline{\delta}_\mu^\nu &= \delta_\alpha^\beta \frac{\partial u^\alpha}{\partial \overline{u}^\nu} \frac{\partial \overline{u}^\mu}{\partial u^\beta} \\ &= \frac{\partial u^\alpha}{\partial \overline{u}^\nu} \frac{\partial \overline{u}^\mu}{\partial u^\alpha} \\ &= \frac{\partial \overline{u}^\mu}{\partial \overline{u}^\nu} \\ &= \delta_\nu^\mu \end{aligned} \tag{110}$$

Thus the Kronecker delta tensor has the same components in all coordinate systems. In Cartesian coordinate systems where there is no difference between contravariant and covariant components it is usually written with both indices subscript $\delta_{\mu\nu}$.

4.3 Matrix multiplication and contraction

When doing calculations with 1:st and 2:nd order tensors, it sometimes simplifies calculations to take advantage of matrix multiplication. The following problem illustrates this.

Problem 21 Let $A_{\alpha\beta}$ and $B^{\alpha\beta}$ be two 2:nd order tensors given by

$$(A_{\alpha\beta}) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \tag{111}$$

$$(B^{\alpha\beta}) = \begin{pmatrix} 4 & 3 \\ 2 & 1 \end{pmatrix} \tag{112}$$

Calculate $A_{\alpha\gamma}B^{\gamma\beta}$!

Solution 21 One way of solving this problem is to write out what the expression $A_{\alpha\gamma}B^{\gamma\beta}$ is short hand notation for and calculate each case individually:

$$\begin{aligned} &A_{11}B^{11} + A_{12}B^{21} \qquad (\alpha = 1, \beta = 1) \\ &A_{11}B^{12} + A_{12}B^{22} \qquad (\alpha = 1, \beta = 2) \\ &A_{21}B^{11} + A_{22}B^{21} \qquad (\alpha = 2, \beta = 1) \\ &A_{21}B^{12} + A_{22}B^{22} \qquad (\alpha = 2, \beta = 2) \end{aligned} \tag{113}$$

Here is an alternative solution is given. In this problem it's advantegous to make use of matrix multiplication. Assume there is a tensor $C^{\beta}_{\alpha} = A_{\alpha\gamma}B^{\gamma\beta}$, where the indices α and β in C^{β}_{α} are written right on top of each other because it's not yet decided which index denotes the row and column. Writing out what the expression $C^{\beta}_{\alpha} = A_{\alpha\gamma}B^{\gamma\beta}$ is short hand notation for

$$\begin{aligned} C^1_1 &= A_{11}B^{11} + A_{12}B^{21} \\ C^2_1 &= A_{11}B^{12} + A_{12}B^{22} \\ C^1_2 &= A_{21}B^{11} + A_{22}B^{21} \\ C^2_2 &= A_{21}B^{12} + A_{22}B^{22} \end{aligned} \tag{114}$$

Rearranging

$$\begin{aligned} C^1_1 &= A_{11}B^{11} + A_{12}B^{21} \\ C^1_2 &= A_{21}B^{11} + A_{22}B^{21} \\ C^2_1 &= A_{11}B^{12} + A_{12}B^{22} \\ C^2_2 &= A_{21}B^{12} + A_{22}B^{22} \end{aligned} \tag{115}$$

This can be written

$$\begin{pmatrix} C^1_1 \\ C^1_2 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B^{11} \\ B^{21} \end{pmatrix} \tag{116}$$

$$\begin{pmatrix} C_1^2 \\ C_2^2 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B^{12} \\ B^{22} \end{pmatrix} \tag{117}$$

Or even more compact

$$\begin{pmatrix} C_1^1 & C_1^2 \\ C_2^1 & C_2^2 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B^{11} & B^{12} \\ B^{21} & B^{22} \end{pmatrix} \tag{118}$$

Notice that in the matrix with the A:s and B:s, the α denotes the row and β the column in $A_{\alpha\beta}$ and $B^{\alpha\beta}$. Considering the matrix with the C:s, we let α denote the row and β the column in C_α^β and writes this from now on ${C_\alpha}^\beta$. Thus

$$\begin{pmatrix} {C_1}^1 & {C_1}^2 \\ {C_2}^1 & {C_2}^2 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B^{11} & B^{12} \\ B^{21} & B^{22} \end{pmatrix} \tag{119}$$

Or in tensor notation ${C_\alpha}^\beta = A_{\alpha\gamma}B^{\gamma\beta}$. Returning now to the original problem, where the solution is given by

$$\begin{aligned} ({C_\alpha}^\beta) &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 4 & 3 \\ 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 8 & 5 \\ 20 & 13 \end{pmatrix} \end{aligned} \tag{120}$$

Notice that the method used in this problem can be generalized. If instead

$$(A_{\alpha\beta}) = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \tag{121}$$

$$(B^{\alpha\beta}) = \begin{pmatrix} B^{11} & B^{12} & \cdots & B^{1n} \\ B^{21} & B^{22} & \cdots & B^{2n} \\ \vdots & \vdots & \vdots & \vdots \\ B^{n1} & B^{n2} & \cdots & B^{nn} \end{pmatrix} \tag{122}$$

then ${C_\alpha}^\beta = A_{\alpha\gamma}B^{\gamma\beta}$ is given by

$$({C_\alpha}^\beta) = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} B^{11} & B^{12} & \cdots & B^{1n} \\ B^{21} & B^{22} & \cdots & B^{2n} \\ \vdots & \vdots & \vdots & \vdots \\ B^{n1} & B^{n2} & \cdots & B^{nn} \end{pmatrix} \tag{123}$$

If we put two indices, a contravariant and a covariant, equal to each other in a mixed tensor and sum with respect to this pair of indices we obtain a tensor of order one less contravariant and one less covariant. This process is called *contraction* of the given tensor. Contraction of a mixed tensor of second order for example ${a_\alpha}^\beta$ gives a scalar

$${a_\alpha}^\alpha = {a_1}^1 + {a_2}^2 + \cdots + {a_n}^n \tag{124}$$

Rule to memorize If two 2:nd order tensors are multiplied with each other (for example $A_{\alpha\gamma}B^{\gamma\beta}$ or ${A^\alpha}_\gamma B^{\gamma\beta}$) and the column index of the tensor with matrix representation A is contracted with the row index of the tensor with matrix representation B then the resulting tensor can be calculated by matrix multiplication AB.

4.4 The metric tensor

Problem 22 Given the contravariant components a^β and the vectors $\mathbf{r}_\beta$, determine the covariant components a_α.

Solution 22 Using the definitions of contravariant and covariant components, equations (2) and (56) on pages 1 and 10, $a_\alpha = \mathbf{r}_\alpha \cdot \mathbf{v}$ and $\mathbf{v} = a^\beta \mathbf{r}_\beta$

$$a_\alpha = \mathbf{r}_\alpha \cdot \mathbf{v} = \mathbf{r}_\alpha \cdot a^\beta \mathbf{r}_\beta = \mathbf{r}_\alpha \cdot \mathbf{r}_\beta a^\beta \tag{125}$$

Here it is convenient to introduce the *metric tensor* $g_{\alpha\beta}$ which is defined as:

$$g_{\alpha\beta} = \mathbf{r}_\alpha \cdot \mathbf{r}_\beta \tag{126}$$

Thus

$$a_\alpha = g_{\alpha\beta} a^\beta \tag{127}$$

Problem 23 Show that the metric tensor $g_{\alpha\beta}$ is a 2:nd order covariant tensor.

Solution 23 According to the definition of the metric tensor $g_{\alpha\beta}$

$$\begin{aligned} \overline{g}_{\mu\nu} &= \mathbf{r}_{\overline{\mu}} \cdot \mathbf{r}_{\overline{\nu}} \\ &= \left(\mathbf{r}_\alpha \frac{\partial u^\alpha}{\partial \overline{u}^\mu}\right) \cdot \left(\mathbf{r}_\beta \frac{\partial u^\beta}{\partial \overline{u}^\nu}\right) \\ &= \mathbf{r}_\alpha \cdot \mathbf{r}_\beta \frac{\partial u^\alpha}{\partial \overline{u}^\mu} \frac{\partial u^\beta}{\partial \overline{u}^\nu} \\ &= g_{\alpha\beta} \frac{\partial u^\alpha}{\partial \overline{u}^\mu} \frac{\partial u^\beta}{\partial \overline{u}^\nu} \end{aligned} \tag{128}$$

which shows that the metric tensor $g_{\alpha\beta}$ is indeed a 2:nd order covariant tensor.

Problem 24 In problem 22 the covariant and contravariant components are related through the metric tensor as $a_\alpha = g_{\alpha\beta} a^\beta$. When memorizing this formula, how shall one remember that it is $a_\alpha = g_{\alpha\beta} a^\beta$ and not $a_\alpha = g_{\beta\alpha} a^\beta$?

Solution 24 The metric tensor is by definition a symmetric tensor $g_{\alpha\beta} = g_{\beta\alpha}$ and thus both expressions are correct.

Problem 25 Given the covariant components a_β and the vectors $\mathbf{r}_\beta$, determine the contravariant components a^α.

Solution 25 According to problem 22 the connection between contravariant and covariant components is given by $a_\alpha = g_{\alpha\beta} a^\beta$. Assuming we're working on Euclidean 2-space and since the column index β in $g_{\alpha\beta}$ is contracted with the row index β in a^β then $a_\alpha = g_{\alpha\beta} a^\beta$ can be written in matrix notation as

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \begin{pmatrix} a^1 \\ a^2 \end{pmatrix} \tag{129}$$

Problem 26 shows that the matrix $(g_{\alpha\beta})$ is invertible, and thus it's possible to multiply both sides with the inverse matrix $(g_{\alpha\beta})^{-1}$ which gives

$$\begin{pmatrix} a^1 \\ a^2 \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}^{-1} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \tag{130}$$

There is a tensor $g^{\alpha\beta}$ which is defined as the inverse of the metric tensor:

$$(g^{\alpha\beta}) = (g_{\alpha\beta})^{-1} \tag{131}$$

Here

$$\begin{aligned}(g^{\alpha\beta}) &= (g_{\alpha\beta})^{-1} \\ &= \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}^{-1} \\ &= \frac{1}{g_{11}g_{22} - g_{21}g_{12}} \begin{pmatrix} g_{22} & -g_{12} \\ -g_{21} & g_{11} \end{pmatrix}\end{aligned} \tag{132}$$

In the Differential geometry litterature one often finds the notation g for the determinant of the matrix $(g_{\alpha\beta})$. Here it means that $g = g_{11}g_{22} - g_{21}g_{12}$ and one can write

$$(g^{\alpha\beta}) = \frac{1}{g} \begin{pmatrix} g_{22} & -g_{12} \\ -g_{21} & g_{11} \end{pmatrix} \tag{133}$$

which means that $g^{11} = \frac{g_{22}}{g}$, $g^{12} = -\frac{g_{12}}{g}$, $g^{21} = -\frac{g_{21}}{g}$, $g^{22} = \frac{g_{11}}{g}$. With this notation it's possible to write

$$\begin{pmatrix} a^1 \\ a^2 \end{pmatrix} = \begin{pmatrix} g^{11} & g^{12} \\ g^{21} & g^{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \tag{134}$$

Writing this in tensor notation

$$a^\alpha = g^{\alpha\beta} a_\beta \tag{135}$$

Equation (135) is not only valid in Euclidean 2-space but also in Euclidean n-space.

Problem 26 On Euclidean n-space the vectors $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n$ are linearly independent vectors. Show that the $n \times n$-matrix $(g_{\alpha\beta})$ is invertible

$$(g_{\alpha\beta}) = \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix} \tag{136}$$

where $g_{\alpha\beta} = \mathbf{r}_\alpha \cdot \mathbf{r}_\beta$.

Solution 26 The proof is divided into two steps. The first step proofs the assertion when $g_{\alpha\beta}$ is defined on Euclidean 2-space. The second step proofs the assertion when $g_{\alpha\beta}$ is defined on Euclidean n-space.

Step 1: When defined on the Euclidean plane the metric 4 components g_{11}, g_{12}, g_{21}, g_{22} and this is written in matrix notation as:

$$(g_{\alpha\beta}) = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \tag{137}$$

The matrix $(g_{\alpha\beta})$ is invertible if and only if it's column vectors are linearly independent. Assuming that they are linearly dependent will lead to a contradiction. If they are linearly dependent, this means there is a real number $k \neq 0$ such that

$$\begin{pmatrix} g_{11} \\ g_{21} \end{pmatrix} = k \begin{pmatrix} g_{12} \\ g_{22} \end{pmatrix} \tag{138}$$

Notice that $g_{12} = \mathbf{r}_1 \cdot \mathbf{r}_2 = g_{21}$. Notice also that $g_{12} < \sqrt{g_{11}g_{22}}$. That this is so can be seen from the definition of scalar product.

$$\begin{aligned} g_{12} &= \mathbf{r}_1 \cdot \mathbf{r}_2 \\ &= |\mathbf{r}_1||\mathbf{r}_2| \cos\theta \end{aligned} \tag{139}$$

where $0 \leq \theta < \pi$ is the angle between the two vectors $\mathbf{r}_1$ and $\mathbf{r}_2$. Since $\mathbf{r}_1$ and $\mathbf{r}_2$ are linearly independent the angle $\theta \neq 0$ which means that $\cos\theta < 1$ and

$$\begin{aligned} g_{12} &< |\mathbf{r}_1||\mathbf{r}_2| \\ &= \sqrt{|\mathbf{r}_1|^2|\mathbf{r}_2|^2} \\ &= \sqrt{g_{11}g_{22}} \end{aligned} \tag{140}$$

Now using equations

$$\begin{aligned} g_{12} &< \sqrt{g_{11}g_{22}} \\ &= \sqrt{(kg_{12})\left(\frac{g_{21}}{k}\right)} \\ &= \sqrt{g_{12}g_{21}} \\ &= g_{12} \end{aligned} \tag{141}$$

Thus we arrive at the conclusion $g_{12} < g_{12}$ which is obviously not true. Thus the two column vectors in the matrix $(g_{\alpha\beta})$ are linearly independent and the invertibilty of $(g_{\alpha\beta})$ follows.

Step 2: Now to prove the general case when $g_{\alpha\beta}$ is defined on Euclidean n-space. If two column vectors in the matrix $(g_{\alpha\beta})$ are linearly dependent there is a real number $k \neq 0$ such that (here γ is an integer satisfying $1 \leq \gamma \leq n$)

$$\begin{matrix} \text{Row } 1 \\ \vdots \\ \text{Row } \gamma \\ \vdots \\ \text{Row n} \end{matrix} \qquad \begin{pmatrix} g_{1\alpha} \\ \vdots \\ g_{\gamma\alpha} \\ \vdots \\ g_{n\alpha} \end{pmatrix} = k \begin{pmatrix} g_{1\beta} \\ \vdots \\ g_{\gamma\beta} \\ \vdots \\ g_{n\beta} \end{pmatrix} \tag{142}$$

for arbitrary integers α and $\beta \neq \alpha$ such that $1 \leq \alpha, \beta \leq n$. Row α and row β means

$$g_{\alpha\alpha} = kg_{\alpha\beta} \tag{143}$$

$$g_{\beta\alpha} = kg_{\beta\beta} \tag{144}$$

Using $g_{\alpha\beta} = g_{\beta\alpha}$ and $g_{\alpha\beta} < \sqrt{g_{\alpha\alpha}g_{\beta\beta}}$

$$\begin{aligned} g_{\alpha\beta} &< \sqrt{g_{\alpha\alpha}g_{\beta\beta}} \\ &= \sqrt{(kg_{\alpha\beta})\left(\frac{g_{\beta\alpha}}{k}\right)} \\ &= \sqrt{g_{\alpha\beta}g_{\beta\alpha}} \\ &= g_{\alpha\beta} \end{aligned} \tag{145}$$

Thus we arrive at the contradiction $g_{\alpha\beta} < g_{\alpha\beta}$, which means that the column vectors in the matrix $(g_{\alpha\beta})$ must be linearly independent and the matrix $(g_{\alpha\beta})$ is therefore invertible.

Problem 27 Given the contravariant components a^α of a vector $\mathbf{v}$ and the vectors $\mathbf{r}_\alpha$, the vector $\mathbf{v}$ can be written

$$\mathbf{v} = a^\alpha \mathbf{r}_\alpha \tag{146}$$

But if the covariant components a_β of the vector $\mathbf{v}$ were given instead of the contravariant components a^α, the vector $\mathbf{v}$ is also specified and therefore there should exist vectors $\mathbf{r}^\beta$ such that it's possible to write

$$\mathbf{v} = a_\beta \mathbf{r}^\beta \tag{147}$$

Determine the vectors $\mathbf{r}^\beta$!

Solution 27 Since the vectors $\mathbf{r}_\beta$ are given, the metric tensor $g_{\alpha\beta} = \mathbf{r}_\alpha \cdot \mathbf{r}_\beta$ is also given. This in turn implies that the tensor $(g^{\alpha\beta}) = (g_{\alpha\beta})^{-1}$ also is given. The relationship between the contravariant components a^α and the covariant components a_β of a vector is given by

$$a^\alpha = g^{\alpha\beta} a_\beta \tag{148}$$

Assume that the contravariant components of vector $\mathbf{r}^\gamma$ is written $a^{\alpha\gamma}$ and the covariant components of this vector is written ${a_\beta}^\gamma$. This means

$$a^{\alpha\gamma} = g^{\alpha\beta} {a_\beta}^\gamma \tag{149}$$

Since $a^{\alpha\gamma}$ are the contravariant components it's possible to write

$$\mathbf{r}^\gamma = a^{\alpha\gamma} \mathbf{r}_\alpha \tag{150}$$

But

$$\mathbf{r}^\gamma = \delta^\gamma_\beta \mathbf{r}^\beta \tag{151}$$

Thus ${a_\beta}^\gamma = \delta^\gamma_\beta$

$$\begin{aligned} a^{\alpha\gamma} &= g^{\alpha\beta} {a_\beta}^\gamma \\ &= g^{\alpha\beta} \delta^\gamma_\beta \\ &= g^{\alpha\gamma} \end{aligned} \tag{152}$$

$$\mathbf{r}^\gamma = g^{\alpha\gamma} \mathbf{r}_\alpha \tag{153}$$

Problem 28 Since

$$g_{\alpha\beta} = \mathbf{r}_\alpha \cdot \mathbf{r}_\beta \tag{154}$$

one could suspect that

$$g^{\alpha\beta} = \mathbf{r}^\alpha \cdot \mathbf{r}^\beta \tag{155}$$

Show that this is so!

Solution 28 Starting with the expression $\mathbf{r}^\alpha \cdot \mathbf{r}^\beta$

$$\begin{aligned} \mathbf{r}^\alpha \cdot \mathbf{r}^\beta &= g^{\alpha\mu} \mathbf{r}_\mu \cdot g^{\beta\nu} \mathbf{r}_\nu \\ &= g^{\alpha\mu} g^{\beta\nu} g_{\mu\nu} \end{aligned} \tag{156}$$

Here we shall use $g^{\alpha\mu}g_{\mu\nu} = \delta^\alpha_\nu$, which is a relationship between $g^{\alpha\beta}$ and $g_{\alpha\beta}$ that is used a lot in tensor calculations. This follows from $(g^{\alpha\beta}) = (g_{\alpha\beta})^{-1}$, which means that

$$\begin{pmatrix} g^{11} & g^{12} & \cdots & g^{1n} \\ g^{21} & g^{22} & \cdots & g^{2n} \\ \vdots & \vdots & \vdots & \vdots \\ g^{n1} & g^{n2} & \cdots & g^{nn} \end{pmatrix} \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{157}$$

Continuing

$$\begin{aligned} \mathbf{r}^\alpha \cdot \mathbf{r}^\beta &= g^{\alpha\mu}g_{\mu\nu}g^{\beta\nu} \\ &= \delta^\alpha_\nu g^{\beta\nu} \\ &= g^{\beta\alpha} \end{aligned} \tag{158}$$

Since $(g_{\alpha\beta})$ is a symmetric matrix, the inverse matrix $(g^{\alpha\beta})$ is also a symmetric matrix which means $g^{\alpha\beta} = g^{\beta\alpha}$.

Problem 29 The vectors $\mathbf{r}_\beta$ are linearly independent. A natural question thus arises, are the vectors $\mathbf{r}^\gamma$ linearly independent?

Solution 29 Yes, they are linearly independent.

4.5 Raising and lowering an index

We have seen that

$$a^\alpha = g^{\alpha\beta}a_\beta \qquad \mathbf{r}^\alpha = g^{\alpha\beta}\mathbf{r}_\beta \tag{159}$$

$$a_\alpha = g_{\alpha\beta}a^\beta \qquad \mathbf{r}_\alpha = g_{\alpha\beta}\mathbf{r}^\beta \tag{160}$$

Here we show that generally ${a_\mu}^\nu \neq {a^\mu}_\nu$

$$\begin{aligned} {a_\mu}^\nu &= g_{\mu\alpha}a^{\alpha\nu} \\ &= g_{\mu\alpha}g^{\nu\beta}{a^\alpha}_\beta \\ &= (g_{\mu\alpha})\left(({a^\alpha}_\beta)(g^{\nu\beta})\right) \end{aligned} \tag{161}$$

$$(g_{\alpha\beta}) = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \tag{162}$$

$$(g^{\alpha\beta}) = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \tag{163}$$

$$({a^\alpha}_\beta) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \tag{164}$$

$$\begin{aligned} (a^{\alpha\beta}) &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix} \end{aligned} \tag{165}$$

$$\begin{aligned} ({a_\mu}^\nu) &= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 2 \\ 4 & 3 \end{pmatrix} \end{aligned} \tag{166}$$

4.6 Summary

Definition of a tensor of arbitrary order on Euclidean n-space

Let r, s and n be integers ≥ 1.

- A zero order tensor is defined as a scalar.
- A contravariant tensor of order r is a quantity whose n^r components $a^{\alpha_1\alpha_2\cdots\alpha_r}$ are transformed according to

$$\overline{a}^{\beta_1\beta_2\cdots\beta_r} = a^{\alpha_1\alpha_2\cdots\alpha_r}\frac{\partial\overline{u}^{\beta_1}}{\partial u^{\alpha_1}}\frac{\partial\overline{u}^{\beta_2}}{\partial u^{\alpha_2}}\cdots\frac{\partial\overline{u}^{\beta_r}}{\partial u^{\alpha_r}} \tag{167}$$

  under change of coordinate system in Euclidean n-space.

- A covariant tensor of order s is a quantity whose n^s components $b_{\nu_1\nu_2\cdots\nu_s}$ are transformed according to

$$\overline{b}_{\mu_1\mu_2\cdots\mu_s} = b_{\nu_1\nu_2\cdots\nu_s}\frac{\partial u^{\nu_1}}{\partial\overline{u}^{\mu_1}}\frac{\partial u^{\nu_2}}{\partial\overline{u}^{\mu_2}}\cdots\frac{\partial u^{\nu_s}}{\partial\overline{u}^{\mu_s}} \tag{168}$$

  under change of coordinate system in Euclidean n-space.

- A mixed tensor, contravariant to order r and covariant to order s, is a quantity whose n^{r+s} components $c_{\nu_1\nu_2\cdots\nu_s}{}^{\alpha_1\alpha_2\cdots\alpha_r}$ are transformed according to

$$\overline{c}_{\mu_1\mu_2\cdots\mu_s}{}^{\beta_1\beta_2\cdots\beta_r} = c_{\nu_1\nu_2\cdots\nu_s}{}^{\alpha_1\alpha_2\cdots\alpha_r}\frac{\partial u^{\nu_1}}{\partial\overline{u}^{\mu_1}}\frac{\partial u^{\nu_2}}{\partial\overline{u}^{\mu_2}}\cdots\frac{\partial u^{\nu_s}}{\partial\overline{u}^{\mu_s}}\frac{\partial\overline{u}^{\beta_1}}{\partial u^{\alpha_1}}\frac{\partial\overline{u}^{\beta_2}}{\partial u^{\alpha_2}}\cdots\frac{\partial\overline{u}^{\beta_r}}{\partial u^{\alpha_r}} \tag{169}$$