Per Amund Amundsen
Department of Mathematics and Natural Science
UNIVERSITY OF STAVANGER

Notes for FYS500 Classical Mechanics 9.11 2017

Additions and comments to *Classical Mechanics* by H. Goldstein & al. (3 ed. 2002).

Relativistic tensors

In the lecture notes for 23.10 we summarized some of the basic properties for tensors in Euclidean spaces. Here we generalize these result to more general vector spaces, like the Minkowski space-time of special relativity (see also *Goldstein* sect. (7.5)). We have introduced 4-vectors $x^\mu = [c^0 = ct, x^1, x^2, x^3] = [x^0, \mathbf{x}]$, which transforms with Lorentz-transformations (including rotations), as discussed in *Goldstein* sect. (7.2-3). Any set of 4 quantities u^μ transforming under coordinate transformations like x^μ will be called a *contravariant* 4-vector. These will always be written with an *upper index*, also called a *contravariant index.* This includes the 4-velocity, $v^\mu = \mathrm{d}x^\mu/\mathrm{d}\tau$ and the 4-momentum $p^\mu = mv^\mu$. These vectors satisfy the usual algebraic rules of a vector space, *i.e.* they can be added and multiplied by scalars to form new vectors of the same kind.

In the following we shall consider a more general vector space with contravariant vectors that transforms with some *group* of linear transformations that leave some interesting quantities, the *scalars*, invariant. A scalar f and a contravariant vector, v^μ, transforms as:

$$f' = f\,, \qquad v'^\mu = A^\mu{}_\nu\, v^\nu\,, \tag{0.56}$$

respectively, under a coordinate transformation $A = (A^\mu{}_\nu)$. Here we have introduced a modified version of the Einstein summation convention: *We sum over a pair of repeated indices if and only if one is upper and one is lower.* Thus, A is written with one upper and one lower index. Lower indices are also called *covariant indices.* Note in particular the unit matrix: $\mathbb{1}^\mu{}_\nu = \delta^\mu{}_\nu\ (= \delta_{\mu\nu})$. When transposing a matrix using this convention, upper indices remain upper and lower remain lower:

$$\widetilde{A} = \left(\widetilde{A}^\mu{}_\nu\right) = \left(A_\nu{}^\mu\right)\,. \tag{0.57}$$

An equation, or other relation, that retains its form under a coordinate transformation, is called *covariant.* A covariant homogenous linear relation between two vectors can be written as $w = Mv$ in matrix form or $w^\mu = M^\mu{}_\nu v^\nu$ in component form. Under a coordinate change, $v^\mu \to v'^\mu = A^\nu{}_\nu v^\nu$, and similar for $w \to w'$, the relation is transformed into $v'^\mu = M'^\mu_\nu w'^\nu$. Here M' is given by a similarity transformation, which in our new notation can be written:

$$M' = AMA^{-1} \qquad \Longleftrightarrow \qquad M'^\mu\nu = A^\mu{}_\rho M^\rho{}_\sigma\, A^{-1\sigma}{}_\nu = A^\mu{}_\rho \widetilde{A^{-1}}_\nu{}^\sigma M^\rho{}_\sigma\,. \tag{4.41a}$$

We say that M transforms as a *mixed tensor* of rank 2, with one upper and one lower index. Similar to the Euclidean case, this transformation law can be generalized to a tensor of rank $M+N$ or M, N; with M upper (contravariant) indices and N lower (covariant) indices:

$$T'^{\mu_1\cdots\mu_M}{}_{\nu_1\ldots\nu_N} = A^{\mu_1}{}_{\sigma_1}\cdots A^{\mu_M}{}_{\sigma_M}\widetilde{A^{-1}}_{\nu_1}{}^{\rho_1}\cdots\widetilde{A^{-1}}_{\nu_N}{}^{\rho_N}\, T_{\sigma_1\ldots\sigma_M}{}^{\rho_1\cdots\rho_N} \tag{5.10b}$$

Note that upper and lower indices can occur in any order, and the order is important, so *e.g.* $T^{\mu\nu}{}_\lambda \neq T^\mu{}_\lambda{}^\nu$.

Products of matrices satisfy $\widetilde{AB} = \widetilde{B}\widetilde{A}$ and $(AB)^{-1} = B^{-1}A^{-1}$, and, as one easily checks, $\widetilde{A^{-1}} = \widetilde{A}^{-1}$, which means that $\widetilde{(AB)}^{-1} = \widetilde{A^{-1}}\widetilde{B^{-1}}$. Thus $\widetilde{A^{-1}}$, and A satisfy the same group multiplication rules, but they are only identical for orthogonal transformations. For other groups of transformations, we have to distinguish carefully between upper and lower indices.

The transformation matrices satisfy:

$$A^\lambda{}_\mu\widetilde{A^{-1}}_\lambda{}^\nu = A^{-1\nu}{}_\lambda A^\lambda{}_\mu = \mathbb{1}^\nu{}_\mu = \delta^\nu{}_\mu = A^\mu{}_\lambda\widetilde{A^{-1}}_\nu{}^\lambda\,, \tag{0.58}$$

where the last identity follows by transposition. Using this it is easily shown that a *contraction* of an upper and a lower index of a tensor of rank M in the contravariant indices and rank N in the covariant ones yields a tensor of rank $M-1, N-1$. One can contract any upper index with any lower. We show this for a contraction of the last index pair, the proof for any other pair is the same:

$$\begin{aligned}
T'^{\mu_1\ldots\mu_{M-1}\lambda}{}_{\nu_1\ldots\nu_{N-1}\lambda} &= A^{\mu_1}{}_{\sigma_1}\cdots A^{\mu_{M-1}}{}_{\sigma_{M-1}} A^{\lambda}{}_{\sigma_M} \widetilde{A^{-1}}{}_{\nu_1}^{\rho_1}\cdots\widetilde{A^{-1}}{}_{\nu_{N-1}}^{\rho_{N-1}}\widetilde{A^{-1}}{}_{\lambda}^{\rho_N} T_{\sigma_1\ldots\sigma_M}{}^{\rho_1\ldots\rho_N} \\
&= \delta^{\rho_M}{}_{\sigma_M} A^{\mu_1}{}_{\sigma_1}\cdots A^{\mu_{M-1}}{}_{\sigma_{M-1}}\widetilde{A^{-1}}{}_{\nu_1}^{\rho_1}\cdots\widetilde{A^{-1}}{}_{\nu_{N-1}}^{\rho_{N-1}} T_{\sigma_1\ldots\sigma_{M-1}\sigma_M}{}^{\rho_1\cdots\rho_{N-1}\rho_N}\,, \\
&= A^{\mu_1}{}_{\sigma_1}\cdots A^{\mu_{M-1}}{}_{\sigma_{M-1}}\widetilde{A^{-1}}{}_{\nu_1}^{\rho_1}\cdots\widetilde{A^{-1}}{}_{\nu_{N-1}}^{\rho_{N-1}} T^{\sigma_1\ldots\sigma_{M-1}\lambda}{}_{\rho_1\ldots\rho_{N-1}\lambda}\,, \qquad (0.51a)
\end{aligned}$$

which is the correct transformation law. In a similar manner we can also generalize the result of eq. (0.52) from the lectures of 23.10 to mixed tensors. A particularly useful contraction is the *trace* of a mixed tensor of rank 1,1, which is just the trace of the corresponding matrix: $\mathrm{Tr}\,(M^{\mu}{}_{\nu}) = M^{\mu}{}_{\mu}$, which is a scalar.

A tensor of rank 1 with only a lower index is called a *covariant vector*. These form a vector space of their own, which is isomorphous to the space spanned by the set of contravariant vectors, since they have the same number of component, say D, and all real vector spaces of dimension D are isomorphous. But we cannot identify the two types of vectors in the general case, because they transform differently under coordinate transformations. But if we restrict ourself to orthogonal transformations, *i.e.* rotations and mirrorings, then $A^{-1} = \widetilde{A}$, and there is no difference between the transformation properties of co- and contravariant vectors, and we may identify the two. More generally, we can identify co- and contravariant tensor components in this case. This leads to the formalism for Euclidean tensors discussed in the notes for 23.10.

As in the case of Euclidean tensors, we can construct tensors of arbitrary rank by taking direct products of vectors. Thus if $v_1 \ldots v_M$ are M contravariant vectors and $w_1 \ldots w_N$ are N covariant ones, we have:

$$T^{\mu_1\cdots\mu_N}{}_{\nu_1\ldots\nu_N} = v_1^{\mu_1}\cdots v_M^{\mu_N}\, w_{1\nu_1}\cdots w_{N\nu_N}\,, \qquad (0.47a)$$

which is seen to transform correctly. In the same manner we can define the direct product of tensors, producing new tensors. We can, of course, form contractions of such direct products. Particularly important is the contraction of a covariant and a contravariant vector, v^μ and w_μ, which yields the scalar $v^\mu w_\mu$. This we can call a *scalar product*, but note that we so far have no prescription for forming a scalar product of two contravariant, or two covariant, vectors.

An important class of covariant vectors is formed from *scalar fields*, *i.e.* scalars which take different values at different points, so they depend on the coordinates x^μ. If $f(x)$ is such a scalar field, we can always calculate the *gradient* of f, with components written $\partial_\mu f = f_{,\mu}$, like in Euclidean space:

$$\partial_\mu f = \frac{\partial f}{\partial x^\mu}\,. \qquad (0.59)$$

It is easy to check that $\partial_\mu f$ really transforms as a covariant vector, *i.e.* with a lower index, because from the transformation law for x we have:

$$x^\mu = A^{-1}{}^{\nu}{}_{\nu}\, x'^{\nu} \qquad \Longleftrightarrow \qquad \frac{\partial x^\mu}{\partial x'^\nu} = A^{-1}{}^{\mu}{}_{\nu}\,. \qquad (0.60)$$

so from eq. (0.58) and the chain rule:

$$\partial'_\mu f = \frac{\partial f}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial f}{\partial x^\nu} = A^{-1}{}^{\nu}{}_{\mu}\,\partial_\nu f = \widetilde{A^{-1}}{}_{\mu}^{\nu}\,\partial_\nu f\,, \qquad (0.61)$$

which is the correct transformation law for a tensor with one lower index, *i.e.* a covariant vector. With this notation, we can write the differential of f simply as:

$$\mathrm{d}f = \frac{\partial f}{\partial x'^\mu}\mathrm{d}x^\mu = \partial_\mu f\,\mathrm{d}x^\mu \qquad (0.62)$$

This we can read as the contraction of the contravariant vector $\mathrm{d}x^\mu$, which transforms like x^μ, and $\partial_\mu f$, which produces a scalar. Note that there is no minus sign in this formula, even if A is a Lorentz transformation, in which case we have:

$$\mathrm{d}f = \frac{\partial f}{\partial x'^\mu}\mathrm{d}x^\mu = \partial_\mu f\, \mathrm{d}x^\mu = \partial_0 f\, \mathrm{d}x^0 + \boldsymbol{\nabla} f \cdot \mathrm{d}\mathbf{x} = \partial_t f\, \mathrm{d}t + \boldsymbol{\nabla} f \cdot \mathrm{d}\mathbf{x}\,. \tag{0.62a}$$

Eq. (0.62) can be interpreted as an expansion of the *differential form* $\mathrm{d}f$ in terms of the *basic forms* $\{\mathrm{d}x^\mu\}$, with $\partial_\mu f$ as the components of $\mathrm{d}f$ in this basis. Because of this interpretation, covariant vectors are also called *differential forms*, and $\mathrm{d}f$ is interpreted as the coordinate-free way of writing this vector, just as we may write $x = x^\mu\,\mathbf{e}_\mu$ for "normal" contravariant vectors. Note that, since we have so far introduced no concept of a *length* of a vector, there is no assumption that $\mathrm{d}f$ or $\mathrm{d}x^\mu$ are "small" involved in this notation. But we can of course still use eq. (0.62) to calculate the rate of change of f along some parameterized path $x^\mu(\lambda)$ in our vector space from the chain rule:

$$\frac{\mathrm{d}f}{\mathrm{d}\lambda} = \partial_\mu f\,\frac{\partial x^\mu}{\partial\lambda}\,. \tag{0.62b}$$

In particular, for a time-like trajectory in Minkowski space we can use the proper time, τ, as parameter, so we have the rate of variation of f along the trajectory $x^\mu(\tau)$ as:

$$\frac{\mathrm{d}f}{\mathrm{d}\tau} = \frac{\partial f}{\partial x^\mu}\,\frac{\partial x^\mu}{\partial\tau} = \partial_\mu f\, u^\mu = \partial_0 f\, u^0 + \boldsymbol{\nabla} f \cdot \mathbf{u} = \gamma[\partial_t f + \boldsymbol{\nabla} f \cdot \mathbf{v}]\,, \tag{0.62c}$$

where $u^\mu = \mathrm{d}x^\mu/\mathrm{d}\tau = \gamma[c, \mathbf{v}]$ is the 4-velocity. Note that there is no minus sign in this formula either, and that it can be interpreted as the usual substantial derivative multiplied by the time dilatation factor γ. As in the Euclidean case, we can generalize eq. (0.62a) to define the *directional derivative* of f along an arbitrary vector $v = [v^\mu]$:

$$\partial_v f = \partial_\mu f\, v^\mu\,. \tag{7.41}$$

One can, of course, also take partial derivatives of tensor fields, *i.e.* tensors that depend on the coordinates x^μ. This operation increases the covariant rank of a tensor by one, as is easily verified. In particular, if $w^\mu(x)$ is a (contravariant) vector field, $\partial_\mu w^\nu$ is a mixed tensor of rank 2. If we contract the indices of this, we obtain the *divergence* of w as the scalar:

$$\mathrm{Div}\ w = \partial_\mu w^\mu\,. \tag{0.63}$$

In Minkowski space this leads to: $\mathrm{Div}\ w = \partial_0\, w^0 + \boldsymbol{\nabla}\cdot\mathbf{w}$, again without a minus sign. In particular, $\mathrm{Div}\ x = \partial_\mu x^\mu = 4$.

So far, we have worked without introducing the length of a vector. Thus the formulation works for arbitrary linear coordinate transformations, including those that change the length of a vector, like scale transformations. But just like Euclidean space, Minkowski space is endowed with an additional structure, namely a *scalar product* for contravariant vectors. It is then useful to restrict the definition of tensors to transformations that leave the scalar product invariant. In Minkowski space these are the Lorentz transformations, including rotations and mirrorings. Scalar products are generally defined by an expression quadratic in the coordinate vector, x. We therefore consider coordinate transformations that leave a fixed quadratic form invariant. This is defined by a non-singular symmetric matrix, called the *metric* or the *metric tensor*, $g = \widetilde{g} = (g_{\mu\nu})$:

$$x\cdot y = \widetilde{x}gy = x^\mu g_{\mu\nu}\, y^\nu\,, \tag{0.64}$$

The important point is that this form is required to be *invariant*, *i.e.* it is the same in all coordinate systems. In Minkowski space we have $g = \mathrm{Diag}[1, -1, -1, -1] = g^{-1}$. For the scalar product to be invariant, one must have:

$$x'\cdot y' = \widetilde{x}\widetilde{A}\, g\, Ay = \widetilde{x}gy = x\cdot y \qquad \Longleftrightarrow \qquad \widetilde{A}\, g\, A = g \qquad \Longleftrightarrow \qquad g = \widetilde{A^{-1}}gA^{-1}\,, \tag{0.65}$$

for all coordinate transformations A preserving the scalar product. Multiplying this equation from the right with A^{-1}, one finds, using the symmetry, $\widetilde{g} = g \to \widetilde{g^{-1}} = g^{-1}$:

$$g^{-1}\widetilde{A}\,g\,A = \mathbb{1} \qquad \Longleftrightarrow \qquad A^{-1} = g^{-1}\widetilde{A}\,g \qquad \Longleftrightarrow \qquad \widetilde{A^{-1}} = gAg^{-1}\,. \tag{0.66}$$

From eq. (0.65) it follows that $g_{\mu\nu}$ is both invariant and transforms as a covariant tensor of rank 2. ($\widetilde{A^{-1}} = \widetilde{A}^{-1}$):

$$g = \widetilde{A^{-1}}gA^{-1} \qquad \Longleftrightarrow \qquad g_{\mu\nu} = \widetilde{A^{-1}}_{\mu}{}^{\sigma}\,\widetilde{A^{-1}}_{\nu}{}^{\rho}\,g_{\sigma\rho}\,, \tag{0.67}$$

One now *defines* the covariant components of any vector w by:

$$x_\mu = g_{\mu\nu}w^\nu\,. \tag{0.68}$$

In particular for Minkowski space $x_0 = x^0 = ct$, $x_i = -x^i$, so $x_\mu = [ct, -\mathbf{x}]$. It is easily checked that the components of x_μ forms a covariant vector. Using that from eq. (0.65) $gA = \widetilde{A^{-1}}g$, one finds:

$$w'_\mu = g_{\mu\nu}w'^\nu = g_{\mu\nu}A^\nu{}_\lambda\,w^\lambda = \widetilde{A^{-1}}_{\mu}{}^{\nu}\,g_{\nu\lambda}\,w^\lambda = \widetilde{A^{-1}}_{\mu}{}^{\nu}\,w_\nu\,. \tag{0.69}$$

From eq. (0.67) we see that $g^{-1} = (g^{\mu\nu}) = (g^{\nu\mu})$ transforms as an invariant covariant tensor of rank 2:

$$g^{-1} = Ag^{-1}\widetilde{A} \qquad \Longleftrightarrow \qquad g^{\mu\nu} = A^\mu{}_\sigma A^\nu{}_\rho g^{\sigma\rho} \tag{0.67a}$$

Furthermore:

$$g^{\mu\lambda}g_{\lambda\nu} = g^{\mu\lambda}g_{\nu\lambda} = g_{\mu\lambda}g^{\lambda\nu} = g_{\mu\lambda}g^{\mu\lambda} = \mathbb{1}^\mu{}_\nu = \mathbb{1}_\mu{}^\nu = \delta^\mu{}_\nu\,. \tag{0.70}$$

Multiplying eq. (0.68) with g^{-1}, we find:

$$w^\mu = g^{\mu\nu}w_\nu\,. \tag{0.68a}$$

Thus we say that $g_{\mu\nu}$ *lowers* the index of a contravariant vector, while $g^{\mu\nu}$ *raises* the index of a covariant one. By the same procedure one easily shows that $g_{\mu\nu}$ can be used to lower any contravariant tensor index, converting it to a covariant one. Similarly, using that from eq. (0.66) it follows that $g^{-1}\widetilde{A^{-1}} = Ag^{-1}$, one finds that $g^{\mu\nu}$ can be used to raise any covariant index. Note that when rising and lowering indices, their order must remain unchanged. As an example, if $T^{\mu\nu}{}_{\sigma\rho}$ is a tensor, we can form other tensors like:

$$T^\mu{}_{\lambda\sigma\rho} = g_{\lambda\sigma}T^{\mu\nu}{}_{\sigma\rho}\,, \qquad T^{\mu\nu\lambda}{}_\sigma = g^{\lambda\sigma}T^{\mu\nu}{}_{\sigma\rho}\,, \qquad etc.$$

We can, of course, rise and lower several indices at the same time, and rising and lowering performed in different orders commute. In particular, we can rise an index of $g_{\mu\nu}$, or lower one of $g^{\mu\nu}$. Using eq. (0,70) we find:

$$g^\mu{}_\nu = g^{\mu\lambda}g_{\lambda\nu} = \delta^\mu{}_\nu\,. \tag{0.71}$$

Thus the mixed metric tensor is just nothing but the Kronecker symbol, or the unit matrix in matrix notation.