SRT Math: Contra- and Covariant Vectors ★ ref. Jackson, Sec. 11.6

12) We need one more trick to handle 4-vectors with (moderate) impunity, viz. the distinction between "**CONTRAVARIANT**" and "**COVARIANT**" vectors. The distinction relates to a difference in the way in which what we have calling "vectors" can behave under a coordinate transform. This difference is now illustrated.

1. Under ordinary rotation of space coordinates, the 3-vector position changes as:

$\mathbf{r} \rightarrow \mathbf{r}' = \underset{\sim}{R}\,\mathbf{r}$, w/ $\underset{\sim}{R} = (R_{ij})$ the rotation matrix, e.g. // $\underset{\sim}{R} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$ rotn by $\measuredangle\theta$ about z-axis.

i.e., $x_i \rightarrow x_i' = R_{ij}\,x_j$ { for compts of $\mathbf{r}$ (summation convention). (32)

NOTE: $\dfrac{\partial x_i'}{\partial x_k} = R_{ij}\left(\dfrac{\partial x_j}{\partial x_k}\right) = R_{ik}$ { because $\partial x_j/\partial x_k = \delta_{jk}$ (for orthogonal cds x_i)

... the rotation $\boxed{R_{ij} = \dfrac{\partial x_i'}{\partial x_j}}$ is defined by the diffl cd. changes. (33)

Any 3-comp. obect $\mathbb{A} = (A_i)$ is a "3-vector" if it transforms like $\mathbf{r}$:

$$A_i \rightarrow A_i' = R_{ij}\,A_j = (\partial x_i'/\partial x_j)\,A_j \text{, for a 3-vector } \mathbb{A} = (A_i) \text{ [analogous to } \mathbf{r}\text{].} \quad (34)$$

Such a vector has an invariant length under the transfn $\underset{\sim}{R}$, i.e.

$$\rightarrow A_i'^2 = \left[\left(\frac{\partial x_i'}{\partial x_j}\right)A_j\right]\left[\left(\frac{\partial x_i'}{\partial x_k}\right)A_k\right] = A_k^2 \Rightarrow \left(\frac{\partial x_i'}{\partial x_j}\right)\left(\frac{\partial x_i'}{\partial x_k}\right) = \delta_{jk}. \quad (35)$$

This kind of $\mathbb{A} = (A_i)$, transforming per Eq.(34), is a **CONTRAVARIANT** 3-vector.

---

2. Eq.(34) is not the only transfn law possible for vectors, however. E.g.

$$\left[\begin{array}{l} i^{\text{th}} \text{ comp. of} \\ \nabla \text{ operator} \end{array}\right\}\; \nabla_i = \frac{\partial}{\partial x_i} \rightarrow \nabla_i' = \frac{\partial}{\partial x_i'} = \left(\frac{\partial x_j}{\partial x_i'}\right)\frac{\partial}{\partial x_j} = \underbrace{(\partial x_j/\partial x_i')}_{\text{Compare with Eq.(34)}}\nabla_j. \quad (36)$$

Comparison with Eq. (34) shows the transfn coefficients are "upside-down", i.e. here we get $(\partial x_j/\partial x_i')$ rather than above $(\partial x_i'/\partial x_j)$. Even though

---

★ Ref. G. Arfken "Mathl Methods for Physicists" (Acad. Press, 3rd ed., 1985), Chap. 3.

$\nabla = (\partial/\partial x_i)$ transforms differently than $\mathbf{r} = (x_i)$, it is still a vector in that its length $(\partial/\partial x_i)^2$ is invariant under cd. rotation.† So we can define 3-comp. objects $\mathbb{B} = (B_i)$ to be vectors also if they transform like $\partial/\partial\mathbf{r}$:

$$\underline{\underline{B_i \rightarrow B_i' = (\partial x_j/\partial x_i')\, B_j}}, \quad \text{for a 3-vector } \mathbb{B} = (B_i) \;\; \underline{[\text{analogous to } \partial/\partial\mathbf{r}]}. \tag{37}$$

The length $B_i'^2$ is invariant, and this kind of vector is a **COVARIANT** 3-vector.

---

3. To distinguish between the vector transforms just cited, we define a notation:

CONTRAVARIANT VECTOR — write comps. as superscripts:

$$A^\alpha \rightarrow A'^\alpha = \left(\frac{\partial x'^\alpha}{\partial x^\beta}\right) A^\beta \quad \text{prototype is position } \mathbf{r} = (x^\alpha). \tag{38a}$$

COVARIANT VECTOR — write comps. as subscripts:

$$B_\alpha \rightarrow B_\alpha' = \left(\frac{\partial x^\beta}{\partial x'^\alpha}\right) B_\beta \quad \text{prototype is gradient } \partial/\partial\mathbf{r}. \tag{38b}$$

In cartesian cds (x,y,z), there is no distinction between contra- and covariant vectors, because: $(\partial x^\beta/\partial x'^\alpha) \equiv (\partial x'^\alpha/\partial x^\beta)$.¶ **BUT**, in curvilinear cds $(\partial x^\beta/\partial x'^\alpha) \neq (\partial x'^\alpha/\partial x^\beta)$ in general, and the distinction is real.

---

13) The contra-covariant distinction does not affect scalar fields... a scalar transforms into itself (by definition) no matter what cd. system you designate. We have shown above how the distinction works for vector fields (tensor rank one). Next on the list is tensors of rank two (matrices), which we will encounter. We can define three distinct types of 2nd rank tensors by their transf$^{\text{n}}$s, viz.

$$\text{CONTRAVARIANT}: \; F'^{\alpha\beta} = (\partial x'^\alpha/\partial x^\gamma)(\partial x'^\beta/\partial x^\epsilon)\, F^{\gamma\epsilon}; \tag{39a}$$

---

† As an exercise, show $(\partial/\partial x_i)^2$ is in fact invariant under 3D rotation $\underset{\sim}{R}$.

¶ Follows from property of rotation matrix: $\underset{\sim}{R}^{-1}$ (inverse) $= \underset{\sim}{R}^{T}$ (transpose), i.e. $R^{-1}_{ij} = R_{ji}$.

COVARIANT : $G'_{\alpha\beta} = (\partial x^\gamma/\partial x'^\alpha)(\partial x^\epsilon/\partial x'^\beta)\, G_{\gamma\epsilon}$; (39b)

MIXED : $H'^\alpha_\beta = (\partial x'^\alpha/\partial x^\gamma)(\partial x^\epsilon/\partial x'^\beta)\, H^\gamma_\epsilon$. (39c)

In these defns, the summation convention is in force (sum over repeated indices).

In ordinary vector analysis, we formed the scalar product $\mathbb{A}\cdot\mathbb{B} = A_i B_i$ by summing over a repeated index -- this reduces the tensor rank of $\mathbb{A}$ & $\mathbb{B}$ from one & one (vector) to zero (scalar). For 2nd rank tensors like Eqs (39), summation over a repeated index can be done within the tensor itself, i.e. we can form H'^α_α. For such tensors, summing over a repeated is known as "Contraction".

Let us "contract" the mixed tensor H'^α_β in (39c). I.e. we form...

$$\rightarrow H'^\alpha_\alpha = \cancel{\left(\frac{\partial x'^\alpha}{\partial x^\gamma}\right)\left(\frac{\partial x^\epsilon}{\partial x'^\alpha}\right)}^{\,1} H^\gamma_\epsilon = \left(\frac{\partial x^\epsilon}{\partial x^\gamma}\right) H^\gamma_\epsilon = \underbrace{\delta^\epsilon_\gamma}_{\text{Kronecker}} H^\gamma_\epsilon = H^\gamma_\gamma. \tag{40}$$

this $\equiv \mathrm{Tr}\,\underset{\sim}{H}'$; $\mathrm{Tr}\,\underset{\sim}{H}$ is invariant

Such a contraction (over one tensor or a product of such tensors) reduces the rank by two (here 2nd rank → 0th rank scalar). Anyway, any mixed tensor H^γ_ϵ always has $\mathrm{Tr}\,\underset{\sim}{H}$ = invariant under any cd transfn $x^\epsilon \rightarrow x'^\epsilon$.

---

14) We can form a 2nd rank tensor from two 1st rank tensors (vectors) by what is called a "direct product". In ordinary vector analysis, we do

$$\rightarrow \underset{\sim}{T} = \mathbb{B}\otimes\mathbb{A} = (B_1, B_2, B_3)\otimes(A_1, A_2, A_3) = \begin{pmatrix} B_1A_1 & B_1A_2 & B_1A_3 \\ B_2A_1 & B_2A_2 & B_2A_3 \\ B_3A_1 & B_3A_2 & B_3A_3 \end{pmatrix} \text{ i.e. } T_{ij} = B_iA_j \tag{41}$$

Notice that: $\mathrm{Tr}\,\underset{\sim}{T} = B_iA_i = \mathbb{B}\cdot\mathbb{A}$.

The same sort of construction works for the tensors in Eq.(39). For example we can construct a mixed tensor $H^\alpha_\beta = B_\beta A^\alpha$ from the direct product of a covariant vector (B_β) and contravariant vector (A^α). That H_β^α is in fact a qualified 2nd rank mixed tensor can be checked by how it transforms:

$$\rightarrow H'^\alpha_\beta = B'_\beta A'^\alpha = [(\partial x^\epsilon/\partial x'^\beta)B_\epsilon][(\partial x'^\alpha/\partial x^\gamma)A^\gamma] = \left(\frac{\partial x'^\alpha}{\partial x^\gamma}\right)\left(\frac{\partial x^\epsilon}{\partial x'^\beta}\right)H_\epsilon^\gamma, \quad (42)$$

w/ $H_\epsilon^\gamma = B_\epsilon A^\gamma$. This exactly matches the defⁿ in (39c). So $B_\beta A^\alpha$ is OK.

For $H_\beta^\alpha = B_\beta A^\alpha$, the contraction H'^α_α yields an invariant, per Eq. (40), i.e.

$$\boxed{B'_\alpha A'^\alpha \equiv Tr(B_\beta A^\alpha) \equiv B_\gamma A^\gamma}, \text{ invariant.} \quad (43)$$

This result suggests how we should define the vector scalar product in the present notational scheme, viz.

$$\left[\begin{array}{l}\text{SCALAR}\\ \text{PRODUCT}\end{array}\right\} \; \underline{\underline{B\cdot A = B_\alpha A^\alpha}} = \text{invariant.} \quad \int \text{Note: 1st vector is covariant and 2nd vector is contravariant, always.} \quad (44)$$

This is the required scalar invariant under cd transforms $x^\epsilon \rightarrow x'^\epsilon$.

N.B. "Contraction" over an index (like α in Eq. (44)) always means summation on a repeated index which appears once as a contravariant superscript and once as a covariant subscript. Never contract super-super, or sub-sub.¶

15) The above general results apply to our Lorentz 4-vector formalism as follows:

$$\left[\begin{array}{l}\text{invariant spacetime}\\ \text{interval}\end{array}\right\} (ds)^2 = (dx^0)^2 - (dx^k)^2 = g_{\alpha\beta}\, dx^\alpha dx^\beta \quad \int \underline{\alpha \,\&\, \beta = 0,1,2,3}$$

$$\text{where: } g_{\alpha\beta} = g_{\beta\alpha}, \text{ is the covariant "}\underline{\text{metric tensor}}\text{": } (g_{\alpha\beta}) = \begin{pmatrix} 1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix}. \quad (45)$$

This is for "flat" (Minkowski) spacetime. In the "curved" spacetime of

¶ Reason: a contraction like $G_{\alpha\alpha}' = (\partial x^\gamma/\partial x'^\alpha)(\partial x^\epsilon/\partial x'^\alpha)G_{\gamma\epsilon}$, for the covariant $\underset{\sim}{G}$ in Eq. (39b) produces a zero rank tensor which is non-invariant gibberish.

general relativity, $\underset{\sim}{g}$ has off-diagonal entries, and $g_{\alpha\beta}dx^\alpha dx^\beta$ has cross-terms.

For "flat" space, useful properties of $\underset{\sim}{g}$ are...

1. Contravariant form: $\underline{g^{\alpha\beta} \equiv g_{\alpha\beta}}$. Identity: $\underline{g_{\alpha\gamma}\, g^{\gamma\beta} = \delta_\alpha^\beta}$ ← Kronecker delta. (46)

2. Contraction over $\underset{\sim}{g}$ index
$$\begin{cases} g_{\alpha\beta}A^\beta = A_\alpha \leftarrow \text{converts contravariant } A^\beta \text{ to covariant } A_\alpha; \\ g^{\alpha\beta}B_\beta = B^\alpha \leftarrow \text{ " covariant } B_\beta \text{ to contravariant } B^\alpha. \end{cases} \quad (47)$$

What happens here is: if $\underline{\tilde{A}^\beta = (A^0, \mathbb{A})}$ is contravariant, and is made covariant by: $\tilde{A}^\beta \to \tilde{A}_\beta = g_{\beta\alpha}\tilde{A}^\alpha$, then $\underline{\tilde{A}_\beta = (A^0, -\mathbb{A})}$.

3. The ~~vector~~ scalar product of Lorentz 4-vectors $\tilde{B}$ & $\tilde{A}$ is defined, per Eq.(44), as

$$\boxed{\tilde{B}\cdot\tilde{A} = \tilde{B}_\alpha \tilde{A}^\alpha = (B^0, -\mathbb{B})\cdot(A^0, \mathbb{A}) = B^0A^0 - \mathbb{B}\cdot\mathbb{A}}. \quad (48)$$

This will be Lorentz invariant, just as $(ds)^2 = (dx^0)^2 - (dx^k)^2$. **NOTICE**: with the metric tensor $(g_{\alpha\beta})$ defining the contra-covariant relations in (47) [and thereby introducing the required (−) sign] we do not need to include $\underset{\sim}{g}$ explicitly in the def^n of $\tilde{B}\cdot\tilde{A}$, as on pp. SRT 16-17. A notational advantage!

---

16) As an indication of coming attractions, we construct a 4-vector version of ∇. We already know that $\nabla = \partial/\partial\mathbf{r}$ is a covariant 3-vector. Also $\partial/\partial x^0$ ($\mathcal{W}\, x^0 = ct$) is covariant under a Lorentz transform, so we can construct a 4-vector operator:

$$\left.\begin{matrix}\underline{\text{COVARIANT del}} \\ \text{(Jackson's } \partial_\alpha)\end{matrix}\right\} \nabla_\alpha = \frac{\partial}{\partial x^\alpha} = \left(\frac{\partial}{\partial x^0}, \nabla\right) = \underline{\underline{\partial_\alpha}} \quad \Bigg\| \quad \left.\begin{matrix}\text{CONTRA-} \\ \underline{\text{VARIANT del}} \\ \text{(Jackson's } \partial^\alpha)\end{matrix}\right\} \nabla^\alpha = \frac{\partial}{\partial x_\alpha} = \left(\frac{\partial}{\partial x^0}, -\nabla\right) = \underline{\underline{\partial^\alpha}}.$$

1. $\partial_\alpha\partial^\alpha = (\partial/\partial x^0)^2 - \nabla^2 = \square$ (D'Alembertian) is manifestly invariant wave operator.

2. If $\tilde{A}$ is a 4-vector, then the 4-divergence: $\boxed{\partial_\alpha A^\alpha = \frac{\partial A^0}{\partial x^0} + \nabla\cdot\mathbb{A} = \partial^\alpha A_\alpha}$ is a Lorentz invariant (because it is a scalar product in the sense of Eq. (48)).

So: $\partial_\alpha A^\alpha = 0$ is a Lorentz invariant eqtn, e.g. We shall use this many times.

REVIEW

Recent accounts of "Special Math" for SRT

ϕ520: Feb. 1993

1) Notion of 4-vectors: $\tilde{A} = (A_0, \mathbb{A})$.

A. Invariance of Minkowski "length" $\tilde{A}^2 = A_0^2 - \mathbb{A}^2$ under LT's

B. Form-invariant way of writing SRT kinematics & dynamics.

(Get velocity-addition from $\tilde{u} \to \tilde{u}' = \underset{\sim}{\Lambda}\,\tilde{u}$, etc)

(Get $\mathcal{E}^2 = (pc)^2 + (mc^2)^2$ from conserved $\tilde{p}^2$, etc)

---

2) Nature of Minkowski space.

(−) sign is "non-Euclidean"

A. Scalar product: $\tilde{A} \cdot \tilde{B} = A_0 B_0 - (A_1 B_1 + A_2 B_2 + A_3 B_3)$.

B. Can write: $\tilde{A} \cdot \tilde{B} = [\tilde{A}]\,\underset{\sim}{g}\,(\tilde{B})$, $\underset{\sim}{g} = \begin{pmatrix} 1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix}$ metric tensor.

---

3) Construction of most general LT's, as Minkowski "rotations".

A. $\underset{\sim}{\Lambda} = (e^{-\overbrace{\boldsymbol{\omega}}^{\text{ROTATION}} \cdot \underset{\text{set of 3 rotation matrices}}{\underset{\sim}{\mathbb{S}}}})(e^{-\overbrace{\boldsymbol{\zeta}}^{\text{BOOST}} \cdot \underset{\text{set of 3 boost matrices}}{\underset{\sim}{\mathbb{K}}}})$, (from: $\underset{\sim}{\Lambda}\,\underset{\sim}{g}\,\underset{\sim}{\Lambda} = \underset{\sim}{\Lambda}$, $\det \underset{\sim}{\Lambda} = 1$).

---

4) Contravariant & Covariant Vectors & Tensors.

A. position $\mathbb{r}$ transforms like $\underbrace{A_i \to A_i' = (\partial x_i' / \partial x_j)\, A_j}_{\text{CONTRAVARIANT vector}: A^i}$ Summation convention

B. gradient $\partial/\partial\mathbb{r}$ transforms: $\underbrace{B_i \to B_i' = (\partial x_j / \partial x_i')\, B_j}_{\text{COVARIANT vector}: B_i}$

C. We can rewrite scalar products as $A \cdot B = A^i B_i$, w/o $\underset{\sim}{g}$.