Lecture IV: Particle motion in special relativity

(Dated: August 30, 2019)

I. OVERVIEW

This lecture considers particle motion (e.g., 4-velocities) in special relativity. This is covered in Chapter 2 of the book (plus I have dropped in some ideas that we will encounter in Chapter 3), but I will do it from a Lagrangian perspective (the book does not), since I think this will be helpful in understanding the similar problem in general relativity. It will also be helpful in further familiarizing you with the "up and down index" notation.

II. THE 4-VELOCITY AND 4-ACCELERATION

Let's consider the trajectory of a particle through spacetime, which may be described parametrically as $x^\alpha(\sigma)$, where σ is any coordinate describing the path of the particle (it could be laboratory time, it could be proper time, etc.). The proper time along a small segment of the trajectory is

$$d\tau = \sqrt{-ds^2} = \sqrt{-\eta_{\mu\nu}dx^\mu dx^\nu} = \sqrt{-\eta_{\mu\nu}\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma}}\,d\sigma. \tag{1}$$

We can define the *4-velocity* of a particle to be the vector $\boldsymbol{U}$:

$$U^\alpha = \frac{dx^\alpha}{d\tau} = \left(-\eta_{\mu\nu}\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma}\right)^{-1/2}\frac{dx^\alpha}{d\sigma}. \tag{2}$$

By construction, the 4-velocity is normalized:

$$\eta_{\alpha\beta}U^\alpha U^\beta = \left(-\eta_{\mu\nu}\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma}\right)^{-1}\eta_{\alpha\beta}\frac{dx^\alpha}{d\sigma}\frac{dx^\beta}{d\sigma} = -1. \tag{3}$$

In the frame where the particle is instantaneously at rest, $\boldsymbol{U}$ has components $(1,0,0,0)$. We can thus think of $\boldsymbol{U}$ as being the 0th basis vector for an observer riding along with the particle, $\boldsymbol{U} = \boldsymbol{e}_{\bar{0}}$.

Note that the transformation rule for basis vectors then implies:

$$\boldsymbol{U} = \boldsymbol{e}_{\bar{0}} = [\boldsymbol{\Lambda}^{-1}]^\epsilon{}_{\bar{0}}\boldsymbol{e}_\epsilon \quad\to\quad U^\epsilon = [\boldsymbol{\Lambda}^{-1}]^\epsilon{}_{\bar{0}}. \tag{4}$$

One may similarly define a 4-acceleration,

$$a^\alpha = \frac{dU^\alpha}{d\tau}. \tag{5}$$

Note that by differentiating Eq. (3), we find

$$0 = \frac{d}{d\tau}\left(\eta_{\alpha\beta}U^\alpha U^\beta\right) = \eta_{\alpha\beta}U^\alpha\frac{dU^\beta}{d\tau} + \eta_{\alpha\beta}\frac{dU^\alpha}{d\tau}U^\beta = \eta_{\alpha\beta}U^\alpha a^\beta + \eta_{\alpha\beta}a^\alpha U^\beta = 2\boldsymbol{U}\cdot\boldsymbol{a}, \tag{6}$$

so $\boldsymbol{U}\cdot\boldsymbol{a} = 0$. In the frame of the particle, where $U^{\bar{0}} = 1$ and $U^{\bar{i}} = 0$, this says $a^{\bar{0}} = 0$: thus in the particle's own frame, the 4-acceleration can have only spatial components.

III. THE PARTICLE ACTION AND 4-MOMENTUM

In special relativity, like in Newtonian physics, a particle with no forces acting on it will travel in a straight line at constant speed. A variant of this idea, which will carry over to general relativity, is that a particle follows a *geodesic* – a path of locally shortest (or longest, or saddle-point) length. This is similar to the principle of least action that you learned in classical mechanics, and we will use similar mathematical tools to make use of the principle of the geodesic.

We define the *free-particle action* for the path of a particle in special relativity by the equation:

$$S = -m\int_{\rm start}^{\rm finish} d\tau. \tag{7}$$

Here m is a constant (which we will later identify with the particle's mass), $d\tau$ is the differential of proper time along its trajectory, and the $-$ sign is inserted so that the solution in special relativity will be a minimum action rather than a maximum.

A. Solution of the free particle action

We will write S in terms of the parameter σ:

$$S = \int_{\text{start}}^{\text{finish}} -m\sqrt{-\eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma}\frac{\mathrm{d}x^\nu}{\mathrm{d}\sigma}}\,\mathrm{d}\sigma = \int_{\text{start}}^{\text{finish}} f\,\mathrm{d}\sigma; \tag{8}$$

in this sense, the integrand f takes the place of the Lagrangian that you learned about in undergraduate physics when you did the calculus of variations. You know that the minimum action path can be found from the Euler-Lagrange equations:

$$\frac{\mathrm{d}p_\gamma}{\mathrm{d}\sigma} - \frac{\partial f}{\partial x^\gamma} = 0, \quad p_\gamma = \frac{\partial f}{\partial(\mathrm{d}x^\gamma/\mathrm{d}\sigma)}, \tag{9}$$

where f is expressed in terms of x^γ, $\mathrm{d}x^\gamma/\mathrm{d}\sigma$, and σ. The canonical momentum p_μ can be determined from Eq. (8) via

$$\begin{aligned}
p_\gamma &= \frac{\partial f}{\partial(\mathrm{d}x^\gamma/\mathrm{d}\sigma)} \\
&= \frac{\partial}{\partial(\mathrm{d}x^\gamma/\mathrm{d}\sigma)}\left[-m\sqrt{-\eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma}\frac{\mathrm{d}x^\nu}{\mathrm{d}\sigma}}\right] \\
&= -\frac{1}{2}m\left(-\eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma}\frac{\mathrm{d}x^\nu}{\mathrm{d}\sigma}\right)^{-1/2}\left(-\eta_{\mu\nu}\delta^\mu_\gamma\frac{\mathrm{d}x^\nu}{\mathrm{d}\sigma} - \eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma}\delta^\nu_\gamma\right) \\
&= m\left(-\eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma}\frac{\mathrm{d}x^\nu}{\mathrm{d}\sigma}\right)^{-1/2}\eta_{\gamma\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\sigma} \\
&= m\eta_{\gamma\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \\
&= m\eta_{\gamma\nu}u^\nu.
\end{aligned} \tag{10}$$

(I will denote the 4-velocity of the particle here by $\boldsymbol{u}$ rather than $\boldsymbol{U}$, since I may use the latter for an observer.) Since f doesn't depend explicitly on x^γ, we have $\partial f/\partial x^\gamma = 0$, and Eq. (9) implies that p_γ is conserved along a trajectory.

B. Transformation of the canonical momenta

You will notice that I wrote the momentum with a "down" index. The reason for this is how it transforms. If we switch to a barred coordinate, we have

$$x^{\bar\delta} = \Lambda^{\bar\delta}{}_\gamma x^\gamma \quad \text{and} \quad x^\gamma = [\boldsymbol{\Lambda}^{-1}]^\gamma{}_{\bar\delta}x^{\bar\delta}. \tag{11}$$

Therefore:

$$\frac{\mathrm{d}x^\gamma}{\mathrm{d}\sigma} = [\boldsymbol{\Lambda}^{-1}]^\gamma{}_{\bar\delta}\frac{\mathrm{d}x^{\bar\delta}}{\mathrm{d}\sigma} \tag{12}$$

(remember on the right-hand side that because of EΣC we sum over $\bar\delta = 0, 1, 2, 3$). This means that the derivative defining a canonical momentum transforms under the chain rule:

$$\begin{aligned}
p_{\bar\delta} &= \frac{\partial f}{\partial(\mathrm{d}x^{\bar\delta}/\mathrm{d}\sigma)} \\
&= \frac{\partial(\mathrm{d}x^\gamma/\mathrm{d}\sigma)}{\partial(\mathrm{d}x^{\bar\delta}/\mathrm{d}\sigma)}\frac{\partial f}{\partial(\mathrm{d}x^\gamma/\mathrm{d}\sigma)} \\
&= [\boldsymbol{\Lambda}^{-1}]^\gamma{}_{\bar\delta}p_\gamma.
\end{aligned} \tag{13}$$

In matrix notation, the canonical momentum transforms with a factor of $\boldsymbol{\Lambda}^{-1}$: $\bar{\boldsymbol{p}} = \boldsymbol{\Lambda}^{-1\,\mathrm{T}}\boldsymbol{p}$.

A quantity with this transformation law is called a *1-form* or *covector*. We will always write it with a down index.

C. Raising and lowering indices

You might think it is weird that I described momentum as a covector; you learned as an undergraduate that it is a vector, so surely "covectors" (down indices) and "vectors" (up indices) are related. In fact they are, but only through the use of the object $\boldsymbol{\eta}$. If I have a vector $\boldsymbol{v}$ whose components obey the transformation law $v^{\bar\alpha} = \Lambda^{\bar\alpha}{}_\beta v^\beta$, then I can define the lowered-index object

$$v_\zeta = \eta_{\zeta\alpha} v^\alpha. \tag{14}$$

We call this operation "lowering an index." This is a covector because when I apply a Lorentz transformation:

$$v_{\bar\epsilon} = \eta_{\bar\epsilon\bar\alpha} v^{\bar\alpha} = \eta_{\bar\epsilon\bar\alpha}\Lambda^{\bar\alpha}{}_\beta v^\beta = \eta_{\bar\epsilon\bar\alpha}\Lambda^{\bar\alpha}{}_\beta[\boldsymbol{\eta}^{-1}]^{\beta\gamma} v_\gamma = [\bar{\boldsymbol{\eta}}\boldsymbol{\Lambda}\boldsymbol{\eta}^{-1}]_{\bar\epsilon}{}^\gamma v_\gamma = [\boldsymbol{\Lambda}^{-1\,\mathrm{T}}]_{\bar\epsilon}{}^\gamma v_\gamma, = [\boldsymbol{\Lambda}^{-1}]^\gamma{}_{\bar\epsilon} v_\gamma, \tag{15}$$

where we use the transformation rule:

$$\bar{\boldsymbol{\eta}} = \boldsymbol{\Lambda}^{-1\,\mathrm{T}}\boldsymbol{\eta}\boldsymbol{\Lambda}^{-1} \quad\to\quad \bar{\boldsymbol{\eta}}\boldsymbol{\Lambda}\boldsymbol{\eta}^{-1} = \boldsymbol{\Lambda}^{-1\,\mathrm{T}}. \tag{16}$$

Thus Eq. (15) is the transformation rule for covectors: the operation of matrix multiplication by $\boldsymbol{\eta}$ turns a vector into a covector. In special relativity, where

$$\boldsymbol{\eta} = \left(\begin{array}{c|ccc} -1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right), \tag{17}$$

we have simply:

$$v_0 = -v^0, \quad v_1 = v^1, \quad v_2 = v^2, \quad \text{and} \quad v_3 = v^3. \tag{18}$$

That is, lowering an index flips the sign on the time component. The equivalent operation in Euclidean geometry with Cartesian coordiantes (no $-$ sign) does nothing, which is why in your introductory physics classes you never made a distinction between vectors and covectors.

If a vector can be turned into a covector, then a covector can be turned into a vector. The raising operation is the matrix inverse:

$$v^\alpha = [\boldsymbol{\eta}^{-1}]^{\alpha\zeta} v_\zeta. \tag{19}$$

The raising operation uses the inverse of $\boldsymbol{\eta}$. Of course, this matrix is:

$$\boldsymbol{\eta}^{-1} = \left(\begin{array}{c|ccc} -1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right); \tag{20}$$

it happens to have the same entries as $\boldsymbol{\eta}$. We will simply write the matrix inverse with upper indices: "$\eta^{\alpha\zeta}$ will be taken to mean $[\boldsymbol{\eta}^{-1}]^{\alpha\zeta}$. In general relativity, the object $\eta_{\mu\nu}$ (known as the *metric tensor*) will be replaced by a general symmetric matrix $g_{\mu\nu}$; we will continue to use upper indices to denote the matrix inverse, but you should keep in mind that in GR the values are not the same.

With the use of the metric tensor, we can re-write the equation $p_\gamma = m\eta_{\gamma\nu}U^\nu$ as $p_\gamma = mU_\gamma$ or $\boldsymbol{p} = m\boldsymbol{U}$.

IV. PROPERTIES OF THE 4-MOMENTUM

A. Energy and momentum conservation

The conservation of the canonical momenta p_γ is an example of Noether's theorem: they are related to the underlying translation symmetries in space ($\gamma = 1, 2, 3$) and time ($\gamma = 0$). You know that the conserved momentum associated with translations along each of the coordinate axes is normally identified with that component of momentum p_i, and the conserved momentum associated with translations in the time direction is identified as $-E$, where E is the energy: $E = -p_0$. In special relativity, we can view these as the definitions of momentum and energy, and we can write $E = p^0$ (since raising a time index simply flips the sign).

A common problem would be to determine the energy $\bar E$ seen by an observer $\bar{\mathcal{O}}$ with 4-velocity $\boldsymbol{U}$, whose frame is related to the lab frame $\mathcal{O}$ by a Lorentz transformation $\boldsymbol{\Lambda}$. We see that:

$$\bar E = -p_{\bar 0} = -[\boldsymbol{\Lambda}^{-1}]^\gamma{}_{\bar 0} p_\gamma = -U^\gamma p_\gamma = -U^\gamma \eta_{\gamma\delta} p^\delta = -\boldsymbol{U}\cdot\boldsymbol{p}, \tag{21}$$

where we have used Eq. (4).

B. Mass shell constraint

Because the 4-velocity $\boldsymbol{u}$ satisfies $\boldsymbol{u} \cdot \boldsymbol{u} = -1$, the 4-momentum components are not independent. We must have $\boldsymbol{p} \cdot \boldsymbol{p} = -m^2$. This means, in component notation:

$$-(p^0)^2 + (p^1)^2 + (p^2)^2 + (p^3)^2 = -m^2 \tag{22}$$

or

$$E^2 = (p^1)^2 + (p^2)^2 + (p^3)^2 + m^2. \tag{23}$$

This is the familiar relation between energy and momentum.

C. Expression in terms of 3D velocity

Finally, we write the energy and momentum in terms of 3D velocities. We already know that for a particle with 3D velocity $\boldsymbol{v}$ and $\gamma = (1 - v^2)^{-1/2}$, we have

$$u^0 = \frac{\mathrm{d}x^0}{\mathrm{d}\tau} = \frac{\mathrm{d}t}{\mathrm{d}\tau} = \gamma \tag{24}$$

and

$$u^i = \frac{\mathrm{d}x^i}{\mathrm{d}\tau} = \frac{\mathrm{d}x^i}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}\tau} = v^i\gamma. \tag{25}$$

Thus the energy is

$$E = p^0 = mu^0 = m\gamma = m + \frac{1}{2}mv^2 + \frac{3}{8}mv^4 + ... \tag{26}$$

and the 3D momentum is

$$p^i = mu^i = mv^i\gamma. \tag{27}$$

Thus we see that the canonical momentum corresponds to our usual notion of momentum for slow-moving particles. The canonical energy corresponds to a rest mass (m; remember $c^2 = 1$) plus a kinetic energy ($\frac{1}{2}mv^2$) plus higher-order corrections.