The Concept of Four-vectors

V. Devanathan

The Academy of Sciences, Chennai
Department of Nuclear Physics
University of Madras, Guindy Campus
Chennai - 600 025
Four-dimensional World

Special Theory of Relativity
United: Space & Time and Mass & Energy
Energy & Momentum conservation

We live in a four-dimensional world of space-time continuum.

Einstein introduced the concept of four vectors such that the scalar product of any two four-vectors is invariant under Lorentz transformations.

It is similar to the concept that the scalar product of any two three-vectors in the three dimensional space is invariant under rotation of coordinate system.
We list below some of the four-vectors.

\[ x : \quad ct, x, y, z. \quad (x_0, x: \text{time-space}) \]

\[ \frac{\partial}{\partial x} : \quad \frac{1}{c} \frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z}. \quad (\text{time-space gradients}) \]

\[ p : \quad E/c, px, py, pz. \quad (\text{energy-momentum}) \]

\[ J : \quad c\rho, J_x, J_y, J_z. \quad (\text{charge and current densities}) \]

\[ k : \quad \omega/c, k_x, k_y, k_z. \quad (\text{four wave-vector}) \]

\[ A : \quad \phi, A_x, A_y, A_z. \quad (\text{scalar, vector potentials}) \]

We use upright bold letters to denote the four-vector and italic bold letters to denote the three-vectors. A scalar product of any two four-vectors \( a \) and \( b \) is defined by

\[ a \cdot b = a_0b_0 - a \cdot b = a_0b_0 - a_xb_x - a_yb_y - a_zb_z. \]

The first component of the four-vector is usually called the time-component or the zeroth component of the four vector. Note that in defining the scalar product of two four-vectors, we use different signs for the time component and space components.
The scalar product of the four-vectors listed above are invariant under Lorentz transformation. This gives a powerful method of calculating the kinematical variables and their transformation from one inertial coordinate system to another.
Lorentz transformation is common for both Newtonian Mechanics and Maxwell’s theory of Electromagnetism.

The special theory of relativity has brought together Newtonian mechanics and Maxwell’s equations of electromagnetism into one fold, satisfying Lorentz transformations when one goes from one inertial frame to another. In the case of Newtonian mechanics, Galilian transformations still hold true when the motion of particles is much lower than the speed of light. Lorentz transformations reduce to Galililian transformations for velocities much less than the velocity of light.
Advantages of the Four-Vector Concept

Reaction: \[ A + B \rightarrow C + D + E \]

Any reaction obeys two conservation laws:
1. Energy conservation
2. Momentum conservation

Energy & Momentum depends on the frame of reference.

In the Four-Vector formalism, the two laws are merged into one law of Energy-momentum conservation and is much simpler to go from one reference frame to another.
Four-momentum Square

Relativistic Energy-momentum relation for a particle with rest mass $m$

$$E^2 = p^2 c^2 + m^2 c^4$$

If $\mathbf{p}$ is the four momentum, then

$$\mathbf{p}^2 = \frac{E^2}{c^2} - p^2 = m^2 c^2.$$  

In Units with $c = 1$,

$$\mathbf{p}^2 = E^2 - p^2 = m^2.$$  

The square of the four-momentum is equal to the square of the rest mass in natural units. This is independent of the reference frame.
Problem 1:
Proton-antiproton pair production

\[ p + p \rightarrow p + p + p + \bar{p}. \]

What is the threshold energy of the incident proton in laboratory frame for proton-antiproton pair production in proton-proton collision?

\( P_a \): four-momentum of the incident proton
\( P_b \): four-momentum of the target proton
\( P_c \): four-momentum of the aggregate of particles in the final state.

Conservation of energy: \( E_a + E_b = E_c \)
Conservation of momentum: \( p_a + p_b = p_c \)

They can be written jointly as conservation of energy-momentum in the four-vector notation.

\[ P_a + P_b = P_c. \]

Squaring, we get

\[ (P_a + P_b) \cdot (P_a + P_b) = P_c \cdot P_c. \]
In c.m. frame, the four momentum of the final state for the threshold production of $p\bar{p}$ pair is

$$P_c = (E_c, P_c) = (4M, 0),$$

since the minimum energy required for $p\bar{p}$ pair production is $4M$, where $M$ is the mass of the proton (antiproton) and the net momentum $P_c = 0$ in c.m. frame.

$$P_a^2 + P_b^2 + 2P_a \cdot P_b = 16M^2$$
$$2M^2 + 2P_a \cdot P_b = 16M^2, \quad (P_a^2 = P_b^2 = M^2)$$
$$P_a \cdot P_b = 7M^2.$$

In laboratory frame,

$$P_a = (E_a, p_a), \quad P_b = (M, 0).$$

So,

$$P_a \cdot P_b = ME_a.$$

Since the scalar product of two four-vectors is invariant under transformation of coordinate systems,

$$ME_a = 7M^2 \text{ or } E_a = 7M.$$
Threshold energy of incident proton in Lab. for $p\bar{p}$ production

$E_a$: Total energy of the incident proton in Lab. system for $p\bar{p}$ pair production

$E_a = M + T_a$

= Rest mass energy $M$ + Kinetic energy $T_a$

= $7M$

This yields the value $T_a = 6M$, which is threshold kinetic energy for the incident proton in laboratory for the $p\bar{p}$ pair production. Taking the value $M = 938$ MeV, we obtain the kinetic energy of the incident proton in laboratory for threshold production of $p\bar{p}$ as

$T_a = 6 \times 938$ MeV $= 5.628$ GeV.

Thus, we find that only a part of the kinetic energy of the incident proton is used up in the production of new particles.
Problem 2
Decay of charged $\pi$-meson: $\pi^+ \rightarrow \mu^+ + \nu_\mu$

A charged $\pi$-meson (rest mass = 273 $m_e$) at rest decays into a muon (rest mass = 207 $m_e$) and a neutrino (zero rest mass), where $m_e$ denotes the rest mass of electron ($m_e = 0.511 \text{ MeV}/c^2$). What is the energy of the emitted neutrino?

Four-momentum equation for the pion-decay

$$P_\pi = P_\mu + P_\nu \quad \text{or} \quad P_\pi - P_\nu = P_\mu.$$  

Squaring, we get

$$P^2_\pi + P^2_\nu - 2P_\pi \cdot P_\nu = P^2_\mu.$$  

$m^2_\pi + m^2_\nu - 2m_\pi E_\nu = m^2_\mu.$

Since $m_\nu = 0$, we obtain

$$E_\nu = \frac{m^2_\pi - m^2_\mu}{2m_\pi} = \frac{273^2 - 207^2}{2 \times 273} = 58.02 \ m_e$$

$$= 58.02 \times 0.511 = 29.6 \text{ MeV}.$$
Problem 3

Photoproduction of $\pi^+$ from proton:

$$\gamma + p \rightarrow n + \pi^+$$

What is the threshold energy of the photon in MeV for photoproduction of $\pi^+$ from proton, given the masses $M_p = M_n = M = 939$ MeV/$c^2$ and $m_\pi = 139$ MeV/$c^2$?

Energy-momentum conservation

$$P_\gamma + P_p = P_{n\pi},$$

$P_\gamma$: four-momentum of incident photon

$P_p$: four-momentum of the target proton

$P_{n\pi}$: four momentum of final particles $n$ & $\pi$

Squaring, we get (using units with $c = 1$)

$$P_\gamma^2 + P_p^2 + 2P_\gamma \cdot P_p = P_{n\pi}^2$$

$$0 + M^2 + 2E_\gamma M = (M + m)^2$$

$$E_\gamma = \frac{2Mm + m^2}{2M} = m + \frac{m^2}{2M}.$$
We have used above the principle that the scalar product of four-vectors is invariant in all inertial frames.

So, we have evaluated $P_\gamma \cdot P_p$ in laboratory frame and $P_{n\pi}^2$ in centre of momentum (c.m.) frame. In laboratory frame, the target particle is at rest and in c.m. frame, the total momentum of the final state of the particles will be zero.

Substituting the values $M = 939$ MeV and $m = 139$ MeV, we get

$$E_\gamma = 149.29\text{ MeV}.$$  

Thus, we obtain the threshold energy of the incident photon to be 149.29 MeV for the photoproduction of $\pi^+$ from proton.
Problem 4
The Compton Effect

When a photon is scattered by an electron, the shift in the wavelength of the scattered photon depends only on the angle of scattering and not on the wavelength of the incident photon. This is known as the Compton effect.

If \( \lambda_i \) is the wavelength of the incident photon and \( \lambda_f \) is the wavelength of the scattered photon, then the shift in the wavelength \( \Delta \lambda \) is given by

\[
\Delta \lambda = \lambda_f - \lambda_i = \frac{h}{m_0c}(1 - \cos \theta),
\]

where \( \theta \) denotes the angle of scattering of the photon, \( m_0 \) the rest mass of the electron, \( h \) the Planck constant and \( c \) the velocity of light.
The Compton scattering

Four-vector equation for compton scattering.

\[ \mathbf{k}_i + \mathbf{p}_i = \mathbf{k}_f + \mathbf{p}_f \quad \text{or} \quad \mathbf{k}_i - \mathbf{k}_f + \mathbf{p}_i = \mathbf{p}_f. \]

Squaring, we have

\[
(k_i - k_f + p_i)^2 = (p_f)^2
\]

\[
\mathbf{k}_i^2 + \mathbf{k}_f^2 + \mathbf{p}_i^2 - 2\mathbf{k}_i \cdot \mathbf{k}_f + 2\mathbf{k}_i \cdot \mathbf{p}_i - 2\mathbf{k}_f \mathbf{p}_i = \mathbf{p}_f^2.
\]

For photon, the rest mass is zero and for the electron, the rest mass is \( m_0 \).
So, we obtain

\[ k_i^2 = k_f^2 = 0; \quad p_i^2 = p_f^2 = m_0^2; \]
\[ k_i \cdot k_f = E_i E_f - (k_i \cdot k_f) = E_i E_f (1 - \cos \theta); \]
\[ k_i \cdot p_i = E_i m_0, \quad \text{since } p_i = 0; \]
\[ k_f \cdot p_i = E_f m_0, \quad \text{since } p_i = 0. \]

Substituting these values, we get

\[ m_0 (E_i - E_f) = E_i E_f (1 - \cos \theta), \]

Incident photon energy: \( E_i = h \nu_i \)

Scattered photon energy: \( E_f = h \nu_f \)

Angle of scattering: \( \theta \)

\[ \nu_i - \nu_f = \frac{h}{m_0} \nu_i \nu_f (1 - \cos \theta). \]

Since \( \nu = 1/\lambda \), in units \( c = 1 \), the Compton shift in wavelength for the scattered photon, in units with \( c = 1 \).

\[ \Delta \lambda = \lambda_f - \lambda_i = \frac{h}{m_0} (1 - \cos \theta). \]

Restoring to MKS units,

\[ \Delta \lambda = \frac{h}{m_0 c} (1 - \cos \theta). \]
Rotation in three-dimensional space

\[ x' = x \cos \theta + y \sin \theta \]
\[ y' = y \cos \theta - x \sin \theta \]
\[ z' = z \]

Rotation about the z–axis through an angle \( \theta \)

\[ r^2 = x^2 + y^2 + z^2 = x'^2 + y'^2 + z'^2. \]

The length of the vector \( r \) is invariant under rotation of the coordinate system.
Rotation about the common z-axis, leaves the z-coordinate unchanged. So, consider only changes in the x-y plane.
Orthogonal Transformation

We can write the transformation of coordinates in the form of a matrix $R(\theta)$.

\[
\begin{bmatrix}
  x' \\
  y'
\end{bmatrix} = R(\theta) \begin{bmatrix}
  x \\
  y
\end{bmatrix} = \begin{bmatrix}
  \cos \theta & \sin \theta \\
  -\sin \theta & \cos \theta
\end{bmatrix} \begin{bmatrix}
  x \\
  y
\end{bmatrix}.
\]

\[
\begin{bmatrix}
  x \\
  y
\end{bmatrix} = R^{-1}(\theta) \begin{bmatrix}
  x' \\
  y'
\end{bmatrix} = \begin{bmatrix}
  \cos \theta & -\sin \theta \\
  \sin \theta & \cos \theta
\end{bmatrix} \begin{bmatrix}
  x' \\
  y'
\end{bmatrix}.
\]

This is what is called the orthogonal transformation since the transpose $\tilde{R}(\theta)$ of the matrix $R(\theta)$ is equal to its inverse $R^{-1}(\theta)$.

\[
\tilde{R}(\theta) = R^{-1}(\theta).
\]

The above consideration in three-dimensional space can be extended to n-dimensional space wherein a space-point can be represented by a set of n coordinates, $x_1, x_2, x_3 \cdots x_n$. 
Four-dimensional complex Minkowski space

Let the Fourth Coordinate be $x_4 = ict$.

The Minkowski space is a complex four-dimensional space. The Lorentz transformation can be visualized as a rotation in the complex Minkowski space that preserves the length of the vector.

$$x'^2 + y'^2 + z'^2 - c^2 t'^2 = x^2 + y^2 + z^2 - c^2 t^2.$$ 

The length of the vector $(x^2 + y^2 + z^2 - c^2 t^2)^{1/2}$, in this case may be real or imaginary, real if $x^2 + y^2 + z^2 - c^2 t^2$ is positive, or imaginary if $x^2 + y^2 + z^2 - c^2 t^2$ is negative. This means that the Lorentz transformation is an orthogonal transformation in the four-dimensional complex Minkowski space.
Consider Lorentz transformation between any two inertial frames which are in uniform relative motion along the common x-axis. Consider changes in $x_1$ and $x_4$ coordinates only since $x_2$ and $x_3$ remain invariant.

$$
\begin{align*}
x'_1 &= \gamma (x_1 - vt) \quad \rightarrow \quad x'_1 = \gamma (x_1 + i\beta x_4); \\
t' &= \gamma \left(t - \frac{vx_1}{c^2}\right) \quad \rightarrow \quad x'_4 = \gamma (x_4 - i\beta x_1).
\end{align*}
$$

(1)

$x_4 = ict, x'_4 = ict', \gamma = (1 - \beta^2)^{-1/2}, \beta = v/c.$

In matrix form.

$$
\begin{bmatrix}
x'_1 \\
x'_4
\end{bmatrix} =
\begin{bmatrix}
\gamma & i\beta \gamma \\
-i\beta \gamma & \gamma
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_4
\end{bmatrix}.
$$

(2)

The Lorentz transformation matrix

$$
L =
\begin{bmatrix}
\gamma & i\beta \gamma \\
-i\beta \gamma & \gamma
\end{bmatrix},
$$

(3)

is an orthogonal matrix.

$$
\tilde{L} = L^{-1}; \quad \tilde{L}L = L\tilde{L} = 1.
$$

(4)

The matrix $L$ represents a rotation in $x_1 - x_4$ plane of Minkowski’s 4-dimensional space through an angle $\theta$ which is imaginary.
Four-dimensional real Minkowski space

Replace \( x_4 = ict \) by \( x_0 = ct \).
The new coordinates are \( x_0, x_1, x_2, x_3 \).
Lorentz transformation equations in terms of the new set of coordinates are

\[
\begin{align*}
x' &= \gamma(x - vt) \quad \rightarrow \quad x'_1 = \gamma(x_1 - \beta x_0); \\
\quad t' &= \gamma\left(t - \frac{vx}{c^2}\right) \quad \rightarrow \quad x'_0 = \gamma(x_0 - \beta x_1).
\end{align*}
\]

Notations: \( x_0 = ct, \beta = v/c, \gamma = (1 - \beta^2)^{-\frac{1}{2}} \)

The Eq. in matrix form

\[
\begin{bmatrix}
x'_0 \\
x'_1
\end{bmatrix}
= \begin{bmatrix}
\gamma & -\beta \gamma \\
-\beta \gamma & \gamma
\end{bmatrix}
\begin{bmatrix}
x_0 \\
x_1
\end{bmatrix}.
\]

The Lorentz transformation matrix

\[
L = \begin{bmatrix}
\gamma & -\beta \gamma \\
-\beta \gamma & \gamma
\end{bmatrix},
\]

is not an orthogonal matrix.

\[
\tilde{L} \neq L^{-1}; \quad LL \neq 1.
\]

The inverse Lorentz transformation

\[
\begin{bmatrix}
x_0 \\
x_1
\end{bmatrix}
= \begin{bmatrix}
\gamma & \beta \gamma \\
\beta \gamma & \gamma
\end{bmatrix}
\begin{bmatrix}
x'_0 \\
x'_1
\end{bmatrix}.
\]
The metric tensor $g_{\mu\nu}$

In order to define the scalar product of two four-vectors, which is invariant under Lorentz transformation, a metric tensor $g_{\mu\nu}$ ($g^{\mu\nu}$) is introduced.

$$g_{\mu\nu} = g^{\mu\nu} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & -1
\end{bmatrix}.$$

If $\mathbf{x}$ is a four-vector with components $x_0, x_1, x_2, x_3$ and $\tilde{\mathbf{x}}$ is its transpose, then

$$\tilde{\mathbf{x}} g \mathbf{x} = \begin{bmatrix}
x_0 & x_1 & x_2 & x_3
\end{bmatrix} \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & -1
\end{bmatrix} \begin{bmatrix}
x_0 \\
x_1 \\
x_2 \\
x_3
\end{bmatrix} = x_0^2 - x_1^2 - x_2^2 - x_3^2.$$
**Contravariant and covariant four-vectors**

The introduction of metric tensor leads to two different types of four vectors - contravariant and covariant four-vectors.

Let us consider a four-vector $\mathbf{A}$. It has four components, one time component $A_t$ and three space components $A_x, A_y, A_z$. A contravariant four-vector is identified by a superscript.

$$A^\mu : A^0, A^1, A^2, A^3 = A_t, A_x, A_y, A_z,$$

A covariant four-vector is identified by a subscript.

$$A_\mu : A_0, A_1, A_2, A_3 = A_t, -A_x, -A_y, -A_z.$$
Using the metric tensor, one can convert a contravariant four-vector into a covariant four-vector* and vice versa.

\[ A_\mu = \sum_\nu g_{\mu\nu} A^\nu, \quad \mu = 0, 1, 2, 3. \]  
(5)

\[ A^\mu = \sum_\nu g^{\mu\nu} A_\nu, \quad \mu = 0, 1, 2, 3. \]  
(6)

The scalar product of any two four-vectors \( A \) and \( B \) can be written as

\[ A \cdot B = A^\mu B_\mu = A^0 B_0 + A^1 B_1 + A^2 B_2 + A^3 B_3 \]

\[ = A_t B_t - A_x B_x - A_y B_y - A_z B_z \]

\[ = A_t B_t - A \cdot B. \]  
(7)

We adopt the convention of representing four-vectors by upright bold letters \( A, B \) and three-component vectors by italic bold letters \( A, B \).

*The components \( A_t, -A_x, -A_y, -A_z \) of the four-vector \( A_\mu \) have the same signs as the metric tensor \( g_{\mu\nu} \) and so, it is called the covariant four-vector. The components \( A_t, A_x, A_y, A_z \) of the four-vector \( A^\mu \) have signs that are not in conformity with the signs of the metric tensor \( g^{\mu\nu} \) and so it is called the contravariant four-vector.
A contravariant four-vector $A^\mu$ transforms from one coordinate system (unprimed) $x^\mu$ to another coordinate system (primed) $x'^\mu$ according to the formula

$$A'^\mu = \sum_\nu \frac{\partial x'^\mu}{\partial x^\nu} A^\nu; \quad A^\nu = \sum_\mu \frac{\partial x^\nu}{\partial x'^\mu} A'^\mu; \quad (8)$$

On the other hand, a covariant four-vector $A_\mu$ transforms from the unprimed coordinate system to the primed coordinate system and vice versa according to the law

$$A'_\mu = \sum_\nu \frac{\partial x^\nu}{\partial x'^\mu} A_\nu; \quad A_\nu = \sum_\mu \frac{\partial x'^\mu}{\partial x^\nu} A'_\mu; \quad (9)$$

It can be easily checked that the distinction between the contravariant and covariant vector vanishes in the case of complex four-dimensional cartesian Minkowski space since the Lorentz transformation corresponds to an orthogonal transformation. The distinction arises only in the case of real Minkowski space, due to introduction of the metric tensor $g_{\mu\nu}$. 
Lorentz Transformations

Contravariant Vector

\[
\begin{bmatrix}
A'\,^0 \\
A'\,^1
\end{bmatrix}
= \begin{bmatrix}
\gamma & -\beta \gamma \\
-\beta \gamma & \gamma
\end{bmatrix}
\begin{bmatrix}
A\,^0 \\
A\,^1
\end{bmatrix}.
\]

\[
\begin{bmatrix}
A\,^0 \\
A\,^1
\end{bmatrix}
= \begin{bmatrix}
\gamma & \beta \gamma \\
\beta \gamma & \gamma
\end{bmatrix}
\begin{bmatrix}
A'\,^0 \\
A'\,^1
\end{bmatrix}.
\]

Covariant vector

\[
\begin{bmatrix}
A'\,^0 \\
A'\,^1
\end{bmatrix}
= \begin{bmatrix}
\gamma & \beta \gamma \\
\beta \gamma & \gamma
\end{bmatrix}
\begin{bmatrix}
A\,^0 \\
A\,^1
\end{bmatrix}.
\]

\[
\begin{bmatrix}
A\,^0 \\
A\,^1
\end{bmatrix}
= \begin{bmatrix}
\gamma & -\beta \gamma \\
-\beta \gamma & \gamma
\end{bmatrix}
\begin{bmatrix}
A'\,^0 \\
A'\,^1
\end{bmatrix}.
\]
Reference

V. Devanathan - The Special Theory of Relativity - Narosa Publishing House

Thank you