Chapter 16

General Covariance

The term covariance implies a formalism in which the laws of physics maintain the same form under a specified set of transformations.

EXAMPLE: Lorentz covariance implies equations that are constructed in such a way that they do not change their form under Lorentz transformations (three boosts between inertial systems and three rotations).\textsuperscript{a}

\textsuperscript{a}An inertial system is a frame of reference in which Newton’s first law of motion holds. Thus, for example, rotating frames and accelerated frames are not inertial. An inertial system is therefore in uniform translational motion with respect to any other inertial frame.
16.1 Covariance and Poincaré Transformations

• **Lorentz covariance** makes it *manifest* that the principles of special relativity
  
  – Invariance under Lorentz transformations
  – Constant speed of light
  – Time dilation and space contraction
  – Relativity of simultaneity
  – Equivalence of energy and mass

are obeyed by a set of equations.

• **Poincaré covariance:** (Invariance under the six Lorentz transformations plus the four possible uniform space–time translations). Covariance under Poincaré transformations makes it *manifest* that
  
  – Physics does not depend on choice of coordinate system origin, orientation, 
    
  – This implies conservation laws such as those for energy and angular momentum.

In the absence of gravity, all physical systems are thought to be Poincaré invariant (and thus invariant under the Lorentz subgroup of the Poincaré group).
Notice the subtle difference between invariance and covariance with respect to a set of transformations.

- Invariance means that the physical observables of the system are not changed by the transformation (think of the rotation of a featureless sphere).

- Covariance means that the system is invariant, and that this invariance is manifest in the formulation of the theory (can be “seen at a glance”).

Thus, a system that is not formulated in a covariant way might still be invariant, but it may not be obvious that this is so (not manifest) without detailed examination.

Covariance with respect to Poincaré transformations is still insufficient to deal with gravity. We seek a more general covariance that embraces the possibility of non-inertial coordinate systems.

**General Covariance:** a physical equation holds in a gravitational field provided that

- It holds in the absence of gravity (agrees with the predictions of special relativity in flat spacetime).

- It maintains its form under a general coordinate transformation $x \rightarrow x'$ (possibly between accelerated frames).
16.2 Covariance and Tensor Notation

- We shall be concerned generically with a transformation between one set of spacetime coordinates, denoted by

\[ x \equiv x^\mu = (x^0, x^1, x^2, x^3) \]

and a new set

\[ x'^\mu = x'^\mu(x) \quad \mu = 0, 1, 2, 3 \]

where \( x = x^\mu \) denotes the original (untransformed) coordinates.

This notation is an economical form of

\[ x'^\mu = \xi^\mu(x^1, x^2, x^3, x^4) \quad (\mu = 1, 2, \ldots n) \]

where the single-valued, continuously differentiable functions \( \xi^\mu \) assign a new (primed) coordinate \((x'^1, x'^2, x'^3, x'^4)\) to a point of the manifold with old coordinates \((x^1, x^2, x^3, x^4)\). This transformation may be abbreviated to \( x'^\mu = \xi^\mu(x) \) and, even more tersely, to \( x'^\mu = x'^\mu(x) \).

- Coordinates are just labels, so laws of physics cannot depend on them. This implies that the system \( x'^\mu \) is not privileged and therefore this transformation should be invertible.

- Notice carefully that we are talking about the same point described in two different coordinate systems.
As a minimum, we must consider the transformations of

- Fields
- Derivatives of fields
- Integrals of fields.

The first two are necessary to formulate Lagrangians and equations of motion, and the latter enter into various conservation laws.

To facilitate this task, we shall now introduce a set of mathematical quantities called tensors that are a generalization of the idea of scalars and vectors to more components.

16.2.1 Scalar Transformation Law

Simplest possibility: field has a single component (magnitude) at each point that is unchanged by the transformation

$$\varphi'(x') = \varphi(x).$$

Quantities such as $\varphi(x)$ that are unchanged under the coordinate transformation are called scalars.

EXAMPLE: Value of the temperature at different points on the surface of the Earth.
16.2.2 Vectors

The gradient $\frac{\partial \phi}{\partial x^\nu}$ of a scalar field $\phi(x)$ obeys

$$\frac{\partial \phi(x)}{\partial x^{\mu \nu}} = \sum_{\nu} \frac{\partial \phi(x)}{\partial x^\nu} \frac{\partial x^\nu}{\partial x^{\mu \nu}} \equiv \frac{\partial \phi(x)}{\partial x^\nu} \frac{\partial x^\nu}{\partial x^{\mu \nu}}$$

(All partial derivatives understood to be evaluated at same point $P$.)

Einstein summation convention:

- An index that is repeated, once as a superscript and once as a subscript, implies a summation over that index.
- Such an index is a dummy index that is removed by the summation and should not appear on the other side of the equation.
- A repeated (dummy) index may be replaced by any other index not already in use without altering equation: $A_\alpha B^\alpha = A_\beta B^\beta$.
- A superscript (subscript) in a denominator counts as a subscript (superscript) in a numerator.
- Greek indices ($\alpha, \beta, \ldots$) denote the full set of spacetime indices running over 0, 1, 2, 3
- Roman indices ($i, j, \ldots$) denote the indices 1, 2, 3 running only over the spatial coordinates.
- Placement of indices matters: generally $x^\alpha$ and $x_\alpha$ are different quantities.
- At all stages of manipulating equations, the indices on the two sides of an equation (including their up or down placement) must match.
We can classify tensors according to a notation \( t^m_n \), where \( n \) is the number of lower indices and \( m \) is the number of upper indices

- Thus a scalar is a tensor of type \( t^0_0 \), since it carries no indices.
- The sum of \( n \) and \( m \) is the rank of the tensor. A scalar is a tensor of rank zero.

There are two kinds of rank-1 tensors, having the index pattern \( t^0_1 \) and \( t^1_0 \), respectively. The first is called a covariant vector:

**COVARIANT VECTOR:** A tensor having a transformation law that mimics that of the scalar field gradient,

\[
A'_{\mu}(x') = \frac{\partial x^\nu}{\partial x'^{\mu}} A_\nu(x) \quad \text{(covariant vector)}
\]

is of type \( t^0_1 \) and is termed a covariant vector or a 1-form.

**ECONOMY OF NOTATION:** The preceding equation really means four equations:

\[
A'_\mu = \frac{\partial x^0}{\partial x'^{\mu}} A_0 + \frac{\partial x^1}{\partial x'^{\mu}} A_1 + \frac{\partial x^2}{\partial x'^{\mu}} A_2 + \frac{\partial x^3}{\partial x'^{\mu}} A_3 \quad (\mu = 0, 1, 2, 3)
\]

each containing four terms. It is equivalent to the matrix equation

\[
\begin{pmatrix}
A'_0 \\
A'_1 \\
A'_2 \\
A'_3
\end{pmatrix} =
\begin{pmatrix}
\frac{\partial x^0}{\partial x'^0} & \frac{\partial x^0}{\partial x'^1} & \frac{\partial x^0}{\partial x'^2} & \frac{\partial x^0}{\partial x'^3} \\
\frac{\partial x^1}{\partial x'^0} & \frac{\partial x^1}{\partial x'^1} & \frac{\partial x^1}{\partial x'^2} & \frac{\partial x^1}{\partial x'^3} \\
\frac{\partial x^2}{\partial x'^0} & \frac{\partial x^2}{\partial x'^1} & \frac{\partial x^2}{\partial x'^2} & \frac{\partial x^2}{\partial x'^3} \\
\frac{\partial x^3}{\partial x'^0} & \frac{\partial x^3}{\partial x'^1} & \frac{\partial x^3}{\partial x'^2} & \frac{\partial x^3}{\partial x'^3}
\end{pmatrix}
\begin{pmatrix}
A_0 \\
A_1 \\
A_2 \\
A_3
\end{pmatrix}.
\]
CONTRA VARIANT VECTORS: A differential transforms like

\[ dx'\mu = \frac{\partial x'\mu}{\partial x^\nu} dx^\nu, \]

which suggests a second rank-1 transformation rule

\[ A'^\mu(x') = \frac{\partial x'\mu}{\partial x^\nu} A^\nu(x) \]

(contravariant vector).

A tensor that behaves in this way is of type \( t^1_0 \) and is termed a contravariant vector. Note the difference with the covariant vector transformation law:

\[ A'_\mu(x') = \frac{\partial x^\nu}{\partial x'\mu} A^\nu(x) \]

(covariant vector or 1-form)

Therefore, we expect the possibility of two rank-1 tensors:

1. **Covariant vectors (1-forms)**, which carry a lower index and transform like the gradient of a scalar

2. **Contravariant vectors**, which carry an upper index and transform like the coordinate differential

In the general case they must be distinguished (by placement—upper or lower—of indices).
16.2.3 Scalar Product

Covariant and contravariant vectors permit defining a scalar product

\[ A \cdot B \equiv A_\mu B^\mu = A^\mu B_\mu \]

This transforms as a scalar because from

\[ A'_\mu (x') = \frac{\partial x^\nu}{\partial x'^\mu} A_\nu (x) \quad A'^\mu (x') = \frac{\partial x'^\mu}{\partial x^\nu} A^\nu (x). \]

we have that

\[ A' \cdot B' = A'_\mu B'^\mu = \frac{\partial x^\nu}{\partial x'^\mu} A_\nu \frac{\partial x'^\mu}{\partial x^\alpha} B_\alpha = \frac{\partial x^\nu}{\partial x^\alpha} \frac{\partial x'^\mu}{\partial x^\alpha} A_\nu B_\alpha \]
\[ = \frac{\partial x^\nu}{\partial x^\alpha} A_\nu B_\alpha = \delta_\alpha^\nu A_\nu B_\alpha \]
\[ = A_\mu B^\mu = A \cdot B, \]

where we have introduced the Kronecker delta through

\[ \delta_\nu^\mu = \frac{\partial x'^\mu}{\partial x^\nu} = \frac{\partial x^\mu}{\partial x^\nu} = \begin{cases} 1 & \mu = \nu \\ 0 & \mu \neq \nu \end{cases} \]

Eliminating indices by summing over them in tensor products is called contraction. The scalar product has no tensor indices left and is said to be fully contracted.
16.2.4 Rank-2 Tensors

We may distinguish three kinds of rank-2 tensors according to the transformation laws

\[
T'_{\mu \nu} = \frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x^\beta}{\partial x'^\nu} T_{\alpha \beta}
\]

\[
T'_{\nu \mu} = \frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x'^\nu}{\partial x^\beta} T^\beta_{\alpha}
\]

\[
T'^{\mu \nu} = \frac{\partial x'^\mu}{\partial x^\alpha} \frac{\partial x^\nu}{\partial x^\beta} T^{\alpha \beta}
\]

This may easily be generalized to tensors of any rank.

- **Covariant Tensors**: carry only lower indices
- **Contravariant Tensors**: carry only upper indices
- **Mixed Tensors**: carry both upper and lower indices

**EXAMPLE**: the Kronecker delta \(\delta^\nu_{\mu}\) is a mixed tensor of rank 2.

**Memory crutch:**

- Each upper index \(\mu\) on left side requires right-side “factor” of form \(\partial x'^\mu / \partial x^\nu\) (prime in numerator).
- Each lower index \(\nu\) on left side requires right-side “factor” of form \(\partial x^\mu / \partial x'^\nu\) (prime in denominator).
- “Position of index = position of primed coordinate”
- Indices that do not appear on the left side must be “contracted-away” on the right side
NOTE: Not all quantities carrying indices are tensors; *it is the transformation laws that define tensors.*

NOTE: We employ a standard shorthand by using “a tensor $T_{\mu\nu}$” to mean “a tensor with components $T_{\mu\nu}$.”
16.2.5 Metric Tensor

A rank-2 tensor of particular importance is the metric tensor $g_{\mu\nu}$ because it is associated with the line element

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu.$$  

It is symmetric ($g_{\mu\nu} = g_{\nu\mu}$) and satisfies the usual rank-2 tensor transformation law

$$g'_{\mu\nu} = \frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x^\beta}{\partial x'^\nu} g_{\alpha\beta}.$$  

The contravariant metric tensor $g^{\mu\nu}$ is defined by the requirement

$$g_{\mu\alpha} g^{\alpha\nu} = \delta^\nu_\mu.$$  

(Thus, $g_{\mu\nu}$ and $g^{\mu\nu}$ are matrix inverses, which is of considerable practical utility.)

---

We may use contraction with the metric tensor to raise and lower tensor indices; for example

$$A^\mu = g^{\mu\nu} A_\nu \quad A_\mu = g_{\mu\nu} A^\nu.$$  

Thus, the scalar product of vectors may also be expressed as

$$A \cdot B = g_{\mu\nu} A^\mu B^\nu \equiv A_\nu B^\nu.$$
16.2.6 Antisymmetric 4th-Rank Tensor

A rank-4 tensor $\varepsilon^{\alpha \beta \gamma \delta}$ of particular importance may be introduced by the requirement that

- $\varepsilon^{0123} = 1$
- $\varepsilon^{\alpha \beta \gamma \delta}$ be completely antisymmetric in the exchange of any two indices (thus it must vanish if any two indices are the same).

This tensor is commonly called the completely antisymmetric 4th-rank tensor or the Levi–Civita symbol.
16.2.7 Invariant Integration

Change of volume elements for spacetime integration:

\[ d^4x = \det \left( \frac{\partial x}{\partial x'} \right) d^4x' \]

where \( \det(\partial(x)/\partial(x')) \) is the Jacobian determinant of the transformation between the coordinates. The metric tensor transforms as

\[ g'_{\mu\nu} = \frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x^\beta}{\partial x'^\nu} g_{\alpha\beta} \quad \text{(Triple matrix product)} \]

Therefore, since:

\[ \text{determinant of a product} = \text{product of determinants} \]

the determinant of the metric tensor \( g \equiv \det g_{\mu\nu} \) transforms as

\[ g' = \det \left( \frac{\partial x}{\partial x'} \right) \det \left( \frac{\partial x}{\partial x'} \right) g \quad \rightarrow \quad \det \left( \frac{\partial x}{\partial x'} \right) = \frac{\sqrt{|g'|}}{\sqrt{|g|}} \]

which gives when inserted into the first equation

\[ \sqrt{|g|} d^4x = \sqrt{|g'|} d^4x', \]

\( (|g| \text{ because } g \text{ is negative in 4-D spacetime}). \)
This is a scalar expression and thus its value is independent of coordinates. It follows that an integral of the form

\[ I = \int \varphi(x) \sqrt{|g|} d^4x \]

with fixed boundaries is a scalar if \( \varphi(x) \) is a scalar.

In integrals we shall employ

\[ dV = \sqrt{|g|} d^4x \]

as an invariant volume element.
16.2.8 Covariant Derivatives

Let us now consider the derivatives of tensor quantities. First introduce two common compact notations for partial derivatives

\[ \varphi_{,\mu} \equiv \frac{\partial \varphi(x)}{\partial x^\mu} \quad \partial_\mu \varphi \equiv \frac{\partial \varphi(x)}{\partial x^\mu} \]

The derivative of a scalar is a covariant vector and scalars and their derivatives have well-defined tensorial properties. However, for the derivative of a covariant vector, by using the rule for the derivative of a product

\[
A'_{\mu,\nu} \equiv \frac{\partial A'_\mu}{\partial x'^\nu} = \frac{\partial}{\partial x'^\nu} \left( A_{\alpha} \frac{\partial x^\alpha}{\partial x'^\mu} \right) \\
= \frac{\partial A_{\alpha}}{\partial x'^\nu} \frac{\partial x^\alpha}{\partial x'^\mu} + A_{\alpha} \frac{\partial^2 x^\alpha}{\partial x'^\nu \partial x'^\mu} \\
= \frac{\partial A_{\alpha}}{\partial x^\beta \partial x'^\nu} \frac{\partial x^\alpha}{\partial x'^\mu} + A_{\alpha} \frac{\partial^2 x^\alpha}{\partial x'^\nu \partial x'^\mu} \\
= A_{\alpha,\beta} \frac{\partial x^\beta}{\partial x'^\nu} \frac{\partial x^\alpha}{\partial x'^\mu} + A_{\alpha} \frac{\partial^2 x^\alpha}{\partial x'^\nu \partial x'^\mu}
\]

(Tensor) (Not a tensor)

In curved spacetime it is not possible to transform away the second term globally: *Partial differentiation of tensors is NOT a covariant operation in curved spacetime.*
However, if we introduce the Christoffel symbols $\Gamma_{\alpha\beta}^{\lambda}$ and require them to obey a transformation law

$$\Gamma'_{\alpha\beta}^{\lambda} = \Gamma_{\mu\nu}^{\kappa} \frac{\partial x^{\mu}}{\partial x'^{\alpha}} \frac{\partial x^{\nu}}{\partial x'^{\beta}} \frac{\partial x^{\lambda}}{\partial x^{\kappa}} + \frac{\partial^{2} x^{\mu}}{\partial x^{\alpha} \partial x^{\beta}} \frac{\partial x^{\lambda}}{\partial x^{\mu}},$$

we may show that (Exercise),

$$\left( A'_{\mu,\nu} - \Gamma'_{\mu\nu} A'_{\lambda} \right) \equiv T'_{\mu\nu} = \left( A_{\alpha,\beta} - \Gamma_{\alpha\beta}^{\kappa} A_{\kappa} \right) \frac{\partial x^{\alpha}}{\partial x'^{\mu}} \frac{\partial x^{\beta}}{\partial x'^{\nu}},$$

$$\rightarrow T'_{\mu\nu} = \frac{\partial x^{\alpha}}{\partial x^{\mu}} \frac{\partial x^{\beta}}{\partial x^{\nu}} T_{\alpha\beta}.$$

This is the transformation law for a rank-2 covariant tensor. This suggests that we define the covariant derivative of a vector as

$$A_{\mu;\nu} \equiv \left( A_{\mu,\nu} - \Gamma_{\mu\nu}^{\lambda} A_{\lambda} \right),$$

where a subscript comma denotes ordinary partial differentiation and a subscript semicolon denotes covariant differentiation with respect to the variables following it. (We shall also employ the notation $D_{\mu}$ to denote an operator that takes the covariant derivative with respect to $x^{\mu}$.)

The covariant derivative of a covariant vector then transforms as a covariant tensor of rank 2, even though neither of its terms is a tensor.
Likewise, we can introduce the covariant derivative of a contravariant vector

\[ A^\lambda_{;\mu} = A^\lambda_{,\mu} + \Gamma^\lambda_{\alpha \mu} A^\alpha, \]

and the covariant derivatives of the three possible rank-2 tensors through

\[ A_{\mu \nu ; \lambda} = A_{\mu \nu , \lambda} - \Gamma^\alpha_{\mu \lambda} A^\alpha_{\nu \lambda} - \Gamma^\alpha_{\nu \lambda} A_{\mu \alpha}, \]
\[ A^{\mu}_{\nu ; \kappa} = A^{\mu}_{\nu , \kappa} + \Gamma^{\mu}_{\alpha \kappa} A^\alpha_{\nu \kappa} - \Gamma^\alpha_{\nu \kappa} A^{\mu}_{\mu \alpha}, \]
\[ A^{\mu \nu}_{; \kappa} = A^{\mu \nu}_{, \kappa} + \Gamma^{\mu}_{\alpha \kappa} A^{\alpha \nu} + \Gamma^{\nu}_{\alpha \kappa} A^{\mu \alpha}, \]

(which are rank-3 tensors), and so on.

Heuristic rule for constructing the covariant derivative of a tensor having any rank:

- Form the ordinary partial derivative of the tensor
- Add one Christoffel symbol term having the sign and form for a covariant vector for each lower index of the tensor
- Add one Christoffel symbol term having the sign and form for a contravariant vector for each upper index of the tensor
Most rules for partial differentiation carry over with suitable generalization for covariant differentiation.

**EXAMPLE:** covariant derivative of product

\[(A_\mu B_\nu)_;\lambda = A_\mu;\lambda B_\nu + A_\mu B_\nu;\lambda.\]

which is the usual result.

The most important exception concerns the properties of successive covariant differentiations. Although partial derivative operators normally commute, covariant derivative operators generally do not commute with each other.

One important consequence of covariant differentiation (Exercise):

\[D_\alpha g_{\mu\nu} = g_{\mu\nu;\alpha} = 0,\]

Some implications:

- Raising and lowering index by contraction with \(g_{\mu\nu}\) commutes with covariant differentiation.
- This will allow in the Einstein field equations a *vacuum energy term* (accelerated expansion and dark energy).
16.2.9 Invariant Equations

The properties of tensors elaborated above ensure that any equation will be invariant under general coordinate transformations provided that it equates tensors having the same upper and lower indices.

EXAMPLES:

- If the quantities $T_{\nu}^{\mu}$ and $U_{\nu}^{\mu}$ both transform as mixed rank-2 tensors and $T_{\nu}^{\mu} = U_{\nu}^{\mu}$ in the $x$ coordinate system, then in the $x'$ coordinate system $T'_{\nu}^{\mu} = U'_{\nu}^{\mu}$.

- An equation that equates any tensor to zero is invariant under general coordinate transformations.

- Equations such as $T_{\mu}^{\nu} = 42$ or $T_{\mu}^{\mu} = U_{\mu}$ generally are not valid in all coordinate systems because they equate tensors of different kinds.
16.3 Flat Spacetime

- In order to go beyond Newtonian gravitation we must consider, with Einstein, the intimate relationship between the curvature of space and the gravitational field.

- Mathematically, this extension is bound inextricably to the geometry of spacetime, and in particular to the aspect of geometry that permits quantitative measurement of distances.

Let us first consider these ideas within flat space.

16.3.1 Metrics and Distance Intervals

A space that has a prescription associated with it for measuring distances is termed a metric space and the mathematical function that specifies distances is termed the metric for the space.

EXAMPLE: in Euclidean 3-space

\[ d\ell^2 = dx^2 + dy^2 + dz^2 \] ( cartesian)

\[ d\ell^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \] (spherical)
EXAMPLE:

The line element of a 2-sphere is specified by

\[ dl^2 = R^2 d\theta^2 + R^2 \sin^2 \theta d\phi^2, \]

which may be written as the matrix equation

\[
dl^2 = (d\theta \ d\phi) \begin{pmatrix}
R^2 & 0 \\
0 & R^2 \sin^2 \theta \\
\end{pmatrix} \begin{pmatrix}
d\theta \\
d\phi \\
\end{pmatrix}.
\]

The area of the 2-sphere may then be expressed as

\[
A = \int_0^{2\pi} d\phi \int_0^\pi R^2 \sin \theta d\theta
= \int_0^{2\pi} d\phi \int_0^\pi d\theta \sqrt{\det g_{ij}} = 4\pi R^2.
\]

where the metric tensor \( g_{ij} \) is the \( 2 \times 2 \) matrix in the preceding equation.

Let us now generalize these ideas to the 4-dimensional spacetime termed Minkowski space.
16.3.2 Minkowski Space

In a particular inertial frame, introduce unit vectors $e_0$, $e_1$, $e_2$, and $e_3$ that point along the $t$, $x$, $y$, and $z$ axes. Any 4-vector $b$ may be expressed in the form,

$$b = b^0 e_0 + b^1 e_1 + b^2 e_2 + b^3 e_3.$$ 

and the scalar product of 4-vectors is given by

$$a \cdot b = b \cdot a = (a^\mu e_\mu) \cdot (b^\nu e_\nu) = e_\mu \cdot e_\nu a^\mu b^\nu.$$ 

Note that generally we shall use non-bold symbols to denote 4-vectors and reserve bold symbols for 3-vectors. Where there is potential for confusion, we use a notation such as $b^\mu$ to stand generically for all components of a 4-vector.

Introducing the definition

$$\eta_{\mu \nu} \equiv e_\mu \cdot e_\nu,$$

the scalar product may be expressed as

$$a \cdot b = \eta_{\mu \nu} a^\mu b^\nu.$$
and thus the line element becomes

\[ ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2 = \eta_{\mu\nu} dx^\mu dx^\nu, \]

where the metric tensor of flat spacetime may be expressed as

\[ \eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \text{diag}(-1, 1, 1, 1). \]

That is, the line element corresponds to the matrix equation

\[ ds^2 = \begin{pmatrix} cdt \\ dx \\ dy \\ dz \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} cdt \\ dx \\ dy \\ dz \end{pmatrix}, \]

where \( ds^2 \) represents the spacetime interval between \( x \) and \( x + dx \) with

\[ x = (x^0, x^1, x^2, x^3) = (ct, x^1, x^2, x^3). \]

The Minkowski metric is sometimes termed a \textit{pseudo-Euclidean metric} to emphasize that it is Euclidean-like except for the difference in sign between the time and space terms in the line element. It is also called \textit{pseudo-Riemannian} and \textit{Lorentzian}. 
16.3. FLAT SPACETIME

16.3.3 The Light Cone Structure of Minkowski Spacetime

By virtue of the line element

\[ ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2, \]

the Minkowski spacetime may be classified according to the light cone diagram exhibited in Fig. 16.1.

The light cone is a 3-dimensional surface in the 4-dimensional spacetime and events in spacetime may be characterized according to whether they are inside of, outside of, or on the light cone.
The standard terminology:

- If $ds^2 < 0$ the interval is termed *timelike*.

  In that case $ds/c$ is the time measured by a clock moving freely from $x$ to $x + dx$ (the *proper time*; see below).

  If two events differ by a timelike interval, there exists an inertial frame in which they occur at the *same place but different times*.

- If $ds^2 > 0$ the interval is termed *spacelike*.

  Then $ds$ may be interpreted as the length of a ruler with ends at $x$ and $x + dx$, as measured by an observer at rest with respect to the ruler.

  If events differ by a spacelike interval, there exists an inertial frame in which they occur at the *same time but different places*.

- If $ds^2 = 0$ the interval is called *lightlike* (or *null*).

  Then the points $x$ and $x + dx$ are connected by signals moving at light speed.
The light cone classification makes clear the distinction between Minkowski spacetime and a true 4-dimensional Euclidean space

- Two points in the Minkowski spacetime may be separated by a distance whose square could be
  - positive,
  - negative,
  - zero.

- In particular, lightlike particles have worldlines confined to the light cone and the square of the separation of any two points on a lightlike worldline is zero.
16.3.4 Causality and Spacetime

The causal properties of Minkowski spacetime are encoded in its light cone structure, which requires that $v \leq c$ for all signals.

- Each point in spacetime may be viewed as lying at the apex of a light cone ("Now").
- An event at the origin of a light cone may influence any event in its forward light cone (the "Future").
- The event at the origin of the light cone may be influenced by events in its backward light cone (the "Past").
- Events at lightlike separations are causally connected.
- Events at spacelike separations are causally disconnected.
- Events on the light cone are causally connected but only by signals that travel exactly at $c$. 
The light cone is a surface separating the knowable from the unknowable for an observer at the apex of the light cone.

This light cone structure of spacetime ensures that all velocities obey locally the constraint $v \leq c$. Since velocities are defined and measured locally, covariant field theories in either flat or curved space are guaranteed to respect the speed limit $v \leq c$, irrespective of whether globally velocities appear to exceed $c$.

**EXAMPLE:** In the Hubble expansion of the Universe, galaxies beyond a certain distance (the horizon) would recede from us at velocities in excess of $c$. However, all local measurements in that expanding, possibly curved, space would determine the velocity of light to be $c$. 
16.3.5 Geodesics

A metric allows us to define *geodesics* for the corresponding space:

- A geodesic is a path that represents the *shortest distance between any two points*.
- A geodesic may also be viewed as the “*straightest possible path*” between two points.
- More technically, a geodesic is *a curve that parallel-transports its own tangent vector*.

**FLAT SPACE:**

“The shortest distance between two points is a straight line.” Thus, the geodesics in Euclidean space are given by

\[ \ddot{r} = 0 \quad \text{(Newton’s 1st law)} \]

**MINKOWSKI SPACE:**

\[ \frac{d^2 t}{d\tau^2} = 0 \quad \frac{d^2 \mathbf{r}}{d\tau^2} = 0, \]

where \( \tau \) is the proper time (the time that would be measured by a clock carried along a worldline). In both cases, the geodesics are straight lines (generally will not be true in curved spacetime).
16.3.6 Geometrized Units

It is convenient to introduce a new set of units in which $c$ and/or $G$ can be set to unit value so that they do not appear explicitly in equations. These are called geometrized units or $c = G = 1$ units. Geometrized units, and how to convert between standard units and geometrized units, are explained in an Appendix.

From this point onward, we shall commonly work in units for which $c = 1$ or $c = G = 1$, unless the explicit restoration of $c$ or $G$ factors in an equation is desirable for clarity or to make a particular point.
16.3.7 4-Velocities

Particles with finite mass follow *timelike worldlines*. The worldline for a particle is conveniently parameterized in terms of a variable that changes continuously along the worldline. For timelike trajectories the natural choice for this parameter is the *proper time* $\tau$.

The equation of the worldline may then be expressed as

$$x^\mu = x^\mu(\tau)$$

and we may define a velocity 4-vector (the *4-velocity*) by

$$u^\mu = \left(\frac{dx^0}{d\tau}, \frac{dx^1}{d\tau}, \frac{dx^2}{d\tau}, \frac{dx^3}{d\tau}\right).$$

The proper time interval $d\tau$ is related to the spacetime interval $ds$

$$d\tau^2 = -ds^2,$$

and the coordinate time interval $dt$ and the proper time interval $d\tau$ are related through special relativistic time dilation:

$$d\tau = dt \left(1 - v^2\right)^{1/2} = \frac{1}{\gamma} dt \quad \gamma \equiv \left(1 - v^2\right)^{-1/2} \quad \text{(Lorentz \(\gamma\))}$$

where $v$ is the 3-velocity, $v^j = dx^j/dt$

Note: $c = 1$ *units*! This would read $d\tau = dt \left(1 - v^2/c^2\right)^{1/2}$ in standard units.
The 4-velocity is tangent to the worldline of a particle at any point and lies within the forward light cone (Fig. 16.3). Then,

\[ u^0 = \frac{dx^0}{d\tau} = \frac{dt}{d\tau} = \left(1 - v^2\right)^{-1/2}, \quad u^i = \frac{dx^i}{d\tau} = \frac{dx^i}{dt} \frac{dt}{d\tau} = v_i \left(1 - v^2\right)^{-1/2}, \]

so that we may write for the components of the 4-velocity

\[ u^\mu = (\gamma, \gamma v) \quad \gamma = \left(1 - v^2\right)^{-1/2}, \]
The scalar product of \( \mathbf{u} \) with itself gives the normalization

\[
\mathbf{u} \cdot \mathbf{u} = \eta_{\mu \nu} u^\mu u^\nu = \eta_{\mu \nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = -1,
\]

where \( \eta_{\mu \nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \).

A cavalier\(^1\) proof:

\[
ds^2 = -d\tau^2 = \eta_{\mu \nu} dx^\mu dx^\nu,
\]

which gives, upon dividing by \( d\tau^2 \),

\[
-1 = \eta_{\mu \nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = \mathbf{u} \cdot \mathbf{u}.
\]

Note: \( d\tau^2 = -ds^2 > 0 \) for a massive (timelike) particle.

For massive particles we may always invoke the condition \( \mathbf{u} \cdot \mathbf{u} = -1 \).

---

\(^1\text{Definition:} \) (a) A supporter of King Charles I in the English Civil War. (b) A horseman. (c) A small spaniel of a breed with a moderately long, noncurly, silky coat. (d) Showing a lack of proper concern; offhand.
16.3.8 4-Momenta

We may define the 4-momentum by

\[ p^\mu \equiv (E, p) = mu^\mu , \]

where \( m \) is the rest mass. The normalization of the 4-momentum is

\[ p^2 \equiv p \cdot p = m^2 u \cdot u = -m^2 , \]

since \( u \cdot u = -1 \). Because \( u^\mu = (\gamma, \gamma v) \), the components of the 4-momentum are

\[ p^\mu = (E, p) = (\gamma m, \gamma mv) \quad \rightarrow \quad p_\mu = \eta_{\mu \nu} p^\nu = (-E, p) , \]

with \( \gamma = (1 - v^2)^{-1/2} \). Thus, \( p^2 = -m^2 \) implies that

\[ p_\mu p^{\mu} = (-E, p) \begin{pmatrix} E \\ p \end{pmatrix} = -m^2 \quad \rightarrow \quad E = \sqrt{p^2 + m^2} , \]

which is just the familiar Einstein relation

\[ E = \sqrt{p^2 c^2 + m^2 c^4} \quad \rightarrow \quad E = mc^2 \quad (p \rightarrow 0) , \]

written in \( c = 1 \) units.
16.3.9 Principle of Extremal Proper Time

Principle of extremal proper time: *the worldline for free particles between timelike separated points extremizes the proper time between them* (Fig. 16.4).

From \((c = 1 \text{ units})\)

\[
d\tau^2 = -ds^2 = -(-dt^2 + dx^2 + dy^2 + dz^2),
\]

the proper time between the points \(A\) and \(B\) is

\[
\tau_{AB} = \int_A^B d\tau = \int_A^B (dt^2 - dx^2 - dy^2 - dz^2)^{1/2}
\]
We may parameterize the path by a variable $\sigma$ that varies continuously from 0 to 1 as the particle moves from $A$ to $B$ and

$$
\tau_{AB} = \int_0^1 \left[ \left( \frac{dt}{d\sigma} \right)^2 - \left( \frac{dx}{d\sigma} \right)^2 - \left( \frac{dy}{d\sigma} \right)^2 - \left( \frac{dz}{d\sigma} \right)^2 \right]^{1/2} d\sigma
$$

The condition for an extremum is that

$$
\delta \int d\tau = 0,
$$

where the variation is generally of the form

$$
\delta f = \frac{\partial f}{\partial x^\mu} \delta x^\mu
$$

Defining a Lagrangian

$$
L = \left( -\eta_{\mu\nu} \frac{dx^\mu}{d\sigma} \frac{dx^\nu}{d\sigma} \right)^{1/2} \quad \rightarrow \quad \tau_{AB} = \int_0^1 L \, d\sigma
$$

the variation $\delta \int d\tau = 0$ then implies the Euler–Lagrange equations of motion

$$
- \frac{d}{d\sigma} \left( \frac{\partial L}{\partial (dx^\mu/d\sigma)} \right) + \frac{\partial L}{\partial x^\mu} = 0.
$$
EXAMPLE: Consider $x^\mu = x^1$. For constant $\eta_{\mu\nu}$ the Lagrangian $L$ does not depend on $x^1$ and the Euler–Lagrange equation

$$-\frac{d}{d\sigma} \left( \frac{\partial L}{\partial (dx^\mu / d\sigma)} \right) + \frac{\partial L}{\partial x^\mu} = 0.$$ 

reduces to

$$\frac{d}{d\sigma} \left( \frac{1}{L} \frac{dx^1}{d\sigma} \right) = 0.$$

Inserting $1/L = d\sigma / d\tau$ and multiplying by $d\sigma / d\tau$, gives

$$\frac{d^2 x^1}{d\tau^2} = 0$$

Applying similar steps to the other terms then gives the general result (Exercise)

$$\frac{d^2 x^\mu}{d\tau^2} = 0 \quad \rightarrow \quad \text{No curvature for geodesic}$$

The principle of extremal proper time implies that geodesics in Minkowski space are straight lines.
Principle of Extremal Proper Time (Taylor and Wheeler):

“Spacetime shouts ’Go straight!’ The free stone obeys. ... The stone’s wristwatch verifies that its path is straight.”
16.3.10 Light Rays

For particles moving at lightspeed the rest mass is identically zero.

- Photons move on the light cone with the proper time between two points given by
  \[ d\tau^2 = -ds^2 = 0, \]

- Thus photons travel any distance in zero proper time.

The proper time \( \tau \) is not a useful parameterization for the world line of photons and other massless particles.

However, notice that we may write the curve \( x = t \) (corresponding to \( v = c \) expressed in \( c = 1 \) units) parametrically as

\[ x^\mu = u^\mu \lambda \]

where \( u^\mu = (1, 1, 0, 0) \) is a tangent 4-vector,

\[ u^\mu = \frac{dx^\mu}{d\lambda} \]

and \( \lambda \) is a parameter.
With this choice of parameterization the equation of motion for the light ray may be put into the same form as that for a massive particle

\[
\frac{du}{d\lambda} = 0
\]

which is analogous to Newton’s first law.

- Parameters \( \lambda \) for which this is true are termed *affine parameters*.
- Affine parameters are convenient for light rays because they lead to equations of motion that mimic those for timelike particle trajectories.
For massive particles $u \cdot u = -1$, but since for this photon case we can choose

$$u^\mu = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

(On lightcone)

we have

$$u_\mu = \eta_{\mu\nu}u^\nu = (-1, 1, 0, 0)$$

Thus for photons

$$u \cdot u = u_\mu u^\mu = (-1, 1, 0, 0) \times \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = -1 + 1 = 0.$$

The primary differences between equations governing the motion of massive particles and those governing the motion of massless particles (e.g., photons) in gravitational fields will be associated with the difference in 4-velocity normalizations

$$u \cdot u = -1 \quad \text{(massive particles)}$$

$$u \cdot u = 0 \quad \text{(massless particles)}$$

Otherwise their equations of motion will be similar.
For photons we have that the energy $E$ and momentum $p$ are given by

$$E = \hbar \omega \quad p = \hbar k,$$

where $\hbar$ is Planck’s constant and $k$ is the wavevector. Thus,

$$p^\mu = (E, p) = (\hbar \omega, \hbar k) = \hbar k^\mu = \hbar (\omega, k).$$

Since photons are massless, the 4-momentum and 4-wavevector are normalized such that

$$p \cdot p = k \cdot k = 0.$$

(which is $E = pc$ in $c = 1$ units).
### 16.3.11 Observers

- An observer moving through spacetime may be thought of as occupying a local laboratory moving on a (timelike) worldline.

- The observer carries four orthogonal unit vectors $e_0$, $e_1$, $e_2$, and $e_3$ that specify a local, orthonormal coordinate system (Fig. 16.5).

- This coordinate system defines (locally) a time direction and three space directions to which the observer will reference all measurements.

![Unit vectors of a local coordinate system at a point on an observer’s worldline for two space and one time dimension.](image)
• The timelike component $e_0$ will be tangent to the observer’s worldline (the observer’s clock is moving in that direction if it is at rest in the laboratory).

• Since the 4-velocity $u_{\text{obs}}$ of the observer is a unit tangent vector ($u \cdot u = -1$), we have that

$$e_0 = u_{\text{obs}}.$$ 

and the observer may choose any mutually orthogonal set of three unit spatial vectors to complete the set, as long as they are orthogonal to $e_0$.

• Observers refer observations to the axes of their lab and its clocks. Thus, they measure components of 4-vectors along their chosen basis vectors.

• These components may be computed by taking scalar products with the orthonormal basis 4-vectors.

**Example:** For the 4-momentum

$$p = p^\mu e_\mu.$$ 

We have in particular that the energy measured by an observer with 4-velocity $u_{\text{obs}}$ is given by

$$E = p^0 = -p \cdot e_0 = -p \cdot u_{\text{obs}},$$

since $e_0 = u_{\text{obs}}$. 
16.4 Curved Spacetime

The deflection of light in a gravitational field suggests that gravity is associated with the curvature of spacetime. Thus, let us consider the more general issue of\textit{ covariance in curved spacetime.}

16.4.1 Curved Spaces and Gaussian Curvature

• Gauss demonstrated that for 2-surfaces there is a single invariant (\textit{Gaussian curvature}) characterizing the curvature.

• For a 2-D coordinate system \((x^1, x^2)\) having a diagonal metric with non-zero elements \(g_{11}\) and \(g_{22}\), the Gaussian curvature \(K\) is

\[
K = \frac{1}{2g_{11}g_{22}} \left\{ - \frac{\partial^2 g_{22}}{(\partial x^1)^2} - \frac{\partial^2 g_{11}}{(\partial x^2)^2} + \frac{1}{2g_{11}} \left[ \frac{\partial g_{11}}{\partial x^1} \frac{\partial g_{22}}{\partial x^1} + \left( \frac{\partial g_{11}}{\partial x^2} \right)^2 \right] \right. \\
+ \frac{1}{2g_{22}} \left[ \frac{\partial g_{11}}{\partial x^2} \frac{\partial g_{22}}{\partial x^2} + \left( \frac{\partial g_{22}}{\partial x^1} \right)^2 \right] \right\},
\]

which is generally

– Position-dependent

– An intrinsic quantity expressed entirely in terms of the metric for the space and its derivatives
For orthogonal coordinates \((x, y)\),

\[
K(x_0, y_0) = \frac{1}{R_x(x_0)R_y(y_0)},
\]

where \(R_x(x_0)\) is the radius of curvature in the \(x\) direction and \(R_y(y_0)\) is the radius of curvature in the \(y\) direction, both evaluated at a point \((x_0, y_0)\).

**EXAMPLE:** For the special case of a 2-sphere, \(R_x = R_y \equiv R\) and

\[
K = \frac{1}{R^2}
\]

where \(R\) is the radius of the sphere, which is constant.
Consider the 2-sphere of Fig. 16.6 (embedded in 3D euclidean space), defined by

$$x^2 + y^2 + z^2 = R^2,$$

Let us use the circumference of a circle relative to that for flat space to measure deviation from flatness.

- We may define a circle in the 2-dimensional space by marking a locus of points lying a constant distance $S$ from a reference point, which we choose to be the north pole in Fig. 16.6.
- The angle subtended by $S$ is $S/R$ and $r = R \sin(S/R)$. Then the circumference of the circle is

$$C = 2\pi r = 2\pi R \sin\left(\frac{S}{R}\right) = 2\pi S \left(1 - \frac{S^2}{6R^2} + \ldots\right).$$
• If the space were flat, the circumference of the circle would just be $2\pi S$, so the higher-order terms measure the curvature.

• We have that $K = 1/R^2$. Substituting $R^2 = 1/K$ and assuming $S \to 0$,

$$C \simeq 2\pi S \left(1 - \frac{S^2}{6R^2}\right) = 2\pi S \left(1 - \frac{KS^2}{6}\right) \quad (\text{Valid as } S \to 0).$$

Solving for $K$ in the expansion we obtain a general expression for the Gaussian curvature,

$$K = \lim_{S \to 0} \frac{3}{\pi} \left(\frac{2\pi S - C}{S^3}\right).$$

We may define the Gaussian curvature for an arbitrary 2-dimensional surface by measuring the circumference of small circles relative to their radius.

• This example has been for 2-D space. Later we shall generalize the Gaussian curvature parameter for a two-dimensional surface to a set of parameters (elements of the Riemann curvature tensor) that describe the curvature of 4-dimensional spacetime.
We often embed a surface in a higher-dimensional space in order to more easily visualize our arguments. It is important to emphasize that the intrinsic curvature properties of a space can be determined entirely by the properties of the space itself, without reference to a higher-dimensional embedding space.
16.4.2 Distance Intervals in Curved Spacetime

In curved spacetime the interval between two events may be expressed as

$$ds^2 = g_{\alpha\beta}(x)dx^\alpha dx^\beta.$$ 

The metric tensor $g_{\alpha\beta}(x)$ in a curved spacetime generally has a more complicated form than that for Minkowski space, and is a function of the spacetime coordinates.
16.5 Covariant Derivatives and Parallel Transport

• Covariant derivatives have a geometrical interpretation associated with comparison of vectors located at two different space-time points.

• The comparison issue becomes critical for the calculation of derivatives because, by definition, constructing the derivative of a vector requires taking the difference of vectors at two different points within the space.

In curved spaces we do not define a vector in the curved space itself, but rather on the plane tangent to the point on the curved surface (called the tangent space), as illustrated in Fig. 16.7 for the simple case of a sphere.
Since we are dealing with Riemannian spaces for which a locally Euclidean coordinate system can be constructed around an arbitrary point, it is always possible to define such a tangent space.

While the image of Fig. 16.7 is conceptually useful, it is important to appreciate that defining the tangent space by a local flat coordinate system at a point is an intrinsic process with respect to the original manifold and does not require embedding in a higher-dimensional manifold.
Parallel transport of vectors is necessary to compare two vectors at different points (e.g., to define derivatives).

- For a flat space the tangent space corresponds with the space itself and we can just move one vector, keeping its orientation fixed with respect to a global set of coordinate axes, to the position of the other vector and compare.

- On a curved surface this issue is more complicated because the tangent plane also rotates between two points (Fig. 16.7). Natural notion of parallel transport: keep the vector parallel to itself in infinitessimal steps; see Fig. 16.8 (space locally euclidean).

- As Fig. 16.8 illustrates, parallel transport of vectors on a curved surface is generally path-dependent. Hence parallel transport in curved spaces is not unique and requires a prescription.

- The apparent rotation of a vector when parallel transported around a closed path is a measure of the curvature of the space. (On a 2D surface the rotation is proportional to the Gaussian curvature.)
16.5. COVARIANT DERIVATIVES AND PARALLEL TRANSPORT

Parallel transport in flat space

Parallel transport in curved space

Figure 16.9: Illustration of comparing vectors transported on flat and curved surfaces.

In comparing vectors $V_\mu(x)$ and $V_\mu(y)$ at two spacetime points, there are two contributions to any difference $\Delta V_\mu$ (see Fig. 16.9):

\[
\Delta V_\mu = dV_\mu + \delta V_\mu,
\]

where $dV_\mu$ is the change in the same coordinates and $\delta V_\mu$ is the change of the tangent space.

For flat space $\delta V_\mu = 0$ and we obtain the familiar

\[
\Delta V_\mu = dV_\mu = (\partial_\alpha V_\mu)dx^\alpha \quad \text{(covariant deriv = partial deriv)}.
\]

For curved space with infinitesimal separation between $x$ and $y$, $\delta V_\mu$ is linear in $V_\mu$ and $dx^\alpha$,

\[
\delta V_\mu = \Gamma^\nu_{\mu\alpha} V_\nu dx^\alpha,
\]

where $\Gamma^\nu_{\mu\alpha}$ is the Christoffel symbol. Therefore,

\[
\Delta V_\mu = dV_\mu + \delta V_\mu
\]

\[
= (\partial_\alpha V_\mu + \Gamma^\nu_{\mu\alpha} V_\nu) dx^\alpha \quad \text{(curved space)}
\]

\[
= \partial_\alpha V_\mu dx^\alpha \quad \text{(flat space)}
\]
We will see later that

- $\Gamma^\nu_{\mu\alpha}$ is also called the affine connection or the connection coefficient, or the metric connection, or just the connection.

- $\Gamma^\nu_{\mu\alpha}$ can be constructed from the metric tensor and its derivatives.

- $\Gamma^\nu_{\mu\alpha}$ can be chosen to vanish in a space with constant metric.

- $\Gamma^\nu_{\mu\alpha}$ does not follow from the differential geometry of the manifold but is an additional imposed structure that specifies how tangent spaces at different points are related (connected).

- The Riemann curvature tensor describing the local intrinsic curvature of the spacetime may be constructed from the affine connection.

Thus the affine connection (Christoffel symbol) is central to

1. defining the covariant derivative,

2. implementing parallel transport of tensors, and

3. measuring quantitatively the curvature of a manifold.
16.6 Absolute Derivatives

**Absolute derivatives (intrinsic derivatives)** closely related to covariant derivatives.

- Covariant derivatives defined over an entire manifold in terms of ordinary partial derivatives plus correction terms to cancel non-tensorial character.
- Absolute derivatives defined only along paths in manifold in terms of ordinary derivatives plus correction terms to cancel non-tensorial character.

Using $DA/Du$ to denote the absolute derivative along a path parameterized by $u$

\[
\frac{DA_\alpha}{Du} = \frac{dA_\alpha}{du} - \Gamma^\beta_{\alpha\gamma} A_\beta \frac{dx^\gamma}{du} \quad \text{(Covariant vectors)}
\]

\[
\frac{DA^\alpha}{Du} = \frac{dA^\alpha}{du} + \Gamma^\alpha_{\beta\gamma} A^\beta \frac{dx^\gamma}{du} \quad \text{(Contravariant vectors)}
\]

(Generalizations for higher-order tensors similar to that discussed earlier for covariant derivatives.)
Parallel Transport of Vectors: For euclidean (pseudo-euclidean) manifold, parallel transport of a vector along a path means that length and direction of vector (referenced to a universal coordinate system) don’t change, implying that components satisfy

$$\frac{dA^\mu}{du} = 0 \quad \text{(Flat space)}.$$ 

For a more general Riemannian (pseudo-Riemannian) manifold this generalizes to

$$\frac{DA^\mu}{Du} = 0 \quad \text{(Curved space)},$$
16.7 Isometries and Killing Vectors

In differential geometry, Killing vectors are standard tools for analyzing symmetries such as those that arise as conservation laws in the usual Lagrangian or Hamiltonian formulations of mechanics.

• In all spacetimes, whether flat or not, one constant of motion may be deduced from the normalization of the 4-velocity $u^\mu = dx^\mu / d\tau$:

$$g_{\mu\nu} u^\mu u^\nu = -1,$$

corresponding to the preservation of $u \cdot u$.

• If there are additional constants of motion, they must arise from specific symmetries in the problem.

• In ordinary mechanics, continuous symmetries imply conservation laws. Example: conservation of angular momentum follows from a potential that is spherically symmetric.

• If a spacetime metric has a symmetry (termed an isometry), that too will generally imply that some quantity is conserved.
Suppose the metric is independent of one of the spacetime coordinates, say $x^0$, such that

$$x^0 \rightarrow x^0 + \text{constant}$$

leaves the metric unchanged. For such an isometry we define a unit vector pointing along the direction in which the metric is constant,

$$\xi^\mu = (1, 0, 0, 0).$$

The vector $\xi^\mu$ is termed the *Killing vector* associated with the symmetry.

**EXAMPLE:** In flat 3D space

$$ds^2 = dx^2 + dy^2 + dz^2$$

and conservation of the components of linear momentum is associated with three Killing vectors

$$(1, 0, 0) \quad (0, 1, 0) \quad (0, 0, 1)$$

indicating invariance under translations in the $x$, $y$, and $z$ directions, respectively.
Symmetries implied by Killing vectors mean that some quantity is conserved along a geodesic. This quantity may be found using the principle of extremal proper time (Euler–Lagrange equation).
Example: Suppose that the metric is independent of the coordinate $x^1$, corresponding to a Killing vector

$$\xi^\alpha = (0, 1, 0, 0)$$

Then $\partial L/\partial x^1 = 0$ and

$$\frac{\partial L}{\partial (dx^1/d\sigma)} = -\frac{g_{1\mu}}{L} \frac{dx^\mu}{d\sigma} = -g_{\alpha\mu} \xi^\alpha u^\mu = -\xi \cdot u,$$

where we have used

$$L = \left( -g_{\mu\nu} \frac{dx^\mu}{d\sigma} \frac{dx^\nu}{d\sigma} \right)^{1/2} \quad L d\sigma = d\tau \quad g_{1\mu} = g_{\alpha\mu} \xi^\alpha.$$ 

Then Euler–Lagrange

$$-\frac{d}{d\sigma} \left( \frac{\partial L}{\partial (dx^\mu/d\sigma)} \right) + \frac{\partial L}{\partial x^\mu} = 0$$

reduces to

$$\frac{d}{d\sigma} (\xi \cdot u) = 0 \quad \rightarrow \quad \xi \cdot u \text{ conserved on geodesic}.$$ 

Thus $\xi \cdot u$ is conserved along a geodesic if $\xi$ is a Killing vector associated with a symmetry of the spacetime metric.