

Notes: Special Relativity and Flat Spacetime

Abdelrahman Mohamed Anwar

October 15, 2025

Contents

1 Special Relativity and Flat Spacetime	2
1.1 Prelude	2
1.2 Space and Time, Separately and Together	3
1.3 Lorentz Transformations	5
1.4 Vectors	7
1.5 Dual Vectors (One-Forms)	9
1.6 Tensors	10
1.7 Manipulating Tensors	12
1.8 Maxwell's Equations	13
1.9 Energy and Momentum	15
1.10 Classical Field Theory	17

Chapter 1

Special Relativity and Flat Spacetime

1.1 Prelude

Core Concept of General Relativity (GR)

General relativity is Einstein's theory of space, time, and gravitation. The essential idea is that gravity is not a force represented by fields in spacetime but is instead an inherent property of spacetime itself. The phenomenon experienced as "gravity" is a manifestation of the curvature of spacetime. The objective is to understand spacetime, curvature, and the mechanism by which curvature becomes gravity.

Newtonian Gravity: A Review

Newtonian gravity is defined by two fundamental components: an equation for the gravitational field as influenced by matter and an equation for the response of matter to this field. This can be formulated in two equivalent ways.

Force-Based Formulation

Field Creation: The force between two objects of masses M and m is described by the inverse-square law:

$$F = -\frac{GMm}{r^2}\hat{e}_r \quad (1.1)$$

Matter Response: This force imparts an acceleration to a particle of mass m according to Newton's second law:

$$F = ma \quad (1.2)$$

Potential-Based Formulation

Field Creation: The gravitational potential Φ is related to the mass density ρ by Poisson's equation:

$$\nabla^2\Phi = 4\pi G\rho \quad (1.3)$$

Matter Response: The acceleration of a particle is given by the gradient of the potential:

$$a = -\nabla\Phi \quad (1.4)$$

General Relativity: A Preview

General Relativity replaces the two components of Newtonian theory with statements about spacetime curvature.

How Energy-Momentum Curves Spacetime

The response of spacetime curvature to the presence of matter and energy is governed by Einstein's equation:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi GT_{\mu\nu} \quad (1.5)$$

How Matter Responds to Spacetime Curvature

Free particles move along paths of "shortest possible distance," known as geodesics. In a curved spacetime, these are the closest equivalent to straight lines. The parameterized paths $x^\mu(\lambda)$ of these particles obey the geodesic equation:

$$\frac{d^2x^\mu}{d\lambda^2} + \Gamma_{\rho\sigma}^\mu \frac{dx^\rho}{d\lambda} \frac{dx^\sigma}{d\lambda} = 0 \quad (1.6)$$

The Metric Tensor and the Learning Path

The metric tensor $g_{\mu\nu}$ is the fundamental concept that encodes the geometry of a space. Characterizes curvature by expressing deviations from the flat-space Pythagorean theorem. From the metric, one can derive the Riemann curvature tensor (used in Einstein's equation) and the geodesic equation.

1.2 Space and Time, Separately and Together

Special relativity (SR) is a theory describing the fundamental structure of spacetime—the four-dimensional arena in which all physical processes occur. It replaces Newtonian mechanics, which also describes motion in space and time, but treats them as fundamentally separate entities.

Spacetime and Worldlines

Spacetime: A four-dimensional set of points (events), each labeled by three spatial coordinates and one time coordinate.

Worldline: The one-dimensional curve traced by a particle as it moves through spacetime. The key difference between the Newtonian and relativistic views lies in how they treat the relationship between space and time.

Newtonian Spacetime: Absolute Simultaneity

In the Newtonian picture, time flows uniformly and independently of space. Spacetime is imagined as a stack of "time slices," each representing all of space at one instant of universal time. Simultaneity is absolute—observers everywhere agree on which events occur at the same moment.

Particle trajectories move forward through these slices, unconstrained by any upper limit on their velocity.

Relativistic Spacetime: The Light Cone Structure

Special relativity eliminates absolute time and simultaneity. At each event, the invariant structure is instead the *light cone*—the set of all possible directions that light rays through that event can take.

Events inside the future light cone can be influenced by the event, and events inside the past light cone can influence it. Physical trajectories (worldlines of particles) must always remain within the light cone—no signal or matter can travel faster than light.

Conceptual contrast: Newtonian spacetime has absolute simultaneity; relativistic spacetime has causal cones.

The Invariant Spacetime Interval

The unification of space and time into a single four-dimensional entity is supported by the existence of a quantity invariant under all changes of inertial frames.

Analogy with Euclidean Geometry: In a two-dimensional plane,

$$(\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2$$

is invariant under rotation. The projections Δx and Δy change, but the distance Δs does not.

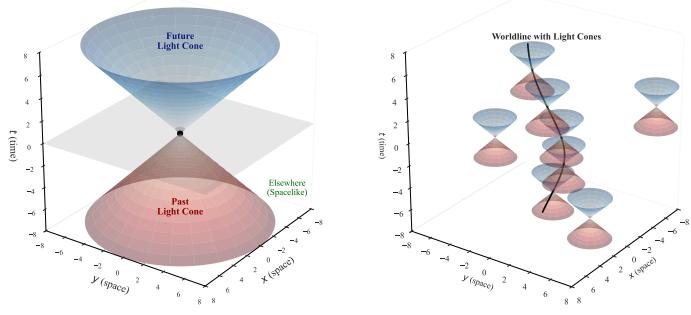


Figure 1.1: Left: Newtonian spacetime slices. Right: Relativistic light cone structure. (See [source notebook](#)).

This invariance defines the geometry of flat spacetime, known as **Minkowski space**.

Minkowski Space: Coordinates and Metric

We define spacetime coordinates x^μ ($\mu = 0, 1, 2, 3$):

$$x^0 = ct, \quad x^1 = x, \quad x^2 = y, \quad x^3 = z$$

and often use units where $c = 1$.

The metric tensor of flat spacetime is

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (1.7)$$

so that

$$(\Delta s)^2 = \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu$$

Causal Structure and Proper Time

Intervals in spacetime fall into three classes:

- **Timelike:** $(\Delta s)^2 < 0$ — events can influence one another (connected by slower-than-light motion).
- **Spacelike:** $(\Delta s)^2 > 0$ — no causal connection possible.
- **Null:** $(\Delta s)^2 = 0$ — connected by light.

For timelike separations, the *proper time* is defined as

$$(\Delta\tau)^2 = -(\Delta s)^2 = -\eta_{\mu\nu} \Delta x^\mu \Delta x^\nu \quad (1.8)$$

It is the time measured by a clock moving along the worldline between two events.

Inertial Frames and Clock Synchronization

Inertial frames are constructed from unaccelerated observers with synchronized clocks. Clocks are synchronized by light signals according to

$$t_2 = \frac{1}{2}(t_1 + t'_1)$$

ensuring that light travel time is equal in both directions.

Likewise, in spacetime the invariant quantity (the *interval*) is

$$(\Delta s)^2 = -(c\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2$$

and remains the same for all inertial observers:

$$(\Delta s)^2 = -(c\Delta t')^2 + (\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2$$

The Twin Paradox

An inertial observer's proper time between two events A and C is simply $\Delta\tau = \Delta t$.

A traveling twin moving out and back with velocity v experiences a total proper time

$$\Delta\tau_{\text{moving}} = \sqrt{(\Delta t)^2 - (\Delta x)^2} = \Delta t \sqrt{1 - v^2} < \Delta t$$

Hence, the accelerated twin ages less. This is not a contradiction but a manifestation of the geometry of spacetime: the straight (inertial) worldline maximizes proper time.

General Trajectories and Proper Time Integral

The infinitesimal spacetime line element is

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu$$

and for any timelike trajectory $x^\mu(\lambda)$, the total proper time is

$$\Delta\tau = \int \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda$$

and for any spacelike trajectory $x^\mu(\lambda)$, the total distance is

$$\Delta s = \int \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda$$

This definition applies even to accelerated motion in flat spacetime.

1.3 Lorentz Transformations

This section provides a formal description of the coordinate transformations between different inertial frames. These are defined as the transformations that leave the spacetime interval, $(\Delta s)^2$, invariant.

Types of Spacetime Transformations

The transformations that preserve the interval consist of translations and linear transformations.

- **Translations:** These are simple shifts of the coordinate origin in space or time.

$$x^{\mu'} = \delta_\mu^{\mu'} (x^\mu + a^\mu) \quad (1.9)$$

Here, a^μ is a constant four-vector representing the shift. Translations leave coordinate differences (Δx^μ) unchanged, so the invariance of the interval under translations is trivial.

- **Linear Transformations:** These include spatial rotations and boosts (changes to a frame with a constant velocity). They are described by multiplying the coordinate vector by a spacetime-independent 4×4 matrix, $\Lambda^{\mu'}_\nu$.

$$x^{\mu'} = \Lambda_\nu^{\mu'} x^\nu \quad (1.10)$$

In matrix notation, this is written as $x' = \Lambda x$.

The Condition for a Lorentz Transformation

To find the properties of the matrix Λ , we impose the condition that the interval must be invariant.

Derivation

1. The spacetime interval is $(\Delta s)^2 = \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu$. In matrix notation, this is $(\Delta s)^2 = (\Delta x)^T \eta (\Delta x)$.
2. In the new coordinate system, the interval is $(\Delta s')^2 = (\Delta x')^T \eta (\Delta x')$.

3. The coordinate differences transform as $\Delta x' = \Lambda \Delta x$. Substituting this into the expression for $(\Delta s')^2$ gives:

$$(\Delta s')^2 = (\Lambda \Delta x)^T \eta (\Lambda \Delta x) = (\Delta x)^T \Lambda^T \eta \Lambda (\Delta x)$$

4. For the interval to be invariant, we must have $(\Delta s)^2 = (\Delta s')^2$ for any arbitrary Δx . This requires that the matrices themselves be equal.

This leads to the fundamental condition that a matrix Λ must satisfy to be a Lorentz transformation.

$$\eta = \Lambda^T \eta \Lambda \quad (\text{Matrix Form}) \quad (1.11)$$

$$\eta_{\rho\sigma} = \Lambda_\rho^{\mu'} \Lambda_\sigma^{\nu'} \eta_{\mu'\nu'} \quad (\text{Index Form}) \quad (1.12)$$

The Lorentz Group

The set of all matrices Λ that satisfy the condition $\eta = \Lambda^T \eta \Lambda$ forms a mathematical group under matrix multiplication, known as the **Lorentz group**.

- **Analogy to the Rotation Group O(3):** There is a close analogy between the Lorentz group and the group of rotations in 3D Euclidean space, O(3). A 3D rotation matrix R preserves the Euclidean distance, which means it satisfies the condition $R^T R = I$, where I is the 3×3 identity matrix. The identity matrix is the metric of flat 3D space. The condition can be written as $I = R^T I R$ to make the analogy clear. The Lorentz group, often denoted **O(3,1)**, is thus a generalization of the rotation group where the Euclidean metric (I) is replaced by the Minkowski metric (η). A metric like η , which has one negative and three positive eigenvalues, is called **Lorentzian**.
- **Subgroups of the Lorentz Group:** The full group O(3,1) includes not only continuous rotations and boosts but also discrete transformations: **parity reversal** (spatial inversion, e.g., $x \rightarrow -x$) and **time reversal** ($t \rightarrow -t$). We can restrict to transformations with determinant $|\Lambda| = 1$. This is the **proper Lorentz group, SO(3,1)**. However, SO(3,1) still includes transformations that reverse the direction of time (e.g., a combined parity and time reversal). To isolate the transformations continuously connected to the identity, we add the condition that the time-time component must be positive, $\Lambda^0_0 \geq 1$. This defines the **proper orthochronous or restricted Lorentz group**, which describes all physical transformations between inertial frames without reversing space or time orientation.
- **The Poincaré Group:** The set of both Lorentz transformations and spacetime translations forms the 10-parameter **Poincaré group**, which represents the full set of spacetime symmetries in special relativity.

Explicit Transformation Matrices

Spatial Rotation: A rotation by an angle θ in the $x - y$ plane has the form:

$$\Lambda^{\mu'}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & \sin \theta & 0 \\ 0 & -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (1.13)$$

Boost: A boost can be thought of as a "rotation" between a time and a space direction. A boost along the x -direction is parameterized by a quantity ϕ called the **rapidity** or **boost parameter**.

$$\Lambda^{\mu'}_\nu = \begin{pmatrix} \cosh \phi & -\sinh \phi & 0 & 0 \\ -\sinh \phi & \cosh \phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (1.14)$$

The rapidity ϕ is related to the relative velocity v between the frames.

Derivation of Velocity-Rapidity Relation

1. The transformed coordinates are $t' = t \cosh \phi - x \sinh \phi$ and $x' = -t \sinh \phi + x \cosh \phi$.
2. The origin of the primed frame is defined by the condition $x' = 0$.
3. Setting $x' = 0$ implies $-t \sinh \phi + x \cosh \phi = 0$.
4. The velocity of this origin as seen from the unprimed frame is $v = x/t$. Solving for this gives:

$$v = \frac{x}{t} = \frac{\sinh \phi}{\cosh \phi} = \tanh \phi$$

Using the identities $\cosh \phi = \gamma$ and $\sinh \phi = v\gamma$ (where $\gamma = 1/\sqrt{1-v^2}$), the boost transformation can be written in the more familiar form:

$$t' = \gamma(t - vx) \quad (1.15)$$

$$x' = \gamma(x - vt) \quad (1.16)$$

Geometric Interpretation of Boosts

On a spacetime diagram, a Lorentz boost appears as a transformation of the coordinate axes.

- **Transformation of Axes:** The axes of the boosted frame are rotated towards each other, appearing to "scissor" together. The new time axis (t' -axis, defined by $x' = 0$) is the line $t = x/v$. The new space axis (x' -axis, defined by $t' = 0$) is the line $t = vx$.
- **Invariance of the Light Cone:** The paths of light rays, $x = \pm t$, are identical to the paths $x' = \pm t'$. The light cone itself is unchanged by the boost. This is the geometric expression of the postulate that the speed of light is the same in all inertial frames.
- **Relativity of Simultaneity:** Events that are simultaneous in the (t, x) frame (i.e., lie on a line of constant t) are not simultaneous in the (t', x') frame (they do not lie on a line of constant t').

1.4 Vectors

To probe the structure of Minkowski space in more detail, it is necessary to introduce the concepts of vectors and tensors. In spacetime, vectors are four-dimensional and are often referred to as **four-vectors**.

Vectors and the Tangent Space

Beyond dimensionality, the most important concept is that each vector is located at a **given point in spacetime**. The notion of "free" vectors that can be moved around is not useful in the context of curved spaces.

- At each point p in spacetime, we associate the set of all possible vectors located at that point. This set is a vector space known as the **tangent space** at p , or T_p .
- **Geometric Intuition:** The name is inspired by imagining the set of vectors at a point on a curved two-dimensional surface as comprising a plane tangent to that point. It is important to think of these vectors as existing only at that single point.

A vector space is a collection of objects (vectors) that can be added together and multiplied by real numbers in a linear way. For any two vectors V and W and real numbers a and b , we have:

$$(a + b)(V + W) = aV + bV + aW + bW \quad (1.17)$$

Basis Vectors and Components

While a vector is a well-defined geometric object, it is often useful to decompose it into components with respect to a set of **basis vectors**. A basis is a set of vectors that is linearly independent and spans the vector space. For the four-dimensional tangent space at a point in Minkowski space, a basis consists of four vectors.

Let us consider a basis of four vectors $\hat{e}_{(\mu)}$, where $\mu \in \{0, 1, 2, 3\}$. Any abstract vector A can be written as a linear combination of these basis vectors:

$$A = A^\mu \hat{e}_{(\mu)} \quad (1.18)$$

The coefficients A^μ are the **components** of the vector A . It is crucial to distinguish between the abstract geometric entity (the vector A) and its components (A^μ), which are merely numerical coefficients in a chosen basis. The parentheses around the index on the basis vectors, $\hat{e}_{(\mu)}$, serve as a reminder that this is a collection of vectors, not the components of a single vector.

Example: Tangent Vector to a Curve

A standard example of a vector in spacetime is the tangent vector to a curve. A parameterized path through spacetime is specified by the coordinates as a function of a parameter, $x^\mu(\lambda)$. The tangent vector $V(\lambda)$ to this curve has components:

$$V^\mu = \frac{dx^\mu}{d\lambda} \quad (1.19)$$

Transformation Properties of Vectors

Under a Lorentz transformation, the coordinates x^μ change, and therefore the components of the tangent vector must also change. The components transform according to the Lorentz transformation matrix Λ :

$$V^{\mu'} = \Lambda^{\mu'}{}_\nu V^\nu \quad (1.20)$$

However, the abstract vector V itself is a geometric object and must be invariant under a change of coordinates. We can use this fact to derive the transformation properties of the basis vectors.

Derivation of Basis Vector Transformation

1. The vector V is invariant, so it can be expressed in either the old basis $\{\hat{e}_{(\mu)}\}$ or the new basis $\{\hat{e}_{(\nu')}\}$:

$$V = V^\mu \hat{e}_{(\mu)} = V^{\nu'} \hat{e}_{(\nu')}$$

2. Substitute the transformation rule for the components, $V^{\nu'} = \Lambda^{\nu'}{}_\mu V^\mu$, into the equation:

$$V = (\Lambda^{\nu'}{}_\mu V^\mu) \hat{e}_{(\nu')}$$

3. Comparing this with the original expression $V = V^\mu \hat{e}_{(\mu)}$, and noting that this must hold for any vector (i.e., for any set of components V^μ), we can equate the basis vectors:

$$\hat{e}_{(\mu)} = \Lambda^{\nu'}{}_\mu \hat{e}_{(\nu')}$$

4. To find how the new basis vectors are expressed in terms of the old ones, we multiply by the inverse of the Lorentz transformation. We denote the inverse matrix by $\Lambda^{\mu}{}_{\nu'}$, such that $\Lambda^{\mu}{}_{\nu'} \Lambda^{\nu'}{}_\rho = \delta^\mu_\rho$. This gives the transformation rule for the basis vectors:

$$\hat{e}_{(\nu')} = \Lambda^\mu{}_{\nu'} \hat{e}_{(\mu)} \quad (1.21)$$

This "contravariant" transformation behavior (one transforms oppositely to the other) ensures that the full geometric object, $V = V^\mu \hat{e}_{(\mu)}$, is invariant under coordinate changes. This is the foundation of tensor notation.

1.5 Dual Vectors (One-Forms)

Once a vector space is established, one can define an associated vector space of equal dimension known as the **dual vector space**. The dual space to the tangent space T_p is called the **cotangent space** and is denoted T_p^* . Its elements are called **dual vectors**, **covariant vectors**, or **one-forms**.

Definition of a Dual Vector

A dual vector $\omega \in T_p^*$ is a **linear map** from the original vector space T_p to the real numbers \mathbb{R} . This means that for any vectors $V, W \in T_p$ and any real numbers a, b :

$$\omega(aV + bW) = a\omega(V) + b\omega(W) \in \mathbb{R} \quad (1.22)$$

The set of these maps forms a vector space itself. If ω and η are dual vectors, then their linear combination acts on a vector V as:

$$(a\omega + b\eta)(V) = a\omega(V) + b\eta(V) \quad (1.23)$$

Basis and Components of Dual Vectors

To make this construction more concrete, we can introduce a set of **basis dual vectors** $\hat{\theta}^{(\nu)}$ by demanding a specific relationship with the basis vectors $\hat{e}_{(\mu)}$ of the original vector space:

$$\hat{\theta}^{(\nu)}(\hat{e}_{(\mu)}) = \delta_{\mu}^{\nu} \quad (1.24)$$

where δ_{μ}^{ν} is the Kronecker delta. Every dual vector can then be written as a linear combination of this basis. The components are labeled with lower indices:

$$\omega = \omega_{\mu} \hat{\theta}^{(\mu)} \quad (1.25)$$

The action of a dual vector on a vector can then be expressed simply in terms of their components.

Derivation of the Component Action

$$\begin{aligned} \omega(V) &= (\omega_{\mu} \hat{\theta}^{(\mu)})(V^{\nu} \hat{e}_{(\nu)}) \\ &= \omega_{\mu} V^{\nu} \hat{\theta}^{(\mu)}(\hat{e}_{(\nu)}) \quad (\text{by linearity}) \\ &= \omega_{\mu} V^{\nu} \delta_{\nu}^{\mu} \\ &= \omega_{\mu} V^{\mu} \end{aligned}$$

This yields the simple and powerful result:

$$\omega(V) = \omega_{\mu} V^{\mu} \in \mathbb{R} \quad (1.26)$$

This also implies that vectors can be seen as linear maps on dual vectors, $V(\omega) \equiv \omega(V)$, meaning the dual of the dual space is the original vector space itself.

Dual Vector Fields and Transformation Properties

In spacetime, we are interested in fields of vectors and dual vectors. The action of a dual vector field on a vector field is not a single number, but a **scalar** (a function) on spacetime. A scalar is a quantity without indices that is unchanged by Lorentz transformations.

The components of a dual vector must transform in such a way that the scalar product $\omega_{\mu} V^{\mu}$ is invariant. Since we know $V^{\mu'} = \Lambda^{\mu'}_{\nu} V^{\nu}$, the components of a dual vector must transform with the inverse matrix:

$$\omega_{\mu'} = \Lambda^{\nu}_{\mu'} \omega_{\nu} \quad (1.27)$$

Correspondingly, the basis dual vectors transform with the original Lorentz matrix:

$$\hat{\theta}^{(\rho')} = \Lambda^{\rho'}_{\sigma} \hat{\theta}^{(\sigma)} \quad (1.28)$$

This is consistent with the general rule of index placement: upper indices transform with Λ , lower indices transform with the inverse.

Example: The Gradient of a Scalar Function

The simplest and most important physical example of a dual vector is the **gradient** of a scalar function ϕ . The gradient $d\phi$ is the dual vector whose components are the partial derivatives of ϕ :

$$d\phi = \frac{\partial \phi}{\partial x^\mu} \hat{\theta}^{(\mu)} \quad (1.29)$$

The transformation of these components under a Lorentz transformation follows directly from the chain rule for partial derivatives:

$$\frac{\partial \phi}{\partial x^{\mu'}} = \frac{\partial x^\nu}{\partial x^{\mu'}} \frac{\partial \phi}{\partial x^\nu} = \Lambda^\nu{}_{\mu'} \frac{\partial \phi}{\partial x^\nu} \quad (1.30)$$

This is precisely the transformation law for the components of a dual vector. Common shorthand notations for the partial derivative are:

$$\frac{\partial \phi}{\partial x^\mu} = \partial_\mu \phi = \phi_{,\mu} \quad (1.31)$$

The natural action of the gradient (a dual vector) on the tangent vector to a curve $x^\mu(\lambda)$ (a vector) is the ordinary derivative of the function along that curve:

$$\partial_\mu \phi \frac{dx^\mu}{d\lambda} = \frac{d\phi}{d\lambda} \quad (1.32)$$

Here is the LaTeX code for Section 1.6, formatted to match your document's style.

1.6 Tensors

A straightforward generalization of vectors and dual vectors is the notion of a tensor. Just as a dual vector is a linear map from vectors to \mathbb{R} , a tensor is a multilinear map from a collection of dual vectors and vectors to \mathbb{R} .

Definition of a Tensor

A **tensor** T of type (or rank) (k, l) is a multilinear map from k copies of the cotangent space T_p^* and l copies of the tangent space T_p to the real numbers:

$$T : \underbrace{T_p^* \times \cdots \times T_p^*}_{k \text{ times}} \times \underbrace{T_p \times \cdots \times T_p}_{l \text{ times}} \rightarrow \mathbb{R} \quad (1.33)$$

Multilinearity means that the tensor acts linearly in each of its arguments. For a tensor of type (1, 1), we have:

$$T(a\omega + b\eta, cV + dW) = acT(\omega, V) + adT(\omega, W) + bcT(\eta, V) + bdT(\eta, W) \quad (1.34)$$

From this point of view: a scalar is a type (0,0) tensor, a vector is a type (1,0) tensor, and a dual vector is a type (0,1) tensor.

Tensor Products and Basis

To construct a basis for the space of all (k, l) tensors, we define the **tensor product** \otimes . If T is a (k, l) tensor and S is an (m, n) tensor, their tensor product $T \otimes S$ is a $(k+m, l+n)$ tensor defined by its action on a collection of dual vectors ($\omega^{(i)}$) and vectors ($V^{(j)}$):

$$(T \otimes S)(\omega^{(1)}, \dots, V^{(1)}, \dots) = T(\omega^{(1)}, \dots, V^{(1)}, \dots) \times S(\omega^{(k+1)}, \dots, V^{(l+1)}, \dots) \quad (1.35)$$

A basis for the space of (k, l) tensors is formed by taking all tensor products of the basis vectors and dual vectors:

$$\hat{e}_{(\mu_1)} \otimes \cdots \otimes \hat{e}_{(\mu_k)} \otimes \hat{\theta}^{(\nu_1)} \otimes \cdots \otimes \hat{\theta}^{(\nu_l)} \quad (1.36)$$

Tensor Components and Transformation

An arbitrary tensor T can be written as a linear combination of basis tensors, where the coefficients are its **components**:

$$T = T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} \hat{e}_{(\mu_1)} \otimes \dots \otimes \hat{e}^{(\nu_l)} \quad (1.37)$$

The components can be found by acting the tensor on the basis vectors and dual vectors:

$$T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} = T(\hat{e}^{(\mu_1)}, \dots, \hat{e}^{(\mu_k)}, \hat{e}_{(\nu_1)}, \dots, \hat{e}_{(\nu_l)}) \quad (1.38)$$

The transformation of tensor components under a Lorentz transformation $\Lambda^{\mu'}{}_\mu$ is a direct generalization of the vector and dual vector cases: each upper index gets transformed like a vector, and each lower index gets transformed like a dual vector.

$$T^{\mu'_1 \dots \mu'_k}{}_{\nu'_1 \dots \nu'_l} = \Lambda^{\mu'_1}{}_{\mu_1} \dots \Lambda^{\mu'_k}{}_{\mu_k} \Lambda^{\nu_1}{}_{\nu'_1} \dots \Lambda^{\nu_l}{}_{\nu'_l} T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} \quad (1.39)$$

Tensors can also be thought of as maps between other tensors. For example, a (1,1) tensor can map a vector to another vector: $T^\mu{}_\nu : V^\nu \rightarrow T^\mu{}_\nu V^\nu$. This viewpoint emphasizes that tensors are geometric objects with a meaning independent of any coordinate system.

Examples of Tensors in Spacetime

- **The Metric Tensor**, $\eta_{\mu\nu}$: A type (0,2) tensor. Its action on two vectors defines the **inner product** (or scalar product):

$$\eta(V, W) = \eta_{\mu\nu} V^\mu W^\nu = V \cdot W \quad (1.40)$$

The norm of a vector is $V \cdot V$. Unlike in Euclidean space, this is not positive definite. A vector V is called:

- **timelike** if $V \cdot V < 0$.
- **lightlike** or **null** if $V \cdot V = 0$.
- **spacelike** if $V \cdot V > 0$.

- **The Inverse Metric**, $\eta^{\mu\nu}$: A type (2,0) tensor defined by the relation:

$$\eta^{\mu\rho} \eta_{\rho\nu} = \delta_\nu^\mu \quad (1.41)$$

In flat spacetime with Cartesian coordinates, the components of $\eta^{\mu\nu}$ are numerically identical to those of $\eta_{\mu\nu}$.

- **The Levi-Civita Symbol**, $\tilde{\epsilon}_{\mu\nu\rho\sigma}$: A type (0,4) tensor defined as:

$$\tilde{\epsilon}_{\mu\nu\rho\sigma} = \begin{cases} +1 & \text{if } \mu\nu\rho\sigma \text{ is an even permutation of 0123} \\ -1 & \text{if } \mu\nu\rho\sigma \text{ is an odd permutation of 0123} \\ 0 & \text{otherwise} \end{cases} \quad (1.42)$$

Note: This object is technically a "tensor density" and will be treated more carefully in the context of general relativity.

- **The Electromagnetic Field Strength Tensor**, $F_{\mu\nu}$: A type (0,2) antisymmetric tensor that unifies the electric and magnetic fields.

$$F_{\mu\nu} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix} = -F_{\nu\mu} \quad (1.43)$$

This tensor structure elegantly shows how electric and magnetic fields transform into one another under Lorentz boosts.

1.7 Manipulating Tensors

With the definition of tensors established, we can now systematize some of their important properties and operations.

Contraction

Contraction is an operation that turns a tensor of type (k, l) into a tensor of type $(k - 1, l - 1)$. It is performed by summing over one upper and one lower index. For example, if we have a tensor $T^{\mu\nu\rho}_{\sigma\lambda}$, we can contract the second upper index (ν) with the second lower index (λ) to form a new tensor:

$$S^{\mu\rho}_{\sigma} = T^{\mu\nu\rho}_{\sigma\lambda} \quad (1.44)$$

It is only permissible to contract an upper index with a lower index; contracting two indices of the same type does not result in a well-defined tensor. Note also that the order of the remaining indices matters, so that in general:

$$T^{\mu\nu\rho}_{\sigma\nu} \neq T^{\mu\rho\nu}_{\sigma\nu} \quad (1.45)$$

Raising and Lowering Indices

The metric $\eta_{\mu\nu}$ and its inverse $\eta^{\mu\nu}$ can be used to raise and lower the indices on any tensor. This effectively changes a tensor of one type to another. Given a tensor $T^{\alpha\beta}_{\gamma\delta}$, we can define new tensors:

$$T^{\alpha\beta\mu}_{\delta} = \eta^{\mu\gamma} T^{\alpha\beta}_{\gamma\delta} \quad (\text{raising the } \gamma \text{ index}) \quad (1.46)$$

$$T_{\mu}^{\beta}{}_{\gamma\delta} = \eta_{\mu\alpha} T^{\alpha\beta}_{\gamma\delta} \quad (\text{lowering the } \alpha \text{ index}) \quad (1.47)$$

$$T_{\mu\nu}^{\rho\sigma} = \eta_{\mu\alpha} \eta_{\nu\beta} \eta^{\rho\gamma} \eta^{\sigma\delta} T^{\alpha\beta}_{\gamma\delta} \quad (1.48)$$

This operation does not change the horizontal position of an index relative to others. For example, vectors and dual vectors can be converted into one another:

$$V_{\mu} = \eta_{\mu\nu} V^{\nu}, \quad \omega^{\mu} = \eta^{\mu\nu} \omega_{\nu} \quad (1.49)$$

Because the metric and its inverse are truly inverses, a pair of indices being contracted over can be raised and lowered simultaneously without changing the result:

$$A^{\lambda} B_{\lambda} = \eta^{\lambda\rho} A_{\rho} \eta_{\lambda\sigma} B^{\sigma} = \delta_{\sigma}^{\rho} A_{\rho} B^{\sigma} = A_{\sigma} B^{\sigma} \quad (1.50)$$

In a Lorentzian spacetime, the components of a vector V^{μ} are not numerically equal to the components of its dual vector V_{μ} :

$$V^{\mu} = (V^0, V^1, V^2, V^3) \implies V_{\mu} = (-V^0, V^1, V^2, V^3) \quad (1.51)$$

This distinction becomes even more critical in curved spacetimes, where the metric components are more complicated.

Symmetry and Antisymmetry

A tensor is **symmetric** in a set of its indices if it is unchanged upon their exchange. For example, $S_{\mu\nu\rho}$ is symmetric in its first two indices if:

$$S_{\mu\nu\rho} = S_{\nu\mu\rho} \quad (1.52)$$

A tensor is **antisymmetric** (or skew-symmetric) in a set of its indices if it changes sign upon their exchange. For example, $A_{\mu\nu\rho}$ is antisymmetric in its first and third indices if:

$$A_{\mu\nu\rho} = -A_{\rho\nu\mu} \quad (1.53)$$

The metric $\eta_{\mu\nu}$ is symmetric, while the electromagnetic field strength tensor $F_{\mu\nu}$ is antisymmetric.

Symmetrization and Antisymmetrization

Any tensor can be symmetrized or antisymmetrized over any number of its upper or lower indices.

- **Symmetrization** (denoted by round brackets) is the average over all permutations of the indices:

$$T_{(\mu_1 \dots \mu_n)} = \frac{1}{n!} (\text{sum of } T \text{ with indices } \mu_1 \dots \mu_n \text{ permuted}) \quad (1.54)$$

- **Antisymmetrization** (denoted by square brackets) is the alternating sum over all permutations:

$$T_{[\mu_1 \dots \mu_n]} = \frac{1}{n!} (\text{alternating sum of } T \text{ with indices } \mu_1 \dots \mu_n \text{ permuted}) \quad (1.55)$$

For example, for three indices:

$$T_{[\mu\nu\rho]} = \frac{1}{6} (T_{\mu\nu\rho} - T_{\mu\rho\nu} + T_{\rho\mu\nu} - T_{\nu\mu\rho} + T_{\nu\rho\mu} - T_{\rho\nu\mu}) \quad (1.56)$$

For any two indices, a tensor can be decomposed into its symmetric and antisymmetric parts: $T_{\mu\nu} = T_{(\mu\nu)} + T_{[\mu\nu]}$. This simple decomposition does not hold for three or more indices due to the existence of more complex symmetry properties.

The Trace

The **trace** of a (1,1) tensor $X^\mu{}_\nu$ is its contraction, which is a scalar:

$$X = X^\lambda{}_\lambda \quad (1.57)$$

For a (0,2) tensor $Y_{\mu\nu}$, the trace is defined as raising one index and then contracting:

$$Y = Y^\lambda{}_\lambda = \eta^{\mu\nu} Y_{\mu\nu} \quad (1.58)$$

Note that this is not the sum of the diagonal components of $Y_{\mu\nu}$. For example, the trace of the metric itself is:

$$\eta^{\mu\nu} \eta_{\mu\nu} = \delta^\mu_\mu = 4 \quad (1.59)$$

A Note on Partial Derivatives

In the special case of **flat spacetime with inertial coordinates**, the partial derivative of a (k, l) tensor is a well-defined $(k, l + 1)$ tensor. For example, if $T^\mu{}_\nu$ is a tensor, then so is:

$$S_\alpha{}^\mu{}_\nu = \partial_\alpha T^\mu{}_\nu \quad (1.60)$$

Important: This property will **fail** in more general (curved) spacetimes or non-inertial coordinate systems. We will later introduce a **covariant derivative** to replace the partial derivative in such cases. The one exception is the gradient of a scalar, $\partial_\alpha \phi$, which is always a well-defined (0,1) tensor. In flat spacetime, partial derivatives commute:

$$\partial_\mu \partial_\nu (\dots) = \partial_\nu \partial_\mu (\dots) \quad (1.61)$$

1.8 Maxwell's Equations

With the necessary tensor know-how accumulated, we can now illustrate these concepts using Maxwell's equations of electrodynamics. This exercise demonstrates the power and economy of the tensor formalism.

Maxwell's Equations in Traditional Notation

In 19th-century 3-vector notation, Maxwell's equations are:

$$\nabla \times \mathbf{B} - \partial_t \mathbf{E} = \mathbf{J} \quad (1.62)$$

$$\nabla \cdot \mathbf{E} = \rho \quad (1.63)$$

$$\nabla \times \mathbf{E} + \partial_t \mathbf{B} = 0 \quad (1.64)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (1.65)$$

Here, \mathbf{E} and \mathbf{B} are the electric and magnetic field 3-vectors, \mathbf{J} is the current 3-vector, and ρ is the charge density. These equations are, in fact, invariant under Lorentz transformations, but they do not look obviously so. Tensor notation makes this invariance manifest.

From 3-Vector to Tensor Formulation

The first step is to rewrite the equations in component notation and then combine them into a single tensor object. We begin by defining the current 4-vector $J^\mu = (\rho, J^x, J^y, J^z)$ and recalling the definition of the antisymmetric field strength tensor $F_{\mu\nu}$.

The goal is to show that the four equations above are equivalent to the following two tensor equations.

1. The Inhomogeneous Equations (with sources)

Derivation:

1. The first two of Maxwell's equations, $\nabla \cdot \mathbf{E} = \rho$ and $\nabla \times \mathbf{B} - \partial_t \mathbf{E} = \mathbf{J}$, involve the sources of the fields, ρ and \mathbf{J} .
2. In component notation (with $x^0 = t$), these are:

$$\partial_i E^i = J^0 \quad \text{and} \quad \tilde{\epsilon}^{ijk} \partial_j B_k - \partial_0 E^i = J^i$$

where $\tilde{\epsilon}^{ijk}$ is the three-dimensional Levi-Civita symbol.

3. We can express the components of the field strength tensor with *upper* indices, $F^{\mu\nu} = \eta^{\mu\alpha}\eta^{\nu\beta}F_{\alpha\beta}$, in terms of the electric and magnetic fields (1.43) :

$$F^{0i} = E^i, \quad F^{ij} = \tilde{\epsilon}^{ijk} B_k \quad (1.66)$$

4. Substituting these into the component equations gives:

$$\partial_i F^{0i} = J^0 \quad \text{and} \quad \partial_j F^{ji} - \partial_0 F^{0i} = J^i \quad (1.67)$$

5. Using the antisymmetry of the field strength tensor (i.e., $F^{\mu\nu} = -F^{\nu\mu}$, which implies $F^{00} = 0$ and $\partial_0 F^{00} = 0$), these two distinct equations can be combined into a single, compact 4-vector equation.

The result is the first of the covariant Maxwell equations:

$$\partial_\mu F^{\nu\mu} = J^\nu \quad (1.68)$$

2. The Homogeneous Equations (sourceless)

A similar line of reasoning reveals that the third and fourth of Maxwell's equations, $\nabla \cdot \mathbf{B} = 0$ and $\nabla \times \mathbf{E} + \partial_t \mathbf{B} = 0$, can also be written in a single tensor form. The result is:

$$\partial_{[\mu} F_{\nu\lambda]} = 0 \quad (1.69)$$

Due to the antisymmetry of $F_{\mu\nu}$, this is equivalent to the expanded form:

$$\partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0 \quad (1.70)$$

Significance of the Covariant Formulation

The four traditional Maxwell equations are replaced by two compact tensor equations, demonstrating the economy of the notation. More importantly, both sides of equations (1.68) and (1.69) are well-defined tensors. An equation that equates two tensors of the same type is said to be **manifestly covariant**. If such an equation is true in one inertial frame, it must be true in any Lorentz-transformed frame. This is why tensors are the natural language for relativity: they allow physical laws to be expressed in a frame-independent way.

- The tensor form is referred to as the **covariant form** of Maxwell's equations.
- The original 3-vector form is referred to as **noncovariant**.

1.9 Energy and Momentum

This section reviews the physics of energy and momentum in Minkowski spacetime, starting with single particles and extending to continuous media.

Particle Kinematics

The worldline of a massive particle is a timelike curve. Since the proper time τ is the time measured by a clock traveling along such a worldline, it is the natural parameter to use for the path, $x^\mu(\tau)$.

Four-Velocity and Four-Momentum

- **Four-Velocity:** The tangent vector to a particle's worldline parameterized by proper time is the **four-velocity**, U^μ .

$$U^\mu = \frac{dx^\mu}{d\tau} \quad (1.71)$$

From the definition of proper time, $d\tau^2 = -\eta_{\mu\nu}dx^\mu dx^\nu$, the four-velocity is automatically normalized:

$$\eta_{\mu\nu} U^\mu U^\nu = -1 \quad (1.72)$$

This reflects the fact that a massive particle always travels through spacetime at a constant "rate." In the particle's rest frame, its four-velocity has components $U^\mu = (1, 0, 0, 0)$.

- **Four-Momentum:** The **momentum four-vector** is defined as the rest mass m (an invariant scalar) times the four-velocity.

$$p^\mu = mU^\mu \quad (1.73)$$

The components of the four-momentum are the energy and the three-momentum: $p^\mu = (E, p^x, p^y, p^z)$.

- In the particle's rest frame, $p^\mu = (m, 0, 0, 0)$, giving $E = m$ (or $E = mc^2$ in standard units).
- For a particle moving with three-velocity $v = dx/dt$ along the x-axis, the components are:

$$p^\mu = (\gamma m, v\gamma m, 0, 0) \quad (1.74)$$

where $\gamma = 1/\sqrt{1-v^2}$.

- The norm of the four-momentum is an invariant:

$$p_\mu p^\mu = \eta_{\mu\nu} p^\mu p^\nu = -m^2 \quad (1.75)$$

This gives the relativistic energy-momentum relation: $E^2 - p^2 = m^2$, or

$$E = \sqrt{m^2 + p^2} \quad (1.76)$$

Four-Force

The relativistic analogue of Newton's Second Law introduces the **force four-vector** f^μ :

$$f^\mu = m \frac{d^2 x^\mu}{d\tau^2} = \frac{dp^\mu}{d\tau} \quad (1.77)$$

For electromagnetism, the Lorentz force law has the unique tensorial generalization:

$$f^\mu = q F^\mu_\lambda U^\lambda \quad (1.78)$$

where q is the particle's charge and F^μ_λ is the electromagnetic field strength tensor.

The Energy-Momentum Tensor

For extended systems like fluids or fields, we use the **energy-momentum tensor** (or stress-energy tensor), $T^{\mu\nu}$.

- **Physical Meaning:** $T^{\mu\nu}$ is the flux of the μ -component of four-momentum (p^μ) across a surface of constant x^ν .

- **Components in the Rest Frame:**

- $T^{00} = \rho$ (Energy density)
- $T^{0i} = T^{i0}$ (Momentum density / Energy flux)
- T^{ij} (Stress, i.e., flux of i -momentum in the j -direction). The diagonal terms T^{ii} (no sum) represent pressure:

$$p_i = T^{ii} \quad (1.79)$$

Examples of Energy-Momentum Tensors

- **Dust:** A collection of non-interacting particles all at rest with respect to each other. It is characterized by a single four-velocity field U^μ and a rest-frame number density n . The **number-flux four-vector** is:

$$N^\mu = n U^\mu \quad (1.80)$$

If each particle has mass m , the rest-frame energy density is $\rho = mn$. The energy-momentum tensor for dust is:

$$T_{\text{dust}}^{\mu\nu} = \rho U^\mu U^\nu \quad (1.81)$$

Dust has zero pressure.

- **Perfect Fluid:** A fluid that is completely characterized by its rest-frame energy density ρ and an isotropic rest-frame pressure p . In its rest frame, its energy-momentum tensor is:

$$T_{\text{rest}}^{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} \quad (1.82)$$

The general, covariant expression that is valid in any inertial frame is:

$$T^{\mu\nu} = (\rho + p) U^\mu U^\nu + p \eta^{\mu\nu} \quad (1.83)$$

The physics of a specific perfect fluid is determined by an **equation of state**, $p = p(\rho)$. An important example is **vacuum energy**, for which $p_{\text{vac}} = -\rho_{\text{vac}}$, leading to $T_{\text{vac}}^{\mu\nu} = -\rho_{\text{vac}} \eta^{\mu\nu}$.

Conservation of Energy-Momentum

The energy-momentum tensor for a closed system is conserved, which is expressed as the vanishing of its four-divergence:

$$\begin{aligned}\partial_\mu T^{\mu\nu} &= \partial_\mu ((\rho + p)U^\mu U^\nu) + \partial^\nu p = 0 \\ \implies 0 &= \partial_\mu(\rho + p)U^\mu U^\nu + (\rho + p)U^\mu \partial_\mu U^\nu + (\rho + p)U^\nu \partial_\mu U^\mu + \partial^\nu p\end{aligned}\quad (1.84)$$

This is a set of four equations. The $\nu = 0$ component expresses conservation of energy, while the $\nu = k$ components express conservation of momentum.

Analysis for a Perfect Fluid: Applying the conservation law to the perfect fluid tensor yields:

$$\partial_\mu((\rho + p)U^\mu U^\nu) + \partial^\nu p = 0$$

This equation can be projected into parts parallel and perpendicular to the fluid's four-velocity U^μ .

Projection Parallel to U^μ (Energy Conservation): Contracting with U_ν (multiplying the equation (1.84)) and setting the result to zero gives the relativistic equation for energy conservation. using the fact that $U_\nu U^\nu = -1$ is constant along the flow, we have:

$$U_\nu \partial_\mu U^\nu = \frac{1}{2} \partial_\mu(U_\nu U^\nu) = 0 \quad (1.85)$$

Which can be shown as follows:

$$\begin{aligned}\partial_\mu(U_\nu U^\nu) &= U_\nu \partial_\mu U^\nu + \partial_\mu U_\nu U^\nu \\ &= U_\nu \partial_\mu U^\nu + \partial_\mu(g_{\nu\lambda} U^\lambda) U^\nu \\ &= U_\nu \partial_\mu U^\nu + \partial_\mu U^\lambda U_\lambda \\ &= 2U_\nu \partial_\mu U^\nu\end{aligned}\quad (1.86)$$

By Changing the dummy index λ to ν in the last term and by the fact that they commute, we see that it is identical to the first term. Thus, $U_\nu \partial_\mu U^\nu = 0$.

$$\begin{aligned}U_\nu \partial_\mu T^{\mu\nu} &= \partial_\mu(\rho + p)U^\mu U_\nu U^\nu + (\rho + p)U^\mu U_\nu \partial_\mu U^\nu + (\rho + p)U_\nu U^\nu \partial_\mu U^\mu + U_\nu \partial^\nu p \\ &= -\partial_\mu(\rho + p)U^\mu - (\rho + p)\partial_\mu U^\mu - U^\nu \partial_\nu p\end{aligned}\quad (1.87)$$

In the non-relativistic limit ($|v^i| \ll 1, p \ll \rho, \gamma \approx 1$), this becomes the familiar **continuity equation**: by applying the limits we get:

$$\begin{aligned}U_\nu \partial_\mu T^{\mu\nu} &\approx \partial_\mu(\rho)U^\mu + \rho \partial_\mu U^\mu \\ &= \partial_\mu(\rho U^\mu) \\ &= \partial_0(\rho U^0) + \partial_i(\rho U^i) \\ &= \partial_0(\rho\gamma) + \partial_i(\rho\gamma v^i) \\ &= \partial_t\rho + \nabla \cdot (\rho\mathbf{v})\end{aligned}\quad (1.88)$$

Projection Perpendicular to U^μ (Momentum Conservation): Projecting the conservation law onto the space orthogonal to U^μ (using the projection tensor $P^\sigma_\nu = \delta^\sigma_\nu + U^\sigma U_\nu$) gives the equations for momentum conservation. In the non-relativistic limit, this yields the **Euler equation** for fluid dynamics:

$$\rho[\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{v}] = -\nabla p \quad (1.89)$$

1.10 Classical Field Theory

When we transition to general relativity, the metric $\eta_{\mu\nu}$ is promoted to a dynamical tensor field, $g_{\mu\nu}(x)$. GR is thus a particular example of a classical field theory. We can build intuition by first considering classical fields on the fixed background of flat spacetime.

From Mechanics to Field Theory

The dynamics of a system can be derived from the **principle of least action**.

- **Point-Particle Mechanics:** For a single particle with coordinate $q(t)$, the action S is the time integral of the Lagrangian, $L(q, \dot{q})$.

$$S = \int L(q, \dot{q}) dt \quad (1.90)$$

The Lagrangian is typically the kinetic energy minus the potential energy, $L = K - V$. The trajectory that a particle follows is one that makes the action stationary ($\delta S = 0$), which leads to the **Euler-Lagrange equations**:

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) = 0 \quad (1.91)$$

- **Classical Field Theory:** We replace the single coordinate $q(t)$ with a set of spacetime-dependent fields, $\Phi^i(x^\mu)$. The action becomes an integral of the **Lagrange density**, \mathcal{L} , over all of spacetime.

$$S = \int \mathcal{L}(\Phi^i, \partial_\mu \Phi^i) d^4x \quad (1.92)$$

The Lagrange density \mathcal{L} must be a Lorentz scalar.

The Euler-Lagrange Equations for Fields

Derivation:

1. We require the action S to be stationary under an infinitesimal variation of the fields, $\Phi^i \rightarrow \Phi^i + \delta\Phi^i$.
2. The change in the action, δS , is found by expanding the Lagrangian:

$$\delta S = \int d^4x \left[\frac{\partial \mathcal{L}}{\partial \Phi^i} \delta\Phi^i + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi^i)} \partial_\mu (\delta\Phi^i) \right]$$

3. We integrate the second term by parts, $\int u dv = uv - \int v du$. Let $u = \frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi^i)}$ and $dv = \partial_\mu (\delta\Phi^i)$.

$$\int d^4x \frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi^i)} \partial_\mu (\delta\Phi^i) = - \int d^4x \left[\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi^i)} \right) \right] \delta\Phi^i + \text{surface term}$$

4. Assuming the variations vanish at the boundary, the surface term is zero. Substituting this back gives:

$$\delta S = \int d^4x \left[\frac{\partial \mathcal{L}}{\partial \Phi^i} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi^i)} \right) \right] \delta\Phi^i$$

5. For $\delta S = 0$ for any arbitrary variation $\delta\Phi^i$, the expression in the brackets must vanish.

This gives the Euler-Lagrange equations for a field theory:

$$\frac{\partial \mathcal{L}}{\partial \Phi^i} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi^i)} \right) = 0 \quad (1.93)$$

Example 1: Real Scalar Field

Consider a single real scalar field $\phi(x^\mu)$. The Lagrangian density is constructed by generalizing "kinetic minus potential energy."

$$\mathcal{L} = -\frac{1}{2}\eta^{\mu\nu}(\partial_\mu\phi)(\partial_\nu\phi) - V(\phi) \quad (1.94)$$

Here, the first term is the Lorentz-invariant combination of kinetic ($\frac{1}{2}\dot{\phi}^2$) and gradient ($-\frac{1}{2}(\nabla\phi)^2$) energy density. Applying the Euler-Lagrange equations:

- $\frac{\partial\mathcal{L}}{\partial\phi} = -\frac{dV}{d\phi}$
- $\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} = -\eta^{\mu\nu}\partial_\nu\phi$

The equation of motion is:

$$\eta^{\mu\nu}\partial_\mu\partial_\nu\phi - \frac{dV}{d\phi} = 0 \quad \text{or} \quad \square\phi - \frac{dV}{d\phi} = 0 \quad (1.95)$$

where $\square \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu$ is the d'Alembertian operator. For a simple harmonic oscillator potential, $V(\phi) = \frac{1}{2}m^2\phi^2$, this becomes the famous **Klein-Gordon equation**:

$$\square\phi + m^2\phi = 0 \quad (1.96)$$

Example 2: Electromagnetism

The dynamical field for electromagnetism is the vector potential A_μ . Physical observables are expressed in terms of the gauge-invariant field strength tensor:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (1.97)$$

The Lagrangian density for electromagnetism coupled to a 4-current source J^μ is:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + A_\mu J^\mu \quad (1.98)$$

Applying the Euler-Lagrange equations for the field A_ν :

- $\frac{\partial\mathcal{L}}{\partial A_\nu} = J^\nu$
- A careful calculation shows $\frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)} = -F^{\mu\nu}$

Plugging these into the Euler-Lagrange equation gives:

$$J^\nu - \partial_\mu(-F^{\mu\nu}) = 0$$

Using the antisymmetry of $F^{\mu\nu}$, this becomes the inhomogeneous Maxwell's equation:

$$\partial_\mu F^{\nu\mu} = J^\nu \quad (1.99)$$

The other Maxwell equation, $\partial_{[\mu}F_{\nu\lambda]} = 0$, is automatically satisfied by the definition of $F_{\mu\nu}$ in terms of A_μ , since partial derivatives commute.

Energy-Momentum Tensors from the Action

The action principle provides a direct procedure for deriving the energy-momentum tensor for any field theory. For the examples above, the results are:

- **Scalar Field:**

$$T_{\text{scalar}}^{\mu\nu} = (\partial^\mu\phi)(\partial^\nu\phi) - \eta^{\mu\nu} \left[\frac{1}{2}(\partial_\lambda\phi)(\partial^\lambda\phi) + V(\phi) \right] \quad (1.100)$$

- **Electromagnetism:**

$$T_{\text{EM}}^{\mu\nu} = F^{\mu\lambda}F^\nu{}_\lambda - \frac{1}{4}\eta^{\mu\nu}F^{\lambda\sigma}F_{\lambda\sigma} \quad (1.101)$$

Using the equations of motion for each theory, these energy-momentum tensors can be shown to be conserved ($\partial_\mu T^{\mu\nu} = 0$).