Relativity and Classical Field Equations

by
Ekrem Demirboğa

Supervisor: Tekin Dereli

Department of Physics
Koç University

Relativity and Classical Field Equations

Ekrem Demirboğa

Abstract

Einstein has shown in his special theory of relativity that the time and Newtonian 3-dimensional space are not absolute on their own. We live and give our physical theories in a 4-dimensional Minkowski space-time. The aim of the present study is first to learn the differential geometry of Minkowski space-time. Then tensors as multi-linear maps on space-time will be defined and the algebra of tensors in general will be given. Equipped with this mathematical background, it is planned to learn physically relevant classes of relativistic wave equations.Namely, (i) Klein-Gordon equation satisfied by real or complex scalar fields; (ii) Maxwell equations satisfied by a real massless vector field (i.e the photon field) and its gauge covariance; Proca equations satisfied by massive vector fields; finally (iv) Einstein field equations satisfied by second rank covariant,symmetric tensor field. In order to reach this final stage, Einstein tensor in a curved pseudoRiemannian space-time will be constructed. It is purely geometrical and put on the left hand side of the field equations. The right hand side of Einstein's equations is reserved for distributions of matter and radiation. We will discuss the construction of conserved stress-energy-momentum tensors of the scalar and vector fields given above.

Contents

1 Manifolds

Introduction

In physics and mathematics, manifolds are one of the most important concepts. Relativity is a theory of geometry. Therefore it requires to deal with some ideas like manifolds,vectors,tensors,etc. In this chapter, a general summary of these topics will be given, then we will briefly discuss the special relativity.

1.1 Definition of a Manifold

Manifolds are basically topological spaces that look like euclidean space in local regions. In other words for each point of a manifold we have a neighborhood on the manifold which resembles the euclidean space $\mathbb{R}^{\mathbf{n}}$. Manifolds might have some complicated topological properties in general but it is constructed by smoothly sewing these euclidean local regions together. It is important to point out that the dimension n of every local region must be the same. Therefore we call this type of a manifold an $n-$manifold and usually denote by $\mathcal{M}$.

In order to give a more rigorous definition we require some other definitions:

A map $\phi : \mathbb{R}^n \to \mathbb{R}^m$ simply takes $n-$tuple $(x^1, \ldots, x^n)$ to $m-$tuple $(x^1, \ldots, x^m)$ and can can be written as a collection of m functions ϕ^i of n variables. If these all these functions has at least have pth derivative, then we refer to this map as $\mathbf{C^p}$. If a map can be differentiated infinitely many times, it is called C^∞.

We call two set A and B **diffeomorphic** if there exist a C^∞ map $\phi : A \to B$ as well as an inverse $\phi^{-1} : B \to A$; where the map ϕ is called diffeomorphic.

A **coordinate chart** or **coordinate system** is a way of expressing points on a Manifold $\mathcal{M}$ belong to a small neighborhood. Technically, it is a map $\phi : U \to \mathbb{R}^n$ such that the image of ϕ is open in $\mathbb{R}^n$ where n is the dimension of the manifold. Basically, as in usual coordinate systems it assigns a point in $\mathbb{R}^n$ for every point on a manifold. (FIGURE)

Finally, a C^∞ **atlas**, as the name suggests, is a collection of coordinate systems U_α, ϕ_α with the following conditions:

I Union of U_α's are equal to M,i.e., U covers M.

II They are smoothly sewn together. More precisely, if two charts overlaps, $U_\alpha \cap U_\beta \neq 0$, then the map $(\phi_\alpha \phi_\beta^{-1})$ takes points in $\phi_\beta(U_\alpha \cap U_\beta)$ onto an open set $\phi_\alpha(U_\alpha \cap U_\beta)$, and all these maps should be C^∞. This can be easily understood from the figure.

Now we have everything we need to give a proper definition of a manifold: a C^∞ **n−manifold** $\mathcal{M}$ is a set along with a maximal atlas which contains every possible compatible chart. Also, it is important to point out that this definition is intrinsic, i.e., it does not require a higher dimensional space and embedding the manifold into that space. As we will see in next chapter and chapter 7, Spacetime is a 4-dimensional manifold and is not embedded in a higher dimensional spaces, although, there are some other theories suggest that.

1.2 Vectors and Tensors

To investigate the structure of a manifold, (or in particular the spacetime) in more detail, we need to introduce concepts of scalars, vectors and tensors. A **scalar** ϕ is a quantity that does not change when we change the coordinates.

$$\phi \to \phi^{'} = \phi \tag{1.1}$$

Where the prime denotes the transformed coordinates.

A *vector* is an element of a vector space which is a set of objects satisfying some certain axioms along with two defined operator, namely multiplication and addition. There is two type of vectors we will be dealing with.

Consider the change of coordinates $x^\mu \to (x^\mu)^{'}$ where $\mu = 0, 1, \ldots, n$ and n is the dimension of the manifold. Then, V is called a **contravariant**

vector (or just **vector**) if it satisfies the following transformation

$$V^{\mu} \rightarrow (V^{\mu})' = \frac{\partial (x^{\mu})'}{\partial x^{j}} V^{j} \tag{1.2}$$

And is called **Covariant Vector** (or **Covector** or **Dual Vector**) if it satisfies the following transformation

$$V_{\mu} \rightarrow (V_{\mu})' = \frac{\partial (x^{\mu})}{\partial (x^{j})'} V^{j} \tag{1.3}$$

The indices placement were not arbitrary, we use upper indices for contravariant vectors and lower indices for covariant vectors.

The name Contravariant and covariant comes from their response to the change in the basis vectors. For example, think of simple Cartesian coordinate system. If we double the length of the basis vectors, then the components of the vectors in the new basis will be half in size, therefore they "*contra*" varied. On the other hand, since covariant vectors transform in the opposite way, i.e., they uses the inverse transformation matrix, their size would also be doubled up and hence "*co*" varied. We will be mostly using vectors instead of contravariant and dual vector for covariant.

However, once we introduce curvature to spacetime, the idea of vectors as arrows pointing from one location to another (and dual vectors are usually visualized as bunch of stacks rather than arrows.) becomes vague. It is more appropriate to think that every vector is located at one point in spacetime. Therefore, we come up with the idea of **Tangent Space** T_p of a point p. It is the set of all possible vectors located at the point p. We may think of it as a surface that is tangent to a $2-$sphere as in figure. But This idea requires embedding of higher dimensions, therefore it really is a better idea to think as every vector located at one point.

(FIGURE)

We also have a dual vector space T_p^* to a tangent space T_p called **Cotangent Space** where dual vectors live. It is the space of all linear maps from tangent vector space T_p to real numbers $\mathbb{R}$. In other words, a dual vector $\omega \in T_p^*$ acts as a function which takes vector and gives a real number

$$\omega : T_p \rightarrow \mathbb{R} \tag{1.4}$$

For example, gradient of a scalar function f, denoted $\mathrm{d}f$ is a dual vector as we can check, it transforms according to 1.3. This notation may seem strange, but it will become more clear as we discuss exterior derivatives in Chapter 5.

So far we have talked about tangent spaces and cotangent spaces but how do we construct these spaces without an embedding to higher dimensions? We can talk about all possible curves γ that passes through p and we can say that tangent space is just the collection of the tangent vectors $dx^\mu/d\lambda$ to those curves at p. However, this approach depends on coordinate system. In order to achieve a coordinate independent definition that we consider the following: let $\mathcal{F}$ be the space of all smooth functions on $\mathcal{M}$. Then, we claim that tangent space T_p is the space of directional derivatives operators along curves through p, i.e., instead of $dx^\mu/d\lambda$ we use $d/d\lambda$ as vectors. And as a basis for these space, we consider the most straightforward n directional derivatives, partial derivatives ∂_μ which is called *coordinate basis.* Also with this approach Transformation laws for the coordinate system becomes more obvious

$$\partial_\mu^{'} = \frac{\partial x^\mu}{\partial x^{\mu'}}\partial_\mu \tag{1.5}$$

Tensors are just the collection of vector and dual vectors combined together using tensor product. More precisely, a tensor T of rank (l,m) is a multilinear map from a collection of vectors and dual vectors to $\mathbb{R}$

$$T : \underbrace{T_p^* \times \cdots \times T_p^*}_{l} \times \underbrace{T_p \times \cdots \times T_p}_{m} \rightarrow \mathbb{R} \tag{1.6}$$

and denoted like $T^{a_1\cdots a_l}{}_{b_1\cdots b_m}$ where $\times$ means Cartesian product. Multilinearity means that tensors acts each of its components linearly. For example if we have a rank $(1,1)$ tensor T, Multilinearity states that

$$\begin{aligned} T(a\omega_1 + b\omega_2, cV_1 + dV_2) =& acT(\omega_1, V_1) + adT(\omega_1, V_2) \\ &+ bcT(\omega_2, V_1) + bdT(\omega_2, V_2) \end{aligned} \tag{1.7}$$

Having constructed this definition, we can now say that a scalar is a $(0,0)$ tensor, a vector is a $(1,0)$ tensor, dual vector is a $(0,1)$ tensor. Also it is straightforward to show that a rank (l,m) tensor $T^{a_1\cdots a_l}{}_{b_1\cdots b_m}$ transforms as

$$T^{a_1^{'}\cdots a_l^{'}}{}_{b_1^{'}\cdots b_m^{'}} = \frac{\partial x^{\mu_1^{'}}}{\partial x^{\mu_1}} \cdots \frac{\partial x^{\mu_l^{'}}}{\partial x^{\mu_l}} \frac{\partial x^{\mu_1}}{\partial x^{\mu_1^{'}}} \cdots \frac{\partial x^{\mu_m}}{\partial x^{\mu_m^{'}}} T^{a_1\cdots a_l}{}_{b_1\cdots b_m} \tag{1.8}$$

And we define a tensor product of a rank (k, l) tensor T and a rank (m, n) tensor G as a new rank $(k+m, l+n)$ tensor $T \otimes G$ where $\otimes$ denotes the tensor product. It is also important ot point out that $T^{a_1 \cdots a_l}{}_{b_1 \cdots b_m}$ is just the representation of the components of the tensor. In other words

$$T = T^{a_1 \cdots a_l}{}_{b_1 \cdots b_m} \, \hat{e}_{(a1)} \otimes \cdots \otimes \hat{e}_{(al)} \otimes \hat{\theta}_{b_1} \otimes \cdots \otimes \hat{\theta}_{b_m} \tag{1.9}$$

where $\hat{e}$s are the basis vectors and $\hat{\theta}$ are the dual basis vectors. But we usually omit the basis vectors while writing a tensor and just write it as its components. This is also true for vectors and dual vectors.

An example of a tensor is the **metric tensor** which is (0,2) tensor denoted by η in special relativity and g in general relativity for more general cases. It is a mathematical object that measures distance. It can be thought of as a matrix whose elements the dot product of basis vectors. Although, it has some other deeper meanings it is enough for us to know this much for now.

1.3 Differential Forms

Differential Forms are a special class of Tensors. Simply, an anti-symmetric $(0, p)$ tensor is called a $-p$ form. The space of all $-p$ forms is denoted by Λ^p. Therefore, scalars are 0-forms and dual vectors are simply one forms.

We mentioned that directional derivatives are vectors and gradients are dual vectors. A dual vector acting on a vector should produce a real number from definition. Therefore, looking at the action of a gradient of a scalar $\mathrm{d}f$ on a vector $\frac{d}{d\lambda}$

$$\mathrm{d}f \frac{d}{d\lambda} = \frac{df}{d\lambda} \tag{1.10}$$

Just as partial derivatives along coordinate axes provide a basis for vectors, gradients of the coordinate functions x^μ provide a basis for dual vectors. In general, we construct basis for cotangent space by demanding $\theta^\mu(e_\nu) = \delta^\nu_{nu}$. Hence,

$$\mathrm{d}x^\mu(\partial_{[\nu}) = \frac{\partial x^\mu}{\partial x^\nu} = \delta^\mu_\nu \tag{1.11}$$

We can expand any one form into its components as $\omega = \omega_\mu \mathrm{d}x^\mu$.

Also Given a -p form A and a -q form B, we can construct a (p+q) form by **wedge product** $A \wedge B$ which means taking their antisymmetrized tensor product:

$$(A \wedge B)_{\mu_1 \dots \mu_{p+q}} = \frac{(p+q)!}{p!q!} A_{[\mu_1 \dots \mu_p} B_{\mu_{p+1} \dots \mu_{p+q}]}. \qquad (1.12)$$

1.4 Integration on Manifolds

Let us introduce the levi-civita symbol defined as

$$\tilde{\epsilon}_{\mu_1 \mu 2 \dots \mu_n} = \begin{cases} +1 & \text{even permutation} \\ -1 & \text{odd permutation} \\ 0 & \text{otherwise} \end{cases} \qquad (1.13)$$

~ symbol stands to point out that it is not a tensor. Given any $n \times n$ matrix $M^{\mu}_{\mu'}$ we can write

$$\tilde{\epsilon}_{\mu'_1 \mu'_2 \dots \mu'_n} |M| = \tilde{\epsilon}_{\mu_1 \mu 2 \dots \mu_n} M^{\mu_1}_{\mu_1} \cdots M^{\mu_n}_{\mu_n} \qquad (1.14)$$

If we set $M^{\mu}_{\mu'} = \frac{\partial x^{\mu}}{\partial x^{\mu'}}$ we end up with

$$\tilde{\epsilon}_{\mu'_1 \mu'_2 \dots \mu'_n} = \left| \frac{\partial x^{\mu}}{\partial x^{\mu'}} \right| \tilde{\epsilon}_{\mu_1 \mu 2 \dots \mu_n} \frac{\partial x^{\mu_1}}{\partial x^{\mu'_1}} \cdots \frac{\partial x^{\mu_n}}{\partial x^{\mu'_n}} \qquad (1.15)$$

Hence, it transform a tensor except for the determinant factor. Objects transforming this way is called **tensor densities**. The power of the determinant is called the weight of the tensor density and in this case it is one. If we multiply by $|g|^{w/2}$ with a tensor density where w is the weight of the tensor density we get a tensor. Therefore we can define a levi-civita tensor as

$$\epsilon_{\mu_1 \mu 2 \dots \mu_n} = \sqrt{|g|} \epsilon_{\mu_1 \mu \tilde{2} \dots \mu_n}. \qquad (1.16)$$

Since we defined tensor densities, we can now define integrals on manifolds. An integral over an n dimensional region $\Sigma \subset M$ is a map from an n-Sform field ω to real numbers

$$\int_{\Sigma} : \omega \to \mathbb{R} \qquad (1.17)$$

We also need to define a volume element $d^n x = \mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^{n-1}$. But this definition is not a tensor, it is a tensor density. Therefore we make it into a tensor by multiplying with $\sqrt{|g|}$.

$$\sqrt{|g|}d^n x = \sqrt{|g|}\mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^{n-1} \tag{1.18}$$

In fact, if we write the volume element in explicit basis one forms, we see that it is equal to levi-civita tensor :

$$\begin{aligned}
\epsilon =& \epsilon_{\mu_1\mu 2\ldots\mu_n} \mathrm{d}x^{\mu_1} \otimes \cdots \otimes \mathrm{d}x^{\mu_n} && (1.19)\\
=& \frac{1}{n!}\sqrt{|g|}\tilde{\epsilon}_{\mu_1\mu_2\ldots\mu_n} \mathrm{d}x^{\mu_1} \wedge \cdots \wedge \mathrm{d}x^{\mu_n} && (1.20)\\
=& \sqrt{|g|}\mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^{n-1} && (1.21)\\
=& \sqrt{|g|}d^n x && (1.22)
\end{aligned}$$

Therefore we can write the integral of a scalar function ϕ in n dimensional manifold as

$$\int \phi(x)\sqrt{|g|}d^n x \tag{1.23}$$

as well as

$$\int \phi(x)\epsilon \tag{1.24}$$

1.5 Review of Special Relativity

In Special relativity, we have 4 dimensional manifold called Minkowski spacetime. Minkowski spacetime is flat, i.e., no curvature exist (Although we do not explain the proper definition of curvature, intuitively we can think of what curvature is.). Therefore, in cartesian coordiantes, we can write the metric η for Minkowski spacetime as

$$\eta_{\mu\nu} = diag(-1, 1, 1, 1) \tag{1.25}$$

where *diag* means the diagonal elements. In other words, we can write

$$ds^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = -dt^2 + dx^2 + dy^2 + dz^2 \tag{1.26}$$

Usually, many books on SR refers to ds^2 as the line element or spacetime interval as if these dx's represent the so called infinitesimal distances. This is not a problem in flat spacetime. However, as we mentioned in previous chapters, these dx's are nothing but basis for cotangent vector space. Therefore writing $\eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$ is just expressing the metric in its components. Hence ds^2 or η are both the same thing.

$$\eta = ds^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu \tag{1.27}$$

1.5.1 Lorentz Transformations

Let us now consider the coordinate transformation in spacetime. We are looking for the relations between to inertial frames. This kind of transformations are called **Lorentz Transformations**. For example **translations** which just shifts the coordinates in space or time are Lorentz Transformation.

Now, Let's look at more general Lorentz Transformations including spatial **rotations** and offsets by a constant velocity vector, or **boosts**. These are linear transformations, hence can be tought as multiplying the coordinates with some matrix:

$$x^{\mu^{'}} = \Lambda^{\mu^{'}}_{\nu} x^\nu \tag{1.28}$$

In order to find the matrices that leaves the space-time interval (or metric) invariant (which is the same thing as saying speed of light measured is the same) we substitute this into space time interval and we find that

$$\eta_{\rho\sigma} = \Lambda^{\mu^{'}}_{\rho}\Lambda^{\nu^{'}}_{\sigma}\eta_{\mu^{'}\nu^{'}} \tag{1.29}$$

If (1.29) is satisfied, that matrix is called *Lorentz Transformations.* For example, a boost in the -x direction can be written as

$$\Lambda^{\mu^{'}}{}_{\nu} = \begin{bmatrix} cosh\phi & -sinh\phi & 0 & 0 \\ -sinh\phi & cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{1.30}$$

It is obvious that boost corresponds to changing the coordinates to another one moving with a constant velocity. Relations between coordinates are

$$t' = t cosh\phi - x sinh\phi \tag{1.31}$$
$$x' = -t sinh\phi + x cosh\phi \tag{1.32}$$

from this, we can see the the point $x' = 0$ is moving with velocity

$$v = \frac{x}{t} = \frac{sinh\phi}{cosh\phi} = tanh\phi \tag{1.33}$$

By taking $\phi = tanh^{-1}v$ we can write

$$\begin{aligned} t' &= \gamma(t - vx) \\ x' &= \gamma(x - vt) \end{aligned} \tag{1.34}$$

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

1.5.2 Energy and Momentum

In this chapter, We'll briefly discuss the physics in Minkowski Spacetime. First, we introduce the four velocity of an object

$$U^\mu = \frac{dx^\mu}{d\tau} \tag{1.35}$$

where τ is the proper time. Since $d\tau^2 = -\eta_{\mu\nu}dx^\mu dx^\nu$ four-velocity is automatically normalized to -1.

$$\eta_{\mu\nu}U^\mu U^\nu = -1 \tag{1.36}$$

This means that four-velocity is not a velocity in space like ordinary velocity, but a velocity through spacetime and always has the same magnitude. Since we define it for timelike trajectories it is negative. Second, we introduce the momentum four-vector

$$p^\mu = mU^\mu \tag{1.37}$$

where m is the mass of the particle which is the same in all inertial frames. From this definition, energy of a particle is $E = p^0$. For example, for a particle at rest $p^0 = m$ (since we take $c = 1$). For a particle moving in the x direction with a velocity v, momentum four-vector becomes $p^\mu = (\gamma m, v\gamma m, 0, 0)$. For small v we can write $p^0 = m + \frac{1}{2}mv^2$ and as expected. We can also write without approximation

$$p^\mu p_\mu = -m^2 \tag{1.38}$$

which can also be written as

$$E = \sqrt{m^2 + \mathbf{p^2}} \tag{1.39}$$

At this point we can also define a force four-vector similar to Newton's second law

$$f^\mu = m\frac{d^2}{d\tau^2}x^\mu(\tau) \tag{1.40}$$

For example 3 dimensional Lorentz force $\mathbf{F} = \mathbf{q}(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ becomes

$$f^\mu = -qU^\lambda F^\mu_\lambda \tag{1.41}$$

where F is the so called Faraday Tensor and will be discussed in Chapter 3.

Momentum four-vector provides enough information for a single particle's energy and momentum. But we need more than that for complex systems with many particles. We usually think such systems as *fluids* and define an energy-momentum tensor $T^{\mu\nu}$. It is usually defined as *the flux of momentum four-vector p^μ across a surface of constant x^ν*. From this definition it follows that: T^{00} is the flux of p^0 in the x^0 direction, or energy in time direction which simply is the energy denstiy ρ. $T^{0i} = T^{i0}$ is the momentum density and T^{ij} are the momentum flux or stress.

One of the most important property of this tensor is that it is conserved which can be expressed as

$$\partial_\mu T^\mu \nu = 0 \tag{1.42}$$

for $\nu = 0$ equation becomes energy conservation while other values of ν corresponds to momentum conservation. This ideas become more important as we move into General Relativity in Chapter 6.

2 Klein-Gordon Equations

Klein-Gordon Equation is a relativistic wave equation satisfied by real or complex scalar fields. It is related to Schrodinger Equation. In this chapter, the classical field equation will be discussed. Then, Klein-Gordon Equation will be derived as a classical field equation. Then its relation to the Schrodinger equation will be studied.

2.1 Classical Field Theory

We know that in classical mechanics for a particle with coordinate $q(t)$ we can derive the equations of motion using the *least action principle* where we define the action as

$$S = \int dt L(q, \dot{q}) \tag{2.1}$$

where L is the lagrangian and typically given as $L = K - V$ where K is the kinetic energy and V is the potential energy. Then, by using variational calculus we search for the critical points in the action and end up with *euler-lagrange equations*:

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\left(\frac{\partial L}{\partial(\dot{q})}\right) = 0 \tag{2.2}$$

Similarly, in field theory instead of coordinates $q(t)$ we have a set of spacetime dependent fields $\Phi^i(x^\mu)$. Then we define the lagrangian L as an integral of a **Lagrange Density** $\mathcal{L}$ over all space where $\mathcal{L}$ is a function of the fields and their derivatives

$$L = \int d^3x \mathcal{L} \tag{2.3}$$

Therefore the action becomes

$$S = \int dt L = \int d^4x \mathcal{L}(\Phi^i, \partial_\mu \Phi^i) \tag{2.4}$$

Normally "Lagrange Density" is called just "Lagrangian" as we will do throughout this paper.

In order to derive the Euler-Lagrange equations for a field, we require action to be unchanged under small variations

$$\Phi^i \rightarrow \Phi^i + \delta\Phi^i \tag{2.5}$$

$$\partial_\mu \Phi^i \rightarrow \partial_\mu \Phi^i - \partial_\mu(\delta\Phi^i) \tag{2.6}$$

Then, we can Taylor expand the Lagrangian as

$$\mathcal{L}(\Phi^i + \delta\Phi^i, \partial_\mu\Phi^i + \partial_\mu\delta\Phi^i) = \mathcal{L}(\Phi^i, \partial_\mu\Phi^i) + \frac{\partial\mathcal{L}}{\partial\Phi^i}\delta\Phi^i + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi^i)}\partial_\mu(\delta\Phi^i) \tag{2.7}$$

Similarly we write $S \rightarrow S + \delta S$ where

$$\delta S = \int d^4x \left[\frac{\partial\mathcal{L}}{\partial\Phi^i}\delta\Phi^i + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi^i)}\partial_\mu(\delta\Phi^i)\right] \tag{2.8}$$

Then, we write the second term as

$$\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi^i)}\partial_\mu(\delta\Phi^i) = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi^i)}\delta\Phi^i\right) - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi^i)}\right)\delta\Phi^i \tag{2.9}$$

as a consequence of chain rule. Then, plugging this into (2.8), we end up

$$\delta S = \int d^4x \left[\frac{\partial\mathcal{L}}{\partial\Phi^i} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi^i)}\right)\right]\delta\Phi^i + \int d^4x \left[\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi^i)}\delta\Phi^i\right)\right] \tag{2.10}$$

But the second term is a total derivative. Therefore can be transformed to a surface integral by Stoke's Theorem. Since we are dealing with the variations, we can choose variations that can vanish at boundaries. Therefore we left with

$$\delta S = \int d^4x \left[\frac{\partial\mathcal{L}}{\partial\Phi^i} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi^i)}\right)\right]\delta\Phi^i \tag{2.11}$$

Finally, by claiming that this integral should vanish for the critical points we end up with the Euler-Lagrange Equation for field theory

$$\frac{\partial\mathcal{L}}{\partial\Phi^i} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi^i)}\right) = 0 \tag{2.12}$$

2.2 Lagrangian Formalism of Klein-Gordon Equation

One of the simplest example of a field is real scalar field. They basically associates a scalar value to every point in a space:

$$\phi(x^\mu) : (spacetime) \to \mathbb{R} \tag{2.13}$$

So, we are trying to understand the classical mechanics of a single scalar field. We will have an energy density as a function of spacetime which includes kinetic energy term $\frac{1}{2}\dot{\phi}^2$, a gradient energy term $\frac{1}{2}(\boldsymbol{\nabla}\phi)^2$ and a potential energy term $V(\phi)$. We can combine these three term to make a Lorentz invariant Lagrangian $\mathcal{L}$ as

$$\mathcal{L} = -\frac{1}{2}\eta^{\mu\nu}(\partial_\mu\phi)(\partial_\nu\phi) - V(\phi) \tag{2.14}$$

analogous to $L = K - V$. We can derive the equation of motions from Euler-Lagrange Equation (2.12). We have

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\phi} &= -\frac{dV(\phi)}{d\phi}\\ \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} &= -\eta^{\mu\nu}\partial_\nu\phi\end{aligned} \tag{2.15}$$

putting these into Euler-Lagrange equations, we end up with our equation of motion

$$\eta^{\mu\nu}\partial_\mu\partial_\nu\phi - \frac{dV}{d\phi} = 0 \tag{2.16}$$

It is usually written as

$$\Box\phi - \frac{dV}{d\phi} = 0 \tag{2.17}$$

where we defined the operator $\Box = \eta^{\mu\nu}\partial_\mu\partial_\nu$ which is called **d'Alembertian**. Note that our metric convention is $(-+++)$, therefore, in a flat spacetime this equation is equivalent to

$$\ddot{\phi} - \nabla^2\phi + \frac{dV}{d\phi} = 0 \tag{2.18}$$

If we choose a potential of simple harmonic oscillator, i.e., $V(\phi) = \frac{1}{2}m^2\phi^2$ where the parameter m is the mass of the field. The reason of a field have a mass can be understood quantum mechanically. When we quantize the field, we find that momentum eigenstates are collection of particles each with mass m. For now, it is simply mass. Substituting this potential we have

$$\Box\phi - m^2\phi = 0 \tag{2.19}$$

which is the **Klein-Gordon Equation**.

2.3 A Quantum Mechanical Approach

This approach is more physically understandable and actually how Klein-Gordon approach the problem in the first place. We know that non-relativistic energy of a free particle is

$$\frac{p^2}{2m} = E \tag{2.20}$$

And bu quantizing this, we end up with the Schrodinger equation for a free particle

$$\frac{\hat{p}^2}{2m}\psi = \hat{E}\psi \tag{2.21}$$

where $\hat{p} = -i\hbar\boldsymbol{\nabla}$ and $\hat{E} = i\hbar\frac{\partial}{\partial t}$. However, Schrödinger Equation is not Lorentz Invariant. Therefore, in order to achieve Lorentz Invariant we need to work with relativistic energy relation

$$\sqrt{p^2c^2 + m^2c^4} = E \tag{2.22}$$

But quantizing this gives troubles. Klein and Gordon, instead tried to quantize the square of the energy as

$$[(i\hbar\boldsymbol{\nabla})^2c^2 + m^2c^4]\psi = -\hbar^2\frac{\partial^2}{\partial t^2}\psi \tag{2.23}$$

Rearranging terms yields

$$\frac{1}{c^2}\frac{\partial^2}{\partial t}\psi - \nabla^2\psi + \frac{m^2c^2}{\hbar^2}\psi = 0 \tag{2.24}$$

which can be written as

$$(\Box + \mu^2)\psi = 0. \tag{2.25}$$

3 Maxwell's Equations

Maxwell's equations are set of differential equations that governs the nature of electromagnetism and relate the electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$ together. In vector calculus notation the equations are

$$
\begin{aligned}
\boldsymbol{\nabla} \cdot \mathbf{E} =& \rho \\
\boldsymbol{\nabla} \times \mathbf{B} - \partial_{\mathbf{t}} \mathbf{E} =& J \\
\boldsymbol{\nabla} \cdot \mathbf{B} =& 0 \\
\boldsymbol{\nabla} \times \mathbf{E} + \partial_{\mathbf{t}} \mathbf{B} =& 0
\end{aligned} \tag{3.1}
$$

where ρ is the charge density, $\mathbf{J}$ is the current density and $\boldsymbol{\nabla}\times$ and $\boldsymbol{\nabla}\cdot$ are the curl and divergence operators. These equations are Lorentz invariant. In fact, this is the one of the main motivation for the development of a new theory, namely Relativity. But the Theory of Relativity helps to improve the Laws of Electromagnetism as much as Electromagnetism helps the Relativity to emerge. For instance, as we will discover in General Relativity, matter and energy cause spacetime to have some curvature, therefore, Maxwell's Equations require some modifications. But we will not be dealing with it in this paper. For now, we will consider flat spacetime with Minkowski metric η.

3.1 Covariant Form of Maxwell's Equations

As we stated earlier, Maxwell's Equations are Lorentz invariant. But it is not easy to grasp it as we look at in Eq (3.1). Therefore we will try to write down the equation in Tensor notation. First, Let's rewrite the equations component-wise

$$
\begin{aligned}
\partial_i E^i &= J^0 \\
\tilde{\epsilon}^{ijk}\partial_j B_k - \partial_0 E^i &= J^i \\
\partial_i B^i &= 0 \\
\tilde{\epsilon}^{ijk}\partial_j E_k + \partial_0 B^i &= 0
\end{aligned} \tag{3.2}
$$

where we introduce 3−dimensional Levi-Civita symbol $\tilde{\epsilon}$ with the same rules as 4−dimensional one with one index absent. Notice that we are working in 3−dimensional euclidean space where metric is δ^i_j as well its inverse. Therefore, lower and upper indices can be used interchangeably. Also we write J^0 for charge density because we define a 4-vector $J^\mu = J(\rho, J^1, J^2, J^3)$.

Now we introduce an antisymmetric rank $(0,2)$ **electromagnetic field strength tensor** (sometimes called as *Faraday tensor*) F as

$$
F_{\mu\nu} = -F_{\nu\mu} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix} \tag{3.3}
$$

Using the definition (3.3) and equation (3.2) we can construct a tensorial formalization of Maxwell's Equations. We can also raise the indices of the Faraday Tensor such that

$$
F^{\mu\nu} = \eta^{\mu\sigma}\eta^{\nu\rho}F_{\sigma\rho} = \begin{pmatrix} 0 & +E_1 & +E_2 & +E_3 \\ -E_1 & 0 & B_3 & -B_2 \\ -E_2 & -B_3 & 0 & B_1 \\ -E_3 & B_2 & -B_1 & 0 \end{pmatrix} \tag{3.4}
$$

Note that $F^{0i} = E^i$ and $F^{ij} = \epsilon^{ijk}B_k$ where Latin letters run from 1 to 3 whereas Greek letters run from 0 to 3 as before. Using these two, we can write the first two equation as

$$
\begin{aligned}
\partial_j F^{ij} - \partial_0 F^{0i} =& J^i \\
\partial_i F^{0i} =& J^0
\end{aligned} \tag{3.5}
$$

adding two equation side by side and using the antisymmetry property of the Faraday tensor, we can combine these two equation as

$$
\partial_\mu F^{\nu\mu} = J^\nu \tag{3.6}
$$

Similarly for the last two equations, we can write $-\frac{1}{2}\epsilon^{ijk}F_{jk} = B^i$ and $F_{0i} = -E_i$. Then the last two equation can be written as

$$\partial_{[\mu}F_{\nu\sigma]} = \partial_\mu F_{\nu\sigma} + \partial_\nu F_{\sigma\mu} + \partial_\sigma F_{\mu\nu} = 0 \tag{3.7}$$

where bracket ([ ]) means antisymmetrization. It basically makes a tensor antisymmetric by permuting the indexes between brackets. Since the tensor F is antisymmetric, it can be written as in equation (3.7). In general, we have written the Maxwell's equation in tensor notation in equation (3.6) and (3.6). Since they are tensor equations, it is clear that they are Lorentz invariant. Usually we call (3.6) and (3.7) together as **Covariant form of the Maxwell's equations**. The name covariant here is different from the one we had before. It refers to the transformation rules for the equation is the same in both sides. Maxwell's equations is also covariant in vector calculus notation of course, however, it is just a matter of jargon that people usually say the tensor notation is covariant because it is easy to see.

3.2 Lagrangian Formalism of Maxwell's Equations

Another way to obtain Maxwell's equation is to use the Lagrangian formalism as introduced in Chapter 2 by choosing a correct Lagrangian. Before we start, we know from classical electromagnetism that electric and magnetic field can be expressed as

$$\begin{aligned}\mathbf{E} &= -\boldsymbol{\nabla}\cdot\phi - \frac{\partial A}{\partial t}\\ \mathbf{B} &= \boldsymbol{\nabla}\times A\end{aligned} \tag{3.8}$$

where A is the vector potential and ϕ is a scalar field called the electric potential. At this point, we introduce a four vector $A^\mu = (\phi, A^1, A^2, A^3)$. With this definition using the equations (3.8) we can express the Faraday Tensor as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{3.9}$$

From this expression, we can easily see that, Faraday tensor is *gauge invariant*, i.e., if we perform a gauge transformation $A_\mu \to A_\mu + \partial_\mu f(x)$ it will remain unchanged:

$$(\partial_\mu A_\nu - \partial_\nu A_\mu) \to (\partial_\mu A_\nu - \partial_\nu A_\mu) + \partial_\mu \partial_\nu f(x) - \partial_\nu \partial_\mu f(x) = (\partial_\mu A_\nu - \partial_\nu A_\mu)$$
$$F_{\mu\nu} \to F_{\mu\nu} + \partial_\mu \partial_\nu f(x) - \partial_\nu \partial_\mu f(x) = F_{\mu\nu} \tag{3.10}$$

which follows from the fact that order of the partial derivatives do not matter. We can now move forward to derive Maxwell's equations (3.6) and (3.7) from our new expression of Faraday tensor and using Lagrangian formalism. The equation (3.7) can be seen easily:

$$\partial_{[\mu} F_{\nu\sigma]} = \partial_{[\mu} \partial_\nu A_{\sigma]} - \partial_{[\mu} \partial_\sigma A_{\nu]} = 0 \tag{3.11}$$

because again the order of partial derivatives doesn't affect the result. Equation (3.6) comes from the Euler-Lagrange equations:

$$\frac{\partial \mathcal{L}}{\partial A_\nu} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \right) = 0 \tag{3.12}$$

where we choose Lagrangian to be

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + A_\mu J^\mu \tag{3.13}$$

It is easy to see that the first term of the Euler-Lagrange equation gives:

$$\frac{\partial \mathcal{L}}{\partial A_\nu} = J^\nu \tag{3.14}$$

For the second term, we need to F tensor with lower indices since we are taking derivative with respect to $\partial_\mu A_\nu$. Therefore we write the first term in the lagrangian as

$$-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} = -\frac{1}{4} \eta^{\mu\alpha} \eta^{\nu\beta} F_{\mu\nu} F_{\alpha\beta} = -\frac{1}{4} \eta^{\sigma\alpha} \eta^{\rho\beta} F_{\sigma\rho} F_{\alpha\beta} \tag{3.15}$$

where we also changed the dummy indices $\mu\nu$ to $\sigma\rho$ since we will be differentiating with respect to $\partial_\mu A_\nu$. Then using chain rule we can write

$$-\frac{1}{4} \frac{\partial F_{\sigma\rho} F^{\sigma\rho}}{\partial(\partial_\mu A_\nu)} = -\frac{1}{4} \eta^{\sigma\alpha} \eta^{\rho\beta} \left[\left(\frac{\partial F_{\sigma\rho}}{\partial(\partial_\mu A_\nu)} \right) F_{\alpha\beta} + F_{\sigma\rho} \left(\frac{\partial F_{\alpha\beta}}{\partial(\partial_\mu A_\nu)} \right) \right] \tag{3.16}$$

Also we have

$$\frac{\partial F_{\sigma\rho}}{\partial(\partial_\mu A_\nu)} = \frac{\partial(\partial_\sigma A_\rho - \partial_\rho A_\sigma)}{\partial(\partial_\mu A_\nu)} = \delta^\sigma_\mu \delta^\rho_\nu - \delta^\rho_\mu \delta^\sigma_\nu \tag{3.17}$$

plugging (3.17) into (3.16) we have

$$\begin{aligned} -\frac{1}{4}\frac{\partial F_{\sigma\rho} F^{\sigma\rho}}{\partial(\partial_\mu A_\nu)} =& -\frac{1}{4}\eta^{\sigma\alpha}\eta^{\rho\beta}\big[(\delta^\sigma_\mu \delta^\rho_\nu - \delta^\rho_\mu \delta^\sigma_\nu)F_{\alpha\beta} + F_{\sigma\rho}(\delta^\alpha_\mu \delta^\beta_\nu - \delta^\beta_\mu \delta^\alpha_\nu)\big] \\ =& -\frac{1}{4}\big[(\eta^{\mu\alpha}\eta^{\nu\beta} - \eta^{\nu\alpha}\eta^{\mu\beta})F_{\alpha\beta} + F_{\sigma\rho}(\eta^{\sigma\mu}\eta^{\rho\nu} + \eta^{\sigma\nu}\eta^{\rho\mu})\big] \\ =& -\frac{1}{4}[F^{\mu\nu} - F^{\nu\mu} + F^{\mu\nu} - F^{\nu\mu}] \\ =& -F^{\mu\nu} \end{aligned} \tag{3.18}$$

Combining (3.14) and (3.18) we end up with the second Maxwell's equation (3.6)

$$\partial_\mu F^{\nu\mu} = J^\nu. \tag{3.19}$$

4 Proca Equations

Proca equation is an extended version of Maxwell's Equations and named after Romanian physicist Alexandru Proca. It describes a massive spin-1 field of mass m in Minkowski spacetime.

4.1 Lagrangian Formalism of Proca Equation

The Lagrangian is given as

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m^2 A^{\mu}_{\mu} + A_{\mu}J^{\mu} \tag{4.1}$$

This is the same Lagrangian we introduced for Maxwell's Equations with an extra term $\frac{1}{2}m^2 A^{\mu}_{\mu}$. Therefore we just going to modify Euler-Lagrange Equation from the previous section. We have

$$\frac{\partial \mathcal{L}}{\partial A_{\nu}} = J^{\nu} - m^2 A^{\nu} \tag{4.2}$$

$$\frac{\partial \mathcal{L}}{\partial(\partial_{\mu} A_{\nu})} = -F^{\mu\nu} \tag{4.3}$$

Therefore Euler-Lagrange Equation gives

$$m^2 A^{\nu} - \partial_{\mu}F^{\mu\nu} = j^{\nu} \tag{4.4}$$

which can be also written as

$$-\Box A^{\nu} + \partial^{\nu}(\partial_{\mu}A^{\mu}) + m^2 A^{\nu} = j^{\nu} \tag{4.5}$$

note that minus sign is due to metric convention. Hence, it is possible to see the equation as

$$\Box A^{\nu} - \partial^{\nu}(\partial_{\mu}A^{\mu}) + m^2 A^{\nu} = j^{\nu} \tag{4.6}$$

This equation is called **Proca Equation**. And for $m = 0$, it reduces to Maxwell's equations.

5 Curvature

We have studied the flat spacetime special theory of relativity. But in order to generalize the laws of physics for more general curved spacetime and to understand general relativity, we need to study curvature. Therefore, in this chapter, more rigours definition of curvature will be given as well as some important concepts in differential geometry.

5.1 Exterior Derivative

Partial derivatives are essential operators for study of tensors. However, partial derivative of a tensor are not a tensor in general. Let's look at the transformation of the partial derivative of a one-form

$$\partial_{\mu'}\omega_{\nu'} = \frac{\partial x^{\mu}}{\partial x^{\mu'}}\partial_{\mu}\left(\frac{\partial x^{\nu}}{\partial x^{\nu'}}\omega_{\nu}\right) \tag{5.1}$$

$$= \frac{\partial x^{\mu}}{\partial x^{\mu'}}\frac{\partial x^{\nu}}{\partial x^{\nu'}}\left(\partial_{\mu}\omega_{\nu}\right) + \omega_{\nu}\frac{\partial x^{\mu}}{\partial x^{\mu'}}\partial_{\mu}\frac{\partial x^{\nu}}{\partial x^{\nu'}} \tag{5.2}$$

As can be seen, there is an extra term which we don't have in tensor transformation law. There we define new derivative operators which are tensors. First one is **exterior derivative**. It acts on -p form fields and gives (p+1)-form fields. Usually the symbol d is used for exterior derivative. It is defined as

$$(\mathrm{d}A)_{\mu_1\cdots\mu_{p+1}} = (p+1)\partial_{[\mu_1}A_{\mu_2\cdots\mu_{p+1}} \tag{5.3}$$

The idea to show the gradient with this symbol makes much more sense, because exterior derivative of scalar is just gradient.

$$(\mathrm{d}\phi)_{\mu} = \partial_{\mu}\phi \tag{5.4}$$

5.2 Covariant Derivative

In flat spacetime partial derivatives are maps from (k,l) tensor fields to (k,l+1) tensor fields. We would like to have a similar map for more general cases, in other words, coordinate independent as oppose to partial derivatives. To achieve this we define a covariant derivative operator ∇ defined as

$$\nabla_\mu V^\nu = \partial V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda \tag{5.5}$$

the Γ symbol also known as **connection coefficients** in the second term is basically the correction for partial derivatives to make them transform like a tensor. The reason it is called a symbol is that it is not a tensor and by demanding that covariant derivative transforms like

$$\nabla_{\mu'} V^{\nu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} \nabla_\mu V^\nu \tag{5.6}$$

we can show that connection coefficients transforms like

$$\Gamma^{v'}_{\mu'\lambda'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\lambda}{\partial x^{\lambda'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \Gamma^\nu_{\mu\lambda} V^\lambda + \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\lambda}{\partial x^{\lambda'}} \frac{\partial^2 x^{\mu'}}{\partial x^\mu \partial x^\lambda} \tag{5.7}$$

Hence, it is not a tensor. The second term and the extra term in partial derivative transformation exactly cancels and therefore makes a tensor.S We also write a general formula for a one form as

$$\nabla_\mu \omega_\nu = \partial_\mu \omega_\nu + \tilde{\Gamma}^\lambda_{\mu\nu} \omega_\lambda \tag{5.8}$$

Also, It easy to check that covariant derivative obeys Leibniz rule and it is linear. We also want to add to more properties to it: 1) it should commute with contractions $\nabla_\mu (T^\lambda_\rho) = (\nabla T)^\lambda_{\mu\rho}$. 2) It should reduce to partial derivative on scalars. With these to prorperties, it is easy to show that

$$\tilde{\Gamma}^\lambda_{\mu\nu} = -\Gamma^\lambda_{\mu\nu} \tag{5.9}$$

Hence, using the same connection coefficient we can write a covariant derivative formula for differential forms.

$$\nabla_\mu \omega_\nu = \partial_\mu \omega_\nu - \Gamma^\lambda_{\mu\nu} \omega_\lambda \tag{5.10}$$

The name connection coefficients comes from the fact that, by differentiating we move the vectors from one tangent space to another. In other words, they simply tells the connection between tangent spaces. However, it turns out there in general relativity the metric defines a unique connection. In order to find it we need to add more properties to our connection. First, let's define a **torsion tensor** as

$$T^{\lambda}_{\mu\nu} = \Gamma^{\lambda}_{\mu\nu} - \Gamma^{\lambda}_{\nu\mu} = 2\Gamma^{\lambda}_{[\mu\nu]} \tag{5.11}$$

Be aware that difference of two different connections is in fact a tensor. At this point we add two more properties: 1) Torsion-free $\Gamma^{\lambda}_{\mu\nu} = \Gamma^{\lambda}_{\nu\mu}$. 2) Metric compatibility $\nabla_{\rho} g_{\mu\nu}$. We claim that there is one unique torsion-free connection on a manifold that is compatible with some metric. And by simple calculation using the fact that derivative of metric is zero we can find that

$$\Gamma^{\sigma}_{\mu\nu} = \frac{1}{2} g^{\rho\sigma} (\partial_{\mu} g_{\nu\rho} + \partial_{\nu} g_{\rho\nu} - \partial_{\rho} g_{\mu\nu}) \tag{5.12}$$

This specific connection which general relativity is based on is called **Christoffel Symbol**.

A useful example is the covariant divergence of a vector which can be written as

$$\nabla_{\mu} V^{\mu} = \partial_{\mu} V^{\mu} + \gamma^{\mu}_{\mu\lambda} V^{\lambda} \tag{5.13}$$

where we can write

$$\Gamma^{\mu}_{\mu\lambda} = \frac{1}{\sqrt{|g|}} \partial_{\lambda} \sqrt{|g|} \tag{5.14}$$

Hence,

$$\nabla_{\mu} V^{\mu} = \frac{1}{\sqrt{|g|}} \partial_{\mu} (\sqrt{|g|} V^{\mu}) \tag{5.15}$$

Using covariant divergence we can write curved spacetime version of Stoke's theorem on region Σ with boundary $\partial\Sigma$ as

$$\int_{\Sigma} \nabla_{\mu} V^{\mu} \sqrt{|g|} d^n x = \int_{\partial\Sigma} = n_{\mu} V^{\mu} \sqrt{|\gamma|} d^[n - 1] x \tag{5.16}$$

where $n_m u$ is normal to $\partial\Sigma$ and γ is the induced metric on $\partial\Sigma$.

5.3 Geodesic Equation and Parallel Transport

Parallel Transport is the curved space generalization of keeping a vector constant as move on a manifold. It is obvious how to keep a tensor consant in a flat spacetime. Simple take the derivative and equate it to zero. In order to generalize the idea we replace the derivative with covariant derivative and define a *directional covariant derivative*:

$$\frac{D}{d\lambda} = \frac{dx^\mu}{d\lambda}\nabla_\mu \tag{5.17}$$

Then, we say that *a tensor is parallel transported along a curve if the directional covariant derivative of that tensor is zero along the curve.* For example, for a vector the equation becomes

$$\frac{d}{d\lambda}V^\mu + \Gamma^\mu_{\sigma\rho}\frac{dx^\sigma}{d\lambda}V^\rho = 0 \tag{5.18}$$

which is called **equation of parallel transport**.

Another concept is the **geodesics**. It is the curved space generalization of the euclidean straight line. Alternativly, we can define it as a curve along which the tangent vector is parallel transported. Hence we write

$$\frac{D}{d\lambda}\frac{dx^\mu}{d\lambda} = 0 \tag{5.19}$$

or equivalently,

$$\frac{d^2x^\mu}{d\lambda} + \Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\lambda}\frac{dx^\sigma}{d\lambda} = 0 \tag{5.20}$$

This equation is called **geodesic equation**. Another approach is to take the proper time functional

$$\tau = \int \sqrt{-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}}d\lambda \tag{5.21}$$

and find the critical points by simple variational calculus method, i.e., use euler-lagrange equation. With a little bit of work it can be seen that it will also lead to the same equation.

5.4 Reimann Curvature Tensor

We know that flat space has some properties such as if we move a vector in loop it will stay constant. This is not true in general for curved spaces. Consider an infinitesimal parallelepiped loop defined by two vectors A^μ and $B\nu$. And we move a vector V^σ along the loop to where it started. We say that there should be a tensor that says how much the vector V has changed as

$$\delta V^\sigma = R^\rho_{\sigma\mu\nu} A^\mu B^\nu V^\sigma \tag{5.22}$$

Where R is the Reimann Curvature Tensor. It is antysymmetric in the last two indices since interchanging the vectors A and B gives the inverse of the original answer. This is not a proper definition yet it is good for understanding what really Reimann Tensor measures. The real definition comes from the commutator of two covariant derivatives

$$\begin{aligned}
[\nabla_\mu, \nabla_\nu] V^\rho =& \nabla_\mu \nabla_\nu V^\rho - \nabla_\nu \nabla_\mu V^\rho & (5.23) \\
=& (\partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\nu \Gamma^\rho_{\nu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\sigma_{\nu\lambda}\Gamma^\lambda_{\mu\sigma}) V^\rho - 2\Gamma^\lambda_{[\mu\nu]} \nabla_\lambda V^\rho \\
=& R^\rho_{\sigma\mu\nu} V^\sigma - T^\lambda_{\mu\nu} \nabla_\lambda V^\rho & (5.24)
\end{aligned}$$

where we subtittue the torsion tensor and defined

$$R^\rho_{\sigma\mu\nu} = \partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\nu \Gamma^\rho_{\nu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\sigma_{\nu\lambda}\Gamma^\lambda_{\mu\sigma} \tag{5.25}$$

This is true for any connection, torsion-free or not. But we will be concerned with Christoffel connection. Therefore, connection can be derived from metric and curvature can be thought of as due to metric. Hence, looking at the definition of the Reimann Tensor we can say that *If we can write the components of the metric as constants in some coordinate system Reimann tensor will vanish and vice versa.* Also it is important to point out that these statements holds for simply connected regions of a manifold.

5.5 Ricci Tensor and Einstein Tensor

Ricci Tensor is obtained by taking the contraction of the Reimann Tensor

$$R_{\mu\nu} = R^\lambda_{\mu\lambda\nu} \tag{5.26}$$

For connection other than the Christoffel connection, there are several contraction we can take. However, for the Christoffel connection we only ave this one. All other contractions are either disappears or they are related to this one.

Ricci Tensor associated with Christoffel connection is symmetric

$$R_{\mu\nu} = R_{\nu\mu} \tag{5.27}$$

the trace of the Ricci Tensor is the Ricci scalar:

$$R = R^{\mu}_{\mu} = g^{\mu\nu} R_{\mu}\nu \tag{5.28}$$

Finally without giving any deeper explanation for this chapter, we define the Einstein Tensor as

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu}\nu \tag{5.29}$$

6 Einstein's Field Equations

Having equipped with the notion of curvature and all necessary mathematical tool, we will in this chapter begin to construct a general theory of relativity and derive the Einstein's Equation.

6.1 Curvature and Gravity

General Relativity basically tells us how the curvature acts on matter and manifest as gravity and how energy and momentum creates curvature and hence determines the gravitational field. In Newtonian Physics we have the acceleration of a body in a gravitational potential Φ as

$$\mathbf{a} = -\boldsymbol{\nabla}\boldsymbol{\Phi} \tag{6.1}$$

And we have the Poisson's equation that relates matter density and gravitational potential as

$$\nabla^2\Phi = 4\pi G\rho \tag{6.2}$$

Einstein Equivalence Principle states that "In small regions of spacetime, the laws of physics reduce to those of special relativity and it is impossible to detect the gravity by local experiments". The Idea behind these statement is that gravity is universal. Therefore it is not a force like any other, it the fundamentalW feature pf the fabric of the spacetime where all the matter fields propagate.

At this point what we are going to do is to generalize the laws of physics for more general curved spacetime. To do this, all we need to do is, we write the laws of physics in tensorial form for flat spacetime and generalize it by using metric g instead of Minkowski metric η and using covariant derivative instead of partial derivatives and so on. There are some problems with this approach but it is good enough for present purposes.

For example, an unaccelerated (freely falling) objects in flat spacetime moves in straight lines $x^\mu(\lambda)$ as

$$\frac{d^2x^\mu}{d\lambda^2} = 0 \tag{6.3}$$

writing in covariant derivative we get

$$\frac{d^2x^\mu}{d\lambda^2} = \frac{dx^\mu}{d\lambda}\partial_\nu\frac{dx^\mu}{d\lambda} \to \frac{dx^\mu}{d\lambda}\nabla_\nu\frac{dx^\mu}{d\lambda} \tag{6.4}$$

$$= \frac{d^2x^\mu}{d\lambda} + \Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\lambda}\frac{dx^\sigma}{d\lambda} \tag{6.5}$$

Therefore, we see that free particles move along the geodesics.

Another example is the energy-momentum conservation. We had in special relativity

$$\partial_\mu T^{\mu\nu} = 0 \tag{6.6}$$

Which for curved spacetime will become

$$\nabla_\mu T^{\mu\nu} = 0 \tag{6.7}$$

Therefore, we can generalize our equations for more general curved spacetime. How can we say that the result describes gravity. To be satisfied with this we can show that in Newtonian limit the gravity fits into picture. Consider, slow moving particle (w.r.t speed of light), weak gravitational field (so that it can be thought of as the perturbation to the flat space) and field is static. since we are moving slowly we can write the geodesic equation as (we choose the parameter λ as proper time)

$$\frac{d^2x^\mu}{d\tau} + \gamma^\mu_{00}\left(\frac{dt}{d\tau}\right)^2 = 0 \tag{6.8}$$

where $\Gamma^\mu_{00} = -\frac{1}{2}g^{\mu\lambda}\partial_\lambda g_{00}$. Assuming a weak gravitation allows us the write our metric as Minkowski metric plus some small perturbation:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{6.9}$$

where $|h| \ll 1$. Therefore to first order in h we can write

$$\Gamma^{\mu}_{00} = -\frac{1}{2}\eta^{\mu\lambda}\partial_{\lambda}h_{00} \tag{6.10}$$

Then geodesic equation becomes

$$\frac{d^2x^{\mu}}{d\tau^2} = \frac{1}{2}\eta^{\mu\nu}\partial_{\lambda}h_{00}\left(\frac{dt}{d\tau}\right)^2 \tag{6.11}$$

$\mu = 0$ tells us that $dt/d\tau$ is constant. And spacelike components gives

$$\frac{d^2x^{i}}{dt^2} = \frac{1}{2}\partial_{i}h_{00} \tag{6.12}$$

which is very similar to Newton's gravity formula. In fact if we take $h_{00} = -2\Phi$ they become exactly the same. Therefore

$$g_{00} = -(1 + 2\Phi) \tag{6.13}$$

This shows that curvature of spacetime is actually sufficient to describe gravity in the limit. Only thing left is to find the field equations.

6.2 Einstein's Equation

We want to find an equation that replaces Poisson equation (6.2) for Newtonian potential. For the right hand side we have a generalization of mass density which is the energy momentum tensor $T + \mu\nu$. And we know that gravitational potential should be replaces by metric tensor as in (6.13). We need to have tensor that has second derivative of the metric in it. We may try the d'Alembertian operator $\Box = \nabla^{\mu}\nabla_{\mu}$, however, by metric compatibility this gives zero. But we have another tensor with a second derivative of the metric in it which is Reimann tensor. But the indices does not much, therefore we can use Ricci tensor and write

$$R_{\mu\nu} = \alpha T_{\mu\nu} \tag{6.14}$$

where α is some constant. Actually, Einstein made this proposition at some point. But there is problem with the conservation of energy-momentum. If we want to conserve energy, i.e.

$$\nabla^{\mu}T_{\mu\nu} = 0 \tag{6.15}$$

we also need to say

$$\nabla^{\mu} R_{\mu\nu} = \tag{6.16}$$

Yet this is not always true. Also, we can write from our equation as

$$R = \alpha g^{\mu\nu} T_{\mu\nu} = \alpha T \tag{6.17}$$

this implies that

$$\nabla_{\mu} T = 0 \tag{6.18}$$

meaning T is constant everywhere, which is not true either. $T = 0$ in vacuum and $T \neq 0$ in matter. Therefore, we need another guess which is of course the Einstein tensor

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \tag{6.19}$$

which always satisfies $\nabla^{\mu} G_{\mu\nu} = 0$, therefore

$$G_{\mu\nu} = \alpha T_{\mu\nu} \tag{6.20}$$

All we need to do is to find the proportionality constant α. To de this, we look whether this equation reduces to Poisson equation (6.2) in the limit. For this, we consider a perfect-fluid source of energy-momentum

$$T_{\mu\nu} = (\rho + p) U_{\mu} U_{\nu} + p g_{\mu\nu} \tag{6.21}$$

where U is the fluid four-velocity and ρ and p are the rest frame energy and momentum densities. Since we consider Newtonian limit we may neglect the pressure which becomes unimportant as the particles moves slowly relative to that of light. Therefore,

$$T_{\mu\nu} = \rho U_{\mu} U_{\nu} \tag{6.22}$$

which is the energy-momentum tensor of a dust which in this case a massive body. Again, in the newtoinan limit just like in (6.13) we write

$$g_{00} = -1 + h_{00} \tag{6.23}$$

Then, to the first order in h we get (we work in the rest frame of the body)

$$U^0 = 1 + \frac{1}{2}h_{00} \tag{6.24}$$

Since, we are going to plug this into (6.22) ρ is small we can simply take $U^0 = 1$ and $U_0 = -1$. Then all components vanish except

$$T_{00} = \rho \tag{6.25}$$

As expected, in Newtonian limit rest mass is much larger than other terms, therefore we neglected them all. We plug this into our field equations and get

$$R_{00} = \frac{1}{2}\alpha\rho \tag{6.26}$$

We need to evaluate $R_{0\lambda} = R^0_{0\lambda 0}$ to find the explicit formula. $R^0_{000} = 0$ therefore we get

$$R_{00} = R^i_{0i0} \tag{6.27}$$

$$= \partial_i\big[\frac{1}{2}g^{i\lambda}(\partial_0 g_{\lambda 0} + \partial_0 g_{0\lambda} - \partial_\lambda g_{00})\big] \tag{6.28}$$

$$= \frac{1}{2}\delta^{ij}\partial_i\partial_j h_{00} \tag{6.29}$$

$$= \frac{1}{2}\nabla^2 h_{00} \tag{6.30}$$

Hence, the field equation became

$$\nabla^2 h_{00} = -\alpha\rho \tag{6.31}$$

Since $h_{00} = -2\Phi$ if we choose $\alpha = 8\pi G$ this equation becomes exactly the Poisson equation (6.2). Therefore **Einstein's Equation** becomes

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{6.32}$$

7 Conclusion

In this study, we derived the classical field equations namely 1) Klein-Gordon Equation; 2) Maxwell's Equations; 3) Proca Equations and 4) Einstein Equations and particularly studied the subject of Relativity. After studying the subjects like Integration on manifolds, tensors and differential forms as well as giving a brief review of the Special Theory of Relativity in Chapter 1; and explaining the Classical Field Theory in Chapter 2, we gave the derivation and physical meaning of the e Klein-Gordon,Maxwell and Proca Equation. Then, in Chapter 5 we studied the subject of differential geometry and curvature allowing us the derive Einstein's Equations of General Relativity.

Bibliography

I **Spacetime and Geometry: An Introduction to General Relativity**, S.Carroll (Addison-Wesley,2003)

II **Calculus on Manifolds**, M.Spivak (Benjamin,1965)

III **Special Relativity**, N.M.J.Woodhouse (Springer,1992)

IV **The Geometry of Minkowski Spacetime**, G.L.Naber (Springer,1992)

V **Relativity, Gravitation and Cosmology**, Ta-Pei Cheng (Oxford University Press, 2005)

VI **A First Course in General Relativity**, Bernard F. Schutz (Cambridge University Press 1985)