Special Relativity: A (Very) Brief Introduction

1 Introduction:

This will serve as a basic introduction to the essential elements of special relativity necessary to fully understand QFT. I begin with the key postulates and mathematical structure used to describe Einstein's very special theory,before heading into further topics such as Lorentz invariance and some EM theory.

Each topic will be kept self-contained and as brief as possible. Note that I will provide historical context for some of the things discussed in these notes. But, only where it is deemed interesting and insightful (by me).

2 Minkowski Space:

This section of the text will set the stage for quite literally everything to come. We shall quickly review the historical context behind this beautiful mathematical structure before looking (in)formally at its contents.

2.1 Hermann Minkowski:

Hermann Minkowski is an interesting historical figure (check his wiki page) and was a prominent mind within mathematics and mathematical physics during the early 20th century. He was, funnily enough, a former tutor of Einstein at Zurich (with Einstein deciding to skip most of his mathematics lectures), and found it surprising that his former pupil, who was arguably bad at maths, was able to come up with such a profound theory using simple algebraic logic. Minkowski was the first to make Einstein's rules rigorous, turning his algebraic work into a fully geometric picture. Let us now review this topic.

2.2 Geometry and Spacetime:

Before we begin, I should mention that we shall not cover curved spacetimes at all within this text. They are not needed to understand QFT and are only applicable if you were to look at theories defined on curved background geometries.

I shall now introduce you to some rather basic differential geometry that is required to fully understand special relativity in a way that is applicable to QFT. Let's begin.

2.2.1 Smooth Manifolds:

In simple terms, a smooth (or differentiable) manifold is one that locally "looks like" (technically, it must be locally homeomorphic to) $\mathbb{R}^n$ with enough structure to do calculus on. This means that we are able to define continuous functions on it: such as tensor fields. Note, there is, of course, a very thorough topological definition which requires certain conditions like 'second countable' and 'Hausdorff', but we will act like good physicists and ignore the mathematicians' pleas for rigour.

Given a smooth manifold $\mathcal{M}$, we can define a generalisation of a tangent to a curve via the following definition:

Definition 2.1. For any $p \in \mathcal{M}$ the *tangent space* at p, denoted $T_p\mathcal{M}$, is the set of all tangent vectors to $\mathcal{M}$.

The tangent space naturally forms a real vector space, allowing us to take linear combinations and scalar multiples of tangent vectors. Tangent vectors can also be viewed as smooth linear operators that act on functions defined on $\mathcal{M}$ that are required to satisfy the Leibniz rule (we will ignore this). If we want to look at the tangent spaces $\forall\, p \in \mathcal{M}$ a natural object is the tangent bundle,

---

Definition 2.2. The *Tangent bundle*, denoted $T\mathcal{M}$, is given by the disjoint union of the tangent spaces at all points $p \in \mathcal{M}$,

$$T\mathcal{M} = \bigsqcup_{p\,\in\mathcal{M}} T_p\mathcal{M}\,.$$

Think of this as just the set of all tangent vectors to the manifold. An element of $T\mathcal{M}$ is a pair (p, X), where $p \in \mathcal{M}$ and $X \in T_p\mathcal{M}$ is a vector. As with any vector space, we may find its' dual; namely the set of linear maps from the vector space to the base field we are working over (usually $\mathbb{R}$ or $\mathbb{C}$):

Definition 2.3. The *cotangent space* at a point p is the dual vector space to the tangent space at p, and is denoted $T_p^*\mathcal{M}$. An element of this space is a covector and is the map

$$\omega : T_p\mathcal{M} \to \mathbb{R}\,(\text{usually, but can be different})\,.$$

In a similar fashion to our definition of the tangent bundle, we may define:

Definition 2.4. The *cotangent bundle* of a smooth manifold $\mathcal{M}$, denoted $T^*\mathcal{M}$, is the disjoint union of the cotangent spaces at all points $p \in \mathcal{M}$,

$$T^*\mathcal{M} = \bigsqcup_{p \in \mathcal{M}} T_p^*\mathcal{M}.$$

An element of $T^*\mathcal{M}$ is a pair (p, ω), where $p \in \mathcal{M}$ and $\omega \in T_p^*\mathcal{M}$ is now a covector (or 1-form) at p. A *1-form field* on $\mathcal{M}$ is a smooth section of the cotangent bundle.

As we have now introduced both the tangent and cotangent bundles, we may define what is meant by a tensor field and, more generally, the tensor bundle. First, recall that a vector field is simply the assignment of a vector to a specific point in space, hence the tensor field may be defined naturally as

Definition 2.5. Let $\mathcal{M}$ be a smooth manifold. The *tensor bundle* of type (r, s) is

$$T_s^r\mathcal{M} := \underbrace{T^*\mathcal{M} \otimes \cdots T^*\mathcal{M}}_{r \text{ times}} \otimes \underbrace{T\mathcal{M} \otimes \cdots \otimes T\mathcal{M}}_{s \text{ times}},$$

whose fibre (just a selection of the specific tangent and cotangent spaces) at $p \in \mathcal{M}$ is the vector space of type (r, s) tensors (a map from r copies of $T_p^*\mathcal{M}$ and s copies of $T_p\mathcal{M}$ to $\mathbb{R}$) over $T_p\mathcal{M}$ and $T_p^*\mathcal{M}$.

With this, a *tensor field* of type (r, s) on $\mathcal{M}$ is a smooth section (a choice of one element from each fibre varying smoothly with p) of $T_s^r\mathcal{M}$, i.e. a smooth assignment

$$p \mapsto T(p) \in T_s^r\mathcal{M}.$$

Simple examples of tensor fields include scalar fields, $T_0^0\mathcal{M}$, and vector fields, $T_0^1\mathcal{M}$. For our purposes, we shall only concern ourselves with tensor fields of, at most, type $(0, 2)$.

With these definitions, which will hopefully give you a sense of the underlying structure, the metric tensor can be defined via

Definition 2.6. The *metric tensor*, $g \in T_2^0\mathcal{M}$, must satisfy the following conditions:
  • *Symmetry* - Given $X, Y \in T_p\mathcal{M}$

$$g(X, Y) = g(Y, X).$$

  • *Non-degeneracy* - Given $X, Y \in T_p\mathcal{M}$

$$g(X, Y) = 0, \forall Y \in T_p\mathcal{M} \iff X = 0.$$

Non-degeneracy simply means that the only vector orthogonal to all others is the zero vector.

2.2.2 The Minkowski metric:

With these basic definitions we can now discuss Minkowski space proper. The thing that distinguishes Minkowski space from regular Euclidean 4-space, is the Minkowski metric.

Definition 2.7. The *Minkowski metric*, canonically denoted η, is a metric tensor of signature $(-,+,+,+)$. In the standard orthonormal basis, its' components take the form

$$\eta_{\mu\nu} = \text{diag}(-,+,+,+) = \text{diag}(-1,1,1,1).$$

A manifold that has the above signature is called *Lorentzian*. We have now "rigorously" defined everything we need to completely define Minkowski space:

Definition 2.8. *Minkowski space*, is the Lorentzian manifold $M = (\mathbb{R}^4, \eta)$, where η is the Minkowski metric. Since our manifold is flat, we may choose global *inertial coordinates* such that $x \in M$ takes the form $x = (t, x^1, x^2, x^3)$. We also often write M as $\mathbb{R}^{1,3}$.

The differing sign means that the Minkowski metric is not positive definite (> 0), like the Euclidean metric, but actually indefinite and has 3 distinct possibilities, depending on the sign of the inner product. As per usual, the inner product is defined via the metric tensor of the corresponding space, for example, in Euclidean space the inner product of two vectors, $v, w \in \mathbb{R}^n$ can be written as $\delta_{ij}v^i w^j$. We may also do the same in Minkowski space, where we now take our indices to be Greek (I assume you know Einstein's summation convention).

This structure allows us to encode causality in a simple way. Given $X \in T_p M$ then:

- If $\eta_{\mu\nu} X^\mu X^\nu > 0$, then X is said to be *spacelike*.
- If $\eta_{\mu\nu} X^\mu X^\nu = 0$, then X is said to be *null*.
- If $\eta_{\mu\nu} X^\mu X^\nu < 0$, then X is said to be *timelike*.

The metric also induces an isomorphism between the tangent space and its dual, by "raising" and "lowering" indices, via

$$X_\mu(\in T_p^* M) = \eta_{\mu\nu} X^\nu(\in T_p M).$$

If we define the inverse metric, written $\eta^{\mu\nu}$ via,

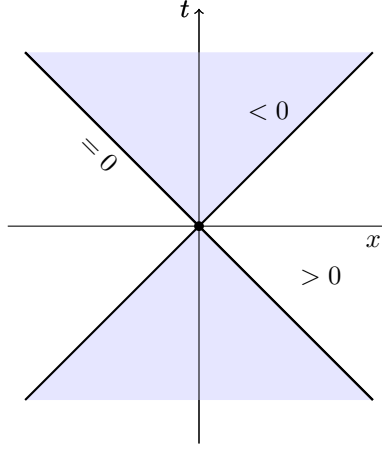$$\eta^{\mu\alpha}\eta_{\mu\beta} = \delta^\alpha_\beta\,,$$

we then find

Figure 1: Light cone in 1+1 dimensions with $c = 1$. Causal influence lies on or within the cone; massless signals follow the null lines, i.e. move at c.

$$X^\mu = \eta^{\mu\nu} X_\nu \,.$$

As previously mentioned, we may choose a global orthonormal basis such that η takes the regular form. We are, of course, well within our right to choose a different basis for our space. This observation leads to the Lorentz group - which we shall now cover.

2.3   The Lorentz Group:

The Lorentz Group, named after Dutch physicist Hendrik Lorentz, actually predates special relativity. It was initially formulated as the set of transformations that leaves Maxwell's equations invariant under changes of frame. Einstein subsequently derived these transformations independently in his paper on special relativity. Minkowski was the first to define the group of all such transformations.

As a motivational example we shall look initially at $SO(n)$, before looking at the Lorentz group. Recall that $SO(n)$ is a subgroup of the *isometry group* of $\mathbb{R}^n$, this means that the quadratic form is preserved. Namely, given the quadratic form

$$Q(v) = g(v, v) = v^{\mathrm{T}} g \, v \,,$$

where g is the metric, a linear operator A is said to *preserve* Q iff

$$Q(Av) = Q(v)$$

which implies that $A^{\mathrm{T}}gA = g$. This is exactly the requirement for the matrix Lie group $O(n)$ (when $g = I$). Think of these transformations as rotations; the vector's norm hasn't changed, it has simply been rotated. If we wish to preserve the orientation of the vector, i.e. no flips, we require only proper rotations, our group then becomes $SO(n)$ - the group of matrices in $O(n)$ with det $= 1$.

This is essentially exactly the case with Minkowski space, however, our space is not $\mathbb{R}^4$, it is $\mathbb{R}^{1,3}$. Thus, such a subgroup of the group of isometries for Minkowski space is often called the *Pseudo-Orthogonal Group* or *Lorentz group* and denoted $SO(1,3)$. Elements of this group are Lorentz transformations which satisfy

$$\Lambda^{\mathrm{T}}\eta\Lambda = \eta\,.$$

There is also another nuance; time has a different sign to all other components in our space. In order to impose time orientation (since det $= 1$ imposes spatial orientation) of our space we also require that the time component of our transformation, $\Lambda^0_0 \geq 1$. With this time orientation, our group is then $SO^+(1,3)$ the *Proper Orthochronus Lorentz Group.*

The next logical place for us to cover is the Lie algebra of the Lorentz group, I shall assume a certain amount of knowledge of Lie algebras, but, don't worry it won't be too in depth - I won't be covering anything like Cartan Matrices, Cartan-Weyl bases, or anything other than the most basic representation theory.

2.3.1   The Lorentz Algebra:

To begin I will firstly introduce the, hopefully familiar, definition of a Lie algebra:

Definition 2.9. Let G be a matrix Lie group, its corresponding Lie algebra, $\mathfrak{g}$, is a vector space along with the product $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathbb{F}$. It admits a basis of (in general) matrices, often called the *generators* of the Lie algebra.

To find the corresponding algebra of a Lie group we linearise near the identity of our group manifold. The tangent space at the identity is imbued (it follows from the group multiplication law and the Becker-Campbell-Hausdorff formula) with an anti-symmetric, bilinear, binary operation: the Lie bracket. For Minkowski space with the isometry group $SO^+(1,3)$ we let $\Lambda \in SO^+(1,3)$ and write an infinitesimal Lorentz transformation as

$$\Lambda^\mu_{\ \nu} = \delta^\mu_{\ \nu} + \omega^\mu_{\ \nu} + \mathcal{O}(\omega^2)\,.$$

The condition $\Lambda^T\eta\Lambda = \eta$ requires that, to first order in ω,

$$\eta_{\mu\rho}\,\omega^\rho_{\ \nu} + \eta_{\nu\rho}\,\omega^\rho_{\ \mu} = 0 \quad \Longleftrightarrow \quad \omega_{\mu\nu} = -\,\omega_{\nu\mu}.$$

Thus the Lie algebra $\mathfrak{so}(1,3)$ consists of antisymmetric matrices with six independent parameters (three rotations, and three boosts. This number comes from counting free components and imposing the appropriate constraints).

Let me now also introduce some representation theory that will be useful to know.

Definition 2.10. Given a Lie group G, a representation of G is a linear map $D : G \to GL(n, V)$. Where V is the vector space over which the matrices act, i.e.

$$v \mapsto D(g)v \in V \,,$$

with $v \in V$, $g \in G$. V is often called representation space.

Think of a representation as simply taking an element of our group and "representing" it with a matrix from the general linear group. There is also an analogous definition for Lie algebras. I will not cover this. In physics, we often deal with representations that cannot be split into sub-representations. It will hence be important to quickly define what we mean by this:

Definition 2.11. A representation $D : G \to GL(n, V)$ is said to be *irreducible* if it does not possess any non-trivial invariant subspaces.

The obvious trivial invariant subspaces are $\emptyset$ and V itself.

We now form a basis of generators for our Lie algebra by defining the generators $M^{\mu\nu} = -M^{\nu\mu}$ such that

$$\Lambda = \exp\left(\tfrac{1}{2}\omega_{\rho\sigma}\, M^{\rho\sigma}\right) \,, \quad (M^{\rho\sigma})^\mu{}_\nu := \eta^{\rho\mu}\delta^\sigma{}_\nu - \eta^{\sigma\mu}\delta^\rho{}_\nu.$$

This relation between the group elements and Lie algebra elements is called the *exponential map*. These generators satisfy the Lorentz algebra commutator relations:

$$[M^{\mu\nu},\, M^{\rho\sigma}] = \eta^{\nu\rho}M^{\mu\sigma} - \eta^{\mu\rho}M^{\nu\sigma} - \eta^{\nu\sigma}M^{\mu\rho} + \eta^{\mu\sigma}M^{\nu\rho} \,.$$

We often split this basis into rotations and boosts via the following relations Let

$$J_i := \tfrac{1}{2}\,\epsilon_{ijk}\, M^{jk} \text{ (rots)}, \qquad K_i := M^{0i} \text{ (boosts)} \qquad (i = 1, 2, 3).$$

Then

$$[J_i, J_j] = \epsilon_{ijk}J_k, \qquad [J_i, K_j] = \epsilon_{ijk}K_k, \qquad [K_i, K_j] = -\epsilon_{ijk}J_k.$$

Notice that the rotational generators J_i obey the same algebras as angular momentum operators in QM. This should be unsurprising. The boosts, however, do not form the same algebra. We can make them do so by extending our field to $\mathbb{C}$, i.e. we now take linear combinations of vectors with complex coefficients. This leads to the *Complexified Lie algebra* $\mathfrak{so}(1,3)_\mathbb{C}$. We take the combinations

$$A_i := \tfrac{1}{2}\left(J_i + iK_i\right), \qquad B_i := \tfrac{1}{2}\left(J_i - iK_i\right),$$

which obey

$$[A_i, A_j] = \epsilon_{ijk} A_k, \qquad [B_i, B_j] = \epsilon_{ijk} B_k, \qquad [A_i, B_j] = 0.$$

Note that I have not included the typical factor of i. The fact that A and B commute implies that the subspaces they span form an object known as an *ideal* such that the total algebra can be written as the direct sum of the two sub-algebras. The sub-algebra is $\mathfrak{su}(2)$, since it is the angular momentum algebra. Thus we may write the complexified Lie algebra as

$$\mathfrak{so}(1,3)_{\mathbb{C}} \cong \mathfrak{su}(2) \oplus \mathfrak{su}(2).$$

Let us now take a brief detour to introduce some 'highest weight theory'. But, firstly we must use an important relation, namely the fact that

$$\mathfrak{sl}(2, \mathbb{C}) \cong \mathfrak{su}(2)_{\mathbb{C}}.$$

I shall quickly prove this.

Proposition 2.1. The complexified Lie algebra $\mathfrak{su}(2)_{\mathbb{C}}$ is isomorphic to the algebra $\mathfrak{sl}(2, \mathbb{C})$.

Proof. To show there exists an isomorphism we must show that $\exists$ a bijection between the algebras. Let's firstly see exactly what we mean by $\mathfrak{su}(2)_{\mathbb{C}}$:

$$\mathfrak{su}(2)_{\mathbb{C}} = \mathfrak{su}(2) \otimes_{\mathbb{R}} \mathbb{C}$$

Now, we should construct a bijection (injective and surjective) between the two sets. I shall define the map $\eta : \mathfrak{su}(2)_{\mathbb{C}} \to \mathfrak{sl}(2, \mathbb{C})$ by

$$\eta(X \otimes z) \mapsto zX \ , \ X \in \mathfrak{su}(2), \ z \in \mathbb{C}.$$

This map is $\mathbb{R}$-linear in the first argument (since $\mathfrak{su}(2)$ is a real vector space) and $\mathbb{C}$-linear in the second, i.e.

$$\eta((aX + bY) \otimes z) = z(aX + bY)$$
$$\eta(X \otimes (z + w)) = (z + w)X.$$

For it to be a valid Lie algebra isomorphism, it must preserve the Lie bracket: let's check

$$\eta\left([X \otimes z, Y \otimes w]\right) = \eta\left([X, Y] \otimes zw\right)$$
$$= zw[X, Y]$$
$$= [zX, wY]$$
$$= [\eta(X \otimes z), \eta(Y \otimes w)]$$

It can be then shown (I exclude for brevity) that this map is also injective and surjective. $\square$

Definition 2.12. The Lie algebra $\mathfrak{sl}(2, \mathbb{C})$ is generated by $\{H, E_{\pm}\}$ with

$$[H, E_{\pm}] = \pm 2E, \qquad [E_+, E_-] = H.$$

A representation of $\mathfrak{sl}(2, \mathbb{C})$ is said to be of *highest weight* ℓ if there exists a non-zero vector $v_0 \in V$ (the rep space) such that

$$Hv_m = mv_m, \qquad E_- v_0 = 0, \qquad E_+ v_\ell = 0,$$

and V is spanned by repeated applications of E_+:

$$v_k := (E_+)^k v_0, \quad k = 0, 1, \ldots, \ell.$$

v_ℓ is then called the highest weight state. For $\ell \in \mathbb{N}$, this construction yields a finite–dimensional irreducible representation of dimension $\ell + 1$. Every finite–dimensional irreducible representation of $\mathfrak{sl}(2, \mathbb{C})$ arises in this way.

Given these facts, we are then free to write the two distinct sets of spin-quantum numbers of the two algebras in $\mathfrak{so}(1,3)_{\mathbb{C}} = \mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})$ as (j_L, j_R). Where in physics we define $j_L/j_R = \frac{\ell_{L/R}}{2}$. These numbers encode differing representations of the Lorentz group.

The Lorentz group is actually what's known as non-compact (imagine infinite volume), thus it can be shown (take my word for it) that there exists no non-trivial finite-dimensional unitary representations. This is bad for any quantum theory since we require our theories to be unitary. This means that particles can't be described with this alone; we actually need the *Poincaré* group.

The Poincaré group is the Lorentz group + translations (technically it is a semi-direct product group - but we won't worry about that here), with single particle states being described by unitary irreducible reps of the Poincaré group. They are indexed by both spin and mass. This topic is slightly beyond the main aim of this text; it is, however, very interesting, so I suggest reading further about it.

2.3.2 Representations of the Lorentz Group:

We shall now quickly look at certain representations of the Lorentz group i.e. different values of (j_L, j_R):

- $(0,0)$ - This representation is the called the scalar representation. It is not very interesting. In terms of Lie theory it is called the trivial representation. These representations are carried by scalar field theories.

- $(1/2, 0)$ and $(0, 1/2)$ - These are the left and right handed Weyl spinor representations. They are useful in supersymmetric theories.

- $(1/2, 1/2)$ - This is the vector representation of the Lorentz group. Some objects that transform under this representation are 4-position x_μ, derivatives ∂_μ, and the 4-momentum p_μ, along with others.

These representations are only a few of the many allowed possibilities. In terms of field theories that have multiple different types of field e.g. imagine having a scalar field including an interaction term which involves a vector field like A_μ, the fields live in different representation spaces of the Lorentz group (technically the full Poincaré group), hence, transform under separate representations. Let's look quickly at the scalar representation to hopefully make this more concrete. A scalar is simply a 1×1 matrix, hence, our representation would be $D : SO^+(1,3) \to GL(1, \mathbb{C})$. Now, it should hopefully be clear to see that $GL(1, \mathbb{C}) \cong \mathbb{C}^\times$. With this, it turns out that the only representation that is 1-dimensional and continuous for this group is the trivial one (such that all elements of G map to the identity of $GL(n, V)$). Thus, $D(\Lambda) = 1$, $\forall \Lambda$. This is exactly how scalar fields transform in QFT (i.e. the field is unchanged under Lorentz transformations). So, if you want to sound clever, simply say that a scalar field transforms under the trivial representation of the Lorentz (actually Poincaré) group.

3 4-Vectors and Dynamics:

Now that we have covered the group theory of the Lorentz group, let's now actually look at some relativistic dynamics. Note, being as terrible with constants as I am, I shall use natural units: we set $c = 1$. With this choice of units, Einstein's famous relation is now just $E = m$. This is often the case in most theory texts as it makes things a lot cleaner. It can, however, lead to issues if you forget to put them back, when all said and done.

3.1 4-Vectors:

Given our Minkowski space $\mathbb{R}^{1,3}$ a point on the manifold is called an event or 4-position and (as stated earlier) takes the form $x^\mu = (t, x^1, x^2, x^3)$. As you know, time in relativity is relative. Hence, it would not make sense to define the usual things like velocity or acceleration using the regular coordinate time. This would mean that the velocity of an object would be frame dependent i.e. a race-car

could be one velocity for the driver and another velocity for an observer. We instead use something called the proper time, which I shall now introduce:

Definition 3.1. A *world-line* is the curve $\Gamma : I \subset \mathbb{R} \to \mathbb{R}^{1,3}$. It is said to be *timelike* if $\eta_{\mu\nu}\dot{\Gamma}^\mu\dot{\Gamma}^\nu < 0 \,, \forall\, t \in I$.

Given such a world-line, the proper time is simply the actual length of the curve. Which is invariant.

Definition 3.2. The *proper time* $\tau : I \to J \subset \mathbb{R}$ is given by the arc length along the curve Γ,

$$\tau(t) = \int_{t_0}^t ds \,\sqrt{-\eta_{\mu\nu}\dot{\Gamma}^\mu(s)\dot{\Gamma}^\nu(s)}\,.$$

For the proper time to be non-complex (which would be odd) for a timelike curve, we require the condition $\eta_{\mu\nu}\dot{\Gamma}^\mu(\tau)\dot{\Gamma}^\nu(\tau) = -1$. This then means that the curve Γ is parametrised by the proper time. With this, we can define the 4-velocity of the curve (Imagine a particle whose position is dependent on time).

Definition 3.3. The *4-velocity* of a timelike curve Γ is simply

$$U^\mu = \frac{\partial \Gamma^\mu(\tau)}{\partial \tau} = \frac{\partial (t(\tau), x^1(\tau), x^1(\tau), x^2(\tau))}{\partial \tau}\,.$$

Where we have used local coordinates to describe the curve Γ.

The rate of change of coordinate time with respect to the proper time is a special factor: called the Lorentz factor, γ. It encodes how much coordinate time, t, differs from the proper time, τ. Let's use the normalisation condition,

$$
\begin{aligned}
-1 &= -(U^0)^2 + \|\boldsymbol{U}\|^2 \\
&= -(U^0)^2 + |U^0|^2\|\boldsymbol{v}\|^2 \\
\implies 1 &= (U^0)^2\left(1 - \|\boldsymbol{v}\|^2\right) \\
\implies U^0 &= \frac{1}{\sqrt{1 - \|\boldsymbol{v}\|^2}} = \gamma\,.
\end{aligned}
$$

Just as we have defined the 4-velocity we are free to define the 4-acceleration:

Definition 3.4. The *4-acceleration* of a timelike curve Γ is simply

$$A^\mu = \frac{\partial U^\mu(\tau)}{\partial \tau}\,.$$

Given the normalisation condition on the 4-velocity we find,

$$\frac{\partial}{\partial \tau}\left(\eta_{\mu\nu}U^\mu U^\nu\right) = 0,$$

$$\implies \eta_{\mu\nu}U^\mu A^\nu = 0\,.$$

This then implies the 4-velocity and 4-acceleration are orthogonal. This means essentially that the 4-acceleration is purely spatial. To see this, swap to a simple rest frame of the particle and use the orthogonality condition to find that the timelike component of A is 0.

3.1.1 Curves of Extremal Proper Time:

I hope you're familiar with the action functional and the concept that the equations of motion come from extremising said action. Essentially, in this case, we are looking for a curve such that $\delta\tau = 0$.

Let's look at how this would work if we vary the action with respect to the curve $\Gamma(s)$. For this, we need functional calculus. Just treat it the same as regular calculus, but now we are operating on functions of functions.

$$\tau[t] = \int \mathrm{d}s \, \sqrt{-\eta_{\mu\nu}\dot\Gamma^\mu(s)\dot\Gamma^\nu(s)}$$

$$\frac{\delta\tau}{\delta\dot\Gamma^\beta(\sigma)} = \int \mathrm{d}s \frac{1}{\sqrt{-\eta_{\mu\nu}\dot\Gamma^\mu(s)\dot\Gamma^\nu(s)}} \cdot \delta^\nu_\beta \, \eta_{\mu\nu}\dot\Gamma^\mu(s)\delta(s-\sigma)$$

$$\frac{\delta\tau}{\delta\dot\Gamma^\beta(\sigma)} = \frac{1}{\sqrt{-\eta_{\mu\nu}\dot\Gamma^\mu(\sigma)\dot\Gamma^\nu(\sigma)}} \cdot \dot\Gamma_\beta(\sigma)\,.$$

Similar to finding the extrema in regular calculus we demand that the variation, $\delta\tau = 0$. Thus, using the Euler-Lagrange equations, we find

$$\frac{\partial}{\partial\sigma}\left(\frac{1}{\sqrt{-\eta_{\mu\nu}\dot\Gamma^\mu(\sigma)\dot\Gamma^\nu(\sigma)}} \cdot \dot\Gamma_\beta(\sigma)\right) = 0$$

$$\implies \ddot\Gamma_\beta(\sigma) = 0\,.$$

where the derivative of the square root vanished due to the normalisation condition $\eta_{\mu\nu}\dot\Gamma^\mu(\tau)\dot\Gamma^\nu(\tau) = -1$. Since, for a timelike curve, we are free to set $\sigma = \tau$, we have (raising the index)

$$\ddot\Gamma^\mu(\tau) = 0 \implies \Gamma^\mu(\tau) = a^\mu\tau + b^\mu\,.$$

12

This means that curves are straight lines (affinely parametrised): i.e., particles follow straight paths, at least for flat space. Note, this equation can also be derived from the action functional of a relativistic particle.

3.2   Energy-Momentum relations:

You are probably aware that energy and momentum are often combined into a single tensor (along with shear and pressure): the Energy-momentum tensor, $T_{\mu\nu}$. In special relativity (i.e. ignoring any curvature) this can be defined as follows:

Definition 3.5. The *Energy-Momentum* tensor of a Lagrangian density $\mathcal{L}$, which is dependent on the fields ϕ_a, $a = 1, \cdots, n$, and their derivatives, is

$$T^\mu_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \partial_\nu \phi_a - \delta^\mu_\nu \mathcal{L}\,.$$

Let's look at the case where we have a relativistic point particle of mass m parametrized by τ, and localized at a point y. In this case we have the energy-momentum tensor,

$$T^{\mu\nu} = m \int \mathrm{d}\tau \; U^\mu(\tau) U^\nu(\tau) \delta^4(x - y(\tau))\,.$$

It is a distribution with support along the world-line parametrized by τ. Using the energy-momentum tensor, we are able to form the 4-momentum,

---

Definition 3.6. Given a stress-energy tensor $T^{\mu\nu}(x)$ on Minkowski space, the total *4-momentum* of a localized system at time t is

$$P^\nu(t) := \int_{\Sigma_t} T^{0\nu}(t, \boldsymbol{x})\, d^3x,$$

where $\Sigma_t = \{(t, \boldsymbol{x}) \,|\, \boldsymbol{x} \in \mathbb{R}^3\}$ is a constant-time hypersurface. In particular,

$$E = P^0 = \int T^{00}\, \mathrm{d}^3x, \qquad P^i = \int T^{0i}\, \mathrm{d}^3x\,.$$

This then allows us to compute the 4-momentum for the above energy-momentum

tensor, via

$$E = \int_{\mathbb{R}^3} d^3x \ T^{00}$$

$$= m \int_{\mathbb{R}^3} \int d^3x \, d\tau \ U^0 U^0 \delta(t - y^0(\tau)) \delta^3(\boldsymbol{x} - \boldsymbol{y}(\tau))$$

$$= m \int_{\mathbb{R}^3} d^3x \, d\tau \ \frac{\delta(\tau - \tau')}{\left| \frac{dy^0(\tau')}{d\tau} \right|} U^0 U^0 \delta^3(\boldsymbol{x} - \boldsymbol{y}(\tau))$$

$$= m U^0 = m\gamma \,,$$

where τ' is a root of the function $f(\tau) = t - x^0(\tau)$. Via a similar computation, we find the 3-momentum to be $P^i = m U^i = \gamma m v^i$ (check yourself). If we then take the inner product of the 4-momentum, we find,

$$P^\mu P_\mu = -E^2 + \|\boldsymbol{p}\|^2 \,.$$

Now, what is $P^\mu P_\mu$? Let's see.

Proposition 3.1 (Mass shell condition)**.** The 4-momentum P^μ of a particle of rest mass m satisfies

$$\eta_{\mu\nu} P^\mu P^\nu = -m^2.$$

Proof. For a point particle $P^\mu = m U^\mu$ with $U^\mu = \frac{dx^\mu}{d\tau}$. Since the 4-velocity is normalized $\eta_{\mu\nu} U^\mu U^\nu = -1$, it follows that

$$P_\mu P^\mu = m^2 \, U_\mu U^\mu = -m^2.$$

$\square$

Thus, we see that the inner product is

$$E^2 = m^2 + \|\boldsymbol{p}\|^2 \,.$$

Which is the famous relation (for the rest frame of the particle, where $\boldsymbol{p}$ is zero). For particles that do not obey the normalization condition, for example, virtual particles, the mass-shell condition does not hold. These particles are said to be *off-shell*. As this quantity is a Lorentz scalar, it is clearly invariant under Lorentz transformations (check this).

If we take the massless limit, we have $P_\mu P^\mu = 0$. This is due to the fact that, for a null world-line, $\eta_{\mu\nu} k^\mu k^\nu = 0$, where k^μ is the tangent vector to the curve. Which is parametrised by an arbitrary affine parameter, since the proper time does not exist.

3.3 The Centre of Momentum Frame:

A convenient frame of reference to use in particle collisions is the *centre of momentum frame.*

Definition 3.7. Suppose we have n particles each with an associated 3-momentum $\boldsymbol{p}_n$, then the *centre of momentum frame* is the inertial frame such that,

$$\boldsymbol{p}_{\text{tot}} = \sum_n \boldsymbol{p}_n = 0\,.$$

In such a frame, the norm of the total momentum is clearly zero. Hence, for a collection of particles, the centre of momentum frame is also the rest frame. This frame is useful in multi-particle collisions since we may set the sum of all momenta to be zero, and thus simplify our analysis. While it is good to know that it exists, it does not come up (I think at all) in the courses you wish to read.

4 Some Classical Electrodynamics:

The theory of Electrodynamics is one of the great theories of classical physics, formulated by the brilliant James Clerk Maxwell. It was, in fact, the first fully relativistic theory. Before Einstein's own theory. This is due to the formulation requiring that the speed of light be constant for all observers. This suggested, at the time, the existence of an "aether". Thus, explaining the need for the famous Michelson–Morley experiment. This issue was then later resolved by Einstein. His theory eliminated the need for an aether and showed that the invariance of the speed of light is a fundamental principle of nature.

4.1 Maxwell Theory:

To discuss classical Electrodynamics, we must first discuss Maxwell's theory of electromagnetism. As well as its gauge symmetries.

Definition 4.1. *Maxwell's Equations* are the following set of PDEs for the electric field $\boldsymbol{E}$ and magnetic field $\boldsymbol{B}$:

$$\nabla \cdot \boldsymbol{E} = \rho, \qquad\qquad \nabla \cdot \boldsymbol{B} = 0,$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \qquad\qquad \nabla \times \boldsymbol{B} = \boldsymbol{J} + \frac{\partial \boldsymbol{E}}{\partial t}\,,$$

where I have set $\varepsilon_0 = \mu_0 = 1$.

The divergence relation for the $\boldsymbol{B}$ field famously says there exists *no magnetic monopoles.* This, however, disagrees with Dirac's famous quantization condition:

which states that the quantization of the electric charge can be explained exactly if there were to exist a single monopole. There has, of course, never been any monopole detected. Despite this, Dirac's theoretical explanation remains the most widely accepted for providing a reason to why the electric charge is quantized.

From general definitions in vector calculus it is clear that the combination $\nabla \cdot (\nabla \times \boldsymbol{X}) = 0$. If you want to see this explicitly, write it in index notation and, assuming $\boldsymbol{X}$ is sufficiently smooth, notice that due to the anti-symmetry in the cross product and symmetry in partials, the total must be zero. This insight allows us to write $\mathbf{B}$ in terms of a more fundamental object, the magnetic potential $\boldsymbol{A}$, as

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}\,.$$

Similarly, the Electric field can be written as

$$\boldsymbol{E} = -\nabla\phi - \frac{\partial \boldsymbol{A}}{\partial t}\,.$$

The ϕ term is the electric potential. It is a scalar field, similar in nature to the formulation of force in terms of the potential energy, $\boldsymbol{F} = -\nabla U$. We can actually package it all up neatly into a single 4-vector, the *Gauge potential*,

$$A^\mu = (\phi, \boldsymbol{A})\,.$$

It is called this since it is the prototypical example of a gauge theory. In fact, all gauge theories follow essentially the same logic as EM. Now, a gauge is merely a redundancy in our description. The underlying physics should not depend on the choice of gauge. Think, as an example, of a weighing scale. It does not matter whether we assign a function to measure the weight of someone in Kg or lbs, the underlying physical laws should remain unchanged. This is to say that the "observables" of the theory (things we can actually measure in our lab) have to be gauge invariant. I should emphasize that a gauge transformation is *not* a symmetry. It is a redefinition of the fields themselves.

In this manner, the gauge potential is not unique. We are free to add a smoothly differentiable function $\omega(t, \boldsymbol{x})$, such that

$$A^\mu \to A^\mu + \partial^\mu \omega\,,$$

where

$$\partial^\mu = (-\partial_t, \nabla)\,,$$

16

the 4-derivative. Does this transformation really leave the physics unchanged? Well, yes. I will quickly show this.

Proposition 4.1. Both the $\boldsymbol{E}$ and $\boldsymbol{B}$ fields are unchanged under this (gauge) transformation.

Proof. Under the transformation

$$A^\mu \to A'^\mu = A^\mu + \partial^\mu \omega \,,$$

we can see that the $\boldsymbol{E}$ field transforms as

$$\begin{aligned} \boldsymbol{E} \to \boldsymbol{E}' &= -\nabla(\phi - \partial_t \omega) - \partial_t(\boldsymbol{A} + \nabla\omega) \\ &= -\nabla\phi - \partial_t \boldsymbol{A} + \partial_t(\nabla\omega - \nabla\omega) \\ &= \boldsymbol{E} \,. \end{aligned}$$

It is a similar story for the $\boldsymbol{B}$ field, it transforms as

$$\begin{aligned} \boldsymbol{B} \to \boldsymbol{B}' &= \nabla \times (\boldsymbol{A} + \nabla\omega) \\ &= \nabla \times \boldsymbol{A} + (\nabla \times \nabla\omega) \\ \implies \boldsymbol{B} &= \boldsymbol{B}' \,, \end{aligned}$$

since $\nabla \times \nabla\omega = 0$. $\qquad\qquad\square$

This is good! The things we can actually measure using EMF devices are completely gauge invariant. This redundancy allows us to the *choose a gauge*. Certain gauges will make calculations much simpler. Common gauge include:

- The *Lorenz Gauge*,
$$\partial_\mu A^\mu = 0.$$

  We are always free to choose this gauge, to see this notice that if we are given a gauge field A'^μ such that $\partial_\mu A'^\mu = f(t, \boldsymbol{x})$, we can then choose ω such that
$$\partial_\mu \partial^\mu \omega = -f \,.$$

  Our gauge potential is then
$$A^\mu = A'^\mu + \partial^\mu \omega,$$

  with $\partial_\mu A^\mu = 0$. These equations always have solutions. The Lorenz gauge has the advantage that it is Lorentz invariant.

- The *Coulomb Gauge*,
$$\nabla \cdot \boldsymbol{A} = 0.$$

  We are free to choose this gauge using the same argument as for the Lorenz gauge. It turns out that the A_0 component is purely specified by $\nabla \cdot \boldsymbol{A}$, we may set
$$A_0 = 0 \,.$$

This condition is called the *temporal gauge.* In general the Coulomb gauge requires $A_0 =$ fixed, by Gauss's Law. The remaining three components of A^μ satisfy one constraint leaving only two degrees of freedom (dof). In terms of EM theory, these two dof align with the two polarisation states of the photon. Sadly, unlike the Lorenz gauge, the Coulomb gauge is not Lorentz invariant.

Within field theory it is typical to package the Electric and magnetic fields into a single object defined as follows:

Definition 4.2. The *Field-Strength tensor* is an anti-symmetric $(2,0)$ tensor, written in terms of the gauge potential as

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu .$$

It should be clear to see that this tensor is completely gauge invariant (please check yourself). In terms of $\boldsymbol{E}$ and $\boldsymbol{B}$, the (contravariant) Field-Strength tensor can be written as

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} .$$

Maxwell's equations can then be written neatly in terms of $F^{\mu\nu}$ via the introduction of a *4-current,*

$$J^\mu = (\rho, \boldsymbol{J}) ,$$

such that

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

Notice, however, that $F^{\mu\nu}$ satisfies

$$\partial_\mu \partial_\nu F^{\mu\nu} = 0 .$$

Thus the 4-current must satisfy the continuity condition

$$\partial_t \rho + \nabla \cdot \boldsymbol{J} = 0 .$$

Written another way, this implies

$$\partial_\mu J^\mu = 0.$$

A current that satisfies this is said to be *conserved*. This relation can be derived more thoroughly from, arguably, the most important theorem in physics: Noether's Theorem. I wont cover Noether's theorem here as it is covered extensively and excellently in David Tong's lecture notes on QFT. But, it essentially boils down to the following,

$$Continuous\ Symmetry \implies Conserved\ Current.$$

Note, this is not a two way relation, i.e. the converse is not always true. Also, this relation only holds classically and on-shell. In the quantum theory, classically conserved currents can acquire *anomalies*.

For the remaining Maxwell equations we use something called the *Bianchi Identity* to write them as

$$\partial_{[\mu}F_{\nu\rho]} = 0\,.$$

4.2   EM is a Lorentz Covariant Theory:

As we alluded to earlier, the theory of Electromagnetism is a relativistic theory. This becomes obvious once we introduce the gauge potential and field-strength tensor. This is because they transform naturally under Lorentz transformations. For example, the gauge potential transforms as

$$A^{\mu} \to A'^{\mu} = \Lambda^{\mu}_{\nu}A^{\nu}\,.$$

In the language of representation theory, we would say it transforms under the vector (sometimes called *defining*) rep of $SO^{+}(1,3)$.

Remark 4.1. *If you want the formal reason for this, it is because the gauge potential is a 1-form, so can be written as*

$$A = A_{\mu}dx^{\mu}\,,$$

where dx^{μ} is a dual basis (of 1-forms). In this way we have

$$dx^{\mu} \to \Lambda^{\mu}_{\nu}dx^{\nu}\,,$$

hence A transforms in the way seen above.

Now, in a similar fashion the field-strenth tensor also transforms under Lorentz transformations. But, not in the same way. Since it is a (2,0) tensor, we see it takes two covectors as inputs,

$$F^{\mu\nu} = F(dx^{\mu}, dx^{\nu})\,,$$

under a Lorentz transformation we then have

$$F^{\mu\nu} \to F'^{\mu\nu} = \Lambda^\mu_\rho F^{\rho\eta} \Lambda^\nu_\eta \,.$$

In vector notation we have $F' = \Lambda F \Lambda^{-1}$. It is said to transform in the *adjoint* rep of $SO^+(1,3)$. Given that they transform under Lorentz transformations, we can form Lorentz invariant quantities. Which is great when we want to consider field theories. Since, the Lagrangian densities (and hence actions) we build should be completely Lorentz invariant. This fact also constrains the form of the action to Lorentz invariants built from the field-strength.

It is likely best too see and example. One common Lagrangian seen in QFT is given by,

$$\mathcal{L} = F_{\mu\nu} F^{\mu\nu} \,.$$

This quantity is a Lorentz scalar, which is exactly what we need. This is the Lagrangian density used in QED (up to constants).

5 Conclusion:

These are essentially all the tools you need to jump straight into QFT and possibly AQFT if you're feeling brave. There is of course, many things that I have left out of these notes to avoid bloat. But, this is about all the material you need to be in line with the requirements given on the Part III webpage. To test your understanding I suggest doing this exercise sheet for fun.