Introduction to Relativistic Particle Mechanics

Lawrence Frolov

November 2022

1 Special Relativity

1.1 Introduction to Special Relativity

Like all other physicists, each day I wake up and pray to the gods slumbering below that the laws of physics have not changed overnight. Down would be up, toasters would cease to function, and my local bagel place would be forced to shutdown. There is one law of physics that particularly keeps me up at night. In vacuum, electromagnetic fields satisfy the wave equation with propagation speed $c = 1$, i.e

$$\partial_t^2 \vec{E} + \Delta \vec{E} = 0. \tag{1}$$

What's terrifying about this? Well, according to the laws of electromagnetism, this equation must hold in all frames of reference. Which means electromagnetic disturbances will seem to propogate at the speed of light, no matter what frame of reference you're in.

To drive this insanity home, consider a lake which has been perturbed by a single raindrop. To a person sitting on the shore, the surface of the water satisfies the 2-dimensional wave equation, so the droplet will create a ripple that propagates at some speed in all directions. However, a person sailing on a boat will see different parts of the ripple moving at different speeds. For instance, if the sailor is moving at the same speed as the ripple, some parts of the ripple will appear stationary while the other side will appear to be moving twice as fast.

Maxwell's equation defies common sense. It says regardless the frame of reference you're in, every part of the ripple will move at the same speed, the speed of light. It also defies the principles of Galilean relativity, it is not invariant under Galilean transformations.

Definition 1.1. *Let $(t, \vec{x})$ be the standard coordinate system for $\mathbb{R} \times \mathbb{R}^n$. Then* ***Galilean transformations*** *are given by translations, rotations, and reflections in space, along with the transformation corresponding to uniform motion of velocity $\vec{v}$:*

$$(t, \vec{x}) \to (t, \vec{x} + t\vec{v}) = (t', \vec{x}'). \tag{2}$$

So, at this point we have two options. Throw away Galilean relativity, or allow my local bagel place to shut down. The choice is clear, we will replace Galilean transformations/Galilean Relativity with Lorentz transformations/Special Relativity.

Definition 1.2. *Let $(t, \vec{x})$ be the standard coordinate system for $\mathbb{R} \times \mathbb{R}^n$. All* ***Lorentz transformations*** *are given by translations, rotations, and reflections in space, along with the transformation corresponding to uniform motion of velocity $\vec{v}$*

$$(t, \vec{x}) \to \gamma(t - \vec{v} \cdot \vec{x}, \vec{x} - \vec{v}t) = (t', \vec{x}') \tag{3}$$

where $\gamma := \frac{1}{\sqrt{1-|\vec{v}|^2}}$, and the speed of light has been set to 1.

Lorentz transformations are derived by taking the speed of light to be constant in all reference frames, a good video about them can be found here: https://www.youtube.com/watch?v=feBT0Anpg4A. Special relativity takes its laws of nature to be invariant under Lorentz transformations. Maxwell's laws of electromagnetism are one example. Special relativity is a bit of a departure from Galilean relativity, and is much easier to digest after one throws away the idea that space and time are separate. As we can see from the Lorentz transformation, space and time coordinates can "mix" together in the same way that space coordinates "mix" together under spatial rotations. In fact, one can view Lorentz transformations as hyperbolic rotations in space-time by introducing $\cosh(\alpha) = \gamma$, $\vec{\tanh}(\alpha) = \vec{v}$. Lorentz transformations can then be written as

$$\begin{pmatrix} t \\ \vec{x} \end{pmatrix} \to \begin{pmatrix} \cosh(\alpha) & -\vec{\sinh}(\alpha) \\ -\vec{\sinh}(\alpha) & \cosh(\alpha) \end{pmatrix} \begin{pmatrix} t \\ \vec{x} \end{pmatrix} \tag{4}$$

Uniform motion now translates into a "hyperbolic rotation" of space-time! Unlike spatial rotations, Lorentz transformations do not preserve lengths in space or time, but they do preserve something akin to a "hyperbolic space-time length".

Proposition 1.1. *The proper time length given by $d\tau = \sqrt{dt^2 - dx^2}$ is preserved under all Lorentz transformations.*

We call this the "proper time" because if the space-like separation between two events is 0, then $d\tau = dt$ measures the time passed. However, when one views the events from a different frame of reference, the space-like separation is no longer zero, and the time-like separation has also changed due to the nature of Lorentz transformations. Proper time lengths do not change under Lorentz transformations, which means all observers will agree upon the proper time length between any two events in space-time. Proper time is now the only way to objectively talk about distances in special relativity, and is frequently used to parameterize the trajectories or "world-lines" of particles.

Proposition 1.2. *Let $(t(\theta), \vec{x}(\theta)) : \mathbb{R} \to \mathbb{R} \times \mathbb{R}^n$ be an arbitrary parametization of a particle's trajectory in space-time. Since the particle's speed is capped by* *the speed of light, we have that* $|\frac{d\vec{x}}{d\theta}| < \frac{dt}{d\theta}$. *So the proper time function defined by*

$$\tau(\theta) := \int_0^\theta \sqrt{(\frac{dt}{d\theta})^2 - |\frac{d\vec{x}}{d\theta}|^2} d\theta' \quad (5)$$

is strictly increasing and thus a valid parameterization. It is independent of reference frame.

Remark 1.1. *The proper time function measures the amount of time the particle experiences in its own frame of reference.*

1.2 Special Relativity in the language of Manifolds

Where do particles, fields, and all other things which make up the universe live? Typically we think of massive objects as existing in space, as subsets of $\mathbb{R}^3$. Special relativity forces us to abandon this mindset, and invites us to try thinking about time in the same way we think about space. There is a sense in which we can think of objects not as subsets of $\mathbb{R}^3$, but as subsets of $\mathbb{R} \times \mathbb{R}^3$. For instance, instead of thinking of a particle as a single point in space, you can think of it as a line in space-time tracing its trajectory. The position of a particle is not just its location in space, but also its location in time. Thus, we are motivated to start thinking of the universe as a four dimensional space-time manifold.

Definition 1.3. *We define* $\mathbb{R}^{1,n}$ *as the set of all space-time positions*

$$\mathbb{R}^{1,n} := \{x^\mu | x^0 \in \mathbb{R}, \vec{x} \in \mathbb{R}^n\}. \quad (6)$$

Each relativistic position vector has $n+1$ *components, with "time" as the zeroth component.*

Take a point in a manifold $\mathcal{M}$, and consider the set of vectors emanating from that point which still lies tangent to the manifold. We call this the tangent space at that point $T_p\mathcal{M}$. For instance, if your manifold is the surface of a sphere embedded in 3 dimensional space, and your point is the north pole, then any vector pointed in the z direction won't be tangent to the manifold. But any vector residing in the plane tangent to the sphere's surface at the north pole will be.

Since our manifold is secretly just $\mathbb{R}^{n+1}$, its tangent space is also $\mathbb{R}^{n+1}$. Vectors now not only have a direction in space, but in time as well. A change in reference frame causes the components of a $n+1$ vector to transform according to the Lorentz transformations described in the last section.

To define a notion of distance, we equip our manifold with a metric invariant under Lorentz transformations.

Definition 1.4. *Define the bilinear form* $\eta : T_p\mathbb{R}^{1,n} \times T_p\mathbb{R}^{1,n}$ *via*

$$\eta(v, w) := \eta_{\mu\nu} v^\mu w^\mu := v^0 w^0 - \vec{v} \cdot \vec{w}. \quad (7)$$

This metric is essentially an inner (dot) product for space-time vectors. We define the magnitude of a space-time vector via $||v^\mu||^2 := \eta_{\mu\nu} v^\mu v^\nu$.

2 Particle Dynamics in Special Relativity

2.1 Notation

$$\mathbb{R}^{1,n} := \{x^\mu | x^0 \in \mathbb{R}, \vec{x} \in \mathbb{R}^n\} \tag{8}$$

$$\eta_{\mu\nu} := \eta^{\mu\nu} := \begin{pmatrix} 1 & 0 \\ 0 & -\mathbb{I}_n \end{pmatrix} \tag{9}$$

Whenever you see an index repeated, with one upper and one lower, it means we are summing over the index. We call this contraction. So

$$v^\nu w_\nu := \sum_{\nu=1}^{n} v^\nu w_\nu. \tag{10}$$

Given a vector v^μ, we define its dual covector v_ν by contracting with the metric

$$v_\nu := \eta_{\mu\nu} v^\mu = \sum_{\mu=1}^{n} \eta_{\mu\nu} v^\mu. \tag{11}$$

The magnitude of a vector and its dual is given by

$$||v^\mu|| := \sqrt{\eta_{\mu\nu} v^\mu v^\nu} = \sqrt{v_\nu v^\nu} = ||v_\mu||. \tag{12}$$

The components of a vector v^μ abide a certain transformation law under changes in frame of reference. Since the quantity $v_\nu v^\nu$ has no free indices it is invariant under changes in frame of reference.

Lastly, for a given field $\phi : \mathbb{R}^{1,n} \to \mathbb{R}$ define its gradient via

$$\partial_\mu \phi := (\partial_0 \phi, \partial_1 \phi, ..., \partial_n \phi) \tag{13}$$

Notice that $\partial_\mu \phi$ is a covector because given a vector v^μ, the derivative of ϕ in the direction of v^μ is invariant of reference frame and satisfies

$$\frac{d}{dt}(\phi(x^\mu + tv^\mu))\big|_{t=0} = \sum_{\mu=1}^{n} v^\mu \partial_\mu \phi = v^\mu \partial_\mu \phi. \tag{14}$$

Therefore, we may raise the index of $\partial_\mu \phi$ by contracting with the metric $\eta^{\mu\nu}$.

2.2 Motion

We define a point particle as a line, typically referred to as a "world-line" $\Gamma \subset \mathbb{R}^{1,3}$, equipped with some mass parameter m. We may parameterize this line, and write

$$\Gamma = \{z^\mu(\theta) | \theta \in \mathbb{R}\}. \tag{15}$$

Differentiating z^μ with respect to θ returns something we'd like to think of as the particle's velocity. However, the definition suffers from its dependence on the parameterization. For particles traveling under the speed of light, there is a canonical parameterization of the particle's world-line which will return a very intuitive definition for velocity.

Proposition 2.1. *Let $z^\mu(\theta) : \mathbb{R} \to \mathbb{R}^{1,3}$ be an arbitrary parametization of a particle's trajectory in space-time. Since the particle's speed is capped by the speed of light, we have that $|\frac{d\vec{z}}{d\theta}| < \frac{dz^0}{d\theta}$. So the proper time function defined by*

$$\tau(\theta) := \int_0^\theta \sqrt{\eta_{\mu\nu}\dot{z}^\mu \dot{z}^\mu} d\theta' \tag{16}$$

is strictly increasing and thus a valid parameterization. It is independent of reference frame and original parameterization (up to translation).

Remark 2.1. *Proper time measures the amount of time the particle experiences in its own frame of reference. It's powerful in that it can be computed in any reference frame.*

Definition 2.1. *Define the particle's four velocity as $u^\mu := \frac{dz^\mu}{d\tau}$.*

We have established a notion of velocity, but it is unclear what it means for a velocity to have four components. Three components indicate the particle's rate of change in spatial position, while the 0th component represents the rate of temporal change. To illustrate this, consider fixing a frame of reference, so that we can parameterize our world-line with respect to z^0.

Proposition 2.2. *Fix a frame of reference, and let $v^\mu := \frac{dz^\mu}{dz^0}$. Then*

$$u^\mu = \frac{1}{\sqrt{1 - |\vec{v}|^2}} v^\mu. \tag{17}$$

Thus, u^0 is always greater than or equal to 1.

Remark 2.2. *When $u^0 > 1$, the rate at which time is experienced by the particle is less than the rate at which the observer experiences time. This gives rise to the phenomenon known as time-dilation.*

Equation (17) follows from the fact that u^μ always satisfies $||u^\mu|| = \sqrt{\eta_{\mu\nu}u^\mu u^\nu} = 1$. This is perhaps not too surprising, given that $u^\mu = (1, 0, 0, 0)$ in the particle's frame of reference, and magnitude is invariant of reference frame.

Remark 2.3. *Notice that point particle dynamics in special relativity do not have more degrees of freedom than in the non-relativistic setting. Although our particle's velocity now has four components we wish to solve for, it only has three degrees of freedom.*

This also places a vital restriction on the acceleration that any particle undergoes.

Definition 2.2. *Define the particle's four-acceleration as $a^\mu = \frac{du^\mu}{d\tau}$.*

Proposition 2.3. *To preserve the length of the particle's four velocity, we must have that $\eta(a^\mu, u^\mu) = \eta_{\mu\nu}a^\mu u^\nu = 0$. In other words, the particle's four acceleration must always be perpendicular to its four-velocity.*

2.3 Energy-Momentum

With our notion of velocity in hand, we may begin to discuss momentum in special relativity.

Definition 2.3. *Define the particle's four momentum vector as* $p^\mu := mu^\mu$.

Once again, it's time to play the game of "what does the 0th component represent?" The prize will be a deeper understanding of the universe. Since $||u^\mu||^2 = 1$, we have that $(p^0)^2 - |\vec{p}|^2 = m^2$. So whatever p^0 is, it can be related to $\vec{p}$ via $p^0 = \sqrt{m^2 + |\vec{p}|^2}$. We will take two approaches to uncovering the meaning of p^0.

Firstly, we will consider the non-relativistic limit, i.e fix a frame where $|\vec{v}| << 1$. Now, let's Taylor expand p^0 in terms of $|\vec{v}|$, and only consider the lowest order term. Doing this returns $p^0 = m + \frac{1}{2}m|\vec{v}|^2 + O(|\vec{v}|^2)$.

Alternatively, consider a relativistic particle undergoing no forces. In any fixed frame, we should have $\frac{d\vec{v}}{dz^0} = 0$. However, since this is true for all reference frames, we must have that $\frac{du^\mu}{d\tau} = 0$. Multiplying this equation by m, we see that the spatial parts of that equation tells us that momentum $\vec{p}$ is conserved. But, we also now have that for free particles, p^0 is conserved. What else, other than momentum, could it be?

Energy, it's energy. But the fact that it's energy is surprising, because in the non-relativistic limit we got an additional m term. This tells us that relativistic particles at rest have energy proportional to their mass.

To recap, energy and momentum are part of the same four-vector. We have always known that there exists a relationship between the two, but their connection is now deeper. In the same way that time and space mix under changes in frame of reference, so do energy and momentum.

2.4 Relativistic Forces and Orthogonality

Definition 2.4. *The relativistic force acting on a particle is a four-vector satisfying* $F^\mu = \frac{dp^\mu}{d\tau}$.

From our earlier discussion, F^μ must satisfy $\eta_{\mu\nu}F^\mu u^\nu = 0$ to be an admissible relativistic force law. In principle, you can think of this restriction as saying that the rates of change of momentum and energy must be related.

Corollary 2.1. *In a fixed frame of reference, the four force must be of the form* $F^\mu = \frac{1}{\sqrt{1-v^2}}(\vec{v}\cdot\vec{f}, \vec{f})$. *The rates of change of energy and momentum are related via* $\frac{dE}{dz^0} = \vec{v}\cdot\vec{f}$ *and* $\frac{d\vec{p}}{dz^0} = \vec{f}$. *This is recognizable as the equation for power.*

Remark 2.4. *It is possible to write down a force law which seems Lorentz invariant, but which does not satisfy the orthogonality condition. Physicists often forget this.*

An example of a force law which does satisfy this condition is the Lorentz force. Given a vector potential $A_\mu : \mathbb{R}^{1,3} \to \mathbb{R}^{1,3}$, we may define its Faraday tensor $F_{\mu\nu}$ via

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{18}$$

Remark 2.5. *By the anti-symmetry of $F_{\mu\nu}$, the Lorentz force law given by $F^{\mu} = eF^{\mu\nu}u_{\nu}$ is admissible.*

This shows us that there's a way to couple a relativistic particle to a vector potential, but what about a scalar potential? Let $\phi : \mathbb{R}^{1,3} \to \mathbb{R}$ be a scalar field (in physics this is typically associated with gravitational fields) and a be the particle's scalar charge. What kind of force law is admissible? In any fixed frame, we probably want something along the lines of $\vec{f} = -a\vec{\nabla}\phi$, like we do for Newtonian gravity.

Remark 2.6. *$F^{\mu} = -a\partial^{\mu}\phi$ is not an admissible force law for a relativistic particle, because there's no gaurantee that $\eta_{\mu\nu}u^{\mu}\partial^{\nu}\phi = 0$.*

This is a bit of an issue. At this point we can either throw our hands in the air and say that there is no scalar field theory of gravity which is compatible with special relativity, or we can do something weird. First, notice that $\eta_{\mu\nu}u^{\mu}\partial^{\nu}\phi = \frac{d\phi(z^{\mu}(\tau))}{d\tau}$.

Proposition 2.4. *$F^{\mu} = -a\partial^{\mu}\phi + au^{\mu}\frac{d\phi(z^{\mu}(\tau))}{d\tau}$ is an oddly specific but admissible force law for a relativistic particle coupled to a scalar field.*

The proposition above follows from the fact that $\eta_{\mu\nu}u^{\mu}u^{\nu} = 1$. But, I think most people will agree that the additional term seems highly unmotivated. We will write down the correct force law for a relativistic particle coupled to a scalar field after giving a brief introduction to relativistic Lagrangian mechanics.

3 Relativistic Lagrangian Mechanics

Lagrangian mechanics has already been adapted for particles that live on general manifolds in the non-relativistic setting. Previously, we considered an action functional $\mathcal{A}$ defined on a space of functions mapping a time interval onto some curve on a manifold $\mathcal{M}$. We showed that critical points of certain action functionals represent trajectories which satisfy Newton's laws of motion, and studied properties of these critical points. In the relativistic setting, trajectories will be replaced with world-lines on space-time, and $\mathcal{M}$ will be four dimensional. The "time interval" will now denote the interval over which we parameterize the particle's world-line $z^{\mu}(\theta)$. Our goal will be to study world-lines as critical points of some functional.

3.1 Brief Recap of Lagrangian Mechanics

Lagrangian mechanics is a framework which studies the trajectories of particles by viewing them as critical points of certain action functionals. Suppose your particle lives on a manifold $\mathcal{M}$ (which you can think of as being $\mathbb{R}^n$). Then the trajectory of your particle $z(t) : [0, T] \to \mathcal{M}$ lives in a function space.

Definition 3.1. *Denote the space of weakly once-differentiable maps from $[0, T]$ to $\mathcal{M}$ by $W^{1,1}([0, T], \mathcal{M})$.*

Consider a newtonian particle which is being acted upon by a conservative force of the form $\vec{f} = -\vec{\nabla} V$.

Definition 3.2. *We define the Lagrangian of this system as*

$$L(z, \vec{v}) := \frac{1}{2}\vec{v} \cdot m\vec{v} - V(z). \tag{19}$$

It turns out that critical points of action functionals constructed using the Lagrangian obey Newton's law of motion.

Definition 3.3. *Given z_0, z_T, define the action functional $\mathcal{A} : W \to \mathbb{R}$ via*

$$\mathcal{A}(z(\cdot)) = \int_0^T L(z(t), \frac{dz}{dt}(t))dt \tag{20}$$

where W is defined as

$$W := \{z(\cdot) \in W^{1,1}([0, T], \mathcal{M}) | z(0) = z_0, z(T) = z_T\}. \tag{21}$$

Theorem 3.1. *The critical points of the action functional given above satisfy the Euler-Lagrange equations*

$$\frac{d}{dt}\partial_{\vec{v}} L(z(t), \frac{dz}{dt}(t)) = \partial_z L(z(t), \frac{dz}{dt}(t)). \tag{22}$$

Proof. This proof will cover the case that $\mathcal{M} = \mathbb{R}^n$. $z(\cdot)$ is a critical point iff

$$\frac{d}{d\epsilon} A(z(\cdot) + \epsilon w(\cdot))\big|_{\epsilon=0} = 0 \tag{23}$$

for all $w(\cdot) \in W^{1,1}$ such that $w(0) = 0 = w(T)$. Those boundary conditions are necessary to ensure $z(\cdot) + \epsilon w(\cdot) \in W$. Using the integral definition of $\mathcal{A}$ and integrating by parts returns

$$0 = \int_0^T w(t) \cdot \big(\partial_z L(z(t), \frac{dz}{dt}(t)) - \frac{d}{dt}\partial_{\vec{v}} L(z(t), \frac{dz}{dt}(t))\big)dt. \tag{24}$$

We conclude by adhering to a cool lemma:

Lemma 3.1. *If a function u satisfies*

$$\int_0^T \phi \cdot u dt = 0 \tag{25}$$

for all smooth, compactly supported functions ϕ on $(0, T)$, then $u = 0$.

□

For the Lagrangian of a particle in a potential V, the Euler-Lagrange equations reduce to

$$\frac{d}{dt}(m\frac{dz}{dt}) = -\nabla V \tag{26}$$

as desired. So, one way of studying trajectories of particles is by viewing them as critical points of action functionals. In practice we use this formulation to study conservation laws for quantities such as momentum and energy.

Definition 3.4. *The canonical momentum of a point particle with dynamics given by Lagrangian $L(z, \vec{v})$ is*

$$p = \partial_{\vec{v}} L. \tag{27}$$

3.2 Formulation of Relativistic Lagrangian Mechanics

Consider a free relativistic particle with arbitrary parameterization variable θ.

Definition 3.5. *We define the Lagrangian of the relativistic particle as*

$$L_p(z^\mu, \dot{z}^\mu) := -m\sqrt{\eta_{\mu\nu}\dot{z}^\mu \dot{z}^\nu}. \tag{28}$$

Definition 3.6. *Given z_1, z_2, we define the action functional $\mathcal{A} : W_p \to \mathbb{R}$ via*

$$\mathcal{A}(z^\mu(\cdot)) := \int_{\theta_1}^{\theta_2} L(z^\mu(\theta), \dot{z}^\mu(\theta))d\theta \tag{29}$$

where

$$W_p = \{z(\cdot) \in W^{1,1}([\theta_1, \theta_2], \mathbb{R}^{1,3}) | z(\theta_1) = z_1, z(\theta_2) = z_2\} \tag{30}$$

Remark 3.1. *The action functional of a free particle is proportional to the amount of proper time passed during the interval $[\theta_1, \theta_2]$.*

Proposition 3.1. *The critical points of the action functional satisfy the Euler Lagrange equations*

$$\frac{d}{d\theta}(\partial_{\dot{z}^\mu} L(z^\mu(\theta)), \dot{z}^\mu(\theta)) = \partial_z L(z^\mu(\theta), \dot{z}^\mu(\theta)) \tag{31}$$

Proof. Same as the non-relativistic case. □

Proposition 3.2. *For the free particle Lagrangian, the Euler-Lagrange equations read*

$$\frac{d}{d\theta}\left(\frac{m\dot{z}_\mu}{\sqrt{\eta_{\alpha\beta}\dot{z}^\alpha \dot{z}^\beta}}\right) = 0. \tag{32}$$

The world-line which acts as a critical point for the action functional satisfies equation (32). To see that this represents a free particle, consider reparameterizing this world-line by its proper-time function. Reparameterizing returns

$$\frac{d}{d\tau}(mu^\mu) = 0 \tag{33}$$

as desired. But how do we add in a conservative force? We'd like our final equation of motion to be

$$\frac{d}{d\theta}\left(\frac{m\dot{z}_\mu}{\sqrt{\eta_{\alpha\beta}\dot{z}^\alpha \dot{z}^\beta}}\right) = -\partial_\mu \phi \tag{34}$$

for some field ϕ. Working backwards, the Euler Lagrange equations imply that

$$L(z^\mu, \dot{z}^\mu) = -m\sqrt{\eta_{\mu\nu}\dot{z}^\mu \dot{z}^\nu} + \phi. \tag{35}$$

3.3 Impossibility of conservative forces in relativistic lagrangian mechanics

Definition 3.7. *Define the proper-time parameterization function* $\tau : W_p \times [\theta_1, \theta_2] \to \mathbb{R}$ *via*

$$\tau(z^\mu, \theta) = \int_{\theta_1}^{\theta} \sqrt{\eta_{\mu\nu}\dot{z}^\mu \dot{z}^\nu} d\theta. \tag{36}$$

Definition 3.8. *Define the space of weakly differentiable time-like world-lines via*

$$\mathcal{T}_p = \{z^\mu(\cdot) \in W_p | \tau(z^\mu, \theta) \text{ strictly increasing in } \theta\} \tag{37}$$

Theorem 3.2. *The action functional* $\mathcal{A} : \mathcal{T}_p \to \mathbb{R}$ *given by*

$$\mathcal{A}(z^\mu(\cdot)) := \int_{\theta_1}^{\theta_2} -m\sqrt{\eta_{\mu\nu}\dot{z}^\mu \dot{z}^\nu} + \phi(z(\theta))d\theta \tag{38}$$

has no critical points in its domain for most scalar fields ϕ*.*

Proof. Suppose for the sake of contradiction that there exists a critical point $w^\mu(\cdot)$ of this functional. Then it must satisfy the Euler-Lagrange equations, which for this Lagrangian returns

$$\frac{d}{d\theta}\left(\frac{m\dot{w}_\mu}{\sqrt{\eta_{\alpha\beta}\dot{w}^\alpha \dot{w}^\beta}}\right) = -\partial_\mu \phi \tag{39}$$

Notice that the vector $\frac{\dot{w}_\mu}{\sqrt{\eta_{\alpha\beta}\dot{w}^\alpha \dot{w}^\beta}}$ must always be of unit length, so the equation can only hold if ϕ satisfies the orthogonality condition $\dot{w}^\mu \partial_\mu \phi = \frac{d\phi(w^\mu(\theta))}{d\theta} = 0$ over the interval. In other words, the only critical points of the action functional are world-lines along which ϕ are constant. □

4 Consequences of Relativistic Lagrangian Mechanics

4.1 Canonical Energy-Momentum

We define the canonical energy-momentum covector for relativistic point particles in the same way we do in the non-relativistic case:

Definition 4.1. *Define the canonical Energy-Momentum co-vector* $p_\mu = -\frac{\partial L}{\partial \dot{z}^\mu}$

Remark 4.1. *The negative sign is due to our choice of metric signature.*

Proposition 4.1. *For the free particle Lagrangian we have that*

$$p_\mu = m\frac{\dot{z}_\mu}{\sqrt{\eta_{\alpha\beta}\dot{z}^\alpha \dot{z}^\beta}} = mu_\mu. \tag{40}$$

4.2 Electromagnetically Charged Particles

The action for a point particle of charge e coupled to an electromagnetic vector potential $A_\mu = (\phi, \vec{A})$ is given by

$$L = L_p + L_{\text{int}} := -m\sqrt{\eta_{\mu\nu}\dot{z}^\mu \dot{z}^\nu} - eA_\mu \dot{z}^\mu \tag{41}$$

The Euler-Lagrange equations then read

$$\frac{d}{d\theta}(m\frac{\dot{z}_\mu}{\sqrt{\eta_{\mu\nu}\dot{z}^\mu \dot{z}^\nu}} + eA_\mu) = e\dot{z}^\nu \partial_\mu A_\nu \tag{42}$$

Notice that the canonical energy-momentum vector is not equal to mu_μ. Most physicists tend to ignore this using arguments regarding gauge invariance. Anyways, re-arranging the terms of the Euler-Lagrange equations and reparameterizing with respect to proper time returns the Lorentz force law

$$\frac{d}{d\tau}(mu_\mu) = eF_{\mu\nu}u^\nu. \tag{43}$$

4.3 Scalar Charged Particles

The action for a point particle of scalar charge a coupled to a scalar potential ϕ is given by

$$L = L_p + L_{\text{int}} := -m\sqrt{\eta_{\mu\nu}\dot{z}^\mu \dot{z}^\nu} + a\phi\sqrt{\eta_{\mu\nu}\dot{z}^\mu \dot{z}^\nu}. \tag{44}$$

This coupling is not as well-known as it should be, but is standard in the physics research community. It is one of the few ways to couple a scalar potential to a point particle in a way that is invariant of parameterization, and is Lorentz invariant. For instance, if we were to take $L_{\text{int}} = \dot{z}_\mu \partial^\mu \phi$, then this would return a theory equivalent to a particle coupled to a vector potential but with the Lorentz force being identically zero. The Euler-Lagrange equations return

$$\frac{d}{d\tau}((m - a\phi(z))u_\mu) = -a\partial_\mu \phi \tag{45}$$

Notice that this force law is essentially conservative, if you recall that canonical momentum is given by $(m - \phi(z^\mu))u_\mu = m(\tau)u_\mu$. In the context of scalar particles, we take the form of this canonical momentum as telling us something fundamental regarding the particle's mass $m(\tau)$. Mass is not constant for scalar particles, and their interaction with the field contributes to some amount of their mass. This is connected to why the strong force, something we believe originates from particles interacting with scalar fields, generates most of the mass of particles such as protons and nuetrons.

You'll notice that I originally motivated scalar potentials in the context of gravity, but have switched to discussing strong forces. We see that because interactions with scalar fields generate mass for point particles, this is not really the way to go about constructing a relativistic theory of gravity.

5 Joint Particle-Field Evolution Problems

We conclude this paper by discussing the ill-posedness of popular joint particle-field evolution problems. These joint evolution problems are studied when one wants to consider a dynamical system composed of a point particle acting as a source in some field, while the field acts on the particle via a force law.

5.1 Lagrangian Density Formulation For Relativistic Fields

Similarly to how we can view world-lines satisfying a force law as being critical points of an actional functional, the same can be done for fields which evolve according to certain PDEs.

Definition 5.1. *Given a subset Ω of $\mathbb{R}^{1,3}$, and boundary condition g on $\partial\Omega$, define the action functional $\mathcal{A} : W \to \mathbb{R}$ via*

$$\mathcal{A}(\phi) := \int_\Omega \mathcal{L}(\phi, \partial_\mu \phi)(x)\sqrt{-\eta}dx^4 \tag{46}$$

where W_ϕ is defined as

$$W_\phi := \{\phi \in W^{1,1}(\Omega) | \phi = g \text{ on } \partial\Omega\} \tag{47}$$

and $\mathcal{L}(\Phi, \Pi_\mu) : \mathbb{R} \times \mathbb{R}^{1,3} \to \mathbb{R}$ is a given function referred to as the Lagrangian density.

Theorem 5.1. *Critical points ϕ of the action functional $\mathcal{A}$ must satisfy the partial differential equation given by the Euler-Lagrange equations*

$$\partial_\mu\left(\frac{\partial \mathcal{L}}{\partial \Pi_\mu}(\phi, \partial_\mu \phi)\right) = \frac{\partial \mathcal{L}}{\partial \Phi}(\phi, \partial_\mu \phi). \tag{48}$$

Proof. Identical to the proof of Theorem (3.1). □

Proposition 5.1. *The Lagrangian density associated with a given scalar charge current $J^\mu = (\rho, \vec{J})$ acting as a source for a scalar field ϕ is*

$$\mathcal{L}_\phi(\Phi, \Pi_\mu) := \frac{1}{2}(\Pi^\mu)(\Pi_\mu) + \Phi\sqrt{J^\mu J_\mu}. \tag{49}$$

Proof. Calculating each side of the Euler-Lagrange equations returns

$$\partial_\mu\left(\frac{\partial \mathcal{L}_\phi}{\partial \Pi_\mu}(\phi, \partial_\mu \phi)\right) = \partial^\mu \partial_\mu \phi, \quad \frac{\partial \mathcal{L}_\phi}{\partial \Phi}(\phi, \partial_\mu \phi) = \sqrt{J^\mu J_\mu}. \tag{50}$$

So ϕ satisfies the wave equation with a source term

$$\partial^\mu \partial_\mu \phi = \sqrt{J^\mu J_\mu}. \tag{51}$$

□

5.2 Coupling Particles and Fields

At first glance, the Lagrangian formulation for particles and fields do not seem compatible with each other. For particles, we examine a Lagrangian L which is integrated over a parameterization parameter θ, while the field Lagrangian density is integrated over a volume form. Amazingly, we can transform the particle Lagrangian L into a Lagrangian density using singular delta distributions.

Definition 5.2. *Define the particle's Lagrangian density $\mathcal{L}_p$ via*

$$\mathcal{L}_p(x) := \frac{1}{\sqrt{\eta}} \int_{\theta_1}^{\theta_2} -m\delta^{(4)}(x-z)\sqrt{\eta_{\mu\nu}\dot{z}^\mu \dot{z}^\nu} d\theta \tag{52}$$

where $\delta^{(4)}(x-z) = \Pi_{i=1}^4 \delta^\mu(x^\mu - z^\mu)$ is the four-dimensional delta distribution.

Proposition 5.2. *The action functional associated with the particle's Lagrangian density is*

$$A(z(\cdot)) = \int_\Omega \mathcal{L}_p(x)\sqrt{-\eta} dx^4 = \int_{\theta_1}^{\theta_2} -m\sqrt{\eta_{\mu\nu}\dot{z}^\mu \dot{z}^\nu} d\theta \tag{53}$$

where we've used that $z(\cdot)$ is wholly contained in Ω. In particular, we brought the Ω integral inside the θ integral, and integrated out the delta function.

With this in hand, we can finally write down an action principle which fully describes a coupled particle-field system.

Definition 5.3. *Define the particle-field action functional $\mathcal{A} : \mathcal{W} \to \mathbb{R}$*

$$\mathcal{A}(z(\cdot), \phi) := \int_\Omega \mathcal{L}\sqrt{-\eta} dx^4 := \int_\Omega \mathcal{L}_p + \mathcal{L}_{int} + \mathcal{L}_\phi \sqrt{-\eta} dx^4 \tag{54}$$

where

$$\mathcal{W} := W_p \times W_\phi \tag{55}$$

and $\mathcal{L}_{int}$ is an interaction term coupling the particle to the field of general form

$$\mathcal{L}_{int} = \frac{1}{\sqrt{-\eta}} \int_{\theta_1}^{\theta_2} \delta^{(4)}(x-z) L_{int}(z, \dot{z}, \phi, \partial_\mu \phi) d\theta \tag{56}$$

Theorem 5.2. *Critical points of $\mathcal{A}$ is a pair $(z(\cdot), \phi)$ which satisfy their respective Euler-Lagrange equations*

Proof. A particle-field pair $(z(\cdot), \phi)$ is a critical point if they jointly satisfy

$$\frac{d}{d\epsilon} A(z(\cdot) + \epsilon w(\cdot), \phi)\big|_{\epsilon=0} = 0, \quad \frac{d}{dt} A(z, \phi + t\varphi)\big|_{t=0} = 0 \tag{57}$$

for all $w(\cdot) \in W^{1,1}$ satisfying $w(\theta_1) = 0 = w(\theta_2)$, and all $\varphi \in W^{1,1}(\Omega)$ satisfying $\varphi = 0$ on $\partial\Omega$. To evaluate the first derivative, we notice that we can re-write the action as

$$A(z(\cdot), \phi) = \int_{\theta_1}^{\theta_2} L_p + L_{\text{int}} d\theta + G(\phi). \tag{58}$$

We did this because the $G(\phi)$ term will drop out when we perform the differentiation with respect to ϵ. Thus, the critical point $z(\cdot)$ will satisfy the Euler-Lagrange equations with $L = L_p + L_{\text{int}}$. Similar arguments will show that ϕ satisfies the Euler-Lagrange equations with $\mathcal{L} = \mathcal{L}_\phi + \mathcal{L}_{\text{int}}$. $\square$

5.3 Joint Evolution Equations for Scalar Particle

The Lagrangian interaction for a particle with charge a coupled to a scalar field is given by

$$L_{\text{int}} = a\phi\sqrt{\eta_{\mu\nu}\dot{z}^\mu\dot{z}^\nu}. \tag{59}$$

The associated Lagrangian density is

$$\mathcal{L}_{\text{int}} = \frac{1}{\sqrt{-\eta}}\int_{\theta_1}^{\theta_2} \delta^{(4)}(x-z)a\phi(x)\sqrt{\eta_{\mu\nu}\dot{z}^\mu\dot{z}^\nu}d\theta. \tag{60}$$

The Euler-Lagrange equations for the particle's world-line returns

$$\frac{d}{d\theta}((m-a\phi(z))\frac{\dot{z}_\mu}{\sqrt{\eta_{\mu\nu}\dot{z}^\mu\dot{z}^\nu}}) = \frac{d}{d\tau}((m-\phi(z))u_\mu) = -a\partial_\mu\phi \tag{61}$$

exactly like before. However, now the evolution of ϕ is coupled to $z(\cdot)$ through its Euler Lagrange equations

$$\partial^\mu\partial_\mu\phi = \frac{1}{\sqrt{-\eta}}\int_{\theta_1}^{\theta_2} a\delta^{(4)}(x-z(\theta))\sqrt{\eta_{\mu\nu}\dot{z}^\mu\dot{z}^\nu}d\theta = \int_{\tau_1}^{\tau_2} a\delta^{(4)}(x-z(\tau))d\tau. \tag{62}$$

Lemma 5.1. *If $g(\tau)$ has only one root at τ_0, then $\delta(g(\tau)) = \frac{1}{g'(\tau_0)}\delta(\tau-\tau_0)$.*

For time-like world-lines, $x^0 - z^0(\tau)$ has a root τ_0 for each x^0. So $\delta(x^0 - z^0(\tau)) = \frac{1}{u^0(\tau_0)}\delta(\tau-\tau_0)$, and we can rewrite the evolution equation for ϕ as

$$\partial^\mu\partial_\mu\phi = \frac{1}{u^0}\delta(\vec{x}-\vec{z}). \tag{63}$$

We see that the point particle acts a singularity in the second derivatives of the field, while the field acts on the particle through a force law.

Theorem 5.3. *The action functional $\mathcal{A} : \mathcal{W} \to \mathbb{R}$ which represents a particle interacting with a scalar field has no critical points in its domain.*

Proof. Suppose for sake of contradiction that there are critical points of this action functional. Then $z(\cdot)$ would satisfy the force law (61) while ϕ satisfies the partial differential equation given by (63). But the delta singularity in the second derivatives of ϕ necessarily implies that the first derivatives of ϕ will not be defined along the path of the particle, and the force law becomes ill-defined. We come to the conclusion that these equations cannot give rise to a joint evolution, and thus the action functional has no critical points. $\square$

Remark 5.1. *The action functional for a electromagnetic vector potential interacting with a point particle fails to have critical points for the same reasoning.*