Notes on Special Relativity, Curved Space, and Geodesics

Mathew Calkins

Center for Cosmology and Particle Physics,
New York University

`mc8558@nyu.edu.`

January 19, 2026

Abstract

Some supplementary discussion of geodesics, plus reviews of special relativity and curved spacetime

Contents

1 Review of Special Relativity

1.1 Lorentz and Poincaré Transformations

We first learned Newtonian mechanics in the background of Galilean relativity. Here every two inertial references frames are equally valid and mutually indistinguishable, and the transformations that relate two such frames are arbitrary compositions of rotations, Galilean boosts, and translations.

Rotations are implemented by matrices. For example, if an observer in frame S' is rotated by an angle θ relative to an observer in frame S, their coordinates are related as

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

or if we include all of our spacetime coordinates

$$\begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix}.$$

Galilean boosts are also implemented by matrices. If the primed observer is moving relative to the unprimed observer with velocity $\vec{v} = v\hat{x}$, their coordinates are related as

$$\begin{bmatrix} t' \\ x' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -v & 1 \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix}.$$

which can similarly be padded to give a full map $(t, x, y, z) \mapsto (t', x', y', z')$. Finally, if the primed observer takes their origin displaced by a constant vector $\vec{a}$ and starts their stopwatch at a time a_t (both relative to the unprimed observer), then the primed and unprimed coordinates are related as

$$t' = t - a_t,$$
$$\vec{x}' = \vec{x} - \vec{a}.$$

The transformation laws for special relativity work exactly the same way, except that Galilean boosts are replaced by Lorentz boosts. For the same example given above, we have

$$\begin{bmatrix} t' \\ x' \end{bmatrix} = \gamma_v \begin{bmatrix} 1 & -v \\ -v & 1 \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix} \quad \text{where } \gamma_v := \frac{1}{\sqrt{1-v^2}}.$$

Exercise Restore explicit factors of c and show that, in the limit $v/c \to 0$ at v fixed, Lorentz boosts approach Galilean boosts.

An arbitrary transformation built from rotations, Lorentz boosts, and translations is called a *Poincaré transformation*. If we exclude translations (if only because they're less interesting) the result is called a *Lorentz transformation*.

Since rotations and boosts act on the coordinates by matrix multiplication and translations act by adding a constant vector, we can write an arbitrary Poincaré transformation as

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\nu$$

2

as long as Λ corresponds to something built from rotations and Lorentz boosts. Then restricting to Lorentz transformations is as simple as taking $a = 0$.

1.2   The Minkowski Metric and Spacetime

We can derive all this by one of two equivalent routes. One is to take Einstein's two classical axioms

1. Any two inertial frames are equally valid.

2. Any two inertial frames agree on the speed of light.

and perform some choice thought experiments involving observers with flashbulbs on and off trains and so on. More abstractly, we could start by declaring that spacetime has a natural background metric η which, in every inertial frame, has constant components

$$
\eta_{\mu\nu} = \begin{cases} -1 & \mu = \nu = 0, \\ 1 & \mu = \nu \neq 0, \\ 0 & \text{otherwise.} \end{cases}
$$
$$
= \operatorname{diag}(-1, +1, +1, +1).
$$

We can use this object to compute the inner product between arbitrary four-vectors, written variously as

$$
A \cdot B = \eta(A, B) = A^\mu \eta_{\mu\nu} B^\nu = A^0 B^0 - \vec{A} \cdot \vec{B}.
$$

We can in particular use it to measure the squared distance between any two events. If events A, B have coordinates x_A, x_B and relative locations $\Delta x := x_B - x_A$, then the *invariant interval* between the two events is by definition

$$
\Delta s^2 := \Delta x^\mu \eta_{\mu\nu} \Delta x^\nu = -(\Delta t)^2 + \Delta \vec{x} \cdot \Delta \vec{x}.
$$

If you sat down and found all of the coordinate transformations which always preserve this invariant interval between events (and which can be continuously built up from infinitesimal transformations, unlike time-reversal or certain reflections), you would find exactly the Lorentz transformations plus translations. In fact, in terms of this metric we can write down another even more abstract characterization of a Lorentz transformation, namely a coordinate transformation $x \mapsto x' = \Lambda x$ where Λ satisfies

$$
\Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu \eta_{\rho\sigma} = \eta_{\mu\nu}.
$$

Exercise Show that $x' \mapsto \Lambda x$ preserves the invariant interval between arbitrary events exactly when this latest condition is satisfied.

It is useful to classify a general four-vector based on the sign of its squared length, written variously as $A^2 = \|A\|^2 = A^\mu \eta_{\mu\nu} A^\nu$. We say that

$$
A \text{ is } \begin{cases} \text{timelike} & \text{if } \|A\|^2 < 0, \\ \text{spacelike} & \text{if } \|A\|^2 > 0, \\ \text{null/lightlike} & \text{if } \|A\|^2 = 0. \end{cases}
$$

Applying this to two events as above, we say that

$$x_A \text{ and } x_B \text{ are } \begin{cases} \text{timelike} & \text{if } \Delta s^2 < 0, \\ \text{spacelike} & \text{if } \Delta s^2 > 0, \\ \text{null/lightlike} & \text{if } \Delta s^2 = 0. \end{cases}$$
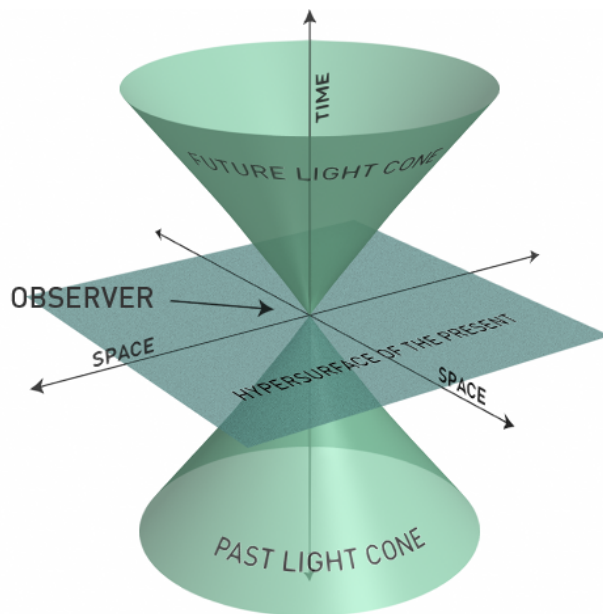

Figure 1: A lightcone in a (2+1)-dimensional spacetime.

This turns to out essentially classify the possible causal relationships between the two events. It is helpful to consider the diagram in Fig. 1., where for simplicity we place x_A at the origin and consider all of the different regions where x_B could be found. Some essential features of this diagram are:

- If the points are spacelike separated, by invariance of the invariant interval they will remains spacelike separated under a change of inertial reference frame.

  - In this case and only in this case, we can perform a Lorentz boost to arbitrarily place event B in the future, past, or exact present of event A. That is, it is exactly in this case that simultaneity is relative.

  - In this case, the two events are too far spatially separated relative to their temporal separation for a signal to propagate from A to B. Physically, this means there can be no causal influence between the events. This is why the ambiguity in their temporal ordering gives no paradox.

- If the points are timelike separated, they will remain timelike separated under a change of inertial reference frame.

  - In this case x_B must geometrically be inside either the future lightcone or the past lightcone of x_A. No Lorentz boost can take x_B from the future interior lightcone to the past interior lightcone nor vice versa.

4

– Physically, these events are close enough in space relative to their temporal displacement that a signal can travel from one to the other. This is why it **must** be impossible to change their temporal ordering by a Lorentz transformation. We would run into paradoxes if two events that can influence one another had amibiguous ordering.

- If the points are lightlike separated, they will will remain lightlike separated under a change of inertial reference frame.

  – Here we say that x_B lies exactly on the lightcone (either past or future) of x_A, not in its interior.

  – This is exactly on the threshold of causal influence. If the events were any further apart in space or closer in time, they would be unable to influence each other. Because they are in causal contact, just like in the timelike case we must have unambiguous causal ordering, *i.e.*, all observers must agree on the sign of $t_B - t_A$.

1.3   Relativistic Kinematics

To lift Newtonian mechanics up to a relativistic theory, we must take its basic characters (trajectories, forces, momenta, and so on) and lift them up to nice covariant objects. The first step to doing so is to stop thinking of a trajectory as a position $\vec{x}$ parametrized by time t, and start thinking of it as a path through spacetime where both $\vec{x}, \vec{t}$ are functions of some other parameter.

$$(t, \vec{x}) = (t(\lambda), \vec{x}(\lambda)) = (x^\mu(\lambda)).$$

The path through spacetime that a particle traces out over its history is called its *worldline*. With respect to this parametrization our particle has velocity

$$V = \frac{dx^\mu}{d\lambda}.$$

But measuring the change in spacetime position per change in λ value gives a notion of velocity which depends on how we parametrize, in much the same way that our Newtonian velocity will take a different numerical value if we measure the change in location per minute or per second. To make things unambiguous we can parametrize by $\lambda = \tau$, the proper time along the particle's worldline. This gives the *four-velocity*

$$U^\mu := \frac{dx^\mu}{d\tau}.$$

There is still the much smaller ambiguity of what point along the trajectory we take to be at $\tau = 0$ (physically, when the particle's internal clock is set to midnight), but this doesn't affect the definition of $d\tau$.

We know that dt and $d\tau$ are related as

$$d\tau = \sqrt{1 - v^2}dt \implies dt = \gamma d\tau.$$

which is exactly the statement of time dilation. Plugging this into our definition of four-velocity gives equivalently

$$(U^\mu) = \left(\frac{dt}{d\tau}, \frac{dx}{d\tau}, \frac{dy}{d\tau}, \frac{dz}{d\tau} \right) = \left(\gamma, \gamma\frac{dx}{dt}, \gamma\frac{dy}{dt}, \gamma\frac{dz}{dt} \right) = \left(\gamma, \gamma\dot{\vec{x}} \right)$$

where we use Newton's flyspeck notation to denote total derivatives with respect to coordinate time. We also know that the squared norm of a four-vector must give a Lorentz-invariant scalar. In this case that scalar comes out to

$$
\begin{aligned}
U^\mu U^\nu \eta_{\mu\nu} &= -(U^0)^2 + \vec{U} \cdot \vec{U} \\
&= -\gamma^2 \left(1 + \dot{\vec{x}} \cdot \dot{\vec{x}}\right) \\
&= -1
\end{aligned}
$$

where in the last step we apply the definition of γ. This property that $U \cdot U = -1$ is another characterization of proper time.

Differentiating again with respect to proper time gives the four-acceleration

$$
A^\mu := \frac{dU^\mu}{d\tau} = \frac{d^2 x^\mu}{d\tau^2}
$$

and so on[1]. We can similarly lift a massive particle's linear three-momentum $\vec{p} = m\dot{\vec{x}}$ up to a four-vector by defining

$$
(p^\mu) := (mU^\mu) = \left(\gamma m, \gamma m \dot{\vec{x}}\right)
$$

from which we identify $p^0 = \gamma m \equiv E$ as the particle's intrinsic mechanical energy. In every frame we have

$$
p^\mu = mU^\mu \implies p \cdot p = m^2 U \cdot U = -m^2
$$

so in the particle's rest frame $\dot{\vec{x}} = 0$ we have in particular

$$
-m^2 = -E^2
$$

or restoring explicit factors of c
$$
E = mc^2.
$$

If our particle is subject to a four-force with components F^μ, then Newton's second law takes the covariant form
$$
F^\mu = \frac{d}{d\tau} p^\mu.
$$

If our particle is influenced only by its interaction with some conservative potential Φ (which is not always the case, as you saw in your homework problems involving the vector potential $\vec{A}$) then we will have

$$
\frac{d}{d\tau} p^\mu = -\frac{\partial \Phi}{\partial x^\mu} =: -\partial_\mu \Phi.
$$

---

[1]One could in principle continue to define four-jerk, four-jounce, and so on but these won't be used in this course.

1.4 Lagrangians

Rather than using Newton-style framing, we can formulate particle dynamics in special relativity (as well as in Galilean or even general relativity, as we'll soon see) in terms of Lagrangians and action principles. To do so, we use the worldline picture and write down a Lagrangian

$$L = L\left(x(\lambda), \frac{dx(\lambda)}{d\lambda}, \lambda\right)$$

then use this to define an action, which is a machine that takes in a worldline between to fixed values λ_i, λ_f and returns a number.

$$S[x] = \int_{\lambda_i}^{\lambda_f} L\left(x(\lambda), \frac{dx(\lambda)}{d\lambda}, \lambda\right).$$

This encodes dynamics in the sense that a particle's equations of motion are equivalent to the condition that the trajectory $x(\lambda)$ must extremize the action, within the space of all trajectories with fixed initial and final points $x(\lambda_i), x(\lambda_f)$.

$$x \text{ satisfies equations of motion} \iff \delta S[x] = 0.$$

Practically, we extract the equations of motion by noting that this extremization condition is equivalent to the Euler-Lagrange equations

$$\delta S[x] = 0 \iff \frac{\partial L}{\partial x^\mu} = \frac{d}{d\lambda}\frac{\partial L}{\partial \left(\frac{dx^\mu}{d\lambda}\right)}.$$

This also gives us a natural language for talking about symmetries. Any way of changing our trajectory (say, by performing an active rotation) is said to be a *symmetry* if it always takes solutions of the equations of motions to other solutions. Phenomenologically, this means an observer will see the particle obeying the same laws before and after the transformation.

In terms of action principles, this means taking extrema of S to other extrema of S. Generally this is accomplished by cooking up a transformation which preserves the variation of S.

$$\delta S[x] = \delta S[x'].$$

At the Lagrangian level, for this to happen it is sufficient if the Lagrangian changes by at most a total λ-derivative

$$L\left(x'(\lambda), \frac{dx'(\lambda)}{d\lambda}, \lambda\right) = L\left(x(\lambda), \frac{dx(\lambda)}{d\lambda}, \lambda\right) + \frac{d}{d\lambda}F\left(x(\lambda), \frac{dx(\lambda)}{d\lambda}, \lambda\right)$$

to leading order in infinitesimal $x' - x$. Then the corresponding actions will differ by the constant $F(\lambda_f) - F(\lambda_i)$, which doesn't affect the variation since these by definition leave the endpoints fixed.

For a free particle in special relativity, the action is exactly the invariant length (*i.e.*, the proper time) of the relevant segment of the particle's worldline.

$$S_{\text{free}} = \int d\tau = \int \sqrt{1 - \dot{\vec{x}} \cdot \dot{\vec{x}}}dt$$

though in the presence of outside forces the action will acquire more terms.

1.5 Transformation Laws

Any vector V has components V^μ which transform under a change of reference frame as

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

We could find this by remembering that, without indices, the transformation law must look like matrix multiplication

$$A' = \Lambda A$$

then adding in indices in the unique way allowed by our summation conventions. But remember that the vector itself doesn't change. What's happening is that our change of reference frame brings a change of basis, hence a change in the coefficients when we expand in that basis. But the vector itself is unchanged. If we write the basis vectors before and after the change as $(\mathbf{e}_\mu)$ and $(\mathbf{e}'_\mu)$, then we have

$$A = A^\mu \mathbf{e}_\mu = A'^\mu \mathbf{e}'_\mu$$

which in turn implies that the new basis vectors must be related to the old ones as

$$\mathbf{e}'_\mu = \Lambda^\nu{}_\mu \mathbf{e}_\nu.$$

Covectors, at the present level of abstraction, are objects B whose components B_μ carry a single lowered index, such as that gradient of a scalar field

$$B_\mu = \frac{\partial \phi}{\partial x^\mu}.$$

These transform as

$$B'_\mu = \Lambda^\nu{}_\mu B_\nu.$$

There are more exotic objects with multiple indices, such as the Maxwell field strength

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

This transforms like a four-vector except now we have two Λ factors, one for each raised index.

$$F'^{\mu\nu} = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma F^{\rho\sigma}.$$

2 Review of Curved Spacetime

In classical electromagnetism we learned two qualitatively different sets of laws. On one hand, background electromagnetic fields accelerate charged particles according to the Lorentz force law. Conversely, charged particles source the electromagnetic fields according to the Maxwell equations, though the fields have their own nontrivial dynamics even in the absence of sources.

Something analogous happens in general relativity, which is Einstein's successful unification of gravity with special relativity. Here we have one set of laws called the *Einstein*

field equations describing how matter and energy source the gravitational field, while a certain action principle will turn out to describe how particles move about in a given gravitation background.

For now we will only concern ourselves with this latter principle. That is, we will study how gravitational effects arise as manifestations of the nontrivial geometry of spacetime, and for now we won't worry about how exactly spacetime acquired that geometry. In short, we will study particle mechanics in the background of some fixed spacetime metric.

2.1 Spacetime Geometry

We can most easily describe general relativity by saying how it departs from special relativity, since we are now experts on the latter. Speaking loosely and intuitively, the principle departure is that spacetime is now curved, more like a sphere or some wobbly higher-dimensional surface than like Euclidean or Minkowski space. More formally, this means that in place of our constant metric η we have some background metric g whose components in some coordinates x measure infinitesimal distances near a point p as

$$ds^2\big|_p = g_{\mu\nu}(x(p)) \, dx^\nu\big|_p \, dx^\mu\big|_p \, .$$

Besides measuring distances this gives us an inner product between any two vectors A_p, B_p based at the point p. We can write this variously as

$$A_p \cdot B_p = ds^2(A_p, B_p) = g(A_p, B_p) = g_{\mu\nu}A_p^\mu B_p^\nu.$$

As before, this quantity is a scalar, taking the same value no matter what coordinates x we chose. In terms of the coordinate basis $(\mathbf{e}_\mu(x))$ for vectors based at p, we can also write the metric as

$$g_{\mu\nu}(x(p)) = g\left(\mathbf{e}_\mu(x), \mathbf{e}_\nu(x)\right) = \mathbf{e}_\mu(x) \cdot \mathbf{e}_\nu(x).$$

Moving forward we will often be too lazy to explicitly write out the point at which everything is being measured.

One measure of the departure of our general geometry from that of Minkowski space is that, not only do we have $g_{\mu\nu} \neq \eta_{\mu\nu}$ in general, we cannot even pick a coordinate system that makes this true everywhere simultaneously. The best we can do is to pick *local inertial coordinates* around an event p_0. By definition these are coordinates in which $g_{\mu\nu} \neq \eta_{\mu\nu}$ exactly at p_0, with all first derivatives vanishing so that corrections to this equality as we move away from p_0 arise only at second order. In terms of the Taylor series, we can equivalently say that there always exists some set of coordinates ξ such that

$$g_{\mu\nu}(\xi(p)) = \underbrace{g_{\mu\nu}(\xi(p_0))}_{=\eta_{\mu\nu}} + (\xi^\rho(p) - \xi^\rho(p_0)) \underbrace{\frac{\partial g_{\mu\nu}}{\partial \xi^\rho}(\xi(p))}_{=0} + (\text{quadratic or higher in } \xi(p) - \xi(p_0))$$

Conversely, if you find a set of coordinates ξ such that $g_{\mu\nu}(\xi) = \eta_{\mu\nu}$ everywhere, you have discovered that your spacetime is indeed flat, even if it was originally written in a coordinate system that obscured this fact.

2.2 Transformation Laws

As in special relativity, we can classify the various kinds of geometric objects we encounter by the ways that they transform if we move from one coordinate system x to another $\tilde{x}$. For example, a vector's components in the two coordinate systems are related as

$$\tilde{A}^\mu = \frac{\partial \tilde{x}^\mu}{\partial x^\nu} A^\nu$$

so that invariance

$$A = A^\mu \mathbf{e} = \tilde{A}^\mu \tilde{\mathbf{e}}_\mu$$

(here we don't bother writing the explicit point at which everything is being evaluated) then uniquely fixes the transformation laws for the vectors in the coordinate basis to be

$$\tilde{\mathbf{e}}_\mu = \frac{\partial x^\nu}{\partial \tilde{x}^\mu} \mathbf{e}_\nu.$$

A covector, as you might guess from special relativity, transforms as

$$\tilde{B}_\mu = \frac{\partial x^\nu}{\partial \tilde{x}^\mu} B_\nu$$

so that if we take a vector and a covector together and sum over their joint index, we get a scalar.

$$A^\mu B_\mu = \tilde{A}^\mu \tilde{B}_\mu = \text{same value for all coordinate systems.}$$

Objects with even more indices just get more Jacobian factors. For example, the metric transforms as

$$\tilde{g}_{\mu\nu} = \frac{\partial x^\rho}{\partial \tilde{x}^\mu} \frac{\partial x^\sigma}{\partial \tilde{x}^\nu} g_{\rho\sigma}.$$

This holds for an arbitrary coordinate transformation $x \to \tilde{x}$, so we can in particular let our first coordinate system be the locally inertial coordinates ξ and write x for the new coordinates this law. Then this becomes

$$g_{\mu\nu} = \frac{\partial \xi^\rho}{\partial x^\mu} \frac{\partial \xi^\sigma}{\partial x^\nu} \eta_{\rho\sigma}.$$

Conversely, the Maxwell stress tensor transforms as

$$\tilde{F}^{\mu\nu} = \frac{\partial \tilde{x}^\mu}{\partial x^\rho} \frac{\partial \tilde{x}^\nu}{\partial x^\sigma} F^{\rho\sigma}.$$

The general rule is that each raised index brings with it a Jacobian factor $\partial \tilde{x}/\partial x$ while each lowered index brings an inverse Jacobian factor $\partial x/\partial \tilde{x}$.

3 Actions and Geodesics

3.1 Worldlines Extremize Proper Time

Just as in special relativity, we can describe particle dynamics in curved spacetimes via an action principle. Even better, in the absence of forces other than gravity (which should be thought of as in some sense always present, since it manifests through the

background geometry) the action of is still the invariant length of the segment of the worldline that we are considering.

$$S = \int d\tau = \int \sqrt{-ds^2}.$$

If we work in coordinates x, parametrize as $x = x(\lambda)$, and keep all arguments totally explicit, we can write this as

$$S = \int \sqrt{-g_{\mu\nu}dx^\mu dx^\nu} = \int_{\lambda_i}^{\lambda_f} \sqrt{-g_{\mu\nu}(x(\lambda))\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}}d\lambda$$

from which we can identify our Lagrangian as

$$L\left(x, \frac{dx}{d\lambda}\right) = \sqrt{-g_{\mu\nu}(x(\lambda))\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}}.$$

In a general geometry, a path from point A to point B which extremizes the invariant length is called a *geodesic*. In Minkowski or Euclidean space, we can run through Euler-Lagrange directly to confirm our intuition that geodesics in that edge case are exactly straight lines.

Exercise Do this calculation, and explain why the result describes a straight line.

So conceptually this isn't so different from the action principle in special relativity. The new practical complication is that now we are in a general metric, so extremizing this functional is mechanics more difficult.

In the case of a background Minkowski metric

$$S_{\text{Minkowski}} = \int \sqrt{-\eta_{\mu\nu}dx^\mu dx^\nu} = \int \sqrt{dt^2 - d\vec{x} \cdot d\vec{x}}$$

we can differentiate directly, or else simply guess from the covariant form of Newton's second law, to find that that the equations come out to be

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

In a general background metric things are just a bit more complicated. As it turns out, if we work out the equations of motion

$$\frac{\partial L}{\partial x^\mu} = \frac{d}{d\lambda}\frac{\partial L}{\partial\left(\frac{dx^\mu}{d\lambda}\right)}$$

and then at the last step rephrase everything in terms of proper time, we will find that the geodesic equation of motion take the form

$$\frac{d^2 x^\mu}{d\tau^2} = -\Gamma^\mu_{\nu\rho}\frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau}.$$

where the position-dependent coefficients $\Gamma^\mu_{\nu\rho}(x) = \Gamma^\mu_{\rho\nu}(x)$ are called the *Christofell symbols* and characterize the background geometry. To convince yourself of this, it's best to work through an example. This will also serve as a practical guide to the kinds of techniques you might use to find geodesics in a given background.

3.1.1 Subtleties on Parametrization

Before we start calculation we make some careful notes on parametrization. The form we just wrote is that taken if we parametrize by proper time (that is, by the path's invariant length), but we can relax that condition slightly and still get an equation of motion of the same form. If we instead use a parameter ω which is τ up to a constant rescaling $a \neq 0$ and a constant shift b

$$\omega = a\tau + b \implies d\omega = a d\tau$$

then the factors of a cancel on each side and we get a structurally identical geodesic equation.

$$\frac{d^2 x^\mu}{d\omega^2} = -\Gamma^\mu_{\nu\rho} \frac{dx^\nu}{d\omega} \frac{dx^\rho}{d\omega}.$$

What can we say about such a parameter ω for general a, b? If τ parametrizes by invariant length, then ω is a general parametrization whose increments are proportionate to invariant length. If we imagine placing tick marks along a trajectory at fixed values $\omega = 0, 1, 2, 3$ and so on, then as measured by our metric these tick marks would be at equally spaced points along the trajectory. Such a parametrization is called an *affine parametrization*[2]

This is not true of a general parametrization, and indeed a general parametrization will give a messier equation of motion upon varying our action. We conventionally reserve the term *geodesic equation* for the neat equation of motion above that we get in an affine parametrization.

3.2 The Wormhole Geometry: Naive Lagrangian

The *wormhole geometry* is the spacetime with metric

$$ds^2 = -dt^2 + dr^2 + (b^2 + r^2)(d\theta^2 + \sin^2\theta d\phi^2)$$

for b some real constant. In this geometry our action principle is governed by the Lagrangian

$$L\left(x, \frac{dx}{d\lambda}\right) = \sqrt{-g_{\mu\nu}(x(\lambda)) \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}}$$

$$= \sqrt{\left(\frac{dt}{d\lambda}\right)^2 - \left(\frac{dr}{d\lambda}\right)^2 - (b^2 + r^2)\left[\left(\frac{d\theta}{d\lambda}\right)^2 + \sin^2\theta\left(\frac{d\phi}{d\lambda}\right)^2\right]}$$

$$\equiv \left[-\frac{dx}{d\lambda} \cdot \frac{dx}{d\lambda}\right]^{1/2}.$$

where as always $\cdot$ denotes the inner product between vectors as computed by our metric g. As already noted, the equations of motion are

$$\frac{\partial L}{\partial x^\mu} = \frac{d}{d\lambda} \frac{\partial L}{\partial \left(\frac{dx^\mu}{d\lambda}\right)}$$

---

[2]Like many other ideas we have come across, this also have a more abstract definition, but for now this will suffice.

To illustrate our ideas, it will be enough to work out in detail the equations of motion for just one component of (x^μ). We consider in particular the equation of motion for r.

$$\frac{\partial L}{\partial r} = \frac{d}{d\lambda} \frac{\partial L}{\partial \left(\frac{dr}{d\lambda}\right)}.$$

To massage this into a useful form, we only partially work out the partial derivative on the LHS, rewriting it as a partial derivative of the quantity sitting under the square root. Explicitly, we have

$$\frac{\partial L}{\partial r} = \frac{\partial}{\partial r} \left(\left[-\frac{dx}{d\lambda} \cdot \frac{dx}{d\lambda} \right]^{1/2} \right)$$

$$= \frac{1}{2} \left[-\frac{dx}{d\lambda} \cdot \frac{dx}{d\lambda} \right]^{-1/2} \frac{\partial}{\partial r} \left[-\frac{dx}{d\lambda} \cdot \frac{dx}{d\lambda} \right]$$

$$= \frac{1}{2L} \frac{\partial}{\partial r} \left[-\frac{dx}{d\lambda} \cdot \frac{dx}{d\lambda} \right].$$

The key step here is that the square root structure of the Lagrangian gives us a $1/L$ term all the way on the outside. So far this is totally general. Now we can finish working out the derivative using the particular metric we're working in.

$$\frac{\partial L}{\partial r} = \frac{1}{2L} \frac{\partial}{\partial r} \left[\left(\frac{dt}{d\lambda}\right)^2 - \left(\frac{dr}{d\lambda}\right)^2 - (b^2 + r^2) \left[\left(\frac{d\theta}{d\lambda}\right)^2 + \sin^2\theta \left(\frac{d\phi}{d\lambda}\right)^2 \right] \right]$$

$$= -\frac{r}{L} \left[\left(\frac{d\theta}{d\lambda}\right)^2 + \sin^2\theta \left(\frac{d\phi}{d\lambda}\right)^2 \right].$$

Now we perform a trick, so neat it almost seems like cheating. By definition our action is the particle's proper time, so we have

$$\int d\tau = \int L d\lambda \implies d\tau = L d\lambda.$$

Plugging this in carefully, we have equivalently

$$\frac{\partial L}{\partial r} = -rL \left[\left(\frac{d\theta}{L d\lambda}\right)^2 + \sin^2\theta \left(\frac{d\phi}{L d\lambda}\right)^2 \right]$$

$$= -rL \left[\left(\frac{d\theta}{d\tau}\right)^2 + \sin^2\theta \left(\frac{d\phi}{d\tau}\right)^2 \right]$$

So one side of our Euler-Lagrange equation for r is now in terms of τ rather than λ. The innermost partial derivative on the RHS works similarly, and we can use the same trick.

$$\frac{\partial L}{\partial \left(\frac{dr}{d\lambda}\right)} = \frac{1}{2L} \frac{\partial}{\partial \left(\frac{dr}{d\lambda}\right)} \left[-\frac{dx}{d\lambda} \cdot \frac{dx}{d\lambda} \right]$$

$$= \frac{1}{2L} \frac{\partial}{\partial \left(\frac{dr}{d\lambda}\right)} \left[-\left(\frac{dr}{d\lambda}\right)^2 - (b^2 + r^2) \left[\left(\frac{d\theta}{d\lambda}\right)^2 + \sin^2\theta \left(\frac{d\phi}{d\lambda}\right)^2 \right] \right]$$

$$= -\frac{1}{L} \frac{dr}{d\lambda}$$

$$= -\frac{dr}{d\tau}.$$

Great! We're nearly done. Now our Euler-Lagrange equation for r has simplified to

$$-rL\left[\left(\frac{d\theta}{d\tau}\right)^2 + \sin^2\theta\left(\frac{d\phi}{d\tau}\right)^2\right] = \frac{d}{d\lambda}\left(-\frac{dr}{d\tau}\right)$$

Canceling minus signs, dividing L over, and again applying the relation $Ld\lambda = d\tau$ then gives

$$\frac{d^2r}{d\tau^2} = r\left[\left(\frac{d\theta}{d\tau}\right)^2 + \sin^2\theta\left(\frac{d\phi}{d\tau}\right)^2\right].$$

Now we have our equation of motion for r purely in terms of τ, even though we started with λ.

From this we get automatically some of the components of our Christofell symbols. The geodesic equation we wrote earlier has r component

$$\frac{d^2r}{d\tau^2} = -\Gamma^r_{\nu\rho}\frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau}.$$

By comparison with our equation of motion for r, we can then read off

$$\Gamma^r_{\theta\theta} = -r,$$
$$\Gamma^r_{\phi\phi} = -r\sin^2\theta$$

with all other components $\Gamma^r_{\nu\rho}$ vanishing.

Finding the remaining equations of motion proceeds by identical logic, and gives us the components $\Gamma^\mu_{\nu\rho}$ for $\mu = t, \theta, \phi$. The full results are given in Hartle Examples 8.2 and 8.3.

3.3   The Wormhole Geometry: Hook Lagrangian

We can do the same calculation in a slightly slicker way using the Hook Lagrangian that we have seen in lecture. Recall that this is simply the square of our Lagrangian above.

$$\begin{aligned}
L_{\text{Hook}}\left(x, \frac{dx}{d\lambda}\right) &:= -g_{\mu\nu}(x(\lambda))\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} \\
&= \left(\frac{dt}{d\lambda}\right)^2 - \left(\frac{dr}{d\lambda}\right)^2 - (b^2 + r^2)\left[\left(\frac{d\theta}{d\lambda}\right)^2 + \sin^2\theta\left(\frac{d\phi}{d\lambda}\right)^2\right] \\
&= \frac{dx}{d\lambda}\cdot\frac{dx}{d\lambda}.
\end{aligned}$$

The great utility of the Hook Lagrangian is that its equations of motion are exactly what the naive Lagrangian yields in the special case that the latter is written in terms of an affine parameter. Practically, this means we can write out our Hook Lagrangian in terms of proper term and vary directly, without having to entertain the $Ld\lambda = d\tau$ game we played above. To illustrate this, we again proceed by working through the calculation. In terms of proper time $\lambda = \tau$ our Hook Lagrangian here is

$$L_{\text{Hook}}\left(x, \frac{dx}{d\tau}\right) = \left(\frac{dt}{d\tau}\right)^2 - \left(\frac{dr}{d\tau}\right)^2 - (b^2 + r^2)\left[\left(\frac{d\theta}{d\tau}\right)^2 + \sin^2\theta\left(\frac{d\phi}{d\tau}\right)^2\right]$$

and the Euler-Lagrange equation of motion for r is

$$\frac{\partial L_{\text{Hook}}}{\partial r} = \frac{d}{d\tau} \frac{\partial L_{\text{Hook}}}{\partial \left(\frac{dr}{d\tau}\right)}.$$

The LHS derivative is

$$\frac{\partial L_{\text{Hook}}}{\partial r} = \frac{\partial}{\partial r}\left[\left(\frac{dt}{d\tau}\right)^2 - \left(\frac{dr}{d\tau}\right)^2 - (b^2 + r^2)\left[\left(\frac{d\theta}{d\tau}\right)^2 + \sin^2\theta\left(\frac{d\phi}{d\tau}\right)^2\right]\right]$$

$$= -2r\left[\left(\frac{d\theta}{d\tau}\right)^2 + \sin^2\theta\left(\frac{d\phi}{d\tau}\right)^2\right].$$

The innermost RHS derivative is

$$\frac{\partial L_{\text{Hook}}}{\partial \left(\frac{dr}{d\tau}\right)} = \frac{\partial}{\partial \left(\frac{dr}{d\tau}\right)}\left[\left(\frac{dt}{d\tau}\right)^2 - \left(\frac{dr}{d\tau}\right)^2 - (b^2 + r^2)\left[\left(\frac{d\theta}{d\tau}\right)^2 + \sin^2\theta\left(\frac{d\phi}{d\tau}\right)^2\right]\right]$$

$$= -2\frac{dr}{d\tau}$$

So our equations of motion for r are

$$\frac{d^2 r}{d\tau^2} = r\left[\left(\frac{d\theta}{d\tau}\right)^2 + \sin^2\theta\left(\frac{d\phi}{d\tau}\right)^2\right].$$

3.4 The Wormhole Geometry: The Metric

We learned in lecture that there is another characterization of the Christofell symbols in terms of the metric, namely

$$g_{\mu\sigma}\Gamma^{\sigma}_{\nu\rho} = \frac{1}{2}\left(\partial_\rho g_{\mu\nu} + \partial_\nu g_{\mu\rho} - \partial_\mu g_{\nu\rho}\right).$$

where as before we use the shorthand

$$\partial_\mu := \frac{\partial}{\partial x^\mu}.$$

This uniquely defines Γ since g is invertible as a matrix, but computing Γ this way is often much more tedious than varying the Hook Lagrangian. For completeness we will do the calculation nonetheless. Still working in the wormhole geometry

$$ds^2 = -dt^2 + dr^2 + (b^2 + r^2)(d\theta^2 + \sin^2\theta d\phi^2) = g_{\mu\nu}dx^\mu{}^\nu$$

and ordering our coordinates as $(x^\mu) = (t, r, \theta, \phi)$, we read off our metric components as

$$(g_{\mu\nu}) = \begin{bmatrix} g_{tt} & g_{tr} & g_{t\theta} & g_{t\phi} \\ g_{rt} & g_{rr} & g_{r\theta} & g_{r\phi} \\ g_{\theta t} & g_{\theta r} & g_{\theta\theta} & g_{\theta\phi} \\ g_{\phi t} & g_{\phi r} & g_{\phi\theta} & g_{\phi\phi} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & b^2 + r^2 & 0 \\ 0 & 0 & 0 & (b^2 + r^2)\sin^2\theta \end{bmatrix}.$$

Now we do our index manipulations carefully. First, we take our equation above for the Christofell symbols and set the free index μ to be r. This gives

$$g_{r\sigma}\Gamma^{\sigma}_{\nu\rho} = \frac{1}{2}\left(\partial_\rho g_{r\nu} + \partial_\nu g_{r\rho} - \partial_r g_{\nu\rho}\right).$$

Now the LHS has a sum over the dummy index σ, but we see that the components of g form a diagonal matrix, so $g_{r\sigma}$ is nonzero only at $\sigma = r$ at which point we have $g_{rr} = 1$. So we can do the LHS sum and find

$$\Gamma^r_{\nu\rho} = \frac{1}{2} \left(\partial_\rho g_{r\nu} + \partial_\nu g_{r\rho} - \partial_r g_{\nu\rho} \right).$$

The last step is to work out the RHS for each pair ν, ρ, though with experience one can sometimes spot a symmetry that tells us in advance that some components will vanish. For example, we can set $\nu = \rho = \theta$ to find

$$\begin{aligned}
\Gamma^r_{\theta\theta} &= \frac{1}{2} \left(\partial_\theta g_{r\theta} + \partial_\nu g_{r\theta} - \partial_r g_{\theta\theta} \right) \\
&= \frac{1}{2} \left(\partial_\theta(0) + \partial_\nu(0) - \partial_r(b^2 + r^2) \right) \\
&= -r
\end{aligned}$$

as we found earlier. Repeating this process for all pairs ν, ρ yields

$$\begin{aligned}
\Gamma^r_{\theta\theta} &= -r, \\
\Gamma^r_{\phi\phi} &= -r \sin^2 \theta
\end{aligned}$$

with all other components $\Gamma^r_{\nu\rho}$ vanishing, as we found before. Now that we have all of our $\Gamma^r_{\nu\rho}$ components, we can plug them into our geodesic equation

$$\frac{d^2 x^\mu}{d\tau^2} = -\Gamma^\mu_{\nu\rho} \frac{dx^\nu}{d\tau} \frac{dx^\rho}{d\tau}$$

at $\mu = r$ to find

$$\begin{aligned}
\frac{d^2 r}{d\tau^2} &= -\Gamma^r_{\nu\rho} \frac{dx^\nu}{d\tau} \frac{dx^\rho}{d\tau} \\
&= - \left(\Gamma^r_{\theta\theta} \frac{d\theta}{d\tau} \frac{d\theta}{d\tau} + \Gamma^r_{\phi\phi} \frac{d\phi}{d\tau} \frac{d\phi}{d\tau} + \text{vanishing terms} \right) \\
&= r \left[\left(\frac{d\theta}{d\tau} \right)^2 + \sin^2 \theta \left(\frac{d\phi}{d\tau} \right)^2 \right]
\end{aligned}$$

just as we found before.

The moral here is that we can go either way: we can derive the geodesic equation by variational methods and from that read off the Christofell symbols, or we can compute the Christofell symbols directly and from them build the geodesic equation. In practice, for a geometry simple enough to work out by hand, the simplest method will usually be to vary the Hook Lagrangian.

3.5   Conserved Quantities and Killing Vectors

We have seen before that writing down a physical model in terms of an action principle makes it more natural to talk about symmetries and conserved quantities. We learned about the general machinery of Noether's theorem and so on, but here we will

do something slightly simpler, though in exchange it will be less systematic.

The appropriate notion of a Noether charge in solving the geodesic equation is a *Killing vector*. Loosely speaking[3] a Killing vector is a vector field ξ such that an active transformation along the direction ξ doesn't change the metric. Of course at this point we don't know systematically what it means to do an active transformation on a general tensor, so for now we restrict ourselves to the easiest-to-spot flavor of killing vector.

Here is the practical rule: if someone hands you a metric, and the metric doesn't explicitly depend on some coordinate, then the corresponding coordinate basis vector is a Killing vector. For example, in the wormhole geometry we see that nothing actually depends on t, so the vector

$$(\xi^\mu) = (\xi^t, \xi^r, \xi^\theta, \xi^\phi) = (1, 0, 0, 0)$$

is a Killing vector. The same reasoning applies to the coordinate ϕ. As we noted already, this isn't quite systematic enough, and not every Killing vector will take this form. But even with its limits this method is useful enough that we note it now.

But why, in turn, is it useful to spot a Killing vector? Just as we found earlier in slightly different language, Noether tells us that continuous symmetries give rise to conservation laws. In the language of geodesics, this manifests itself in the fact that, given any Killing vector ξ, the quantity

$$\xi \cdot u = \xi^\mu g_{\mu\nu} \frac{dx^\nu}{d\tau}$$

is constant along any geodesic. For the Killing vector $(\xi^\mu) = (1, 0, 0, 0)$ in the wormhole geometry for example, we proceed carefully to find

$$
\begin{aligned}
\xi \cdot \mu &= \xi^\mu g_{\mu\nu} \frac{dx^\nu}{d\tau} \\
&= \xi^t g_{t\nu} \frac{dx^\nu}{d\tau} \text{ (because } \xi \text{ has only one nonzero component)} \\
&= g_{t\nu} \frac{dx^\nu}{d\tau} \\
&= g_{tt} \frac{dt}{d\tau} \text{ (because } g_{\mu\nu} \text{ is diagonal)} \\
&= \frac{dt}{d\tau}.
\end{aligned}
$$

So any solution of the geodesic equation in the background of the wormhole geometry obeys the constraint

$$\frac{dt}{d\tau} = a$$

for a a constant. We can then integrate to find

$$t(\tau) = a\tau + b$$

---

[3]yet another thing we'll have to make more precise later

for b a constant of integration. And just like that, we have solved for $t(\tau)$. For a messier component like $r(\tau)$ it may take a bit more work and explicit reference to the geodesic equation, but the general principle stands: This is incredibly useful if you attempt to actually solve the geodesic equation, for exactly the same reason that solving equations of motion in Newtonian mechanics becomes easier or sometimes even totally trivial if you have enough independent conserved quantities.