Spacetime and Geometry - Sean M. Carroll

Notes by Gene

2026

Contents

1 Special Relativity and Flat Spacetime (pp. 1 - 47) 1
1.1 Introductory Concepts 1
1.2 Special Relativity 2
1.2.1 Basics 2
1.2.2 The Twin Paradox 3
1.2.3 Lorentz Transformations 4
1.2.4 4-Vectors 4
1.2.5 Dual Vectors 5

1 Special Relativity and Flat Spacetime (pp. 1 - 47)

1.1 Introductory Concepts

General relativity: basically gravity and geodesics.

- Most forces are embedded in spacetime (e.g. Electromagnetic field).
- "Gravity" isn't a force, but the *curvature* of spacetime. It's an *inherent* feature.

Note: We adopt the convention $c = 1$ for the speed of light.

We begin from Newton's law of gravitation: $F = GMm/r^2$. This can be rewritten as Poisson's Equation, describing gravity in terms of a potential.

$$\underbrace{\nabla^2\Phi}_{curvature} = \underbrace{4\pi G\rho}_{source}$$

- Φ is the gravitational potential, a scalar field of potential energy per mass.
- ρ is the mass density, i.e. the distribution of mass in space. This is the source that "generates" the gravitational potential.
- Acceleration is simple: $a = \nabla\Phi$, the gradient.

Roughly, mass affects the shape of the gravitational potential field. Similarly, in general relativity, spacetime curves in the presence of matter and energy.

However, Poisson's equation (and Newton's law) describes an instantaneous transfer of information: a change in mass immediately affects gravity. This can't be true, if special relativity is accepted. So the equations must be modified.

Later we'll derive

$$
\begin{gathered}
R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi G T_{\mu\nu} \\
\frac{d^2 x^\mu}{d\lambda^2} + \Gamma^\mu_{\rho\sigma} \frac{dx^\rho}{d\lambda} \frac{dx^\sigma}{d\lambda} = 0
\end{gathered}
$$

The first equation is Einstein's equation. The second describes a geodesic, where $x^\mu(\lambda)$ is the parametrized path.

1.2 Special Relativity

1.2.1 Basics

(Largely a review) Special relativity is a theory of 4D spacetime, the "Minkowski Space" (a special case of 4D manifolds). The distinction between "space" and "time" is arbitrary — you have to think of "spacetime" as a whole entity, not the sum of its parts.

In Newtonian dynamics, we have a well-defined notion of "simultaneity." In SR, we don't. Particles stay within their light cones in a spacetime diagram.

In Cartesian coordinates, distance is invariant under coordinate changes:

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

In SR, the "spacetime interval" is invariant under coordinate changes:

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

We sum over Greek indices for spacetime coordinates and Latin indices for space coordinates. We also define the spacetime metric $\eta_{\mu\nu}$.

$$
x^\mu : \begin{cases} x^0 = ct \\ x^1 = x \\ x^2 = y \\ x^3 = z \end{cases} \quad , \quad x^i : \begin{cases} x^1 = x \\ x^2 = y \\ x^3 = z \end{cases}
$$

$$
\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
$$

Then we can rewrite the spacetime interval more compactly. The summation convention is that indices in both superscript & subscript are summed over.

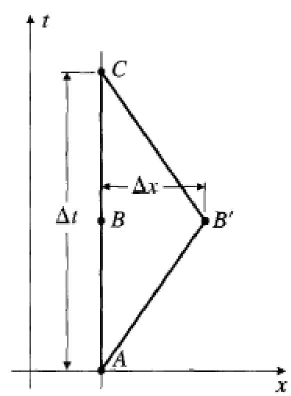
$$(\Delta s)^2 = \eta_{\mu\nu}\Delta x^\mu \Delta x^\nu$$

$$\begin{aligned}
(\Delta s)^2 &< 0 : \text{timelike separated} \\
(\Delta s)^2 &= 0 : \text{null} \\
(\Delta s)^2 &> 0 : \text{spacelike separated}
\end{aligned}$$

Define "proper time" τ to measure the time elasped as seen by an observer moving in a straight path between two events. I.e. This is the time recorded by the *observer*. We add a minus sign because otherwise, timelike separated events would be negative and that's weird.

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

1.2.2 The Twin Paradox



The Twin Paradox

Let $B = \left(\frac{1}{2}\Delta t, 0\right)$ and $B' = \left(\frac{1}{2}\Delta t, \Delta x\right)$ with $\Delta x = \frac{1}{2}v\Delta t$.

Alice travels straight from A to B. Bob travels from A to B'. The proper times (with $c = 1$ and $v \leq c$) are

$$\begin{aligned}
\text{Alice} : \Delta\tau_{AB} &= \frac{1}{2}\Delta t \\
\text{Bob} : \Delta\tau_{AB'} &= \sqrt{\left(\frac{1}{2}\Delta t\right)^2 - (\Delta x)^2} = \frac{1}{2}\Delta t\sqrt{1 - v^2}
\end{aligned}$$

Bob has aged less than Alice! Because a nonstraight path in *spacetime* has a different interval than a straight path. Additionally, the elasped time for the observer depends on the path you take.

1.2.3 Lorentz Transformations

Transformations in spacetime. The basic example is a translation with δ as the potential. We "shift" one variable, either spatial or temporal. Note that the distance Δx^μ between two points remains the same, thus the interval $(\Delta s)^2$ is invariant.

$$x^\mu \longrightarrow x^{\mu'} = \delta^{\mu'}_{\mu}\left(x^\mu + a^\mu\right), \quad \delta^{\mu'}_{\mu} = \begin{cases} 1 & \mu' = \mu \\ 0 & \mu' \neq \mu \end{cases}$$

Other examples are rotations and boosts (offsets by a constant velocity vector). We model these with a matrix Λ.

$$x^{\mu'} = \Lambda^\mu_\nu x^\nu, \quad \text{i.e. } \; x' = \Lambda x$$

The interval $(\Delta s)^2$ must remain invariant, so Δx^μ must not change. Thus we have

$$\begin{gathered} (\Delta s)^2 = (\Delta x)^T \eta (\Delta x) = (\Delta x')^T \eta (\Delta x') \\ (\Delta x)^T \eta (\Delta x) = (\Delta x)^T \Lambda^T \eta \Lambda (\Delta x) \\ \eta = \Lambda^T \eta \Lambda \end{gathered}$$

The Λ that satisfies the above is a Lorentz transformation. Under matrix multiplication, the Lorentz transformations form a Lorentz group. The set of translations and Lorentz transformations form a 10-parameter, non-Abelian group known as the *Poincaré group*.

As an example, consider a boost in the x-direction (which can be thought of as rotations between space and time directions).

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

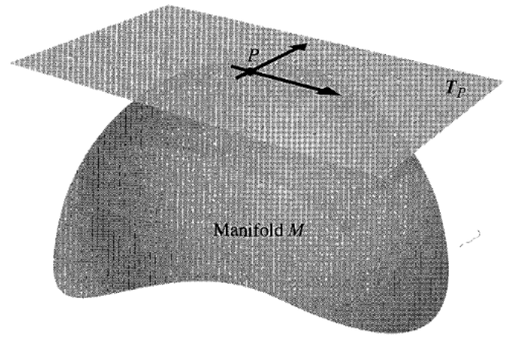
By applying Λ, we get $t' = t\cosh\phi - x\sinh\phi$ and $x' = -t\sinh\phi + x\cosh\phi$. Focusing on the point at x', we see that it moves according to $t\sinh\phi = x\cosh\phi$. Thus $v = x/t = \tanh\phi$.

Using $\phi = \tanh^{-1} v$, we can reconstruct the usual Lorentz transformations. Set $\gamma = \frac{1}{\sqrt{1-v^2}}$ and $c = 1$ to get

$$\begin{gathered} t' = \gamma(t - vx) \\ x' = \gamma(x - vt) \end{gathered}$$

1.2.4 4-Vectors

4-vectors are located at a single point in spacetime, not stretched from A to B. At each point p, the set of al possible vectors located at p is T_p, the tangent space of p. The set of all tangent spaces of an n-dimensional manifold M can be assembled into a $2n$-dimensional manifold $T(M)$, the "tangent bundle."



Tangent Space and Manifold Diagram (it doesn't exactly look like this though)

A vector A can be represented as a linear combination of basis vectors $\hat{e}_{(\mu)}$ and coefficients A^μ:

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

The basis vectors transform as expected with Lorentz transformations:

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

A tangent vector $V(\lambda)$ to a curve in spacetime parametrized by $x^\mu(\lambda)$ can be represented

$$V^\mu = \frac{dx^\mu}{d\lambda} \qquad \text{(components)}$$
$$V = V^\mu \hat{e}_{(\mu)}$$

1.2.5 Dual Vectors

Kind of a complement to a vector space. The dual vector of the tangent space T_p is denoted T_p^*, the "cotangent space." A vector ω is in T_p^* iff:

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

Essentially, the dual space is the space of all linear maps from T_p to $\mathbb{R}$. A quick and dirty way to think about it: upper indices denote vectors, lower indices denote dual vectors.

Examples of dual vectors include the gradient, the row vector (corresponding to column vectors), etc.

We can define a "basis dual vector" $\hat{\theta}^{(\nu)}$ in order to represent the dual vector ω:

$$\hat{\theta}^{(\nu)}\left(\hat{e}_{(\mu)}\right) = \delta^\nu_\mu$$
$$\omega = \omega_\mu \hat{\theta}^{(\mu)}$$

Then we can find the effect of a dual vector on a vector:

$$\begin{aligned}\omega(V) &= \omega_\mu \hat{\theta}^{(\mu)} \left(V^\nu \hat{e}_{(\nu)}\right) \\ &= \omega_\mu V^\nu \hat{\theta}^{(\mu)} \left(\hat{e}_{(\nu)}\right) \\ &= \omega_\mu V^\nu \delta^\mu_\nu \\ &= \omega_\mu V^\mu \in \mathbb{R}\end{aligned}$$

A similar calculation shows that vectors are also linear maps on dual vectors, which is a nice symmetry.

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

A simple example of a dual vector in spacetime is the gradient of a scalar function. We denote $d\phi$ as the set of partial derivatives w.r.t. spacetime coordinates.

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

The chain rule is intuitive.

$$\frac{\partial \phi}{\partial x^{\mu'}} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial \phi}{\partial x^\mu} = \Lambda^\mu_{\mu'} \frac{\partial \phi}{\partial x^\mu}$$

We define the shorthand $\partial_\mu \phi = \frac{\partial \phi}{\partial x^\mu}$. The gradient of a tangent vector along the parametrized curve $x^\mu(\lambda)$ is

$$\frac{d\phi}{d\lambda} = \partial_\mu \phi \frac{\partial x^\mu}{\partial \lambda}$$