Spacetime and Geometry - Sean M. Carroll

Notes by Gene

2026

Contents

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

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

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 $\eta_{\mu\nu}$.

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

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

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

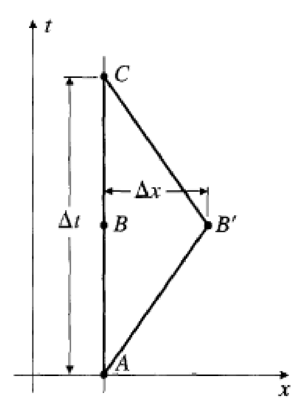
$$(\Delta s)^2 = \eta_{\mu\nu}\Delta x^\mu \Delta x^\nu$$

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

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.

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

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

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

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).

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

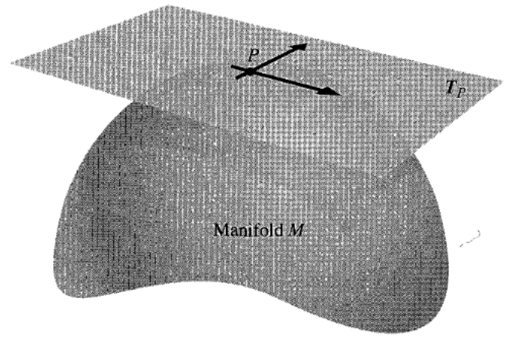
By applying Λ, we get $t' = t\cosh\phi - x\sinh\phi$ and $x' = -t\sinh\phi + x\cosh\phi$. Focusing on the point at x', we see that it moves according to $t\sinh\phi = x\cosh\phi$. Thus $v = x/t = \tanh\phi$.

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

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

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

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

1.3 Tensors

Tensors are a generalization of vectors and dual vectors. Recall that a dual vector is a map from vectors to $\mathbb{R}$. Tensors are a map from a collection of vectors & dual vectors to $\mathbb{R}$. For a tensor of rank (k, l), we map the vectors in T_p^* and dual vectors in T_p to $\mathbb{R}$.

$$T : \underbrace{T_p^* \times \cdots \times T_p^*}_{\text{k terms}} \times \underbrace{T_p \times \cdots \times T_p}_{\text{l terms}} \longrightarrow \mathbb{R}$$

$\times$ is the Cartesian product. $T_p \times T_p$ is the space of the ordered pair of vectors.

A tensor is *multilinear*, meaning it acts linearly in each argument. So for example, a tensor of rank $(1, 1)$ would act like

$$T(a\omega + b\eta, cV + dW) = acT(\omega, V) + adT(\omega, W) + bcT(\eta, V) + bdT(\eta, W)$$

We can further say that a scalar is a tensor of rank (0, 0); a vector is a tensor of rank (1, 0); and a dual vector is a tensor of rank (0, 1). The space of all tensors of rank (k, l) form a vector space, where we can define the *tensor product* $\otimes$ as

$$\begin{gathered}
T : \text{tensor of rank } (k, l) \\
S : \text{tensor of rank } (m, n) \\
T \otimes S \left(\omega^{(1)}, \cdots, \omega^{(k)}, \cdots, \omega^{(k+m)}, V^{(1)}, \cdots, V^{(l)}, \cdots, V^{(l+n)}\right) \\
= T\left(\omega^{(1)}, \cdots, \omega^{(k)}, \cdots, \omega^{(k+m)}\right) \times \\
S\left(V^{(1)}, \cdots, V^{(l)}, \cdots, V^{(l+n)}\right)
\end{gathered}$$

$T \otimes S$ is a rank $(k+m, l+n)$ tensor.

In spacetime, for a rank (k, l) tensor, there's 4^{k+l} basis tensors. We can represent this simply by summing over each bases.

$$T = T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l} \hat{e}_{(\mu_1)} \otimes \cdots \otimes \hat{e}_{(\mu_k)} \otimes \hat{\theta}^{(\nu_1)} \otimes \cdots \otimes \hat{\theta}^{(\nu_k)}$$

Similarly, the action of tensors on a st of vectors and dual vectors is given below. Note that each a (k, l) tensor has k upper and l lower indices. Upper indices transform like vectors, while lower indices transform like dual vectors.

$$T\left(\omega^{(1)}, \cdots, \omega^{(k)}, V^{(1)}, \cdots, V^{(l)}\right) = T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l}\ \omega^{(1)}_{\mu_1} \cdots \omega^{(k)}_{\mu_k} V^{(1)\nu_1} \cdots V^{(l)\nu_l}$$

Tensors transform as expected under Lorentz:

$$T^{\mu'_1 \cdots \mu'_k}{}_{\nu'_1 \cdots \nu'_l} = \Lambda^{\mu'_1}{}_{\mu_1} \cdots \Lambda^{\mu'_k}{}_{\mu_k} \Lambda^{\nu_1}{}_{\nu'_1} \cdots \Lambda^{\nu_l}{}_{\nu'_l}\ T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l}$$

We're also under no obligation to transform every component. For example, we could just transform a single component of a (1, 1) tensor to make it a vector $\to$ vector map.

$$T^{\mu}{}_{\nu} : V^{\nu} \longrightarrow T^{\mu}{}_{\nu} V^{\nu}$$

It gets complicated, but just try to keep track of the indices. Also note that tensors $\neq$ vectors. Tensors are geometric entities independent of the coordinate system.

1.3.1 Examples of Tensors

Metric The $\eta_{\mu\nu}$ is a (0, 2) tensor. Its action on two vectors is known as the "inner product":

$$\eta(V, W) = \eta_{\mu\nu} V^{\mu} W^{\nu} = V \cdot W$$

If the inner product vanishes, the two vectors are orthogonal. Note that the inner product is a scalar, thus invariant under Lorentz transformations. So a set of orthogonal vectors in one inertial Cartesian frame is also orthogonal after some Lorentz transforms, even though they may not look like it.

The norm of a vector is its inner product with itself:

$$\eta(V, V) = \eta_{\mu\nu} V^{\mu} V^{\nu} \begin{cases} < 0 & V^{\mu} \text{ is timelike} \\ = 0 & V^{\mu} \text{ is lightlike/null} \\ > 0 & V^{\mu} \text{ is spacelike} \end{cases}$$

Kronecker Delta The Kronecker delta δ^μ_ρ is another example of a tensor. It's an identity map from a vector to the same vector. Note that we put the indices in the same "column." That's because order doesn't matter, $\delta^\mu_\rho = \delta^\mu{}_\rho = \delta_\rho{}^\mu$.

The Kronecker delta is defined as the metric times its inverse:

$$\delta^\mu_\rho = \eta^{\mu\nu}\eta_{\nu\rho} = \eta_{\rho\nu}\eta^{\nu\mu}$$

Levi-Civita The Levi-Civita tensor has rank $(0,4)$ in spacetime. We define it as

$$\tilde{\epsilon}_{\mu\nu\rho\sigma} = \begin{cases} +1 & \text{if } \mu\nu\rho\sigma \text{ is an even permutation of 0123} \\ -1 & \text{if } \mu\nu\rho\sigma \text{ is an odd permutation of 0123} \\ 0 & \text{otherwise} \end{cases}$$

Electromagnetic Field Strength This is a rank $(0,2)$ tensor composed of the E_i and B_i vectors. Note the Latin indices indicating spacelike 1, 2, 3 components. By Lorentz transforming this tensor, we can get the E and B fields in a different reference frame.

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

So for example, the field tensor transforms under a Lorentz boost as

$$F^{\mu'\nu'} = \Lambda^{\mu'}{}_\mu \Lambda^{\nu'}{}_\nu \, F^{\mu\nu}$$

1.3.2 Manipulating Tensors

Contraction A contraction of a tensor reduces its rank from $(k,l) \longrightarrow (k-1, l-1)$. To do this, we sum over the upper and lower indices.

$$S^{\mu\rho}{}_\sigma = T^{\mu\nu\rho}{}_{\sigma\nu}$$

Order matters. In general $T^{\mu\nu\rho}{}_{\sigma\nu} \neq T^{\mu\rho\nu}{}_{\sigma\nu}$.

Raising & Lowering Indices We can also raise and lower indices in tensors using the metric and its inverse. For example,

$$T^{\alpha\beta\mu}{}_\delta = \eta^{\mu\gamma}\, T^{\alpha\beta}{}_{\gamma\delta}$$
$$T_\mu{}^\beta{}_{\gamma\delta} = \eta_{\mu\alpha}\, T^{\alpha\beta}{}_{\gamma\delta}$$

Since the metric and its inverse are, well, inverse of each other, we also have the following:

$$A^\lambda B_\lambda = \eta^{\lambda\rho} A_\rho \eta_{\lambda\sigma} B^\sigma = \delta^\rho_\sigma A_\rho B^\sigma = A_\sigma B^\sigma$$

Symmetric & Antisymmetric Tensors A tensor is called symmetric if it remains unchanged under exchange of indices. In the following, S is symmetric in μ and ν.

$$S_{\mu\nu\rho} = S_{\nu\mu\rho}$$

We can symmetrize a tensor by summing over its permutations. The symmetric part is denoted by parentheses.

$$T_{(\mu_1\mu_2\cdots\mu_n)\rho}{}^{\sigma} = \frac{1}{n!}\left(T_{\mu_1\mu_2\cdots\mu_n\rho}{}^{\sigma} + \text{ sum over permutations of } \mu_1 \cdots \mu_n\right)$$

A tensor is called antisymmetric if it changes sign under exchange of indices. In the following, A is symmetric in μ and ρ.

$$A_{\mu\nu\rho} = -A_{\rho\nu\mu}$$

We can antisymmetrize a tensor by taking the alternating sum over its permutations. The antisymmetric part is denoted by a bracket. "Alternating sum" means that the permutations that result from an odd number of exchanges are given a negative sign.

$$T_{[\mu_1\mu_2\cdots\mu_n]\rho}{}^{\sigma} = \frac{1}{n!}\left(T_{\mu_1\mu_2\cdots\mu_n\rho}{}^{\sigma} + \text{ alternating sum over permutations of } \mu_1 \cdots \mu_n\right)$$

When contracting over a symmetric tensor, only the symmetric parts matter. Similar for antisymmetric tensors.

$$X^{(\mu\nu)}Y_{\mu\nu} = X^{(\mu\nu)}Y_{(\mu\nu)}$$
$$X^{[\mu\nu]}Y_{\mu\nu} = X^{[\mu\nu]}Y_{[\mu\nu]}$$

We can also decompose a tensor into a sum for *two indices*. For 3 and above, we generally can't.

$$T_{\mu\nu\rho\sigma} = T_{(\mu\nu)\rho\sigma} + T_{[\mu\nu]\rho\sigma}$$

Trace The trace of a tensor, denoted X is given by

$$X = X^{\lambda}{}_{\lambda}$$

That's the simple case of a (1, 1) tensor. For a (0, 2) tensor $Y_{\mu\nu}$, we raise indices first.

$$Y = Y^{\lambda}{}_{\lambda} = \eta^{\mu\nu}Y_{\mu\nu}$$

1.4 Tensor Form of Maxwell's Equations

The classical differential form of Maxwell's Equations:

$$
\begin{cases}
\nabla \times \mathbf{B} - \partial_t \mathbf{E} = J & (1) \\
\nabla \cdot \mathbf{E} = \rho & (2) \\
\nabla \times \mathbf{E} + \partial_t \mathbf{B} = 0 & (3) \\
\nabla \cdot \mathbf{B} = 0 & (4)
\end{cases}
$$

We convert to component notation. Spatial indices (Latin i, j, k) are raised and lowered at will since the metric δ_{ij} is an identity on flat 3D space. We also replace J with the current 4-vector $J^\mu = (\rho, J^x, J^y, J^z)$.

$$
\begin{cases}
\tilde{\epsilon}^{ijk} \partial_j B_k - \partial_0 E^i = J^i & (5) \\
\partial_i E^i = J^0 & (6) \\
\tilde{\epsilon}^{ijk} \partial_j E_k + \partial_0 B^i = 0 & (7) \\
\partial_i B^i = 0 & (8)
\end{cases}
$$

We use the electromagnetic field strength tensor from before.

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

Now note that $F^{0i} = \eta^{00}\eta^{ii}F_{0i} = E^i$ (we raise and lower indices at will). Also, $F^{ij} = \tilde{\epsilon}^{ijk}B_k$. Thus the first two of Maxwell's equations can be rewritten as

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

We can merge the two equations into one:

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

Similarly, the 3rd and 4th of Maxwell's equations can be merged into one, giving us the two equations that describe E and B:

$$
\partial_\mu F^{\nu\mu} = J^\nu
$$
$$
\partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0
$$