2. General Relativity

Reading: Chapter 3

An accurate theory of cosmology was really not possible prior to the discovery of GR.

We will largely use pseudo-Newtonian approximations to GR for our calculations in this course, but one can only see what those approximations should be thanks to GR.

Special Relativity

Postulates of theory:

1. There is no state of "absolute rest."
2. The speed of light in vacuum is constant, independent of state of motion of emitter.

Implies: Simultaneity of events and spatial separation of events depend on state of motion of observer.

Observers in relative motion disagree on the time separation Δt and the spatial separation $\Delta l = \left[(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2\right]^{1/2}$ between the same pair of events.
But they agree on the "spacetime interval" $(\Delta s)^2 = -c^2(\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2$ between events.
Analogy: Stationary observers with rotated reference frames disagree on Δx, Δy but agree on $\Delta l = \left[(\Delta x)^2 + (\Delta y)^2\right]^{1/2}$.

The Equivalence Principle

"Special" relativity restricted to uniformly moving observers. Can it be generalized?

Newtonian gravity: $\mathbf{a} = \mathbf{F}/m, \qquad \mathbf{F} = -GMm\hat{\mathbf{r}}/r^2.$
Why is this unsatisfactory?

Implicitly assumes infinite speed of signal propagation.

Coincidental equality of inertial and gravitational mass.

Einstein, 1907: "The happiest thought of my life." If I fall off my roof, I feel no gravity.

Equivalence between uniform gravitational field and uniform acceleration of frame. True in mechanics. *Assume* exact equivalence, i.e., for electrodynamics as well.
Equivalence principle implies gravitational and inertial masses *must* be equal.
Allows extension of relativity to accelerating frames.
Implies that extension of relativity *must* involve gravity.

From the equivalence principle alone, can infer
(a) that gravity should bend light
(b) gravitational redshift.
From gravitational redshift, one can demonstrate that gravity affects the flow of time and that energy has an effective *gravitational* mass $m = E/c^2$.

Restatement of equivalence principle: In the coordinate system of a freely falling observer, special relativity always holds locally (to first order in separation). No gravity.

Over larger scales (second order in separation), gravity doesn't vanish in a freely falling frame — tidal effects. E.g., freely falling objects in an inhomogeneous gravitational field may accelerate towards or away from each other.

Gravitational Redshift Thought Experiment

Consider a box of height h in a gravitational field of acceleration g.

If you shoot an object of mass m from the bottom of the box to the top, how much kinetic energy does it lose?

What happens if you send a photon of frequency ν from the bottom of the box to the top? Not obvious.

Apply the equivalence principle, considering a box that is accelerating upward with acceleration of g in empty space.

The time required for the photon to traverse the box is $t = h/c$.

In this time the box acquires an upward velocity $v = gt = gh/c$.

The usual Doppler formula tells us that the frequency of the photon at the top is therefore decreased (wavelength increased) by a fractional amount $\Delta\nu/\nu = v/c = gh/c^2$.

Since $E = h_P\nu$ (where h_P is Planck's constant, not the height of the box), $\Delta E/E = \Delta\nu/\nu$.

Therefore $\Delta E = (gh/c^2)E = (E/c^2)gh$, the same as for a massive object of mass $m = E/c^2$.

Summary of General Relativity

With aid of equivalence principle, can change relativity postulate from "There is no absolute rest frame" to "There is no absolute set of inertial frames."

More informally, "There is no absolute acceleration."

Uniform acceleration can be treated as uniform gravitational field; the two are *indistinguishable*.

A *geodesic path* is a path of shortest distance, e.g.,

In flat space, a straight line.

On a sphere, a great circle.

In relativity, a geodesic path is a path of shortest *spacetime interval*.

In flat spacetime, a straight line at constant velocity.

Freely falling particles move along these geodesics in flat spacetime.

GR description of gravity:

All freely falling particles follow geodesic paths in curved spacetime.

Distribution of matter (more generally, stress-energy) determines spacetime curvature.

Misner, Thorne, and Wheeler's catchy summary of GR:

Spacetime tells matter how to move. (Along geodesic paths.)

Matter tells spacetime how to curve. (Field equation.)

Compare to equivalent description of Newtonian gravity:

Gravitational force tells matter how to accelerate. ($F = ma$.)

Matter tells gravity how to exert force. ($F = GMm/r^2$.)

The "equivalence of inertial and gravitational mass" in Newton's description is not a coincidence but a *necessary consequence* of the assumption that all freely falling particles follow geodesic paths.

Space curvature and the spatial metric

On a flat, two-dimensional surface, the angles of a triangle satisfy

$$\alpha + \beta + \gamma = \pi.$$

The spatial separation dl between two nearby points is

$$dl^2 = dx^2 + dy^2,$$

or, in polar coordinates,

$$dl^2 = dr^2 + r^2 d\theta^2.$$

The total length l of a path can be found by integrating dl along the path.

On the surface of a sphere, the angles of a triangle add to

$$\alpha + \beta + \gamma = \pi + A/R^2,$$

where A is the area enclosed by the triangle and R is the radius of the sphere.
If r is the spatial distance along a great circle from the origin (e.g., the North Pole) and θ the azimuthal angle (e.g., the longitude), then the spatial separation of nearby points is

$$dl^2 = dr^2 + R^2 \sin^2(r/R) d\theta^2.$$

Analogously, on a negatively curved (saddle-like) surface of constant curvature, the angles of a triangle add to

$$\alpha + \beta + \gamma = \pi - A/R^2,$$

and the spatial separation is

$$dl^2 = dr^2 + R^2 \sinh^2(r/R) d\theta^2.$$

Formulas that relate coordinate separations to length separations are called *metrics.* In general, a metric is a matrix (more precisely a tensor) of functions that tells how to go from differential coordinate separations to differential distances.
For more examples and discussion, see my notes on metrics and metric notation.

These results for two-dimensional surfaces can be naturally generalized three-dimensional spaces.

The metrics for flat, positively curved, and negatively curved spaces of constant curvature radius R, in "spherical" coordinates, are, respectively,

$$\begin{aligned}
dl^2 &= dr^2 + r^2 \left[d\theta^2 + \sin^2\theta d\phi^2\right] \\
dl^2 &= dr^2 + R^2 \sin^2(r/R) \left[d\theta^2 + \sin^2\theta d\phi^2\right] \\
dl^2 &= dr^2 + R^2 \sinh^2(r/R) \left[d\theta^2 + \sin^2\theta d\phi^2\right].
\end{aligned}$$

Note that $d\theta^2 + \sin^2\theta d\phi^2$ is just the squared angular separation, $d\Omega^2$ in Ryden's notation.

Positive curvature $\Longrightarrow$geodesics "accelerate" (in 2nd derivative sense) towards each other. Initially "parallel" geodesics converge.
Example: great circles on a sphere.

Zero curvature $\Longrightarrow$no geodesic "acceleration." Initially parallel geodesics stay parallel. Euclidean geometry.

Example: straight lines on a plane.

Negative curvature $\Longrightarrow$geodesics "accelerate" away from each other. Initially parallel geodesics diverge.

Example: geodesics on a saddle.

Example of how following geodesic paths can "look like" gravity: explorers following lines of constant longitude from the equator to the south pole and onwards, vs. two point masses in outer space.

Caution: this is a loose analogy. Freely falling objects follow *spacetime* geodesics, not just *space* geodesics.

Spacetime metric

With special relativity, Einstein showed that one cannot separate space and time in a way that is independent of the observer. We must therefore work with a more general 4-d spacetime.

We can generalize the notion of metric to 4-d spacetime, relating coordinate separation of events to the spacetime intervals between them.

In special relativity (flat spacetime) with Cartesian coordinates, the metric is

$$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$$

or in spherical coordinates

$$ds^2 = -c^2dt^2 + dr^2 + r^2d\Omega^2\ .$$

In a general case, with coordinate separations between two events dx^μ where $\mu = 0, 1, 2, 3$ (with 0 representing time), the spacetime interval is

$$ds^2 = \sum_{\mu,\nu} g_{\mu\nu}dx^\mu dx^\nu,$$

where the quantity $g_{\mu\nu}$ is called the metric.

It is a symmetric, 4×4 matrix, with ten independent components.

Note that $g_{\mu\nu}$ is, in general, a function of spacetime position.

Observers in relative motion, and/or with different coordinate systems, may disagree on dx^μ and $g_{\mu\nu}$, but they will agree on ds^2.

Newtonian Gravity Equations

To make the analogy with GR, it is useful to formulate Newtonian gravity in terms of

An *acceleration equation*

$$\frac{d^2\mathbf{x}}{dt^2} = \mathbf{g} = -\vec{\nabla}\Phi$$

and the *Poisson equation*

$$\nabla^2\Phi = -\vec{\nabla}\cdot\mathbf{g} = 4\pi G\rho\ .$$

The first equation governs the motion of freely falling particles in terms of the gravitational potential $\Phi(\mathbf{x})$.

The second equation determines the potential $\Phi(\mathbf{x})$ from the matter density field $\rho(\mathbf{x})$.
You may be more familiar with the integral version of the Poisson equation

$$\Phi(\mathbf{r}) = -G \int \frac{\rho(\mathbf{x})}{|\mathbf{x} - \mathbf{r}|} d^3r \ ,$$

but the differential form is often more useful.
Note that in Cartesian coordinates

$$\nabla^2 \Phi = \frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} \ .$$

Note that the gravitational potential has units of (velocity)^2, and that traversing a gravitational potential difference $\Delta\Phi$ will typically induce a velocity change $\Delta v^2 \sim \Delta\Phi$.

General Relativity Equations

Mathematically, GR is defined by "two" equations.
The first is the equation for geodesic paths, which gives the equation of motion for freely falling particles (or photons) in a specified coordinate system.
In practice, this equation represents four 2nd-order differential equations that determine $x^\alpha(\tau)$, given initial position and 4-velocity, where τ is proper time measured along path of particle:

$$\frac{d^2 x^\alpha}{d\tau^2} + \sum_{\mu,\nu} F\,[\text{metric}]\, \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = 0.$$

The geodesic equation is the relativistic analog of the Newtonian equation $\mathbf{g} = -\vec{\nabla}\Phi$.
The metric $g_{\mu\nu}$ is the relativistic generalization of the gravitational potential. Note that every component of $g_{\mu\nu}$ may change as a function of spacetime location.
When the Newtonian limit is accurate, it is typically g_{00} that is equivalent to the Newtonian Φ.

The second is the *Einstein Field Equation*, which relates the curvature of spacetime to the distribution of matter and energy.
This is the analog of Poisson's equation $\nabla^2\Phi = 4\pi G\rho$.
The Field Equation is usually written

$$G_{\mu\nu} = 8\pi G_{\text{Newton}} T_{\mu\nu}.$$

$G_{\mu\nu}$ is the *Einstein tensor*, built from $g_{\mu\nu}$ and its derivatives up to second order. (Like $\nabla^2\Phi$.)
$T_{\mu\nu}$ is the *stress energy tensor*, the relativistic generalization of density.
For an ideal fluid at rest, $T_{\mu\nu} = \text{diag}(\rho, p/c^2, p/c^2, p/c^2)$, where ρc^2 is the energy density.
The constant $8\pi G_{\text{Newton}}$, where G_{Newton} is Newton's gravitational constant, is determined by demanding correspondence to Newtonian gravity in the appropriate limit.

Solutions of the field equation

Note that $G_{\mu\nu} = 8\pi G_{\text{Newton}} T_{\mu\nu}$ is a set of ten, second-order differential equations for the ten components of $g_{\mu\nu}$.

Second-order $\Longrightarrow$

boundary conditions matter

spacetime can be curved even where $T_{\mu\nu} = 0$

propagating wave solutions exist

Nonlinear $\Longrightarrow$hard to solve.

Some exact solutions, e.g.

$\mathbf{T} = 0$ everywhere $\longrightarrow$ flat spacetime, "Minkowski space"

Spherically symmetric, flat at ∞, point mass at $r = 0 \longrightarrow$ Schwarzschild solution

Generalization to include angular momentum $\longrightarrow$ Kerr solution

Homogeneous cosmologies, which we will study

In other cases, approximate, by considering small departures from an exact solution (perturbation theory).

Recall that the (Newtonian) gravitational potential Φ has units of velocity^2.

The "weak field" limit of GR corresponds to $\Phi \ll c^2$. Spacetime curvature is weak; photons travel on nearly straight paths.

The combination of the weak field limit and $v \ll c$ leads to the Newtonian limit, in which GR approaches Newtonian gravity.

Tests of GR

High-Precision Quantitative Tests

- Yields Newtonian gravity in appropriate limit
- Precision tests of equivalence principle
- Precession of Mercury – the key from Einstein's point of view
- Bending of light – historically important
- Gravitational redshift
- Higher-order solar system tests $\Longrightarrow$measured values of "post-Newtonian parameters" agree with GR predictions
- Post-Newtonian effects (up to $\sim (v/c)^3$) measured for pulsars in binary systems – precession, gravitational time delay

Lower precision qualitative tests

- Gravitational lenses
- Black holes: existence of dark massive objects, producing strong gravitational redshifts, apparently with event horizons rather than hard surfaces
- The Event Horizon Telescope has imaged the shadow of the event horizon and strong light bending effects of the central supermassive black holes in M87 and the Milky Way (in 2019 and 2021, respectively), with the expected properties.
- Successes of the big bang theory built on GR; GR predicts expansion of the universe, curvature of space

Gravitational waves

As noted by Einstein already in 1916, GR permits the existence of propagating spacetime ripples, generated by accelerating masses.

- Binary pulsar orbits shrink because energy radiated in gravitational waves; theory and measurement agree at 1% level (1974 onward)
- Direct detection of gravitational waves from merging black holes (at a distance of about 1 Glyr); properties agree with model predictions for black holes (2015)
- Direct detection of gravitational waves from a merging binary neutron star, also seen in electromagnetic radiation; speed of light and speed of gravitational waves are identical to high precision (2017)
- Now many dozens of gravitational wave mergers detected, consistent with GR predictions for merging black holes at the few percent precision level.

Cosmology and GR

Despite these impressive tests, application to cosmology requires gigantic extrapolation in length and time scale.

Can't rest comfortably on empirical basis of small-scale tests.

Cosmological models based on GR are impressively successful, but they require two strange ingredients: dark matter and dark energy.

Existence of these ingredients could be an indication that GR is breaking down in some way on cosmological scales, though we will generally take the view that it is not.