1 The problem of quantum gravity

These are lectures on quantum gravity. To start, we better understand clearly what problem we are trying to solve when we say 'quantum gravity.' At low energies, the classical action is

$$S = \frac{1}{16\pi G_N} \int \sqrt{-g}\,(R - 2\Lambda + \mathcal{L}_{matter})\ . \tag{1.1}$$

Why not just quantize this action? The answer of course is that it is not renormalizable. This does not mean it is useless to understand quantum gravity, it just means we have to be careful about when it is reliable and when it isn't. In this first lecture we will consider gravity as a low-energy effective field theory,[1] see when it breaks down, and make some general observations about what we should expect or not expect from the UV completion.

1.1 Gravity as an effective field theory

The rules of effective field theory are:

1. Write down the most general possible action consistent with the symmetries;

2. Keep all terms up to some fixed order in derivatives;

3. Coefficients are fixed by dimensional analysis, up to unknown order 1 factors (unless you have a good reason to think otherwise);

4. Do quantum field theory using this action, including loops;

5. Trust your answer only if the neglected terms in the derivative expansion are much smaller than the terms you kept.

This works for renormalizable or non-renormalizable theories. Let's follow the steps for gravity. Our starting assumption is that nature has a graviton — a massless spin-2 field. This theory can be consistent only if it is diffeomorphism invariant.

---

[1]Donaghue gr-qc/9512024.

Counting metric degrees of freedom

This can be argued various ways[2]; we'll just count degrees of freedom. In 4D, a massless particle has two degrees of freedom (2 helicities). Similarly, the metric $g_{\mu\nu}$ has[3]

$$10 \text{ components} - 4 \text{ diffeos} - 4 \text{ non-dynamical} = 2 \text{ dof} . \tag{1.2}$$

In D dimensions, we count dof of a massless particle by looking at how the particle states transform under $SO(D-2)$, the group of rotations that preserve a null ray.[4] A spin-2 particle transforms in the symmetric traceless tensor rep of $SO(D-2)$, which has dimension $\frac{1}{2}(D-2)(D-1) - 1 = \frac{1}{2}D(D-3)$. Similarly, assuming diffeomorphism invariance, the metric has

$$\tfrac{1}{2}D(D+1) - D - D = \tfrac{1}{2}D(D-3) \tag{1.3}$$

degrees of freedom.

Note that in $D = 3$, the metric has no (local) dof. It turns out that it does have some nonlocal dof; this will be useful later in the course.

Back to effective field theory: Steps 1 and 2, the derivative expansion

The only things that can appear in a diff-invariant Lagrangian for the metric are objects built out of the Riemann tensor $R_{\mu\nu\rho\sigma}$ and covariant derivatives ∇_μ. Each Riemann contains $\partial\partial g$, so the derivative expansion is an expansion in the number of R's and ∇'s. Up to 4th order in derivatives,

$$S = \frac{1}{16\pi G_N} \int \sqrt{-g} \left(-2\Lambda + R + c_1 R^2 + c_2 R_{\mu\nu} R^{\mu\nu} + c_3 R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} + \cdots \right) \tag{1.4}$$

So the general theory is the Einstein-Hilbert term plus higher curvature corrections. We have ignored the matter terms $\mathcal{L}_{matter}$ and matter-curvature couplings, like ϕR.

---

[2]See Weinberg QFT V1, section 5.9, and the discussion of the Weinberg-Witten theorem below, and Weinberg Phys. Rev. 135, B1049 (1964).

[3]In more detail: a 4x4 symmetric matrix has 10 independent components. In 4D we have 4 functions worth of diffeomorphisms, $x^\mu \to x^{\mu'}(x^\mu)$. And $\ddot{g}_{0\mu}$ cannot appear in a 2nd order diff-invariant equation of motion, so these components are non-dynamical. For more details, see the discussion of gravitational waves in any introductory GR textbook, which should show that in transverse-traceless gauge the linearized Einstein equation have two independent solutions (the '+' and '×' polarizations).

[4]See Weinberg QFT V1, section 2.5.

Step 3: Coefficients = scale of new physics

Coefficients should be fixed by dimensional analysis, up to $O(1)$ factors. This doesn't work for the cosmological constant: experiment (ie the fact the universe is not Planck-sized) indicates that Λ is unnaturally small. This is the cosmological constant problem. We will just sweep this under the rug, take this fine-tuning as an experimental fact, and proceed to higher order.

For these purposes let's take the coordinates to have dimensions of length, so the metric is dimensionless, and R has mass dimension 2. The action should be dimensionless (since $\hbar = 1$). Looking at the Einstein-Hilbert term, that means $[G_N] = 2 - D$, so in terms of the Planck scale,

$$\frac{1}{G_N} \equiv (M_P)^{D-2} \ . \tag{1.5}$$

In $D > 2$, this term is not renormalizable. This means that the theory is strongly coupled at the Planck scale. If we try to compute scattering amplitudes using Feynman diagrams, we would find non-sensical, non-unitary answers for $E \gtrsim M_P$. The rules of effective field theory tell us that we must include the R^2 terms, with coefficients $c_{1,2,3} \sim 1/M_P^2$. Higher curvature terms should also be included, suppressed by more powers of M_P. More generally, the rule is that these coefficients should be suppressed by the scale of new physics, which we will call M_s. New physics must appear at or below the Planck to save unitarity, so $M_s < M_P$, but it's possible that $M_s \ll M_P$. So to allow for this possibility, we set

$$c_{1,2,3} \sim \frac{1}{M_s^2} \ . \tag{1.6}$$

In string theory M_s would be the mass of excited string states:

$$M_s \sim \frac{1}{\ell_s} \ , \qquad c_{1,2,3} \sim \alpha' \tag{1.7}$$

where ℓ_s is the string length and $\alpha' = \ell_s^2$ is the string tension. In this context the R^2 and higher curvature terms in the action are called 'stringy corrections.'

Steps 4 and 5: Do quantum field theory, but be careful what's reliable

In this section to be concrete we will work in $D = 4$. To do perturbation theory (about

flat space) with the action (1.4), we set

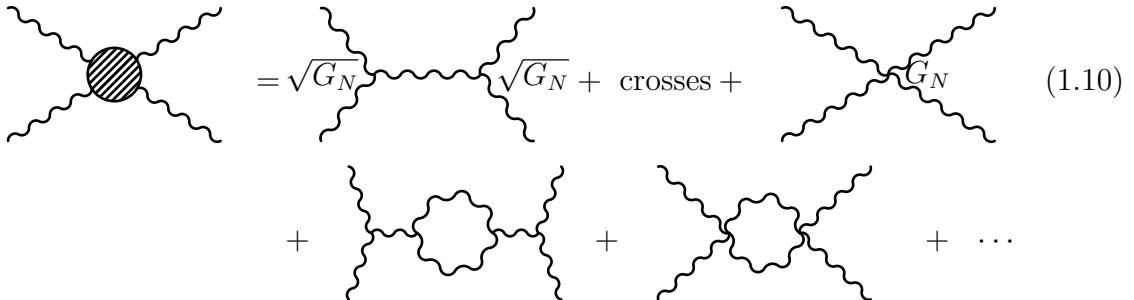$$g_{\mu\nu} = \eta_{\mu\nu} + \frac{1}{M_P} h_{\mu\nu} \tag{1.8}$$

and expand in h, or equivalently in $1/M_P$. The factor of M_P is inserted here so that the quadratic action is canonically normalized; schematically, the perturbative action looks like

$$S \sim \int \partial h \partial h + \frac{1}{M_P} h \partial h \partial h + \cdots + \frac{1}{M_s^2} \left(\partial^2 h \partial^2 h + \frac{1}{M_P} h \partial^2 h \partial^2 h + \cdots \right) \tag{1.9}$$

where the first terms come from expanding the Einstein action, and the other terms come from the higher curvature corrections.[5] In curved space, the higher curvature terms would also contribute to the terms like $h \partial h \partial h$ since $R \sim \text{const} + \partial^2 h + \cdots$.

Scattering and the strong coupling scale

As expected in a non-renormalizable theory, the perturbative expansion breaks down at high energies. First consider the case where we set $M_s = M_P$ (or, we keep only the Einstein term in the Lagrangian) and calculate the amplitude for graviton scattering in perturbation theory:

This is an expansion in the coupling constant $\sqrt{G_N} = 1/M_P$; but this is dimensionful, so it must really be an expansion in E/M_P. That is, each diagram contributes something

---

[5]The term written here comes from a term in the action $\left(\frac{M_s}{M_P} \right)^2 R^2$, where we pick off the terms $\sqrt{-g} \sim 1 + \delta g$, $R \sim \partial^2 \delta g$, then rescale $\delta g = \frac{1}{M_P} h$.

of order

$$\left(\frac{E^2}{M_P^2}\right)^{1+\text{number of loops}} . \tag{1.11}$$

So the strong coupling scale, where loop diagrams are the same size as tree diagrams, is

$$E_{strong} \sim M_P . \tag{1.12}$$

Below the strong coupling scale, this is a perfectly good quantum theory. We can use it to make reliable predictions about graviton-graviton scattering, including calculable loop corrections.

If there are two different scales M_s and M_P with $M_s \ll M_P$ the situation is slightly more complicated. The Einstein term is strongly coupled at the Planck scale, but looking at (1.9), the higher curvature terms become strongly coupled at a lower scale somewhere between M_s and M_P,

$$E_{strong} \sim M_P^x M_s^{1-x} \tag{1.13}$$

for some $x \in (0,1)$. (This is a known number that you can find by examining all the diagrams.) It is important, however, that interactions in the higher curvature terms still come with powers of $1/M_P$, so even in this case E_{strong} contains some factor of M_P (ie, $x > 0$): the theory is still weakly coupled at the scale of new physics, $E \sim M_s$.

Classical corrections to the Newtonian potential and ghosts

Returning to the classical theory, consider for example the theory with just the first term in (1.4), so $c_1 = (M_s)^{-2}$ and $c_2 = c_3 = 0$. The equations of motion derived from this action are schematically

$$\Box h + \left(\frac{1}{M_s}\right)^2 \Box\Box h = 8\pi G_N T . \tag{1.14}$$

Going to momentum space $\Box \sim E^2$, so clearly the higher curvature term is negligible at low energies $E \ll M_s$. The propagator looks like

$$\frac{1}{q^2 + M_s^{-2}q^4} = \frac{1}{q^2} - \frac{1}{q^2 + M_s^2} \tag{1.15}$$

The 2nd term looks like a massive field with mass M_s, but the wrong sign. It is a new, non-unitary degree of freedom, or 'ghost', in the classical theory. Although the only field is still the metric, it makes sense that we've added a degree of freedom because we need more than 2 functions worth of initial data to solve the 4th order equation of motion (1.14).

The 'ghost' should not bother us, because it appears at the scale M_s. This is the scale of new physics where we should not trust our effective field theory anyway. And at energies $E \ll M_s$, the ghost has no effect on classical gravity. To see this, let's compute the classical potential between two massive objects. The first term in (1.15) gives the Newtonian $1/r$ potential. The second term looks like a massive Yukawa force, so the classical potential is

$$V(r) = -G_N m_1 m_2 \left[\frac{1}{r} - \frac{e^{-rM_s}}{r} \right] . \tag{1.16}$$

This tiny for distances $r > 1/M_s$.

Loop corrections to the Newtonian potential

The 2pt function has calculable, reliable quantum corrections:

$$\sim \quad \frac{1}{q^2} - \frac{1}{M_s^2} + \frac{1}{q^2} \frac{a}{M_P^2} q^4 \log \frac{q^2}{\Lambda^2} \frac{1}{q^2} + \cdots \tag{1.17}$$

The first two terms are the classical part,

$$\frac{1}{q^2 + M_s^{-2}q^4} = \frac{1}{q^2} - \frac{1}{M_s^2} + \cdots , \tag{1.18}$$

and the log term is the loop diagram (including external legs!). a is an order 1 number that can been calculated from this diagram, and we've dropped some terms to simplify the discussion (see Donaghue for details). To calculate the attractive potential between

two stationary masses, we set the frequency to zero $q^0 = 0$ and go to position space,[6]

$$\int d^3\vec{q} \left(\frac{1}{\vec{q}^2} + \frac{1}{M_s^2} - \frac{1}{M_P^2} \log \vec{q}^2 + \cdots \right) e^{i\vec{q}\cdot\vec{x}} \sim \frac{1}{r} + \frac{1}{M_s^2}\delta(r) + \frac{1}{M_P^2 r^3} \qquad (1.19)$$

This Fourier transform is just done by dimensional analysis (we've dropped numerical coefficients). The first term is the classical Newtonian potential, the second term is the classical higher-curvature correction, and the last term is a quantum correction. The delta function does not matter at separated points; it is UV physics and does not affect the potential. It came from the same physics as the Yukawa term e^{-rM_s} in our discussion above— the difference is that the Yukawa term is the exact classical contribution whereas the delta function comes from expanding out the propagator in a derivative expansion.

The last term in (1.19) is a reliable prediction of quantum gravity, with small but non-zero effects at low energies.

Does $M_s = M_P$?

So does $M_s = M_P$, or is there a new scale $M_s \ll M_P$? This is basically the question of whether the new UV physics that fixes the problems of quantum gravity is weakly coupled ($M_s \ll M_P$) or strongly coupled (no new scale). Both options are possible, and both are realized in different corners of string theory.

If we ask the analogous question about other effective field theories that exist in nature, then sometimes the new physics is strongly coupled (for example, QCD as the UV completion of the pion Lagrangian) and sometimes it's weakly coupled (for example, electroweak theory as the UV completion of Fermi's theory of beta decay).

Breakdown

As argued above, the effective field theory breaks down at (or below) M_P. It is conceivable that this is just a problem with perturbation theory, and that the theory makes sense non-perturbatively, for example by doing the path integral on a computer. The problem is that the theory is UV-divergent and must be renormalized; this means we do not have the option of just plugging in a particular action, say just the Einstein term

---

[6]See for example Peskin and Schroeder section 4.7.

$\int \sqrt{-g}R$, and using this to define a quantum theory. We must include the full series of higher curvature terms. Each comes with a coupling constant, so we have an infinite number of tunable parameters and lose predictive power. The only way this theory can make predictions is if these infinite number of running couplings flow under RG to a UV fixed point with a finite number of parameters. This idea is called 'asymptotic safety,' and although it's a logical possibility there is little evidence for it. It is not how other effective fields theories in nature (eg pions) have been UV completed.

1.2 Quantum gravity in the Ultraviolet

We have seen that at $E \ll M_P$, ordinary methods in quantum field theory can be applied to gravity without any problems. So the real problem, of course, is how to find a UV completion that reduces at low energy to the effective field theory we've just described. This problem is unsolved, but there has been a lot of progress. The point of this section is to make a few comments about what we should and should not expect in a theory of quantum gravity, and to introduce the idea of emergent spacetime.

Quantum gravity has no local observables

Gauge symmetry is not a symmetry. It is a fake, a redundancy introduced by hand to help us keep track of massless particles in quantum field theory. All physical predictions must be gauge-independent.

In an ordinary quantum field theory without gravity, in flat spacetime, there two types of physical observables that we most often talk about are correlation functions of gauge-invariant operators $\langle O_1(x_1) \cdots O_n(x_n) \rangle$, and S-matrix elements. The correlators are obviously gauge-independent. S-matrix elements are also physical, even though electrons are not gauge invariant. The reason is that the states used to define the S-matrix have particles at infinity, and gauge transformations acting at infinity are *true* symmetries. They take one physical state to a different physical state — unlike local gauge transformations, which map a physical state to a different description of the same physical state.

In gravity, local diffeomorphisms are gauge symmetries. They are redundancies. This means that local correlation functions like $\langle O_1(x_1) \cdots O_n(x_n) \rangle$ are not gauge invariant,

and so they are not physical observables.[7] On the other hand, diffeomorphisms that reach infinity (like, say, a global translation) are physical symmetries — taking states in the Hilbert space to different states in the Hilbert space — so we get a physical observable by taking the insertion points to infinity. This defines the S-matrix, so it is sometimes said that 'The S-matrix is the only observable in quantum gravity.'

This is not quite true, since there are also non-local physical observables. For example, suppose we send in an observer from infinity, along a worldline $x_0^\mu(\tau)$, with τ the proper time along the path. Although the coordinate value $x_0^\mu(\tau)$ depends on the coordinate system, it unambiguously labels a physical point on the manifold; that is, $x_0^\mu(\tau_1)$ labels the same physical point as $x^{\mu'}(\tau_1)$ in some other coordinates. Therefore $\langle O_1(x_0(\tau_1)) \cdots O_n(x_0(\tau_n)) \rangle$ should be a physical prediction of the theory, which answers the physical question 'If I follow the path $x_0^\mu(\tau)$, carrying an O-meter, what do I measure?'. So apparently, to construct diff-invariant physical observables, we need to tie them to infinity. Although this sounds like a straightforward fix, it is actually a radical departure from ordinary, local quantum field theory.

The graviton is not composite (Weinberg-Witten theorem)
In QCD, there are quarks at high energies, and pions are composite degrees of freedom that appear at low energy where the quarks are strongly coupled. The pion Lagrangian is non-renormalizable; it breaks down at the QCD scale and must be replaced by the full UV-complete theory of QCD.

Based on this analogy, we might guess that the UV completion of gravity is an ordinary, $D = 4$ quantum field theory with no graviton, and that the graviton is a emergent degree of freedom at low energies. This is wrong. The graviton may be an emergent degree of freedom, but it cannot come from an ordinary $D = 4$ quantum field theory in the UV. The reason is the Weinberg-Witten theorem:[8]

---

[7]This is true in the effective field theory of gravity too, not just in the UV. However in perturbation theory it is not a problem. In perturbation theory, coordinates x^μ label points on the fixed background manifold, which are meaningful. It is only when we allow the geometry to fluctuate wildly that this really becomes a problem.

[8] See the original paper for the proof, it is short and clear. The basic idea is to first argue that, since states carry energy, $\langle p|T^{00}(0)|p \rangle \neq 0$. However, if these are single-particle states for a particle with helicity ± 2, then there is just now way for $\langle p'|T_{\mu\nu}|p \rangle$ to transform properly under rotations unless it vanishes.

A 4D Lorentz-invariant QFT with a conserved, gauge-invariant stress tensor $T_{\mu\nu}$ cannot have massless particles with spin > 1.

This does not rule out general relativity itself because $T_{\mu\nu}$ is not gauge invariant in GR (or, equivalent, the physical part of $T_{\mu\nu}$ is not really a Lorentz tensor). It does rule out a composite graviton: If gravity emerged from an ordinary QFT, then in the UV there is no diffeomorphism symmetry, the stress tensor is gauge invariant, so there can be no graviton in the spectrum.

Emergent spacetime

So the theory of the graviton is sick in the UV, but if we stick to ordinary QFT we cannot eliminate the graviton in the UV. This leaves two possibilities. One is that the graviton appears in the UV theory, along with other degrees of freedom which cure the problems seen in effective field theory. The other is that the graviton is an emergent degree of freedom, but the UV theory is not an ordinary 4D QFT. These are not mutually exclusive, and in fact both of these possibilities are realized in string theory (simultaneously!).

In this course we will focus on the second possibility. We will discuss models where not only the graviton, but spacetime itself is emergent. The fundamental degrees of freedom of the theory do not live in the same spacetime as the final theory, or in some cases do not live in any spacetime at all. Spacetime is an approximate, collective description of these underlying degrees of freedom, and makes sense only the infrared. The graviton is emergent, but evades Weinberg-Witten because the way it emerges is outside the usual framework of QFT.

A look ahead

There are many ways to approach this subject. In this course we will take a route that begins and ends with black holes. Unlike other EFTs (eg the pion Lagrangian), the Einstein action contains an enormous amount of information about the UV completion — infrared hints about the ultraviolet. Much of this information is encoded in the thermodynamics of black holes, so that is our starting point, and will be the basis of the first half of the course. As it turns out, black holes also lead to emergent spacetime and the AdS/CFT correspondence, which are the topics of the second half of the course.

1.3 Homework

Review the chapter on black holes in Carroll's textbook (or online lecture notes) on General Relativity.