2+1-Dimensional Gravity Revisited

Edward Witten
PiTP
July 26, 2007

Three-dimensional pure quantum gravity, with the Einstein-Hilbert action

$$I = \frac{1}{16\pi G}\int d^3x\sqrt{g}\left(R + 2/\ell^2\right)$$

has been studied from many points of view, but its status is fundamentally unclear.

I will be making a highly *tentative* effort to reconsider it. (thanks to J. Maldacena)

My main motivation for doing so is the BTZ black hole. With negative cosmological constant, there are black holes in 2+1 dimensions. Since pure gravity in 2+1 dimensions is "trivial," i.e. it has no propagating modes, we might hope to get an exact description of the quantum theory of these black holes. At least above 1+1 dimensions, where propagating modes have complicated interactions, there is no hope to get an equally precise black hole model in a theory that is not "trivial."

The first observation that we might make about this theory is that even though on dimensional grounds it appears to be unrenormalizable by power counting,

$$I = \frac{1}{16\pi G} \int d^3x \sqrt{g} \left(R + 2/\ell^2 \right)$$

it is actually finite in perturbation theory – modulo a field redefinition and a renormalization of the cosmological constant.

This is so for the same reason that the theory is “trivial.” In 2+1 dimensions, the Riemann tensor R_{ijkl} can be expressed in terms of the Ricci tensor R_{ij} which in turn, using Einstein’s equations, can be expressed as a multiple of the metric tensor g_{ij} So finally, on shell, the only possible counter-term is the volume of spacetime, that is, a renormalization of the cosmological constant.

What I have just said is valid regardless of how one formulates perturbation theory, but actually, there is a natural formulation in which no renormalization or field redefinition is needed.

This comes from the fact that classically, 2+1-dimensional gravity can be expressed in terms of gauge theory. The spin connection ω is an $SO(2,1)$ gauge field. ω can be combined with the “vierbein” e to make a gauge field of gauge group $SO(2,2)$ if the cosmological constant is negative (and a similar group otherwise). We just combine ω and e into a 4 x 4 matrix

$$A = \begin{pmatrix} \omega & e \\ -e & 0 \end{pmatrix}$$

As long as the vierbein is invertible, the usual transformations under local Lorentz transformations and diffeomorphisms combine together as gauge transformations of A. This statement actually has an analog in any dimension. What is special in 2+1 dimensions is that the Einstein-Hilbert action can be expressed in a gauge-invariant form

$$I = \frac{k}{4\pi}\int \mathrm{Tr}\,\left(A\wedge dA + \frac{2}{3}A\wedge A\wedge A\right)$$

as a Chern-Simons interaction.

(Achucarro and Townsend 1987, EW 1988)

From this point of view, perturbation theory is renormalizable by power counting, and is actually finite, since there are no possible local counterterms (and the cosmological constant is a structure constant of the gauge group).

It is pretty clear that the gauge theory description of gravity is valid in perturbation theory – since it is valid classically – and perturbation theory will not take us out of the classical region of invertible vierbein.

But nonperturbatively there is a real question, as in gauge theory we will have to allow a non-invertible vierbein.

My own view in 1988 was that the gauge theory description was correct nonperturbatively, and one had to allow a degenerate vierbein to make sense of the quantum theory.

This view was criticized fairly convincingly, especially by N. Seiberg, who argued that in 0+1 and 1+1 dimensions, where we do know how to make sense of quantum gravity, we take seriously the invertibility of the vierbein at the quantum level.

There is, however, another problem with the idea “gauge theory=gravity in 2+1 dimensions” that seems even more acute.

A few years after the gauge theory interpretation of 2+1-dimensional gravity was proposed, it was appreciated that for the case of negative cosmological constant, there is a black hole in this “trivial” theory

(Banados, Teitelboim, Zanelli, 1992)

and developments in the AdS/CFT correspondence (beginning with Strominger 1997) made it clear within a few years that this should be taken seriously.

The BTZ black hole has a horizon of positive circumference and a corresponding Bekenstein-Hawking entropy. If, therefore, pure 2+1-dimensional gravity does correspond to a quantum theory, this theory ought to have a huge degeneracy of black hole states -- which we are not going to be able to get in a reasonable way from topological field theory (though some attempts have been made).

Before going on, let us discuss what we are going to aim for in trying to solve 2+1-dimensional gravity.

First of all, I am only going to consider the case of negative cosmological constant.

Currently there is some suspicion that quantum gravity with $\Lambda > 0$ doesn't exist nonperturbatively (in any dimension) with positive cosmological constant. One reason for this is that it does not appear to be possible, with $\Lambda > 0$, to define precise observables. This is natural if it is the case that a world with positive cosmological constant (like the one we may be living in) is always unstable.

If that is so, then a world with $\Lambda > 0$ doesn't really make sense as an exact theory in its own right but (like an unstable particle) must be studied as part of a larger system.

Whether that is the right interpretation or not, since I do not know how to define any precise observables, I don't know what it would mean to try to solve 2+1-dimensional gravity with $\Lambda > 0$, since it isn't clear what we'd want to compute.

With zero cosmological constant, above 2+1 dimensions, there is a meaningful observable: the S-matrix.

In 2+1 dimensions without matter fields, we have no local propagating fields, and also no black holes if the cosmological constant is zero. So there is no S-matrix, and again, there is no clear picture of what one wants to calculate in trying to solve the theory.

With negative cosmological constant, there is an analog of the S-matrix, namely the dual conformal field theory. It captures the asymptotic information that is analogous to the S-matrix in the $\Lambda = 0$ case. Not only does this make sense in 2+1 dimensions, but in fact one of the precursors of the AdS/CFT correspondence was the work of Brown and Henneaux (1986) on 2+1-dimensional gravity.

Brown and Henneaux showed that the Hilbert space of 2+1-dimensional gravity, with

$$I = \frac{1}{16\pi G}\int d^3x \sqrt{g}\left(R + 2/\ell^2 + \ldots\right)$$

-- the ellipses refer to the fact that this result isn't affected by matter fields – has an action of a left- and right- moving pair of Virasoro algebras with $c_L = c_R = 3\ell/2G$

In our modern understanding, this is part of a much richer structure – the dual CFT

What it means to solve quantum gravity with $\Lambda < 0$ is to find the dual CFT.

And to repeat, we focus on the case $\Lambda < 0$ because that is the only case in which we know what it would mean to solve the theory.

This formulation makes obvious a statement that from a classical point of view looks rather surprising. When we look at the classical action,

$$I = \frac{1}{16\pi G} \int d^3x \sqrt{g} \left(R + 2/\ell^2 \right)$$

it appears that ℓ, which is the cosmological constant in Planck units, is a free parameter. But the formula for the central charge $c_L = c_R = 3\ell/2G$ shows that this cannot be the case.

According to the Zamolodchikov c-theorem, in any continuously variable family of conformal field theories in 1+1 dimensions, the central charge c is a constant.

Hence it cannot depend on a variable parameter such as ℓ/G . It must be that 2+1-dimensional quantum gravity makes sense *at most* only for certain values of ℓ/G .

Of course, Zamolodchikov's theorem has an important technical assumption – the theory must have a normalizable and SL(2)-invariant ground state. This assumption is valid in 2+1-dimensional gravity, with Anti de Sitter space being the classical approximation to the relevant quantum state.

(This assumption in Zamolodchikov's theorem has been overlooked in some claims about 2+1-dimensional gravity.)

I should remark that the statement that ℓ/G cannot be continuously varied is not limited to pure gravity – it holds for the same reason in any theory of 2+1-dimensional gravity plus matter that has a sensible Anti de Sitter vacuum. For example, in the string theory models whose CFT duals are known, ℓ/G is expressed in terms of integer-valued fluxes, which gives a direct explanation of why it cannot be varied.

So we are only going to aim to solve the theory for negative cosmological constant, and even then, only for certain values of ℓ/G But what are the right values?

I don't have any rigorous way to determine this. But there is a simple picture that gives us a plausible heuristic way to try to find the right values, and it turns out that this gives interesting values.

We are just going to take at face value the gauge theory description of 2+1-dimensional gravity. First of all, in addition to ℓ/G, there is really a second dimensionless parameter, since one can add to the action a multiple of the Chern-Simons invariant of the spin connection:

$$I' = \frac{k}{4\pi}\int \mathrm{Tr}\,\left(\omega d\omega + \frac{2}{3}\omega^3\right)$$

Here k is an integer for topological reasons.

So the theory really depends on two parameters, namely ℓ/G and k.

We understand why the second has to be an integer, but we want to know why the first one cannot vary continuously (which would contradict the c-theorem).

If we just take at face value the gauge theory interpretation of 2+1-dimensional gravity, it gives an explanation of why the parameters only take special values, and the values it selects prove to be interesting.

We use the fact that the gauge group $SO(2,2)$ is essentially the same as $SO(2,1) \times SO(2,1)$

The two $SO(2,1)$ gauge fields are $A_\pm = \omega \pm e$ and the action is a sum

$$I = \frac{k_L}{4\pi} \int \mathrm{Tr} \left(A_- dA_- + \frac{2}{3} A_-^3 \right) + \frac{k_R}{4\pi} \int \mathrm{Tr} \left(A_+ dA_+ + \frac{2}{3} A_+^3 \right)$$

If we take the gauge theory description literally, then both k_L and k_R are quantized to integer values for topological reasons. Moreover, from the formulas of Brown and Henneaux, combined with the gravity/Chern-Simons relation, the central charges turn out to be

$$(c_L, c_R) = 24(k_L, k_R)$$

This is an interesting result, because holomorphic factorization is possible in 2d CFT precisely for these values of c.

A further hint of holomorphic factorization is simply the fact that the Chern-Simons action is the sum of a “left” part and a “right” part.

So in continuing, we are going to assume holomorphic factorization, and we will just try to describe the holomorphic part of the theory – a holomorphic CFT with central charge $c = 24k$ for some integer k

Now one simple fact is that the ground state energy of such a theory is $-c/24 = -k$

What other states are there? Naively, none at all, since 2+1-dimensional gravity is "trivial." However, this isn't right. At least there are the boundary excitations that lead to the Virasoro algebra of Brown and Henneaux.

If there were no other excitations, the partition function in genus 1 would be

$$Z(q) = q^{-k} \prod_{n=2}^{\infty} \frac{1}{1-q^n}$$

This function counts the excitations of the vacuum that can be made by acting repeatedly with the stress tensor and its derivatives.

But that cannot be the full answer, since this function is not modular invariant.

There must be some additional primaries, apart from the identity. The partition function will then be

$$Z(q) = q^{-k}\left(\Pi_{n=2}^{\infty}\frac{1}{1-q^n} + \mathcal{O}(q^l)\right)$$

for some l .

Now we are going to interpret the fact that the theory is classically “trivial” to mean that we should make l as large as possible.

It turns out that the largest that l can be is $l = k + 1$, and if this is the right value, then the partition function $Z(q)$ is uniquely determined.

My proposal is that this gives the partition function of the dual CFT, including the black holes.

Riemann surfaces of genus 1 are parametrized by the "j-function"

$$j(q) = q^{-1} + 744 + 196884q + \ldots$$

It actually is more convenient to use

$$J = j - 744 = q^{-1} + 196884q + \ldots$$

The fact that the partition function is modular invariant and has its only pole at $q = 0$ means that it is a polynomial in $J(q)$

More specifically, $Z(q)$ is a polynomial in J of order k, since its pole at $q = 0$ is of order k. So we have

$$Z(q) = \sum_{r=0}^{k} a_r J(q)^r$$

with some coefficients a_r We can pick these coefficients to make $Z(q)$ agree with the naïve function $q^{-k} \prod_{n=2}^{\infty} \frac{1}{1-q^n}$

up to order q^0 .

But then we have no control over the term of order q^1. This is above the ground state energy by k+1 units, so it means that there will be primary fields of dimension

$$h = k + 1$$

We interpret them as black holes, since the theory is “trivial” except for the black holes.

In fact, something nice happens. The minimum classical black hole mass, in these units, is $M = k$. But the Bekenstein-Hawking entropy vanishes if the mass is precisely k . So black holes of positive entropy exist precisely if the mass is greater than k , in perfect accord with the fact that, according to our proposal for the quantum theory, the lowest dimension of a primary (other than the identity) is $k + 1$

Moreover, the partition function that I’ve described how to calculate, though I didn’t quite write down an explicit formula, gives a result for the black hole degeneracies that agrees perfectly with the Bekenstein-Hawking entropy. From the way I’ve explained things, this may sound like a miracle, but it will seem less surprising if one is familiar with the “Farey tale” of

Dijkgraaf, Maldacena, Moore, and Verlinde (1997)

- Let us give an example. If $k = 1$, the partition function is simply the J-function itself, so

$$Z(q) = q^{-1} + 196884q + \dots$$

- The number of black hole primaries of mass 2 is therefore 196883. The black hole entropy is therefore log(196883)=12.19…
- The classical entropy of a black hole with k=1 and mass 2 is 4π=12.57... So we are off by just a few percent.

This is the worst case. If we increase k or the black hole mass, the semi-classical approximation to the black hole entropy quickly improves and it becomes asymptotically exact for large k.

But let us go back to k=1.

The number 196883 of black hole primaries at this low mass is a very special number.

Back in the 1970’s, the last of the sporadic (or exceptional) finite simple groups was constructed – the Fischer-Griess “monster” or “friendly giant.” The lowest dimension of an irreducible representation of the monster is 196883. John McKay noticed that this is very close to the coefficient of the third term in

$$j(q) = q^{-1} + 744 + 196884q + \ldots$$

After higher coefficients of the j-function were similarly found to be related to the monster, Frenkel, Lepowsky, and Meurman (1985) were led to conjecture the existence of, and to construct, a holomorphic conformal field theory with partition function our friend $J(q)$

They also conjectured that their theory is the *unique* holomorphic CFT with this partition function.

If so – and the conjecture hasn't been entirely proved – then the Frenkel-Lepowsky-Meurman monster theory must be the dual CFT at k=1.

Thus, we can interpret the monster group as the symmetry of 2+1-dimensional black holes, at least at this value of k.

What about higher values of k? It is not hard to show that the coefficients in the q-expansion of the partition function can always be expressed in terms of dimensions of monster representations. So it is conceivable that the monster is a symmetry of 2+1-dimensional quantum gravity at every value of k. But I don’t know how to show that this is true, or even to show that the CFT’s with the claimed partition functions do exist.

There is a close analog of all this in the supersymmetric case. The allowed values of the central charge (in the holomorphic sector) are now $c=12\,k$, where again k should be an integer. The partition function can be uniquely determined for every k, by imitating the reasoning we used in the bosonic case. It turns out that candidates for the superconformal field theories of interest exist in the superconformal case for $k=1$ and 2.

The candidate for k=1 was constructed by Frenkel, Lepowsky, and Meurman (1985) who also conjectured its uniqueness, and understood more fully by Duncan (2005), while the candidate at k=2 was constructed by Dixon, Ginsparg, and Harvey (1988).

The k=1 and k=2 theories have in common a discrete symmetry group that is smaller than the monster but is again one of the sporadic finite groups – the “Conway group,” in this case.