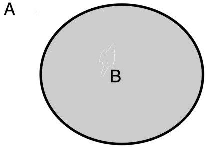
An Algebra of Observables for de Sitter Space

Edward Witten, IAS

I will be talking about a recent paper arXiv:2206.10780 that has the same title as the talk, with V. Chandrasekharan, G. Penington, and R. Longo, as well as arXiv:1209.10454 with Chandrasekharan and Penington as well as an earlier paper "Gravity and the Crossed Product" arXiv.2112.12828. The original inspiration came from two papers by Liu and Leutheusser arXiv:2110.05497 and 2112.12156.

The idea that the Bekenstein-Hawking entropy of a black hole should be understood in terms of entanglement entropy was apparently first put forward by R. Sorkin in 1983 (in a paper that attracted only modest attention at the time). The idea was just the following. In a quantum field theory, divide space into two regions A and B



Let Ψ be a state of the system, and ρ_A the "reduced density matrix" of the state Ψ for measurements in region A. One can try to calculate the von Neumann entropy $S(\rho) = -\mathrm{Tr}\,\rho_A \log \rho_A$ of this density matrix. One finds that it is ultraviolet divergent (regardless of Ψ) and the coefficient of the leading divergence is proportional to the area A of the boundary between regions A and B. Sorkin hoped that gravity would somehow cut off this divergence and lead to the Bekenstein-Hawking result A$/4G$.

Susskind and Uglum (1993) advanced the subject with a simple observation. If in the Bekenstein formula for the generalized entropy

$$S_{\rm gen} = \frac{\mathsf{A}}{4G\hbar} + S_{\rm out}$$

we interpret $S_{\rm out}$ as von Neumann entropy of the density matrix, then $S_{\rm gen}$ is better defined that either term is separately. The second term has an ultraviolet divergence that Sorkin had noted. The first term has a similar problem, because there is an ultraviolet divergence in the relation between the bare Newton constant G_0 and the physical, observed Newton constant G:

$$\frac{1}{G\hbar} = \frac{1}{G_0\hbar} + c\Lambda^2 + \cdots .$$

Here Λ is an ultraviolet cutoff and c is a constant (at 1-loop level, c is independent of $\hbar$). Susskind and Uglum argued that the ultraviolet divergences in $S_{\rm out}$ cancel those in $1/G$ (and these arguments were refined later). So $S_{\rm gen}$ is better-defined than either of the two terms on the right hand side of the formula.

At this level, we could be talking about either a black hole horizon or a cosmological horizon.

In this talk, I will give a slightly abstract explanation of "why" entropy is better defined when gravity is included, at least in the case of the black hole and de Sitter space.

First of all, in ordinary quantum mechanics, when one considers the entanglement between two systems A and B, one normally assumes at the start that each system has its own Hilbert space $\mathcal{H}_A$ or $\mathcal{H}_B$. The combined system then has a tensor product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. A state ψ_{AB} in this combined Hilbert space might be a simple tensor product of states ψ_A and ψ_B:

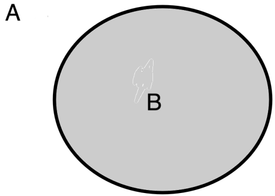
$$\psi_{AB} = \psi_A \otimes \psi_B.$$

In that case, systems A and B can separately be described by pure states ψ_A and ψ_B, and there is no entanglement entropy. But more generally we may have

$$\psi_{AB} = \sum_i \sqrt{p_i}\psi_A^i \otimes \psi_B^i,$$

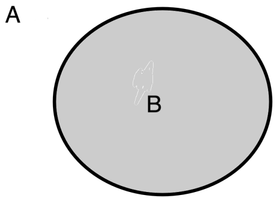
in which case we say that systems A and B are "entangled" and system A (or B) has a nonzero von Neumann entropy.

The point is that in ordinary quantum mechanics, whether or not a state has a nonzero entanglement and entanglement entropy is a property of the state. That is not so for entanglement entropy between different regions in quantum field theory.



The divergence found by Sorkin was an ultraviolet divergence, so it does not depend on the state: every state looks like the vacuum at short distances.

The root of the problem is that it is not true



that there are separate Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ for the "inside" and "outside" regions. There is only a combined Hilbert space $\mathcal{H}$ for the whole system. What the separate regions A and B have are not Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, but only algebras of observables $\mathcal{A}$ and $\mathcal{B}$. These algebras act on $\mathcal{H}$ so they can be defined to be von Neumann algebras (a von Neumann algebra is an algebra of bounded operators on a Hilbert space that is closed under a certain type of limiting operation).

There are three types of von Neumann algebra:

(I) A Type I algebra is the algebra of all operators on a Hilbert space. In ordinary quantum mechanics, when we discuss a system A, it has a Hilbert space $\mathcal{H}_A$ and the algebra of observables of the system is the algebra of all (possibly bounded or self-adjoint) operators on $\mathcal{H}_A$. This algebra is of Type I. If a system is described by a Type I algebra $\mathcal{A}$, then the system can have quantum mechanical pure states – namely states in the Hilbert space on which $\mathcal{A}$ is the algebra of observables. One can also define density matrices and entropies for a system that has such an algebra of observables.

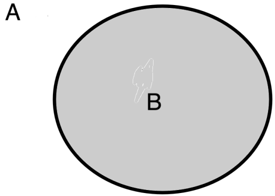
The other types are less familiar. But first the bottom line:

(II) A Type II algebra does not have pure states, but there is a notion of density matrix and entropy for a system in which the algebra of observables is of Type II.

(iii) A Type III algebra is the "worst" type – a system whose observables form a Type III algebra does not have pure states and also does not have density matrices or entropies.

By now you might anticipate the bad news:

In quantum field theory, the algebra of observables of a region of spacetime



is always of Type III. So to a region, one can never associate a pure state, or a density matrix or entropy. The Type III nature of the algebra is the "reason" for the universal ultraviolet divergence of the entanglement entropy.

However, it turns out that including gravity in a semiclassical way changes the picture: at least in the case of the black hole or de Sitter space, including gravity at a semiclassical level changes the algebra of the region outside the horizon from Type III to Type II. So when gravity is turned on semiclassically, the region outside the black hole or de Sitter horizon is described by an algebra in which the notion of entropy is well-defined, though there is no notion of a quantum mechanical microstate. We get a Type II_1 algebra for de Sitter space, and a Type II_∞ algebra for the black hole.

Roughly, Type II and Type III algebras are the natural operator algebras that act on a system A that has an infinite amount of entanglement with another system B. That is the situation in quantum field theory for operators in a given region of spacetime; it is also the situation in quantum statistical mechanics if we consider an infinite system at positive temperature. Without gravity, those problems lead to Type III algebras. A Type II algebra is more special.

A Type II_1 algebra is most simply described as the algebra that acts on an infinite collection of qubits that are in an almost maximally mixed state. Consider a system A of N qubits that is maximally entangled with a second system B also consisting of N qubits:

$$\Psi = \frac{1}{2^{N/2}} \bigotimes_{n=1}^{N} \left(\sum_{i=1,2} |i\rangle_{A,n} \otimes |i\rangle_{B,n} \right).$$

Let a, a' be operators that act only on the first k spins of system A, for some $k \leq N$. Define a function

$$F(\mathrm{a}) = \langle \Psi | \mathrm{a} | \Psi \rangle.$$

Since the density matrix of system A is $\rho = 2^{-N}\mathrm{Id}$, we have

$$F(\mathsf{a}) = \mathrm{Tr}\,\rho\mathsf{a} = 2^{-N}\mathrm{Tr}\,\mathsf{a}$$

and hence

$$F(\mathsf{a}\mathsf{a}') = F(\mathsf{a}'\mathsf{a}) = 2^{-N}\mathrm{Tr}\,\mathsf{a}\mathsf{a}'.$$

Also

$$F(1) = 1.$$

And the function $F(\mathsf{a})$ has a thermodynamic limit because it is unchanged if we add more maximally entangled spins to the system (with the given operator a not acting on the added spins).

For $N \to \infty$, the function $F(\mathsf{a})$ can be defined for any operator a that acts on any finite set of qubits in system A and of course it still satisfies

$$F(1) = 1$$

and

$$F(\mathsf{a}\mathsf{a}') = F(\mathsf{a}'\mathsf{a}).$$

So far we have defined F on the whole algebra $\mathcal{A}_0$ of all operators that act on only finitely many qubits in system A. By taking the "closure" of $\mathcal{A}_0$ in von Neumann's sense (this means we allow certain operators that act on any number k of qubits, but with matrix elements that decay rapidly for large k) we can complete $\mathcal{A}_0$ to a von Neumann algebra $\mathcal{A}$, still with a function $F(\mathsf{a})$ that has the same properties I've stated. Since $F(\mathsf{a}\mathsf{a}') = F(\mathsf{a}'\mathsf{a})$ this function is usually called a trace: We formally define

$$F(\mathsf{a}) = \mathrm{Tr}\,\mathsf{a}$$

but $\mathrm{Tr}\,\mathsf{a}$ is *not* the trace of a in any Hilbert space representation. It is more like a renormalized trace in which we removed an infinite factor $2^N\big|_{N\to\infty}$. Note that

$$\mathrm{Tr}\,1 = 1.$$

There is a more elementary example of an infinite dimensional algebra with a trace – the Type I algebra $\mathcal{B}$ of all operators on an infinite-dimensional Hilbert space $\mathcal{H}$. In this example, however, while we can define a trace on elements of $\mathcal{B}$, it is not defined for all elements of $\mathcal{B}$, only for those that are "trace class." For example, the identity element of $\mathcal{B}$ does not have a trace (unless one wants to allow $\mathrm{Tr}\,1 = \infty$). By contrast, from the infinite system of qubits, we constructed an algebra $\mathcal{A}$ in which every element has a trace. Clearly then it is an essentially new type of algebra. This is, in fact, the simplest example of a Type II algebra – it is said to be of Type II_1. (It is called the Type II_1 factor of Murray and von Neumann.)

If $\mathcal{A}$ is the Type II$_1$ algebra that we just constructed, and $\mathcal{B}$ is the Type I algebra of all operators on an infinite-dimensional Hilbert space, then we can make a third type of algebra by simply taking their tensor product:

$$\mathcal{C} = \mathcal{A} \otimes \mathcal{B}.$$

This new algebra $\mathcal{C}$ still has a trace (since each factor does) but it is not defined for all elements (because of the factor $\mathcal{B}$). In fact, $\mathcal{C}$ is a new kind of algebra, said to be of Type II$_\infty$. It turns out that $\mathcal{C}$ is related to the black hole and $\mathcal{A}$ is related to de Sitter space.

To construct an algebra of Type III, we make a similar construction, starting with a state that is “fully” but not maximally entangled:

$$\Psi = \otimes_{k=1}^{N} \frac{1}{(1+e^{-\beta/2})^{1/2}} \left(|\uparrow\rangle_{A,k} |\uparrow\rangle_{B,k} + e^{-\beta/2} |\downarrow\rangle_{A,k} |\downarrow\rangle_{B,k} \right).$$

We can still define the function $F(\mathrm{a}) = \langle\Psi|\mathrm{a}|\Psi\rangle$ and as before it has a thermodynamic limit. The important difference is that now $F(\mathrm{a}\mathrm{a}') \neq F(\mathrm{a}'\mathrm{a})$. For $N \to \infty$, we can define an algebra $\mathcal{A}_0$ consisting of operators that act on any finite set of qubits of the A system, and its completion is now a von Neumann algebra of Type III. In the infinite volume limit at temperature $T > 0$, the algebra of observables of any quantum thermal system is of Type III.

Algebras of Type II or Type III do not have an irreducible representation in a Hilbert space; whenever such an algebra acts on a Hilbert space $\mathcal{H}$, it always commutes with another algebra of the same type. For example, we constructed our Type II and Type III algebras as algebras of operators on the "A" part of a bipartite system AB, so in that construction they commute with an identical algebra that acts on system B.

The difference between a Type II algebra and a Type III algebra is that a Type II algebra has a trace, and a Type III algebra does not.

Moreover, in a Type II algebra, the trace is nondegenerate in the sense that if $F(\mathsf{a})$ is any linear function of $\mathsf{a} \in \mathcal{A}$, we have

$$F(\mathsf{a}) = \mathrm{Tr}\, \mathsf{a}\mathsf{a}'$$

for some unique $\mathsf{a}' \in \mathcal{A}$. In particular if $\mathcal{A}$ acts on a Hilbert space $\mathcal{H}$, and Ψ is a state in $\mathcal{H}$, we can consider the linear function $\mathsf{a} \to \langle \Psi | \mathsf{a} | \Psi \rangle$. It will be $\mathrm{Tr}\, \rho \mathsf{a}$ for some "density matrix" $\rho \in \mathcal{A}$:

$$\langle \Psi | \mathsf{a} | \Psi \rangle = \mathrm{Tr}\, \rho \mathsf{a}.$$

Thus a state of a Type II algebra has a density matrix.

Once we have density matrices, we can also define entropies;

$$S(\rho) = -\mathrm{Tr}\,\rho \log \rho.$$

So a state of a Type II algebra has an entropy.

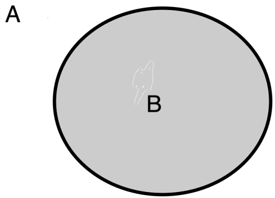
However, in physical terms, the entropy of a state of a Type II algebra is a sort of renormalized entropy from which an infinite constant has been subtracted. For example, let us go back to the system A of N qubits maximally entangled with another such system B. The A system has entropy N, infinite in the large N limit. Suppose instead we disentangle k of the N qubits (where we will keep k fixed as $N \to \infty$). The entropy is now $N - k$. Entropy of a Type II_1 algebra is defined by subtracting N before taking $N \to \infty$. So the maximally mixed state has entropy 0, and the state with k qubits disentangled has entropy $-k$.

More formally, we defined the trace by $\mathrm{Tr}\,\mathsf{a} = \langle\Psi|\mathsf{a}|\Psi\rangle$, where Ψ is the maximally mixed state, so the maximally mixed state has density matrix $\rho = 1$ (this is indeed a density matrix since $\mathrm{Tr}\,1 = 1$). So the von Neumann entropy of the maximally mixed state is

$$S(\rho) = -\mathrm{Tr}\,1\log 1 = 0,$$

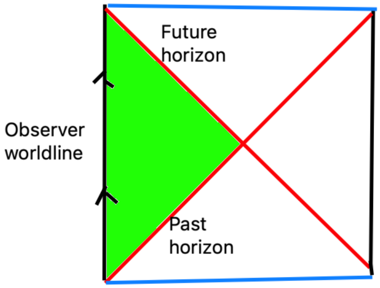
and it is not hard to prove that any other density matrix has strictly negative entropy.

As I have already explained, in ordinary quantum field theory the algebras



are Type III. But it turns out (at least for the black hole and de Sitter space) that when we include gravity, things are different: gravitational effects even for very weak coupling convert the Type III algebras into Type II algebras. This can be viewed as an abstract explanation of why entropy is better defined in the presence of gravity. The details are somewhat different in the two cases and I will begin with de Sitter space.

Here is the setup:



The green region is called a "static patch." There is a Killing vector field of "time translations" that is future directed timelike in the static patch (it is past directed timelike at regions spacelike separated from the static patch). Let H be the generator of time translations.

In ordinary quantum field theory in de Sitter space (and also in the presence of semiclassical gravity) there is a natural de Sitter state $\Psi_{\rm dS}$ which can be obtained by analytic continuation from Euclidean signature. Correlation functions in the state $\Psi_{\rm dS}$ have a thermal interpretation at the de Sitter temperature $T_{\rm dS} = 1/\beta_{\rm dS}$, where $\beta_{\rm dS} = 2\pi r_{\rm dS}$ ($r_{\rm dS}$ is the de Sitter radius). A slightly abstract way to describe this thermal interpretation is to say that the "modular Hamiltonian" of the state $\Psi_{\rm dS}$ is

$$H_{\rm mod} = \beta_{\rm dS} H.$$

In ordinary quantum field theory, we would associate to the static patch a Type III algebra of observables. Including weakly coupled gravitational fluctuations does not qualitatively change the picture, but what does really change the picture is that in a closed universe, such as de Sitter space, the isometries have to be treated as constraints. This means that we should replace $\mathcal{A}_0$ by $\mathcal{A}_0^H$, its invariant subalgebra. But that does not work: the invariant subalgebra is trivial. Basically, anything that commutes with H can be averaged over all the thermal fluctuations and replaced by its thermal average, a c-number.

To get a reasonable algebra of observables, we include an observer in the analysis. Of course, in principle an observer should really be described by the theory, not injected from outside. What it really means to include an observer is that we consider a "code subspace" of states in which an observer is present in the static patch, and then we consider operators that can be defined in the low energy effective field theory in this code subspace, though they are not well-defined on the whole Hilbert space.

Should we be surprised that we need to include the observer in the analysis to get a sensible answer? In ordinary quantum mechanics without gravity, one can consider the observer who studies a quantum system to be external to the system. WIth gravity included, the observer inevitably gravitates and cannot truly be considered external to the system. However, in an open universe – for example one that is asymptotically flat – the gravity of the observer can be neglibile. It is in a closed universe that it may be impossible to ignore the gravity of the observer. That is exactly the situation that we are in here because de Sitter space is a simple model of a closed universe, that is, a universe with compact spatial sections. And indeed we find that to get a sensible result we need to take into account the gravity of the observer.

As a minimal model of the observer, we consider a clock with Hamiltonian

$$H_{\rm obs} = q.$$

It is physically reasonable to assume that the observer's energy is bounded below by 0, so we assume $q \geq 0$. Thus the effect of including the observer is to modify the Hilbert space by

$$\mathcal{H}_0 \to \mathcal{H}_0 \otimes L^2(\mathbb{R}_+).$$

(Positive half-line since $q \geq 0$.) The algebra is likewise extended from $\mathcal{A}_0$ to

$$\mathcal{A}_1 = \mathcal{A}_0 \otimes B(L^2(\mathbb{R}_+)).$$

The last factor is the Type I algebra of all bounded operators on $L^2(\mathbb{R}_+)$); it is generated by q and by $p = -\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}q}$.

Finally the constraint becomes the total Hamiltonian of the quantum fields plus the observer:

$$H \to \widehat{H} = H + H_{\mathrm{obs}}.$$

The "correct" algebra of observables taking account of the presence of the observer is therefore

$$\mathcal{A} = \mathcal{A}_1^{\widehat{H}},$$

that is, the $\widehat{H}$-invariant part of $\mathcal{A}_1$.

Once an observer is present, we can “gravitationally dress” any operator to the observer’s world-line. For any $\mathsf{a} \in \mathcal{A}_0$, the operator

$$\widehat{\mathsf{a}} = e^{\mathrm{i}pH} \mathsf{a} e^{-\mathrm{i}pH}$$

commutes with the constraint $\widehat{H} = H + q$. One more operator that commutes with the constraint is q itself (or equivalently $-H$). It follows from classic results of Connes and Takesaki from the 1970’s that (1) there are no more operators that commute with the constraint, and (2) the algebra $\mathcal{A}$ that is generated by $\widehat{\mathsf{a}},\ \ \mathsf{a} \in \mathcal{A}_0$ along with q is of Type II.

To be more precise, the algebra $\mathcal{A}$ we get this way is of Type II_∞ if we do not stipulate that the observer energy q is bounded below. If we do impose this condition, we get an algebra of Type II_1. It is believed that de Sitter space in the presence of gravity has a state of maximum entropy – the state Ψ_{dS} that can be defined by analytic continuation from Euclidean signature and represents "empty de Sitter space." So to get a model of de Sitter space, it is important to assume that the observer energy is bounded below.

Once we get a Type II_1 algebra, there is going to be a state of maximum entropy, with density matrix $\rho = 1$. It is not difficult to identify this state:

$$\Psi_{\max} = \Psi_{\mathrm{dS}} \sqrt{\beta_{\mathrm{dS}}} e^{-\beta_{\mathrm{dS}} q/2}.$$

In other words, the state of maximum entropy is the state Ψ_{dS} that represents empty de Sitter space, tensored with a thermal state of the observer at the de Sitter temperature $T_{\mathrm{dS}} = 1/\beta_{\mathrm{dS}}$.

We can draw a few easy conclusions, which harmonize with claims made in the past by others (such as Banks; Susskind; Dong, Silverstein, and Torroba). First of all, since the maximum entropy state has $\rho = 1$, it has a "flat entanglement spectrum" (all eigenvalues of the density matrix are equal) and accordingly the Rényi entropies are constant:

$$S_\alpha(\rho) = \frac{1}{1-\alpha} \log \mathrm{Tr}\, \rho^\alpha = 0.$$

Given the assertion that de Sitter space has a state of maximum entropy, this is what one should expect: In ordinary quantum mechanics, the maximum entropy state of a system is "maximally mixed," with a "flat entanglement spectrum" (the density matrix is a multiple of the identity and all its eigenvalues are equal) and its Rényi entropies are independent of α.

Now, suppose that the observer makes a measurement with two outcomes that correspond to the projection operators Π and $1-\Pi$. The probabilities of the two outcomes are $\mathrm{Tr}\,\Pi$ and $\mathrm{Tr}\,(1-\Pi) = 1 - \mathrm{Tr}\,\Pi$. All values $0 \leq \mathrm{Tr}\,\Pi \leq 1$ are possible. If the outcome corresponding to Π is observed, then after this measurement, the density matrix is

$$\sigma = \frac{1}{\mathrm{Tr}\,\Pi}\Pi.$$

Since the two eigenvalues of σ are 0 and $1/\mathrm{Tr}\,\Pi$, one has $\sigma \log \sigma = \sigma \log(1/\mathrm{Tr}\,\Pi)$ so the entropy after the observation is

$$S(\sigma) = -\mathrm{Tr}\,\sigma \log \sigma = -\log(1/\mathrm{Tr}\,\Pi).$$

The entropy reduction from knowing the outcome is therefore $\Delta S = \log(1/\mathrm{Tr}\,\Pi)$, and this is related to the probability $p = \mathrm{Tr}\,\Pi$ of the given outcome by

$$p = e^{-\Delta S}.$$

However, the probability of a (low entropy) energy E fluctuation of the static patch is

$$p = e^{-\beta_{\mathrm{dS}} E},$$

according to the thermal interpretation of de Sitter space. Since also $p = e^{-\Delta S}$, we must have for consistency of the two descriptions

$$e^{-\beta_{\mathrm{dS}} E} = e^{-\Delta S}.$$

In other words, “‘thermal” suppression of a fluctuation can be understood as purely entropic suppression. This is surprising, but it has been argued before on other grounds, notably by considering the case that the “fluctuation” is a small black hole at the center of the static patch.

Which part of this is surprising? The formula $p = e^{-\Delta S}$ for the probability of an outcome is an inevitable consequence of having a maximum entropy state in which all states are equally probable. In other words, if all states are equally likely, then the probability of a given outcome is just proportional to the number of microstates that are compatible with that outcome. Here I am using language appropriate for an ordinary quantum system with a finite-dimensional Hilbert space. A moment ago, I explained how to reach the same conclusion in the context of a Type II_1 algebra.

The surprise is not that $p = e^{-\Delta S}$, which one should expect for a maximum entropy state, but that the maximum entropy state also has a thermal interpretation. Let us discuss how to see this in the context of the Type II_1 algebra. First of all, ignoring the constraint for the moment, the time dependence of an operator $\mathsf{a} \in \mathcal{A}_0$ is defined in the usual way by

$$\mathsf{a}(t) = e^{\mathrm{i}Ht}\mathsf{a}e^{-\mathrm{i}Ht}.$$

Then time-dependent correlations such as

$$\langle\Psi_{\mathrm{dS}}|\mathsf{a}(t_1)\mathsf{a}'(t_2)|\Psi_{\mathrm{dS}}\rangle$$

have thermal properties that reflect the fact that these correlation functions can be computed by analytic continuation from Euclidean signature.

After imposing the constraint, we replace a with the dressed version $\widehat{\mathsf{a}} = e^{\mathrm{i}pH}\mathsf{a}e^{-\mathrm{i}pH}$, and again we define its time dependence by

$$\widehat{\mathsf{a}}(t) = e^{\mathrm{i}Ht}\widehat{\mathsf{a}}e^{-\mathrm{i}Ht}.$$

Then, because

$$H\Psi_{\mathrm{dS}} = 0,$$

we rather trivially find

$$\langle\Psi_{\mathrm{max}}|\widehat{\mathsf{a}}(t_1)\widehat{\mathsf{a}}'(t_2)|\Psi_{\mathrm{max}}\rangle = \langle\Psi_{\mathrm{dS}}|\mathsf{a}(t_1)\mathsf{a}'(t_2)|\Psi_{\mathrm{dS}}\rangle.$$

So correlators of gravitationally dressed operators after imposing the constraints have the same thermal properties that correlators of "bare" operators had before imposing the constraints.

Thus weakly coupled gravity does not disturb the thermal interpretation of de Sitter space, but it leads to a new interpretation, which we would not have without gravity:

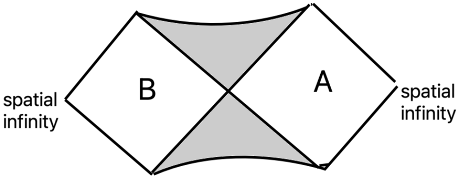
The natural de Sitter state is a maximally mixed state of maximum possible entropy.

Before comparing to the analogous construction for a black hole, it is convenient to put the answer for de Sitter space in a slightly different form. We had the algebra $\mathcal{A}$ generated by $e^{\mathrm{i}pH}\mathsf{a}e^{-\mathrm{i}pH}$, $\mathsf{a} \in \mathcal{A}_0$, along with the observer Hamiltonian q. Instead we can conjugate by $e^{-\mathrm{i}pH}$ and say that $\mathcal{A}$ is generated by a and $q - H$. Thus to include the observer and the constraint, we do two things:

(1) Slightly enlarge the Hilbert space by $\mathcal{H} \to \mathcal{H} \otimes L^2(\mathbb{R}_+)$

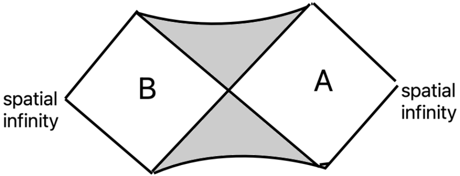
(2) Add one more generator $H - q$ to the algebra.

Now consider a black hole say in asymptotically flat spacetime:



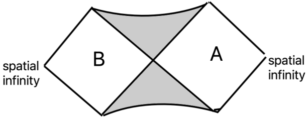
We want to construct an algebra $\mathcal{A}$ of observables for an observer in region A.

In ordinary quantum field theory, we would construct a Hilbert space $\mathcal{H}_0$ of quantum states in this spacetime:



and a Type III algebra $\mathcal{A}_0$ of operators in region A acting on $\mathcal{H}_0$. Adding small gravitational fluctuations that can be observed locally does not qualitatively change the picture. What does qualitatively change the picture is that there is a mode, the relative time-shift between the two sides, that cannot be measured locally. Hence the Hilbert space is $\mathcal{H} = \mathcal{H}_0 \otimes L^2(\mathbb{R})$, where $L^2(\mathbb{R})$ consists of functions of the relative time-shift.

There is also one extra operator that can be measured in the presence of gravity, and does not have an analog in ordinary QFT. This is simply the ADM energy at infinity. Now let us discuss the bulk picture. The eternal black hole spacetime



has a Killing vector field of time translations (forward on right, backward on left) and a corresponding conserved charge H. But in ordinary QFT, there is no formula

$$H \stackrel{?}{=} H_R - H_L$$

where H_R and H_L are energies to the right and left of the horizon.

However, in the presence of gravity, with H_R and H_L understand as the ADM Hamiltonians defined at infinity, we do have

$$H = H_R - H_L,$$

and H_R is an observable for an observer at infinity in region A. Defining $x = H_L$, this means that the algebra $\mathcal{A}$ of observables in region A outside the horizon is generated by the naive algebra $\mathcal{A}_0$ and one more generator

$$H_R = H + H_L = H + x.$$

Thus coupling to gravity modifies the answer we would have in ordinary QFT by

(1) Slightly enlarge the Hilbert space $\mathcal{H} \to \mathcal{H} \otimes L^2(\mathbb{R})$.

(2) Add one more generator $H + x$.

This is the same as we had for de Sitter space, if we identify $q = -x$ except that in de Sitter space we had a constraint $q \geq 0$, which has no analog for the black hole.

To prove that the algebras are Type II, we can define the trace. First consider the black hole. Since the algebra is generated by $\mathsf{a} \in \mathcal{A}_0$ along with $H + x$, a rather general element of the algebra is

$$\widehat{\mathsf{a}} = \int_{-\infty}^{\infty} \mathrm{d}u \, \mathsf{a}(u) e^{\mathrm{i}u(H+x)}.$$

The trace of this element is defined as

$$\mathrm{Tr}\, \widehat{\mathsf{a}} = \int_{-\infty}^{\infty} \mathrm{d}x \mathrm{d}u \, e^{x} e^{\mathrm{i}ux} \langle \Psi_{\mathrm{dS}} | \mathsf{a}(u) | \Psi_{\mathrm{dS}} \rangle.$$

With $\widehat{\mathsf{a}}(x) = \int_{-\infty}^{\infty} \mathrm{d}u \, e^{\mathrm{i}ux} \mathsf{a}(u)$, the trace is

$$\mathrm{Tr}\, \widehat{\mathsf{a}} = \int_{-\infty}^{\infty} \mathrm{d}x \, e^{x} \langle \Psi_{\mathrm{dS}} | \widehat{\mathsf{a}}(x) | \Psi_{\mathrm{dS}} \rangle.$$

To prove that $\mathrm{Tr}\, \widehat{\mathsf{a}}\widehat{\mathsf{a}}' = \mathrm{Tr}\, \widehat{\mathsf{a}}'\widehat{\mathsf{a}}$, one has to use the fact that H is the modular Hamiltonian of the Hartle-Hawking thermal state of the black hole.

The integral in

$$\mathrm{Tr}\,\widehat{\mathsf{a}} = \int_{-\infty}^{\infty} \mathrm{d}x\, e^x \langle \Psi_{\mathrm{dS}} | \widehat{\mathsf{a}}(x) | \Psi_{\mathrm{dS}} \rangle$$

doesn't always converge, because of the factor e^x. Remember that in a Type II_∞ algebra, the trace is only defined for some elements of the algebra. To get the Type II_1 algebra appropriate for de Sitter space, we impose a constraint $q \geq 0$, or equivalently $x \leq 0$. So the formula for the trace becomes

$$\mathrm{Tr}\,\widehat{\mathsf{a}} = \int_{-\infty}^{0} \mathrm{d}x\, e^x \langle \Psi_{\mathrm{dS}} | \widehat{\mathsf{a}}(x) | \Psi_{\mathrm{dS}} \rangle$$

and now there is no problem with convergence of the integral. The trace is defined for all elements of the algebra, and the algebra is of Type II_1.

In sum, I have given a somewhat abstract explanation of why, at least for de Sitter space and the black hole, entropy is better-defined in the presence of gravity than it would be in ordinary quantum field theory. However, a natural question is: Does the entropy defined this way agree with the usual generalized entropy

$$S_{\rm gen} = \frac{\mathcal{A}}{4G\hbar} + S_{\rm out}.$$

We only expect this to work for semiclassical states (states that can be described to good approximation by specifying a state of the quantum fields in a definite spacetime background), because the usual field theoretic discussions of gravitational entropy are only valid for states of this type. It is indeed possible to show that in the case of a semiclassical state, the entropy of a state of the Type II algebra agrees with the generalized entropy – up to an additive constant that is independent of the state. This additive constant is unavoidable because entropy of a Type II algebra is a renormalized entropy with a divergent constant subtracted.