An Algebra of Observables for de Sitter Space

Edward Witten, IAS

In this talk, I will provide a somewhat abstract answer to the question "why is entropy better defined in the presence of gravity than in ordinary quantum field theory?" in the sense that the generalized entropy

$$S_{\mathrm{gen}} = \frac{A}{4G} + S_{\mathrm{out}}$$

is better defined than either term on the right hand side. In the process, we will also give an answer to the question, "In what sense is the generalized entropy an entropy of something?"

(Based on arXiv:2206.10780, with V. Chandrasekharan, G. Penington, and R. Longo, and arXiv.2112.12828.)

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:

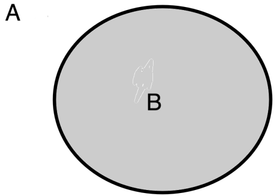
$$\psi_{AB} = \psi_A \otimes \psi_B.$$

More generally, it might be entangled

$$\psi_{AB} = \sum_{i=1}^{k} \sqrt{p_i} \psi_A^i \otimes \psi_B^i, , \quad k > 1$$

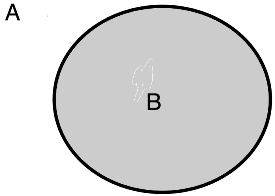
in which case system A (or B) is in a mixed state with a nonzero von Neumann entropy.

Thus 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 entanglement entropy between two regions is ultraviolet divergent and the leading divergence 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. This is the case we assume in ordinary quantum mechanics. Pure states, density matrices, and entropies all exist.

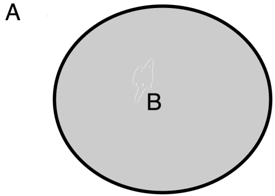
The other types are less familiar::

(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.

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.

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

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

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).

In the limit $N \to \infty$, the function $F(\mathsf{a})$ can be be defined on the whole algebra $\mathcal{A}$ of (bounded) operators on system A, still obeying

$$F(1) = 1, \quad F(\mathsf{a}\mathsf{a}') = F(a'a).$$

Because of the latter property, the function F 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

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

What we have just described is the original Murray-von Neumann algebra of Type II_1, which I claim is isomorphic to the natural algebra of observables in de Sitter space.

The other important von Neumann algebras are all constructed similarly:

To get a Type III algebra (relevant in ordinary QFT), we start with a state that is "fully" but not maximally entangled.

To get a Type II_∞ algebra (relevant to a black hole), we include infinitely many unentangled qubits.

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 algebra $\mathcal{A}$ as the algebra of operators on the "A" part of a bipartite system AB, so in that construction it commutes 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 $G(\mathsf{a})$ is any linear function of $\mathsf{a} \in \mathcal{A}$, we have

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

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.

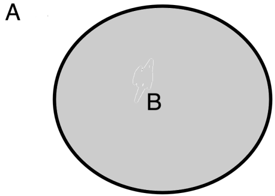
However, in physical terms, the entropy of a state of a Type II algebra is a sort of renormalized entropy from which an infinite constant has been subtracted. For example, let us go back to the system A of N qubits maximally entangled with another such system B. The A system has entropy N, infinite in 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

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

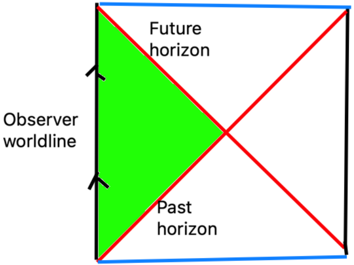
and it is not hard to prove that any other density matrix has strictly negative entropy.

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 – though they can also be presented in parallel – and today I will concentrate on the case of 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.

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 $p = -{\rm i}\frac{\rm d}{{\rm d}q}$.

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

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

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.

The algebra we get this way is isomorphic to the Type II_1 algebra that I described before – acting on the infinite system of almost maximally entangled qubits – if and only if we impose the constraint that the observer energy is bounded below. (Otherwise, we get an algebra of Type II_∞ – appropriate for a black hole but not for de Sitter space.)

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:

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

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

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.

Since the density matrix is 1, all states are equally likely and the probability to observe a given fluctuation is

$$p = \exp(-\Delta S),$$

where ΔS is the entropy reduction if that outcome is observed.

How can this be compatible with the thermal interpretation of de Sitter space, according to which the probability of a fluctuation with energy E is $p = e^{-\beta E}$? In fact, even though de Sitter space is a maximally mixed state of maximum possible entropy, correlation functions in this state are thermal at the usual de Sitter temperature. 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 – as has been suggested by other authors in the past from a different point of view – empty de Sitter space is a maximally mixed state of maximum possible entropy, but nonetheless correlations in this state are thermal at the de Sitter temperature.

To conclude I will just mention a few of the other topics that you can read about in the papers:

1) A similar treatment of the black hole;

2) A proof that in the case of a semiclassical state, the entropy $-\mathrm{Tr}\,\sigma\log\sigma$ of a state of the Type II algebra agrees with the usual generalized entropy $S_{\mathrm{gen}} = A/4G + S_{\mathrm{out}}$ (up to an additive constant independent of the state);

3) A concrete formula for the trace in the Type II algebra.