A Background Independent Algebra in Quantum Gravity

Edward Witten

Lemaitre Conference, June 2024

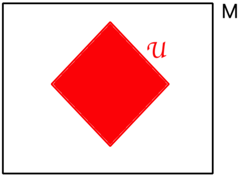
In ordinary quantum mechanics, we usually consider the observer to be outside the system that is being observed.

In ordinary quantum mechanics, we usually consider the observer to be outside the system that is being observed. This is problematical in the presence of gravity, In ordinary quantum mechanics, we usually consider the observer to be outside the system that is being observed. This is problematical in the presence of gravity, most obviously in the case of a closed universe:

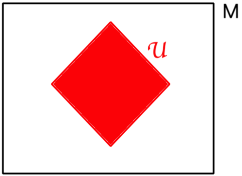
In ordinary quantum mechanics, we usually consider the observer to be outside the system that is being observed. This is problematical in the presence of gravity, most obviously in the case of a closed universe: No one can look at a closed universe from outside.

In ordinary quantum field theory, we can pick an arbitrary region $\mathcal{U}$ in spacetime and define am algebra of observables in region $\mathcal{U}$:

In ordinary quantum field theory, we can pick an arbitrary region $\mathcal{U}$ in spacetime and define am algebra of observables in region $\mathcal{U}$:

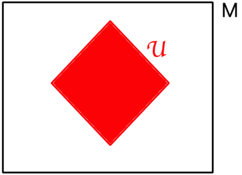


In ordinary quantum field theory, we can pick an arbitrary region $\mathcal{U}$ in spacetime and define am algebra of observables in region $\mathcal{U}$:



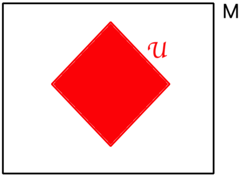
There are problems with this in the presence of gravity:

In ordinary quantum field theory, we can pick an arbitrary region $\mathcal{U}$ in spacetime and define am algebra of observables in region $\mathcal{U}$:



There are problems with this in the presence of gravity: With spacetime fluctuating, it is hard to explain what we mean by the region $\mathcal{U}$?

In ordinary quantum field theory, we can pick an arbitrary region $\mathcal{U}$ in spacetime and define am algebra of observables in region $\mathcal{U}$:



There are problems with this in the presence of gravity: With spacetime fluctuating, it is hard to explain what we mean by the region $\mathcal{U}$? but anyway why do we want to define an algebra unless it is the algebra of observables available to someone who lives in the spacetime?

So I want to construct an algebra that describes the measurements made by an observer.

So I want to construct an algebra that describes the measurements made by an observer. I will assume that the observer knows the laws of nature but has no knowledge of the state of the universe except whatever is gleaned from observation.

So I want to construct an algebra that describes the measurements made by an observer. I will assume that the observer knows the laws of nature but has no knowledge of the state of the universe except whatever is gleaned from observation. (The second part fits our situation in the universe, but the first does not, since in the last few centuries we have been using our observations to learn the laws of nature as well as learning the state of the universe, i.e. part of what is usually called cosmology.

So I want to construct an algebra that describes the measurements made by an observer. I will assume that the observer knows the laws of nature but has no knowledge of the state of the universe except whatever is gleaned from observation. (The second part fits our situation in the universe, but the first does not, since in the last few centuries we have been using our observations to learn the laws of nature as well as learning the state of the universe, i.e. part of what is usually called cosmology. I make the first assumption because it would be much harder to model an observer who is trying to learn the laws of nature.)

The algebra will depend on the laws of nature, but it is required to be *universal* and *background independent*, meaning that is is defined once and for all without any knowledge of the specific spacetime in which the observer is living.

Some relevant papers:

Some relevant papers:

Algebras of operators outside a black hole horizon

Leutheusser and Liu (2021)

Some relevant papers:

Algebras of operators outside a black hole horizon

Leutheusser and Liu (2021)
EW (2021), Chandrasekharan, Penington, and EW (2022)

Some relevant papers:

Algebras of operators outside a black hole horizon

Leutheusser and Liu (2021)
EW (2021), Chandrasekharan, Penington, and EW (2022)

Algebra for a static patch in de Sitter space:

(*) Chandrasekharan, Longo, Penington, and EW (2022)

Some relevant papers:

Algebras of operators outside a black hole horizon

Leutheusser and Liu (2021)
EW (2021), Chandrasekharan, Penington, and EW (2022)

Algebra for a static patch in de Sitter space:

(*) Chandrasekharan, Longo, Penington, and EW (2022)

In JT gravity with negative cosmological constant

Penington and EW (2023), Kolchmeyer (2023)

Some relevant papers:

Algebras of operators outside a black hole horizon

Leutheusser and Liu (2021)
EW (2021), Chandrasekharan, Penington, and EW (2022)

Algebra for a static patch in de Sitter space:

(*) Chandrasekharan, Longo, Penington, and EW (2022)

In JT gravity with negative cosmological constant

Penington and EW (2023), Kolchmeyer (2023)

In a general diamond-like region

Jensen, Sorce, and Speranza (2023)

Some recent developments

C. H. Chen and Penington (2024)

Kudler-Flam, Leutheusser, and Satishchandran (2024).

My starting point for today will be to reinterpret the CLPW paper on the static patch in de Sitter space.

My starting point for today will be to reinterpret the CLPW paper on the static patch in de Sitter space. In that paper, the goal was to define an algebra for the static patch.

My starting point for today will be to reinterpret the CLPW paper on the static patch in de Sitter space. In that paper, the goal was to define an algebra for the static patch. Because of the symmetries of the static patch, it was necessary to assume that there was an observer in the static patch: then one could define an algebra by "gravitationally dressing" an operator to the observer's worldline.

My starting point for today will be to reinterpret the CLPW paper on the static patch in de Sitter space. In that paper, the goal was to define an algebra for the static patch. Because of the symmetries of the static patch, it was necessary to assume that there was an observer in the static patch: then one could define an algebra by "gravitationally dressing" an operator to the observer's worldline. This logic does not apply for more general spacetimes: in a generic spacetime with less symmetry, one could "gravitationally dress" an operator to a feature of the spacetime, or of the state.

My starting point for today will be to reinterpret the CLPW paper on the static patch in de Sitter space. In that paper, the goal was to define an algebra for the static patch. Because of the symmetries of the static patch, it was necessary to assume that there was an observer in the static patch: then one could define an algebra by "gravitationally dressing" an operator to the observer's worldline. This logic does not apply for more general spacetimes: in a generic spacetime with less symmetry, one could "gravitationally dress" an operator to a feature of the spacetime, or of the state. Today we will consider the same construction with a different motivation: background independence.

We expect that in a full theory of quantum gravity, an observer cannot be introduced from outside but must be described by the theory.

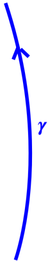
We expect that in a full theory of quantum gravity, an observer cannot be introduced from outside but must be described by the theory. What it means then to assume the presence of an observer is that we define an algebra that makes sense in a subspace of states in which an observer is present.

We expect that in a full theory of quantum gravity, an observer cannot be introduced from outside but must be described by the theory. What it means then to assume the presence of an observer is that we define an algebra that makes sense in a subspace of states in which an observer is present. We don't try to define an algebra that makes sense in all states.

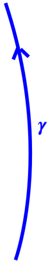
First let us describe the situation in the absence of gravity.

First let us describe the situation in the absence of gravity. The observer propagates in a spacetime M on a geodesic γ:

First let us describe the situation in the absence of gravity. The observer propagates in a spacetime M on a geodesic γ:

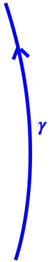


First let us describe the situation in the absence of gravity. The observer propagates in a spacetime M on a geodesic γ:



The worldline is parametrized by proper time τ.

First let us describe the situation in the absence of gravity. The observer propagates in a spacetime M on a geodesic γ:

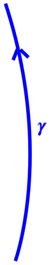


The worldline is parametrized by proper time τ. As in classic work of Unruh (1976), the observer measures along γ, for example, a scalar field ϕ, or the electromagnetic field $F_{\mu\nu}$, or the Riemann tensor $R_{\mu\nu\alpha\beta}$, as well as their covariant derivatives in normal directions.

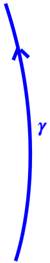
Focus on a particular observable, say $\phi(x(\tau))$ for a scalar field ϕ; I will abbreviate this as $\phi(\tau)$.

Focus on a particular observable, say $\phi(x(\tau))$ for a scalar field ϕ; I will abbreviate this as $\phi(\tau)$. When we take gravity to be dynamical, we have to consider that the same worldline can be embedded in a given spacetime in different ways, differing by $\tau \to \tau + \mathrm{constant}$:

Focus on a particular observable, say $\phi(x(\tau))$ for a scalar field ϕ; I will abbreviate this as $\phi(\tau)$. When we take gravity to be dynamical, we have to consider that the same worldline can be embedded in a given spacetime in different ways, differing by $\tau \to \tau + \mathrm{constant}$:



Focus on a particular observable, say $\phi(x(\tau))$ for a scalar field ϕ; I will abbreviate this as $\phi(\tau)$. When we take gravity to be dynamical, we have to consider that the same worldline can be embedded in a given spacetime in different ways, differing by $\tau \to \tau + \mathrm{constant}$:



So $\phi(\tau)$ isn't by itself a meaningful observable: we need to introduce the observer's degrees of freedom and define τ relative to the observer's clock.

In a minimal model, we equip the observer with a Hamiltonian $H_{\text{obs}} = mc^2 + q$, and a canonical variable $p = -\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{q}}$.

In a minimal model, we equip the observer with a Hamiltonian $H_{\mathrm{obs}} = mc^2 + q$, and a canonical variable $p = -\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}q}$. However, it turns out that it is better to assume that the observer energy is bounded below, say $q \geq 0$ (so m is the observer's rest mass).

In a minimal model, we equip the observer with a Hamiltonian $H_{\rm obs} = mc^2 + q$, and a canonical variable $p = -{\rm i}\frac{\rm d}{{\rm d}q}$. However, it turns out that it is better to assume that the observer energy is bounded below, say $q \geq 0$ (so m is the observer's rest mass). We then only allow operators that preserve this condition, so for example $e^{{\rm i}p}$, which does not preserve $q \geq 0$, should be replaced with $\Pi e^{{\rm i}p}\Pi$, where $\Pi = \Theta(q)$ is the projection operator onto $q \geq 0$.

We now want to allow only operators that commute with

$$\widehat{H} = H_{\rm bulk} + H_{\rm obs},$$

where $H_{\rm bulk}$ is (any) gravitational constraint operator that generates a shift of τ along the worldline.

We now want to allow only operators that commute with

$$\widehat{H} = H_{\rm bulk} + H_{\rm obs},$$

where $H_{\rm bulk}$ is (any) gravitational constraint operator that generates a shift of τ along the worldline. An operator that commutes with $\widehat{H}$ is invariant under a spacetime diffeomorphism that moves the observer worldline forward in time, together with a time translation of the observer's system.

How do we find operators that commute with $\widehat{H} = H + H_{\mathrm{obs}} = H_{\mathrm{bulk}} + H_{\mathrm{obs}}$?

How do we find operators that commute with $\widehat{H} = H + H_{\mathrm{obs}} = H_{\mathrm{bulk}} + H_{\mathrm{obs}}$? Since

$$[H_{\mathrm{bulk}}, \phi(\tau)] = -\mathrm{i}\dot{\phi}(\tau),$$

we need

$$[q, \phi(\tau)] = \mathrm{i}\dot{\phi}(\tau),$$

which we can achieve by just setting

$$\tau = p$$

or more generally

$$\tau = p + s$$

for a constant s.

So a typical allowed operator is $\phi(p+s)$, or more precisely

$$\widehat{\phi}_s = \Pi\phi(p+s)\Pi = \Theta(q)\phi(p+s)\Theta(q).$$

So a typical allowed operator is $\phi(p+s)$, or more precisely

$$\widehat{\phi}_s = \Pi\phi(p+s)\Pi = \Theta(q)\phi(p+s)\Theta(q).$$

In addition to these operators (with ϕ possibly replaced by any local field along the worldline such as the electromagnetic field or the Riemann tensor) there is one more obvious operator that commutes with $\widehat{H}$, namely q itself.

So a typical allowed operator is $\phi(p+s)$, or more precisely

$$\widehat{\phi}_s = \Pi\phi(p+s)\Pi = \Theta(q)\phi(p+s)\Theta(q).$$

In addition to these operators (with ϕ possibly replaced by any local field along the worldline such as the electromagnetic field or the Riemann tensor) there is one more obvious operator that commutes with $\widehat{H}$, namely q itself. So we define an algebra $\mathcal{A}_{\rm obs}$ that is generated by the $\widehat{\phi}_s$ as well as q.

The setup hopefully sounds “background independent,” since we described it without picking a background.

The setup hopefully sounds "background independent," since we described it without picking a background. However, background independence really depends on interpreting the formulas properly.

The setup hopefully sounds "background independent," since we described it without picking a background. However, background independence really depends on interpreting the formulas properly. We will not get background independence if we interpret $\widehat{\phi}_s$ and q as Hilbert space operators.

The setup hopefully sounds "background independent," since we described it without picking a background. However, background independence really depends on interpreting the formulas properly. We will not get background independence if we interpret $\widehat{\phi}_s$ and q as Hilbert space operators. To get a Hilbert space on which $\widehat{\phi}_s$ and q act, we have to pick a spacetime M in which the observer is propagating.

The setup hopefully sounds "background independent," since we described it without picking a background. However, background independence really depends on interpreting the formulas properly. We will not get background independence if we interpret $\widehat{\phi}_s$ and q as Hilbert space operators. To get a Hilbert space on which $\widehat{\phi}_s$ and q act, we have to pick a spacetime M in which the observer is propagating. Then we won't have background independence.

The setup hopefully sounds "background independent," since we described it without picking a background. However, background independence really depends on interpreting the formulas properly. We will not get background independence if we interpret $\widehat{\phi}_s$ and q as Hilbert space operators. To get a Hilbert space on which $\widehat{\phi}_s$ and q act, we have to pick a spacetime M in which the observer is propagating. Then we won't have background independence.

To get background independence, we have to think of $\mathcal{A}_{\rm obs}$ as an operator product algebra, rather than an algebra of Hilbert space operators.

The setup hopefully sounds "background independent," since we described it without picking a background. However, background independence really depends on interpreting the formulas properly. We will not get background independence if we interpret $\widehat{\phi}_s$ and q as Hilbert space operators. To get a Hilbert space on which $\widehat{\phi}_s$ and q act, we have to pick a spacetime M in which the observer is propagating. Then we won't have background independence.

To get background independence, we have to think of $\mathcal{A}_{\mathrm{obs}}$ as an operator product algebra, rather than an algebra of Hilbert space operators. The algebras for different M's are inequivalent representations of the same underlying operator product algebra.

In the absence of gravity, we would characterize the objects $\phi(\tau)$ by their universal short distance singularities:

$$\phi(\tau)\phi(\tau') \sim C(\tau - \tau' - \mathrm{i}\epsilon)^{-2\Delta} + \cdots .$$

In the absence of gravity, we would characterize the objects $\phi(\tau)$ by their universal short distance singularities:

$$\phi(\tau)\phi(\tau') \sim C(\tau - \tau' - \mathrm{i}\epsilon)^{-2\Delta} + \cdots .$$

This characterization does not require any knowledge about the quantum state.

In the absence of gravity, we would characterize the objects $\phi(\tau)$ by their universal short distance singularities:

$$\phi(\tau)\phi(\tau') \sim C(\tau - \tau' - \mathrm{i}\epsilon)^{-2\Delta} + \cdots .$$

This characterization does not require any knowledge about the quantum state. After coupling to gravity and including the observer and the constraint, the short distance expansion in powers of $\tau - \tau'$ becomes an expansion in $1/q$.

In the absence of gravity, we would characterize the objects $\phi(\tau)$ by their universal short distance singularities:

$$\phi(\tau)\phi(\tau') \sim C(\tau - \tau' - \mathrm{i}\epsilon)^{-2\Delta} + \cdots .$$

This characterization does not require any knowledge about the quantum state. After coupling to gravity and including the observer and the constraint, the short distance expansion in powers of $\tau - \tau'$ becomes an expansion in $1/q$. We characterize $\mathcal{A}_{\rm obs}$ purely by the universal short distance or $1/q$ expansion of operator products.

In the absence of gravity, we would characterize the objects $\phi(\tau)$ by their universal short distance singularities:

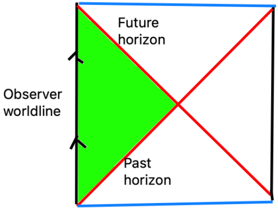
$$\phi(\tau)\phi(\tau') \sim C(\tau - \tau' - \mathrm{i}\epsilon)^{-2\Delta} + \cdots .$$

This characterization does not require any knowledge about the quantum state. After coupling to gravity and including the observer and the constraint, the short distance expansion in powers of $\tau - \tau'$ becomes an expansion in $1/q$. We characterize $\mathcal{A}_{\rm obs}$ purely by the universal short distance or $1/q$ expansion of operator products. With that understanding, $\mathcal{A}_{\rm obs}$ is background-independent.

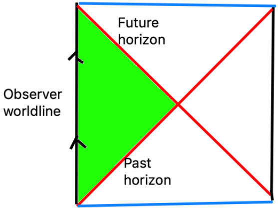
There is a very special case that turns out to be important.

There is a very special case that turns out to be important. This is the case that M is an empty de Sitter space, with some positive value of the effective cosmological constant.

There is a very special case that turns out to be important. This is the case that M is an empty de Sitter space, with some positive value of the effective cosmological constant.



There is a very special case that turns out to be important. This is the case that M is an empty de Sitter space, with some positive value of the effective cosmological constant.



The green region is called a static patch, because it is invariant under a particular de Sitter generator H that advances the proper time of the observer.

In the absence of gravity, there is a distinguished de Sitter invariant state $\Psi_{\rm dS}$ such that correlation functions in this state are thermal at the de Sitter temperature $T_{\rm dS} = 1/\beta_{\rm dS}$ (Gibbons and Hawking; Figari, Nappi, and Hoegh-Krohn).

In the absence of gravity, there is a distinguished de Sitter invariant state $\Psi_{\rm dS}$ such that correlation functions in this state are thermal at the de Sitter temperature $T_{\rm dS} = 1/\beta_{\rm dS}$ (Gibbons and Hawking; Figari, Nappi, and Hoegh-Krohn). For example, this means that two point functions $\langle\Psi_{\rm dS}|\phi(\tau)\phi'(\tau')|\Psi_{\rm dS}\rangle$ have two key properties:

In the absence of gravity, there is a distinguished de Sitter invariant state Ψ_{dS} such that correlation functions in this state are thermal at the de Sitter temperature $T_{\mathrm{dS}} = 1/\beta_{\mathrm{dS}}$ (Gibbons and Hawking; Figari, Nappi, and Hoegh-Krohn). For example, this means that two point functions $\langle\Psi_{\mathrm{dS}}|\phi(\tau)\phi'(\tau')|\Psi_{\mathrm{dS}}\rangle$ have two key properties:

(1) Time translation symmetry:

$$\langle\Psi_{\mathrm{dS}}|\phi(\tau+s)\phi'(\tau'+s)|\Psi_{\mathrm{dS}}\rangle = \langle\Psi_{\mathrm{dS}}|\phi(\tau)\phi'(\tau')|\Psi_{\mathrm{dS}}\rangle.$$

In the absence of gravity, there is a distinguished de Sitter invariant state $\Psi_{\rm dS}$ such that correlation functions in this state are thermal at the de Sitter temperature $T_{\rm dS} = 1/\beta_{\rm dS}$ (Gibbons and Hawking; Figari, Nappi, and Hoegh-Krohn). For example, this means that two point functions $\langle\Psi_{\rm dS}|\phi(\tau)\phi'(\tau')|\Psi_{\rm dS}\rangle$ have two key properties:

(1) Time translation symmetry:

$$\langle\Psi_{\rm dS}|\phi(\tau+s)\phi'(\tau'+s)|\Psi_{\rm dS}\rangle = \langle\Psi_{\rm dS}|\phi(\tau)\phi'(\tau')|\Psi_{\rm dS}\rangle.$$

(2) The KMS condition, which says roughly:

$$\langle\Psi_{\rm dS}|\phi(\tau)\phi'(0)|\Psi_{\rm dS}\rangle = \langle\Psi_{\rm dS}|\phi'(0)\phi(\tau-{\rm i}\beta)|\Psi_{\rm dS}\rangle.$$

In the absence of gravity, there is a distinguished de Sitter invariant state $\Psi_{\rm dS}$ such that correlation functions in this state are thermal at the de Sitter temperature $T_{\rm dS} = 1/\beta_{\rm dS}$ (Gibbons and Hawking; Figari, Nappi, and Hoegh-Krohn). For example, this means that two point functions $\langle\Psi_{\rm dS}|\phi(\tau)\phi'(\tau')|\Psi_{\rm dS}\rangle$ have two key properties:

(1) Time translation symmetry:

$$\langle\Psi_{\rm dS}|\phi(\tau+s)\phi'(\tau'+s)|\Psi_{\rm dS}\rangle = \langle\Psi_{\rm dS}|\phi(\tau)\phi'(\tau')|\Psi_{\rm dS}\rangle.$$

(2) The KMS condition, which says roughly:

$$\langle\Psi_{\rm dS}|\phi(\tau)\phi'(0)|\Psi_{\rm dS}\rangle = \langle\Psi_{\rm dS}|\phi'(0)\phi(\tau-{\rm i}\beta)|\Psi_{\rm dS}\rangle.$$

(A precise statement involves holomorphy of the correlation function in a strip in the complex plane.)

Including gravity and the observer, we define a special state in which the observer energy has a thermal distribution at the de Sitter temperature

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

and we replace operators $\phi(\tau)$ by "gravitationally dressed" operators $\widehat{\phi}_s = \Pi \phi(p+s) \Pi$.

Including gravity and the observer, we define a special state in which the observer energy has a thermal distribution at the de Sitter temperature

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

and we replace operators $\phi(\tau)$ by "gravitationally dressed" operators $\widehat{\phi}_s = \Pi \phi(p+s) \Pi$. Then a straightforward computation shows that

Including gravity and the observer, we define a special state in which the observer energy has a thermal distribution at the de Sitter temperature

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

and we replace operators $\phi(\tau)$ by "gravitationally dressed" operators $\widehat{\phi}_s = \Pi \phi(p+s) \Pi$. Then a straightforward computation shows that

(1′) We still have time-translation symmetry

$$\langle \Psi_{\rm max} | \widehat{\phi}_s \widehat{\phi}'_{s'} | \Psi_{\rm max} \rangle = \langle \Psi_{\rm max} | \widehat{\phi}'_{s+c} \widehat{\phi}'_{s'+c} | \Psi_{\rm max} \rangle, \quad c \in \mathbb{R}.$$

Including gravity and the observer, we define a special state in which the observer energy has a thermal distribution at the de Sitter temperature

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

and we replace operators $\phi(\tau)$ by "gravitationally dressed" operators $\widehat{\phi}_s = \Pi \phi(p+s) \Pi$. Then a straightforward computation shows that

(1′) We still have time-translation symmetry

$$\langle \Psi_{\rm max} | \widehat{\phi}_s \widehat{\phi}'_{s'} | \Psi_{\rm max} \rangle = \langle \Psi_{\rm max} | \widehat{\phi}'_{s+c} \widehat{\phi}'_{s'+c} | \Psi_{\rm max} \rangle, \quad c \in \mathbb{R}.$$

(2′) The KMS condition simplifies:

$$\langle \Psi_{\rm max} | \widehat{\phi}_s \widehat{\phi}'_{s'} | \Psi_{\rm max} \rangle = \langle \Psi_{\rm max} | \widehat{\phi}'_{s'} \widehat{\phi}_s | \Psi_{\rm max} \rangle.$$

Condition (2′) tells us that if, for any $\mathbf{a} \in \mathcal{A}_{\rm obs}$, we define

$$\operatorname{Tr} \mathbf{a} = \langle \Psi_{\max} | \mathbf{a} | \Psi_{\max} \rangle,$$

Condition $(2')$ tells us that if, for any $\mathbf{a} \in \mathcal{A}_{\mathrm{obs}}$, we define

$$\mathrm{Tr}\,\mathbf{a} = \langle \Psi_{\mathrm{max}} | \mathbf{a} | \Psi_{\mathrm{max}} \rangle,$$

then the function Tr does have the algebraic property of a trace:

$$\mathrm{Tr}\,\mathbf{ab} = \mathrm{Tr}\,\mathbf{ba}, \quad \mathbf{a}, \mathbf{b} \in \mathcal{A}_{\mathrm{obs}}.$$

Condition (2′) tells us that if, for any $\mathbf{a} \in \mathcal{A}_{\rm obs}$, we define

$$\mathrm{Tr}\,\mathbf{a} = \langle \Psi_{\rm max} | \mathbf{a} | \Psi_{\rm max} \rangle,$$

then the function Tr does have the algebraic property of a trace:

$$\mathrm{Tr}\,\mathbf{ab} = \mathrm{Tr}\,\mathbf{ba}, \quad \mathbf{a}, \mathbf{b} \in \mathcal{A}_{\rm obs}.$$

This function has the property that $\mathrm{Tr}\,\mathbf{a}^\dagger \mathbf{a} > 0$ for all $\mathbf{a} \neq 0$, meaning in particular that it is "nondegenerate."

Condition (2′) tells us that if, for any $\mathbf{a} \in \mathcal{A}_{\rm obs}$, we define

$$\mathrm{Tr}\,\mathbf{a} = \langle \Psi_{\rm max} | \mathbf{a} | \Psi_{\rm max} \rangle,$$

then the function Tr does have the algebraic property of a trace:

$$\mathrm{Tr}\,\mathbf{ab} = \mathrm{Tr}\,\mathbf{ba}, \quad \mathbf{a}, \mathbf{b} \in \mathcal{A}_{\rm obs}.$$

This function has the property that $\mathrm{Tr}\,\mathbf{a}^\dagger\mathbf{a} > 0$ for all $\mathbf{a} \neq 0$, meaning in particular that it is "nondegenerate." Note that if $\Psi_{\rm max}$ is normalized then

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

Let $\mathcal{H}_{\mathrm{dS}}$ be the Hilbert space that we get by quantizing fields in de Sitter space (in perturbation theory).

Let $\mathcal{H}_{\mathrm{dS}}$ be the Hilbert space that we get by quantizing fields in de Sitter space (in perturbation theory). It is important to understand what $\mathcal{H}_{\mathrm{dS}}$ describes and what it doesn't describe.

Let $\mathcal{H}_{\mathrm{dS}}$ be the Hilbert space that we get by quantizing fields in de Sitter space (in perturbation theory). It is important to understand what $\mathcal{H}_{\mathrm{dS}}$ describes and what it doesn't describe. There are states in $\mathcal{H}_{\mathrm{dS}}$ with any number of graviton excitations, but not a number of them of order $1/\hbar$.

Let $\mathcal{H}_{\rm dS}$ be the Hilbert space that we get by quantizing fields in de Sitter space (in perturbation theory). It is important to understand what $\mathcal{H}_{\rm dS}$ describes and what it doesn't describe. There are states in $\mathcal{H}_{\rm dS}$ with any number of graviton excitations, but not a number of them of order $1/\hbar$. An $\hbar$-independent number of graviton excitations produces a back reaction on the geometry that is of order $\hbar$ – or order G.

Let $\mathcal{H}_{\mathrm{dS}}$ be the Hilbert space that we get by quantizing fields in de Sitter space (in perturbation theory). It is important to understand what $\mathcal{H}_{\mathrm{dS}}$ describes and what it doesn't describe. There are states in $\mathcal{H}_{\mathrm{dS}}$ with any number of graviton excitations, but not a number of them of order $1/\hbar$. An $\hbar$-independent number of graviton excitations produces a back reaction on the geometry that is of order $\hbar$ – or order G. So $\mathcal{H}_{\mathrm{dS}}$ doesn't describe states with an $\mathcal{O}(1)$ change in the geometry, only an $\mathcal{O}(G)$ change.

Let Ψ be any state in $\mathcal{H}_{\rm dS}$ and consider the function $\mathbf{a} \to \langle\Psi|\mathbf{a}|\Psi\rangle$, $\mathbf{a} \in \mathcal{A}_{\rm obs}$.

Let Ψ be any state in $\mathcal{H}_{\rm dS}$ and consider the function $\mathbf{a} \to \langle\Psi|\mathbf{a}|\Psi\rangle$, $\mathbf{a} \in \mathcal{A}_{\rm obs}$. Roughly speaking, because $\mathcal{A}_{\rm obs}$ has the nondegenerate trace Tr, we can hope that there is a "density matrix" $\rho \in \mathcal{A}_{\rm obs}$ such that

$$\langle\Psi|\mathbf{a}|\Psi\rangle = \mathrm{Tr}\,\mathbf{a}\rho, \quad \mathbf{a} \in \mathcal{A}_{\rm obs}.$$

Let Ψ be any state in $\mathcal{H}_{\mathrm{dS}}$ and consider the function $\mathbf{a} \to \langle\Psi|\mathbf{a}|\Psi\rangle$, $\mathbf{a} \in \mathcal{A}_{\mathrm{obs}}$. Roughly speaking, because $\mathcal{A}_{\mathrm{obs}}$ has the nondegenerate trace Tr, we can hope that there is a "density matrix" $\rho \in \mathcal{A}_{\mathrm{obs}}$ such that

$$\langle\Psi|\mathbf{a}|\Psi\rangle = \mathrm{Tr}\,\mathbf{a}\rho, \quad \mathbf{a} \in \mathcal{A}_{\mathrm{obs}}.$$

Rather as in ordinary quantum mechanics, we expect ρ to be a positive element $\rho \in \mathcal{A}_{\mathrm{obs}}$ with $\mathrm{Tr}\,\rho = 1$.

Let Ψ be any state in $\mathcal{H}_{\rm dS}$ and consider the function $\mathbf{a} \to \langle\Psi|\mathbf{a}|\Psi\rangle$, $\mathbf{a} \in \mathcal{A}_{\rm obs}$. Roughly speaking, because $\mathcal{A}_{\rm obs}$ has the nondegenerate trace Tr, we can hope that there is a "density matrix" $\rho \in \mathcal{A}_{\rm obs}$ such that

$$\langle\Psi|\mathbf{a}|\Psi\rangle = \mathrm{Tr}\,\mathbf{a}\rho, \quad \mathbf{a} \in \mathcal{A}_{\rm obs}.$$

Rather as in ordinary quantum mechanics, we expect ρ to be a positive element $\rho \in \mathcal{A}_{\rm obs}$ with $\mathrm{Tr}\,\rho = 1$. For example, let us find the density matrix of the state $\Psi_{\rm max}$.

The definition of the trace makes it clear that the density matrix of the state $\Psi_{\max}$ is $\sigma_{\max} = 1$, since to satisfy

$$\langle \Psi_{\max} | \mathbf{a} | \Psi_{\max} \rangle = \operatorname{Tr} \mathbf{a} \sigma_{\max} \equiv \langle \Psi_{\max} | \mathbf{a} \sigma_{\max} | \Psi_{\max} \rangle,$$

we set

$$\sigma_{\max} = 1.$$

The definition of the trace makes it clear that the density matrix of the state Ψ_{max} is $\sigma_{\mathrm{max}} = 1$, since to satisfy

$$\langle \Psi_{\mathrm{max}} | \mathbf{a} | \Psi_{\mathrm{max}} \rangle = \mathrm{Tr}\, \mathbf{a} \sigma_{\mathrm{max}} \equiv \langle \Psi_{\mathrm{max}} | \mathbf{a} \sigma_{\mathrm{max}} | \Psi_{\mathrm{max}} \rangle,$$

we set

$$\sigma_{\mathrm{max}} = 1.$$

This means that Ψ_{max} is "maximally mixed," similar to a maximally mixed state in ordinary quantum mechanics whose density matrix is a multiple of the identity.

Now if $\mathbf{a} \in \mathcal{A}_{\mathrm{obs}}$ is any operator, consider the state $\Psi_{\mathbf{a}} = \mathbf{a}\Psi_{\max}$.

Now if $\mathbf{a} \in \mathcal{A}_{\mathrm{obs}}$ is any operator, consider the state $\Psi_{\mathbf{a}} = \mathbf{a}\Psi_{\mathrm{max}}$. It has a density matrix $\rho_{\Psi_{\mathbf{a}}} = \mathbf{a}\mathbf{a}^\dagger$, since for any $\mathbf{b} \in \mathcal{A}_{\mathrm{obs}}$,

$$\langle\Psi_{\mathbf{a}}|\mathbf{b}|\Psi_{\mathbf{a}}\rangle = \langle\Psi_{\mathrm{max}}|\mathbf{a}^\dagger\mathbf{b}\mathbf{a}|\Psi_{\mathrm{max}}\rangle = \mathrm{Tr}\,\mathbf{a}^\dagger\mathbf{b}\mathbf{a} = \mathrm{Tr}\,\mathbf{b}\mathbf{a}\mathbf{a}^\dagger.$$

Now if $\mathbf{a} \in \mathcal{A}_{\rm obs}$ is any operator, consider the state $\Psi_{\mathbf{a}} = \mathbf{a}\Psi_{\rm max}$. It has a density matrix $\rho_{\Psi_{\mathbf{a}}} = \mathbf{a}\mathbf{a}^\dagger$, since for any $\mathbf{b} \in \mathcal{A}_{\rm obs}$,

$$\langle \Psi_{\mathbf{a}} | \mathbf{b} | \Psi_{\mathbf{a}} \rangle = \langle \Psi_{\rm max} | \mathbf{a}^\dagger \mathbf{b} \mathbf{a} | \Psi_{\rm max} \rangle = {\rm Tr}\, \mathbf{a}^\dagger \mathbf{b} \mathbf{a} = {\rm Tr}\, \mathbf{b} \mathbf{a} \mathbf{a}^\dagger .$$

But states $\Psi_{\mathbf{a}}$ are dense in $\mathcal{H}_{\rm dS}$ – roughly by the Reeh-Schlieder theorem, which is the fundamental result about entanglement in quantum field theory.

Now if $\mathbf{a} \in \mathcal{A}_{\rm obs}$ is any operator, consider the state $\Psi_{\mathbf{a}} = \mathbf{a}\Psi_{\rm max}$. It has a density matrix $\rho_{\Psi_{\mathbf{a}}} = \mathbf{a}\mathbf{a}^\dagger$, since for any $\mathbf{b} \in \mathcal{A}_{\rm obs}$,

$$\langle \Psi_{\mathbf{a}} | \mathbf{b} | \Psi_{\mathbf{a}} \rangle = \langle \Psi_{\rm max} | \mathbf{a}^\dagger \mathbf{b} \mathbf{a} | \Psi_{\rm max} \rangle = {\rm Tr}\, \mathbf{a}^\dagger \mathbf{b} \mathbf{a} = {\rm Tr}\, \mathbf{b} \mathbf{a} \mathbf{a}^\dagger .$$

But states $\Psi_{\mathbf{a}}$ are dense in $\mathcal{H}_{\rm dS}$ – roughly by the Reeh-Schlieder theorem, which is the fundamental result about entanglement in quantum field theory. So a dense set of states have density matrices.

If we want *all* states to have density matrices, we need to take a useful further step.

If we want *all* states to have density matrices, we need to take a useful further step. The Hilbert space $\mathcal{H}_{\rm dS}$ is the *closure* of a dense set of states $\mathbf{a}\Psi_{\rm dS}$, so if we want *every* state in $\mathcal{H}_{\rm dS}$ to have a density matrix, we have to similarly take a closure of $\mathcal{A}_{\rm obs}$.

If we want *all* states to have density matrices, we need to take a useful further step. The Hilbert space $\mathcal{H}_{\mathrm{dS}}$ is the *closure* of a dense set of states $\mathbf{a}\Psi_{\mathrm{dS}}$, so if we want *every* state in $\mathcal{H}_{\mathrm{dS}}$ to have a density matrix, we have to similarly take a closure of $\mathcal{A}_{\mathrm{obs}}$. This closure, which is no longer background independent, can be defined as the von Neumann algebra generated by bounded operators in $\mathcal{A}_{\mathrm{obs}}$.

If we want *all* states to have density matrices, we need to take a useful further step. The Hilbert space $\mathcal{H}_{\rm dS}$ is the *closure* of a dense set of states $\mathbf{a}\Psi_{\rm dS}$, so if we want *every* state in $\mathcal{H}_{\rm dS}$ to have a density matrix, we have to similarly take a closure of $\mathcal{A}_{\rm obs}$. This closure, which is no longer background independent, can be defined as the von Neumann algebra generated by bounded operators in $\mathcal{A}_{\rm obs}$. I will call the closure $\mathcal{A}_{\rm obs,dS}$.

If we want *all* states to have density matrices, we need to take a useful further step. The Hilbert space $\mathcal{H}_{\rm dS}$ is the *closure* of a dense set of states $\mathbf{a}\Psi_{\rm dS}$, so if we want *every* state in $\mathcal{H}_{\rm dS}$ to have a density matrix, we have to similarly take a closure of $\mathcal{A}_{\rm obs}$. This closure, which is no longer background independent, can be defined as the von Neumann algebra generated by bounded operators in $\mathcal{A}_{\rm obs}$. I will call the closure $\mathcal{A}_{\rm obs,dS}$. Every state in $\mathcal{H}_{\rm dS}$ has a density matrix in (or technically, in general affiliated to) $\mathcal{A}_{\rm obs,dS}$. It is in this step that von Neumann algebras enter the picture.

Once we know that every state has a density matrix, we can define entropies as well.

Once we know that every state has a density matrix, we can define entropies as well. The von Neumann entropy of a density matrix ρ is as usual

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

Once we know that every state has a density matrix, we can define entropies as well. The von Neumann entropy of a density matrix ρ is as usual

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

In ordinary quantum mechanics, a maximally mixed state has a density matrix that is a multiple of the identity, and it has the maximum possible von Neumann entropy.

Once we know that every state has a density matrix, we can define entropies as well. The von Neumann entropy of a density matrix ρ is as usual

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

In ordinary quantum mechanics, a maximally mixed state has a density matrix that is a multiple of the identity, and it has the maximum possible von Neumann entropy. The analog here is Ψ_{max}, with density matrix $\sigma_{\mathrm{max}} = 1$.

Once we know that every state has a density matrix, we can define entropies as well. The von Neumann entropy of a density matrix ρ is as usual

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

In ordinary quantum mechanics, a maximally mixed state has a density matrix that is a multiple of the identity, and it has the maximum possible von Neumann entropy. The analog here is Ψ_{max}, with density matrix $\sigma_{\mathrm{max}} = 1$. It is clear that

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

and by imitating an argument that in ordinary quantum mechanics proves that a maximally mixed state has maximum possible entropy, one can prove that every other density matrix $\rho \neq 1$ has strictly smaller entropy:

$$S(\rho) < 0.$$

One way to make this proof is as follows.

One way to make this proof is as follows. Let $\rho \neq 1$ be any density matrix other than the identity.

One way to make this proof is as follows. Let $\rho \neq 1$ be any density matrix other than the identity. Then for $0 \leq t \leq 1$, $\rho_t = (1-t) + t\rho$ is also a density matrix.

One way to make this proof is as follows. Let $\rho \neq 1$ be any density matrix other than the identity. Then for $0 \leq t \leq 1$, $\rho_t = (1-t) + t\rho$ is also a density matrix. Let $f(t) = S(\rho_t) = -\mathrm{Tr}\,\rho_t \log \rho_t$, so $S(\rho) = f(1)$.

One way to make this proof is as follows. Let $\rho \neq 1$ be any density matrix other than the identity. Then for $0 \leq t \leq 1$, $\rho_t = (1-t) + t\rho$ is also a density matrix. Let $f(t) = S(\rho_t) = -\mathrm{Tr}\,\rho_t \log \rho_t$, so $S(\rho) = f(1)$. Then

$$f'(0) = 0, \quad f''(t) < 0 \;\; \text{for} \;\; 0 \leq t \leq 1.$$

The first statement is almost immediate, and to prove the second, one uses $\log M = \int_0^\infty \mathrm{d}s \left(\frac{1}{s} - \frac{1}{s+M}\right)$, which leads to

$$f''(t) = -\int_0^\infty \mathrm{d}s\, \mathrm{Tr}\, \frac{1}{s+\rho_t}(1-\rho)\frac{1}{s+\rho_t}(1-\rho) = -\int_0^\infty \mathrm{d}s\, \mathrm{Tr}\, B^2 < 0,$$

where B is the self-adjoint operator $B = \left(\frac{1}{s+\rho_t}\right)^{1/2}(1-\rho)\left(\frac{1}{s+\rho_t}\right)^{1/2}$. Since $f'(0) = 0$, $f''(t) < 0$, we get $f(1) < 0$ so

$$S(\rho) < 0.$$

Thus, the system consisting of an observer in a static patch in de Sitter space has a state of maximum entropy

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

consisting of empty de Sitter space with a thermal distribution of the observer energy.

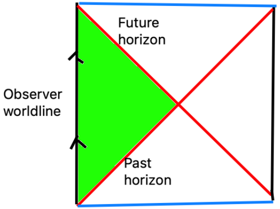
Thus, the system consisting of an observer in a static patch in de Sitter space has a state of maximum entropy

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

consisting of empty de Sitter space with a thermal distribution of the observer energy. Why did this happen?

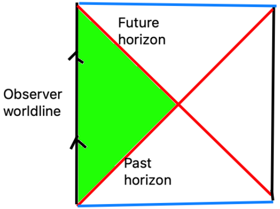
The original argument that empty de Sitter space has maximum entropy is due to Bousso (2000), who argued that this must be true, based on the Second Law of Thermodynamics, because the static patch is empty in the far future:

The original argument that empty de Sitter space has maximum entropy is due to Bousso (2000), who argued that this must be true, based on the Second Law of Thermodynamics, because the static patch is empty in the far future:



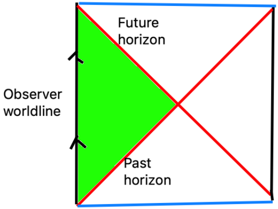
In the present context, we've defined the static patch by the presence of the observer, so by definition the observer doesn't leave the static patch even in the far future.

The original argument that empty de Sitter space has maximum entropy is due to Bousso (2000), who argued that this must be true, based on the Second Law of Thermodynamics, because the static patch is empty in the far future:



In the present context, we've defined the static patch by the presence of the observer, so by definition the observer doesn't leave the static patch even in the far future. But we can expect that in the far future the static patch will be empty except for the presence of the observer, and that the observer will be in thermal equilibrium with the bulk quantum fields, and that is what we see in the state $\Psi_{\max}$.

The original argument that empty de Sitter space has maximum entropy is due to Bousso (2000), who argued that this must be true, based on the Second Law of Thermodynamics, because the static patch is empty in the far future:



In the present context, we've defined the static patch by the presence of the observer, so by definition the observer doesn't leave the static patch even in the far future. But we can expect that in the far future the static patch will be empty except for the presence of the observer, and that the observer will be in thermal equilibrium with the bulk quantum fields, and that is what we see in the state $\Psi_{\rm max}$. So the maximum entropy state that we found is the one suggested by Bousso's argument.

In a more general spacetime, I do not know an equally explicit definition of what would be meant by a state of the observer algebra.

In a more general spacetime, I do not know an equally explicit definition of what would be meant by a state of the observer algebra. However, there is a fairly reasonable conjecture, which is inspired by A. Wall's proof of the generalized second law (2011) and has some support from recent work of Chen and Penington.

In a more general spacetime, I do not know an equally explicit definition of what would be meant by a state of the observer algebra. However, there is a fairly reasonable conjecture, which is inspired by A. Wall's proof of the generalized second law (2011) and has some support from recent work of Chen and Penington. The idea is to interpret the Hartle-Hawking no boundary state as a sort of universal state of maximum entropy (generalizing empty de Sitter space, which has maximum entropy among states in a particular de Sitter spacetime).

In a more general spacetime, I do not know an equally explicit definition of what would be meant by a state of the observer algebra. However, there is a fairly reasonable conjecture, which is inspired by A. Wall's proof of the generalized second law (2011) and has some support from recent work of Chen and Penington. The idea is to interpret the Hartle-Hawking no boundary state as a sort of universal state of maximum entropy (generalizing empty de Sitter space, which has maximum entropy among states in a particular de Sitter spacetime). The expectation value in the no boundary state $\mathbf{a} \to \langle \Psi_{\mathrm{HH}} | \mathbf{a} | \Psi_{\mathrm{HH}} \rangle$ is a state of the observer algebra that I will denote as $\sigma(\mathbf{a})$. Then if $\mathbf{a} \to \rho(\mathbf{a})$ is any state of the observer algebra, then I suggest that the relative entropy between ρ and σ for the observer algebra gives - up to sign - a definition of the entropy of the state seen by the observer:

$$S(\rho) = -S(\rho|\sigma).$$

A lot of things are missing, for example:

A lot of things are missing, for example:

* What about an observer (or a civilization) that did not always exist?

A lot of things are missing, for example:

* What about an observer (or a civilization) that did not always exist?

* In this presentation, I used a field theory language; now can the discussion be generalized to string/M-theory?

A lot of things are missing, for example:

* What about an observer (or a civilization) that did not always exist?

* In this presentation, I used a field theory language; now can the discussion be generalized to string/M-theory?

* Though the definitions make sense regardless, I want to remark that the presentation that I've given seems most natural for an observer who because of black hole or cosmological horizons cannot see the whole universe.