

Entanglement in Field Theory and Holography

Matthew Headrick*

Brandeis University

E-mail: headrick@brandeis.edu

These lecture notes aim to provide an introduction to entanglement entropies in quantum field theories, including holographic ones. We explore basic properties and simple examples of entanglement entropies, mostly in two dimensions, with an emphasis on physical rather than formal aspects of the subject. In the holographic case, the focus is on how that physics is realized geometrically by the Ryu-Takayanagi formula. In order to make the notes somewhat self-contained for readers whose background is in high-energy theory, a brief introduction to the relevant aspects of quantum information theory is included.

Theoretical Advanced Study Institute Summer School 2017 "Physics at the Fundamental Frontier"
4 June - 1 July 2017
Boulder, Colorado

*Speaker.

1. Introduction

The phenomenon of *entanglement* is fundamental to quantum mechanics, and one of the features that distinguishes it sharply from classical mechanics. Although sometimes viewed as an exotic phenomenon, entanglement is in fact generic and ubiquitous in quantum systems. Given a quantum system in some state, if there is a natural way to divide the system into two parts, then it is natural to ask the question, “How entangled are they?” This turns out to be a very fruitful question to ask in any area of physics where quantum mechanics plays a role, including the study of quantum many-body systems, quantum field theories, and even quantum gravity. The study of entanglement therefore sits at the intersection of three different areas of theoretical physics: quantum information theory, condensed matter theory, and high-energy theory.

The aim of these notes is to give a basic introduction to some of the most important and interesting ideas in the study of entanglement in quantum field theories, including both general quantum field theories (in section 3) and holographic ones (in section 4). No attempt is made to be comprehensive or to bring the reader “up to date”. Since the target audience is students with a background in high-energy theory, who may not have studied classical or quantum information theory, we begin in section 2 with a brief review of the most relevant aspects of those subjects. On the other hand, for the same reason, the reader is assumed to have a working knowledge of quantum field theory and holography.

There are many places to learn more about this subject. A discussion of entanglement and von Neumann entropy can be found in any of the many textbooks on quantum information theory, such as the one by Nielsen and Chuang [1], as well as many articles such as the classic review by Wehrl [2] and the recent introduction by Witten [3]. Review articles on entanglement entropy in field theory and holography include those by Casini and Huerta [4], Calabrese and Cardy [5], Nishioka, Ryu, and Takayanagi [6], and Nishioka [7]. There is also the book devoted to holographic entanglement entropy by Rangamani and Takayanagi [8]. Finally, chapter 27 of the book by Nastase [9] covers holographic entanglement entropy.

2. Entropy and entanglement

2.1 Shannon entropy

We start with a little bit of classical information theory, because many of the concepts that carry over to the quantum case can be understood more simply here.

We consider a classical system with a discrete state space labelled by a . Although classical systems usually have continuous state (or phase) spaces, we will stick to the discrete case both to avoid certain mathematical complications and because that will provide the relevant math when we turn to the quantum case.

Our knowledge (or lack thereof) of the state of the system is described by a probability distribution \vec{p} , where

$$p_a \geq 0, \quad \sum_a p_a = 1. \quad (2.1)$$

(For example, the system could be a message about which we have partial information.) The expectation value of an observable \mathcal{O}_a (or other function of a ; an observable is a function that does

not depend on \vec{p}) on the state space is

$$\langle \mathcal{O}_a \rangle_{\vec{p}} = \sum_a \mathcal{O}_a p_a. \quad (2.2)$$

The *Shannon entropy* [10] is

$$S(\vec{p}) := \langle -\ln p_a \rangle_{\vec{p}} = -\sum_a p_a \ln p_a \quad (2.3)$$

(where $0 \ln 0$ is defined to be 0). The entropy detects indefiniteness of the state in the following sense:

$$S(\vec{p}) = 0 \Leftrightarrow p_a = \delta_{aa_0} \text{ for some } a_0 \Leftrightarrow \text{all observables have definite values,} \\ \text{otherwise } S(\vec{p}) > 0. \quad (2.4)$$

(By ‘‘definite value’’ we mean vanishing variance, i.e. $\langle \mathcal{O}_a^2 \rangle = \langle \mathcal{O}_a \rangle^2$.)

More than just detecting indefiniteness, the Shannon entropy is a quantitative measure of the *amount* of our ignorance, in other words of how much information we’ve gained if we learn the state. This assertion is backed up by two basic facts.

First, the Shannon entropy is *extensive*. This means that if A, B are *independent*, so that the joint distribution \vec{p}_{AB} is

$$\vec{p}_{AB} = \vec{p}_A \otimes \vec{p}_B, \quad \text{i.e. } (p_{AB})_{ab} = (p_A)_a (p_B)_b, \quad (2.5)$$

then the entropies add:

$$S(\vec{p}_{AB}) = S(\vec{p}_A) + S(\vec{p}_B). \quad (2.6)$$

In particular, the total entropy of N independent copies of A with identical distributions is $NS(\vec{p}_A)$.

Second, the *Shannon noiseless coding theorem* states that the state of our system can be specified using a binary code requiring on average $S(\vec{p})/\ln 2$ bits. For example, for a 3-state system with $p_1 = 1/2, p_2 = p_3 = 1/4$, then $S(\vec{p}) = (3/2)\ln 2$, the following coding uses $3/2$ bits on average:

$$1 \mapsto 0, \quad 2 \mapsto 10, \quad 3 \mapsto 11. \quad (2.7)$$

(More precisely, the theorem states that the state of N independent copies of the system, all with the same distribution, can be coded with $NS(\vec{p})/\ln 2$ bits in the limit $N \rightarrow \infty$.) Thus, all forms of information are interconvertible, as long as the entropies match. The impressive thing about this theorem is that, once we’ve computed $S(\vec{p})$, we know many bits we will need *without having to actually construct a coding* (and, if we do construct a coding, we know how far it is from being optimal). The ability to straightforwardly calculate a quantity of information-theoretic interest is unfortunately the exception rather than the rule: many interesting quantities in information theory are simply *defined* as the optimal rate at which some task can be performed, and there is usually no practical way to calculate them. As physicists, we are mostly interested in quantities that, like the Shannon entropy, are *calculable*.

2.2 Joint distributions

Above we saw that the entropy is additive for independent systems. Now we will consider a general joint probability distribution \vec{p}_{AB} , and ask what we can learn from the Shannon entropy.

The *marginal* distribution \vec{p}_A is defined by

$$(p_A)_a = \sum_b (p_{AB})_{ab} \quad (2.8)$$

and likewise for \vec{p}_B . (Note that, in the special case of independent systems, $\vec{p}_{AB} = \vec{p}_A \otimes \vec{p}_B$, the notation is consistent, in the sense that \vec{p}_A is indeed the marginal distribution as defined by (2.8).) \vec{p}_A is the distribution that gives the right expectation value for any observable that depends only on a , i.e. $\mathcal{O}_{ab} = \mathcal{O}_a$:

$$\langle \mathcal{O}_{ab} \rangle_{\vec{p}_{AB}} = \langle \mathcal{O}_a \rangle_{\vec{p}_A}. \quad (2.9)$$

For a state b of B with $(p_B)_b \neq 0$, the conditional probability on A , denoted $\vec{p}_{A|b}$, is

$$(p_{A|b})_a = \frac{(p_{AB})_{ab}}{(p_B)_b}. \quad (2.10)$$

A few lines of algebra reveal that its entropy, averaged over b , can be written in a very simple form:

$$\langle S(\vec{p}_{A|b}) \rangle_{\vec{p}_B} = S(\vec{p}_{AB}) - S(\vec{p}_B). \quad (2.11)$$

From now on, we will abbreviate $S(\vec{p}_{AB})$ as $S(AB)$, etc. The quantity on the right-hand side of (2.11) is called the *conditional entropy*, and denoted $H(A|B)$:

$$H(A|B) := S(AB) - S(B). \quad (2.12)$$

This is how much, on average, you *still* don't know about the state of A even after learning the state of B . Note that, as the expectation value of a non-negative function, it is non-negative:

$$H(A|B) \geq 0. \quad (2.13)$$

This implies that, if $S(AB) = 0$ then $S(B) = 0$, in other words if the full system is in a definite state then so are all of its parts. We emphasize this rather obvious point since it fails in the quantum setting.

The amount of information about A you've gained (on average) from learning the state of B is

$$S(A) - H(A|B) = S(A) + S(B) - S(AB) =: I(A : B). \quad (2.14)$$

$I(A : B)$ is called the *mutual information*. Note that it is symmetric between A and B . By extensivity of the entropy, the mutual information vanishes between independent systems:

$$\vec{p}_{AB} = \vec{p}_A \otimes \vec{p}_B \quad \Rightarrow \quad S(AB) = S(A) + S(B) \quad \Rightarrow \quad I(A : B) = 0. \quad (2.15)$$

(Clearly, if the systems are independent, then learning b tells you nothing about a .) It turns out that in the presence of correlations the mutual information is necessarily positive:¹

$$\vec{p}_{AB} \neq \vec{p}_A \otimes \vec{p}_B \quad \Rightarrow \quad S(AB) < S(A) + S(B) \quad \Rightarrow \quad I(A : B) > 0. \quad (2.16)$$

¹For this and other claims made without proof in this section, the proofs can be found in [2], [1], or most other quantum information textbooks.

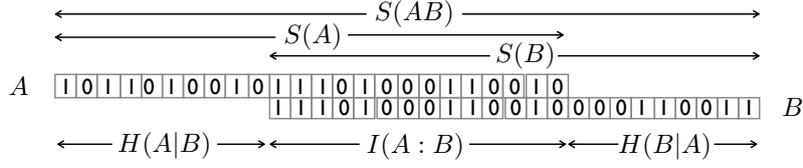


Figure 1: Schematic illustration of the mutual information and conditional entropy in a classical bipartite system: Two correlated systems A and B can be encoded into $S(A)$ and $S(B)$ bits respectively, such that $I(A : B)$ bits of each are perfectly correlated, $H(A|B)$ bits of A are uncorrelated with those of B , and $H(B|A)$ bits of B are uncorrelated with those of A . (Factors of $\ln 2$ have been dropped for clarity.)

For example, for a perfectly correlated pair of bits, with $p_{00} = p_{11} = 1/2$,

$$I(A : B) = \ln 2. \quad (2.17)$$

The inequality

$$S(AB) \leq S(A) + S(B) \quad (2.18)$$

is called *subadditivity* of Shannon entropy.

Correlations can also be detected by correlation functions between observables on A and on B :

$$\langle \mathcal{O}_a \mathcal{O}'_b \rangle_{\vec{p}_{AB}} - \langle \mathcal{O}_a \rangle_{\vec{p}_A} \langle \mathcal{O}'_b \rangle_{\vec{p}_B}. \quad (2.19)$$

The correlation functions vanish for *all* observables $\mathcal{O}_a, \mathcal{O}'_b$ if and only if the systems are independent. In practice, one is unlikely to compute the correlation function for *all* possible observables. The correlations may be *hidden*, in the sense that they may be invisible to the accessible set of observables. The advantage of the mutual information, if it can be computed, is that it detects correlations of any form.

There is a simple picture, shown in fig. 1, which helps to understand the meaning of the conditional entropy and mutual information. AB is coded into $S(AB)/\ln 2$ bits, divided into three groups of $H(A|B)/\ln 2$, $I(A : B)/\ln 2$, and $H(B|A)/\ln 2$ bits respectively. The first two groups together are a coding of A , while the last two are a coding of B . Thus $H(A|B)/\ln 2$ bits belong purely to A , $H(B|A)/\ln 2$ bits belong purely to B , and $I(A : B)/\ln 2$ bits are perfectly correlated between the two systems. (More precisely, this is a coding of N independent identical copies of AB , with the number of bits in each group multiplied by N .) Thus, $I(A : B)$ does not just detect the *existence* of correlation, but quantifies the *amount* of correlation.

2.3 Von Neumann entropy

We now move on to quantum mechanics. One of the many strange features of quantum mechanics is that it does not allow us to separate the notion of “the actual state of the system” from the notion of “our knowledge of the state of the system”. Together, they are encoded in what is simply called a *state*, or *density matrix*², which is an operator ρ with the following properties:

$$\rho^\dagger = \rho, \quad \rho \geq 0, \quad \text{Tr} \rho = 1. \quad (2.20)$$

²“State” is more common in the quantum information literature; “density matrix” is more common in the physics literature, where “state” often implicitly means pure state.

The expectation value of an observable (or other operator) \mathcal{O} is

$$\langle \mathcal{O} \rangle_\rho = \text{Tr}(\mathcal{O}\rho). \quad (2.21)$$

The most “definite” states are the projectors, of the form

$$\rho = |\psi\rangle\langle\psi|. \quad (2.22)$$

Such a state is called *pure*; the rest are *mixed*. Whereas, in classical mechanics, in a definite state all observables have definite values, in quantum mechanics by the uncertainty principle even in a pure state not all observables have definite values.

More generally, based on (2.20), ρ can be diagonalized,

$$\rho = \sum_a p_a |a\rangle\langle a|, \quad (2.23)$$

and its eigenvalues p_a form a probability distribution. The Shannon entropy of this distribution,

$$S(\rho) := S(\vec{p}) = -\text{Tr} \rho \ln \rho = \langle -\ln \rho \rangle_\rho, \quad (2.24)$$

is the *von Neumann entropy*. From (2.4), we have $S(\rho) \geq 0$, and $S(\rho) = 0$ if and only if ρ is pure. Thus the von Neumann entropy is a diagnostic of mixedness. An important property, obvious from its definition, is invariance under unitary transformations:

$$S(U\rho U^{-1}) = S(\rho). \quad (2.25)$$

When we consider joint systems, much but—as we will see—not all of the classical discussion carries through. The Hilbert space for the joint AB system is the tensor product of the respective A , B Hilbert spaces,

$$\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B. \quad (2.26)$$

Given a state ρ_{AB} , the effective state on A which reproduces the expectation value of an arbitrary observable of the form $\mathcal{O}_A \otimes I_B$ is given by the partial trace on B :

$$\rho_A := \text{Tr}_B \rho_{AB} \quad (2.27)$$

(where Tr_B is shorthand for $\text{Tr} \mathcal{H}_B$). This is called the *marginal state* or *reduced density matrix* (the former is more common in the QI literature, and the latter in the physics literature). Again, if ρ_{AB} is a product state, $\rho_{AB} = \rho_A \otimes \rho_B$, then all correlation functions of operators on A and B vanish:

$$\langle \mathcal{O}_A \otimes \mathcal{O}_B \rangle_{\rho_{AB}} - \langle \mathcal{O}_A \rangle_{\rho_A} \langle \mathcal{O}_B \rangle_{\rho_B} = 0; \quad (2.28)$$

in this case we say that A and B are uncorrelated.

Like the Shannon entropy, the von Neumann entropy is extensive and subadditive:

$$\rho_{AB} = \rho_A \otimes \rho_B \Leftrightarrow S(AB) = S(A) + S(B) \quad (2.29)$$

$$\rho_{AB} \neq \rho_A \otimes \rho_B \Leftrightarrow S(AB) < S(A) + S(B) \quad (2.30)$$

Hence the mutual information, defined again by

$$I(A : B) := S(A) + S(B) - S(AB), \quad (2.31)$$

detects correlation. It can also be viewed as a quantitative measure of the amount of correlation. This is more subtle than in the classical case, since, as we will see, in general there is no simple coding of the state like the one described in fig. 1. However, there are two facts that support the interpretation of the mutual information as the amount of correlation. The first is that it bounds correlators:

$$\frac{1}{2} \left(\frac{\langle \mathcal{O}_A \otimes \mathcal{O}_B \rangle - \langle \mathcal{O}_A \rangle \langle \mathcal{O}_B \rangle}{\|\mathcal{O}_A\| \|\mathcal{O}_B\|} \right)^2 \leq I(A : B), \quad (2.32)$$

where $\|\cdot\|$ is the operator norm [11, 12]. (Unfortunately, this bound is not so useful in the field-theory context, where the observables of interest are rarely bounded operators.) The second is that it is non-decreasing under adjoining other systems to A or B :

$$I(A : BC) \geq I(A : B). \quad (2.33)$$

In terms of the entropy, this inequality becomes

$$S(AB) + S(BC) \geq S(B) + S(ABC), \quad (2.34)$$

called *strong subadditivity* (SSA) [13]. SSA is a cornerstone of quantum information theory.

The entropies of subsystems also obey two other useful inequalities:

$$S(AB) \geq |S(A) - S(B)|, \quad (2.35)$$

called the *Araki-Lieb* inequality, and

$$S(AB) + S(BC) \geq S(A) + S(C), \quad (2.36)$$

called either the *second form of SSA* or *weak monotonicity*. A useful special case of Araki-Lieb occurs when AB is pure; then it implies

$$S(A) = S(B). \quad (2.37)$$

For any ρ_A , we can define a formal ‘‘Hamiltonian’’ H_A , called the *modular Hamiltonian*, with respect to which it is a Gibbs state:

$$\rho_A = \frac{1}{Z} e^{-H_A}. \quad (2.38)$$

Here the ‘‘temperature’’ is by convention set to 1. H_A is defined up to an additive constant. (If ρ_A has vanishing eigenvalues, then H_A has infinite eigenvalues.) From (2.24), we have

$$S(A) = \langle H_A \rangle + \ln Z. \quad (2.39)$$

(Note that the expectation value can be evaluated either on A or on the full system.) The fact that any subsystem can be thought of as being a canonical ensemble allows us to employ our intuition from statistical physics. As we will see, there are also a few interesting cases where H_A can be written in closed form and $S(A)$ can be evaluated using (2.39).

2.4 Rényi entropies

The *Rényi entropies* (sometimes called α -entropies) S_α are a very useful one-parameter generalization of the Shannon entropy, defined for $\alpha \in [0, \infty]$. Since we will apply them in the quantum context, we will define them directly in terms of ρ :

$$S_\alpha := \frac{1}{1-\alpha} \ln \text{Tr} \rho^\alpha = \frac{1}{1-\alpha} \ln \left(\sum_a p_a^\alpha \right) \quad (\alpha \neq 0, 1, \infty) \quad (2.40)$$

$$S_0 := \lim_{\alpha \rightarrow 0} S_\alpha = \ln(\text{rank} \rho) \quad (2.41)$$

$$S_1 := \lim_{\alpha \rightarrow 1} S_\alpha = S \quad (2.42)$$

$$S_\infty := \lim_{\alpha \rightarrow \infty} S_\alpha = -\ln \|\rho\| = -\ln(\max_a p_a). \quad (2.43)$$

S_α , for any fixed α , shares several important properties with S . First, it is unitarily invariant. Second, it is non-negative and vanishes if and only if ρ is pure. Therefore it detects mixedness.³ Third, it is extensive. Therefore the Rényi mutual information,

$$I_\alpha(A : B) := S_\alpha(A) + S_\alpha(B) - S_\alpha(AB), \quad (2.44)$$

detects correlations, in the sense that if it is non-zero then A and B are necessarily correlated. However, the Rényi entropy is not subadditive (except for $\alpha = 0, 1$), so the converse does not hold; $I_\alpha(A : B)$ can be zero, or even negative, in the presence of correlations. Related to this, S_α does not obey strong subadditivity (except for $\alpha = 1$), so $I_\alpha(A : B)$ can decrease under adjoining other systems to A or B . For these reasons (and another one given below), $I_\alpha(A : B)$ is a poor quantitative measure of the amount of correlation.

S_α also has some useful properties as a function of α (for fixed ρ), for example

$$\frac{dS_\alpha}{d\alpha} \leq 0, \quad \frac{d^2 S_\alpha}{d\alpha^2} \geq 0. \quad (2.45)$$

Our interest in the Rényi entropies stems mainly from the fact that some of them are much more easily computed than the von Neumann entropy. A direct computation of the von Neumann entropy its definition requires an explicit representation of the operator $\ln \rho$, which essentially requires diagonalizing ρ . On the other hand, computing $\text{Tr} \rho^2$, $\text{Tr} \rho^3$, . . . , and therefore S_2 , S_3 , . . . , is relatively straightforward in any basis. What can we learn from these Rényis? First, as discussed above, we learn whether ρ is pure, and in the case of joint systems, we may learn whether they are correlated. Second, as we will discuss below, on physical grounds we expect S_α to be an analytic function of α for sufficiently “well-behaved” systems. Therefore, we may be able extrapolate S_α from its values for $\alpha = 2, 3, \dots$ to its value for general α by fitting to an analytic function of α . By then taking the limit $\alpha \rightarrow 1$, we learn the von Neumann entropy. In many cases, especially in field theories, this is the most practical analytic route to the von Neumann entropy [14]. Knowing S_α for all α also in principle determines, via an inverse Laplace transform, the full spectrum of ρ .

³ S_α is also a good quantitative measures of mixedness in the sense that it increases under mixing, in other words replacing ρ by $\sum_i p_i U_i^\dagger \rho U_i$, where $\{p_i\}$ is a probability distribution and the U_i 's are unitaries. See [2] for details.

Let us consider two examples. First, if ρ is proportional to a rank- n projector P ,

$$\rho = \frac{1}{n}P, \quad (2.46)$$

then the Rényis are independent of α :

$$S_\alpha = \ln n. \quad (2.47)$$

This would apply, for example, to the microcanonical ensemble

$$\rho = \frac{1}{n(E)} \sum_{E \leq E_a \leq E + \Delta E} |a\rangle\langle a|, \quad (2.48)$$

(where the $|a\rangle$ are energy eigenstates), in which case

$$S_\alpha = S(E) := \ln n(E), \quad (2.49)$$

where $S(E)$ is the microcanonical entropy.

In the canonical ensemble, or Gibbs state,

$$\rho = \frac{1}{Z_\beta} e^{-\beta H}. \quad (2.50)$$

we find instead

$$S_\alpha = \frac{\alpha\beta}{1-\alpha} (F_\beta - F_{\alpha\beta}) \quad (\alpha \neq 0, 1, \infty), \quad (2.51)$$

where Z_β and F_β are the partition function and free energy respectively at temperature $1/\beta$.

These two examples illustrate an important fact about Rényi entropies, which explains why they are not very familiar to most physicists: they are not thermodynamic quantities. By this we mean the following. In a thermodynamic limit, with N degrees of freedom, extensive quantities (such as the free energy and von Neumann and Rényi entropies) are, to leading order in $1/N$, proportional to N . Furthermore, for certain quantities, this leading term depends only on the macroscopic state; it is independent of the particular ensemble—microcanonical, canonical, grand canonical, etc.—used to describe that state. These leading parts are the thermodynamic quantities, the ones we learn about in high-school physics. The subleading ($O(1)$ etc.) terms, meanwhile, are sensitive to statistical fluctuations and therefore depend on the choice of ensemble. From (2.49) and (2.51), we immediately see that the leading term in $S_\alpha(\rho)$ for $\alpha \neq 1$ is of order N but depends on the ensemble, and is therefore *not* a thermodynamic quantity.

We can say this another way: Thermodynamics is related to statistical mechanics essentially by a saddle-point approximation. The operations of calculating a Rényi entropy and making a saddle-point approximation do not commute, since taking the α power of ρ will typically shift the saddle-point. This can be seen in (2.51), where S_α depends on the free energy at the temperature $1/(\alpha\beta)$, rather than the system's actual temperature $1/\beta$. The shift in the saddle-point has several implications: (1) for fixed α , S_α will have phase transitions at different values of the temperature and other parameters than the actual system; (2) for fixed parameters, S_α may have phase transitions (i.e. be non-analytic) as a function of α ; (3) for joint systems that are correlated only via their

fluctuations (for example, a cylinder with two chambers separated by a piston), such that $I(A : B)$ is $O(1)$, $I_\alpha(A : B)$ is typically $O(N)$ —again, a poor measure of the amount of correlation.⁴

All of which is to say that, when working in a thermodynamic limit, the Rényis should be treated with extreme caution.

2.5 Entanglement

So far, we have emphasized the parallels between the concepts of entropy in the classical and quantum settings. Now we will discuss some of their crucial differences. We will start by consider the *pure* states of a bipartite system AB , where these differences are dramatically illustrated. The analogous states in the classical setting are, as we saw, quite boring: a definite state of a bipartite system is necessarily of the form $\vec{p}_{AB} = \vec{p}_A \otimes \vec{p}_B$, with no correlations; moreover, each marginal state \vec{p}_A, \vec{p}_B is necessarily itself definite.

In contrast, consider a *Bell pair*, a two-qubit system in the following pure state:

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}} (|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B). \quad (2.52)$$

This state clearly cannot be factorized, i.e. written in the form

$$|\psi\rangle_{AB} = |\psi'\rangle_A \otimes |\psi''\rangle_B \quad (2.53)$$

for any $|\psi'\rangle_A, |\psi''\rangle_B$. Furthermore, its marginals,

$$\psi_A = \text{Tr}_B |\psi\rangle_{AB} \langle \psi|_{AB} = \frac{1}{2} (|0\rangle_A \langle 0|_A + |1\rangle_A \langle 1|_A), \quad (2.54)$$

$$\psi_B = \text{Tr}_A |\psi\rangle_{AB} \langle \psi|_{AB} = \frac{1}{2} (|0\rangle_B \langle 0|_B + |1\rangle_B \langle 1|_B) \quad (2.55)$$

are mixed!

A pure state that cannot be factorized is called *entangled*. As mentioned at the very start of these lectures, entanglement is one of the key features that distinguishes quantum from classical mechanics. It is at the root of many seemingly exotic quantum phenomena, such as violation of the Bell inequalities, and it plays a crucial role in essentially all potential quantum technologies, including quantum cryptography and quantum computation.

The example of the Bell pair suggests a connection between entanglement and mixedness of the two subsystems. For a factorized state, (2.53), the subsystems are clearly pure:

$$\psi_A = |\psi'\rangle_A \langle \psi'|_A, \quad \psi_B = |\psi''\rangle_B \langle \psi''|_B. \quad (2.56)$$

The converse is easily established. Using the singular value decomposition, an arbitrary pure state can be written in the form

$$|\psi\rangle_{AB} = \sum_{ab} \sqrt{p_a} |a\rangle_A \otimes |a\rangle_B, \quad (2.57)$$

where $\{|a\rangle_A\}, \{|a\rangle_B\}$ are orthonormal sets of vectors in \mathcal{H}_A and \mathcal{H}_B respectively, and $\{p_a\}$ is a probability distribution. (2.57) is called the *Schmidt decomposition*. (It is unique up to a unitary acting jointly on any subsets $|a\rangle_A, |a\rangle_B$ with degenerate p_a s.) The marginals are then

$$\psi_A = \sum_a p_a |a\rangle_A \langle a|_A, \quad \psi_B = \sum_a p_a |a\rangle_B \langle a|_B. \quad (2.58)$$

⁴See [15] for further discussion.

Thus ψ_A is pure if and only if ψ_B is pure, and this is true if and only if $|\psi\rangle_{AB}$ is factorized.

We can therefore use any Rényi entropy $S_\alpha(A)$ ($= S_\alpha(B)$ by (2.58)) to detect entanglement between A and B . For this reason, physicists often refer to the entropies of subsystems as *entanglement entropies* (EE)—even when the full system is *not* (as we have been discussing here) in a pure state, and therefore $S(A)$ does *not* necessarily reflect entanglement! For this and other reasons, this terminology is not ideal, and it is generally frowned upon by quantum information theorists. Nonetheless, since it is common in the physics literature, we will follow it to some extent in these notes.

The von Neumann entropy $S(A)$ not only detects entanglement, but is also a quantitative measure of the *amount* of entanglement, in the following sense. A unitary acting only on A or B cannot change $S(A)$ or $S(B)$, and therefore cannot turn a factorized state into an entangled one. More generally, no combination of *local operations* on A and B and *classical communication* between A and B (abbreviated LOCC) can create a pure entangled state from a factorized one. (We will consider the mixed case below.) Therefore, if A and B are separated, creating entanglement requires actually sending some physical system between them (or from a third party). Therefore, entanglement between separated systems should be viewed as a scarce resource. Now, LOCC can transform entangled states into other entangled states, for example a general entangled state $|\psi\rangle_{AB}$ into a set of Bell pairs. More precisely, N identical copies of $|\psi\rangle_{AB}$ can be transformed into $NS(A)/\ln 2$ Bell pairs and vice versa, in the limit $N \rightarrow \infty$. This is very similar in spirit to the Shannon noiseless coding theorem. It shows that different forms of entanglement are interconvertible under LOCC, as long as the *amount* of entanglement matches, as quantified by the EE. As with the Shannon theorem, the utility of this statement derives in large part to the fact that the entropy is an independently defined, and often computable, quantity.

The existence of entangled states is fairly obvious mathematically from the definition of a tensor product of Hilbert spaces. Nonetheless, it is worth stepping back to ask how it possible physically for a state to be pure with respect to the whole system, yet mixed with respect to a subsystem. As we emphasized above, a classical state that is definite with respect to the whole system has definite values for *all* observables, including all the observables on the subsystem; therefore, it must be definite with respect to the subsystem. The point where this logic breaks down in quantum mechanics is that, by the uncertainty principle, even in a pure state only some observables have definite values. In particular, the observables necessary to establish that the state is pure may not reside entirely in the subalgebra of A -observables. If this is true, then to an A -observer, the state is effectively mixed.

2.6 Purification

Now suppose we have a single system A , and a mixed state ρ_A on it. Whatever the physical origin of ρ_A , we can always mathematically construct a larger system AB and a pure state $|\psi\rangle_{AB}$ on it, such that $\psi_A = \rho_A$. This is called *purification*, and it is accomplished by reverse-engineering the Schmidt decomposition. Write ρ_A in diagonal form,

$$\rho_A = \sum_a p_a |a\rangle_A \langle a|_A \quad (2.59)$$

(where $p_a \neq 0$). Let \mathcal{H}_B be a Hilbert space whose dimension is at least the rank of ρ_A (for example, it can be a copy of \mathcal{H}_A), and let $\{|a\rangle_B\}$ be an orthonormal set of vectors. Then

$$|\psi\rangle_{AB} := \sum_a \sqrt{p_a} |a\rangle_A \otimes |a\rangle_B \quad (2.60)$$

is the desired purification.

Purification shows that mixedness of any form is indistinguishable from that arising from entanglement. Thus, “entanglement” and “entropy” are effectively interchangeable concepts (making the phrase “entanglement entropy” redundant; quantum information theorists just call it “entropy”). Thinking of entropy as resulting from entanglement marks a radical change from the classical conception of entropy as fundamentally a measure of ignorance. In classical mechanics, you can never “unmix” a system by adjoining another one.

Purification is also a handy technical device. For example, by purifying the AB system, the subadditivity inequality $S(AB) \leq S(A) + S(B)$ is seen to imply the Araki-Lieb inequality (2.35) and vice versa; similarly, the two forms of SSA (2.34), (2.36) are equivalent.

As an example of purification, consider the Gibbs state

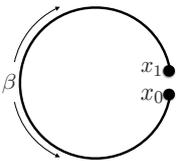
$$\rho_A = \frac{1}{Z} e^{-\beta H_A} = \frac{1}{Z} \sum_a e^{-\beta E_a} |a\rangle_A \langle a|_A, \quad (2.61)$$

where $|a\rangle_A$ are the energy eigenstates. Then we can choose \mathcal{H}_B to be a copy of \mathcal{H}_A ; the purification is

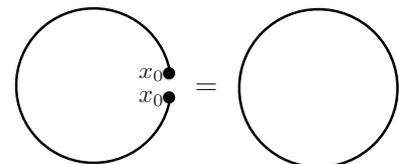
$$|\psi\rangle_{AB} = \frac{1}{\sqrt{Z}} \sum_a e^{-\beta E_a/2} |a\rangle_A \otimes |a\rangle_B. \quad (2.62)$$

This is called the *thermofield double* (TFD) state.

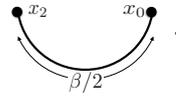
The TFD state can be represented in terms of a Euclidean path integral as a semicircle of length $\beta/2$. It is worth going over this in detail, as it provides good practice for the path-integral manipulations we will do in field theories. We start by noting that the operator $e^{-\beta H_A}$ can be represented by a Euclidean path integral on an interval of length β ; concretely, its matrix element between position eigenstates is given by

$$\langle x_0 | e^{-\beta H_A} | x_1 \rangle_A = \int_{x_0}^{x_1} \mathcal{D}x \quad (2.63)$$


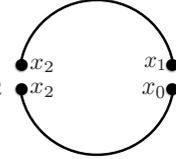
where we’ve chosen to draw the interval as an open circle, and the path integral is evaluated with boundary conditions $x = x_0$ and $x = x_1$ on the two endpoints. The partition function is the trace of this operator:

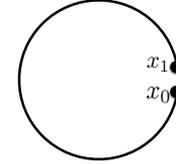
$$Z = \text{Tr} e^{-\beta H_A} = \int dx_0 \langle x_0 | e^{-\beta H_A} | x_0 \rangle_A = \int dx_0 \quad (2.64)$$


integrating over the boundary conditions “sews together” the endpoints, leaving a circle of circumference β . The TFD state (2.62) is then given in terms of a path integral as a semicircle of length $\beta/2$, meaning that its components in the position basis are given as follows:

$$\langle x_0|_A \otimes \langle x_2|_B |\psi\rangle_{AB} = \frac{1}{\sqrt{Z}} \times \text{Diagram} \quad (2.65)$$


To check that this reproduces the Gibbs state on A when traced over B , we take the outer product with the Hermitian conjugate $\langle \psi|_{AB}$, represented by an upper semicircle, and sew together the two B ends:

$$\langle x_0|_A (\text{Tr}_B |\psi\rangle_{AB} \langle \psi|_{AB}) |x_1\rangle_A = \frac{1}{Z} \int dx_2 \text{Diagram} \quad (2.66)$$


$$= \frac{1}{Z} \times \text{Diagram} = \langle x_0|_A \rho_A |x_1\rangle_A \quad (2.67)$$


2.7 Entangled vs. separable states

The discussion in subsection 2.5 assumed that the joint system AB is in a pure state and explained that $S(A)$ is a measure of the amount of entanglement between A and B . It is clear that this interpretation is lost once we allow AB to be a mixed state. For example, $S(A)$ may be non-zero even in a product state $\rho_{AB} = \rho_A \otimes \rho_B$. (Consider, for example, the Gibbs state of two non-interacting systems.) This is another reason why the phrase “entanglement entropy” is misleading.

Can we nonetheless define what it means for A and B to be entangled, and quantify the amount of entanglement, when AB is in a mixed state? First, we define a *separable*, or *classically correlated*, state as a convex mixture of product states:

$$\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i, \quad (2.68)$$

where $\{p_i\}$ is a probability and ρ_A^i, ρ_B^i are states. This is precisely the class of states that can be obtained by LOCC starting from a product state. A state that is not separable is then considered *entangled*. It is easy to see that this definition reduces to the previous one in the case that ρ_{AB} is pure.

As noted in subsection 2.3, the mutual information $I(A : B)$ quantifies the amount of correlation. This includes both classical correlation and entanglement. For example, a maximally classically correlated pair of qubits

$$\rho_{AB} = \frac{1}{2} (|00\rangle\langle 00| + |11\rangle\langle 11|) \quad (2.69)$$

has $I(A : B) = \ln 2$, while a Bell pair has $I(A : B) = 2 \ln 2$ (thereby confirming the intuition that entanglement is a stronger form of correlation than classical correlation). However, it does not

distinguish between the two forms of correlation; for example, two classically correlated pairs of qubits have the same mutual information as a single Bell pair.

How then can we tell whether a given state is separable or entangled? This turns out to be a very hard (in fact, NP-hard) problem. Here we will discuss two tests for entanglement that are useful in the field-theory context. First, we saw that, for classical probability distributions,

$$H(A|B) \geq 0. \quad (2.70)$$

In fact, (2.70) holds generally for separable states. However, it does not always hold in the presence of entanglement; for example, for an entangled pure state,

$$H(A|B) = -S(B) < 0. \quad (2.71)$$

Therefore, negativity of the conditional entropy can be used to test for the presence of entanglement. However, the test has “false negatives”; it’s easy to construct entangled states with positive conditional entropy.

Another calculable test for entanglement involves the partial transpose [16, 17]. This is based on the fact that the transpose of a density matrix is also a density matrix. (The transpose depends on the basis, but this is true in any given basis.) Suppose we define $\tilde{\rho}$ to be the operator obtained by transposing ρ_{AB} only on the B indices:

$$(\tilde{\rho})_{ab,a'b'} := (\rho_{AB})_{ab',a'b}. \quad (2.72)$$

If ρ_{AB} is separable, (2.68), we have

$$\tilde{\rho} = \sum_i p_i \rho_A^i \otimes \rho_B^{iT}. \quad (2.73)$$

For all i , ρ_B^{iT} is still a density matrix, so $\tilde{\rho}$ is a density matrix as well. However, if ρ_{AB} is entangled, then nothing guarantees that $\tilde{\rho}$ is a density matrix: while $\tilde{\rho}$ is Hermitian and has trace 1, it may not be positive. Therefore the presence of negative eigenvalues of $\tilde{\rho}$ establishes that ρ_{AB} is entangled. (While $\tilde{\rho}$ depends on the choice of basis used in (2.72), its spectrum does not.) This is called the *positive partial trace* criterion. In particular, if $\tilde{\rho}$ has any negative eigenvalues, then $\text{Tr}|\tilde{\rho}| \neq \text{Tr}\tilde{\rho} = 1$; hence the *logarithmic negativity* (or *entanglement negativity*) $\ln \text{Tr}|\tilde{\rho}|$ detects entanglement [18]. However, like the conditional entropy, it is not reliable detector, in the sense that an entangled state can have vanishing logarithmic negativity. These are the only two mixed-state entanglement detectors that have been calculated in field theories.

Quantifying the *amount* of entanglement is similarly hard, and there exists a large body of theory around this problem. Unlike in the pure-state case, there is no single, canonical measure of entanglement; rather, many different measures of entanglement have been defined, which are useful in different operational and theoretical senses. These include:

- the *entanglement of formation*, the number of Bell pairs required to create ρ_{AB} by LOCC;
- the *distillable entanglement*, the number of Bell pairs that can be created out of ρ_{AB} by LOCC;

- the *squashed entanglement*, the minimum of $I(A : B|C)$ over ρ_{ABC} such that $\text{Tr}_C \rho_{ABC} = \rho_{AB}$, where the $I(A : B|C) := S(AC) + S(BC) - S(C) - S(ABC)$ is the *conditional mutual information*;
- the *relative entropy of entanglement*, the minimum of $S(\rho_{AB} || \sigma_{AB})$ over separable states σ_{AB} , where $S(\rho || \sigma) := \text{Tr}(\rho \ln \rho - \rho \ln \sigma)$ is the *relative entropy*;
- and the *logarithmic negativity*, defined above.

See [19] for an overview and references. Again, except for the logarithmic negativity, they are hardly ever computable in practice, so we will not dwell on them.

2.8 Entanglement, decoherence, and monogamy

Although it is sometimes depicted as an exotic phenomenon, entanglement is the rule rather than the exception. Mathematically, given a bipartite system, factorized states are a measure-zero subset of the pure states, and separable states are measure-zero subset of the mixed states. Physically, interactions between two systems nearly always lead to entanglement.

One of the most important physical implications of this is that interactions with the environment cause *decoherence* of a system's state. Consider, for example, a spin-1/2 system A interacting with some environmental degree of freedom B . Label the initial state of E is $|0\rangle_B$, and suppose the interaction causes E to change to $|1\rangle_B$ if A 's spin is down:

$$|\uparrow\rangle_A \otimes |0\rangle_B \rightarrow |\uparrow\rangle_A \otimes |0\rangle_B, \quad |\downarrow\rangle_A \otimes |0\rangle_B \rightarrow |\downarrow\rangle_A \otimes |1\rangle_B. \quad (2.74)$$

It would appear that the state of A is not affected. Suppose, however, that A starts in an arbitrary pure state:

$$\alpha|\uparrow\rangle + \beta|\downarrow\rangle. \quad (2.75)$$

After interaction with B , this becomes

$$\alpha|\uparrow\rangle \otimes |0\rangle_B + \beta|\downarrow\rangle \otimes |1\rangle_B. \quad (2.76)$$

If the environmental degree of freedom is inaccessible, then ρ_A has undergone the following transformation (in the $|\uparrow\rangle_A, |\downarrow\rangle_A$ basis):

$$\begin{pmatrix} |\alpha|^2 & \alpha^* \beta \\ \alpha \beta^* & |\beta|^2 \end{pmatrix} \rightarrow \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix}. \quad (2.77)$$

Assuming neither α nor β is zero, the state has become mixed! In general, entanglement with inaccessible environmental degrees of freedom causes the off-diagonal matrix elements of ρ_A (in a basis determined by the interactions) to decrease in magnitude, which increases $S(A)$.⁵

⁵It is important to note that entanglement alone does not produce decoherence; rather, the environmental degrees of freedom need to be both inaccessible and physically independent. Slave degrees of freedom, such as the fast modes in a system with a separation of energy or time scales (such as the electrons in a Born-Oppenheimer approximation, or field-theory modes that have been integrated out in a Wilsonian scheme), do not decohere the slow modes (the nuclei, or the fields kept), although the fast and slow modes are entangled. For example, the quarks may be inaccessible to a low-energy nuclear physicist, but they do not decohere the state of the nucleons. For this reason, it is not generally useful to think of the slow modes as being in a mixed state.

Environmental decoherence is the basic reason it's harder to build a quantum than a classical computer: not only must the qubits be protected from being disturbed by the environment (as the bits in a classical computer must be), but they must also be protected from *influencing* the environment, even microscopic environmental degrees of variable (stray photons, etc.).

It has also been argued that environmental decoherence explains why the world appears to be classical (i.e. lack superpositions) to macroscopic observers such as ourselves, and that decoherence due to interactions between a system and measurement apparatus is behind the apparent collapse of the wave function upon measurement [20].

Paradoxically, while entanglement causes decoherence, decoherence destroys entanglement. Consider, for example, the following pure state on three qubits, called the *GHZ state*:

$$|\text{GHZ}\rangle_{ABC} = \frac{1}{\sqrt{2}} (|000\rangle_{ABC} + |111\rangle_{ABC}). \quad (2.78)$$

Are A and B entangled? At first sight, it may look like they are. But, as we emphasized above, this is really a question about ρ_{AB} , which is given by the separable state (2.69). So, while A is entangled with BC , it is *not* entangled with B , due to decoherence by C . Thus, although ubiquitous, entanglement is also rather fragile.

More generally, entanglement within a system excludes entanglement (or even classical correlation) with other systems. This property is called *monogamy of entanglement*. It is quite different from the classical case, where nothing stops A from being simultaneously correlated with B and with BC (in fact, if it is correlated with B then it *must* be correlated with BC). Monogamy of entanglement plays an important role in many aspects of quantum information theory. For example, in quantum cryptography one establishes entanglement between A and B ; monogamy then guarantees that any eavesdropping—which necessarily requires creating some correlation with AB —destroys the entanglement and is therefore detectable. Monogamy also appears in many ideas concerning the black-hole information paradox, such as the firewall argument [21].

2.9 Subsystems, factorization, and time

We noted above that interactions between two systems nearly always lead them to become entangled. Specifically, for a generic Hamiltonian H_{AB} on a joint system coupling A and B , if at time t_0 the system is in a pure, unentangled state, then at another time t_1 they will—absent some fine-tuning—be entangled, just because almost all pure states in the joint Hilbert space \mathcal{H}_{AB} are entangled. In the Schrödinger picture, it is clear that the state evolves from an unentangled to an entangled one, and more generally the entanglement entropy $S(A)$ evolves with time. In the Heisenberg picture (which we will be using in the next section, when we discuss field theories), however, this is confusing: the state doesn't evolve, so how can it become entangled? The answer is that the *factorization* of the Hilbert space into $\mathcal{H}_A \otimes \mathcal{H}_B$ evolves! More precisely, instead of (2.26), we really should have written

$$\mathcal{H}_{AB} \cong \mathcal{H}_A \otimes \mathcal{H}_B, \quad (2.79)$$

where \cong means isomorphic. Normally we wouldn't bother with such a fine point, but in the Heisenberg picture the isomorphism is *time-dependent*. Whether a given state $|\psi\rangle_{AB}$ in \mathcal{H}_{AB} is isomorphic to a tensor product $|\psi'\rangle_A \otimes |\psi''\rangle_B \in \mathcal{H}_A \otimes \mathcal{H}_B$ depends on the isomorphism and therefore on the

time. Thus, in the Heisenberg picture, when factorizing a Hilbert space into $\mathcal{H}_A \otimes \mathcal{H}_B$ and computing ρ_A , $S(A)$, etc., we need to specify not only the subsystems A and B but also the *time* at which we are looking at them.

3. Entropy, entanglement, and fields

In this section, our aim will be to understand a few basic facts about entanglement entropies in *relativistic quantum field theories*. In subsection 3.1 we consider field theories in any dimension, but when once we consider concrete examples starting in subsection 3.2, we will mostly restrict ourselves to $(1+1)$ -dimensional field theories. This will suffice for the limited selection of topics we'll cover, although obviously there are many interesting new phenomena in higher dimensions.

3.1 What is a subsystem?

The first thing we need to decide is what kind of *subsystem* we want to consider. Obviously there are many different ways to divide up a Hilbert space, but some of those more natural physically than others. A defining feature of quantum field theories is the existence of local observables living at points in spacetime and obeying the following rule:

$$\text{Local observables at spacelike-separated points commute.} \quad (3.1)$$

If we consider all the local operators on a given time slice, two features stand out: (1) By (3.1), the operators at different points commute. (2) They are complete in the sense that a local operator at any point in the spacetime can be written in terms of them using the (Heisenberg) equations of motion. These two facts suggest that the points on the time slice host independent degrees of freedom. (This intuition is supported by the fact that many field theories can be understood as continuum limits of lattice systems with independent spins or other degrees of freedom on the lattice sites or links.) This suggests that a good way to divide up the system is by *regions on a given time slice*. Letting A be a region on a given slice Σ and $A^c := \Sigma \setminus A$ its complement, the claim is that the Hilbert space admits a decomposition

$$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{A^c} \quad (3.2)$$

in which the local operators on A act only on \mathcal{H}_A and those in A^c act only on \mathcal{H}_{A^c} . For a given state ρ , one can then define the reduced density matrix ρ_A and entropy $S(A)$.

A more careful analysis shows that (3.2) does not *quite* hold in field theories due to subtleties at the boundary between A and A^c , called the *entangling surface*. There is some interesting physics in this non-factorization, especially in gauge theories, but nothing that affects what we will say in the rest of these lectures. Suffice it to say that, with a bit of care, one can nonetheless define the reduced density matrix ρ_A and entanglement entropy $S(A)$. In essence, factorization (3.2) is just a stand-in for spacelike commutation (3.1), which is what really allows one to define ρ_A and $S(A)$.

As discussed in subsection 2.9, in the Heisenberg picture the factorization of the Hilbert space associated to a given subsystem depends on the time. So a region A_0 of the time slice $t = t_0$ and its time translation A_1 to the time $t = t_1$ imply different factorizations of the form (3.2). For this reason, A_0 and A_1 may just as well be thought of as *different subsystems*, rather than the same subsystem

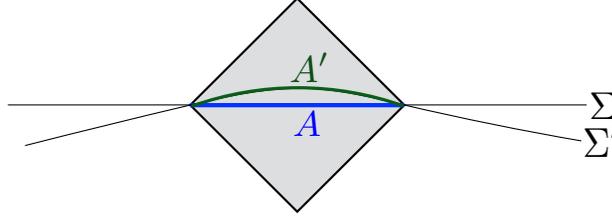


Figure 2: A region A of a Cauchy slice Σ and a region A' of another Cauchy slice Σ' with the same causal domain (shown in gray), $D(A') = D(A)$.

at different times. Indeed, in a different reference frame A_0 and A_1 do not even necessarily cover the same region of space.

Causality in a relativistic field theory has a further implication for the definition of a *subsystem*. To understand this, it is useful to broaden our perspective a bit. So far we have implicitly been working in Minkowski space. More generally, a relativistic field theory can be defined on any *globally hyperbolic* manifold, which is a Lorentzian manifold M that admits a Cauchy slice. A Cauchy slice is the natural generalization to a curved manifold of a “time slice” in Minkowski space. It is a subset Σ of M such that *every inextendible causal curve intersects Σ exactly once*. This means that (assuming the field equations are causal) the fields on Σ provide complete initial data, so that the local observables at any point in M can be written in terms of those on Σ . It also means that distinct points on Σ are spacelike-related (not connected by a causal curve), and therefore the corresponding local operators commute. So in this more general context we associate a subsystem to a region A of a Cauchy slice Σ .

But we can refine this association even more. Given a region A of a Cauchy slice Σ , we define its *causal domain* $D(A)$ as the set of points $p \in M$ such that every inextendible causal curve through p intersects A . By construction, A is a Cauchy slice for $D(A)$, so the latter is a globally hyperbolic spacetime in its own right. Therefore the fields on $D(A)$ define a self-contained field theory, with Hilbert space \mathcal{H}_A . A region A' of a different Cauchy slice Σ' may have the same causal domain, $D(A') = D(A)$ (see fig. 2), in which case $\mathcal{H}_{A'} = \mathcal{H}_A$. For a given state ρ of the full field theory on M , the expectation values of operators in $D(A)$ computed from either ρ_A or $\rho_{A'}$ must agree with those computed from ρ , and therefore with each other, so $\rho_{A'} = \rho_A$. In other words, a *subsystem* is not really associated (as we claimed before) with a region of a Cauchy slice, but rather with the *causal domain* of such a region. This fact plays an important role in all studies of entanglement entropies in relativistic quantum field theories.

3.2 General theory: half-line

The simplest non-trivial subset of the line $t = 0$ we can consider is the half-line:

$$A = \{t = 0, x \geq 0\}. \quad (3.3)$$

(See fig. 3.) Its causal domain is called the *Rindler wedge*, or—since it is a globally hyperbolic spacetime in its own right— $(1+1)$ -dimensional *Rindler space*:

$$D(A) = \{|t| \leq x\}. \quad (3.4)$$

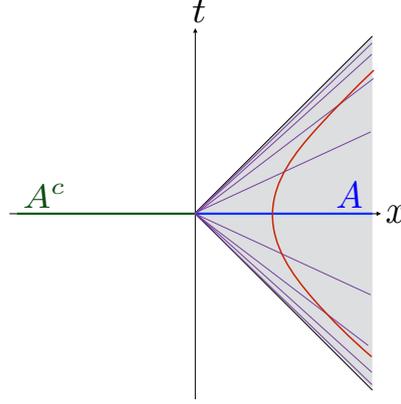


Figure 3: The half-line A (3.3), its complement A^c , and its causal domain $D(A)$, the Rindler wedge (3.4) (shaded), in $(1+1)$ -dimensional Minkowski space. Also shown are some lines of constant χ in the coordinates (χ, r) of (3.13) (purple), each of which is a Cauchy slice for the Rindler wedge, as well as a trajectory of constant r .

This preserves the boost subgroup of the Poincaré group. The entangling surface is just the origin. For the state we choose the vacuum, which is also boost-invariant:

$$\rho = |0\rangle\langle 0|. \quad (3.5)$$

We will see that this state is highly entangled. Therefore, for an observer who spends her whole life inside the Rindler wedge, the world appears to be in a mixed state.

The boost symmetry will allow us to solve this example, in the sense of writing down a closed-form expression for ρ_A in terms of the stress tensor. This will allow us to gain a lot of intuition. Also, the derivation illustrates a very useful technique using Euclidean path integrals.

We formally write the full set of fields as ϕ , and work in a “position” basis $|\phi_0\rangle$, where $\phi_0(x)$ is a field configuration on a fixed time slice. The matrix elements of ρ are given as the $\beta \rightarrow \infty$ limit of the matrix elements of the Gibbs state $e^{-\beta H}/Z$. The latter are calculated as in (2.63) by a Euclidean path integral on a strip of height β (in the Euclidean time direction), which we draw as a cylinder cut along the $\tau = 0$ axis, with boundary conditions dictated by the bra and ket:

$$\langle \phi_0 | e^{-\beta H} | \phi_1 \rangle = \beta \int_{\text{cylinder}} \mathcal{D}\phi \quad (3.6)$$

Now take the limit $\beta \rightarrow \infty$. What is left is the (τ, x) plane cut open along $\tau = 0$:

$$\langle \phi_0 | 0 \rangle \langle 0 | \phi_1 \rangle = \langle \phi_0 | \rho | \phi_1 \rangle = \frac{1}{Z} \times \int_{\text{rectangle}} \mathcal{D}\phi \quad (3.7)$$

Factorizing the left- and right-hand sides, we see that the vacuum wave functional is the path integral over the lower half-plane:

$$\langle \phi | 0 \rangle = \frac{1}{\sqrt{Z}} \times \boxed{\phi = \phi_0} . \quad (3.8)$$

To calculate ρ_A , we trace ρ over \mathcal{H}_{A^c} . This amounts to writing

$$|\phi_0\rangle = |\phi_0^{A^c}\rangle_{A^c} \otimes |\phi_0^A\rangle_A, \quad (3.9)$$

where $\phi_0^{A^c}$ and ϕ_0^A represent the function $\phi_0(x)$ restricted to $x < 0$ and $x \geq 0$ respectively, and summing the matrix element (3.7) over $\phi_0^{A^c}$. The sum over the field value on A^c “glues” the top and bottom sheets together on A^c , leaving a path integral on the plane cut only along A , i.e. the half-line $\tau = 0, x \geq 0$, with boundary conditions given by the field values appearing in the bra and ket:

$$\langle \phi_0^A | {}_A \rho_A | \phi_1^A \rangle_A = \frac{1}{Z} \times \boxed{(0,0) \begin{array}{l} \phi = \phi_1^A \\ \phi = \phi_0^A \end{array}} . \quad (3.10)$$

Looking at the right-hand side of (3.10), it makes sense to change from Cartesian coordinates (τ, x) to polar coordinates (θ, r) , since then the cut plane can be described as $r \in [0, \infty)$, $\theta \in [0, 2\pi]$. Normally, in polar coordinates, for a given r the two points $(\theta = 0, r)$ and $(\theta = 2\pi, r)$ would be identified, but the cut along $\tau = 0, x \geq 0$ makes them distinct. In Euclidean space, we have some freedom in choosing which coordinate we consider the “Euclidean time” coordinate. Let us change our point of view and consider θ , rather than τ , to be the Euclidean time coordinate. From this point of view, the cut plane is simply a time interval of length 2π over the half-line $r \geq 0$. The metric is

$$ds_{\mathbb{E}}^2 = r^2 d\theta^2 + dr^2 . \quad (3.11)$$

If we Wick rotate with respect to θ , writing

$$\theta = i\chi , \quad (3.12)$$

we find the Lorentzian metric

$$ds_{\mathbb{L}}^2 = -r^2 d\chi^2 + dr^2 . \quad (3.13)$$

With $r \geq 0$ and χ running from $-\infty$ to ∞ , this is precisely the Rindler wedge! Thus ρ_A is indeed a state of the field theory on Rindler space.

What state is it? According to (2.63), (3.10) says that ρ_A is a Gibbs state, with $\beta = 2\pi$ and the role of the Hamiltonian being played by the generator of χ -translations, in other words the boost generator K [22]:

$$\rho_A = \frac{1}{Z} e^{-2\pi K} . \quad (3.14)$$

Thus the modular Hamiltonian is

$$H_A = 2\pi K. \quad (3.15)$$

In general, the conserved quantity associated to a Killing vector k^μ is $\int_\Sigma \sqrt{h} k^\mu n^\nu T_{\mu\nu}$, where Σ is a Cauchy slice, h is the induced metric, and n is the unit normal. The Killing vector for boosts is

$$\frac{\partial}{\partial \chi} = x \frac{\partial}{\partial t} + t \frac{\partial}{\partial x}, \quad (3.16)$$

so, choosing $\chi = 0$ as our Cauchy slice for the Rindler wedge, the boost generator is

$$K = \int_0^\infty dr \frac{1}{r} T_{\chi\chi}(r) = \int_0^\infty dx x T_{tt}(x). \quad (3.17)$$

According to (2.62), the thermofield double state is constructed by a path integral in which the Euclidean time runs over an interval $\beta/2$. Here this means that θ runs over an angle π , giving a half-plane. (3.8) shows that this state is nothing but the vacuum. We thus recover the vacuum as the thermofield double of the Rindler state ρ_A , with A^c being the purifying system.

Eq. (3.14) contains a lot of physics. Recall that we are just talking about the vacuum of a field theory. Eq. (3.14) tells us that, to an observer confined to the Rindler wedge, that state appears to be a thermal state with respect to the boost generator (its only continuous symmetry) at a temperature $T_\chi = 1/(2\pi)$. An example of such an observer is one who follows a worldline of constant r (see fig. 3). In the original, inertial coordinates (t, x) , this trajectory is

$$x(t) = \sqrt{r^2 + t^2}, \quad (3.18)$$

which is accelerating with constant proper acceleration $a = 1/r$. To get the physical temperature experienced by such an observer, we have to multiply by the local redshift factor:

$$T_{\text{phys}}(r) = (-g_{\chi\chi})^{-1/2} T_\chi = \frac{1}{2\pi r} = \frac{a}{2\pi}. \quad (3.19)$$

This can also be understood from the fact that the physical inverse temperature β_{phys} is the proper length of the circle of constant r in the Euclidean plane, which is $2\pi r$. The fact that an observer in flat space accelerating at a proper rate a experiences a temperature $a/(2\pi)$ is called the *Unruh effect* [23]. We see that, close to the entangling surface (small r), the fields are very *hot*. Roughly speaking, the observer sees the modes that have been decohered by tracing over A^c , and the closer to the entangling surface, the more UV modes are seen to be decohered. Note that, although the physical temperature is spatially varying, the system is in perfect thermal equilibrium; this is possible due to the spatially varying gravitational potential.

We can use (3.19) to estimate the entanglement entropy $S(A)$ by adding up local thermal entropies. This will give a crude estimate, since the thermal entropy density $s(T)$ is defined in flat space at *constant* temperature, whereas here the temperature is spatially varying. Nonetheless we can get useful intuition from this estimate. For a field of mass m , the entropy density $s(T)$ essentially vanishes at temperatures below m , as the field is frozen out. Thus, much past a distance $r = \xi$, where $\xi := 1/m$ is the correlation length, the field does not contribute significantly to the entropy. In other words, the field is entangled only within a neighborhood $r \lesssim \xi$ of the entangling surface. This very intuitive fact holds in any dimension and for any entangling surface.

On the other hand, for temperatures well above m , by dimensional analysis $s(T) \propto T$. Thus for $r \ll \xi$, the entropy density is diverging like $1/r$. The integral of $1/r$ diverges, so we need to impose a UV cutoff at $r = \varepsilon$. We thus estimate $S(A)$ as follows:

$$S(A) \approx \int_{\varepsilon}^{\infty} dr s(T_{\text{phys}}(r)) \propto \int_{\varepsilon}^{\xi} dr \frac{1}{r} = \ln \frac{\xi}{\varepsilon}. \quad (3.20)$$

We can be more precise by assuming that the theory has a UV fixed point which is a CFT. The entropy density of a CFT with central charge c is

$$s(T) = \frac{2\pi c}{6} T, \quad (3.21)$$

so we get

$$S(A) \approx \int_{\varepsilon}^{\infty} dr s(T_{\text{phys}}(r)) \approx \int_{\varepsilon}^{\xi} dr \frac{c}{6r} = \frac{c}{6} \ln \frac{\xi}{\varepsilon}. \quad (3.22)$$

(In addition, of course, there are ε -independent terms, which depend on the details of the theory as well as on the precise form of the UV regulator.) We learn from (3.22) that the entropy has a logarithmic UV divergence proportional to the central charge of the fixed point, and that in a massive theory it is IR-finite.

Exercise: Repeat the analysis of this subsection for the half-space

$$A = \{t = 0, x^1 \geq 0\}. \quad (3.23)$$

in a $D > 2$ dimensional spacetime. In particular, find the dependence of $S(A)$ on the cutoff ε . You will have to impose an IR regulator on the other spatial directions x^2, \dots, x^{D-1} , for example by making them periodic. You should find

$$S(A) \propto \frac{\sigma}{\varepsilon^{D-2}}, \quad (3.24)$$

where σ is the area of the entangling surface. (Unlike the logarithmic leading term in $D = 2$, the coefficient of this term depends on the precise form of the UV regulator, and so is not physically meaningful.)

For any spatial region with a smooth boundary, very close to the entangling surface it looks like a plane, so we would expect the leading divergence of the entanglement entropy to be given by (3.24). This is called the “area law” divergence, and is a universal feature of entanglement entropies in field theories. There are also subleading divergent terms that depend on the particular field theory, as well as the geometry of A . The divergent terms arise from local UV physics near the entangling surface, and are therefore generally assumed to be independent of the state (for valid field-theory states, which have energy densities well below the cutoff scale), and to take the form of integrals of local geometric invariants. Because of the second property, they cancel when computing for example the mutual information of separated regions, leaving a finite and cutoff-independent quantity; we will see an example in subsection 3.5.

3.3 CFT: interval

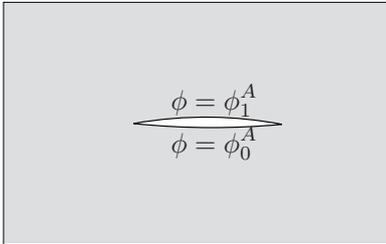
We can get even further by assuming more symmetries for our system. We will therefore next consider a conformal field theory in $1+1$ dimensions.

If we take the limit $m \rightarrow 0$ or $\xi \rightarrow \infty$ in (3.22), we get, in addition to the UV divergence, an IR divergence. To cut this off, instead of a half space, we consider a finite interval. The entangling surface now consists of the two endpoints. By translational symmetry, the entropy can depend only on their relative position, the length L of the interval. One might think that by conformal symmetry the entropy (which is dimensionless) cannot even depend on L . However, as we already saw the entropy is UV divergent and therefore depends on the UV cutoff ε . This breaks the conformal symmetry and allows the entropy to depend on the ratio L/ε . A quick and dirty derivation of this dependence can be made along the same lines as the calculation (3.22). There we found that the UV-divergent part of the entropy near each endpoint is $-(c/6)\ln\varepsilon$. Taking into account the fact that the interval has two endpoints, the total UV divergent part is $-(c/3)\ln\varepsilon$. Since, as explained above, by conformal invariance, the entropy can only depend on L/ε , the ε dependence determines the L dependence, we must have

$$S(A) = \frac{c}{3} \ln \frac{L}{\varepsilon} + \text{constant}, \quad (3.25)$$

where the constant part is non-universal, i.e. it depends on the precise form of the UV regulator.

The result (3.25) can be confirmed by an honest calculation using path integral and CFT techniques [14, 24]. By the same reasoning as in subsection 3.2, the matrix elements of ρ_A are given by the Euclidean path integral on the plane cut along the interval A on the $\tau = 0$ axis:

$$\langle \phi_0^A | \rho_A | \phi_1^A \rangle_A = \frac{1}{Z} \times \left[\text{Diagram} \right]. \quad (3.26)$$


There is a continuous family of conformal transformations of the plane that takes the top of the cut to the bottom, in the same way that a 2π rotation takes the top of the cut in (3.10) to the top. The generator of this conformal rotation corresponds in the Lorentzian spacetime to a conformal boost, acting within the diamond $D(A)$. With the endpoints of A at $x = 0, L$, the conformal boost generator is

$$K = \frac{x(L-x)}{L} T_{tt}, \quad (3.27)$$

and the modular Hamiltonian is $H_A = 2\pi K$. Notice that, near the endpoints, K is approximately equal to the boost generator (3.17).

From the example of Rindler space and this one, one might get the impression that the modular Hamiltonian is always a weighted integral of the stress tensor. This is false. In fact, the two examples we have studied are almost the *only* examples in field theories where the modular Hamiltonian can be expressed as an integral of a local operator. Typically it is some non-local operator, and is not the generator of any geometric symmetry.

Using (3.26), we can compute $\text{Tr} \rho_A^\alpha$ for $\alpha = 2, 3, \dots$ and thereby the Rényi entropies. The matrix elements of ρ_A^α are given by taking α copies of the cut plane and gluing the top of the slit on each sheet to the bottom on the next sheet. The trace is computed by gluing the bottom of the slit on the first sheet to the top on the last. This gives a closed Riemann surface, which is the α -fold branched cover, with branch cut A . This technique for computing $\text{Tr} \rho_A^\alpha$ using copies of the path integral representing ρ_A is called the *replica trick*, and can be used very generally.

To compute the partition function on this α -fold branched cover, we note that it is topologically a sphere, and can therefore be related by a Weyl transformation to a standard unit sphere. The Weyl anomaly is proportional to the central charge c but does not depend on any other feature (operator spectrum, OPE coefficients, etc.) of the CFT. When applying the replica trick in more complicated circumstances (for example, when A consists of multiple intervals), this is not the case, and the results are far more complicated. We omit the calculation of the partition function, but the final result for the Rényi entropy is

$$S_\alpha(A) = \frac{c}{6} \left(1 + \frac{1}{\alpha}\right) \ln \frac{L}{\varepsilon} + \text{finite}. \quad (3.28)$$

The ε -dependence arises because the conical singularities at the endpoints of A in the branched cover lead to singularities in the Weyl transformation taking the cover to the unit sphere and therefore to divergences in the partition function.

It is an easy step to find an analytic function of α that matches (3.28) at $\alpha = 2, 3, \dots$, since the function as written is already analytic. The von Neumann entropy is then just given by (3.25), as anticipated.

We will now give two straightforward generalizations of (3.25), which can be obtained by the same method [24]. The first is to put the full system at finite temperature T . To do this, we simply *don't* take the limit $\beta \rightarrow \infty$ in going from (3.6) to (3.7). As a result, ρ_A is computed by the path integral on a cylinder of circumference $1/T$ in the Euclidean time direction with a slit along A . Via the replica trick, we obtain the following entropy:

$$S(A) = \frac{c}{3} \ln \frac{\sinh(\pi T L)}{\pi T \varepsilon}. \quad (3.29)$$

For L small compared to the thermal correlation length $1/T$, the result reduces to the zero-temperature one (3.25), but for L large compared to $1/T$ the growth with L switches from being logarithmic to linear, with slope $2\pi c T/6$, precisely the thermal entropy density at temperature T . Roughly speaking, the entropy is receiving two contributions: a divergent “area-law” contribution due to entanglement across the entangling surface, and an extensive “volume-law” thermal contribution. In this case, the term “entanglement entropy” for $S(A)$, although commonly used, is particularly inapt, since it is not accounted for entirely by entanglement with the rest of the system.

The second generalization is to put the CFT on a circle of length $2\pi R$. Here again we work in the vacuum. (If we put the CFT on a circle at finite temperature, then the Euclidean surface which computes the matrix elements of ρ_A is a torus, and the non-trivial topology makes the calculation much more difficult, and can only be carried out in certain theories or limits.) The result is very simple:

$$S(A) = \frac{c}{3} \ln \left(\frac{2R}{\varepsilon} \sin \frac{L}{2R} \right). \quad (3.30)$$

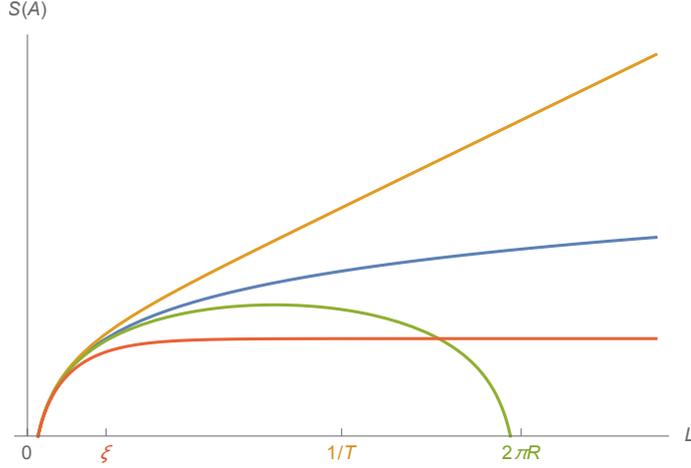


Figure 4: $S(A)$ vs. L for a single interval. Blue: vacuum of CFT on a line, (3.25). Orange: thermal state at temperature T of CFT on a line, (3.29); the slope approaches a positive constant for $L \gg 1/T$. Green: vacuum of CFT on a circle of radius R , (3.30). Red: vacuum of theory with correlation length ξ on a line; the value asymptotes to $(c/3) \ln(\xi/\varepsilon)$ for $L \gg \xi$.

Again, in the limit $R \rightarrow \infty$ we recover the Minkowski-space result (3.25). Two features are noteworthy here: First, the complement A^c is also an interval, of length $L^c = 2\pi R - L$, and we can easily check that $S(A^c) = S(A)$; this is expected from (2.37), since the full system is in a pure state. Second, the entropy *decreases* with L for $L > \pi R$. If we consider an interval B adjacent to A , then AB is also an interval. The conditional entropy $H(B|A) := S(AB) - S(A)$ is then *negative* (and finite), a hallmark of entanglement as discussed in subsection 2.7 above.

The entropies for the three cases discussed here are plotted in fig. 4.

3.4 General theory: interval

In a general $1+1$ dimensional QFT, we still have the formal expression (3.26) for the reduced density matrix, but we no longer have the conformal symmetry that allows us to write down the modular Hamiltonian explicitly and compute the Rényi and von Neumann entropies as in the CFT case. However, we can still make some statements about the qualitative features of the entropy. Consider a theory with a UV fixed point and correlation length ξ . We can guess that for intervals much shorter than the correlation length, the physics is controlled by the UV fixed point and so the entropy is given approximately by (3.25):

$$S(A) \approx \frac{c}{3} \ln \frac{L}{\varepsilon} \quad (L \ll \xi). \quad (3.31)$$

On the other hand, for an interval much longer than ξ , each endpoint looks locally like Rindler space, with the fields farther than a distance ξ from each endpoint frozen out, so we expect the entropy to saturate at a value twice that of Rindler space, (3.22), [24]

$$S(A) \approx \frac{c}{3} \ln \frac{\xi}{\varepsilon} \quad (L \gg \xi). \quad (3.32)$$

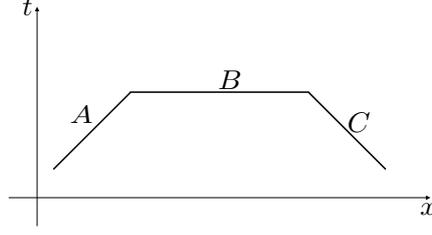


Figure 5: Null trapezoid configuration of intervals in Minkowski space used in the entropic proof of the C-theorem (3.42).

So the entropy as a function of L should like qualitatively like the red curve of fig. 4. While $S(A)$ has not been calculated exactly in a massive theory, semi-analytic calculations for free bosons and fermions confirm this picture [24, 25].

Notice that all the curves in fig. 4 are concave. This is not a coincidence. In fact, it is a consequence of strong subadditivity (2.34). Choose A, B, C to be adjacent intervals, write the inequality in the form

$$(S(ABC) - S(BC)) - (S(AB) - S(B)) \leq 0, \quad (3.33)$$

divide by $L_A L_C$ so it becomes

$$\frac{\frac{S(ABC) - S(BC)}{L_A} - \frac{S(AB) - S(B)}{L_A}}{L_C} \leq 0, \quad (3.34)$$

and finally take the limit $L_A, L_C \rightarrow 0$. We get

$$\frac{d^2 S(B)}{dL^2} \leq 0 \quad (3.35)$$

with $L := L_B$. (Note that in this argument we are never evaluating $S(A)$ or $S(C)$, so it's okay that we are taking their lengths to 0.)

The concavity result (3.35) is a useful constraint on the entropy for a general translationally-invariant state. However, by combining strong subadditivity with Lorentz symmetry, we can make a much more powerful statement [26]. For this purpose we need to work in the vacuum on Minkowski space, since then both the spacetime and state are invariant under the full Poincaré group. This implies that $S(A)$ is Poincaré invariant, and in particular if A is a single interval on any time slice, then $S(A)$ can only be a function of the proper distance between its endpoints. We now consider the “null trapezoid” configuration shown in fig. 5, in which A, B, C are adjacent intervals and A and C are null.⁶ A short exercise in Lorentzian geometry shows that the lengths of B, AB, BC, ABC obey

$$L_{AB} L_{BC} = L_B L_{ABC}, \quad (3.36)$$

or in other words

$$\ln L_{AB} + \ln L_{BC} = \ln L_B + \ln L_{ABC}. \quad (3.37)$$

⁶Strictly speaking, since it includes null segments, this configuration does not lie on a Cauchy slice as we defined it in subsection 3.1. However, we can consider it as the limit of a sequence of Cauchy slices with A and C tending to become null; on each one the inequality must be obeyed, so in the limit it must also be obeyed. Note that in the following argument we are never evaluating $S(A)$ or $S(C)$.

By the same logic as for (3.35) but working in terms of the logarithmic variables, strong subadditivity implies that $S(B)$ is a concave function of $\ln L$ (again $L := L_B$):

$$\frac{d^2}{(d \ln L)^2} S(B) \leq 0. \quad (3.38)$$

This means that the function

$$C(L) := 3 \frac{dS(B)}{d \ln L}, \quad (3.39)$$

is a non-increasing function of L . If the theory has a UV fixed point, then as argued above, for small L the entropy should be that of the fixed point, (3.25) with $c = c_{\text{UV}}$. This implies

$$\lim_{L \rightarrow 0} C(L) = c_{\text{UV}}. \quad (3.40)$$

On the other hand, if the theory has an IR fixed point, then for large L the entropy should again be given by (3.25), with $c = c_{\text{IR}}$ and a different constant than in the UV. Essentially, for large L , the whole RG flow can just be considered as a particular regulator for the IR CFT. We thus have

$$\lim_{L \rightarrow \infty} C(L) = c_{\text{IR}}. \quad (3.41)$$

The monotonicity of $C(L)$ then implies the C-theorem:

$$c_{\text{IR}} \leq c_{\text{UV}}. \quad (3.42)$$

This proof of the C-theorem has the same ingredients as other proofs, such as the original one by Zamolodchikov [27], namely unitarity, locality, and relativity. But the way those ingredients are used seems very different. In particular, the C -function that interpolates monotonically between the UV and IR central charges is different from Zamolodchikov's.

The function $C(L)$ is very useful for diagnosing criticality and identifying the IR CFT in spin chains. In a numerical simulation of a spin chain, it is relatively straightforward to compute the entropy of an interval, and thereby compute $C(L)$. If the spin chain is critical, then it should approach a constant for long intervals, and that constant equals the central charge of the CFT, information which is difficult to obtain by other means.

3.5 CFT: two intervals

As a final example, we consider two separated intervals A, B on a line in the vacuum of a CFT. Their union is the simplest example of a disconnected region. As discussed in subsection 2.3, the mutual information

$$I(A : B) := S(A) + S(B) - S(AB) \quad (3.43)$$

quantifies the total amount of correlation between A and B . Unfortunately, while we know $S(A)$ and $S(B)$, it is very hard to compute $S(AB)$, and indeed it has been computed analytically only for *one* theory, namely the free massless Dirac fermion [26, 25]. The result in that case is very simple:

$$I(A : B) = \frac{1}{2} \ln \frac{(L_A + \Delta)(L_B + \Delta)}{\Delta(L_A + L_B + \Delta)}, \quad (3.44)$$

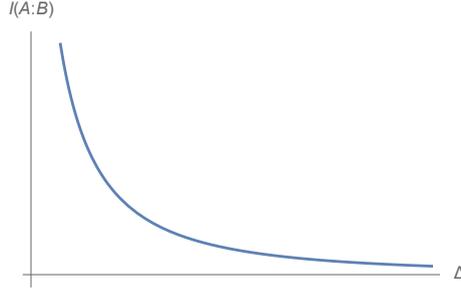


Figure 6: Mutual information $I(A : B)$, given in (3.44), for two intervals separated by Δ for a free massless Dirac fermion.

where Δ is the separation between A and B . It is plotted as a function of Δ in fig. 6. In other CFTs, such as the compact boson, while it is possible to compute the Rényi entropies $S_\alpha(AB)$ for $\alpha = 2, 3, \dots$ exactly, it is not known how to extrapolate the resulting values to $\alpha = 1$ to obtain the von Neumann entropy [28, 29, 30, 31].

Although the massless Dirac fermion is the only case for which we know the mutual information exactly, many of its qualitative features can be shown to hold for the mutual information of separated intervals in the vacuum of any CFT, and some of them even more generally than that:

1. It is finite (independent of ε). This is because the divergences are local on the endpoints of the intervals, and cancel in the mutual information (as long as the intervals are separated). This holds for separated regions in any state in any field theory in any dimension.
2. It is positive. By subadditivity, it must be non-negative, and if it vanished then we would have $\rho_{AB} = \rho_A \otimes \rho_B$, which would imply that all connected correlators between operators in A and in B would vanish, and we know this is not true in a field theory. Therefore, like property 1, this property holds in any state in any field theory.
3. It is conformally invariant. This is because it is finite and therefore independent of ε , so it can only depend on L_A , L_B , and Δ in a conformally-invariant manner. Specifically, it must be a function of the cross-ratio of the four endpoints of the intervals. (The argument of the logarithm in (3.44) is precisely the cross-ratio.)
4. It is a non-decreasing function of L_B . This follows from strong subadditivity, which says that, if we let C be an interval adjacent to B and define $B' = BC$, then

$$I(A : B') = I(A : BC) \geq I(A : B). \quad (3.45)$$

Similarly, it is a non-decreasing function of L_A .

5. It is a non-increasing function of Δ . This follows from the last two facts. It also makes sense intuitively: we would expect the amount of correlation to decrease as the separation increases.
6. It goes to ∞ as $\Delta \rightarrow 0$.

7. It goes to 0 as $\Delta \rightarrow \infty$, and does so at the rate Δ^{-4d} , where d is the scalest dimension of the lightest non-trivial operator [29]. This fact follows from applying the replica trick to the computation of $S(AB)$ and using an OPE expansion to compute the relevant partition function at large Δ . In the case of the Dirac fermion, $d = 1/2$.

Although the examples we have discussed have just scratched the surface of the topic of entanglement in quantum field theories, they have hopefully served to give some idea of how entanglement entropies reflect and illuminate some of the most physics in quantum field theories.

4. Entropy, entanglement, fields, and gravity

Two things are hopefully clear to the reader who has read the previous section: First, entanglement entropies provide significant physical insight into quantum field theories. Second, they are difficult or impossible to calculate except in a few simple cases. This is obviously a somewhat frustrating situation.

There is, however, a surprising exception to this rule. In a *holographic* theory—a field theory admitting a dual description in terms of a higher-dimensional gravitational theory—it is possible to compute an entanglement entropy just by solving a certain classical geometry problem, which is to find the area of a minimal surface with a given boundary. This statement is the celebrated *Ryu-Takayanagi (RT) formula* (4.7) [32, 33]. The RT formula is much more than a calculational tool, however. It is a profound statement about quantum gravity, one that we do not yet fully understand but that has already been fruitfully exploited to advance our understanding of the mysteries of holography. In fact, the RT formula provides such a direct link between spatial entanglement and bulk geometry that it has been suggested that entanglement is the basic ingredient in the field theory from which the bulk spacetime emerges [34].

We will not discuss such investigations here. Instead, our focus will be on how the RT formula works in practice. Specifically, after stating the formula and giving it some motivation and context in subsection 4.2, we will show in subsection 4.3 how the behavior of the entanglement entropies explored in the previous section is realized geometrically by minimal surfaces.

4.1 Holographic dualities

We start with a brief review of the basic facts we will need about holographic dualities. Such a duality relates a D -dimensional quantum field theory to a $D + 1$ -dimensional quantum gravity theory with negative cosmological constant. The gravity theory is subject to asymptotically AdS boundary conditions on the metric, with conformal boundary equal to the manifold M on which the field theory lives, meaning that near the boundary the metric takes the form

$$ds^2 = \frac{\ell^2}{z^2} (ds_M^2 + dz^2) \quad (4.1)$$

where ds_M^2 is the metric on M , ℓ is the AdS radius (related to the cosmological constant by $\Lambda = -\frac{1}{2}D(D-1)\ell^{-2}$), and $z > 0$ with the boundary at $z \rightarrow 0$.⁷ There are also boundary conditions on the

⁷For simplicity, in this description we have performed a Kaluza-Klein reduction on any compact manifold in the bulk, such as the S^5 in $\text{AdS}_5 \times S^5$, so the modes of the fields on this manifold become fields on the AdS_{D+1} spacetime.

other bulk fields depending on the values of the coupling constants in the field theory; the details here will not be important to us.

In the limit of a large number of field-theory degrees of freedom (e.g. large N for a gauge theory, large c for a 2-dimensional CFT), the gravitational theory becomes classical, with that number controlling the hierarchy between the AdS radius ℓ and the Planck length $(G_N \hbar)^{1/(D-1)}$. Specifically, for a two-dimensional CFT,

$$\frac{\ell}{G_N \hbar} = \frac{2c}{3}, \quad (4.2)$$

while for a gauge theory

$$\frac{\ell^{D-1}}{G_N \hbar} \propto N^2. \quad (4.3)$$

In a further limit of the field theory parameters (generally a strong-coupling limit), the low-energy sector of the gravitational theory becomes Einstein gravity coupled to some matter fields. We will be working in both of these limits. States of the system that are classical (from the bulk point of view) are described by asymptotically AdS_{D+1} solutions to the Einstein-matter equations. Depending on the bulk matter content and boundary conditions, the ground state may be AdS_{D+1} (or a patch thereof), in which case the theory is conformal, or it may be more complicated, in which case the theory has a non-trivial RG flow. Roughly speaking, the region of the bulk “close to” the boundary corresponds to the UV of the field theory while the region “far” from the boundary corresponds to the IR. A UV cutoff on the field theory can be implemented by cutting off the bulk geometry at $z = \varepsilon$.

The large- N (or large- c) limit is a kind of thermodynamic limit, and the large number of field-theory degrees of freedom allows phase transitions to occur even in finite volume. For example, there is typically such a phase transition as a function of temperature from a “confined” phase with free energy and entropy of order 1 to a “deconfined” phase with free energy and entropy of order N^2 or c . In the confined phase the bulk geometry is the same as in the ground state (with the entropy accounted for by a thermal gas of gravitons and other light particles), while in the deconfined phase there is a black hole in the bulk (with the order- N^2 part of the entropy accounted for by its Bekenstein-Hawking entropy). For example, a CFT on $\mathbf{R} \times S^{D-1}$ undergoes a first-order transition, called the Hawking-Page transition, from a low-temperature state described by global AdS_{D+1} to a high-temperature one described by an AdS-Schwarzschild black hole [35].

4.2 Ryu-Takayanagi formula

4.2.1 Motivation: Bekenstein-Hawking as entanglement entropy

The Ryu-Takayanagi (RT) formula relates entanglement entropies in holographic theories to the areas of minimal surfaces in the bulk. We will motivate by starting with the Bekenstein-Hawking entropy, interpreting it as an entanglement entropy, noting that the horizon is a minimal surface, and then writing down the most natural generalization consistent with this example.

Consider a holographic field theory in its deconfined phase. The bulk spacetime is a static, asymptotically AdS black hole. The entropy S_{bh} is given, to leading order in $G_N \hbar$ (i.e. to leading order in $1/N$ or $1/c$) by the Bekenstein-Hawking entropy S_{BH} ,

$$S_{\text{bh}} = S_{\text{BH}} + O(G_N^0), \quad (4.4)$$

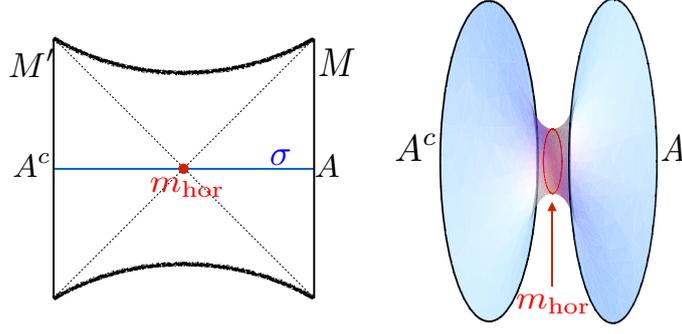


Figure 7: Left: Two-sided static asymptotically AdS black hole. The bulk Cauchy slice σ is time-reflection-invariant and passes through the bifurcate horizon m_{hor} . Its conformal boundaries A, A^c are Cauchy slices for the boundaries M, M' of the spacetime respectively. Right: The induced geometry on σ is an Einstein-Rosen bridge connecting A and A^c . The bifurcate horizon, in red, is the minimal surface on σ between A and A^c .

which is one-quarter the area of the (bifurcate) horizon m_{hor} in Planck units (from now on we suppress \hbar):

$$S_{\text{BH}} = \frac{1}{4G_{\text{N}}} \text{area}(m_{\text{hor}}). \quad (4.5)$$

S_{bh} is a thermal entropy. But, as we discussed in subsection 2.6, the entropy of any mixed state can be understood as an “entanglement” entropy by purifying the state, and for a thermal state a particularly natural purification is the thermofield double state (2.62). The thermofield double is a pure state on two copies of the system, in this case two copies of the field theory, or, to say it differently, the field theory on two copies M and M' of the boundary spacetime. The thermofield double state is represented holographically by the symmetric two-sided extended black hole spacetime, which includes the original black hole representing the thermal state as one of its exterior regions; see fig. 7 [36]. The two exterior regions are static, but the extended spacetime is not globally static, so we cannot define canonical “constant-time” slices. Nonetheless it has distinguished Cauchy slices, namely those that are invariant under a time reflection symmetry of the spacetime. Such an invariant slice σ passes through the bifurcate horizon m_{hor} , and its restriction to either exterior region is a constant-time slice there. It intersects the full boundary $M \cup M'$ on a Cauchy slice Σ which is the union of Cauchy slices A for M and A^c for M' . By the definition of the thermofield double state, the reduced density matrix ρ_A is the same as the thermal density matrix on one copy of the field theory, so $S(A)$ is the black hole entropy (4.4); working to leading order in G_{N} ,

$$S(A) = S_{\text{bh}} = S_{\text{BH}} = \frac{1}{4G_{\text{N}}} \text{area}(m_{\text{hor}}). \quad (4.6)$$

Now we want to generalize (4.6), so we need some bulk surface that will stand in for m_{hor} in a more general situation. The bifurcate horizon is distinguished in various ways, for example, it is the locus of fixed points of a Killing vector (which is timelike in the exterior regions and spacelike in the interior regions). But we can also characterize m_{hor} without appealing to the continuous symmetries of the spacetime: it is a *minimal-area* surface on σ , and more specifically it is the minimal surface that divides the bulk into two regions, one bounded by A and the other by A^c .

4.2.2 Statement

We now guess that this *minimal-area* property is the key feature that associates the surface m_{hor} with the entropy $S(A)$, and that this holds more generally than in the thermofield double example. For a holographic spacetime with a time-reflection symmetry, let σ be a reflection-invariant bulk Cauchy slice and let A be a region of its conformal boundary Σ . The natural generalization of the thermofield double example is to look for the minimal surface $m(A)$ in the bulk such that there is a bulk region $r(A) \subseteq \sigma$ whose boundary is $A \cup m(A)$, and assert that its area gives the entropy of A :

$$S(A) = \frac{1}{4G_{\text{N}}} \text{area}(m(A)). \quad (4.7)$$

This is the RT formula. The gist of it is that the reduced density matrix ρ_A can be treated as if it were described by a black hole with horizon $m(A)$, *even when there is no Killing horizon*. RT thus asserts that the Bekenstein-Hawking formula applies far beyond the domain where it was originally derived.

It is useful to be slightly more careful in the definition of $m(A)$, by keeping track of orientations. Assign an orientation to the bulk slice σ . The boundary region $A \subseteq \partial\sigma$ inherits an orientation from σ . We say an oriented surface m is *homologous* to A , and write $m \sim A$, if there exists a bulk region $r \subseteq \sigma$ such that $\partial r = A \cup -m$, where $-m$ is the surface m with its orientation reversed. $m(A)$ is then the minimal-area surface homologous to A ; we call the corresponding region $r(A)$ the *homology region*. In this notation we can also write (4.7) as

$$S(A) = \frac{1}{4G_{\text{N}}} \min_{m \sim A} \text{area}(m). \quad (4.8)$$

Two comments are in order. First, the requirement of time symmetry was necessary to pick out a distinguished slice, as we can see in the thermofield-double case: a general bulk Cauchy slice of the extended black hole spacetime shown in fig. 7 does not pass through the bifurcate horizon, and the minimal surface on it does not have area equal to S_{bh} . In practice this is not such an onerous requirement, as one is very often interested in entropies in *static* spacetimes, where any constant-time slice is automatically time-reflection invariant; all the examples we discuss in the rest of this section are like this. However, as we will mention in subsection 4.2.4, this restriction can be relaxed, allowing one to compute entropies in time-dependent spacetimes.

Second, in the thermofield double example we used to motivate RT, the region A was an entire boundary so there was no entangling surface ∂A . However, as in the previous section one is often interested in studying the entropy of a region A with an entangling surface. The homology condition then has an important implication for the minimal surface. Since

$$A \cup -m(A) = \partial r(A), \quad (4.9)$$

and $\partial r(A)$ is necessarily closed, A and $m(A)$ must have the same boundary:

$$\partial m(A) = \partial A. \quad (4.10)$$

So the minimal surface must reach all the way to the conformal boundary, which it must touch on the entangling surface; see fig. 8 for an illustration. Given the $1/z^2$ factor in the metric (4.1), its area must diverge. Since the small- z region of the bulk corresponds to the UV of the field theory, this is a UV divergence, and (as we will see in the next subsection) it accounts precisely for the expected UV divergence in $S(A)$.

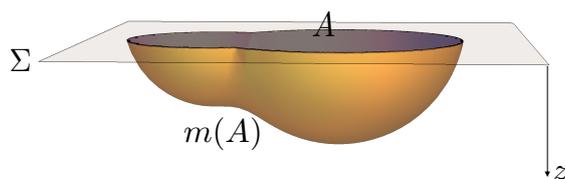


Figure 8: Illustration of RT formula (4.7). A is a region on the boundary Cauchy slice Σ . The space below Σ is the time-reflection-invariant bulk Cauchy slice σ . $m(A)$ is the minimal surface in the bulk homologous to A . In particular, being homologous to A implies that $m(A)$ is anchored on the entangling surface ∂A . It “hangs down” into the bulk in order to minimize its area, subject to that boundary condition, given that the prefactor $1/z^2$ in the metric makes areas smaller deeper in the bulk.

4.2.3 Checks

Should you believe such a bold generalization of the Bekenstein-Hawking formula? Yes, you should. It has been checked and confirmed in many ways, some of which we will review in these lectures:

- RT agrees with calculations of entanglement entropies or parts thereof (e.g. certain divergent terms) in cases where these can be calculated from first principles in the field theory [far too many to cite].
- RT obeys highly non-trivial properties of entanglement entropies, such as strong subadditivity [37, 15]. In fact, it obeys all known applicable properties of entanglement entropies [38].
- RT is consistent with holographic calculations of Rényi entropies carried out using the replica trick in the framework of Euclidean quantum gravity [39, 40, 41].⁸

In addition, RT is implicitly tested by its numerous applications and extensions (again, there are far too many papers to cite here):

- It has been applied and given reasonable results in hundreds of different spacetimes (in many cases informing our understanding of field-theory entanglement entropies more generally).
- It admits sensible generalizations, briefly discussed in subsection 4.2.4, which have also been checked similarly to (thought not quite as extensively as) RT itself.
- It sits at the center of a broader (but still incomplete) theoretical framework concerning the detailed relation between the bulk and boundary theories, which includes such topics as:

⁸In holographic theories it is much harder to calculate Rényi than von Neumann entropies, the opposite of the usual situation in field theories. This is related to the fact that holographic theories are in a certain sense thermodynamic, with the bulk geometry representing the macrostate. Von Neumann entropies are (as discussed in subsection 2.4) good thermodynamic quantities and are therefore captured by this bulk geometry. The Rényi entropies are not thermodynamic quantities, so their calculation requires re-solving the saddle-point equation, in this case the Einstein equation with different boundary conditions. See [15] for further discussion on this point.

derivation of the Einstein equation from RT; subregion duality and entanglement wedge reconstruction; relation between bulk and boundary modular Hamiltonians; holographic error-correcting codes; bit threads; tensor networks; hole-ography, differential entropy, and kinematic space; other entanglement-related quantities such as logarithmic negativity and entanglement of purification; and much more.

4.2.4 Generalizations

The RT formula is limited in three key ways: It requires the bulk theory to be *classical Einstein gravity* and the bulk spacetime to possess a *time-reflection symmetry* under which the boundary spatial region A is invariant. All three of these restrictions can be relaxed:

- For a general bulk spacetime, without assuming any symmetries, and a general boundary spatial region A , the entropy is given by the area of the minimal bulk *extremal* spacelike surface homologous to A [42]. (Extremal means an extremum of the area functional.) This generalization is called the Hubeny-Rangamani-Takayanagi (HRT) formula.
- Moving away from Einstein gravity means including higher-derivative (e.g. stringy) corrections to the bulk gravitational action. This can be taken into account by correcting the functional of the surface to be minimized to include terms beyond the area [43]. In particular, for a Lovelock-type term in the action, the correction takes a particularly simple form: it is the next-lower Lovelock term applied to the surface [44, 45]. For example, a Gauss-Bonnet term in the gravitational action adds an Einstein-Hilbert term (integrated induced Ricci scalar) to the functional to be minimized. A generalization of RT has also been conjectured for three-dimensional Chern-Simons gravity [46].
- Moving away from the classical limit means including G_N corrections to the RT formula. At order G_N^0 , this correction is given by treating the bulk fields (including metric perturbations) as quantum fields on a fixed background and computing the entanglement entropy of the homology region $r(A)$ [47]. This is called the Faulkner-Lewkowycz-Maldacena (FLM) formula. An all-orders generalization is conjectured in [48]. Little is known about non-perturbative quantum corrections to RT, aside from the fact that they play the important role of smoothing out the phase transitions predicted by RT, discussed in subsections 4.3.4 and 4.3.6 below.

Unfortunately, we will not have time in the current lectures to discuss these generalizations further.

4.3 Examples

In section 3, we either estimated $S(A)$ or computed it exactly for a variety of configurations and states of two-dimensional field theories, and saw in each case how the result reflected interesting qualitative features of the physics. To gain practice and intuition with the RT formula, in this subsection we will calculate the same quantities in the holographic setting. We will see that the results agree with those derived earlier from first principles, providing a highly non-trivial check on the formula.⁹ In fact, it seems almost miraculous how the minimal surface manages to realize each aspect of the physics.

⁹The calculations in subsections 4.3.1, 4.3.2, and 4.3.3 were first done in the original RT paper [33].

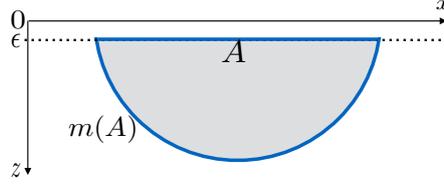


Figure 9: The geodesic $m(A)$ on the hyperbolic plane (4.11) between the endpoints of a boundary interval A is a coordinate semi-circle, truncated at the cutoff surface $z = \varepsilon$. The homology region $r(A)$ that interpolates between A and $m(A)$ is shown in gray.

The bulk in each case is a static asymptotically AdS_3 spacetime, so the constant-time slice σ is an asymptotically hyperbolic two-dimensional space and the “minimal-area surface” $m(A)$ is a minimal-length geodesic. We will also make a few comments about the higher-dimensional case as we go along.

4.3.1 CFT vacuum: interval

The metric dual to the vacuum of a two-dimensional CFT on the line is (the Poincaré patch of) AdS_3 , whose constant-time slice is the hyperbolic plane, with metric

$$ds^2 = \frac{\ell^2}{z^2} (dx^2 + dz^2). \quad (4.11)$$

The geodesics in this space are easily found to be coordinate semicircles centered on the boundary,

$$(x - x_0)^2 + z^2 = r^2. \quad (4.12)$$

If we put the endpoints of the interval at $x = \pm L/2$, then, in order to have the endpoints meet the cutoff surface $z = \varepsilon$ at the endpoints, we should choose

$$x_0 = 0, \quad r^2 = \left(\frac{L}{2}\right)^2 + \varepsilon^2. \quad (4.13)$$

(See fig. 9.) Note that this geodesic is (rather trivially) homologous to the interval A , with the region $r(A)$ being the half-disk bounded by A and $m(A)$. Its length is easily computed to be $2\ell \ln(L/\varepsilon)$ (dropping terms of order ε). (**Exercise:** Check these statements.) Using (4.2), RT says

$$S(A) = \frac{c}{3} \ln \frac{L}{\varepsilon}, \quad (4.14)$$

in agreement with (3.25).

The interesting thing here is how the UV divergence arises. The blow-up of the metric near the boundary forces the geodesic to hit it orthogonally. Given that the length element is $ds \sim \ell dz/z$, the total length is logarithmically divergent, with the correct coefficient. In fact, this generalizes to higher dimensions. Due to the $1/z^2$ factor in the near-boundary metric (4.1), the minimal surface $m(A)$ approaches ∂A vertically as shown in fig. 8. Using (4.1), this gives us the leading divergent part of the area:

$$\int_{m(A)} \sqrt{h} \approx \ell^{D-1} \int_{\partial A} \sqrt{h_{\partial A}} \int_{\varepsilon} dz z^{1-D} = \frac{\ell^{D-1} \sigma}{(D-2) \varepsilon^{D-2}}, \quad (4.15)$$

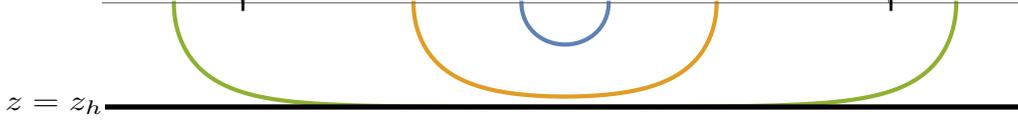


Figure 10: Geodesics on a spatial slice of the planar BTZ black hole geometry (4.16) for various boundary intervals. The separation between the ticks is the thermal length $1/T$. For intervals much shorter than that, the geodesic is close to being a semi-circle as in the vacuum case of fig. 9; the geodesic stays close to the boundary, where the metric is approximately the same as (4.11). On the other hand, for intervals longer than $1/T$, the geodesic skirts the horizon over a long distance, resulting in the linear “volume-law” behavior of the entropy, with slope equal to the thermal entropy density. (In the figure, the curve appears, due to its finite thickness, to coincide with the horizon; in fact it comes very close—in this case, to $z = 0.999z_h$ —but does not touch the horizon.)

where h is the determinant of the induced metric on $m(A)$, $h_{\partial A}$ is the determinant of the metric on ∂A induced from ds_M^2 (the boundary metric), and σ is the total area of the entangling surface (with respect to ds_M^2). By RT, using (4.3), and dropping constants, we find that the leading divergence in $S(A)$ is $N^2\sigma/\varepsilon^{D-1}$, in agreement with (3.24). In other words, in RT the area-law UV divergence comes from the part of the minimal surface near the boundary; this part of the minimal surface represents the short-distance entanglement of nearby degrees of freedom on either side of the entangling surface. Furthermore, the number of degrees of freedom entangled in this way is N^2 , the number of *elementary* (“gluon”) fields.

4.3.2 CFT thermal state: interval

At finite temperature, the metric is replaced by that of the planar BTZ black hole. (Since the theory is conformal and we are in infinite space, the Hawking-Page transition occurs at zero temperature.) The spatial metric is

$$ds^2 = \frac{\ell^2}{z^2} \left(dx^2 + \frac{dz^2}{f(z)} \right), \quad f(z) = 1 - \frac{z^2}{z_h^2}, \quad (4.16)$$

where the horizon position z_h is related to the temperature by

$$T = \frac{1}{2\pi z_h}. \quad (4.17)$$

Again, the geodesic can easily be found, and the entropy (3.29) reproduced; see fig. 10. **Exercise:** Find the geodesics and reproduce (3.29).

As noted in subsection 3.3, for intervals much shorter than the thermal length $1/T$, the entropy is close to that in the vacuum. Here this is due to the fact that the corresponding geodesics stay close to the boundary, where the metric is very close to the vacuum metric (4.11).

Another feature noted in subsection 3.3 is the linear, “volume-law” part of the entropy for $L \gg 1/T$, with slope equal to the thermal entropy density. Again, this property of the entropy is manifested in the behavior of the geodesic. As can be seen in fig. 10, for large L , the geodesic basically has three parts: a part that goes from the boundary almost to the horizon and which in the $L \rightarrow \infty$ limit is independent of L , a part that is very close to (but not intersecting) the horizon,

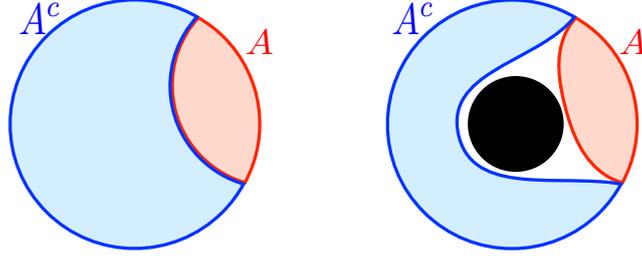


Figure 11: Left: In a pure state, the bulk does not contain a horizon so A and A^c are homologous to each other, since together they bound the entire bulk. This implies that any surface homologous to A is homologous to A^c and vice versa, so by RT they must have the same entropy, in agreement with (2.37). Right: On the other hand, in a mixed state described by a black hole, A and A^c are *not* homologous, so the surfaces homologous to A are different from the ones homologous to A^c , so in general $S(A) \neq S(A^c)$. (Just one topological configuration of the geodesics is shown; others are possible, for example one minimal surface may have a component that wraps the horizon.)

and a part that goes from almost the horizon back to the boundary and again is independent of L . The first and last parts are UV divergent and contribute the $\ln \varepsilon$ part of the entropy, while the part that hugs the horizon contributes an amount equal to L times the thermal entropy density. This is a general phenomenon: The minimal surface for a large region in a deconfined state hugs the horizon, reproducing the expected volume-law thermal part of the entanglement entropy.

4.3.3 CFT on a circle: interval

Finally, the vacuum of the CFT on a circle is described by global AdS_3 . The geometry of the spatial slice is the hyperbolic disk, with metric

$$ds^2 = \frac{\ell^2}{z^2} \left(dx^2 + \frac{dz^2}{f(z)} \right), \quad f(z) = 1 + \frac{z^2}{R^2} \quad (4.18)$$

where $x \sim x + 2\pi R$. **Exercise:** Find the geodesics in the metric (4.18), compute the entanglement entropy of an interval, and show that they agree with (3.30).

We noted below (3.30) that in this case A^c is also an interval, and $S(A^c) = S(A)$, in agreement with (2.37). We can see why this happens from the viewpoint of RT: both quantities are computed by the same geodesic, namely the minimal-length curve in the bulk that separates A and A^c . It is clear that this phenomenon is general: the minimal surface separating A and A^c calculates both $S(A)$ and $S(A^c)$, so they must agree. We can also argue this in the language of homology: A is homologous to A^c , since their union, which is the boundary Cauchy slice Σ , bounds a region, namely the bulk Cauchy slice σ . Therefore any surface m homologous to A is also homologous to A^c and vice versa. (More concretely, if r is the region bounded by $A \cup m$, then $r^c := \sigma \setminus r$ is bounded by $A^c \cup m$, so m is homologous to A^c .) Therefore in (4.8) we are minimizing the area over the same set of surfaces, so we must get the same minimum. (See the left side of fig. 11.)

However, (2.37) assumes that the full system is in a pure state, and in a mixed state we can certainly have $S(A) \neq S(A^c)$. How does this happen holographically? There was a hidden assumption

in the above argument, which is that σ does not have *another* boundary besides Σ (the boundary where the field theory lives). But suppose that the theory is in a deconfined phase, represented holographically by a black hole. Then A^c is *not* homologous to A , and they must have different minimal surfaces, and in general we will have $S(A) \neq S(A^c)$. This can be confirmed by a fairly straightforward calculation in the BTZ black hole background; see right side of fig. 11 for an illustration. Thus, the RT formula cleverly builds in the fact that, in a pure state (where the bulk cannot have a horizon), $S(A) = S(A^c)$, while in a mixed state this is not necessarily true [15].

4.3.4 Gapped theory: half-line

In the last two examples (CFT at finite temperature and on a circle), conformal invariance was broken either by the state or by the geometry of the space on which the field theory lived. Conformal invariance can also be broken by the theory itself. A holographic theory that flows from a UV fixed point is represented by a gravitational theory with boundary conditions such that the ground state (on Minkowski space, say) is not AdS_{D+1} but rather some asymptotically AdS_{D+1} spacetime, on which typically some scalar field depends on the holographic coordinate z . If the theory has a non-trivial IR fixed point, then the bulk has a domain wall parallel to the boundary which interpolates between the UV AdS_{D+1} region near the boundary and another one (with different cosmological constant) in the interior. A simpler case, however, is a wall on which space simply ends. This represents an RG flow to a trivial theory with no degrees of freedom in the deep IR, in other words a gapped or massive theory (also known in the gauge theory context as a confining theory). There exist many holographic constructions of gapped theories, of which some are honest solutions to known quantum gravity theories and others are constructed in an ad hoc fashion, often as models of real-world phenomena.

Since our purpose here is merely to illustrate the qualitative behavior of entanglement entropies in gapped theories, we will consider the simplest possible ad hoc theory.¹⁰ This is a two-dimensional theory on Minkowski space, whose vacuum is represented by the AdS_3 metric, but where the z coordinate extends only from ε to some finite value ξ , at which point the bulk space simply ends:

$$ds^2 = \frac{\ell^2}{z^2} (-dt^2 + dx^2 + dz^2), \quad \varepsilon < z \leq \xi. \quad (4.19)$$

(In honest models of gapped theories, the radial direction ends in a more graceful manner, for example with a compact internal manifold smoothly capping off.) Since as a general rule the region farther from the boundary represents the IR of the field theory, it is intuitive that if the space ends then the theory would have no degrees of freedom in the IR. More quantitatively, by putting appropriate boundary conditions on the bulk fields at $z = \xi$, one can show that correlation functions of boundary fields die off at distances larger than ξ like $e^{-\Delta x/\xi}$ or faster, so ξ is the correlation length.

We emphasized in the previous subsection that a horizon in the bulk influenced the homology condition, so that A and A^c would not have the same minimal surface, reflecting the fact that the full system is not in a pure state. Given the presence of the confining wall, the same argument would

¹⁰The first application of RT to a gapped theory and observation of the qualitative features we explore in this subsection was done in [49].

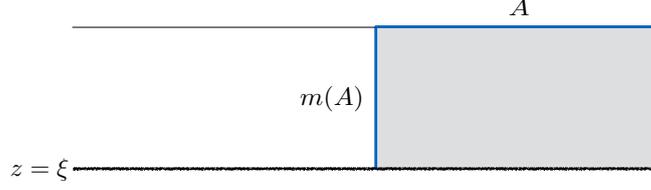


Figure 12: The minimal surface $m(A)$ for the half-line $A = \{x \geq 0\}$ is the straight line in the bulk at $x = 0$, extending from the boundary to the wall.

seem to lead to the same conclusion, yet the system is in the vacuum, a pure state. The resolution to this puzzle is that, unlike a horizon, a confining wall X (an internal wall on which space ends) does *not* count as a “boundary” for the purpose of imposing the homology constraint. In other words, we only require that (4.9) is satisfied modulo X . (Technically, we work in homology *relative to* X .) In the case where the wall is due to an internal compact manifold capping off, this is very natural, since from the higher-dimensional viewpoint there is indeed no boundary at X . More generally, the physical reason for this rule is that a wall, unlike a horizon, does not carry entropy. For example, if we apply RT to the whole time slice Σ , we find that it is null-homologous, $\Sigma \sim \emptyset$, with $r(\Sigma) = \sigma$, the whole bulk time slice, giving the correct answer $S(\Sigma) = 0$.

A more interesting case is the half-line, $A = \{x \geq 0\}$, studied in a general field theory in subsection 3.2. To apply RT, we look for a geodesic in the geometry (4.19) that begins at $(x, z) = (0, \varepsilon)$ and does not end anywhere else on the boundary, so it must end on the wall. As discussed in subsection 4.3.1, such geodesics are coordinate semicircles (see (4.12)). The shortest one is just the straight line (a “semicircle” with infinite radius),

$$m(A) = \{x = 0, \varepsilon < z \leq \xi\} \quad (4.20)$$

(see fig. 12). Its length is

$$\int_{\varepsilon}^{\xi} dz \frac{\ell}{z} = \ell \ln \frac{\xi}{\varepsilon}. \quad (4.21)$$

Using (4.2), we find

$$S(A) = \frac{c}{6} \ln \frac{\xi}{\varepsilon}, \quad (4.22)$$

in agreement with the general result (3.22). (The fact that we got the value on the right-hand side of (3.22) *exactly* is due to the artificial simplicity of our spacetime and should not be taken as significant.) The homology region is the half-space

$$r(A) = \{x \geq 0, \varepsilon < z \leq \xi\}. \quad (4.23)$$

4.3.5 Gapped theory: interval

We now move on to the case of a finite interval, $A = \{-L/2 \leq x \leq L/2\}$. For a general massive theory, we concluded in subsection 3.4 that $S(A)$ was given for short and long intervals by (3.31) and (3.32) respectively. Here we encounter for the first time the phenomenon of *competing saddles*. One *locally* minimal surface homologous to A is the union of vertical segments connecting each

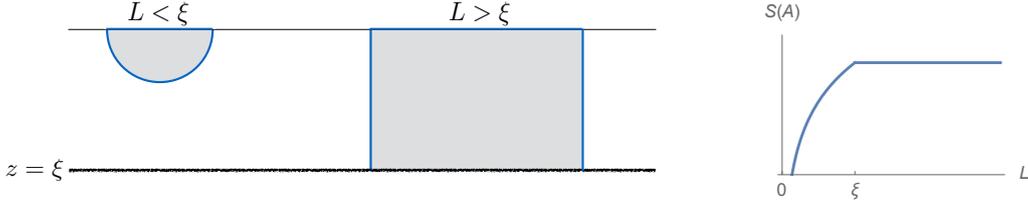


Figure 13: Left: In a bulk with a confining wall at $z = \xi$, the minimal surface $m(A)$ for an interval A of length L takes two different topologies depending on L : for $L < \xi$ it is a semicircle connecting the endpoints of the interval, while for $L > \xi$ it is two line segments connecting each endpoint to the wall. Right: As a result, the entropy $S(A)$ shows a first-order transition (continuous with discontinuous first derivative) at $L = \xi$ from a logarithmic growth to a constant; see (4.29).

endpoint to the wall:

$$m_1 = \left\{ x = -\frac{L}{2}, \varepsilon < z \leq \xi \right\} \cup \left\{ x = \frac{L}{2}, \varepsilon < z \leq \xi \right\}. \quad (4.24)$$

The corresponding homology region is the rectangle

$$r_1 = \left\{ -\frac{L}{2} \leq x \leq \frac{L}{2}, \varepsilon < z \leq \xi \right\}. \quad (4.25)$$

m_1 has length

$$l_1 = 2\ell \ln \frac{\xi}{\varepsilon}. \quad (4.26)$$

On the other hand, for sufficiently short intervals, there is another homologous geodesic, namely the semicircle connecting the endpoints:

$$m_2 = \left\{ x^2 + z^2 = \left(\frac{L}{2}\right)^2 + \varepsilon^2 \right\}. \quad (4.27)$$

m_2 is only locally minimal for $L < 2\xi$. For $L \geq 2\xi$, this semicircle collides with the wall; the arcs in the physical region $0 < z \leq \xi$ are not minimal-length since they can be shortened by straightening them out, returning us to m_1 . For $L < 2\xi$ it has length

$$l_2 = 2\ell \ln \frac{L}{\varepsilon}. \quad (4.28)$$

We see that m_2 is shorter for $L < \xi$ while m_1 is shorter for $L > \xi$, giving

$$S(A) = \frac{c}{3} \times \begin{cases} \ln \frac{L}{\varepsilon}, & L < \xi \\ \ln \frac{\xi}{\varepsilon}, & L > \xi \end{cases}. \quad (4.29)$$

So as we vary L , the entropy is continuous but has a discontinuous first derivative at $L = \xi$; we can think of this as a phase transition, as discussed below. Basically, the short- and long-interval estimates (3.31) and (3.32) become *exact* in this holographic set-up. The behavior is shown in fig. 13.

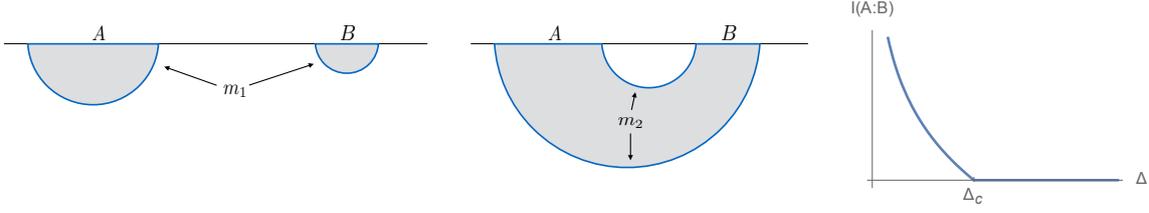


Figure 14: Left and middle: Candidate minimal surfaces for two intervals in a CFT vacuum. m_1 is the union of the minimal surfaces for A and B respectively, while m_2 connects A and B . Right: Mutual information (4.33) calculated from these minimal surfaces as a function of the separation Δ between the intervals.

While the details of the calculation depend on the bulk geometry, the existence of a phase transition does not: in any gapped two-dimensional holographic theory, $S(A)$ undergoes a phase transition as we increase L past a critical value L_c which is of the order of the correlation length. For $L < L_c$, $m(A)$ connects the endpoints of A , while for $L > \xi$ it consists of segments extending directly from each endpoint to the wall. In higher dimensions, analogous behavior is seen: roughly speaking, when A is small compared to the correlation length, $m(A)$ has the same topology as A , hanging down in to the bulk but not touching the wall; when A is large, $m(A)$ has the topology of ∂A times an interval, and directly connects ∂A to the wall.

We can address this behavior of the entanglement entropy using our intuition from statistical mechanics by writing ρ_A as the thermal state at unit temperature with respect to the modular Hamiltonian H_A , $\rho_A = e^{-H_A}/Z$. As we vary L , the parameters in this Hamiltonian change, changing the entropy. (Note that, unlike for a half-line or an interval in a CFT, H_A is not the integral of a local operator, but rather some complicated non-local operator which we don't have an explicit expression for. This does not affect the discussion that follows.) Based on the behavior of $S(A)$, we conclude that the system must undergo a first-order phase transition at $L = L_c$. This is surprising, since the system has finite spatial extent and therefore (in the presence of the UV cutoff ε) would seem to have a finite number of degrees of freedom. The resolution to this puzzle is that we are working in the *classical* limit $G_N \rightarrow 0$ of the bulk, which represents the limit $c \rightarrow \infty$ in the field theory. This is indeed a kind of thermodynamic limit—not the infinite-volume thermodynamic limit typically encountered for example in lattice condensed-matter systems, but a limit with an infinite *density* of degrees of freedom. A more familiar example is the Hawking-Page transition undergone by a holographic CFT on a circle as a function of temperature, mentioned in subsection 4.1. Such transitions cannot occur at finite c , so they must be smoothed out by $1/c$ corrections, which translate holographically to quantum corrections. More precisely, only *non-perturbative* corrections can smooth out a phase transition. In the case of the Hawking-Page transition, these corrections can be computed using the framework of Euclidean quantum gravity. Unfortunately, we do not yet know how to compute non-perturbative quantum corrections to holographic entanglement entropies.

4.3.6 CFT: Two intervals

We conclude this subsection with the simplest example of a disconnected region, namely the union of two disjoint intervals A, B on a line in a CFT vacuum [39]. This was discussed in the

context of a general field theory in subsection 3.5. In computing $S(AB)$, a locally minimal surface consists of geodesics connecting the four endpoints of AB . According to the boundary condition (4.10), and keeping track of orientations, each geodesic must begin on the left endpoint of an interval and end on the right endpoint of an interval. There are two possible configurations, depending on whether each geodesic connects the endpoints of the same or different intervals (see fig. 14):

1. One geodesic connects the two endpoints of A and the other the two endpoints of B ; this is the union of the respective minimal surfaces for A and B , $m_1 = m(A) \cup m(B)$, and it is homologous to AB via the union of the two homology regions, $r_1 = r(A) \cup r(B)$. The “area” (or length) of m_1 is the sum of the lengths of $m(A)$ and $m(B)$:

$$\text{area}(m_1) = \frac{\ell}{2} \ln \frac{L_A}{\varepsilon} + \frac{\ell}{2} \ln \frac{L_B}{\varepsilon}. \quad (4.30)$$

2. The geodesics connect the left endpoint of A to the right endpoint of B and vice versa. Here the homology region r_2 connects A to B . We have

$$\text{area}(m_2) = \frac{\ell}{2} \ln \frac{\Delta}{\varepsilon} + \frac{\ell}{2} \ln \frac{L_A + L_B + \Delta}{\varepsilon}. \quad (4.31)$$

So, as in the single interval in the gapped theory, we see a phase transition: m_1 is the minimal surface if $\Delta > \Delta_c$ while m_2 is the minimal surface if $\Delta < \Delta_c$, where the critical separation is

$$\Delta_c := \sqrt{\left(\frac{L_A + L_B}{2}\right)^2 + L_A L_B} - \frac{L_A + L_B}{2}. \quad (4.32)$$

From this we easily calculate the mutual information,

$$I(A : B) = \begin{cases} \frac{c}{3} \ln \frac{L_A L_B}{\Delta(L_A + L_B + \Delta)}, & \Delta \leq \Delta_c \\ 0, & \Delta \geq \Delta_c \end{cases}, \quad (4.33)$$

plotted in fig. 14. As in the gapped theory, this phase transition must be smoothed out at finite c by non-perturbative quantum corrections to the RT formula.

In subsection 3.5, we made several predictions for properties of $I(A : B)$ that should hold in any CFT. We see that the holographic mutual information (4.33) obeys all but one of them, namely number 2: it vanishes for finite separations. But this is impossible, since vanishing mutual information implies $\rho_{AB} = \rho_A \otimes \rho_B$, implying that all correlators factorize, $\langle \mathcal{O}_A \mathcal{O}_B \rangle = \langle \mathcal{O}_A \rangle \langle \mathcal{O}_B \rangle$, which is obviously not true. The resolution lies in remembering that the RT formula only tells us the *leading*, order c , entropy in a $1/c$ expansion. So all we know from this calculation is that the order- c part of the mutual information vanishes for $\Delta \geq \Delta_c$. The order-1 term arises from a one-loop correction to the RT formula. (It can also be computed directly in the CFT using large- c techniques.) The result depends on the matter content of the bulk theory, but it is non-zero in any theory, as we would expect.

The basic pattern observed here holds for separated regions A, B in any holographic theory. It can be shown that the homology regions $r(A), r(B)$ don't intersect [15]. Therefore, one candidate minimal surface for AB is the union of their separate minimal surfaces, $m_1 = m(A) \cup m(B)$, which

is homologous to AB via the region $r_1 = r(A) \cup r(B)$. m_1 is locally minimal, but there may also be another locally minimal surface m_2 that connects A and B . Typically, m_1 has smaller area for large separations while m_2 does for small separations, leading to a phase transition.

Notice an important fact: since m_1 is always a candidate minimal surface, and since its area is the sum of the areas of $m(A)$ and $m(B)$, we have

$$\begin{aligned} S(AB) &= \frac{1}{4G_N} \min_{m \sim AB} \text{area}(m) \\ &\leq \frac{1}{4G_N} \text{area}(m_1) \\ &= \frac{1}{4G_N} (\text{area}(m(A)) + \text{area}(m(B))) \\ &= S(A) + S(B). \end{aligned} \tag{4.34}$$

Thus, the RT formula automatically obeys the subadditivity property. By a more elaborate argument, involving cutting and pasting minimal surfaces, it can also be shown to obey the strong subadditivity property (2.34) [37, 15]. It also obeys many other inequalities, including some that, unlike subadditivity and strong subadditivity, do *not* hold for general quantum states, indicating that holographic states are rather special. The simplest of these is the *monogamy of mutual information* inequality [38]:

$$S(AB) + S(AC) + S(BC) \geq S(A) + S(B) + S(C) + S(ABC). \tag{4.35}$$

(An example of a state that violates (4.35) is

$$\rho_{ABC} = \frac{1}{2} (|000\rangle\langle 000| + |111\rangle\langle 111|). \tag{4.36}$$

More complicated inequalities obeyed by RT can be found in [50].

Acknowledgments

The author's work was supported by the National Science Foundation through Career Award No. PHY-1053842; by the Simons Foundation through *It from Qubit: Simons Collaboration on Quantum Fields, Gravity, and Information* and through a Simons Fellowship in Theoretical Physics; and by the Department of Energy through grant DE-SC0009987. I would also like to thank the MIT Center for Theoretical Physics, where much of these notes were written, for hospitality throughout the year 2017. Finally, I would like to thank the organizers of TASI 2017 as well as the students, whose many interesting questions and comments improved the lectures and my own understanding of the subject.

References

- [1] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*. Cambridge University Press, Cambridge, 2000.
- [2] A. Wehrl, *General properties of entropy*, *Rev. Modern Phys.* **50** (1978), no. 2 221–260.

- [3] E. Witten, *A Mini-Introduction To Information Theory*, [arXiv:1805.11965](https://arxiv.org/abs/1805.11965).
- [4] H. Casini and M. Huerta, *Entanglement entropy in free quantum field theory*, *J. Phys.* **A42** (2009) 504007, [[arXiv:0905.2562](https://arxiv.org/abs/0905.2562)].
- [5] P. Calabrese and J. Cardy, *Entanglement entropy and conformal field theory*, *J. Phys.* **A42** (2009) 504005, [[arXiv:0905.4013](https://arxiv.org/abs/0905.4013)].
- [6] T. Nishioka, S. Ryu, and T. Takayanagi, *Holographic Entanglement Entropy: An Overview*, *J. Phys.* **A42** (2009) 504008, [[arXiv:0905.0932](https://arxiv.org/abs/0905.0932)].
- [7] T. Nishioka, *Entanglement entropy: holography and renormalization group*, [arXiv:1801.10352](https://arxiv.org/abs/1801.10352).
- [8] M. Rangamani and T. Takayanagi, *Holographic Entanglement Entropy*, *Lect. Notes Phys.* **931** (2017) pp.1–246, [[arXiv:1609.01287](https://arxiv.org/abs/1609.01287)].
- [9] H. Nastase, *Introduction to the ADS/CFT Correspondence*. Cambridge University Press, 2015.
- [10] C. E. Shannon, *A mathematical theory of communication*, *Bell System Tech. J.* **27** (1948) 379–423, 623–656.
- [11] F. Hiai, M. Ohya, and M. Tsukada, *KMS condition and relative entropy in von Neumann algebras*, *Pacific Journal of Mathematics* **96** (1981) 99–109.
- [12] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, *Area laws in quantum systems: Mutual information and correlations*, *Phys. Rev. Lett.* **100** (Feb, 2008) 070502, [[arXiv:0704.3906](https://arxiv.org/abs/0704.3906)].
- [13] E. H. Lieb and M. B. Ruskai, *Proof of the strong subadditivity of quantum-mechanical entropy*, *J. Mathematical Phys.* **14** (1973) 1938–1941. With an appendix by B. Simon.
- [14] C. Holzhey, F. Larsen, and F. Wilczek, *Geometric and renormalized entropy in conformal field theory*, *Nucl. Phys.* **B424** (1994) 443–467, [[hep-th/9403108](https://arxiv.org/abs/hep-th/9403108)].
- [15] M. Headrick, *General properties of holographic entanglement entropy*, *JHEP* **03** (2014) 085, [[arXiv:1312.6717](https://arxiv.org/abs/1312.6717)].
- [16] A. Peres, *Separability criterion for density matrices*, *Phys. Rev. Lett.* **77** (Aug, 1996) 1413–1415.
- [17] M. Horodecki, P. Horodecki, and R. Horodecki, *Separability of mixed states: necessary and sufficient conditions*, *Physics Letters A* **223** (1996), no. 1 1 – 8.
- [18] G. Vidal and R. F. Werner, *Computable measure of entanglement*, *Phys. Rev. A* **65** (Feb, 2002) 032314, [[quant-ph/0102117](https://arxiv.org/abs/quant-ph/0102117)].
- [19] M. B. Plenio and S. Virmani, *An Introduction to entanglement measures*, *Quant. Inf. Comput.* **7** (2007) 1–51, [[quant-ph/0504163](https://arxiv.org/abs/quant-ph/0504163)].
- [20] W. Zurek, *Quantum Darwinism, classical reality, and the randomness of quantum jumps*, *Phys. Today* **67** (2014) 10,44, [[arXiv:1412.5206](https://arxiv.org/abs/1412.5206)].
- [21] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully, *Black Holes: Complementarity or Firewalls?*, *JHEP* **02** (2013) 062, [[arXiv:1207.3123](https://arxiv.org/abs/1207.3123)].
- [22] J. J. Bisognano and E. H. Wichmann, *On the Duality Condition for Quantum Fields*, *J. Math. Phys.* **17** (1976) 303–321.
- [23] W. G. Unruh, *Notes on black-hole evaporation*, *Phys. Rev. D* **14** (Aug, 1976) 870–892.
- [24] P. Calabrese and J. L. Cardy, *Entanglement entropy and quantum field theory*, *J. Stat. Mech.* **0406** (2004) P002, [[hep-th/0405152](https://arxiv.org/abs/hep-th/0405152)].

- [25] H. Casini, C. D. Fosco, and M. Huerta, *Entanglement and alpha entropies for a massive dirac field in two dimensions*, *Journal of Statistical Mechanics: Theory and Experiment* **7** (2005) 7–+, [[cond-mat/0](#)].
- [26] H. Casini and M. Huerta, *A Finite entanglement entropy and the c-theorem*, *Phys. Lett.* **B600** (2004) 142–150, [[hep-th/0405111](#)].
- [27] A. B. Zamolodchikov, *Irreversibility of the Flux of the Renormalization Group in a 2D Field Theory*, *JETP Lett.* **43** (1986) 730–732. [*Pisma Zh. Eksp. Teor. Fiz.*43,565(1986)].
- [28] P. Calabrese, J. Cardy, and E. Tonni, *Entanglement entropy of two disjoint intervals in conformal field theory*, *J. Stat. Mech.* **0911** (2009) P11001, [[arXiv:0905.2069](#)].
- [29] P. Calabrese, J. Cardy, and E. Tonni, *Entanglement entropy of two disjoint intervals in conformal field theory II*, *Journal of Statistical Mechanics* **1101** (2011) P01021, [[1011.5482](#)].
- [30] C. De Nobili, A. Coser, and E. Tonni, *Entanglement entropy and negativity of disjoint intervals in CFT: Some numerical extrapolations*, *J. Stat. Mech.* **1506** (2015), no. 6 P06021, [[arXiv:1501.04311](#)].
- [31] P. Ruggiero, E. Tonni, and P. Calabrese, *Entanglement entropy of two disjoint intervals and the recursion formula for conformal blocks*, [arXiv:1805.05975](#).
- [32] S. Ryu and T. Takayanagi, *Holographic derivation of entanglement entropy from AdS/CFT*, *Phys. Rev. Lett.* **96** (2006) 181602, [[hep-th/0603001](#)].
- [33] S. Ryu and T. Takayanagi, *Aspects of Holographic Entanglement Entropy*, *JHEP* **08** (2006) 045, [[hep-th/0605073](#)].
- [34] M. Van Raamsdonk, *Building up spacetime with quantum entanglement*, *Gen. Rel. Grav.* **42** (2010) 2323–2329, [[arXiv:1005.3035](#)]. [*Int. J. Mod. Phys.D*19,2429(2010)].
- [35] E. Witten, *Anti-de Sitter space, thermal phase transition, and confinement in gauge theories*, *Adv. Theor. Math. Phys.* **2** (1998) 505–532, [[hep-th/9803131](#)]. [,89(1998)].
- [36] J. M. Maldacena, *Eternal black holes in anti-de Sitter*, *JHEP* **04** (2003) 021, [[hep-th/0106112](#)].
- [37] M. Headrick and T. Takayanagi, *A Holographic proof of the strong subadditivity of entanglement entropy*, *Phys. Rev.* **D76** (2007) 106013, [[arXiv:0704.3719](#)].
- [38] P. Hayden, M. Headrick, and A. Maloney, *Holographic Mutual Information is Monogamous*, *Phys. Rev.* **D87** (2013), no. 4 046003, [[arXiv:1107.2940](#)].
- [39] M. Headrick, *Entanglement Renyi entropies in holographic theories*, *Phys. Rev.* **D82** (2010) 126010, [[arXiv:1006.0047](#)].
- [40] T. Faulkner, *The Entanglement Renyi Entropies of Disjoint Intervals in AdS/CFT*, [arXiv:1303.7221](#).
- [41] A. Lewkowycz and J. Maldacena, *Generalized gravitational entropy*, *JHEP* **08** (2013) 090, [[arXiv:1304.4926](#)].
- [42] V. E. Hubeny, M. Rangamani, and T. Takayanagi, *A Covariant holographic entanglement entropy proposal*, *JHEP* **07** (2007) 062, [[arXiv:0705.0016](#)].
- [43] X. Dong, *Holographic Entanglement Entropy for General Higher Derivative Gravity*, *JHEP* **01** (2014) 044, [[arXiv:1310.5713](#)].

- [44] J. de Boer, M. Kulaxizi, and A. Parnachev, *Holographic Entanglement Entropy in Lovelock Gravities*, *JHEP* **07** (2011) 109, [[arXiv:1101.5781](#)].
- [45] L.-Y. Hung, R. C. Myers, and M. Smolkin, *On Holographic Entanglement Entropy and Higher Curvature Gravity*, *JHEP* **04** (2011) 025, [[arXiv:1101.5813](#)].
- [46] J. de Boer and J. I. Jottar, *Entanglement Entropy and Higher Spin Holography in AdS₃*, *JHEP* **04** (2014) 089, [[arXiv:1306.4347](#)].
- [47] T. Faulkner, A. Lewkowycz, and J. Maldacena, *Quantum corrections to holographic entanglement entropy*, *JHEP* **11** (2013) 074, [[arXiv:1307.2892](#)].
- [48] N. Engelhardt and A. C. Wall, *Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime*, *JHEP* **01** (2015) 073, [[arXiv:1408.3203](#)].
- [49] I. R. Klebanov, D. Kutasov, and A. Murugan, *Entanglement as a probe of confinement*, *Nucl. Phys.* **B796** (2008) 274–293, [[arXiv:0709.2140](#)].
- [50] N. Bao, S. Nezami, H. Ooguri, B. Stoica, J. Sully, and M. Walter, *The Holographic Entropy Cone*, *JHEP* **09** (2015) 130, [[arXiv:1505.07839](#)].