Markov Property of the Conformal Field Theory Vacuum and the $a$ Theorem

Horacio Casini, Eduardo Testé, and Gonzalo Torroba
Centro Atómico Bariloche and CONICET, S.C. de Bariloche, Río Negro R8402AGP, Argentina

(Received 24 April 2017; published 30 June 2017)

We use strong subadditivity of entanglement entropy, Lorentz invariance, and the Markov property of the vacuum state of a conformal field theory to give new proof of the irreversibility of the renormalization group in $d = 4$ space-time dimensions—the $a$ theorem. This extends the proofs of the $c$ and $F$ theorems in dimensions $d = 2$ and $d = 3$ based on vacuum entanglement entropy, and gives a unified picture of all known irreversibility theorems in relativistic quantum field theory.

DOI: 10.1103/PhysRevLett.118.261602

Introduction.—The central idea of the renormalization group is that the change of physics with scale in a quantum field theory (QFT) can be assimilated to a change in the parameters of the Hamiltonian describing the relevant degrees of freedom. This flow in the space of theories brings us from the ultraviolet (UV) fixed point at short scales to the infrared (IR) one at large scales. At the fixed points, the physics stops changing, and we focus on relativistic systems in $d$ spacetime dimensions, where the end points of the flow are conformal field theories (CFTs).

It has long been known that the renormalization group (RG) is irreversible in two spacetime dimensions [1]. This result, known as the $c$ theorem, shows that the conformal anomaly $c$ (a dimensionless quantity depending on the CFT) decreases between the UV and IR fixed points. The value of $c$ at conformal fixed points is thus interpreted as a precise measure of the number of field degrees of freedom; Zamolodchikov’s theorem then realizes the intuitive idea that this number should decrease at larger scales due to the decoupling of massive modes. It also establishes an ordering of CFTs: theories with smaller $c$ in the UV cannot flow to theories with larger $c$ in the IR, and the renormalization group is irreversible.

In four spacetime dimensions, Cardy [2] gave arguments suggesting that a particular coefficient of the conformal anomaly, the $a$ coefficient of the Euler term, should also decrease under the RG. After long being sought, the $a$ theorem was proved by Ref. [3].

For odd dimensions the situation was initially unclear because there are no conformal anomalies. Based on RG irreversible quantities in holography, Ref. [4] proposed that in odd dimensions the relevant monotonic quantity is the constant term of the entanglement entropy of a sphere. This conjecture, now known as the $F$ theorem, was established for $d = 3$ in Ref. [5], extending the proof [6] of the $c$ theorem in $d = 2$. The crucial property here is the strong subadditivity of entropy, which ultimately gives a different perspective on unitarity and irreversibility. In a related development in supersymmetric QFTs, Ref. [7] conjectured that the constant term in the free energy of a 3-sphere is monotonic—hence the name $F$. In fact, this quantity is the same as the constant term of the entanglement entropy of a sphere [8], and the proposals of Refs. [4] and [7] actually coincide.

These developments suggest that in any dimension the monotonic quantity is the universal part of the entanglement entropy of a sphere. This is proportional to the Euler anomaly for even dimensions. While this points to some underlying principle behind the irreversibility of the RG across dimensions (see, e.g., Ref. [9]), so far the techniques employed have been quite specific to each particular dimension. Only an entropic proof exists for $d = 3$, and so far only a proof based on local field theoretic quantities has been known in $d = 4$; both entropic and correlator techniques can be used to prove the theorem in $d = 2$. An important difficulty for proofs based on correlation functions in odd dimensions is that the $F$ quantity is, in contrast to anomalies, a rather nonlocal object.

In this work we prove the $a$ theorem using entropic techniques, and provide a unifying approach to the irreversibility of the RG. The new key ingredient here will be the recently discovered Markovian property of the vacuum state of a CFT [10]. Based on this we will extend the approach in Ref. [5] to $d = 4$, resolving previous obstacles from problematic terms in the entanglement entropy (EE) of unions and intersections of spheres.

The setup.—We consider a RG flow between UV and IR CFT fixed points in $d$ spacetime dimensions. The flow is triggered by a perturbation with some relevant operator $O$ of dimension $\Delta < d$.

$$S_1 = S_0 + \int d^dx g O(x).$$  \hspace{1cm} (1)

The theory at the UV fixed point is denoted by $T_0$, while $T_1$ is the theory Eq. (1). In order to understand the irreversibility of the RG, we will study the entanglement entropy on spheres. Let $\rho_X$ be the reduction of the global state to the region $X$ and $S(X) = -\mathrm{Tr}(\rho_X \log \rho_X)$ its von Neumann entropy. This is the entanglement entropy between $X$ and the complementary region $\bar{X}$, which we seek to compute.
For the vacuum state of a QFT, the EE of a sphere is in general a complicated function of the radius $r$, a distance cutoff $\epsilon$, and the dimensionful parameters of the theory. At fixed points and for a sufficiently geometric cutoff (such as Ref. [11]) the entropy simplifies to

$$S(r) = \mu_{d-2} r^{d-2} + \mu_{d-4} r^{d-4} + \cdots + \left\{ \begin{array}{ll} (d/2) - 14 A \log(R/\epsilon) & \text{deven} \\ (-)(d-1)/2 F & \text{dodd} \end{array} \right.$$  (2)

See, e.g., Refs. [12–14]. The last term gives the universal part of the EE. $A$ is the Euler trace anomaly coefficient for even dimensions [15], and $F$ is the constant term of the free energy of a $d$-dimensional Euclidean sphere.

The reason for this expression is that the large distance entanglement does not change with dilatations at a fixed point (with the exception of the anomaly term), and hence the $r$ dependence comes from contributions that are local on the entangling surface, i.e., integrals of curvature tensors. Curvature tensors with an odd number of dimensions change sign when they are evaluated on the two sides of the entangling surface and cannot appear in the expansion because of the identity of entropies for complementary regions $S(X) = S(\bar{X})$. Hence, only powers below the area term differing by an even number appear in Eq. (2).

The coefficients $\mu_{d-k}$ have dimension $d - k$. For a CFT (such as $T_0$ above), the only dimensionful parameter is the cutoff $\epsilon$, so that $\mu_{d-k} \sim \epsilon^{-(d-k)}$. For the theory $T_1$ with the relevant perturbation Eq. (1) the situation is richer. For small spheres $r \sim 0$, where we can apply conformal perturbation theory near the UV, we expect

$$\mu_{d-k}^{UV} \sim \epsilon^{-(d-k)} + g^2 \epsilon^{-(d-k) + 2(d - \Delta)} + \cdots$$  (3)

This is UV divergent (and perturbatively computable) for $\Delta \geq (d + k)/2$. Additionally, for small $r$ we expect finite perturbative corrections to the entropy of the form $S(r) \sim g^2 r^{2(d - \Delta)}$, which are nonlocal. See Ref. [13] for holographic examples. On the other hand, taking $r \to \infty$ the IR fixed point is approached; besides the UV divergent terms just discussed, the EE will contain finite renormalizations to $\mu_{d-k}^{IR}$. These contributions, which should be regularization independent, depend on the full RG flow, and are generally nonperturbative. Nonlocal corrections, however, are absent at the IR fixed point.

**Irreversibility from strong subadditivity.**—The idea is to relate EE coefficients of the UV and IR fixed points using a property of entropy called the strong subadditivity inequality (SSA) [16]. For two regions $A$ and $B$ it reads

$$S(A) + S(B) \geq S(A \cap B) + S(A \cup B).$$  (4)

This motivates the construction in Ref. [5] of the geometrical setup illustrated in Fig. 1. A large number of rotated copies $X_i$, $i = 1, \ldots, N$ of a boosted sphere are placed on a null cone. All these spheres are chosen to have the same radius $\sqrt{Rr}$, and are equally distributed in the angular directions. The $t = 0$ projection of these spheres lies between radii $r$ and $R$. Repeated use of the SSA gives

$$\sum_i S(X_i) \geq S(\bigcup_i X_i) + S(\bigcup_{i,j} (X_i \cap X_j))$$

$$+ S(\bigcup_{i,j,k} (X_i \cap X_j \cap X_k)) + \cdots + S(\bigcap_i X_i).$$  (5)

There are $N$ terms on each side of Eq. (5). The right-hand side contains entropies of regions that approach spheres for large $N$ but have wiggly boundaries in a null direction. The aim is to get inequalities involving only spheres in the limit.

The main question is then how to relate entropies of wiggly spheres with those of regular spheres. Since the surfaces are on the light cone, the area term along the boundary of a wiggly sphere matches that of a regular sphere passing through the middle of the wiggles; see Fig. 1. However, the local curvature is different, and so generically we do not expect the entropies to agree (we will see an example below). Unfortunately, a direct calculation of the wiggly contributions seems too complicated, and a different route is needed. It is important to realize, however, that the differences in the EE of wiggly and regular spheres are purely UV at large $N$. If we managed to subtract the UV contributions while still maintaining strong subadditivity, the wiggly contributions would go smoothly to regular contributions. This is the point where the recently discovered Markov property [10] comes into play.

For any two regions $A$ and $B$ with a boundary lying on the light cone, the CFT vacuum in any dimension is a Markov state, namely, it saturates the SSA inequality [10]

$$S(A) + S(B) - S(A \cap B) - S(A \cup B) = 0.$$  (6)

This follows from the form of the modular Hamiltonian on the light cone, as well as from algebraic QFT methods, but...
at first it looks rather surprising. Indeed, intersections and
unions of regions contain additional local singularities that
may produce divergent terms in the entropy, see, e.g.,
Refs. [17,18].

Let us then briefly describe how this works out in $d = 4$,
where all interesting features already appear. The area term
always cancels in the combination Eq. (6), as is the case for
the log($\epsilon$) term coming from a local integral of the
curvature outside the singular points from intersections.
This would also hold for spheres in a plane. A new feature
comes from the intersection of two spheres; it gives a term
that scales with the length $\ell$ of the line of intersection as
$\ell^d/\epsilon^d$. This must be an integral along this line that is locally
the same as the one of the intersection of two planes tangent
to the spheres at a point of the intersection. These two
spatial planes are contained in a null hyperplane
invariant. There is then no local notion of angle
between the two planes—this feature cannot contribute
since we have no local geometrical quantity to distinguish it
from two parallel planes. Next, the intersection lines are
curved and can produce a log($\epsilon$) contribution times a line
integral of the curvature. This cannot be eliminated by
boosting but we note that it is a signed curvature; the union
and intersection of two spheres have exactly opposite
contributions of this form and hence cancel out. Finally,
we have the vertices where three spheres intersect. This
triangular angle is immersed in a null hyperplane, and does
not contribute by the same boost argument as before.

Because of the Markov property, the difference in EE
between the CFT $T_0$ and the theory $T_1$ along the flow,
\begin{equation}
\Delta S(r) = S_\rho(r) - S_\rho'(r)
\end{equation}
still satisfies the strong subadditivity Eq. (4), and Eq. (5)
applies to $\Delta S$. In this way, all UV effects associated to
wiggles cancel out from the inequality (recall that we take
$N \rightarrow \infty$ with fixed coupling $g$) and $\Delta S_{\text{wiggly}}$ can be replaced by $\Delta S_{\text{regular}}$ inside the SSA formula.

The wiggly spheres lie approximately on constant $t$
planes, with radius $l$ ranging from $r$ to $R$. Let $l_k$ be the
radius of the wiggly sphere of order $k$, that is, the one
formed by the union of the intersections of $k$ spheres.
Defining the density of wiggly spheres
\begin{equation}
\beta(l) = \frac{1}{N} \frac{dk}{dl},
\end{equation}
the geometry gives [5]
\begin{equation}
\beta(l) = \frac{2^{d-3} \Gamma[(d-1)/2]}{(2\pi)^{(d-2)/2}} \frac{r R^{(d-3)/2}}{l^{d-2}} \frac{d^{(d-4)/2}}{R^{d-3}}.
\end{equation}
Hence, the inequality becomes
\begin{equation}
\Delta S(\sqrt{rR}) \geq \frac{1}{N} \sum_{k=1}^{N} \Delta S_k \approx \int_r^R dl \beta(l) \Delta S(l),
\end{equation}
where at large $N$ the sum is replaced by an integral, and we
have already replaced the contribution $\Delta S$ from wiggly
spheres by that of regular spheres. Finally, expanding for
small $R - r$ we arrive at our main result,
\begin{equation}
\Delta S' - (d-3)\Delta S' \leq 0.
\end{equation}
The entropic a theorem.—Before proving the a theorem,
let us discuss the implications of this inequality in lower
dimensions.

For $d = 2$ Eq. (11) gives
\begin{equation}
(r\Delta S')' \leq 0.
\end{equation}
In fact, this is valid directly for $S(r)$ since wiggly spheres
are just ordinary intervals. Defining $\Delta c(r) = c(r) - c_{\text{UV}} =
\Delta S'(r)$, this gets the coefficient of the logarithmic term in
the entropy for fixed points. Since it decreases with size,
Eq. (12) gives a proof of the $c$ theorem.

For $d = 3$, Eq. (11) becomes $(\Delta S'(r))'' \leq 0$ and this has
two implications. First, it gives an “area theorem,” implying
that the quantity
\begin{equation}
\Delta \tilde{\mu}_1(r) \equiv \Delta S'(r)
\end{equation}
decreases along the flow. This is finite for $\Delta < 5/2$, and
coincides with the subtracted area coefficient at fixed
points. Hence, $\Delta \tilde{\mu}_1^{\text{IR}} \leq \Delta \tilde{\mu}_1^{\text{UV}}$. For larger $\Delta$, the nonlocal
UV term discussed below Eq. (3) dominates, making $\Delta \tilde{\mu}_1$
diverge as $r \rightarrow 0$. (The area theorem in $d$ dimensions was
proved using positivity of the relative entropy in Ref. [19].)
The other consequence of the inequality is that
\begin{equation}
[r\Delta S'(r) - \Delta S(r)]' \leq 0.
\end{equation}
The CFT contribution drops out (both the area and constant
term cancel out), and hence the quantity $F(r) = rS'(r) -
S(r)$ decreases monotonically and agrees with $F$ at fixed
points. This gives a proof of the $F$ theorem; it agrees with
that in Ref. [5], where the wiggly circles were replaced by
regular ones because in $d = 3$ the wiggles do not contribute
to the SSA inequality.

Finally, let us consider $d = 4$. The CFT contribution is
\begin{equation}
S_\rho(r) = \mu_0^2 r^2 - 4A_{\text{UV}} \log(r/\epsilon).
\end{equation}
where $A_{\text{UV}}$ is the $a$-anomaly coefficient of the UV fixed
point. Replacing this into Eq. (11) obtains
\begin{equation}
rS'' \rho(r) - S' \rho(r) \leq \frac{8A_{\text{UV}}}{r}.
\end{equation}
Evaluating the left-hand side at the IR fixed point gives
\[ A_{\text{IR}} \leq A_{\text{UV}}. \tag{17} \]

This completes our proof of the a theorem using entropic techniques.

Let us emphasize that this is the point where the Markov property of the CFT plays a key role. Had we just replaced wiggly contributions by regular contributions to the entropy (instead of doing it for \( \Delta S \)), we would have obtained that the left-hand side in Eq. (16) is nonpositive. And this is violated at fixed points. Therefore, we see explicitly in this case that the entropy contributions of wiggly spheres do not tend smoothly to those of regular spheres. With our present approach we have avoided this problem by using the strong subadditivity property of \( S_{\rho} \). Therefore the Markov property of the CFT vacuum is essential for obtaining the a theorem.

Let us end with two remarks. First, an analog to a \( c \) function can be written as \( \Delta c(r) = r \Delta S'(r) - 2 \Delta r \). It is decreasing, it vanishes at the UV, and at the IR it approaches
\[ \Delta c \approx 8(A_{\text{IR}} - A_{\text{UV}}) \log(Mr), \tag{18} \]
where \( M \) is some scale of the RG. It shows the decrease of \( A \); however, it does not converge to a finite value for large \( r \). Finally, as for \( d = 3 \) we have here also an area theorem. Defining the quantity
\[ \Delta \mu_2(r) = \frac{\Delta S'(r)}{2r} = \frac{1}{2r} (S'_{\rho}(r) - S'_{\rho_0}(r)), \tag{19} \]

this is always decreasing \( \Delta \mu_2(r) \leq 0 \). For \( \Delta < 3 \) it is finite and approaches the subtracted area coefficient at fixed points. Hence \( \Delta \mu_2^{\text{IR}} \leq \Delta \mu_2^{\text{UV}} \). In \( d = 2 \) the area theorem coincides with the \( c \) theorem, as discussed in Ref. [19].

**Extension to higher dimensions and final remarks.**—For dimensions higher than \( 4 \) we have more than two coefficients of the entropies \( S_{\rho} \) and \( S_{\rho_0} \) in the IR. Equation (11) gives two relevant inequalities. The first is for the area term. This follows from the interpolating quantity
\[ \Delta \hat{\mu}_{d-2}(r) = \frac{\Delta S'(r)}{(d - 2)r^{d-3}} \tag{20} \]
that always decreases. From Eq. (2), the structure of UV divergences ignoring order one coefficients is
\[ \Delta \hat{\mu}_{d-2}(r) = g^2 \epsilon^{d+2-2\Delta} \left(1 + \frac{\epsilon^2}{r^2} + \cdots \right) + \text{finite}. \tag{21} \]

In the \( UV \) \( r \ll g^{-1/(d-\Delta)} \), the finite term is of order \( g^2 r^{d+2-2\Delta} \). Near the IR fixed point we expect, on dimensional grounds, a finite term of order \( g^{(d-2)/(d-\Delta)} \).

Therefore, \( \Delta \hat{\mu}_{d-2}(r) \) is finite for \( \Delta < (d + 2)/2 \) and interpolates between area terms, so \( \Delta \mu_{d-2}^{\text{IR}} \leq \Delta \mu_{d-2}^{\text{UV}} \). However, if \( \Delta \geq (d + 2)/2 \), \( \Delta \hat{\mu}_{d-2}(r) \) is divergent, while its change with \( r \) can still be finite if \( \Delta < (d + 4)/2 \). The total running of this quantity from \( r = 0 \) to \( r = \infty \) is infinite due to the finite terms in the UV.

The other inequality comes from observing the IR value of Eq. (11). This is dominated by the next leading term proportional to \( r^{d-4} \) in the entropies and gives
\[ \Delta \mu_{d-4}^{\text{IR}} \geq 0. \tag{22} \]

This is finite or not according to whether \( \Delta < (d + 4)/2 \) or \( \Delta \geq (d + 4)/2 \), respectively. For \( d = 4 \) this gives the a theorem discussed before.

The area term is related to the renormalization of Newton’s constant. Along similar lines, it would be interesting to analyze the implications of Eq. (22) for gravitational corrections.

It seems strong subadditivity does not allow us to examine the other terms—in particular we cannot get to the terms that are universal for CFTs in \( d \geq 5 \). However, this suggests that the renormalization of \( \Delta \mu_{d-4} \) may have alternating signs \((-1)^{d/2}\). We have shown this for \( k = 2, 4 \), that in \( d \leq 4 \) give the \( c \), \( F \), and \( a \) theorems. The statement for the last term in the expansion of the entropies of spheres corresponds to the irreversibility of the RG in any dimension. This sign is in agreement with the expected alternating signs of the universal coefficients.

Let us conclude by discussing the connection with relative entropy. The Markov property is equivalent to the cancellation
\[ H_A + H_B - H_{A \cap B} - H_{A \cup B} = 0 \tag{23} \]

of modular Hamiltonians for a CFT [10]. Hence, \(-\Delta S\) can be replaced by the relative entropy \( \mathcal{S}_{\text{rel}}(\rho^1|\rho^0) \) without modifying the outcome of the inequalities. We hope to revisit these results in terms of relative entropies, extending previous work on the RG flow [19]. This would also include, in the same scheme, the \( g \) theorem for CFTs with defects [20].

We thank Marina Huerta for discussions and encouragement. This work was partially supported by CONICET, CNEA, and Universidad Nacional de Cuyo, Argentina. H.C. acknowledges an “It From Qubit” grant of the Simons Foundation. G.T. is also supported by ANPCYT PICT Grant No. 2015-1224.


