Black Holes Conserve Information in Curved-Space Quantum Field Theory

Christoph Adami[1,3] and Greg L. Ver Steeg[2]
[1] *Jet Propulsion Laboratory 126-347, California Institute of Technology, Pasadena, CA 91109*
[2] *California Institute of Technology 103-33, Pasadena, CA 91125*
[3] *Keck Graduate Institute, Claremont, CA 91711*

(Dated: August 17, 2004)

We show that black hole formation and evaporation in curved-space quantum field theory is unitary if *stimulated* (as well as spontaneous) emission at the event horizon is taken into account. In particular, we show that the entropy accreted by a black hole when particles cross the event horizon is exactly balanced by a commensurate entropy increase of the rest of the universe, owing to the quantum entanglement between the black hole, Hawking radiation, and scattered radiation (including stimulated emission). As a consequence, the emitted radiation is *non-thermal*, and information can be retrieved using standard error corrections methods for noisy quantum channels.

PACS numbers: 04.70Dy,03.67-a,03.65.Ud

Black hole evaporation poses a serious challenge to theoretical physics because it is a problem at the intersection of general relativity and quantum mechanics, precisely where our understanding is the weakest. From the moment Hawking discovered the eponymous radiation effect [1], many researchers have concluded that black holes introduce an intrinsic *irreversibility* into the laws of physics, together with an intrinsic *unpredictability* [2]. The reasons for such a radical departure from conventional physics at first seem incontrovertible. If a black hole's entropy is given approximately by the logarithm of the number of possible different initial configurations of matter that collapsed into the black hole, then predictability is lost *even before evaporation*, because no observer can perform measurements within the event horizon of the black hole. But black hole evaporation poses an even more vexing problem than the loss of predictability, because it seems to imply that quantum mechanical pure states can evolve into mixed states, implying a violation of unitarity and probability conservation in general.

It is the purpose of this letter to point out that none of these problems in fact occur if black holes are described within quantum information theory rather than thermodynamics. As is well-known, because black holes have negative heat capacity (the more heat they radiate, the hotter they become), they can never be in a state of thermodynamical equilibrium with an infinite reservoir or heat bath. Instead, a description should be used that does *not* assume that "all 'fast' things have happened and all the 'slow' things not" [3], and in which probability distributions can be arbitrary rather than of the Boltzmann type. Such a description is information theory (see, e.g., Ref. [4] for its relation to statistical physics), and its quantum version (see, e.g., [5]) should be used to describe black hole dynamics.

In units where $\hbar = c = G = k = 1$, a non-charged, non-rotating Schwarzschild black hole has an entropy given by $S_{\rm BH} = 4\pi M^2$, where M is the mass of the black hole, and a temperature $T_{\rm BH} = (8\pi M)^{-1}$, the Hawking temperature. Black holes are formed in stellar collapse of stars of sufficient mass, and can subsequently accrete particles (see Fig. 1a). Through the process of virtual pair formation near the event horizon, with one of the pair's particles disappearing behind the event horizon and the other going off to future infinity, a black hole loses mass (by providing it to the virtual pair which, in return, goes on mass shell), and ultimately disappears. The problem of loss of predictability is easily seen in Fig. 1(a), where different trajectories of particles accreting onto the black hole are outlined, and arbitrarily labelled. If the identity of these labels is lost behind the horizon, we are faced with a many-to-one mapping and concurrent loss of predictability (even if this information is released at a later time) because the future state of the black hole cannot be ascertained given the labels of the accreting particles. If the labels are destroyed in the singularity and the information is thus never released, then we would be forced to assume that, to make matters worse, the laws of physics allow non-unitary dynamics, and do not conserve probabilities.

Hawking radiation occurs because the gravitational field at the event horizon polarizes the vacuum, leading to spontaneous emission of particles that can be detected by stationary detectors at infinity. From a quantum field-theoretic point of view, the scattering cross section for spontaneous particle emission would simply be proportional to the imaginary part of the polarization operator in a background gravitational field, if we were

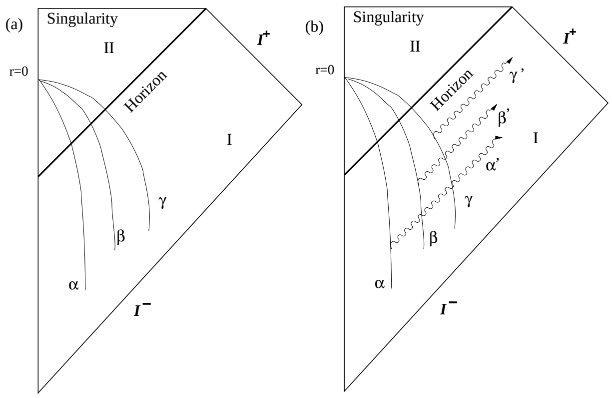

FIG. 1: (a) Penrose diagram of the accretion of arbitrarily labelled particles from past infinity (I^-) onto a black hole. Once the particles cross the event horizon into region II, they are indistinguishable (leading to a loss of information) unless they leave a signature outside (in region I) via stimulated emission, as indicated in (b). This radiation escapes to future infinity (I^+) along with the spontaneous Hawking radiation (not shown).

able to calculate that. By continuing this reasoning, we see that absorption of particles near the event horizon should involve the imaginary part of the self-energy operator in the external field, which describes stimulated emission of radiation. Indeed, standard arguments going back to Einstein ensure that the transition probability for stimulated emission is directly related to that of spontaneous emission. Even though it has been realized before that black holes emit stimulated radiation (see, e.g., [6, 7, 8, 9, 10, 11]), it has apparently been overlooked that stimulated emission guarantees that black holes evolve in a unitary manner, as we now show.

In a quantum theory of scalar fields in a gravitational background, the Minkowski vacuum is a non-trivial state due to the separation of space-time into two regions, generally depicted as the "outside" and "inside" region, or region I and region II, respectively (Fig. 1). The operators that annihilate this vacuum, A_k and B_k (A_k annihilates the "outside" vacuum while B_k annihilates "inside"), are related to the flat space operators a_k and b_k by a Bogoliubov transformation [12]

$$A_k \;=\; e^{iH} a_k e^{-iH} = \alpha a_k - \beta b^\dagger_{-k} \qquad (1)$$

$$B_k \;=\; e^{iH} b_k e^{-iH} = \alpha b_k - \beta a^\dagger_{-k} \qquad (2)$$

where $a^\dagger_k$ creates a particle of mode k while a_{-k} annihilates an antiparticle of the same mode. Let H be a Hamiltonian describing the interaction of particles and antiparticles outside and inside the horizon (for mode k),

$$H = ig(a^\dagger_k b^\dagger_{-k} - a_k b_{-k} + k \to -k) \; . \qquad (3)$$

The Minkowski vacuum is then obtained by mapping from the trivial vacuum $|0\rangle_g$ with a unitary operator $U = e^{-iH}$

$$|0\rangle_M = e^{-iH} |0\rangle_g \; . \qquad (4)$$

In terms of the interaction strength g, it follows that

$$\alpha^2 \;=\; \cosh^2 g = \frac{1}{1 - e^{-\omega/T}} \qquad (5)$$

$$\beta^2 \;=\; \sinh^2 g = \frac{1}{e^{\omega/T} - 1} \; , \qquad (6)$$

where $\omega = |k|$ is the frequency associated to mode k and T is the Hawking temperature. Indeed, as is well known, the Bogoliubov transformation (1,2) leads to a thermal spectrum

$${}_M\langle 0|a^\dagger_k a_k|0\rangle_M = \beta^2 = \frac{1}{e^{\omega/T} - 1} \; , \qquad (7)$$

i.e., the Planckian distribution of spontaneous emission: the Hawking radiation. The Minkowski vacuum for all modes k can be written in terms of a superposition of particle and anti-particle states living in region I and II respectively, as (see, e.g., [13])

$$|0\rangle_M = \Pi_k \frac{1}{\alpha_k^2} \sum_{n_k n'_k} e^{-\frac{(n_k + n'_k)\omega_k}{2T}} |n_k, n'_{-k}\rangle_I |n'_k, n_{-k}\rangle_{II} \; . \qquad (8)$$

In the following, we will focus on a single mode k because density matrices for different modes factor. Thus, for example, we obtain the density matrix of mode k in the outside region (when no particles are incident) as

$$\rho_I \;=\; \mathrm{Tr}_{II} |0\rangle_M\langle 0| = \rho_{k|0} \otimes \rho_{-k|0} \qquad (9)$$

with particle viz. antiparticle thermal density matrices given by

$$\rho_{k|0} = (1 - e^{-\omega/T}) \sum_{n_k=0}^{\infty} e^{-n_k\omega/T} |n_k\rangle_g\langle n_k| \; , \qquad (10)$$

and where $k|0$ indicates that these are obtained without incoming particles. From (9) we can easily calculate the von Neumann entropy of a mode k of the inside or outside region as

$$S(\rho) = 2\left[\frac{\omega/T}{e^{\omega/T} - 1} + \log(1 - e^{-\omega/T})\right] \qquad (11)$$

with equal contributions from particles and antiparticles. This entropy (sometimes called "entanglement entropy") is generated entirely by the existence of the event horizon separating regions I and II, and indeed appears to be bounded by the area of the horizon, rather than the volume [13, 14]. It is clear, however, that such a calculation needs to be supplemented by a rigorous regularization scheme that removes both infrared and ultraviolet divergencies [15, 16, 17] in the sum over modes before we can compare this expression to the Bekenstein-Hawking entropy.

While an identification of the black hole entropy with entanglement entropy is appealing, it does not solve the information paradox. In particular, the entropy (11) is clearly that of perfectly thermal radiation. To solve the information paradox, we need to show that entropy is conserved in the accretion and evaporation process, and that the emitted radiation is indeed non-thermal. It is sufficient to show this for a single mode k because modes of different k do not interfere ($[a_k, a_m^\dagger] = \delta_{km}$), and because an arbitrary amount of information can be encoded into a stream of particles of a single mode accreting onto a black hole.

To describe black hole dynamics, we construct a final state $|n_k\rangle_M$ obtained from the interaction of n_k incident particles with a region II bounded by an event horizon:

$$|n_k\rangle_M = \frac{1}{\sqrt{n!}}(A_k^\dagger)^n|0\rangle_M \; . \tag{12}$$

In order to describe absorption, we also need to consider the elastic scattering of particles off the black hole by introducing a scattering Hamiltonian H_s so that

$$e^{iH_s}a_k e^{-iH_s} = \sqrt{1-\Gamma_0}\; a_k + \sqrt{\Gamma_0}\; b_k \; , \tag{13}$$

where $1-\Gamma_0$ and Γ_0 stand for the elastic scattering and absorption probabilities for a single mode, respectively. Thus, using $H_{\rm tot} = H + H_s$, we obtain

$$|\psi_n\rangle = |n_k\rangle_M = \frac{1}{\sqrt{n!}}(e^{iH_{\rm tot}}a_k^\dagger e^{-iH_{\rm tot}})^n|0\rangle_M =$$
$$\frac{1}{\sqrt{n!}}\left[\sqrt{1-\Gamma_0}\;(\alpha a_k^\dagger - \beta b_{-k}) + \sqrt{\Gamma_0}(\alpha b_k^\dagger - \beta a_{-k})\right]^n|0\rangle_M \; . \tag{14}$$

From this expression, we can calculate the density matrix of region I after n particles interacted by tracing over the degrees of freedom of region II:

$$\rho_I^{(n)} = {\rm Tr}_{II}|\psi_n\rangle\langle\psi_n| \tag{15}$$

with entropy

$$S(\rho_I^{(n)}) = -{\rm Tr}_I(\rho_I^{(n)}\log\rho_I^{(n)}) \; . \tag{16}$$

Because the entropy of the pure state $|\psi_n\rangle\langle\psi_n|$ vanishes, we are assured that the marginal entropy of region I and II are always equal, for any n. Thus, entropy is conserved during accretion and evaporation.

Calculating the mean number of particles N_I emitted into region I when n particles of mode k are incident is instructive, as

$$N_I = \langle\psi_n|a_k^\dagger a_k|\psi_n\rangle = (1-\Gamma_0)n_k + \beta^2\left[1 + n_k(1-\Gamma_0)\right] \; . \tag{17}$$

The first term in (17) is the number of elastically scattered particles, while the second one (proportional to β^2) contains the contribution from spontaneous and stimulated emission, respectively. Note that the fraction Γ_0 of particles that disappear behind the horizon leaves a signature of stimulated *anti*-particles in the outside region, as schematically indicated in Fig. 1b.

Black hole dynamics can now be cast into the language of quantum information theory. We can follow the fate of information interacting with a black hole by using a *preparer* to encode information into quantum states that are then sent into the event horizon. For example, we can imagine a preparer X who sends packets of n_k particles with probability $p(n)$. The internal state of the preparer can be described by the density matrix $\rho_X = \sum_n p(n)|n\rangle\langle n|$, with entropy $S(\rho_X) = H[p] = -\sum_n p(n)\log p(n)$. After the particles interact with the black hole, our preparer is now correlated with it because the final state is now the density matrix

$$\rho_{I,II,X} = \sum_n p(n)|n\rangle_M\langle n| \otimes |n\rangle_X\langle n| \; . \tag{18}$$

Tracing over the black hole interior (region II), we obtain the joint density matrix of the radiation field in region I and the preparer X:

$$\rho_{I,X} \;=\; \sum_n p(n)\rho_I^{(n)} \otimes |n\rangle_X\langle n| \tag{19}$$

with entropy

$$S(\rho_{I,X}) = H[p] + \sum_n p(n)S(\rho_I^{(n)}) \tag{20}$$

owing to the block-diagonal form of Eq. (19). The mutual entropy between radiation field and preparer is then simply given by

$$\begin{aligned} H(X:I) \;&=\; S(\rho_I) + S(\rho_X) - S(\rho_{I,X}) \\ &=\; S(\rho_I) - \sum_n p(n)S(\rho_I^{(n)}) \; , \end{aligned} \tag{21}$$

which is known as the Holevo bound. The latter constitutes the maximum amount of classical information that can be extracted from a quantum measurement [18], and its maximum (over the probability distribution of signal states) turns out to be the capacity of a quantum channel to transmit classical information [19].

We can calculate as an example the simple binary case where the preparer either sends no particle (with probability $1-p$) or one particle (with probability p) into the black hole. In order for this information to be recovered, an outside observer must be able to make measurements on the radiation field that betray the preparer's decision.

In the case of pure scattering or pure absorption, $\rho_I^{(1)}$ is diagonal. For pure scattering ($\Gamma_0 = 0$) we find

$$\rho_{k|1} = \frac{1}{\alpha^4}\sum_m (m+1)e^{-m\omega/T}|m+1\rangle\langle m+1| \tag{22}$$

As noted before, $S(\rho_{k|0}) = S_{\rm therm}$ is thermal. $S(\rho_{k|1})$, on the other hand, is not. Instead, we find

$$S(\rho_{k|1}) = 2S_{\rm therm} - F(\omega/T)\;, \tag{23}$$

with a positive definite quantity

$$F(\omega/T) = \frac{1}{\alpha^2\beta^2}\sum_{m=1}^{\infty} e^{-m\omega/T} m\log m\;. \tag{24}$$

In Fig. 2, we show the Holevo bound

$$\begin{aligned} H(I:X) &= S[(1-p)\rho_{k|0} + p\rho_{k|1}] \\ &\quad -(1-p)S(\rho_{k|0}) - pS(\rho_{k|1}) \end{aligned} \tag{25}$$

as a function of the probability p (note that the contribution from $\rho_{-k|0}$ cancels because it is thermal). Due to the quantum noise created by the Hawking radiation, $H(I:X)$ does not reach the initial entropy $H[p]$, but approaches it very quickly as ω/T becomes larger (see Fig. 2). This loss of information is entirely due to the degradation of the signal that is sent through a noisy quantum channel, and can be restored using standard quantum error correction schemes (see, e.g., [5]).

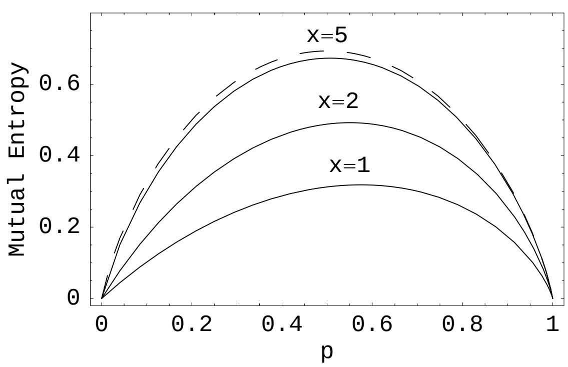

FIG. 2: Mutual entropy $H(X:I)$ between preparer and radiation field as a function of the probability p in the binary black hole channel. The Holevo information $H(X:I)$ approaches the maximal amount of information that can be extracted, $H[p]$ (dashed line) as $x = \omega/T$ increases, i.e., as the Hawking noise decreases.

The case for full absorption can be treated in the same manner (now $\rho^{(1)} = \rho_{k|0} \otimes \rho_{-k|1}$, so replace $k \to -k$ in Eq. 25) except that $\rho_{-k|0}$ and $\rho_{-k|1}$ are more and more difficult to distinguish as $T \to 0$. However, they can still be discerned probabilistically as long as $T > 0$, which is sufficient to ensure full information retrieval with appropriate quantum error correction. While the case for general Γ_0 is more complicated because $\rho^{(1)}$ is not diagonal in the number basis, a numerical diagonalization shows that the capacity of the black hole communication channel is always non-zero, and decreases as Γ_0 increases [20]. Note also that because the capacity of the channel does not depend on previous uses (except for a change in the black hole temperature), the precise nature of the black hole history is immaterial to information transmission.

In conclusion, we found that a consistent treatment of black hole dynamics requires the presence of stimulated radiation outside the event horizon beyond the usual Hawking radiation. The radiation field in this region is non-thermal, while accretion and evaporation described by the Hamiltonian $H_{\rm tot} = H + H_s$ is unitary. This guarantees that any change in the entropy of the inside region is exactly balanced by a commensurate change in the outside. Consequently, while particles can cross the event horizon, information cannot. Instead, the information dynamics of black holes turns out to be that of a very standard noisy quantum channel that is used to transmit classical information, where Hawking radiation provides the noise source. We further note that this solution to the black hole information paradox does not require to go beyond curved-space quantum field theory.

Acknowledgements We are grateful to N.J. Cerf and C.O. Wilke for crucial discussions and comments on the manuscript, as well as to J.P. Dowling, G.M. Hockney, P. Kok, H. Lee, and U. Yurtsever for discussions. This work was carried out in part at the Jet Propulsion Laboratory (California Institute of Technology) under a contract with NASA, with support from the Army Research Office's grant # DAAD19-03-1-0207.

[1] S. W. Hawking, Commun. Math. Phys. **43**, 199 (1975).
[2] S. W. Hawking, Phys. Rev. D **14**, 2460 (1976).
[3] R. P. Feynman, *Statistical Mechanics: A Set of Lectures* (Addison Wesley, Reading, MA, 1972).
[4] E. T. Jaynes, Phys. Rev. **106**, 620 (1957).
[5] J. Preskill, *Quantum Computation and Information*, Phy 219 Lecture Notes (Caltech) (1998).
[6] R. M. Wald, Phys. Rev. D **13**, 3176 (1976).
[7] J. D. Bekenstein and A. Meisels, Phys. Rev. D **15**, 2775 (1977).
[8] P. Panangaden and R. M. Wald, Phys. Rev. D **16**, 929 (1977).
[9] R. D. Sorkin, Class. Quant. Grav. **5**, L149 (1987).
[10] J. Audretsch and R. Müller, Phys. Rev. D **45**, 513 (1992).
[11] M. Schiffer, Phys. Rev. D **48**, 1652 (1993).
[12] N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge U. Press, Cambridge, 1982).
[13] A. Iorio, G. Lambiase, and G. Vitiello, Ann. Phys. **309**, 151 (2004).
[14] S. Mukhoyama, Ph.D. thesis, Kyoto University (1998), gr-qc/9809050.
[15] L. Bombelli, R. K. Koul, J. Lee, and R. D. Sorkin, Phys. Rev. D **34**, 373 (1986).
[16] C. Callan and F. Wilczek, Phys. Lett. B **333**, 55 (1994).
[17] C. Holzhey, F. Larsen, and F. Wilczek, Nucl. Phys. B **424**, 443 (1994).
[18] A. S. Holevo, Probl. Inf. Transm. (USSR) **9**, 177 (1973).
[19] A. S. Holevo, IEEE Trans. Info. Theor. **44**, 269 (1998).
[20] C. Adami and G. L. Ver Steeg, forthcoming (2004).