



This is an author produced version of *Information erasure and the generalized second law of black hole thermodynamics*.

White Rose Research Online URL for this paper:  
<http://eprints.whiterose.ac.uk/10553/>

---

**Article:**

Song, D.D. and Winstanley , E. (2008) Information erasure and the generalized second law of black hole thermodynamics. *International Journal of Theoretical Physics* , 47 (6). pp. 1692-1698. ISSN 1572-9575

<http://dx.doi.org/10.1007/s10773-007-9610-0>

---



*promoting access to  
White Rose research papers*

[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<http://eprints.whiterose.ac.uk/>

# Information Erasure and the Generalized Second Law of Black Hole Thermodynamics

**David D. Song**

*Centre for Quantum Computation, Clarendon Laboratory  
University of Oxford, Parks Road, Oxford OX1 3PU, U.K.  
d.song@qubit.org*

**Elizabeth Winstanley**

*Department of Applied Mathematics, University of Sheffield,  
Hicks Building, Hounsfield Road, Sheffield S3 7RH, U.K.  
E.Winstanley@sheffield.ac.uk*

## ABSTRACT

We consider the generalized second law of black hole thermodynamics in the light of quantum information theory, in particular information erasure and Landauer's principle (namely, that erasure of information produces at least the equivalent amount of entropy). A small quantum system outside a black hole in the Hartle-Hawking state is studied, and the quantum system comes into thermal equilibrium with the radiation surrounding the black hole. For this scenario, we present a simple proof of the generalized second law based on quantum relative entropy. We then analyze the corresponding information erasure process, and confirm our proof of the generalized second law by applying Landauer's principle.

PACS: 04.70.Dy, 03.67-a.

The correspondence between the laws of thermodynamics and black hole mechanics was noted, as a curiosity without physical implications, in a seminal paper by Bardeen, Carter and Hawking [1]. At around the same time, Bekenstein [2] was advocating a rather more radical approach. Noting the area theorem of black holes, which states that the total area of black hole event horizons can never decrease, he observed that this is analogous to the ordinary second law of thermodynamics, i.e. the total entropy of a closed system never decreases. He proposed that, multiplied by appropriate powers of the Planck length, Boltzmann constant and some dimensionless constant of order unity, the black hole area should be interpreted as its physical entropy. This proposal was given physical support by the discovery of Hawking [3] that black holes radiate at a temperature

$$T_{bh} = \frac{\kappa}{2\pi} \tag{1}$$

where  $\kappa$  is the surface gravity (here, and throughout the paper, we work in Planck units, in which  $\hbar = c = G = k_B = 1$ , where  $k_B$  is Boltzmann's constant). This, coupled with the first law of black hole mechanics [1], gives the value of the numerical constant in Bekenstein's conjecture of black hole entropy to be  $\frac{1}{4}$ .

Wheeler provided the initial motivation for Bekenstein's black hole entropy proposal [4]. Wheeler suggested a creature, subsequently called Wheeler's demon, which could violate the ordinary second law of thermodynamics by dropping entropy into a black hole, producing a decrease in the entropy outside the black hole. This led Bekenstein to conjecture that the black hole itself has an entropy (proportional to the area of the event horizon) and, furthermore, the sum of the entropy outside the black hole and the black hole entropy must not decrease,

$$\Delta S_{out} + \Delta S_{bh} \geq 0. \tag{2}$$

This generalized second law has been widely discussed in the literature, and there are proofs due to Frolov and Page [5], and, more recently Mukohyama [6]. Both these proofs make use of quantum field theory in curved space, and apply to quasistationary black holes. Frolov and Page's proof is applicable to eternal black hole space-times, whilst Mukohyama considers black holes arising from gravitational collapse.

In this paper we wish to consider the generalized second law from another point of view, namely quantum information theory and, in particular, Landauer's principle of information erasure [7]. We will consider a quantum system outside a black hole, which then comes into thermal equilibrium with the Hawking radiation surrounding the black hole. This scenario is different from those that have been considered previously in proofs of the generalized second law, giving further weight to its validity.

We firstly discuss Maxwell's demon [8], which is the analogue in ordinary thermodynamics of Wheeler's demon. Consider a container of gas which is divided into two halves, left and right, by a partition. Imagine now that there is a demon sitting on the

partition, who is able to measure the velocities of individual molecules in the gas. If the demon let the fast molecules move to the right container while keeping the slower ones to the left, then this would create a temperature difference and violate the second law of thermodynamics. Bennett noted [9] that in order to do free work, the demon has to record its measurement result, and then its memory needs to be erased in order to do the next measurement. Landauer's principle states that in order to erase a certain amount of information at least the same amount of entropy must be generated. Therefore, the erasure of the memory of the demon generates an entropy greater than or equal to the amount of recorded memory, which preserves the second law. Bennett's classical analysis of Maxwell's demon was later confirmed quantum mechanically [10, 11, 12]. This process resembles Wheeler's demon, who is trying to erase information by dropping an object into a black hole. This necessarily creates an increase of black hole entropy by at least the same amount as the dropped entropy, according to the generalized second law. In this paper, we give a simple proof of the generalized second law for our model, using known results on quantum relative entropy. This will be confirmed by our analysis of the corresponding information erasure process using Landauer's principle. The generalized second law has recently been considered in the context of quantum information theory (concentrating on the entanglement of states inside and outside the black hole event horizon) by Hosoya and collaborators [13].

Let us consider a black hole in thermal equilibrium with a heat bath at the Hawking temperature  $T_{bh}$ . This is the Hartle-Hawking state [14], and can be rendered stable by placing the black hole in a cavity whose dimensions are very much larger than the radius of the black hole event horizon, thereby forming a closed system. We consider a small quantum system outside the black hole, having Hamiltonian  $H$  and initially in a quantum state described by the density matrix  $\rho_i$ . We then suppose that the small quantum system comes into thermal equilibrium, so that its final state is the thermal state  $\rho_f = Z^{-1}e^{-\beta_{bh}H}$ , where  $Z = \text{tr}[e^{-\beta_{bh}H}]$  and  $\beta_{bh} = \frac{1}{T_{bh}}$  (we have set  $k_B$ , the Boltzmann constant, equal to unity). If the cavity is sufficiently large and the quantum system small, we may suppose that the black hole temperature is not affected by this process, so that we are concerned only with quasistationary black holes. Then the change of entropy outside the black hole is simply the entropy difference between the initial and final states which is

$$\Delta S_{out} = S(\rho_f) - S(\rho_i) = \text{tr}[-\rho_f \log \rho_f + \rho_i \log \rho_i]. \quad (3)$$

In other words, the amount of entropy  $\Delta S_{out}$  has been dropped into the black hole.

In order to evaluate the change of black hole entropy, we first calculate the change in energy outside the black hole. Then, by conservation of energy, this will be minus the change in energy of the black hole. Then, using the usual first law of thermodynamics, dividing by the black hole temperature (and the Boltzmann constant), the change in black hole entropy is given by

$$\Delta S_{bh} = -\beta_{bh} \text{tr}[H(\rho_f - \rho_i)]. \quad (4)$$

Since  $\rho_f$  is a thermal state,  $H = -\beta_{bh}^{-1} \log(Z\rho_f)$ , which gives

$$\begin{aligned} \Delta S_{bh} &= tr[\rho_f \log(Z\rho_f)] - tr[\rho_i \log(Z\rho_f)] \\ &= -tr[(\rho_i - \rho_f) \log \rho_f] + [tr(\rho_f) - tr(\rho_i)] \log Z \\ &= -tr[(\rho_i - \rho_f) \log \rho_f]. \end{aligned} \tag{5}$$

The final line follows by conservation of probability. Note that this does not assume that the states evolve unitarily.

Therefore the total change in entropy can be written as follows

$$\Delta S_{out} + \Delta S_{bh} = tr[\rho_i \log \rho_i - \rho_i \log \rho_f]. \tag{6}$$

At this stage it is important to note that, in common with other proofs of the generalized second law, we have had to use the first law. Here we have used the ordinary first law of thermodynamics, although this is directly analogous to the first law of black hole mechanics for quasistationary black holes, and gives the same result.

We should also emphasise at this stage that the process we are considering here is different from the usual gedanken experiment of Bekenstein [2] in which a system containing entropy is dropped down the black hole horizon. Here we consider instead a system which comes into thermal equilibrium with the radiation outside the black hole event horizon. This system will be pertinent to our subsequent consideration of information erasure and Landauer's principle.

The quantity (6) is known as quantum relative entropy which is defined as  $S(\sigma||\rho) = tr[\sigma \log \sigma - \sigma \log \rho]$ . Quantum relative entropy  $S(\sigma||\rho)$  has been shown [15] to be always non-negative and is zero if and only if  $\sigma = \rho$ . Therefore (6) is non-negative and this proves the generalized second law.

Quantum relative entropy has been shown to have various applications in quantum information theory (see [16], for example) including entanglement quantification. We now give a simple example to illustrate this concept. The unit of quantum information is called a quantum bit or qubit. A qubit is a superposition of  $|0\rangle$  and  $|1\rangle$ , an orthonormal basis in a two-dimensional Hilbert space. For example, a spin- $\frac{1}{2}$  state can be considered as a qubit where  $|0\rangle$  and  $|1\rangle$  are spin up and down states. An entanglement of two qubits in subsystems  $A$  and  $B$  can be written as  $|\psi\rangle_{AB} = a|00\rangle_{AB} + b|11\rangle_{AB}$  where  $a^2 + b^2 = 1$  (if we assume  $a, b$  are real). Then the von Neumann entropy of the reduced density matrix of  $|\psi\rangle_{AB}$ , which is  $-a^2 \log a^2 - b^2 \log b^2$ , yields a good measure of the entanglement. However it is not easy to determine the amount of entanglement for mixed states with von Neumann entropy. Relative entropy,  $S(\sigma||\rho)$ , has been shown [17] to be useful in quantifying entanglement for both pure and mixed states. For pure states, relative entropy would reduce to the von Neumann entropy. Let us consider, as an example, a  $|\psi\rangle_{AB}$  which in density matrix form is written as

$$\sigma_{AB} = \begin{pmatrix} a^2 & ab \\ ab & b^2 \end{pmatrix}. \tag{7}$$

In order to give a correct measure of entanglement,  $\rho_{AB}$  in  $S(\sigma_{AB}||\rho_{AB})$  satisfies the following conditions: (1) it is disentangled (i.e.  $\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i$ ) and (2)  $S(\sigma_{AB}||\rho_{AB})$  is minimal. For pure states, a  $\rho_{AB}$  satisfying both these conditions can always be found as  $\sigma_{AB}$  with off-diagonal terms set to zero, i.e.

$$\rho_{AB} = \begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix}. \quad (8)$$

With  $\sigma_{AB}$  and  $\rho_{AB}$  in (7) and (8), we could calculate  $tr[\sigma_{AB} \log \sigma_{AB} - \sigma_{AB} \log \rho_{AB}]$  where  $tr[\sigma_{AB} \log \sigma_{AB}]$  is zero since  $\sigma_{AB}$  is a pure state. The second term,  $tr[-\sigma_{AB} \log \rho_{AB}]$  yields  $-a^2 \log a^2 - b^2 \log b^2$  which is same as the von Neumann entropy of  $|\psi\rangle_{AB}$ .

We shall now relate our black hole process and the generalized second law to Landauer's principle of information erasure. First, we briefly review some of the key ideas, and a mechanism for the erasure of information. Landauer's principle of information erasure (that the erasure of a certain amount of information produces at least the equivalent amount of entropy) has been used to explain some of the fundamental aspects in quantum information theory. The entanglement shared by two parties can be manipulated into another state by local operation and classical communication (LOCC). However it is known that local operation cannot increase the entanglement shared by two separated parties. For example, the conversion from  $a|00\rangle_{AB} + b|11\rangle_{AB}$ , where  $a, b \neq \frac{1}{\sqrt{2}}$ , to  $\frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB})$  (known as entanglement purification) cannot be done with probability 1 by LOCC because the entanglement has increased from  $-a^2 \log a^2 - b^2 \log b^2$  to  $\log 2$ . Vidal [18] has shown that Landauer's principle yields an upper bound for entanglement purification, linking no local increase of entanglement to the second law of thermodynamics. Another fundamental idea in quantum information theory is the Holevo bound [19] which limits the amount of classical information encoded in quantum mixed states that can be recovered. Plenio showed [20] how this Holevo bound may be illustrated using Landauer's principle. In the following, we show how information erasure may be realized physically, following Lubkin's method [11], as presented in [18, 20]. We refer the reader to [20] for details of the method.

Let us consider a quantum state of a system  $S$

$$|\psi\rangle_S = \sum_i \sqrt{\lambda_i} |a_i\rangle, \quad (9)$$

and an apparatus  $M$ , initially in some pure state. In order to make a measurement,  $M$  interacts with the state  $|\psi\rangle_S$  and entangles itself as

$$\sum_i \sqrt{\lambda_i} |a_i\rangle_S |m_i\rangle_M. \quad (10)$$

The state of the apparatus  $M$  can be obtained by tracing over the system  $S$  in (10), which yields

$$\rho = \sum_i \lambda_i |m_i\rangle \langle m_i|, \quad (11)$$

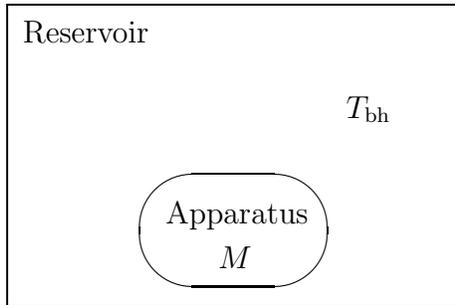


Figure 1: The apparatus  $M$  is thrown into a reservoir with black hole temperature  $T_{bh}$  and then reaches thermal equilibrium.

i.e. with probability  $\lambda_i$  the apparatus is in state  $|m_i\rangle$ . After the measurement, the apparatus will therefore be in one of these pure states, with the associated probability. The general way to erase the information of apparatus is to put the apparatus into a thermal reservoir. The apparatus reaches thermal equilibrium with the reservoir and we then bring in another pure state to perform the next measurement. Then the erasure entropy has two parts: one is the entropy change of apparatus due to its change of state from one of the pure states in (11) to a state which is in thermal equilibrium, and the other is the entropy change of the reservoir due to the apparatus.

In order to make this information erasure process applicable to black holes, we choose a reservoir which is at the black hole temperature,  $T_{bh}$ . Then, as shown in Figure 1, the apparatus  $M$  in state  $\rho$  is thrown into the reservoir with temperature  $T_{bh}$ . After the apparatus is thrown into the reservoir, it reaches thermal equilibrium and the state becomes

$$\omega = Z^{-1}e^{-\beta_{bh}H} \quad (12)$$

where  $\beta_{bh} = \frac{1}{T_{bh}}$ , with  $H$  the Hamiltonian of the apparatus, and the partition function is  $Z = \text{tr}[e^{-\beta_{bh}H}]$ . We imagine the process divided into two parts, firstly the entropy of the apparatus is reduced to zero (destroying the information in the apparatus), and then increased by the erasure process from zero to its final value. This is equivalent to the process described in [20], where the apparatus is in a pure state before the erasure takes place. The erasure entropy is given by the sum of the entropy changes of the apparatus and the reservoir. The entropy change of the apparatus is simply [20] (since we have already reduced its entropy to zero)

$$\Delta S_{app} = -\text{tr}[\omega \log \omega]. \quad (13)$$

The entropy change of the reservoir can be obtained by a method similar to that used previously for the black hole, again assuming that the apparatus is much smaller than the reservoir so that the temperature is not altered. We evaluate the heat change of the reservoir and then divide by the black hole temperature (and Boltzmann constant), to

give

$$\begin{aligned}
\Delta S_{res} &= -\beta\{tr[\omega H] - tr[\rho H]\} \\
&= tr[\omega \log(\omega Z)] - tr[\rho \log(\omega Z)] \\
&= tr[(\omega - \rho) \log \omega],
\end{aligned} \tag{14}$$

where we have again used conservation of probability. The entropy of erasure is then

$$\Delta S_{era} = \Delta S_{app} + \Delta S_{res} \tag{15}$$

$$= -tr[\rho \log \omega]. \tag{16}$$

Therefore if we consider the entropy of the lost information as  $\Delta S_{inf} = 0 - (-tr[\rho \log \rho])$ , then

$$\Delta S_{inf} + \Delta S_{era} = tr[\rho \log \rho - \rho \log \omega]. \tag{17}$$

With the identification of  $\rho$  and  $\omega$  as  $\rho_i$  and  $\rho_f$  in the black hole case, respectively, (17) yields the same result as in (6).

Application of Landauer's principle then tells us that (17) must be positive, confirming our proof of the generalized second law. However note that  $\Delta S_{era}$  is not equal to  $\Delta S_{bh}$ , contrary to intuition. The reason for this lies in the details of the erasure process [20]. As described above, the apparatus is in a pure state before the erasure process takes place, so that the change in entropy of the apparatus due to the erasure process (13) involves only the change between this pure state and the thermal state  $\omega$ , rather than between the thermal state and the initial state  $\rho$ .

Our two approaches in this paper are therefore complementary. In the first method, we change the entropy of the system outside the black hole, and there is a corresponding change in the entropy of the black hole. In the second scenario, we work out how much information we are losing in terms of destroying the initial state, and then calculate the amount of entropy required in order to erase this amount of information. Note that in the black hole situation, the information is effectively destroyed because the final state of the quantum system is the same thermal state as the surrounding radiation. We also emphasise that in the first scenario all the entropy lost goes down the black hole event horizon, and the entropy of the thermal radiation surrounding the black hole does not change.

In conclusion, in this paper we have considered the generalized second law from the point of view of quantum information theory, especially information erasure and Landauer's principle. We have considered a quantum system outside a black hole, which comes into thermal equilibrium with the radiation surrounding the event horizon. For this situation, we have been able to give a simple proof of the generalized second law of black hole thermodynamics by appealing to known results on quantum relative entropy. This result is confirmed by an analysis of the corresponding information erasure process

using Landauer's principle. This illustrates the power of quantum information theoretic ideas when applied to black hole processes.

**Acknowledgment** DS is grateful to Sougato Bose, Lucien Hardy, Vlatko Vedral and Ernesto Galvão for valuable discussions.

## References

- [1] J. Bardeen, B. Carter and S.W. Hawking, *Comm. Math. Phys.* **31**, 161 (1973).
- [2] J.D. Bekenstein, *Lett. Nuovo Cimento* **4**, 737 (1972); *Phys. Rev. D* **7**, 2333 (1973).
- [3] S.W. Hawking, *Nature* **248**, 30 (1974); *Comm. Math. Phys.* **43**, 199 (1975).
- [4] J.D. Bekenstein, *Sci. Am.*, January (1980).
- [5] V.P. Frolov and D.N. Page, *Phys. Rev. Lett.* **71**, 3902 (1993).
- [6] S. Mukohyama, *Phys. Rev. D* **56**, 2192 (1997).
- [7] R. Landauer, *IBM J. Rev. Develop.* **5**, 183 (1961).
- [8] J.C. Maxwell, *Theory of Heat* (Appleton, London, 1871).
- [9] C.H. Bennett, *Int. J. Theor. Phys.* **21**, 905 (1982).
- [10] W.H. Zurek, *Maxwell's demon, Szilard's engine and quantum measurements*, in G.T. Moore and M.O. Scully, *Frontiers of Non-equilibrium Statistical Physics*, (Plenum Press, NY, 1984).
- [11] E. Lubkin, *Int. J. Theor. Phys.* **26**, 523 (1987).
- [12] S. Lloyd, *Phys. Rev. A* **56**, 3374 (1996).
- [13] A. Hosoya, A. Carlini and T. Shimomura, Preprint *gr-qc/0008045*.
- [14] J.B. Hartle and S.W. Hawking, *Phys. Rev. D* **13**, 2188 (1976).
- [15] A. Wehrl, *Rev. Mod. Phys.* **50**, 221 (1978).
- [16] B. Schumacher and M.D. Westmoreland, Preprint *quant-ph/0004045*.
- [17] V. Vedral, M.B. Plenio, M.A. Rippin and P.L. Knight, *Phys. Rev. Lett.* **78**, 2275 (1997);  
V. Vedral and M.B. Plenio, *Phys. Rev. A* **57**, 1619 (1998).
- [18] V. Vedral, *Proc. R. Soc. Lond. A* **456**, 969 (2000).
- [19] A.S. Holevo, *Probl. Pereda. Inf.* **9**, 3 (1973) [ *Probl. Inf. Transm.* **9**, 177 (1973) ].
- [20] M. Plenio, *Phys. Lett. A* **263**, 281 (1999).