

# Entropy in thermodynamics and information theory

From Wikipedia, the free encyclopedia

There are close parallels between the mathematical expressions for the thermodynamic entropy, usually denoted by  $S$ , of a physical system in the statistical thermodynamics established by Ludwig Boltzmann and J. Willard Gibbs in the 1870s; and the information-theoretic entropy, usually expressed as  $H$ , of Claude Shannon and Ralph Hartley developed in the 1940s. Shannon, although not initially aware of this similarity, commented in it upon publicizing information theory in *A Mathematical Theory of Communication*.

This article explores what links there are between the two concepts, and how far they can be regarded as connected.

## Contents

- 1 Equivalence of form of the defining expressions
  - 1.1 Discrete case
  - 1.2 Continuous case
- 2 Theoretical relationship
- 3 Information is physical: (1) Szilard's engine
- 4 Information is physical: (2) Landauer's principle
- 5 Negentropy
- 6 Black holes
- 7 Quantum theory
- 8 The fluctuation theorem
- 9 Topics of recent research
  - 9.1 Is information quantized?
- 10 See also
- 11 External links
- 12 References

## Equivalence of form of the defining expressions

### Discrete case

The defining expression for entropy in the theory of statistical mechanics established by Ludwig Boltzmann and J. Willard Gibbs in the 1870s, is of the form:

$$S = -k \sum_i p_i \log p_i,$$

where  $p_i$  is the probability of the microstate  $i$  taken from an equilibrium ensemble.

The defining expression for entropy in the theory of information established by Claude E. Shannon in 1948 is of the form:



$$H = - \sum_i p_i \log p_i,$$

Boltzmann's tombstone,  
featuring his equation  $S = k \log W$

where  $p_i$  is the probability of the message  $m_i$  taken from the message space  $M$ .

If all the microstates are equiprobable (a microcanonical ensemble), the statistical thermodynamic entropy reduces to the form on Boltzmann's tombstone,

$$S = k \log W$$

where  $W$  is the number of microstates.

If all the messages are equiprobable, the information entropy reduces to the Hartley entropy

$$H = \log |M|$$

where  $|M|$  is the cardinality of the message space  $M$ .

The logarithm in the thermodynamic definition is the natural logarithm. It can be shown that the Gibbs entropy formula, with the natural logarithm, reproduces all of the properties of the macroscopic classical thermodynamics of Clausius. (See article: Entropy (statistical views)).

The logarithm can also be taken to the natural base in the case of information entropy. This is equivalent to choosing to measure information in nats instead of the usual bits. In practice, information entropy is almost always calculated using base 2 logarithms, but this distinction amounts to nothing other than a change in units. One nat is about 1.44 bits.

The presence of Boltzmann's constant  $k$  in the thermodynamic definitions is a historical accident, reflecting the conventional units of temperature. It is there to make sure that the statistical definition of thermodynamic entropy matches the classical entropy of Clausius, thermodynamically conjugate to temperature. For a simple compressible system that can only perform volume work, the first law of thermodynamics becomes

$$dE = pdV + TdS$$

But one can equally well write this equation in terms of what physicists and chemists sometimes call the 'reduced' or dimensionless entropy,  $\sigma = S/k$ , so that

$$dE = pdV + kTd\sigma$$

Just as  $S$  is conjugate to  $T$ , so  $\sigma$  is conjugate to  $kT$  (the energy that is characteristic of  $T$  on a molecular scale).

## Continuous case

The most obvious extension of the Shannon entropy is the differential entropy,

$$H[f] = - \int_{-\infty}^{\infty} f(x) \log f(x) dx,$$

But it turns out that this is *not* in general a good measure of uncertainty or information. For example, the

differential entropy can be negative; also it is not invariant under continuous co-ordinate transformations.

More useful for the continuous case is the **relative entropy** of a distribution, defined as the Kullback-Leibler divergence from the distribution to a reference measure  $m(x)$ ,

$$D_{\text{KL}}(f(x)\|m(x)) = \int f(x) \log \frac{f(x)}{m(x)} dx$$

(or sometimes the negative of this).

The relative entropy carries over directly from discrete to continuous distributions, and is invariant under co-ordinate reparametrisations. For an application of relative entropy in a quantum information theory setting, see eg [1] (<http://arxiv.org/abs/math-ph/0007010/>).

## Theoretical relationship

Despite all that, there is an important difference between the two quantities. The information entropy  $H$  can be calculated for *any* probability distribution (if the "message" is taken to be that the event  $i$  which had probability  $p_i$  occurred, out of the space of the events possible). But the thermodynamic entropy  $S$  refers to thermodynamic probabilities  $p_i$  specifically.

Furthermore, the thermodynamic entropy  $S$  is dominated by different arrangements of the system, and in particular its energy, that are possible on a molecular scale. In comparison, information entropy of any macroscopic event is so small as to be completely irrelevant.

However, a connection can be made between the two, if the probabilities in question are the thermodynamic probabilities  $p_i$ : the (reduced) Gibbs entropy  $\sigma$  can then be seen as simply the amount of Shannon information needed to define the detailed microscopic state of the system, given its macroscopic description. Or, in the words of G. N. Lewis writing about chemical entropy in 1930, "Gain in entropy always means loss of information, and nothing more". To be more concrete, in the discrete case using base two logarithms, the reduced Gibbs entropy is equal to the minimum number of yes/no questions that need to be answered in order to fully specify the microstate, given that we know the macrostate.

Furthermore, the prescription to find the equilibrium distributions of statistical mechanics, such as the Boltzmann distribution, by maximising the Gibbs entropy subject to appropriate constraints (the Gibbs algorithm), can now be seen as something not unique to thermodynamics, but as a principle of general relevance in all sorts of statistical inference, if it desired to find a maximally uninformative probability distribution, subject to certain constraints on the behaviour of its averages. (These perspectives are explored further in the article *Maximum entropy thermodynamics*).

## Information is physical: (1) Szilard's engine

A neat physical thought-experiment demonstrating how just the possession of information might in principle have thermodynamic consequences was established in 1929 by Szilard, in a refinement of the famous Maxwell's demon scenario.

Consider Maxwell's set-up, but with only a single gas particle in a box. If the supernatural demon knows which half of the box the particle is in (equivalent to a single bit of information), it can close a shutter between the two halves of the box, close a piston unopposed into the empty half of the box, and then extract  $k_B T \ln 2$  joules of useful work if the shutter is opened again. The particle can then be left to isothermally expand back to its original equilibrium occupied volume. In just the right circumstances therefore, the possession of a single bit of Shannon information (a single bit of negentropy in Brillouin's

term) really does correspond to a reduction in physical entropy, which theoretically can indeed be parlayed into useful physical work.

## Information is physical: (2) Landauer's principle

In fact one can generalise: any information that has a physical representation must somehow be embedded in the statistical mechanical degrees of freedom of a physical system.

Thus, Rolf Landauer argued in 1961, if one were to imagine starting with those degrees of freedom in a thermalised state, there would be a real reduction in thermodynamic entropy if they were then re-set to a known state. This can only be achieved under information-preserving microscopically deterministic dynamics if the uncertainty is somehow dumped somewhere else — ie if the entropy of the environment (or the non information-bearing degrees of freedom) is increased by at least an equivalent amount, as required by the Second Law, by gaining an appropriate quantity of heat: specifically  $kT \ln 2$  of heat for every 1 bit of randomness erased.

On the other hand, Landauer argued, there is no thermodynamic objection to a logically reversible operation potentially being achieved in a physically reversible way in the system. It is only logically irreversible operations — for example, the erasing of a bit to a known state, or the merging of two computation paths — which must be accompanied by a corresponding entropy increase.

Applied to the Maxwell's demon/Szilard engine scenario, this suggests that it might be possible to "read" the state of the particle into a computing apparatus with no entropy cost; but *only* if the apparatus has already been SET into a known state, rather than being in a thermalised state of uncertainty. To SET (or RESET) the apparatus into this state will cost all the entropy that can be saved by knowing the state of Szilard's particle.

## Negentropy

Shannon entropy has been related by physicist Léon Brillouin to a concept sometimes called negentropy. In his 1962 book *Science and Information Theory*, Brillouin described the Negentropy Principle of Information or NPI, the gist of which is that acquiring information about a system's microstates is associated with a decrease in entropy (work is needed to extract information, erasure leads to increase in thermodynamic entropy).<sup>[1]</sup> There is no violation of the second law of thermodynamics, according to Brillouin, since a reduction in any local system's thermodynamic entropy results in an increase in thermodynamic entropy elsewhere.

## Black holes

Stephen Hawking often speaks of the thermodynamic entropy of black holes in terms of their information content. Do black holes destroy information? See *Black hole entropy* and *Black hole information paradox*.

## Quantum theory

Hirschman showed in 1957, however, that Heisenberg's uncertainty principle can be expressed as a particular lower bound on the sum of the entropies of the *observable* probability distributions of a particle's position and momentum, when they are expressed in Planck units. (One could speak of the "joint entropy" of these distributions by considering them independent, but since they are not jointly observable, they cannot be considered as a joint distribution.)

# The fluctuation theorem

The fluctuation theorem provides a mathematical justification of the second law of thermodynamics under these principles, and precisely defines the limitations of the applicability of that law to the microscopic realm of individual particle movements.

## Topics of recent research

### Is information quantized?

In 1995, Tim Palmer signalled two unwritten assumptions about Shannon's definition of information that may make it inapplicable as such to quantum mechanics:

- The supposition that there is such a thing as an observable state (for instance the upper face a die or a coin) *before* the observation begins
- The fact that knowing this state does not depend on the order in which observations are made (commutativity)

The article *Conceptual inadequacy of the Shannon information in quantum measurement* [2] (<http://scitation.aip.org/getabs/servlet/GetabsServlet?prog=normal&id=PLRAAN000063000002022113000001&idtype=cvips&gifs=yes>), published in 2001 by Anton Zeilinger [3] (<http://www.quantum.univie.ac.at/zeilinger/>) and Caslav Brukner, synthesized and developed these remarks. The so-called Zeilinger's principle suggests that the quantization observed in QM could be bound to *information* quantization (one cannot observe less than one bit, and what is not observed is by definition "random").

But these claims remain highly controversial. For a detailed discussion of the applicability of the Shannon information in quantum mechanics and an argument that Zeilinger's principle cannot explain quantization, see Timpson [4] (<http://www.philosophy.leeds.ac.uk/Staff/CT/Index.htm>) 2003 [5] (<http://arxiv.org/abs/quant-ph/0112178>) and also Hall 2000 [6] (<http://arxiv.org/abs/quant-ph/0007116>) and Mana 2004 [7] (<http://arxiv.org/abs/quant-ph/0302049>)

For a tutorial on quantum information see [8] (<http://members.aol.com/jmtsgibbs/infothry.htm>).

## See also

- Thermodynamic entropy
- Information entropy
- Thermodynamics
- Statistical mechanics
- Information theory
- Physical information
- Fluctuation theorem
- Black hole entropy
- Black hole information paradox

## External links

- Entropy is Simple...If You Avoid the Briar Patches (<http://www.entropysimple.com/content.htm>). Very dismissive of wrong links between information-theoretic and thermodynamic entropy.

## References

- C. H. Bennett, "Logical reversibility of computation," IBM Journal of Research and Development, vol. 17, no. 6, pp. 525-532, 1973.
- Leon Brillouin, *Science and Information Theory*, Mineola, N.Y.: Dover, [1956, 1962] 2004. ISBN 0-486-43918-6
- Michael P. Frank, "Physical Limits of Computing" (<http://www.cise.ufl.edu/research/revcomp/physlim/plpaper.html>), *Computing in Science and Engineering*, **4**(3):16-25, May/June 2002.
- Andreas Greven, Gerhard Keller, and Gerald Warnecke, editors. *Entropy*, Princeton University Press, 2003. ISBN 0-691-11338-6. (A highly technical collection of writings giving an overview of the concept of entropy as it appears in various disciplines.)
- I. Hirschman, *A Note on Entropy*, American Journal of Mathematics, 1957.
- R. Landauer, *Information is Physical* Proc. Workshop on Physics and Computation PhysComp'92 (IEEE Comp. Sci.Press, Los Alamitos, 1993) pp. 1-4.
- R. Landauer, *Irreversibility and Heat Generation in the Computing Process* (<http://domino.research.ibm.com/tchjr/journalindex.nsf/c469af92ea9eceac85256bd50048567c/8a9d4> OpenDocument) IBM J. Res. Develop. Vol. 5, No. 3, 1961
- H. S. Leff and A. F. Rex, Editors, *Maxwell's Demon: Entropy, Information, Computing*, Princeton University Press, Princeton, NJ (1990). ISBN 0-691-08727-X
- Claude E. Shannon. *A Mathematical Theory of Communication*. Bell System Technical Journal, July/October 1948.

Retrieved from "[http://en.wikipedia.org/wiki/Entropy\\_in\\_thermodynamics\\_and\\_information\\_theory](http://en.wikipedia.org/wiki/Entropy_in_thermodynamics_and_information_theory)"

Categories: Pages needing expert attention from Science experts | Articles with sections needing expansion | Entropy | Philosophy of thermal and statistical physics

- 
- This page was last modified 15:44, 2 December 2006.
  - All text is available under the terms of the GNU Free Documentation License. (See **Copyrights** for details.) Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a US-registered 501(c)(3) tax-deductible nonprofit charity.