LibreTexts™

4.4: Entropy and Information

Gibbs Entropy

For a system with a countable number of microstates an ensemble entropy can be defined by a weighted sum over entropies of all microstates that are in turn expressed as $-k_\mathrm{B} \ln P_i$, which is analogous to Boltzmann's entropy definition for a macrostate.

$$S = -k_\mathrm{B} \sum_i P_i \ln P_i \ . \tag{4.4.1}$$

This is the definition of Gibbs entropy, while Boltzmann entropy is assigned to an individual microstate. Note that we have used a capital S because Gibbs entropy is a molecular entropy. Using Equation ???, we obtain for the system entropy $s = NS$,

$$s = -k_\mathrm{B} N \sum_i P_i \left(-\frac{\epsilon_i}{k_\mathrm{B}T} - \ln Z \right) \tag{4.4.2}$$

$$= \frac{u}{T} + k_B \ln z \ , \tag{4.4.3}$$

where we have assumed distinguishable particles, so that $\ln z = N \ln Z$. We have recovered Equation ??? that we had derived for the system entropy starting from Boltzmann entropy and assuming a canonical ensemble. For a canonical ensemble of distinguishable particles, either concept can be used. As noted above, Gibbs entropy leads to the paradox of a positive mixing entropy for combination of two subsystems made up by the same ideal gas. More generally, Gibbs entropy is not extensive if the particles are indistinguishable. The problem can be solved by redefining the system partition function as in Equation ???.

This problem suggests that entropy is related to the information we have on the system. Consider mixing of $^{13}\mathrm{CO}_2$ with $^{12}\mathrm{CO}_2$.[15] At a time when nuclear isotopes were unknown, the two gases could not be distinguished and mixing entropy was zero. With a sufficiently sensitive spectrometer we could nowadays observe the mixing process by ^{13}C NMR. We will observe spontaneous mixing. Quite obviously, the mixing entropy is not zero anymore.

This paradox cautions against philosophical interpretation of entropy. Entropy is a quantity that can be used for predicting the outcome of physical experiments. It *presumes an observer* and depends on the information that the observer has or can obtain.[16] Statistical mechanics provides general recipes for defining entropy, but the details of a proper definition depend on experimental context.

Unlike the system entropy derived from Boltzmann entropy via the canonical ensemble, Gibbs entropy is, in principle, defined for non-equilibrium states. Because it is based on the same probability concept, Gibbs entropy in an isolated system is smaller for non-equilibrium states than for equilibrium states.

Von Neumann Entropy

The concept of Gibbs entropy for a countable set of discrete states and their probabilities is easily extended to continuous phase space and probability densities. This leads to the **von Neumann entropy**,

$$S = -k_\mathrm{B} \mathrm{Trace} \left\{ \rho \ln \rho \right\} \ , \tag{4.4.4}$$

where ρ is the density matrix. Some physics textbooks don't distinguish von Neumann entropy from Gibbs entropy . Von Neumann entropy is a constant of motion if an ensemble of classical systems evolves according to the *Liouville equation* or a quantum mechanical system evolves according to the *Liouville-von-Neumann equation*. It cannot describe the approach of an isolated system to equilibrium. Coupling of the quantum mechanical system to an environment can be described by the **stochastic Liouville equation**

$$\frac{\partial \hat{\rho}}{\partial t} = -\frac{i}{\hbar} \left[\widehat{\mathcal{H}}, \hat{\rho} \right] + \widehat{\widehat{\Gamma}} \left(\hat{\rho} - \hat{\rho}_\mathrm{eq} \right) \ , \tag{4.4.5}$$

where $\widehat{\widehat{\Gamma}}$ is a Markovian operator and ρ_eq the density matrix at equilibrium. This equation of motion can describe quantum dissipative systems, i.e., the approach to equilibrium, without relying explicitly on the concept of entropy, except for the computation of ρ_eq, which relies on generalization of the Boltzmann distribution (see Section [subsection:q_partition]). However, to derive the Markovian operator $\widehat{\widehat{\Gamma}}$, explicit assumptions on the coupling between the quantum mechanical system and its environment must be made, which is beyond the scope of this lecture course.

Shannon Entropy

The concept of entropy has also been introduced into information theory. For any discrete random number that can take values a_j with probabilities $P(a_j)$, the **Shannon entropy** is defined as

$$H_{\text{Shannon}}(a) = -\sum_j P(a_j) \log_2 P(a_j).$$

(4.4.6)

A logarithm to the basis of 2 is used here as the information is assumed to be coded by binary numbers. Unlike for discrete states in statistical mechanics, an event may be in the set but still have a probability $P(a_j) = 0$. In such cases, $P(a_j) \log_2 P(a_j)$ is set to zero. Shannon entropy is the larger the 'more random' the distribution is, or, more precisely, the closer the distribution is to a uniform distribution. Information is considered as deviation from a random stream of numbers or characters. The higher the information content is, the lower the entropy.

Shannon entropy can be related to reduced Gibbs entropy $\sigma = S/k_{\text{B}}$. It is the amount of Shannon information that is required to specify the microstate of the system if the macrostate is known. When expressed with the binary logarithm, this amount of Shannon information specifies the number of yes/no questions that would have to be answered to specify the microstate. We note that this is exactly the type of experiment presumed in the second Penrose postulate (Section [Penrose_postulates]). The more microstates are consistent with the observed macrostate, the larger is this number of questions and the larger are Shannon and Gibbs entropy. The concept applies to non-equilibrium states as well as to equilibrium states. It follows, what was stated before Shannon by G. N. Lewis: "Gain in entropy always means loss of information, and nothing more". The equilibrium state is the macrostate that lacks most information on the underlying microstate.

We can further associate order with information, as any ordered arrangement of objects contains information on how they are ordered. In that sense, loss of order is loss of information and increase of disorder is an increase in entropy. The link arises via probability, as the total number of arrangements is much larger than the number of arrangements that conform to a certain order principle. Nevertheless, the association of entropy with disorder is only colloquial, because in most cases we do not have quantitative descriptions of order.