

Rényi Entropy and Free Energy

John C. Baez

Centre for Quantum Technologies
National University of Singapore
Singapore 117543

and

Department of Mathematics
University of California
Riverside CA 92521

email: baez@math.ucr.edu

February 15, 2011

Abstract

The Rényi entropy is a generalization of the usual concept of entropy which depends on a parameter q . In fact, Rényi entropy is closely related to free energy. Suppose we start with a system in thermal equilibrium and then suddenly divide the temperature by q . Then the maximum amount of work the system can do as it moves to equilibrium at the new temperature, divided by the change in temperature, equals the system's Rényi entropy in its original state. This result applies to both classical and quantum systems. Mathematically, we can express this result as follows: the Rényi entropy of a system in thermal equilibrium is minus the q^{-1} -derivative of its free energy with respect to temperature. This shows that Rényi entropy is a q -deformation of the usual concept of entropy.

1 Introduction

In 1960, Rényi [12] defined a generalization of Shannon entropy which depends on a parameter. If p is a probability distribution on a finite set, its Rényi entropy of order q is defined to be

$$S_q = \frac{1}{1-q} \ln \sum_i p_i^q \quad (1)$$

where $0 < q < \infty$. Of course we need $q \neq 1$ to avoid dividing by zero, but L'Hôpital's rule shows that the Rényi entropy approaches the Shannon entropy as q approaches 1:

$$\lim_{q \rightarrow 1} S_q = - \sum_i p_i \ln p_i.$$

So, it is customary to define S_1 to be the Shannon entropy.

While Shannon entropy has a deep relation to thermodynamics, the Rényi entropy has not been completely integrated into this subject—at least, not in a well-recognized way. While many researchers have tried to *modify* statistical mechanics by changing the usual formula for entropy, so far the most convincing uses of Rényi entropy in physics seem to involve the limiting cases $S_0 = \lim_{q \rightarrow 0} S_q$ and $S_\infty = \lim_{q \rightarrow +\infty} S_q$. These are known as the ‘max-entropy’ and ‘min-entropy’, respectively, since S_q is a decreasing function of q . They show up in studies on the work value of information [5] and the thermodynamic meaning of negative entropy [1]. For other interpretations of Rényi entropy see Harremöes [8], König *et al.* [9], and Uffink [13].

In fact, it is not necessary to modify statistical mechanics to find a natural role for Rényi entropy in physics. Rényi entropy is closely related to the familiar concept of *free energy*, with the parameter q appearing as a ratio of temperatures.

The trick is to think of our probability distribution as a Gibbs state: that is, the state of thermal equilibrium for some Hamiltonian at some chosen temperature, say T_0 . Suppose that all the probabilities p_i are nonzero. Then working in units where Boltzmann's constant equals 1, we can write

$$p_i = e^{-E_i/T_0}$$

for some nonnegative real numbers E_i . If we think of these numbers as the energies of microstates of some physical system, the Gibbs state of this system at temperature T is the probability distribution

$$\frac{e^{-E_i/T}}{Z}$$

where Z is the partition function:

$$Z = \sum_{i \in X} e^{-E_i/T}$$

Since $Z = 1$ when $T = T_0$, the Gibbs state reduces to our original probability distribution p at this temperature.

Starting from these assumptions, the free energy

$$F = -T \ln Z$$

is related to the Rényi entropy as follows:

$$F = -(T - T_0)S_{T_0/T} \tag{2}$$

The proof is an easy calculation:

$$S_{T_0/T} = \frac{1}{1 - T_0/T} \ln \sum_i p_i^{T_0/T} = \frac{T}{T - T_0} \ln \sum_i e^{-E_i/T} = -\frac{F}{T - T_0}.$$

This works for $T \neq T_0$, but we can use L'Hôpital's rule to show that in the limit $T \rightarrow T_0$ both sides converge to the Shannon entropy S_1 of the original probability distribution p .

After the author noticed this result in the special case $T_0 = 1$ [2], Blake Stacey commented that this case was already mentioned in Beck and Schlogl's 1995 text on the thermodynamics of chaotic systems [3]. However, most people using Rényi entropy seem unaware of its connection to free energy, perhaps because they work on statistical inference rather than physics [7]. The connection deserves to be better known.

Shortly after the first draft of this paper was released, Matteo Polettini pointed out the generalization to other values of T_0 [10]. He also gave a nice physical interpretation of Rényi entropy [11]. Eric Downes then made a further generalization [6]. The above argument concerns a system with Gibbs state $p_i = \exp(-E_i/T_0)$ at a chosen temperature T_0 . Such a system automatically has zero free energy at this chosen temperature. Downes generalized the relation between Rényi entropy and free energy to systems whose free energy is not constrained this way. Polettini's physical interpretation of Rényi entropy can easily be extended to these more general systems, and we describe this interpretation in what follows. We also note that Rényi entropy is a ' q -deformation' of the ordinary notion of entropy.

In what follows, we work in a quantum rather than classical context, using a density matrix instead of a probability distribution. However, we can diagonalize any density matrix, and then its diagonal entries define a probability distribution. Thus, all our results apply to classical as well as quantum systems. The quantum generalization of Shannon entropy is, of course, well-known: it is often called 'von Neumann entropy'. The quantum generalization of Rényi entropy is also already known [14].

2 Rényi Entropy as a q -Derivative of Free Energy

Let H be a self-adjoint complex matrix. Thinking of H as the Hamiltonian of a quantum system, and working in units where Boltzmann's constant is 1, we may define the Gibbs state of this system at temperature T to be the density matrix

$$\rho_T = \frac{1}{Z(T)} e^{-H/T} \quad (3)$$

where the partition function

$$Z(T) = \text{tr}(e^{-E/T}) \quad (4)$$

ensures that $\text{tr}(\rho_T) = 1$. The free energy at temperature T is defined by

$$F(T) = -T \ln Z(T). \quad (5)$$

On the other hand, for any density matrix ρ , the quantum generalization of Rényi entropy is defined by

$$S_q(\rho) = \frac{\ln \text{tr}(\rho^q)}{1 - q} \quad (6)$$

since this formula reduces to the usual definition of Rényi entropy, Equation (1), when the probabilities p_i are the eigenvalues of ρ . This formula makes sense when $0 < q < \infty$ and $q \neq 1$, but we can define the quantum Rényi entropy as a limit in the special cases $q = 0, 1, +\infty$. For $q = 1$ this gives the usual von Neumann entropy:

$$S_1(\rho) := \lim_{q \rightarrow 1} S_q(\rho) = -\text{tr}(\rho \ln \rho). \quad (7)$$

Returning to our system with Gibbs state ρ_T at temperature T , let us write $S_q(T)$ for $S_q(\rho_T)$. Computing this Rényi entropy at some temperature T_0 , we find:

$$S_q(T_0) = \frac{1}{1 - q} \ln \text{tr}(\rho_{T_0}^q) = \frac{1}{1 - q} \ln \text{tr} \left(\frac{e^{-qH/T_0}}{Z(T_0)^q} \right).$$

If we define a new temperature T with

$$q = T_0/T, \quad (8)$$

we obtain:

$$\begin{aligned} S_q(T_0) &= \frac{\ln Z(T) - q \ln Z(T_0)}{1 - q} \\ &= \frac{T \ln Z(T) - T_0 \ln Z(T_0)}{T - T_0} \end{aligned}$$

or in short:

$$S_{T_0/T}(T_0) = - \frac{F(T) - F(T_0)}{T - T_0}. \quad (9)$$

This equation, noted by Downes [6] upon reading the first draft of this paper, is the clearest way to express the relation between Rényi entropy and free energy. In the special case where the free energy vanishes at temperature T_0 , it reduces to Equation (2). In the limit $T \rightarrow T_0$, it reduces to

$$S_1(T_0) = - \left. \frac{dF(T)}{dT} \right|_{T=T_0}. \quad (10)$$

Of course, it is already well-known that the von Neumann entropy is the derivative of $-F$ with respect to temperature. What we see now is that the Rényi entropy is the difference quotient approximating this derivative. Instead of the slope of the tangent line, it is the slope of the secant line.

In fact, we can say a bit more: the Rényi entropy is the the ‘ q^{-1} -derivative’ of the negative free energy. For $q \neq 1$, the q -derivative of a function f is defined by

$$\left(\frac{df}{dx}\right)_q = \frac{f(qx) - f(x)}{qx - x}.$$

This reduces to the ordinary derivative in the limit $q \rightarrow 1$. The q^{-1} -derivative is defined the same way but with q^{-1} replacing q . Equation (9) can be rewritten more tersely using this concept as follows:

$$S_q = - \left(\frac{dF}{dT}\right)_{q^{-1}} \quad (11)$$

Here we have made a change of variables, writing T for the variable called T_0 in Equation (9).

The concept of q -derivative shows up in mathematics whenever we ‘ q -deform’ familiar structures, obtaining new ones such as quantum groups. For an introduction, see the text by Cheung and Kac [4]. In some cases q -deformation should be thought of as quantization, with q playing the role of $\exp(\hbar)$. That is definitely not the case here: the parameter q in our formulas is unrelated to Planck’s constant \hbar . Indeed, Equation (11) holds in classical as well as quantum mechanics.

What, then, is the thermodynamic meaning of Rényi entropy? This was nicely explained by Poletini [11]. Start with a physical system in thermal equilibrium at some temperature. Then ‘quench’ it, suddenly dividing the temperature by q . The maximum amount of work the system can do as it moves to thermal equilibrium at the new temperature, divided by the change in temperature, equals the system’s Rényi entropy of order q in its original state. Note that this formulation even accounts for the minus sign in Equation (9), because it speaks of the work the system does, rather than the work done to it.

Acknowledgements

I thank all the members of the Entropy Club at the Centre for Quantum Technologies, and especially Oscar Dahlsten, for exciting discussions in which these ideas emerged. I thank David Corfield and Tom Leinster for useful discussion of Rényi entropy, and Blake Stacey for pointing out references relating it to free energy. I especially thank Matteo Poletini and Eric Downes for suggestions that vastly improved this paper.

References

- [1] J. Aberg, O. Dahlsten, L. del Rio, R. Renner, and V. Vedral, The thermodynamic meaning of negative entropy. Available as [arXiv:1009.1630](https://arxiv.org/abs/1009.1630).
- [2] J. Baez, Rényi entropy and free energy, *Azimuth*, February 10, 2011. Available at <http://johncarlosbaez.wordpress.com/2011/02/10/rnyi-entropy-and-free-energy/>.
- [3] C. Beck and F. Schlögl, *Thermodynamics of Chaotic Systems*, Cambridge U. Press, Cambridge, 1995.
- [4] P. Cheung and V. Kac, *Quantum Calculus*, Springer, Berlin, 2002.
- [5] O. Dahlsten, R. Renner, E. Rieper and V. Vedral, On the work value of information. Available as [arXiv:0908.0424](https://arxiv.org/abs/0908.0424).

- [6] E. Downes, comment on Rényi entropy and free energy, *Azimuth*, February 11, 2011. Available at <http://johncarlosbaez.wordpress.com/2011/02/10/rnyi-entropy-and-free-energy/#comment-4065>.
- [7] D. Erdogmuns and D. Xu, Rényi's entropy, divergence and their nonparametric estimators, in *Information Theoretic Learning: Rényi's Entropy and Kernel Perspectives*, ed. J. Principe, Springer, 2010, pp. 47–102.
- [8] Peter Harremões, Interpretations of Renyi entropies and divergences. Available as [arXiv:math-ph/0510002](https://arxiv.org/abs/math-ph/0510002).
- [9] R. König, R. Renner and C. Schaffner, The operational meaning of min- and max-entropy, *IEEE Trans. Info. Theory* **55** (2009), 4337–4347. Also available as [arXiv:0807.1338](https://arxiv.org/abs/0807.1338).
- [10] M. Polettini, comment on Rényi entropy and free energy, *Azimuth*, February 10, 2011. Available at <http://johncarlosbaez.wordpress.com/2011/02/10/rnyi-entropy-and-free-energy/#comment-4023>.
- [11] M. Polettini, Rényi entropy and free energy, *Matteo*, February 10, 2011. Available at <http://tomate.blogsome.com/2011/02/10/renyi-entropy-and-free-energy/>.
- [12] A. Rényi, On measures of information and entropy, in *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability 1960*, pp. 547–561. Also available at http://digitalassets.lib.berkeley.edu/math/ucb/text/math_s4_v1_article-27.pdf.
- [13] J. Uffink, Can the maximum entropy principle be explained as a consistency requirement?, Sec. 6: Justification by consistency: Shore and Johnson, *Stud. Hist. Phil. Sci.* **B26** (1995), 223–261. Also available at <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.27.6392>.
- [14] W. van Dam and P. Hayden, Rényi-entropic bounds on quantum communication, Sec. 4.1: Rényi entropy. Available as [arXiv:quant-ph/0204093](https://arxiv.org/abs/quant-ph/0204093).