Generalized entropies: what are they good for?

Jan Korbel
Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Břehová 7, 11519, Prague, Czech Republic

A physicist, mathematician and programmer talk about entropy. Somebody asks:

- what is entropy?
  - measure of uncertainty, information, etc.
- how do you define entropy?
  - physicist: $S = \int \frac{dQ}{T}$
  - mathematician: $S = - \sum_i p_i \log_2 p_i$
  - programmer: minimal length of message encoded in a binary code

different definitions and still one quantity
Introduction

- **entropy** is the central concept of thermodynamics, statistical mechanics and information theory
- there are close parallels between thermodynamical entropy and statistical entropy
- nevertheless, there are some differences and it is necessary to distinguish these concepts
- the aim of this talk is comparison of these concepts
- we discuss situations when a generalized form of entropy can be more convenient in description of a system
**Historical Overview**

- **Thermodynamics**
  - 1803: Carnot - first formulation of 2nd law of TD
  - 1854: Clausius - equivalence-value (first formulation of TD entropy)
  - 1865: Clausius - first definition of entropy (he used term *entropy* for the first time)

- **Statistical mechanics**
  - 1877: Boltzman - $S = k_B \ln W$ (microcanonical)
  - 1880’s: Gibbs - $S = -\sum_i p_i \ln p_i$

- **Information theory**
  - 1948: Shannon - mathematical theory of communication - Shannon entropy $S(P) = -\sum_i p_i \log_2 p_i$
  - 1957: Jaynes - MaxEnt principle (connection between information theory and statistical mechanics)
  - 1961: Rényi entropy
  - 1973: Havrda Charvát entropy, 1988 - Tsallis entropy
Thermodynamical entropy

- classical thermodynamics is based on laws of thermodynamics
- the first law of thermodynamics
  \[ dU = \partial Q - \partial W \]
- differential form of heat is not closed, but there exists an integration factor, so
  \[ \partial Q = T dS \]
- thermodynamical entropy is therefore defined as from the previous relation
- as a consequence, entropy is a state variable
STATISTICAL MECHANICS ENTROPY

- Classical thermodynamics works with macroscopic quantities
- Statistical mechanics works with microscopic quantities from which can be the macroscopic quantities derived
- The whole description is based on ensemble description and partition function
- Entropy can be determined as

\[
S = k_B \ln \Omega_{mic} = k_B \ln Z_{can} + \beta < E > = k_B \ln Z_{gc} + \beta ( < E > - \mu < N > )
\]
Informational entropy

- Information theory defines the entropy as a measure of information encoded in an arbitrary probability distribution.
- Entropy represents the amount of information necessary to fully determine a system.
- Exact form of entropy depends on the particular system.
- The most common is the Shannon entropy (SE):

  \[ S(P) = \sum_i p_i \log_2 \left( \frac{1}{p_i} \right) \]
IMPORTANT PRINCIPLES OF ENTROPY

To entropy are connected several important principles, among others:

▶ MaxEnt principle
▶ Legendre structure of TD
▶ Additivity
▶ Extensivity

we briefly discuss the importance of each principle
MaxEnt principle

- Connection between thermodynamical entropy and informational entropy is given by Jaynes' **MaxEnt principle**
- Thermodynamical entropy is obtained from the distribution which maximizes informational entropy under given constrains
- We define Lagrange function

\[
L(P) = S(P) - \sum \lambda_j f_j(P)
\]  

(1)

where \( S \) is the entropy and \( f_j \) are the constraints
- Usually, we consider average energy constraint

\[
\langle E \rangle = \sum_i p_i E_i \quad \text{(canonical ensemble)}
\]

- Corresponding Lagrange parameter represents the inverse temperature \( \beta \)
- Resulting distribution maximizing entropy is the one which is realized in a thermodynamical system
one important consequence of MaxEnt principle is the Legendre structure of thermodynamics

for canonical ensemble represents Lagrange function the free energy of the system

therefore we get the well-known relation

\[ F(T, V) = U(T, V) - TS(T, V) \]

not only for Shannon entropy, but also for other entropy functionals
**Additivity**

- Additivity and extensivity are important properties of statistical systems.
- Nevertheless, they have slightly different meaning.
- Additivity: informational property - entropy of a compound system can be expressed as:

$$S(A \cup B) = S(A) + S(B|A)$$  \hspace{1cm} (3)

- It is possible to introduce a concept of generalized additivity so for independent $A, B$

$$S(A \cup B) = \Phi(S(A), S(B))$$  \hspace{1cm} (4)
Extensivity

- Extensivity is, on the other hand, a thermodynamical property depending on a particular system.
- Entropy is extensive if for state space $W(N)$ is

$$\lim_{N\to\infty} \frac{S(W(N))}{N} = \omega$$

the entropy scales as the number of particles so

$$S(2N) \sim 2S(N)$$

- Shannon entropy is extensive for $W(N) \propto 2^N$ (no correlations)
- Other entropies are extensive systems growing polynomially $N^\alpha$ (Tsallis) or subexponentially $2^{N^\gamma}$
Axiomatic definition of entropy

- Entropy can be straightforwardly defined as a functional on a probability space
- Alternatively, more rigorous definition was done by A. Kolmogorov
- The axiomatic definition of the entropy enables us to understand main properties of entropy
- It is also a good starting point for possible generalizations
Axiomatic definition of Shannon entropy

Shannon entropy is defined by the following four axioms:

1. continuity: $\mathcal{H}(\mathcal{P})$ is continuous in every argument,
2. maximality: $\mathcal{H}(\mathcal{P})$ takes the maximal value for the uniform distribution,
3. expansibility: $\mathcal{H}(p_1, \ldots, p_n, 0) = \mathcal{H}(p_1, \ldots, p_n)$.
4. additivity: $\mathcal{H}(A \cup B) = \mathcal{H}(A) + \mathcal{H}(B \mid A)$, where $\mathcal{H}(B \mid A) = \sum_k p_k \mathcal{H}(B \mid A = a_k)$.

From these four axioms can be determined the form of Shannon entropy

$$\mathcal{H}(\mathcal{P}) = - \sum_k p_k \ln p_k.$$
Generalized entropies

- functional form of Shannon entropy corresponds to Boltzman-Gibbs entropy from statistical mechanics
- information theory provides many more information measures
- one way how to define these entropies is to generalize Shannon axiomatic
- these entropies could be understand as statistical mechanical entropies resulting into a different thermodynamics
- question at stake is if this approach is legitimate and if it brings any advantages
- “everybody can have his/her own entropy”
Rényi entropy
Entropy of multifractal systems

- Shannon entropy is not the only solution for additivity axiom $S(A \cup B) = S(A) + S(B|A)$
- The defining axioms remain the same but the fourth is slightly changed:
  4. Rényi additivity: $\mathcal{I}_q(A \cup B) = \mathcal{I}_q(A) + \mathcal{I}_q(B|A)$, where $\mathcal{I}_q(B|A) = g^{-1} \left[ \sum_k \rho_k(q) g[\mathcal{I}_q(B|A = a_k)] \right]$, $\rho_k(q) = p_k^q / \sum_j p_j^q$, and $g$ is positive and invertible function on $[0, \infty)$.
- Similarly to Shannon entropy, it is possible to find an operational definition of RE (minimal cost)

This leads to the definition of RE in the form:

$$\mathcal{I}_q(\mathcal{P}) = \frac{1}{1-q} \ln \left( \sum_k p_k^q \right)$$
Rényi entropy

\[ I(p) \]

\[ q = 0.3 \quad q = 0.5 \quad q = 1 \quad q = 2 \quad q = 4 \]
APPLICATION OF RE
MULTIFRACTAL THERMODYNAMICS

- Rényi entropy has tight relation to theory of **multifractals**
- fractal systems: characteristic scaling relation and fractal dimension (Koch snowflake, etc.)
- multifractal systems: multiple local scaling exponents, probability distribution $p_i(s) \sim s^{\alpha_i}$ (s - scale)
- distribution of scaling exponents $p(\alpha, s) \sim s^{-f(\alpha)} \text{d}\alpha - f(\alpha)$: multifractal spectrum, fractal dimension of $\alpha$-subset
- partition function $Z(q,s) = \sum_i p_i(s)^q \propto s^{\tau(q)} \sim s^{\alpha q - f(\alpha)}$
- the whole system can be described alternatively by multifractal spectrum $f(\alpha)$ or scaling function $\tau(q)$
- Rényi entropy $I_q = \frac{1}{1-q} \ln \sum_i p_i^q \sim \frac{\tau(q)}{q-1} = D(q)$
The connection to thermodynamics can be done via the partition function $\sum_i p_i^q = \sum_i \exp(-\beta E_i)$.

As a consequence $E_i = -\ln p_i$ and $\beta = q$.

Temperature is equal to parameter $q$ which is called zooming parameter because for $p_i^q$ accentuates different parts of PDF.

Relation between multifractal entropy $I_q = \frac{1}{1-q} \ln Z(q, s)$ and thermodynamical entropy $S_q = (\ln Z(q, s) + \beta < E >)$ is more complicated.
Tsallis entropy generalizes the concept of entropy to non-extensive systems. It was firstly discovered in information theory by Czech mathematicians J. Havrda and F. Charvát. Later, it was introduced to physics by C. Tsallis. The additivity axiom is now changed as follows:

4. Tsallis additivity:

\[ S_q(A \cup B) = S_q(A) + S_q(B|A) + (1 - q)S_q(A)S_q(B|A), \]

where \( S_q(B|A) = \sum_k \rho_k(q)S_q(B|A = a_k) \) and \( \rho_k(q) = \frac{p_k^q}{\sum_j p_j^q} \).

We can deduce that Tsallis entropy has form

\[ S_q(\mathcal{P}) = \frac{1}{1-q} \left( \sum_k p_k^q - 1 \right) \]
Tsallis Entropy

Tsallis entropy

\[ S(p) \]

\[ q=0.3 \quad q=0.5 \quad q=1 \quad q=2 \quad q=4 \]
APPLICATION OF TSALLIS ENTROPY

THERMODYNAMICS WITH FINITE BATH

- Let us consider a system $S$ coupled with a heat bath $B$.
- In microcanonical ensemble is the partition function given by number of states $\Omega_{tot}(E) = \sum_s \Omega_S(E_s)\Omega_B(E - E_s)$.
- For finite bath is $\Omega_B(E - E_s) \propto (E - E_s)^{n_s - 1}$.
- Probability distribution is therefore $p(E_s) = \frac{\Omega_B(E - E_s)}{\sum_k \Omega_B(E - E_k)} = \frac{1}{Z_q} [1 - \beta (q - 1) E_s]^{1/(q - 1)}$.
- Equivalently, Tsallis $q$-gaussian distribution can be derived as a MaxEnt distribution of Tsallis entropy under constraints $\sum_i p_i = 1 \sum_j p_j E_j = <E>$.
- There are problems with temperature definition (self-referential temperature - depending on distribution).
Hybrid Entropy

Overlap between Self–similarity and Non–extensivity

- recently, there have been many attempts on further generalizations - one of them is called hybrid entropy
- hybrid entropy is based on both properties of Rényi entropy and Tsallis entropy
- resulting entropy can describe self–similar and non–extensive systems
- additivity axiom is now changed as

4. J.-A. additivity:

\[ D_q(A \cup B) = D_q(A) + D_q(B|A) + (1 - q)D_q(A)D_q(B|A), \]

where \[ D_q(B|A) = f^{-1} \left( \sum_k \rho_k(q)f[D_q(B|A = a_k)] \right), \]

\[ \rho_k(q) = \frac{p_q^k}{\sum_j p_j^q} \] is escort distribution and \( f \) is a positive and invertible function on \([0, \infty)\).

These axioms define the hybrid entropy in form:

\[ D_q(\mathcal{P}) = \frac{1}{1 - q} \left( e^{-(1-q)\sum_k \rho_k(q)\ln p_k} - 1 \right) = \frac{1}{1 - q} \left( e^{-(1-q)\langle\ln \mathcal{P}\rangle_q} - 1 \right) \]
Hybrid Entropy

MaxEnt distribution

- we demand normalization condition $\sum_k p_k = 1$ and given expectation value of energy $\langle E \rangle_r = \sum_k \rho_k(r) E_k$.
- two most common choices are linear averaging ($r = 1$, $\langle E \rangle = \sum_k p_k E_k$) and $q$-averaging ($r = q$)
- for both cases we find the MaxEnt distributions and discuss their properties
- the maximization under constraints is done via the Lagrange multiplier method:

$$\mathcal{L}_{q,r}(\mathcal{P}) = \mathcal{D}_q(\mathcal{P}) - \Omega \left( \sum_k \rho_k(r) E_k - \langle E \rangle_r \right) - \Phi \left( \sum_k p_k - 1 \right)$$
**Hybrid Entropy**

**MaxEnt for** $r = q$

- condition $\frac{\partial L_{q,q}(\mathcal{P})}{\partial p_i} = 0$ leads to equation

$$\kappa p_i^{1-q} = q \ln p_i + \mathcal{E}_i$$

where $\kappa = \sum_j p_j^q$ and $\mathcal{E}_i = 1 + \frac{q \ln(-\Phi)}{1-q} + \frac{q \Omega}{-\Phi}(E_i - \langle E \rangle_q)$

- previous equation can be solved in terms of Lambert $W$-function defined as $W(x)e^{W(x)} = x$

- we obtain $p_i = \left[\frac{q}{\kappa(q-1)} W\left(\frac{\kappa(q-1)}{q} e^{(q-1)\mathcal{E}_i/q}\right)\right]^{1/(1-q)}$

- Lambert $W$-function is defined only on interval $[-1/e, \infty)$, so there are energy regions with zero probability - energy gaps
Hybrid Entropy

MaxEnt for $r = q$

- Similarly to previous case, from condition $\frac{\partial L_{q,1}(P)}{\partial p_i} = 0$ we obtain

$$\kappa p_i^{1-q} = \frac{\Phi}{\Phi + \Omega(E_i - \langle E \rangle)} \left[ q \ln p_i - q \ln(-\Phi) \frac{q}{q-1} + 1 \right]$$

- The solution of previous equation is

$$p_i = \left[ \frac{q\Phi}{(q-1)\kappa(\Phi + \Omega\Delta E_i)} W \left( -\frac{\kappa(q-1)}{\Phi q} \exp \left( q - 1 \right) \left( 1 + \frac{\Omega}{\Phi} \Delta E_i \right) \right) \right]^{1/(1-q)}$$

- Also in this case we observe energy gaps and forbidden energies
In the derivation of the hybrid entropy was not used the maximality axiom, it is necessary to check its validity.

We focus on the term $\langle \ln P \rangle_q = \sum_k \rho_k(q) \ln p_k$

For $q < 1/2$, $\langle \ln P \rangle_q$ does not become maximal value for equal probabilities.

Thus, the hybrid entropy is properly defined only for $q \geq 1/2$.

Hybrid entropy is for concave for $q \geq 1$ and Schur-concave for $q \geq 1/2$ (weaker version of concavity).
HYBRID ENTROPY

Hybrid entropy

\[ D(p) \]

\[ q = 0.3 \quad q = 0.5 \quad q = 1 \quad q = 2 \quad q = 4 \]
The main question is if generalized entropies are a fundamental concept, similarly to thermodynamical entropy. One could think about systems with $q$-gaussian distributions. MaxEnt distribution does not determine entropy: $q$-gaussian distribution can be obtained from Tsallis entropy under constraints:

$$\sum_i p_i = 1, \sum_i p_i E_i = \langle E \rangle$$

Shannon entropy under constraints:

$$\sum_i p_i = 1, \sum_i p_i^q = N_q, \sum_i E_i \rho_i(q) = \langle E \rangle_q$$

Same distributions - different entropies (correlated walk vs accelerated walk).

Problems with operational definitions (Tsallis entropy cannot be measured) and self-referential entropy.
**Conclusions**

- Entropy is the concept used in many different scientific fields.
- The talk compared its definition and connection to thermodynamics, statistical mechanics, and information theory.
- Several generalized entropies were presented: Rényi entropy, Tsallis entropy.
- We have combined axiomatics of Rényi entropy and Tsallis entropy and obtained a new class of hybrid entropies.
- Generalized entropies can be useful tools for description of various systems.
- Usually, they are not fundamental concepts (contrary to thermodynamical entropy).
REFERENCES


Thank you for your attention.