HAL open science

Entropy in Thermodynamics: from Foliation to Categorization

Radoslaw A. Kycia

▶ To cite this version:

Radoslaw A. Kycia. Entropy in Thermodynamics: from Foliation to Categorization. Communications in Mathematics, 2021, Volume 29 (2021), Issue 1 (Special Issue: Ostrava Mathematical Seminar) (1), pp.49 - 66. 10.2478/cm-2021-0002 . hal-03665006

HAL Id: hal-03665006

https://hal.science/hal-03665006v1

Submitted on 11 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

sciendo

Communications in Mathematics 29 (2021) 49–66
DOI: 10.2478/cm-2021-0002

©2021 Radosław A. Kycia
This is an open access article licensed
under the CC BY-NC-ND 3.0

Entropy in Thermodynamics: from Foliation to Categorization

Radosław A. Kycia

Abstract. We overview the notion of entropy in thermodynamics. We start from the smooth case using differential forms on the manifold, which is the natural language for thermodynamics. Then the axiomatic definition of entropy as ordering on a set that is induced by adiabatic processes will be outlined. Finally, the viewpoint of category theory is provided, which reinterprets the ordering structure as a category of pre-ordered sets.

For Professor Olga Rossi, in memoriam.

1 Introduction

The notion of entropy ('tropos' is Greek word transformation) initially appeared in Thermodynamics to describe the possible direction of the process. At the time the theory was being developed, conception regarding the inner structure of matter, like atoms, was not available, and hence matter was described in terms of macroscopic averaged variables as pressure, volume, temperature, etc. Currently, we know that these variables come from the reduction of a large number of degrees of freedom of particles in a piece of matter (the Avogadro constant $N_A \sim 10^{23}$ atoms which in classical description have 3 numbers describing position coordinates, and 3 numbers describing velocity coordinates) to a few variables mentioned above. The need for pointing out this 'coarse/average' evolution direction was imminent, and to fulfill this need entropy was invented, see Fig. 1. Does it mean that if an entropy in theory appears, then we are dealing with a 'coarse' (not fundamental) description? We do not know yet.

2020 MSC: 80A05, 80-10
Key words: Entropy; Thermodynamics; Contact structure; Ordering; Posets; Galois connection
Affiliation:
Masaryk University, Department of Mathematics and Statistics, Kotlářská 267/2, 611 37 Brno, The Czech Republic & Cracow University of Technology, Faculty of Materials Engineering and Physics, Warszawska 24, Kraków, 31-155, Poland
E-mail: kycia.radoslaw@gmail.com

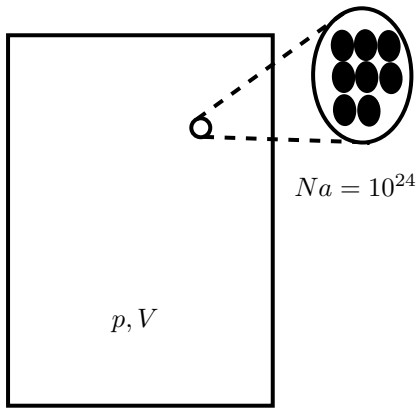

Figure 1: System of large degrees of freedom (atoms) in thermodynamics is reduced to a few variables.

The history of thermodynamics is full of amazing reasonings that finally lead to the correct laws of nature. To mention one, Robert Mayer concluded that heat is energy flow by observing the color of the blood of sailors under different geographic latitudes on the ship on which he was a medical doctor.

Since the pioneering work of Boltzmann that connects thermodynamical entropy with microscopic properties of matter and properties of logarithm function, this notion appeared to be useful in various disciplines like information theory [22] or dynamical systems [9]. The increasing importance of this concept is reflected in bibliometric data of research papers on this subject [24].

This summary will neglect all the classical/physical motivation for thermodynamics, and we go directly to mathematical concepts associated with the notion of entropy. We believe that thanks to such an approach, we avoid mixing assumptions, the result of reasoning, and 'common knowledge' in this theory, which is common in physics literature and which leads to the difficulty in grasping these concepts. Thermodynamics is mature enough to axiomatize fully, and this can be done as will be presented below. This presentation is by no means original research – it is only an overview of the subject and a small part of existing literature. Only the organization of the material is perhaps nonstandard and selective. However, it is hoped it can be treated as a guide for novices (both with mathematical or physical background) to avoid pitfalls common in this subject.

The overview is organized as follows: First, the geometric meaning of entropy close to the original formulation in modern differential geometric terms will be provided. The presentation will be provided with the context, i.e., geometric structure of (phenomenological/equilibrium) thermodynamics. Then the axiomatic approach to entropy will be outlined. Finally, the categorical approach to this subject will be presented. In the Appendix, mathematical preliminaries were collected for the reader's convenience, and we advise the reader to look up the Appendix to be oriented what kind of mathematics is needed to understand the main parts of the paper.

The Ostrava Seminar on Mathematical Physics organized for many years by Diana Barseghyan, Olga Rossi, and Pasha Zusmanovich is a unique platform to exchange knowledge between mathematicians and physicists. There is also a big audience of students that motivates speakers to present material more pedagogically, including also the context of the research. One of my talks, which was a pleasure to deliver, was about entropy and Landauer's principle. This overview paper can be treated as a basic introduction and guide to the subject.

2 Smooth case

We start by describing thermodynamics in the natural setup of smooth manifolds with contact structure. In these terms, although not so precisely due to less developed mathematical language, the fathers of thermodynamics were thinking. The presentation follows closely [2], [3], [4], [7], [8], [18].

2.1 Space

In thermodynamics, we identify some system from the environment by distinguishing some more or less formal boundaries with some specific physical properties (e.g., heat contact or permeability of particles). Such a system should be macroscopically uniform in the sense of its physical and chemical properties - a so-called *simple system*. By distinguishing such a system, we can describe it by some variables depending on the physical context. The common feature of these variables is their uniform behavior under the scaling group $\mathbb{R}^+$, which reflects the physical property that the scaling of the system scales its internal parameters, e.g., scaling the system scales its volume or energy. These variables are called *extensive*. Call these variables $\{X_i\}_{i=0}^n$. If there is another extensive variable X, then it must depend on the previous $X = X(X_0, \ldots, X_n)$ and scales as $X(\lambda X_0, \ldots, \lambda X_n) = \lambda X(X_0, \ldots, X_n)$ where $\lambda \in \mathbb{R}^+$ that it should be truly extensive.

The common choice of extensive variables and the usual symbols (instead of X_i's) designated for them are as follows:

- U – energy of the system;
- V – volume;
- N – number of particles;

The first assumption is that

Assumption 1. *In an equilibrium state, the system is fully described by some set of extensive variables.*

An equilibrium state is attained when the system is left on its own and relaxes attaining this state without any further change of extensive variables. The direction in which a system left on its own evolves is described by entropy, which will be introduced at a later stage.

If a system is composed of more simple systems, then the number of variables multiplies accordingly.

The system (simple or compound) can interact with the environment by exchanging energy. One 'directed' way of transferring energy is work made by the

environment on the system. Various types of work are described by work 1-forms that relate change of extensive parameters of the system with work done on the system. In local coordinates:

$$W = P_i \mathrm{d}X^i \,.$$

The coefficients P_i are called *intensive* variables and describe the 'generalized forces' of environment that act on the system. They do not scale. Note that W does not to be an exact form, and therefore work may depend on the path along which it is integrated.

Common choice of intensive variables are

- p – pressure; associated with change of volume V;
- $-\mu$ – chemical potential describing density of work done by changing the number of particles in the system by adding/removing particles/elements from/to environment or modified by chemical reactions; associated with the number of particles N;

The second way of energy transfer between system and environment is the heat transfer described by a 1-form Q. We will see below that Q can be written in terms of work form: $Q = T\mathrm{d}S$, where T is the absolute temperature (intensive variable), and S is the entropy (extensive variable). In physical terms, Q is the transfer of kinetic energy at the level of atoms/molecules. The detailed description of this transfer in thermodynamics is neglected by reducing microscopic degrees of freedom to a few macroscopic ones. Therefore some additional law has to be introduced that controls such transfer. This is done in terms of entropy and the Second Law of Thermodynamics.

In thermodynamics, the system is described by energy U and $2n$ pairs of associated intensive-extensive variables. These are local coordinates on $2n+1$ dimensional manifold. From physics, it is assumed

Assumption 2. *The equilibrium state is described as a point in $2n+1$ dimensional smooth manifold M called the space of states.*

Local coordinates are usually taken to be $(U, (T, S), (p, V), (\mu, N), \ldots)$, where intensive-extensive pairs were grouped.

In order to compare systems in equilibrium, we introduce the Zeroth Law of Thermodynamics

Axiom 1. *If a system A is with a thermal equilibrium with B and B with C, then A is in thermal equilibrium with C.*

The thermal equilibrium of two systems means that there is no heat flow Q between systems that are connected by thermally conducting material. The Law means that the relation of 'being in thermal equilibrium' is transitive. It is also obviously reflexive and symmetric, and therefore is an equivalence relation. It allows us to define tools/systems called thermometers that measure *empirical temperature*, which represent precisely these equivalence classes. This empirical temperature will be related to the (absolute) temperature T below.

2.2 Processes

The next step is to consider the change in the system that is described by paths in the space of states called thermodynamical processes.

(Equilibrium) thermodynamics is only occupied with *quasi-static processes*, which can be represented by curves in the space of states. In physical terms, they can be considered as physical/chemical processes that occur 'slow enough' that in every step of the process the system and environment are in equilibrium or relax 'fast enough' to equilibrium. It is only an idealization. On the other hand, *non-quasi-static processes* cannot be described as a path in the space of states. They can only be marked as initial and final points if these points are equilibrium states. This peculiarity is connected with the fact that the points in the space of states describe only equilibrium states.

The other distinction is according to reversibility. The process is:

- *reversible* – if it can be conducted in both directions when all variables (intensive of system and extensive of environment) can be returned to initial values in local description;
- *irreversible* – it cannot be reversed;

For quasi-static processes we have a curve γ in the space of states M that we assume to be piecewise smooth which is usual assumption. We can then calculate:

- $\Delta Q(\gamma) := \int_\gamma Q$ – total heat transfer in the process;
- $\Delta W(\gamma) := \int_\gamma W$ – total work done in the process;

Note that these definitions are not valid when there is no curve along which 1-forms Q and W can be integrated, i.e. for non-quasi-static processes.

Some examples of thermodynamic processes are as follows [7]:

- *Quasi-static adiabatic process* – in this case no heat is exchanged, that is $\Delta Q(\gamma) = 0$;
- *Heating at constant volume* – a quasi-static process which for the case of simple system takes place without exchange of particles
$$\Delta W(\gamma) = \int_\gamma p\,\mathrm{d}V = 0\,;$$
- *Non-quasi-static process* – no path in M therefore no ΔW and no ΔW can be calculated. Only the the difference of energy between initial and final state of the process can be defined.

In technical applications, the most important are closed paths that are called *thermodynamical cycles* and describe the cyclic work of engines. They are also crucial in the formulation of the Second Law of Thermodynamics below.

2.3 The First Law of Thermodynamics

The first fundamental law of thermodynamics describes from the physical point of view the conservation of energy during a quasi-static process, namely

Axiom 2 (First Law of Thermodynamics).

$$\mathrm{d}U = Q - W \tag{1}$$

We stated this law as an Axiom since, although on the physics side it is a fundamental law of nature, on the mathematical side it is an unquestionable statement, i.e., an axiom for mathematical formulation of thermodynamics.

For quasi-static processes described by the curve γ in M with the initial point x and the final point y the integrated version of (1) is

$$\Delta U(\gamma) := U(y) - U(x) = \Delta Q(\gamma) - \Delta W(\gamma)\,.$$

This is due to the fact that $\mathrm{d}U$ is exact form and therefore its integral depends only on the endpoints of the curve γ.

In expanded form (1) can be written in local coordinates as

$$\mathrm{d}U = Q - p\mathrm{d}V - \mu\mathrm{d}N\,.$$

In this context we can reinterpret the properties of the processes:

- *Quasi-static adiabatic process* – $\Delta Q(\gamma) = 0$ and therefore $\Delta U = -\Delta W$;
- *Heating at constant volume* – $\Delta W(\gamma) = \int_\gamma p\,\mathrm{d}V = 0$ and therefore $\Delta U = \Delta Q$;

Note that a quasi-static adiabatic process converts all the total energy of the system to the work that can be extracted from or transferred to the system. The restrictions on this process prevents the construction of a 'perpetuum mobile' and is controlled by the Second Law of Thermodynamics described below.

We now turn to finishing the mathematical description of state space. On $2n+1$ dimensional space M we have the form

$$\theta := \mathrm{d}U - Q + W\,. \tag{2}$$

The volume form in M can be given by

$$\theta \wedge (\mathrm{d}\theta)^n \neq 0\,.$$

Therefore θ defines a contact structure on M or equivalently $J^1(N)$, where N (see Appendix) has local coordinates as extensive variables (U, V, N). This leads to the final definition of the space of states for thermodynamics

Definition 1. The space of states in thermodynamics is described by odd dimensional space M with contact form θ that fulfills $\theta \wedge (\mathrm{d}\theta)^n \neq 0$.

We can now reconstruct the conservation law of the First Law of Thermodynamics: Using the Darboux theorem for contact manifolds (see Appendix), there are local coordinates $(X_0, (X_1, P_1), \ldots, (X_n, P_n))$ that the canonical form of θ is

$$\theta = \mathrm{d}X_0 - \sum_{i=1}^{n} P_i \mathrm{d}X^i \,. \tag{3}$$

Comparing with (2) we have that $X_0 = U$ etc.

In this space the submanifold Φ describing the physical system in equilibrium fulfills

$$\Phi^* \theta = 0 \,,$$

that is, physical systems are described by such submanifolds of M that preserve energy/The First Law of Thermodynamics. In the case of a non-degenerate thermodynamical system, it is assumed:

Assumption 3. *The non-degenerate thermodynamical system is described by maximal dimension subspace in the contact space of states of dimension $2n+1$, i.e., Legendre submanifolds of dimension n.*

The Legendre submanifold is defined by providing $X_0 = X_0(X_1, \ldots, X_n)$. Alternatively, using (3), we can provide *equations of state*

$$\begin{cases} P_1 = P_1(X_1, \ldots, X_n) = \frac{\partial X_0}{\partial X_1} \,, \\ \ldots \\ P_n = P_n(X_1, \ldots, X_n) = \frac{\partial X_0}{\partial X_n} \,. \end{cases}$$

This can be viewed as the equivalence of holonomic sections of jet space and Legendre submanifolds on contact space – see Appendix.

The last remaining issue is the direction of heat transfer, which is resolved by the Second Law of Thermodynamics outlined in the next subsection.

2.4 The Second Law of Thermodynamics

We now put some restrictions on quasi-static adiabatic paths/processes γ that are described by $\gamma^* Q = 0$. All tangent vectors to such paths are in $\mathrm{Ker}(Q)$ and define some distribution on M. Since adiabatic processes along arbitrary paths are not present in nature, therefore Caratheodory formulated the following version of the Second Law of Thermodynamics:

Axiom 3 (Second Law of Thermodynamics, Caratheodory). *[7] In a neighbourhood of any state $x \in M$ there is state y that is not accessible from x via quasi-static adiabatic paths γ such that $\gamma^* Q = 0$.*

Using the Caratheodory's theorem on accessibility (see Appendix), we get that the distribution $\mathrm{Ker}(Q)$ is integrable (defines holonomic constraints in M) or, put another way,

$$Q \wedge \mathrm{d}Q = 0 \,.$$

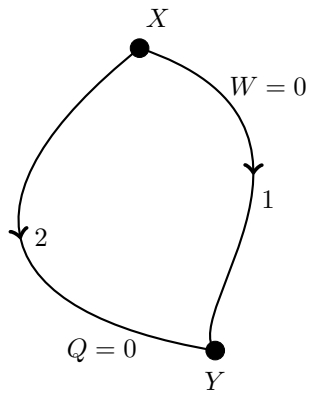

Figure 2: Path 1 is a cooling at constant volume along which no work is made ($W = 0$). The path 2 is a quasi-static adiabatic process ($Q = 0$).

This is also equivalent to the statement that

$$Q = T\mathrm{d}S\,,$$

where T is an integrating factor (a nonsingular function on M) called the *absolute temperature*, and S is called the entropy. It means that $S = \text{const}$ defines a local leaf of the distribution on which quasi-static adiabatic paths lie.

Consider two simple systems with thermal contact (no adiabatic border). It can be shown that these are in equilibrium if their absolute temperatures T are equal [2]. This defines equivalence classes as in the Zeroth Law of Thermodynamics, and therefore absolute temperature can be used as empirical temperature.

There is a stronger version of this law by Kelvin that implies [7] Caratheodory's version, namely,

Axiom 4 (Second Law of Thermodynamics, Kelvin). *In quasi-static cyclic process a quantity of heat cannot be converted entirely into its mechanical equivalent of work.*

This version will be used hereafter.

It can be shown [7] that the foliation exists globally and is not pathological. It relies on the following

Proposition 1. *[7] The state y obtained from x by cooling at constant volume ($W = 0$) cannot be connected again with x by a quasi-static adiabatic process/path.*

Proof. As in [7], consider, on the contrary, the two paths from x to y shown in Fig. 2. The work along 2 is

$$\int_2 W = \int_2 Q - \mathrm{d}U = -\int_2 \mathrm{d}U = -\int_1 \mathrm{d}U = \int_{-1} \mathrm{d}U = \int_{-1} Q\,,$$

as integral of $\mathrm{d}U$ does not depend on the path chosen. Here -1 means the path is followed in the opposite direction than indicated in the figure. Therefore we have that there is a cycle $y \underset{-1}{\longrightarrow} x \underset{2}{\rightarrow} y$ which converts whole heat into work, and this contradicts Axiom 4.

Note also that the above proof is, by contraposition, precisely the statement, that the Second Law of Thermodynamics by Kelvin implies the version of this law by Caratheodory. □

It results from the above Proposition, that the leaf $S = \text{const}$, containing adiabatic processes, is transversal to the paths of the process of cooling at constant volume. This means that the point starting at some leaf of constant entropy has to be taken into another leaf by the cooling at constant volume. Moving along this path we never return to the same leaf. This eliminates pathological situations when, e.g., the leaf winds densely on the manifold, i.e., the cases when leaf makes an initial submanifold [10]. This shows that entropy S, and T are globally defined on M.

The important conclusion that will be a link between classical entropy and its axiomatic definition in the next section is

Theorem 1. *[7] If a state y results from x by any adiabatic process (quasi-static or not), then $S(y) \geq S(x)$.*

We therefore have that in an isolated (i.e., adiabatic) system entropy cannot decrease when achieving equilibrium, which is the commonly known version of the Second Law of Thermodynamics. Here it is presented as a conclusion from a more geometric formulation of this law.

2.5 Symmetries and thermodynamic potentials

Having defined the fundamentals of thermodynamics, we can provide some examples of different choices of variables that do not change thermodynamics. They are useful if we prefer to use different variables to observe the system. Since the coordinate changes should not alter the contact structure, they are contact symmetries mentioned in the Appendix.

The most useful is the Legendre transformation that interchange the role in extensive-intensive pair of variables. This transformation also modifies $X_0 = U$ variable giving a new, so called, *thermodynamic potential*. We present a few examples in case of constant number of particles (system boundaries are not permeable – $\mu = 0$) for simplicity:

- The transformation $p \leftrightarrow V$ gives a thermodynamic potential called the Enthalpy $\tilde{X}_0 = H := U + pV$ and the contact form $\tilde{\theta} = \mathrm{d}H - T\mathrm{d}S - V\mathrm{d}p$. It is useful to observe the system on the submanifold $p = \text{const}$.

- The transformation $T \leftrightarrow S$ gives a thermodynamic potential called the Helmholtz potential/Free energy $\tilde{X}_0 = F := U - TS$ and the contact form $\tilde{\theta} = \mathrm{d}F + S\mathrm{d}T + p\mathrm{d}V$. It is useful to observe the system on the submanifold $V = \text{const}$.

- The transformation $p \leftrightarrow V$ and $T \leftrightarrow S$ gives a potential called the Gibbs potential $\tilde{X}_0 = G := U + pV - TS$ and the contact form $\tilde{\theta} = \mathrm{d}G + S\mathrm{d}T - V\mathrm{d}p$. It is useful to observe the system on the submanifold $T = \text{const}$ and $p = \text{const}$.

2.6 Examples

The thermodynamic relations result from taking the exterior derivative of the contact form pulled-back to the Legendre manifold that describes the thermodynamical system. As an example consider the standard contact form

$$\theta = \mathrm{d}U - T\mathrm{d}S + p\mathrm{d}V\,. \tag{4}$$

The Legendre manifold Φ is given by equations of state $T = T(S,V)$ and $p = p(S,V)$. Then since $\Phi^*\theta = 0$ and $\Phi^*\mathrm{d}\theta = \mathrm{d}\Phi^*\theta = 0$ we get

$$\Phi^*\mathrm{d}\theta = \left(\frac{\partial T}{\partial V} + \frac{\partial p}{\partial S}\right)\mathrm{d}S \wedge \mathrm{d}V = 0\,,$$

which gives one of the Maxwell relations

$$\frac{\partial T}{\partial V} = -\frac{\partial p}{\partial S}\,. \tag{5}$$

Since Φ is also given by $U = U(S,V)$ from (4) we get

$$\Phi^*\theta = \left(\frac{\partial U}{\partial S} - T\right)\mathrm{d}S + \left(\frac{\partial U}{\partial V} + p\right)\mathrm{d}V = 0\,,$$

which means that $T = \frac{\partial U}{\partial S}$ and $p = -\frac{\partial U}{\partial V}$. Then (5) can be written as

$$\frac{\partial^2 U}{\partial S \partial V} = \frac{\partial^2 U}{\partial V \partial S}\,,$$

which is a tautology for smooth U. In general the Maxwell relations can be used as a consistency check of equations of motion – if they define a Legendre submanifold.

Another example is the ideal gas which has the equation of state

$$pV = NRT\,,$$

where N is the number of moles of the gas, and R is the universal gas constant. This is not enough for the definition of a Legendre submanifold, and another relation is provided

$$U = \frac{3}{2}NRT\,.$$

These equations are provided for the Lagrange manifolds given by $S = S(U,V)$, which gives

$$\frac{\partial S}{\partial U} = \frac{1}{T} = \frac{3NR}{2U}\,, \quad \frac{\partial S}{\partial V} = \frac{p}{T} = \frac{NR}{V}\,.$$

One can easily check that the mixed second derivatives agree.

For more examples, one can look e.g., into [2], [18] or for more physical view [4].

3 Axiomatic approach

The above description of entropy can be axiomatized. Our presentation in this section closely follows [16] and [17].

We start from the definition of a simple system, as in the previous section. The states of such a system are points $X, Y, Z, \ldots$ inside the space of states Γ. Then we fix on the set Γ the structure of the space $\mathbb{R}^{2n+1}$, where one variable is the energy U and the remaining $2n$ variables are extensive-intensive pairs.

On such a space we can introduce a scaling by $\lambda, \mu \in \mathbb{R}^+$ that is a multiplication group action $\Gamma^1 = \Gamma$, $(\Gamma^\lambda)^\mu = \Gamma^{\lambda\mu}$. The scaled state λX consists of all extensive variables scaled and all intensive variables unaffected. Two systems Γ_1 and Γ_2 can be composed, and then the composed system is described by points from the Cartesian product $\Gamma_1 \times \Gamma_2$.

The fundamental notion needed for the definition of entropy in [16], [17] is adiabatic accessibility

Definition 2. State Y is *adiabatically accessible* from X if the only result of the transition is a work done. We denote it $X \prec Y$.

This definition does not involve heat since it was not defined yet. Besides, the relation $\prec$ is intended to be some 'ordering' to be specified later.

We can further define

Definition 3.
- *Irreversible adiabatic process*: $X \prec\prec Y$ if $X \prec Y$ and not $Y \prec X$;
- *Adiabatic equivalence*: $X \sim Y$ if $X \prec Y$ and $Y \prec X$;

In order to introduce entropy $S : \Gamma \to \mathbb{R}$ the relation $\prec$ is assumed to fulfill the axioms [16], [17]:

- Monotonicity: $X \sim X$
- Transitivity: If $X \prec Y$ and $Y \prec Z$ then $X \prec Z$
- Consistency: $X \prec X'$ and $Y \prec Y'$ implies $(X, Y) \prec (X', Y')$
- Scaling invariance: $\lambda > 0$ and $X \prec Y$ implies $\lambda X \prec \lambda Y$
- Splitting recombination: $X \sim (\lambda X, (1-\lambda)X)$
- Stability: If $(X, \epsilon Z) \prec (Y, \epsilon Z')$ then $X \prec Y$ for $\epsilon \to 0^+$. This means that a 'small' additional system ϵZ cannot perturb ordering of two systems X, Y.

Up to now the relation $\prec$ is a partial order, however it can be made a total ordering by the following Comparison 'Hypothesis' that can be proved using the definition of a simple systems and the Zeroth Law of Thermodynamics [17]

Definition 4. We say that the *Comparison Hypothesis* (CH) holds for a state-space Γ if all pairs of states in Γ are comparable.

These assumptions/hypothesis imply the existence of entropy

Theorem 2. *[16], [17] A function $S\colon \Gamma \to \mathcal{R}$ called entropy exists under assumption of the above axioms and Comparison Hypothesis, and fulfills:*

- *Monotonicity:* $X \prec Y \quad \Leftrightarrow \quad S(X) \leq S(Y)$
- *Additivity:* $S(X,Y) = S(X) + S(Y)$
- *Extensibility:* $S(\lambda X) = \lambda S(X)$

The last issue is to make consistent all local entropies for subsystems and check if the global entropy can be defined. This is done in the following

Theorem 3. *[16], [17] Assume that CH holds for all compound systems. For each system Γ let S_Γ be some definite entropy function on Γ. Then there are constants a_Γ and $B(\Gamma)$ such that the function S, defined for all states of all systems by affine transformation*

$$S(X) = a_\Gamma S_\Gamma(X) + B(\Gamma)\,,$$

for $X \in \Gamma$, satisfies additivity (2), extensivity (3), and monotonicity (1) in the sense that whenever X and Y are in the same state-space, then

$$X \prec Y \quad \Leftrightarrow \quad S(X) \leq S(Y)\,.$$

The total ordering $\prec$ of adiabatic processes and the existence of entropy S that fulfills monotonicity (for simple systems) establishes a link with Theorem 1 of smooth case. It is also a starting point to define entropy in terms of category theory, which will be the subject of the next section.

4 Categorification

In this section, we review some concepts from [13]. For background from category theory see [23] or [20].

We will consider only a simple (i.e., not compound) systems for simplicity. This approach is based on the definition of Poset (pre-ordered set) as a category:

Definition 5. A poset (pre-ordered set) $(P, \prec)$ is a set P with order relation $\prec$. The arrow $x \to y$ for $x, y \in P$ exists, by definition, when $x \prec y$.

We will use the definition for $\prec$ being a total order since this is the case for entropy from previous sections. Then the ordering relation/the arrow $x \to y$ is defined only when y is adiabatically accessible from x.

If scaling of the system is considered, then instead of Poset, the G-Poset category has to be considered [1]. The first step is to define a set with group action – a G-Set [5] – that accommodates the space of states Γ from the previous sections:

Definition 6. System space is the object of the G-Set category, i.e., $\{\Gamma, (\mathbb{R}^+, \cdot, 1)\}$, where the multiplicative group acts on the set Γ.

In the next step, the definition of G-Poset can be adapted for $P = \Gamma$ – the space of states from the previous section – to define the system with entropy. Under the assumption from the previous section, on the poset the ordering is induced by the entropy $S\colon \Gamma \to \mathbb{R}$, and therefore we can define

Definition 7. [13] The *entropy system* is the object of G-Pos category, which objects are $\mathcal{G} = (\Gamma, \preccurlyeq)$, with preserving ordering group $(\mathbb{R}^+, \cdot, 1)$ action[1], where the (partial or) total order is given by the entropy function $S\colon \Gamma \to \mathbb{R}$.

Hereafter we restrict ourselves only to Posets for simplicity. For the general case of G-Posets, see [13].

Up to now, this is only rephrasing of the previous section in terms of 'abstract nonsense', and it does not introduce anything new. The situation, however, changes when we consider more than one entropy system. In this case, we have two or more posets that can represent different (and not necessary originating from thermodynamic) entropy systems. We can ask what is minimal mapping (Functors between these Posets) that preserves ordering, and therefore entropies that introduce these orderings. It occurs that the minimal 'relation' that preserves these orderings in both directions is the Galois connection [21], [23], which can be seen as a basic example/a 'prototype' of adjoint functors. The Galois connection rewritten in terms of orderings induced by entropy functions has the following form

Definition 8 (The Landauer connection and Landauer's functor). [13] Entropy system $\mathcal{G}_1 = (\Gamma_1, S_1)$ is implemented/realized/simulated in the entropy system $\mathcal{G}_2 = (\Gamma_2, S_2)$ when there is a Galois connection between them, namely, there is a functor $F\colon \mathcal{G}_1 \to \mathcal{G}_2$ and a functor $G\colon \mathcal{G}_2 \to \mathcal{G}_1$ such that $F \dashv G$.

In terms of the entropy it is given as

$$S_2(Fc) \leq S_2(\mathrm{d}) \Leftrightarrow S_1(c) \leq S_1(G\mathrm{d}). \tag{6}$$

We name the functors F and G the Landauer's functors.

The Galois connection usually appears in logical/model theory considerations when we have a Poset of some axioms, and we implement them on a Poset of models that realize these axioms [23]. The ordering is then provided by the 'strength' of axiom and model. In this vein, we can use the Landauer's connection to relate some abstract entropy model with its implementation on the physical system with thermodynamical entropy. If such a connection between these two levels model-realization exists, then the change in entropy at the level of the model is transferred through the Ladauer's connection to the change in entropy in the physical realization level. This was the original idea of Landauer [14], [15], who deduced that any irreversible logical operation at the level of Shanon-entropic system generates a physical heat. In terms of the Landauer's connection this heat is generated by the change of entropy in the physical part of the device that implements a logical system. Therefore, the categorical approach makes a sharp distinction, in which part of the compound entropic system such Landauer's heat is generated. This result also explained Maxwell's demon paradox [13].

This sketch presents only one application of the connection. More details and examples from physics, computer science, and biology can be found in [13].

[1] If for $X, Y \in \Gamma$ there is $X \preccurlyeq Y$, then for $\lambda \in \mathbb{R}^+$ there is $\lambda X \preccurlyeq \lambda Y$.

5 Summary

In this paper, we presented the road from entropy in terms of thermodynamics to its categorification. We started from the foundations of thermodynamics and entropy that rely on contact structure. Having understood the motivation, the axiomatic approach to entropy was presented, which emphasizes the ordering of equilibrium states by adiabatic processes. Finally, this ordering was used to reformulate the system with entropy in terms of pre-ordered sets – Posets. Two such Posets can be Galois connected by functors that preserve orderings, and therefore entropies. This connection can be used in various interesting contexts.

Acknowledgments

I would like to thank Valentin Lychagin for pointing me out this interesting subject and fruitful discussions. This overview was written thanks to the encouragement of Pasha Zusmanovich and warm, positive feedback of Lino Feliciano Reséndis Ocampo. I am also grateful to Referees for their detailed and vital suggestions that help to improve the manuscript. This research was supported by the GACR grant GA19-06357S and Masaryk University grants MUNI/A/1186/2018 and MUNI/A/0885/2019. I also thank the PHAROS COST Action (CA16214) for partial support.

A Differential forms

The mathematical structure underlying equilibrium thermodynamics is the theory of differential forms on contact space and their integrability. This section outline the theory, and the interested reader is referred to various sources, including [6]. All theorems are local, which is convenient for applications. Therefore we restrict ourselves to open subsets of Euclidean space, which are diffeomorphic to open subsets of a manifold N, which will have a (local) coordinate chart $(x^1, \dots, x^{\dim(N)})$.

A.1 Frobenius theorem

The basic problem in exterior calculus is to check complete integrability of an exterior system $\{\omega_1, \dots, \omega_n\}$, that is the existence of a submanifold given locally by n relations $\Phi := \{g_i(x) = c_i, i = 1 \dots n\}$, for constants c_i, on which the exterior system vanishes $\Phi^*\omega_i = 0$ for $i = 1 \dots n$. This is given by

Theorem 4. *[6] The exterior system $\{\omega_1, \dots, \omega_n\}$ is completely integrable iff there exists a nonsingular matrix A_{ij} of 0-forms that $\omega_i = \sum_j A_{ij} \mathrm{d}g_j$.*

For a system given by a 1-form Q complete integrability means that there exists an integrating factor (nonsingular 0-form) T such that $Q = T\mathrm{d}S$. This fact is useful in defining entropy.

This can be reformulated in terms of differential ideals. We say that the set I is the differential ideal defined by the set of 1-forms $\{\omega_1, \dots, \omega_n\}$ if and only if for $\eta \in I$ we have $\eta = \sum_i A_i \omega_i$ for 0-forms A_i. In these terms the Frobenius theorem controls complete integrability of the differential ideal defined by the exterior differential system, namely,

Theorem 5. *[6] The ideal I is integrable iff $\mathrm{d}I \subset I$.*

This means that the ideal I is closed under the exterior derivative, i.e., $\mathrm{d}\eta \in I$ if $\eta \in I$. This also means that $\mathrm{d}\omega_i = \sum_j A_{ij}\omega_j$ for 0-forms A_{ij}, or

$$\mathrm{d}\omega_i \wedge \omega_1 \wedge \ldots \wedge \omega_n = 0$$

for $i = 1 \ldots n$.

An alternative version of the Frobenius theorem is formulated for distributions. Define the vector space $D = \mathrm{Span}(\mathrm{Ker}(\omega_1), \ldots, \mathrm{Ker}(\omega_n))$. This means that at each point of the space we define a vector subspace, and we are asking if these subspaces are tangent to some submanifold that is an integral manifold of the distribution D. Then the Frobenius theorem has the form

Theorem 6. *[6] The distribution D is integrable iff $[D, D] \subset D$.*

This means that taking all possible vectors from the distribution (which can be associated with infinitesimal transformations on the manifold), by making their brackets, we cannot get new vectors (infinitesimal transformations) that are outside the distribution D. This observation gives the Caratheodory's theorem on accessibility:

Theorem 7. *[6], [7] If in the neighborhood of any point there are points not accessible by paths which have tangent vectors in the distribution, then the 1-form θ is integrable ($\theta \wedge \mathrm{d}\theta = 0$).*

Summing up, if the distribution/exterior differential ideal is integrable, then it defines a foliation of the manifold/holonomic constraint. However, this statement is local. Foliation can behaves 'pathologically' forming, e.g., initial submanifold [10]. For defining the global structure of the leaves, and to assure that they are proper submanifolds, some additional work must be done.

The Frobenius theorem is useful in proving the existence of entropy, which is the Second Law of Thermodynamics.

A.2 Darboux theorem

The next important theorem is the Darboux theorem that describes local canonical form of the differential 1-form defining contact and symplectic structures on manifold. We present only version for contact form

Theorem 8. *[6] For a 1-form ω fulfilling $\omega \wedge (\mathrm{d}\omega)^n \neq 0$ and $(\mathrm{d}\omega)^{n+1} = 0$ there exists $n+1$ local functions $\{X_i(x)\}_{i=0}^n$ and n functions $\{P_i(x)\}_{i=1}^n$ on the manifold M with coordinates x^i such that the form ω has representation*

$$\omega = \mathrm{d}X_0 + \sum_{i=1}^{n} P_i \mathrm{d}X_i \,.$$

These functions can be used to introduce new coordinates on the manifold in which ω has a simpler form.

A.3 Contact structure

We define

Definition 9. The pair (M, θ) where M is odd dimensional manifold of dimension $2n+1$ and θ is non-degenerate 1-form that fulfills $\theta \wedge (\mathrm{d}\theta)^n \neq 0$, is called contact manifold.

Since $\dim(M) = 2n+1$ and $\deg(\theta \wedge (\mathrm{d}\theta)^n) = 2n+1$ therefore $(\mathrm{d}\theta)^{n+1} = 0$. We can use the Darboux theorem to conclude that locally we can introduce coordinates that $\theta = \mathrm{d}X_0 + \sum_{i=1}^n P_i \mathrm{d}X_i$.

The contact space and the contact form are in thermodynamics introduced by the First Law of Thermodynamics.

The contact structure is solvable by a submanifold Φ that fulfills $\Phi^*\theta = 0$. We can ask about the maximal dimension of such submanifold. This is controlled by the following:

Theorem 9. *[11] Every maximal submanifold in a $2n+1$ dimensional contact manifold M has dimension n and is called Legendre submanifold.*

A.4 Contact structure vs Jet space

We finish this overview of differential geometry by discussing the rudiments of jet spaces. This presentation is mainly based on [11], [12].

Consider a n dimensional manifold N and smooth functions on the manifold $C^\infty(N)$. In local coordinates $(x_1, \dots x_n)$ define the ideal

$$\mu_a^k := \left\{ f \in C^\infty(N) \;\middle|\; \frac{\partial^{|\sigma|} f}{\partial x^\sigma}(a) = 0, 0 < |\sigma| < k \right\},$$

where multiindices $\sigma = (\sigma_1, \dots, \sigma_n)$, $|\sigma| = \sum_{i=1}^n \sigma_n$, and $\frac{\partial^{|\sigma|} f}{\partial x^\sigma} := \frac{\partial^{|\sigma|} f}{\partial (x^1)^{\sigma_1} \dots (x^n)^{\sigma_n}}$.

Now define the k-th jet of functions at $x = a$ as the quotient

$$J_a^k(N) := C^\infty(N) \,/\, \mu_a^{k+1}.$$

The equivalence classes $[f]_a^k \in J_a^k(N)$ represent the functions that have the same derivatives/contact at $x = a$ up to order k, in other words, their Taylor series at $x = a$ agree up to order $(x-a)^k$. For example for $\dim(N) = 1$, $[x]_0^i = [\sin(x)]_0^i$ for $i = 0, 1, 2$ but disagree for $i = 3$.

The k-jet of functions on N is defined as

$$J^k(N) = \bigcup_{a \in N} J_a^k(N).$$

It is a fiber bundle $\pi\colon J^k(N) \to N$ with the obvious projection.

We can now describe $J^k(N)$ locally by defining the ideal of 1-forms (the Cartan distribution). For simplicity consider $J^1(N)$. The local coordinates are (x^i, y, y_i) where the new coordinates p_i are associates with derivatives $\frac{\partial}{\partial x^i}$ of functions $C^\infty(N)$ by the Cartan distribution

$$\omega = \mathrm{d}y - p_i \mathrm{d}x^i.$$

The distribution is nonintegrable since $\omega \wedge \mathrm{d}\omega \neq 0$. In addition, $\omega \wedge (\mathrm{d}\omega)^n \neq 0$ and $\dim(J^1(N)) = 2n+1$. This is exactly the local form from the Darboux theorem and also from the local definition of a contact form. Therefore the contact space M of dimension $2n+1$ is exactly the 1-jet of smooth functions on N.

The sections of the jet bundle $s\colon N \to J^1(N)$ are called holonomic sections or 1-graphs (in case of $J^k(N)$ are called k-graphs) if 'differential' coordinates are derivatives, i.e.,

$$x^i(s) = x^i, \quad y(s) = f, \quad p_i(s) = \frac{\partial f}{\partial x^i}\,.$$

We can note that the section is a holonomic section iff it is a Legendre submanifold [19]. This means that we can describe a Lagrange submanifold by a function $y(s) = f(x^1, \ldots, x^n)$ and then all p coefficients in the Cartan distribution or P coefficients in a contact form are derivatives

$$\begin{cases} p_1 = p_1(x_1, \ldots, x_n) = \frac{\partial f}{\partial x_1}\,, \\ \ldots \\ p_n = p_n(x_1, \ldots, x_n) = \frac{\partial f}{\partial x_n}\,. \end{cases}$$

Symmetries of contact structure are such transformations that preserve the Cartan distribution [11], [12]. For a diffeomorphism $\phi\colon J^1(N) \to J^1(N)$ the following condition ensures that it is a contact symmetry:

$$\phi^*\omega = \lambda_\phi \omega\,,$$

where λ_ϕ is some smooth non-vanishing function on $J^1(N)$. This condition shows that the kernel of $\phi^*\omega$ is the same as the kernel of ω – they define the same contact distribution.

Apart of simple symmetries like translation $(x^i, y, p_i) \to (x^i + \alpha^i, y + \beta, p_i)$ the most important symmetry in thermodynamics is the Legendre transformation:

$$(x^i, y, p_i) \to (p_i, y - x^i p_i, -x^i)\,,$$

that interchange x^i with corresponding p_i.

$J^1(N)$ is sufficient for thermodynamics, however for general theory of jet spaces consult [10], [11], [12].

References

[1] E. Babson, D.N. Kozlov: Group actions on posets. *Journal of Algebra* 285 (2) (2005) 439–450.

[2] P. Bamberg, S. Sternberg: *A Course in Mathematics for Students of Physics: Volume 2.* Cambridge University Press (1990).

[3] J.B. Boyling: An axiomatic approach to classical thermodynamics. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 329 (1576) (1972) 35–70.

[4] H.B. Callen: *Thermodynamics.* John Wiley & Sons Inc. (1966).

[5] T.T. Dieck: *Transformation Groups and Representation Theory.* Springer, Lecture Notes in Mathematics 766 (1979).

[6] D.G.B. Edelen: *Applied exterior calculus*. Dover (2011).

[7] T. Frankel: *The geometry of physics: An introduction*. Cambridge University Press (2011).

[8] R. Ingarden, A. Jamiołkowski, R. Mrugała: *Fizyka statystyczna*. PWN (1990).

[9] A. Katok, B. Hasselblatt: *Introduction to the modern theory of dynamical systems*. Cambridge University Press (1996).

[10] I. Kolář, P.W. Michor, J. Slovák: *Natural operations in differential geometry*. Springer-Verlag Berlin Heidelberg (1993).

[11] A. Kushner, V. Lychagin, V. Rubtsov: *Contact geometry and nonlinear differential equations*. Cambridge University Press (2007).

[12] A. Kushner, V. Lychagin, J. Slovák: Lectures on Geometry of Monge-Ampère Equations with Maple. In: R. Kycia, M. Ułan, E. Schneider: *Nonlinear PDEs, Their Geometry, and Applications*. Birkhäuser (2019) 53–94.

[13] R.A. Kycia: Landauer's principle as a special case of Galois connection. *Entropy* 20 (12) (2018) 971.

[14] J. Ladyman, S. Presnell, A.J. Short, B. Groisman: The connection between logical and thermodynamic irreversibility. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics* 38 (1) (2007) 58–79.

[15] R. Landauer: Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development* 5 (3) (1961) 183–191.

[16] E.H. Lieb, J. Yngvason: A guide to entropy and the second law of thermodynamics. In: B. Nachtergaele, J.P. Solovej, J. Yngvason: *Statistical Mechanics*. Springer (1998) 353–363.

[17] E.H. Lieb, J. Yngvason: The physics and mathematics of the second law of thermodynamics. *Physics Reports* 310 (1) (1999) 1–96.

[18] V.V. Lychagin: Contact Geometry, Measurement, and Thermodynamics. In: R. Kycia, M. Ułan, E. Schneider: *Nonlinear PDEs, Their Geometry, and Applications*. Birkhäuser (2019) 3–52.

[19] V.V. Lychagin: Contact geometry and non-linear second-order differential equations. *Uspechi Mat. Nauk* 34 (1) (1979) 137–165.

[20] S. Mac Lane: *Categories for the working mathematician*. Springer (1978).

[21] O. Ore: Galois connexions. *Transactions of the American Mathematical Society* 55 (3) (1944) 493–513.

[22] F.M. Reza: *An introduction to information theory*. Dover Publications (1994).

[23] P. Smith: Category theory: A gentle introduction. University of Cambridge (2018)

[24] W. Li, Y. Zhao, Q. Wang, J. Zhou: Twenty years of entropy research: A bibliometric overview. *Entropy* 21 (7) (2019) 694.

Received: 17 August 2019
Accepted for publication: 27 November 2019
Communicated by: Pasha Zusmanovich