The principles of neopositivism and the laws of thermodynamics

D. Lairez[*]
Laboratoire des solides irradiés, École polytechnique, CEA, CNRS, IPP, 91128 Palaiseau, France

(Dated: April 27, 2026)

The second law of thermodynamics, which deals with irreversibility and makes the theory so special, is usually considered empirical. The definition of equilibrium as an attractor, on the other hand, requires a postulate. This article shows that both are actually already contained, even if hidden, in the fundamental principles of neopositivism, which are widely accepted in all fields of science. In particular, from the definition of information as a truth that can only come from an observation but cannot be redundant, we obtain Clausius' inequality.

INTRODUCTION

Symmetries of an experiment in physics are the possible transformations that leave its result unchanged. They are those that are undetectable by simply observing the result of the experiment. A symmetry leaves invariant certain observed regularities of the world and the laws of physics that account for them [1].

For instance, rotation by a given angle is a symmetry for all the laws of physics. But, rotation at a given angular velocity is not. Translation in space is a universal symmetry. Translation in time is also a universal symmetry, but time reversal is not a universal symmetry. And this last point is very confusing and annoying for physicists. What is the problem? It lies in the fact that all the fundamental laws of physics, such as those of electromagnetism and mechanics (classical, relativistic or quantum), are invariant under time reversal, with the exception of the second law of thermodynamics concerning irreversibility. This therefore represents a major challenge, as it is likely that the unification of fundamental interactions (that is, the unification of relativistic and quantum mechanics) will not be sufficient to arrive at the mythical theory of everything. The unification of physics will also require taking the laws of thermodynamics into account. Any new insight into these laws is therefore welcome.

In physics, and more broadly in all scientific fields, statements of a theory belong to two categories : laws and principles [2], also called empirical laws and theoretical laws [3], which are of a different nature. The former come from induction and are simply the generalization of particular observations. They express observed regularities of the world and are true until proven otherwise. The latter concern the rest, that is to say everything that cannot be observed, such as the definition of concepts used to express empirical laws and the way in which these concepts articulate. They are actually conventions or postulates. They are neither true nor false (because they cannot be faced to experiments), they are simply convenient or not with regard to the object of science, namely describing reality which itself remains to be defined by convention. Principles in physics function fundamentally like axioms in mathematics, except that they are not arbitrary but have pragmatic obligations [4].

The first convention preceding any theory concerns the very object and purpose of science. Since science explicitly refers (by its etymology) to knowledge, this first convention also concerns what we mean by "knowledge". The aim of this article is to show that if we adopt the neopositivist [5] conventions on these points (in short : there is no synthetic *a priori* knowledge other than that of logic; the sole object of science is reality understood as that which can be observed) the empirical second law about irreversibility can be deduced. In other words, the empirical second law about irreversibility is already contained in a set of fundamental principles which are common to all fields of physics : *"the scientific conception of the world"* [5].

The derivation places the concepts of knowledge, information and data in a fundamental position. A piece of information is a truth that cannot be redundant but is necessarily carried by data understood as a piece of observation. From this asymmetrical role of information and data originates thermodynamic irreversibility.

The article is organized as follows : in the first section the fundamental principles of neopositivism are presented; the second section deals with the second law of thermodynamics (the first on energy conservation not being specific to it) and the definition of equilibrium; finally in the third section we will see how these laws of thermodynamics are actually contained in the fundamental principles.

I. FUNDAMENTAL PRINCIPLES OF NEOPOSITIVISM

It is impossible to do science without conventions, also called postulates or principles, at least those that concern its object matter and what we mean by explaining and understanding. These epistemological conventions are often overlooked, which leads to useless debates simply because people are not talking about the same thing. This will be avoided by stating them from the outset. My personal observation of the scientific practice shows that these conventions are widely shared and should not offend many people. However, even if this were the case,

---

[*] didier.lairez@polytechnique.edu

these people should not deny that accepting these conventions implies certain things, which is precisely the subject of this article. These conventions are those of the Vienna circle [5] and of its neighbourhood [2–4, 6–9], namely the neopositivism, or logical positivism, or logical empiricism. It is briefly presented below, limiting ourselves to what will be necessary.

The only thing that science can talk about is scientific reality. *"Whereof one cannot speak, thereof one must be silent"* (Wittgenstein [6]), so that there is no other reality in scientific statements. It is defined as :

Principle (reality): *Reality is what can be observed, with observation being understood as an interaction.*

This reality is not made up of objects that could exist independently of us. Reality is inseparable from the observer. This has the consequence, in particular, of requiring us to define what we mean by "objectivity". It is no longer the property of an object, but that of an observation which must be reproducible by anyone and which then becomes true by definition.

Principle (objectivity): *An objective observation can be reproduced by anyone and is true by definition.*

In what follows, we will only consider objective observations and omit the adjective.

We expect science to help us understand reality, and understanding means connecting observations together. These connections can be of different kinds corresponding to different levels of understanding. It is agreed that

Principle (understanding): *The highest level of understanding is reached when all observations, past and future, can be deduced from a finite set of statements that forms a theory.*

This set of statements must obviously contain no contradictions since it is intended to be used by logic.

Any statement is judged with respect to its truth. A statement is true if it relates (narrates) an observation, a fact. The truth of a statement is a quality, not a gradual quantity. That is, a true statement cannot become increasingly true and cannot be truer than another true statement. A statement is false if it contradicts an observation. True and false statements are said to be *a posteriori*. *A priori* statements are those which are independent of any observation. They are neither true nor false.

The statements of the theory are classically divided into two categories : 1) *A priori* statements, called principles, or postulates, or conventions. Neither true nor false, they are convenient. Their convenience is judged by whether or not they allow us to deduce observations from theory; 2) True *a posteriori* statements resulting from generalization of observed regularities. They are called inductive or empirical laws, simply called laws in the following. Laws are true until proven otherwise by the observation of a counter-example. The above can be summarized as follows:

Principle (truth): *The only source of truth is observation and truth is a quality, not a gradual quantity.*

Statements can be deduced from the theory, but their truth is conditioned on conventions; moreover, they can be considered as already contained in the theory, even if hidden, and not as being the source.

For what follows, let us to clarify certain vocabulary points concerning truth, information, and knowledge.

Definition (information): *A piece of information is an unconditional truth supported by an observation.*

Definition (knowledge): *For the observer, to know something is the state of possessing the footprint of a piece of information.*

How exactly knowledge as a state quantity can be evaluated is precisely the contribution of information theory that will be discussed in the following. But we can already say that since truth is a quality that cannot become more and more true, it follows that pieces of information cannot be redundant. Or in other words :

Principles (no redundancy): *Redundant pieces of information do not increase knowledge and counts as one.*

From the definition of reality, it comes that there are no parallel universes. There is no Platonic intelligible and visible world. There is no spiritual and material world. There is no dualism. The universe is one. The deduction on a particular problem from the theory must also be unique (or univocal). But it is like finding the solution to an equation, the solutions are often multiple. Randomness and probabilities are there precisely to address this problem. The solution (the deduction provided by the theory) is then a unique probability distribution.

Principle (univocality): *On a particular problem, deductions from the theory must be univocal in terms of probability distribution.*

Probabilities introduce randomness. Whether randomness is inherent to the nature of things or due to our ignorance leads to exactly the same observations and goes beyond the scope of science. Traditionally, two types of probabilities are distinguished : *a priori* and *a posteriori*. As indicated by their name, the first are completely detached from observation. They are actually conventions. Whereas, the second are relative statistical occurrences of observations. They are empirical laws. Both are mathematically treated in the same manner and both are valid. But what is not permitted is either to say that the probabilities are *a priori* distributed in a certain way, when they have been measured to be distributed in another way (this would be false); or to say that the probabilities are *a posteriori* without having been measured or without the possibility of doing so (this would be inconsistent).

An infinite set of possible observations is deduced from a finite set of statements and, in a certain sense, is contained within it. Theory is therefore a summary of reality. It is an economy of thought [7] and the most economical is the best. This is known as the Occam's razor [10]. We are therefore faced with an optimization problem : *"The basic concepts and laws which are not logically further reducible constitute the indispensable and not rationally deducible part of the theory. It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience"* (Einstein [9]). The minimization procedure applies to the number of statements of the theory (conventions and laws), but it also applies to the laws themselves. An inductive law is similar to inter- or extrapolation of experimental measurements (observations). On the one hand, the law has the constraint to account for all observations, to account for everything that is true, to account for all knowledge. But on the other hand, it must be as simple as possible. The best law must provide the "minimum service" within the imposed constraint of our knowledge. These two ideas about what is optimum, what is the *"best"*, what is the *"supreme goal"* (for the number of laws and for the laws themselves) are actually conventional value judgments. They can be expressed all in one under the form of an additional postulate :

Principle (least-talker): *Theory must minimize what is said (maximize the uncertainty) about what we do not know, with the constraint of saying everything about what we know.*

Here is our final postulate, which sounds like an echo of the first : *"Whereof one cannot speak, thereof one must be silent"* (Wittgenstein [6]). For this reason , the list is certainly not logically irreducible, but by being less concise it is considered clearer.

II. STATEMENTS OF THERMODYNAMICS

By thermodynamics, we mean the initial phenomenological theory (Clausius [11]), which deals only with the macroscopic observable behavior of a system and not with its internal functioning.

Thermodynamics deals with the changes in form of energy. The first statement of the theory is therefore that of its conservation. This statement is actually a convention [2, 12] and more precisely the definition of the concept itself. This principle is not specific to thermodynamics and is in fact common to all of physics; it will therefore not be addressed here and the reader is invited to consult a previous article [13].

In thermodynamics, a system is described by state quantities and the state of a system is a way of being that is supposed to be maintained during a certain duration. Equilibrium is the stationary state for which this duration is not limited and is actually the only object matter of thermodynamics ([14] p.15).

But there is a form of paradox in the intentions. How can we account for the transformations a system can undergo (the changes in form of energy) if we only deal with stationary states ? The solution lies in the notion of quasi-static trasnformation and that of reversibility. Another problem arises : a system out of equilibrium always tends towards equilibrium. Equilibrium is an attractor and this cannot be deduced from its definition as a stationary state.

Reversibility (thus irreversibility) and attractiveness of the equilibrium, are accounted for with the concept of entropy [11], which will be the entry point for the notion of information in the problem. So let us start there.

A. Entropy and reversibility

In thermodynamics, a reversible cycle is a path that leads the system from an initial state through a sequence of other states and finally returns to the starting point, leaving everything (the system and its environment) exactly as it was before. Any process that can be part of such a cycle is said to be reversible. For example, a quasi-static process sufficiently slow relative to any relaxation of the system is reversible. A quasi-static path is a succession of (quasi)equilibrium. It is under control and can be stopped at any time and then resumed or reversed to return to the starting point. Note that, from the outset, we can clearly see the role that knowledge and information will play in thermodynamics: identifying whether everything is as before or being able to control a parameter depend on our knowledge.

Clausius [15] observed that for any reversible cycle :

$$\left(\oint \frac{\mathrm{d}Q}{T}\right)_r = 0 \tag{1}$$

where Q is the heat exchanged by the system with its environment, which sign is relative to the system (positive when received); the subscript "r" stands for "reversible"; and T is the temperature (in Joule) of the system. This observation makes

$$\mathrm{d}S = \left(\frac{\mathrm{d}Q}{T}\right)_r \tag{2}$$

an exact differential that allows us to define a new state-quantity S (that only depend on the state) called entropy (it can also be called Clausius entropy to remove any ambiguity). This is the initial definition of entropy. The difference in entropy between two states A and B is given by :

$$S_B - S_A = \left(\int_A^B \frac{\mathrm{d}Q}{T}\right)_r \tag{3}$$

Basically, the difference in entropy between two states is equal to the net heat exchanged expressed in unit of temperature when the path linking them is reversible. This can be considered an empirical law.

B. Entropy and irreversibility

Reversibility is actually a limiting behavior that is never reached, but towards which quasi-static processes tend when their rate of change tends towards zero. Instead of Eq.1, for any differentiable cycle what is actually observed is that heat is always transferred to the environment

$$\oint \frac{\mathrm{d}Q}{T} \leq 0, \tag{4}$$

so that things are never the same as before.

Instead of 2 and 3, for any differentiable actual process, one has :

$$\mathrm{d}S = \left(\frac{\mathrm{d}Q}{T}\right)_r \quad \geq \quad \frac{\mathrm{d}Q}{T} \tag{5}$$

and

$$S_\mathrm{B} - S_\mathrm{A} = \left(\int_A^B \frac{\mathrm{d}Q}{T}\right)_r \quad \geq \quad \int_A^B \frac{\mathrm{d}Q}{T} \tag{6}$$

For non-differentiable processes, the integral on the right-hand side has to be replaced by a sum. These are different form of the Clausius inequality, which expresses thermodynamic irreversibility and can be considered empirical.

With the definition of entropy (Eq. 2), Clausius inequality forms the second law of thermodynamics (the first being the principle of conservation).

C. Entropy and equilibrium

Let us return to the problem of the attractiveness of equilibrium. As is the case in classical mechanics, the problem is solved if the equilibrium is defined as the state that minimizes a potential [14]. In thermodynamics, this potential always involves the negative of entropy (negentropy). Its very definition depends on the interactions between the system and its environment. For example, for a thermalized system it is the free energy $F = U - TS$, where U is the internal energy, etc.

For simplicity, let us consider the union of a system of interest and its environment, the latter being chosen in such a way that the union can be considered as an isolated system. The system is in a stationary state when its interactions with the environment are also in a stationary state. So that the equilibrium of the system is also that of the union, that is to say it boils down to that of an isolated system. The definition of the equilibrium for the latter is then postulated as :

Principle (of maximum entropy): *The equilibrium of a thermodynamic isolated system is that of maximum entropy.*

For an isolated system, no heat is exchanged with the environment and Eq.6 turns into

$$S_\mathrm{B} - S_\mathrm{A} \geq 0 \tag{7}$$

So that, the entropy of an isolated system can only increase. Conjointly with the above definition, this makes the equilibrium an attractor.

III. THE LAWS OF THERMODYNAMICS FROM THE FUNDAMENTAL PRINCIPLES

As already mentioned, the second law of thermodynamics summarized in §II A and II B originates from observations. It is empirical. The aim of this section is mainly to show that it can be derived from the fundamental principles listed in §I. The derivation passes first through the contribution of statistical mechanics (Gibbs [16]) and that of information theory (Shannon [17]) which provide new interpretations of what entropy is and allowing to link heat (thus energy) and information.

A. Entropy and improbability

With the advent of atoms in physics comes also thermal agitation, probabilities and statistics. The system constantly changes its microscopic configuration (also named microstate, or phase). So that the equilibrium (the purpose of thermodynamics) should be understood as an average state. This is the contribution of statistical mechanics: having found a way to link microscopic to macroscopic scales by expressing state quantities, such as entropy, in terms of average quantities over microstates.

For the calculation, we must first decide on their probability distribution. Knowing that it cannot be measured, this distribution is necessarily *a priori* postulated. *"When one does not know anything the answer is simple. One is satisfied with enumerating the possible events and assigning equal probabilities to them."* (Balian [18]). The distribution to be used is uniform. This is the fundamental postulate of statistical mechanics (a variation of the Laplace's principle of insufficient reason [19]). Against this postulate Ellis wrote : *"It cannot be that because we are ignorant of the matter we know something about it"* [20]. But according to §I, a postulate is not knowledge so that there is actually no inconsistency. As all conventions, this postulate has no obligation to be confirmed by direct observations. Only indirect feedback is required; it will be provided by the agreement between the calculated macroscopic quantities and those measured. And it has to be said, it works !

Let us denote S_Gibbs the average of the logarithm of improbabilities $1/p_i$, where p_i is the probability of microstate i.

$$S_\mathrm{Gibbs} = \sum_{i\in\Gamma} p_i \ln 1/p_i, \tag{8}$$

Γ is the ensemble of possible microstates (the phase space). By starting with a uniform distribution of microstates for an isolated system (the whole) about which nothing is known, we can derive that of any subpart (see e.g. [21]), then deduce the equality :

$$S_{\text{Gibbs}} = S \tag{9}$$

hence the name Gibbs-entropy (or statistical entropy) for S_{Gibbs}. This last equality, which has shed new light on the concept of entropy, is only conditioned on the postulate of a uniform distribution of micro-states for the whole.

B. Entropy and uncertainty

Consider each microstate identified by an integer $i \in [0, \mathscr{W} - 1]$ (with $\mathscr{W}$ is the cardinality of Γ) and encode i in binary. A thermodynamic system can be considered a dynamic random source of these numbers. Independently of thermodynamics and of statistical mechanics, Shannon [17] was interested in the recording space needed to store such random outputs. An initial idea [22] was to allocate the same storage space to everyone, i.e. $\log_2 \mathscr{W}$ bits per output in order to be able to store the largest possible integer when it comes out. But knowledge of the probability distribution of the outcomes, and the allocation of a different storage space for each, make it possible to reserve the largest identifiers for the least probable outcomes and a lossless compression of their storage. Interestingly, Shannon showed that in no case the average number of bit per output can be less than :

$$H = \sum_i p_i \log_2(1/p_i) \tag{10}$$

H was named by Shannon quantity of information emitted by the source. Identification of Eq.10 with the Gibbs formula Eq.8 is immediate, so that

$$S_{\text{Shannon}} = H \times \ln 2 \tag{11}$$

was called Shannon-entropy of the random variable under consideration. Gibbs-entropy is the Shannon-entropy of microstates. To a factor, and only conditioned on the validity of Eq.9, the thermodynamic entropy of a system turns out to be equal to the quantity of information it emits (the quantity of information of microstates).

Let us make a few remarks about the denomination "quantity of information" emitted by the system. If we (the receiver) know nothing about the emission of the source, it must be planed at least $\log_2 \mathscr{W}$ per output. There is no compression. But if we learn that it is now certain that the first bit will not change, then it is no longer necessary to record that bit. This piece of information allowed us to save one bit of storage space. And this, for each bit of information until H is zero when there is no uncertainty about the outcomes. Note that when a piece of information is provided to the receiver, it saves one bit of storage space. It is not necessary to provide this twice; the economy would not increase further. The thing called "information" that is quantified by Shannon, obeys to the principle of no redundancy and is conform to what is usually understood as information. The term "quantity of information" is thus in a certain sense appropriate. But in my opinion, it would have been clearer if the term "information" had remained reserved for a single truth (i.e., a quality), and if the term "knowledge" (instead of "quantity of information") had been used when it came to quantifying it. In any way, H represents the knowledge we lack regarding the current microstate of the system. Or, from another perspective, H is a measure of the uncertainty about it.

Let us go back to the problem of the *a priori* probability distribution of microstates which alone conditions the validity of Eq.9. Consider the probability distribution of random outcomes of which we only have partial knowledge. Shore and Johnson [23] showed that maximizing H (instead of another possible measure of uncertainty) is the only procedure that respects the principle of univocality of §I. If we add that of least-talker, maximizing H is the only procedure to rationally decide for any *a priori* probability distribution. Hence the theorem :

Maximum Shannon-entropy theorem: *The best a priori distribution of any random outcomes is that of maximum Shannon-entropy satisfying the constraint of taking into account our knowledge.*

This is a theorem that can be considered as already contained in the fundamental postulates of §I. In the case where we know nothing, the maximum of Shannon-entropy is obtained for the uniform distribution. Thus, the fundamental postulate of statistical mechanics does not come as an addition to the theory; it was also in fact already contained in the fundamental principles of neopositivism. And this holds *de facto* for Eq.9 and for the equality between Clausius entropy and the Shannon quantity of information about microstates. So that :

$$S = S_{\text{Gibbs}} = S_{\text{Shannon}} \tag{12}$$

This equality is not conditional on any postulates other than those listed in §I.

C. Clausius inequality from the principle of truth

Shannon quantity of information H is actually a quantity of information we (the observer) lack. For our purpose, it is likely clearer to speak in terms of the quantity of information we have, that is to say in terms of our knowledge K :

$$K = -H = -S/\ln 2 \tag{13}$$

To a factor, knowledge K is nothing other than Brillouin's negentropy [24]. Consider first the reversible case, by rewriting Eq.3 as :

$$K_B - K_A = -\int_A^B \frac{\mathrm{d}Q}{T \ln 2} \tag{14}$$

The minus sign in front of the integral indicates that what we receive comes out from the system.

- The left-hand side of Eq.14 refers to the observer only. It is the variation in our knowledge obtained through the observation of this process.

  $K_B - K_A$ is knowledge derived from an observed quantity.

- The right-hand side of Eq.14 refers to the system only. It is the quantity of observations that can be made, from the environment, on the system by interacting with it, by exchanging heat with it.

  $-\int_A^B \frac{\mathrm{d}Q}{T\ln 2}$ is an observable quantity.

The process is reversible in case of equality between these two quantities, between what is observed and what is observable. The process is no longer reversible if it is observed episodically, intermittently, incompletely. But in no case can the observed quantity (and the pieces of information it brings) exceed the observable quantity.

But there is another important feature. Imagine that the reversible process occurs at constant internal energy for the system. All quantity of heat, Q, provided to the environment is compensated by an equivalent quantity of work, W, received: $\mathrm{d}Q = -\mathrm{d}W$. The work provided to the system is an interaction that counts as an observation. So the same observation can be made twice, by measuring heat and by measuring work. But in virtue of the principle of no redundancy, the two count for one single piece of information.

Consider an elementary increase of knowledge $\mathrm{d}K$. On one side, it is necessarily brought to us by the mean of an elementary interaction $-\mathrm{d}Q/T\ln 2$ (an elementary observation), which basically amounts to copy something that comes from the system into our memory. But on the other side, an elementary observation $-\mathrm{d}Q/T\ln 2$ does not necessarily carries an information if it is redundant. $\mathrm{d}K$ is a piece of information we get that cannot increase once it is known. $-\mathrm{d}Q/T\ln 2$ is a data that can be copied, duplicated and redundant as much as we want. From this asymmetrical behavior, it follows that:

$$\mathrm{d}K \quad \leq \quad -\frac{\mathrm{d}Q}{T\ln 2} \tag{15}$$

which is nothing but the Clausius inequality (Eq.5). This stems from the postulate that information necessarily requires observation, but cannot be given twice.

D. Knowledge and equilibrium

The last point to examine is that of the definition of thermodynamic equilibrium. The need for it to be an attractor led us in §II C to define it as the state of maximum entropy. But this can be considered as an additional postulate. Maximum Shannon-entropy theorem (§III B) potentially gives us the possibility of deducing this definition, thus making the economy of a postulate.

However, the Maximum Shannon-entropy theorem does not tell us the distribution of which random variable should be maximized at the thermodynamic equilibrium. Different random variables may lead to different incompatible definitions of the equilibrium, inconsistently with the principle of univocality. Jaynes [25] note the following point. Consider the three physical symmetries: translation, rotation and scaling. Any definition of the equilibrium that is not invariant under these symmetries would actually be equivocal. Or in other words, any definition of equilibrium that depends on these symmetries would say more about equilibrium than they should, inconsistently with the least-talker principle. Actually, a definition of thermodynamic equilibrium that is consistent with the fundamental principles must maximize Shannon-entropy of random variables which distribution is invariant in form upon these symmetries. Examples of these variables include local density and microstate. In virtue of Eq.12, Maximizing Maximum Shannon-entropy of these variables is thus equivalent to maximizing Clausius entropy of the system.

The definition of thermodynamic equilibrium is also already contained in the fundamental principles.

IV. CONCLUSION

In thermodynamics a central state quantity of a system, namely its entropy, is related to our knowledge. A state quantity of a system is in fact also a state quantity of the observer. This was already recognized long before Shannon. In 1878, Maxwell wrote *"The idea of dissipation of energy depends on the extent of our knowledge"* [26]. It is therefore not surprising and likely inevitable to relate thermodynamics to epistemology, i.e. the branch of philosophy which deals with knowledge. Fundamental principles concerning what reality is, what the sole source of truth is, what knowledge is, in fact allow us to deduce the second law of thermodynamics, that of the irreversibility of phenomena, this law which makes thermodynamics so special compared to all other theories of physics. Interestingly, these fundamental principles are not special to thermodynamics, but common to all fields of science and in particular to that of physics. In my opinion, unification of physics lies on this side.

"A reality completely independent of the mind which conceives it, sees or feels it, is an impossibility. A world as exterior as that, even if it existed, would for us be forever inaccessible. But what we call objective reality is, in the last analysis, what is common to many thinking beings, and could be common to all" (Poincaré [4]).

---

[1] R. Feynman, The Feynman lectures on physics vol. I, chap. 52: Symmetry in physical laws (1963).
[2] H. Poincaré, *Science and hypothesis* (The Walter Scott Publishing Co., 1905).
[3] R. Carnap, *Philosophical foundations of physics* (Basic books, New York, 1966).
[4] H. Poincaré, *The value of science* (The science Press, 1907).
[5] H. Hahn, O. Neurath, and R. Carnap, The scientific conception of the world: The Vienna circle, in *Empiricism and Sociology* (Springer Netherlands, Dordrecht, 1973) pp. 299–318.
[6] L. Wittgenstein, Tractatus logico-philosophicus (Kegan Paul, Trench, Trubner and Co. Ltd., 1922).
[7] E. Mach, *History and root of the principle of the conservation of energy* (The open court publishing, 1911).
[8] E. Mach, *The Science of Mechanics* (The Open Court Publishing Company, 1919).
[9] A. Einstein, On the method of theoretical physics, Philosophy of Science **1**, 163 (1934).
[10] H. Hahn, Superfluous entities, or Occam's razor, in *Empiricism, Logic and Mathematics* (Springer Netherlands, 1980) pp. 1–19.
[11] R. Clausius, *The mechanical theory of heat* (Macmillan & Co, London, UK, 1879).
[12] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman lectures on physics* (Addison-Wesley, Reading, MA, 1966) Chap. 4.
[13] D. Lairez, Disentangling Brillouin's negentropy law of information and Landauer's law on data erasure, Entropy **28**, 37 (2025).
[14] H. B. Callen, *Thermodynamics and an introduction to thermostatistics*, 2nd ed. (J. Wiley & sons, 1985).
[15] R. Clausius, Sur diverses formes facilement applicables qu'on peut donner aux équations fondamentales de la théorie mécanique de la chaleur., Journal de Mathématiques Pures et Appliquées , 361 (1865).
[16] J. Gibbs, *Elementary principles in statistical mechanics* (Charles Scribner's sons, 1902).
[17] C. E. Shannon, A mathematical theory of communication, The Bell System Technical Journal **27**, 379 (1948).
[18] R. Balian, *From microphysics to macrophysics* (Springer Berlin Heidelberg, 1991).
[19] H. Dubs, The principle of insufficient reason, Philosophy of Science **9**, 123 (1942).
[20] R. Ellis, Remarks on an alleged proof of the "method of least squares," contained in a late number of the edinburgh review, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **37**, 321 (1850).
[21] D. Lairez, A short derivation of Boltzmann distribution and Gibbs entropy formula from the fundamental postulate, arXiv (2022), 2211.02455v3.
[22] R. V. L. Hartley, Transmission of information, Bell System Technical Journal **7**, 535 (1928).
[23] J. Shore and R. Johnson, Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy, IEEE Transactions on Information Theory **26**, 26 (1980).
[24] L. Brillouin, *Science and Information Theory* (Dover Publications, Mineola, N.Y., 1956).
[25] E. T. Jaynes, The well-posed problem, Foundations of Physics **3**, 477 (1973).
[26] J. C. Maxwell, Diffusion, Encyclopedia Britannica, reproduced in Scientific papers **2**, 625 (1878).