# Historical and Conceptual Foundations of Information Physics

Javier Anta Pulido

UNIVERSITAT DE BARCELONA

Doctoral Thesis

# Historical and Conceptual Foundations of Information Physics

Author: Javier Anta Pulido
Supervisor: Carl Hoefer

PhD Program in Cognitive Science and Language
Facultat de Filosofia

December, 2021

*Para mi madre y mi padre, por hacerlo posible.*

# Index

# *List of Figures, Images and Tables*

# *Glossary of Symbols and Acronyms*

| | |
|---|---|
| TD | Thermodynamics |
| $\Omega$ | State Space |
| $\Omega_A$, $\Omega_B$, … | Thermodynamic State |
| $S_{TD}$ | Thermodynamic Entropy |
| SM | Statistical Mechanics |
| BSM | Boltzmann Statistical Mechanics |
| $A_i$ | Macrovariables |
| $H_B$ | H-quantity |
| $\mu$-space | Individual-Molecules Space |
| D | Individual-Molecules Arrangement |
| $S_{B\mu}$ | Boltzmann Combinatorial Entropy |
| $\Gamma$ | Phase Space |
| $\Gamma_E$ | Energy Hypersurface |
| $x$ | Microstate |
| $\gamma$ | (Micro)trajectory |
| $\mu$ | Probability Measure |
| $\Gamma_M$ | Macrostate-region |
| M | Macrostate |
| $S_B$ | Boltzmann Entropy |
| GSM | Gibbsian Statistical Mechanics |
| $\rho$ | Ensemble Density |
| $\rho_\mu$ | Microcanonical Density |
| $\Gamma_\rho$ | Ensemble-Density Region |
| $S_G$ | Gibbs Entropy |
| $\rho$ | Coarse-Grained Density |
| $\omega_i$ | Gibbsian Cells |
| MEP | Maximum Entropy Principle |
| $SP_B$ | Boltzmann Static Probabilities |
| $SP_G$ | Gibbsian Static Probabilities |
| DP | Dynamic Probabilities |
| $P_{IT}$ | Information Probabilities |
| OI | Ordinary Information |
| IT | Information Theory |
| $s_i$ | Symbol |
| m | Sequence of symbols, message |
| $\Lambda$ | Message space |
| S, D | Source, Destination |
| $H_{IT}$ | Information Entropy |
| $I_{IT}$ | Message Information |
| AP | Averaging Principle |
| AIT | Algorithmic Information Theory |
| K | Algorithmic Information |
| $I_S$ | Semantic Information |
| $\varphi$ | Epistemic State |
| SIT | *Science and Information Theory* |

- - - - - - - - -

# *Thesis Overview*

The main objective of this dissertation is to philosophically assess how the use of informational concepts in the field of classical thermostatistical physics has historically evolved from the late 1940s to the present day. I will first analyze in depth the main notions that form the conceptual basis on which 'informational physics' historically unfolded, encompassing (i) different entropy, probability and information notions, (ii) their multiple interpretative variations, and (iii) the formal, numerical and semantic-interpretative relationships among them. In the following, I will assess the history of informational thermophysics during the second half of the twentieth century. Firstly, I analyse the intellectual factors that gave rise to this current in the late forties (i.e., popularization of Shannon's theory, interest in a naturalized epistemology of science, etc.), then study its consolidation in the Brillouinian and Jaynesian programs, and finally claim how Carnap (1977) and his disciples tried to criticize this tendency within the scientific community.

Then, I evaluate how informational physics became a predominant intellectual current in the scientific community in the nineties, made possible by the convergence of Jaynesianism and Brillouinism in proposals such as that of Tribus and McIrvine (1971) or Bekenstein (1973) and the application of algorithmic information theory into the thermophysical domain. As a sign of its radicality at this historical stage, I explore the main proposals to include information as part of our physical reality, such as Wheeler's (1990), Stonier's (1990) or Landauer's (1991), detailing the main philosophical arguments (e.g., Timpson, 2013; Lombardi et al. 2016a) against those inflationary attitudes towards information. Following this historical assessment, I systematically analyze whether the descriptive exploitation of informational concepts has historically contributed to providing us with knowledge of thermophysical reality via (i) explaining thermal processes such as equilibrium approximation, (ii) advantageously predicting thermal phenomena, or (iii) enabling understanding of thermal property such as thermodynamic entropy. I argue that these epistemic shortcomings would make it impossible to draw ontological conclusions in a justified way about the physical nature of information. In conclusion, I will argue that the historical exploitation of informational concepts has not contributed significantly to the epistemic progress of thermophysics. This would lead to characterize informational proposals as 'degenerate science' (à la Lakatos 1978a) regarding classical thermostatistical physics or as theoretically underdeveloped regarding the study of the cognitive dynamics of scientists in this physical domain.

# Chapter I

# Introduction

Throughout the second half of the twentieth century, certain concepts acquired an indisputable prominence within modern physics, like ‘quarks’, ‘black holes’, or ‘strings’. Among them, one in particular stand out for its indisputable link with the accelerated development not of the physical sciences themselves but of telecommunication and computation; this is undoubtedly the concept (or concepts) of 'information'. There is no doubt that informational notions (i.e., theoretical, technical, or even in its everyday sense) have in recent decades acquired an overflowing presence within a wide variety of social and natural sciences, being especially relevant within evolutionary (Avery 2012) and molecular biology (Yockey 1992). Of course, apart from biology, informational concepts have also conquered virtually all fundamental domains of modern physics, from general relativity (Hosoya and Fujii 2018) to quantum mechanics (Timpson 2013). But this historical emergence to the proliferation of uses of informational notions within physics today does not occur spontaneously and uncoordinated, but rather was intellectually articulated in multiple 'informationalist' streams of physical thought with different scientific motivations and theoretical goals. All these intellectual trends and research programs are encompassed in what has been referred to within the literature as 'informational physics' or 'information physics' (e.g., Maroney and Timpson 2018), which we can define as that interdisciplinary domain of physics in which informational concepts play a central role in the investigation of our physical reality.

## 1.1. Preliminary Considerations and Motivations

Ultimately, the research work that I intend to carry out in the following constitutes a systematic philosophical analysis of informational physics as one of the most important intellectual trends (or even 'research program', in Lakatos' [1978a]) terms that have occurred in the framework of the physical sciences in the last decades. This importance can be expressed in terms of the overwhelming amount of academic output generated (e.g., number of papers, books, conferences, etc.) and the enormous impact of this output within the physics community of the second half of the twentieth century, as we shall see in detail in the following chapters. Despite its decisive weight in the recent development of physics, informational physics as a general theoretical research framework has been deeply neglected in the field of philosophy. On the contrary, within the literature concerning the philosophy of physics various associated local issues or problems of this intellectual movement have been treated, as is the pioneering case of Rudolf Carnap (1977) on informational interpretations of entropy, John Norton (2005, 2011) as well as Meir Hemmo and Orly Shenker (2012, 2019) on the thermophysics of computation, or Christopher Timpson (2013) and Olimpia Lombardi (et al. 2016a) on the physical nature of information. Unlike these previous philosophical analyses, in this research I aim to evaluate these questions not locally, but as part of a comprehensive intellectual movement, historically unfolding over more than seven decades and articulated around certain standardized uses of informational concepts, disciplinary claims, and research strategies. Ultimately, the main motivation for carrying out this task is to achieve a better philosophical understanding of this broad trend of physical thought in its multiple intellectual facets and from a systematic perspective.

## 1.2. Philosophical Objectives

Having stated the initial motivation, I should now state more concretely which are the particular objectives of this analysis. To do so, I must now clarify in what specific sense I intend to provide a 'better understanding' of this current of thought or research program in contemporary physics. Preliminarily, it is tempting to frame informational physics from within philosophy of science as a historical succession of physical theories that share certain common features, regardless of whether I analyse (or even identify) these theories by means of axiomatic sets of logical formulas, set-theoretic structures or even models in a broad sense. Although this may be the case, I will reject this approach for mainly heuristic reasons. First of all, during this evaluation I will constantly analyse scientific products whose contents are subject to rapid historical evolution; for this reason, it would be extremely difficult and highly useless to try to individualize each 'physical-informational' product into distinguishable theories. For this reason, I will prefer to speak of 'theoretical proposals' (rather than 'theories') to refer to a particular scientific-intellectual product of an author or community of authors. Illustratively, Léon Brillouin published his main papers in 1949 and 1951 and a key book in 1956. However, it would be inadequate to claim that (i) his 1949 paper expresses the same (only embryonic) 'physical theory' as that developed in his 1956 book, since both are articulated on significantly different informational notions, or (ii) that his 1949 paper expresses a different 'physical theory' than that of 1956 for this very reason, since underlying both are the same disciplinary goals, i.e., to include the observer within the descriptions generated from classical thermophysics.

Second, the fine-grained notion of 'theoretical proposal' allows us to extend the precision of this philosophical analysis also to the intellectual pretensions and disciplinary motivations underlying each scientific-analytical product I analyse, as opposed to the classical coarse-grained notion of 'theory' restricted to the content (rationally reconstructed or not) of such products. As I will show below, such intellectual pretensions and disciplinary motivations are not mere socio-scientific ornaments to the analysis but serve the explanatory function of allowing us to give a satisfactory account of the emergence, predominance or decline of certain conceptual usages in a certain historical period. Therefore, the analysis of informational physics as a scientific current does not focus on the dynamics of its different theories, but (and here lies the differential value of this work) on the particular manner in which information concepts are exploited descriptively from different theoretical proposals to generate putative knowledge. In this direction, my main intellectual objective with this philosophical thesis on information physics will be (i) to shed light on the evolution of the main uses of informational concepts in physics from the fifties to the present day, and (ii) to evaluate whether these uses along different theoretical proposals have contributed to the progress of physics in obtaining knowledge about our reality.

## 1.3. Historical-Philosophical Methodological Framework

Having stated the main motivation and central objectives, I must now point out the main methodological guidelines. In this sense, the general methodological framework that I will employ as the basis of this project is what is known as 'HPS' (or history and philosophy of science) applied to informational thermophysics as a stream of physical thought, combining a philosophical evaluation of scientific concepts with a historical analysis of how they evolve. In particular, I will adopt an HPS methodology close to what Hasok Chang (2004, p.237-241) called 'complementary science' (also Torretti 2006), as opposed to other HPS methodologies such as Kuhnian-Lakatosian, Latour's sociological-constructivist, etc. Following this methodological proposal, my historical-philosophical study of informational thermophysics as a complementary science starts from a model of close interaction between the results obtained from the (i) historical analysis, and (ii) philosophical evaluation, so that this allows to effectively contribute to the (iii) expert knowledge of this scientific domain. It implies three key consequences for this methodological framework.

First, my historical analysis of this intellectual movement in science is not merely historiographical, i.e., extracting illustrative case studies on which to subsequently develop the philosophical evaluation. The historical analysis that I will develop already possesses an intrinsically philosophical character, required for the following tasks: the individualization of conceptual elements, the specification of the semantic content of these concepts, the standardization of descriptive uses or explanations, etc. Interestingly, in this project I propose to include the main philosophical results obtained on some element of this current of physical thought as a significant part of its historical unfolding. Secondly, my philosophical evaluation of informational thermophysics is intrinsically dependent on the historical deployment of this current of thought, allowing us to trace the subtle weft of influence between theoretical proposals and the progressive evolution of the main conceptual uses under such scientific products, along the lines of works such as those of Arabatzis (2006) or Nersessian (2008). Therefore, the fact that this historical analysis is philosophical (Chapters III and IV) and our philosophical evaluation is historicist (Chapters II and V) allows to unfold our historical-philosophical argument in an integral way, distinguishing both parts in this work only for dialectical reasons. Finally, with my historical-philosophical analysis as a 'complement' (Chang 2004, p.237-241) of the scientific practices carried out in classical thermophysics I modestly intend to contribute to generate critical knowledge on whether the descriptive exploitation of informational concepts has contributed to the scientific progress of this physical discipline.

## 1.4. Scope Restrictions

Before proceeding with the plan and having made explicit my academic motivations, intellectual objectives and HPS methodology, I must point out certain restrictions in the thematic scope that I assume to carry out this project. First, I have previously argued that the objective is focused on the historical-philosophical analysis of 'informational physics' as a current of scientific thought characterized by the application of informational concepts in the field of physics. But as I will try to show below, informational physics encompasses practically all branches of fundamental physics, including quantum mechanics, statistical thermophysics and general relativity. In the literature one can find a systematic philosophical treatment of 'quantum informational physics' by Timpson (2013), but, as I noted previously, there are no systematic conceptual studies of other disciplinary facets of this intellectual current. With the aim of concentrating my argument towards a specific domain and due to its historical importance, in this project I will focus essentially (although sometimes the thematic boundaries are fuzzy) on 'classical informational thermophysics', delimited by the application of informational concepts in classical statistical mechanics, thermodynamics and surrounding fields. Therefore, the range of validity of my historical-philosophical arguments and results will be limited to this domain.

On the other hand, I restrict my analysis to the use of directly informational concepts in this physical domain, thus avoiding the immediate inclusion of computational notions that can be interpreted in terms of informational processing. This leads us to avoid questions related to what is known as Landauer's principle, whose impact in the scientific and philosophical literature (e.g., Norton 2005, 2011a, 2013, 2018; Ladyman et al. 2007, 2013, 2014; Shenker 2020) in the last decades has grown exponentially. Although the current of thought linked to this idea (i.e., that denominated by Charles Bennett [1982] as 'thermodynamics of computation') emerged historically from the informational thermophysics that I will analyse here, because its treatment would imply (at least) doubling the extension of this dissertation, I strongly feel the need to postpone its historical-philosophical evaluation. This thermo-computational current was particularly important from the 1980s onwards, when "computationalism" as an intellectual trend proliferated within the natural sciences, converging synergistically (sometimes indistinctly) with informational research programs. Of course, some of the concepts (e.g., algorithmic information) or intellectual positions (e.g., Lloyd's [2000] pancomputationalism) that the reader will encounter throughout my assessment undeniably possess a certain computational character; however, they will either be informational notions or such positions will be motivated by informationalist pretensions.

## 1.5. Plan for this Dissertation

Finally, the plan for this dissertation is as follows:

The following Chapter II is devoted to delimiting in depth the main concepts that will needed as a basis to unfold the analysis of informational physics. This 'conceptual landscape' will be made up of the multiple notions of entropy, probability and information that nourish the main theoretical proposals that make up this current of scientific thought, as well as the main interpretative axes that allow us to delimit its content. To close this chapter, I will present by means of atlases the main results found in the literature (both scientific and philosophical) on the syntactic, numerical, and semantic relations existing between these concepts.

Chapters III and IV will be exhaustively devoted to a philosophical evaluation of the historical unfolding of information physics. First, in Chapter III I will analyse the historical origin of informational thermophysics in the late 1940s and early 1950s thanks to the interdisciplinary popularization of Shannon's communication theory. From this scenario, I will delimit the two main traditions that emerged from this context. On the one hand, Brillouinism sets out a complementary interpretation between entropy and information; on the other hand, Jaynesianism proposes an information-theoretic framework of probability assignment in statistical mechanics, where entropy and information are identified. Finally, I will explore the main criticisms of the conceptual foundations of these programs raised by Carnap (1977) and his intellectual milieu.

Second, in Chapter IV I assess how informational physics became a predominant current of thought in fundamental physics in the 1990s. Initially I point out that this was made possible by the convergence of Jaynesianism and Brillouinism in proposals such as that of Tribus or Bekenstein (1973) and the application of algorithmic information theory in thermophysics during the 1980s. As a sign of the radicality that this program acquired at the end of the past century, I will explore the main metaphysical frameworks that granted physical nature to information, defended among others by Wheeler (1990), Stonier (1990) or Landauer (1991). Finally, I will analyse the most prominent philosophical arguments (e.g., Timpson, 2013; Lombardi et al. 2016a) in the literature against this kind of inflationary ontological stances regarding information.

Next, I will devote Chapter V to a systematic analysis of the conceptual foundations underlying this current of scientific thought. To this end, I will carefully analyse whether the descriptive exploitation of informational concepts has historically contributed to provide powerful explanation and advantageous prediction of thermal phenomena (e.g., the free expansion of a gas) or to the understanding of a given thermal property (e.g., thermodynamic entropy). Finally, I conclude by defending that the epistemic shortcomings of the main thermo-informational proposals preclude drawing theoretically legitimized ontological conclusions about the physical nature of information.

In the final Chapter VI, I conclude by defending that the historical unfolding of informational thermophysics constitutes a historical process of intellectual inflation, wherein the exploitation of informational concepts since the 1950s has not entailed an effective epistemic progress with respect to thermophysics. In Lakatosian terms (1978a), this lack of progress in the attainment of thermophysical knowledge allows us to characterize the exploitation of informational concepts as a "degenerating scientific program" regarding classical statistical physics. In this sense, I argue that informational program may be theoretically viable in the future as a sort of naturalized epistemology within SM, although much remains to be done in this direction.

# Chapter II

## Conceptual Elements of Information Physics

In order to understand philosophically a current of scientific thought such as information physics, it is essential to account for the main concepts on which its descriptive or theoretical practices are articulated in general. In this direction, the main objective of this chapter will be to carry out an integral analysis of the central concepts on which the physics of information is intellectually based. By 'integral analysis' I mean the evaluation of the syntactic, semantic and pragmatic dimensions of the notions in question. First, due to the formal or mathematically precise definition of most of these concepts, performing an analysis of their strictly syntactic dimension is indispensable for subsequently delimiting their semantic capacity and the pragmatic exploitation of this capacity. As far as this capacity is concerned, philosophical analysis can successfully contribute to specify (i) how the concepts refer (i.e., their representational resources or 'intension'), (ii) what these refer to (i.e., what they refer to or 'extension'), and (iii) how these both can be interpretively modified "A task of philosophy is to expose the cases of meaning reorganization when they are occurring, in order to avoid confusions and promote conceptual clarification" (Lombardi et al. 2016b, 2007). Finally, it is decisive for a deep understanding of these conceptual elements to evaluate under which theoretical objectives and how these are used by epistemic agents in diverse contexts of disciplinary practices. But before carrying out this task, it is necessary to make explicit the conceptual evaluation framework on which I will base my analysis.

### 2.1. A Minimalist Framework for Conceptual Assessment

The systematic assessment of scientific concepts has been one of the central activities for the whole of modern philosophy of science, from the Vienna circle in the 1920s to the present day. At the same time, the rigorous analysis of the semantics and pragmatics of general concepts has also been one of the themes articulating the history of analytic philosophy. Recently, philosophers of scientific concepts such as Nancy Nersessian (2008) (as Carnap [1945] did much earlier) have defended through the so-called 'continuity thesis' that these are not a specific type of concepts with idiosyncratic properties but constitute the result of the technically refined evolution of certain ordinary concepts. Others like Chang (2004) have shown the historical character of the continuity between non-technical notions to technically-defined scientific concepts, as was the transition between certain everyday intuitions about heat and the thermodynamic notion of 'Kelvin temperature'. We subscribe to both ideas as a starting point. In the field of information physics this continuity ('diachronic' or 'synchronic') between ordinary concepts and scientific technical concepts is manifest, as I intend to show below. Illustratively, the scientific and technically-defined concept of 'negentropy' (Brillouin 1956) (central to this stream of scientific thought) is no more than the progressive result of a technical refinement of the everyday concepts of 'order' and 'information' during the period 1865-1956.

At this point it should be noted that this methodological proposal regarding the analysis of ordinary or technical concepts of information physics will be minimalist. This 'minimalism' should properly be understood as a minimal commitment with respect to how (i) scientific theories (syntactic, semantic, pragmatic view, etc.), (ii) scientific and mathematical models (structuralist, formalist, etc.), (iii) scientific concepts, and (iv) the relationship between these previous elements (atomism, holism, etc.) should be philosophically characterized. The justification for this minimalist stance lies precisely in the fact that my aim in this dissertation is not at all to defend any of these philosophical positions or to show their compatibility with certain historical narratives. Although minimalist, I will assume as a starting point a semantic-pragmatic characterization of conceptual elements as representational vehicles whose content is susceptible to inferential exploitation, regardless of their format (symbolic, sentential, diagrammatic, etc.) or their plausible relation to epistemic agents (mental representations, agent-independent, etc.). Note that this semantic-representational (e.g., Nersessian 2008) as well as pragmatic-inferential (e.g., Wilson 2006) characterization of concepts in scientific settings allows us to deal with a vast number of different descriptive practices without trivializing the notion of concept or disregarding its historical development (e.g., Arabatzis 2006). This initial position will be methodologically useful (and therefore pragmatically justified) in this case due to the high degree of fine-grained conceptual analysis required by the subtle interpretative modifications of concepts throughout the historical evolution of this school of thought. In this sense, I will also remain neutral with respect to the debate on the preservation or elimination of the semantic content of scientific concepts throughout history (see Arabatzis 2006), or with respect to their use in other theoretical domains.

Having clarified this point, let us set the stage for the following sections. First, I explore what Gilbert Ryle (1949) called 'logical geography' (to avoid logicist connotations I will use instead the term 'conceptual landscape') in the different scientific domains over which information physics will emerge, as Frigg (2008) did in the particular case of statistical mechanics. Following Wilson (2006), one should not fall into the error of conceiving conceptual landscapes as static quasi-geological entities in which the meaning of their local concepts and their possible interconnections are fixed forever[1]. On the contrary, they are organic and dynamic entities in which their multiple

[1] "The attitude in question I call tropospheric complacency—it represents our native inclination to picture the distribution of properties everywhere across the multifarious universe as if they represented simple transfers of what we experience while roaming the comfortable confines of a temperate and pleasantly illuminated terrestrial crust. In

representational resources (and therefore the meaning of their concepts) are subject to constant reorganization and applicative extension according to the interpretative tactics of their epistemic users. In this chapter I will not aim to present photographs of conceptual landscapes, which would be complacent but philosophically fruitless, but 'maps' that encode their key locations, endowing interpretive coordinates (e.g., subjective-objective, epistemic-ontological, inferential-dynamic [Wallace 2017], etc.) to the reader to guide their intuitions in these convulsive landscapes. In this direction, I aim to map the conceptual landscapes linked to the notions of 'entropy' and 'information', as master pillars of this intellectual trend, as well as to that of 'probability' for its indispensable role in this. Finally, I will combine these maps into 'atlases' to trace the main conceptual connections we find between the previously explored landscapes, which will be indispensable for the subsequent analysis of the historical unfolding of information physics. Let us begin by delving into the rugged landscapes of entropy.

## 2.2. Landscape of Entropy Concepts

The concept of entropy is undoubtedly one of the most ubiquitous scientific concepts in contemporary physics. This does not mean that its meaning is one of the easiest to define and even to understand satisfactorily, not only for the non-scientific population but also within the scientific community. It is said that even John von Neumann, one of the most important physicists and mathematicians of the twentieth century, went so far as to state that "no one knows what entropy really is" (von Neumann, quoted on Tribus and McIrvine 1971, p.180). Even the term 'entropy' is not always used in reference to the same concept. On the contrary, only in the field of classical statistical physics (to which we restrict ourselves here) can we find a rich manifold of entropic concepts, probabilistic and non-probabilistic, with different quantitative properties and whose different ways of interpretatively characterizing their meaning sometimes diverge. Let us now proceed to explore the main paths within this conceptual landscape.

### 2.2.1. Thermodynamic Entropy

The original notion of entropy arises in the field of thermodynamics. Thermodynamics (TD) is the physical discipline devoted to the description of the behavior of macroscopic substances (without considering their microphysical composition) using certain observable quantities, such as volume, pressure or temperature. This small set of macroscopic variables or 'state variables' serves to fully characterize at any given moment the state of thermodynamic systems, such as a thermal machine that exchanges heat Q and work W with its surrounding. The precise relationship between volume V, pressure p and temperature T was historically specified by the empirical laws of Boyle, Charles, and Gay-Lussac, all of them integrated in the so-called equations of state, the most famous being

such a vein, we readily fancy that we already ‘‘know what it is like’’ to be red or solid or icy everywhere, even in alien circumstances subject to violent gravitational tides or unimaginable temperatures” (Wilson 2006, p.55).

the ideal gas law pV = $n$RT for $n$ moles of substance and R being the gas constant. One of the main thermodynamic concepts is that of the thermal equilibrium state, where the value of these macroscopic variables does not change over time. We call 'thermodynamic process' the change of state of a system from $\Omega_A$ to $\Omega_B$, represented by a 'path' in the system's state space Ω between these two points. When $\Omega_A = \Omega_B$, this process constitutes a TD-cycle. If the process from $\Omega_A$ to $\Omega_B$ can be reversed by infinitesimal changes, then this will be 'thermodynamically reversible'[2], otherwise it will be 'irreversible.'. When all intermediate states between $\Omega_A$ and $\Omega_B$ are equilibrium states the process is called 'quasi-static', as is assumed to be the case with delicate heat exchanges between the system and the environment.

One of the pillars of thermodynamics is calorimetry or the study of state changes by adding or subtracting heat Q in the system (Chang 2004), this quantity (regardless of its ontological characterization, i.e., 'caloric') being not a state variable but a 'process variable' specified by differential calculus, just like work W. In fact, during the 1840s scientists like Joule or Mayer (Uffink 2007, p.934) demonstrated the equivalence between Q and W as two distinct forms of energy, so that for all cycles of the equilibrium space state there would exist an 'energy' function U where dU = Q + W and its value is conserved. This constitutes the First Law of thermodynamics. As for the famous Second Law, its content goes back at least to Sadi Carnot in 1824, who specified the efficiency η of certain thermal machines as a function of two temperatures T-, T+ and whose value is the same for all quasi-static cycles. As Chang (2004) argued, this efficiency function was key to Kelvin's derivation of a concept of absolute temperature T. This 'Carnot theorem' originally depended on assuming the conservation of heat, which contradicted the First Law. Subsequently Lord Kelvin (1851) made Carnot's theorem compatible with the First Law, asserting the impossibility of conceiving of a cyclic process in which heat is extracted from an external reservoir and equivalent work is done. This is Kelvin's formulation of the Second Law of thermodynamics. Clausius (1854) formulated this same idea in such a way that heat transfer Q from low-temperature $T_0$ system to a system with high temperature $T_1$ was impossible, from which the following inequality can be defined:

$$\oint_B^A \frac{\mathrm{Q}}{\mathrm{T}} \leq 0 \tag{1}$$

In this same article, Clausius (1854) initially postulated the existence of a 'mechanical equivalent of heat' Q/T with which he subsequently (Clausius 1856) defined the notion (precursor of the concept of 'thermodynamic entropy') of 'equivalence-value' N for any line-integral between any two equilibrium states $\Omega_A$ and $\Omega_B$:

$$\int_{\Omega_A}^{\Omega_B} \frac{\mathrm{Q}}{\mathrm{T}} = -\mathrm{N} \tag{2}$$

[2] Norton (2016) argued that the concept of 'thermodynamically reversible process' is inconsistently based on simultaneously asserting contradictory properties about the system, namely: that the system does not change (because it is in equilibrium) and at the same time it changes (since every process by definition is a change of state). Therefore, this would constitute a non-real process.

In 1865, Clausius used (1) to derive by differential calculus the existence of a state function 'S' over the equilibrium states A and B, whose value is independent of the possible paths between the two states. This is expressed for reversible cycles as follows:

$$S_{TD}(\Omega_{\mathrm{B}}) - S_{TD}(\Omega_{\mathrm{A}}) = \int_{\Omega_A}^{\Omega_B} \frac{Q}{T} \tag{3}$$

This is precisely the original definition of thermodynamic entropy[3]. Although (2) allows us to express the state function $S_{TD}$ as entropy difference (3) from state $\Omega_A$ to state $\Omega_B$, it is possible to define absolute entropy values for each particular state $S_{TD}(\Omega_{A}\mathrm{i})$ in the case of reversible processes (Frigg and Werndl 201, p.116). In this way, the entropy state function also allows us to formulate the fundamental equation of thermodynamics (i.e., dU = TdS - pdV) that defines its interrelations with the other state variables. Clausius (1965) showed that expression (3) becomes an inequality in the case of irreversible cycles, where $\Omega_A$ and $\Omega_B$ can be connected by quasi-static and possibly (see Uffink 2007, p.937) non-quasistatic processes:

$$\mathrm{S}_{TD}(\Omega_B) - \mathrm{S}_{TD}(\Omega_A) \leq \int_{\Omega_A}^{\Omega_B} \frac{Q}{T} \tag{4}$$

In the case where the non-quasistatic processes from $\Omega_A$ to $\Omega_B$ temperature is constant (i.e., 'adiathermal process') or no heat is transferred (i.e., 'adiabatic process') then it is possible to show that $S_{TD}(\Omega_B) \geq S_{TD}(\Omega_A)$. Illustratively, (4) allows us to compute the entropy difference generated during the Joule expansion or free expansion of an ideal gas in a vessel with volume 2V from a subvolume V, i.e., *n*RIn2 in Joules per Kelvin units per *n* moles of substance[4]. Therefore, the concept of entropy, "a real (indirectly) observable property of individual systems" (see Callender 1999, p.350), allows us to elegantly formulate the Second Law in such a way that in adiabatic processes such as a Joule expansion the entropy of the final state $S_{TD}$(A) cannot be less than the entropy of the initial state $S_{TD}$(B). As for the choice of the name 'entropy', Clausius (1865) intended to give physical significance to his concept by means of its semantic connection with energy: “I propose that *S* be taken from the Greek words, `en-tropie' [intrinsic change]. I have deliberately chosen the word entropy to be as similar as possible to the word energy: the two quantities to be named by these words are so closely related in physical significance that a certain similarity in their names appears to be appropriate.” (Clausius 1865).

---

[3] Other closely-related definitions of $S_{TD}$ within the philosophy of TD are the following: “The Second Law states that an extensive state function, the total differential δS = δQ/T, defined only for equilibrium states, is such that ΔS ≥ ∫δQ/T. Loosely put, for realistic systems, this implies that in the spontaneous evolution of a thermally closed system, the entropy can never decrease and that it attains its maximum value for states at equilibrium” (Callender 1999, p.350) and “there exists a function, called entropy S, on the equilibrium states of the system such that (14) $S(s_1) - S(s_2) = \int \delta Q/T$ or, as it more usually known (15) δQ/T = dS. This result is frequently expressed as follows: δQ has an integrating divisor (namely T): division by T turns the inexact (incomplete, non-integrable) differential δQ into an exact (complete, integrable) differential. (…) entropy is defined only for equilibrium states. Therefore it is meaningless within this theory to ask how the entropy of a system changes during a non-quasistatic process” (Uffink 2007, p.936).

[4] The entropy production ΔS in a Joule expansion from V to 2V of an ideal gas can be computed from the fundamental equation of thermodynamics as follows $\mathrm{dS} = \int_{\mathrm{V}}^{2\mathrm{V}} \frac{\mathrm{PdV}}{\mathrm{V}} = \int_{\mathrm{V}}^{2\mathrm{V}} \frac{n\mathrm{RdV}}{\mathrm{V}} = n\mathrm{R} \ln 2$ Joules/Kelvin.

It should be noted that Clausius was one of the main proponents of gas kinetics and advocate of the molecular hypothesis during the 1860s. This is the main reason why this author suggested to understand kinetically the meaning of his concept 'mechanical equivalent of heat' (Clausius 1854) and later 'entropy' (Clausius 1865) by a sort of spatial disaggregation of the molecular components in gaseous substances. However, as Frigg and Werndl (2011, p.117) pointed out, all the kinetic-molecular specifics of the meaning of TD-entropy are irrelevant because this is a strictly phenomenological concept, unlike the notions of entropy we see below. Exploiting its conceptual connection (embedded in its nomenclature, as we see in Clausius' quote above) with the TD-definable concept of energy, some authors such as Wilson (2017) have advocated a strictly phenomenological interpretative strategy for understanding entropy as a way of quantifying the amount of useless energy in the system "The core of thermodynamic thinking (…) invokes the subtle notion of entropy to codify the notion of coherent work capacity lost and regained. (…) Entropy, especially, provides this assistance by numerically quantifying the degradation in energetic coherence" (Wilson 2017, p.143, 145).

### 2.2.2. Boltzmannian Statistical Mechanical Entropy

It was from the 1860s onwards when a significant part of the German-speaking physics community (including Clausius himself) sought to develop purely mechanical descriptions of the macroscopic behavior of material substances not from their macroscopic observable variables $A_i$ (as in TD), but from the dynamic variables of their (then hypothetical) microscopic molecular components. It was the young Austrian mathematician Ludwig Boltzmann himself who first sought to reconceptualize (or even 'reductively explain') the notion of $S_{TD}$ in terms of Hamiltonian analytical mechanics, only seven years after its baptism by Clausius. The development of a strictly mechanical and mathematically sophisticated concept of entropy by Boltzmann (and his immediate followers) will extend over a convulsive intellectual period of more than four decades, marked by his staunch defense of atomism against energeticists such as Ostwald and by refuting the main criticisms of his proposal (Uffink 2007, 2014). This gave rise to not one but several concepts of entropy that we proceed to assess below.

#### *Boltzmann's H-Quantity*

In order to describe the evolution of an ideal gas, Boltzmann (1972) originally employed a distribution function $f(p, q, t)$ over the Cartesian product of all the 6-dimensional phase spaces of individual molecules (or 'μ-space') to specify the relative number of molecules of the kinetic-gas system (conceived as a mathematical hard-sphere colliding model) whose position $p$ and momentum values $q$ lies in an infinitesimal interval dμ at time t. In order to analyse how the function $f$ evolves over time (which is specified by his homonymous equation) Boltzmann originally introduces the famous H-functional:

$$H_B(f) = \int f(d\mu) \log f(d\mu)\, d\mu \qquad (5)$$

Based on the hypothesis of molecular chaos (named 'Stosszahlansatz' by Paul and Tatiana Ehrenfest [1911]), i.e., each pair of molecules becomes statistically correlated when they collide, Boltzmann (1872) derived via his homonymous kinetic equation the H-theorem as a properly mechanical counterpart of the Second Law of TD[5]. This theorem states that the H-quantity (5) decreases monotonically in the system's evolution until it reaches its minimum values at thermal equilibrium, where *f* then corresponds to the Maxwell-Boltzmann distribution. From the H-quantity (5) one could define a 'fine-grained' Boltzmann entropy $S_H$, which is negatively proportional to this first quantity multiplied by the so-called Boltzmann constant $k_B$: $S_H(f) = -k_B\, nH_B(f)$ (see Frigg and Werndl, 2011, p.123).

### *Boltzmann's Combinatorial Entropy*

Due to the famous reversibility objections to this 1872 proposal (Uffink 2007, Section 4.3), Boltzmann developed in 1877 a new theoretical framework from which to model the behavior of molecular systems, paradigmatically gases with a high number n of identical components confined to the volume V of a container and with constant energy E. As in the previous proposal, this was based on considering the Cartesian product of the phase space of individual molecules. As in the previous proposal, this one was based on considering the Cartesian product of the phase space ('*phasenraum*', in Boltzmann's terms) of individual molecules. However, as a novelty, a partition[6] of $\Gamma_{\mu i}$ into rectangular cells $\omega_i$ of equal Lebesgue-measure $\mu(\omega_i)$ is introduced. In this theoretical framework, the μ-microstate of the system or simply 'arrangement' corresponds to the specification of the location of each particle in cells, just as the values of the macroscopic properties $A_i$ (e.g., volume, pressure, temperature, etc.) of the system are represented by the 'distribution' D specifying the number of particles n! located in each cell $\omega_i$. It is in his 1877 where Boltzmann formulates what is known as the 'combinatorial argument', in which different arrangements of the molecules give rise to, or are compatible with, to the same distribution, where the number of compatible

[5] There is controversy in the literature as to whether Boltzmann really intended to argue that the Second Law of TD necessarily followed from his H-theorem of his kinetics of gases, as is generally assumed (Klein, 1972, p.73) or whether he considered notable exceptions (von Plato, 1994, p.81).

[6] The partition of phase space into non-overlapping regions was originally justified as early as 1877 by Boltzmann based on the impossibility of epistemically accessing individual microstates *x*, either by microscopic observation (it is not possible to directly or indirectly observe-measure the position and velocity of $10^{24}$ molecules) or by mathematical specification of such values. As assumed within the literature (McCoy, 2020; Shenker, 2020), the individual microstates of realistic systems ($n = 10^{24}$) are considered epistemically inaccessible precisely because of the practical impossibility of determining a microstate since "requires one to know the initial positions and velocities of all n particles and to follow these motions for all time. Since n is typically of the order of $10^{24}$, this is of course impossible" (Ellis, 2006, p.65). In addition to this pragmatic impossibility, the determination of the Hamiltonian evolution of each individual microstate (specified by the famous Boltzmann kinetic equation) becomes a much more mathematically complex and pragmatically impossible task.

arrangements is determined by $G(D) = n! / n_1!....n_i!$ If we perform a Stirling approximation of this factorial expression, we obtain what could be called 'combinatorial entropy':

$$S_B = k_B \log G(D) \tag{6}$$

wherein $k_B$ is Boltzmann’s constant. On the basis of this argument, Boltzmann (1877) justifies the idea that the distribution with the highest entropy (6) corresponds to that distribution with the highest number of compatible arrangements, assuming that this is the distribution corresponding to thermal equilibrium. His defense that molecular systems evolve from low-entropic to high-entropic distributions was underlain by certain theoretical assumptions not yet explicit in this theoretical framework, i.e., that individual molecules will eventually transition all regions in $\Gamma_\mu$.

*Boltzmann Entropy*

Two intellectual disciples of Boltzmann, Paul Ehrenfest and Tatyana Ehrenfest, attempted to systematize coherently in a 1911 article for the Encyclopedia of Mathematical Sciences their teacher’s theoretical proposal into what we now know as 'Boltzmannian statistical mechanics' (BSM). This is the most relevant Boltzmannian theoretical apparatus nowadays (e.g., Callender, 1999; Albert, 2000; Goldstein et al. 2020). The main conceptual advance they introduce with respect to the 1877 combinatorial framework will be the technically rigorous introduction of the concept of 'Γ-space' (the term 'phase space' first appeared[7] in 1918), defined as the 6*n*-dimensional state space in which the possible position and velocity values of not a single molecule, but of all the n molecules composing the system were encoded. This extremely abstract spatial concept made it possible to define the pivotal SM-concept of Γ-microstate (or simply 'microstate[8]') in which the position and velocity of all molecules at each time t are specified, being represented by individual points in the system’s phase space Γ.

To define the values of observable variables (i.e., volume, pressure, and temperature) by means of microstates, BSM depends theoretically on associating particular values of macrovariables $A_i$ with sets of microstates called 'macrostates' $M_i$. The relationship between macrostates and microstates is one-to-many, where several (observationally indistinguishable and equiprobable) microstates would correspond to the same macrostate as well as each individual microstate *x* can only be contained within a single macrostate $M_i$. Furthermore, within the literature (Frigg, 2008; Shenker, 2020) macrostates are considered to supervene on microstates, indicating that each change in

[7] An extensive history of the concept (and term) of ‘phase space’ was developed by David D. Nolte (2018), who claimed that “By this time [1911], the stigma of using the expression space for n-dimensions had disappeared, and so the Ehrenfests were comfortable using space to define Boltzmann’s n-dimensions. But they dispensed with the term phase, possibly because of its obscurity. Yet, at the very beginning of the encyclopedia article, in order to set the context for the definition of Γ-space, they referred back to Boltzmann’s usage of phase (…) This was not the first use of Phasenraum in print, which appears for the first time in a paper by Paul Hertz in 1910” (Nolte 2018, p.142)

[8] Interestingly, there are some formulations of statistical mechanics in which the microstates do not play any descriptive role. This is the case of the proposal by McCoy (2020), who defended observable variables encoded in probability distributions as the fundamental descriptive elements of this theory.

observable values associated with macrostates $M_i$ will necessarily imply a change in the location of the actual microstate *x* of the system. This implies that each particular macrostate $M_1$ is defined over a particular region of phase space $\Gamma_{M1}$ that will not overlap any other region $\Gamma_{M2}$ associated with a distinct macrostate $M_1$. Thus, a specific set of macrostates $M_i$ would constitute a proper partition of the phase space $\Gamma$ of the system if the (set-theoretical) union of their macrostate-regions $\Gamma_{Mi}$ covers the whole accessible region of its energy hypersurface $\Gamma_{Mi}$. Apart from associating observational descriptions with microstate-based descriptions, the macrostates allow us to reformulate the thermodynamic concept of entropy by means of statistical mechanical resources. With them we can formulate the Boltzmann entropy:

$$\mathrm{S_B} = k_{\mathrm{B}} \log\ (\Gamma_M) \tag{7}$$

Where $k_B$ is Boltzmann constant, $\mu$ is the probability measure[9] used to count the number (or volume, since $\Gamma$ is continuous in classical mechanics) of microstates associated with a macrostate, and $\Gamma_M$ is the macrostate-region in which the actual microstate *x* of the system is located. Thus, although Boltzmann entropy (or simply $S_B$) is defined in (7) by the number of microstates of the macrostate-region in which the actual microstate *x* of the system is located at a particular time, it is generally regarded as a property of individual systems. In BSM it is assumed that the thermal equilibrium state (that state in which the system's macroscopic observables do not vary in time) corresponds to the macrostate with the highest Boltzmann entropy, which will also be associated with the macrostate-region $\Gamma_{Meq}$ that occupies almost its entire energy hypersurface $\Gamma_E$ (see Chapter 2.4.1). With all these technical and theoretical resources, Boltzmannians SM-describe TD-like equilibration processes in such a way that the actual microstate *x* of the system transits through macrostate-regions $\Gamma_{Mi}$ of increasing volume until it reaches $\Gamma_{Meq}$. Of course, due to fluctuations, it is still possible that the microstate *x* will again transit out of the macrostate-region $\Gamma_{Meq}$ once it reaches it, eventually decreasing its entropy against the empirical predictions of thermodynamics.

### *Subjective vs. Objective Interpretations of Boltzmann's Entropy*

The Boltzmann entropy concept just detailed above has usually been characterized in the scientific literature as a 'subjective' or 'anthropomorphic' measure, along the lines of Penrose "There is still something very subjective about this definition of S" (Penrose 2006, p.691) or as we see in Jaynes' famous comment: "entropy is an anthropomorphic concept, not only in the well-known statistical sense that it measures the extent of human ignorance as to the microstate. (...) For it is a property, not of the physical system, but of the particular experiments you or I choose to perform on it" (Jaynes 1965, p.398). Denbigh and Denbigh (1985) carried out an extensive analysis of agent-dependent conceptions of entropy up to the 1980s. The main argument underlying this agent-dependent interpretative strategy with respect to Boltzmann entropy could be reconstructed as

[9] Although 'μ' has been previously used to refer to the state space of a particular molecule, I henceforth will use it for denoting a probability measure.

follows: the value of this notion depends on the particular way in which $\Gamma$ is partitioned into $\Gamma_{Mi}$, and because in the context of the SM-classical there is no objective-universal criterion with which to partition (e.g., like Planck constant in the quantum domain, see Frigg [2008, Section 2.7.]), then the meaningfulness of this concept will ultimately depend on conventional agent-dependent non-objective criteria. Against this subjectivist characterization, Denbigh and Denbigh (1985) argued that the objective character of entropy should be understood precisely through the intersubjective or interagency agreement of the scientific community that employs this concept. Following this line, Callender (1999) defended interpreting the concept of entropy as a non-subjective but 'relational' notion, because "The coarse graining used to define $S_B$ does not imply that $S_B$ is subjective, only that is relational (…) If a Laplacian intelligence informed you of the exact microstate of some macrostate, that would not affect the value of $S_B$ one jot. (…) *How many microstates correspond to a particular thermodynamic description* is still an objective matter, even if a system admits of more than one description" (Callender 1999, p.370-371, Italic are mine)

Another element that contributes to foster agent-dependent interpretations of Boltzmann entropy are the justifications by which the macrostatistical partitions of $\Gamma$ are conceptually legitimized. In this sense, the main justification throughout the history of SM (and in particular BSM) has been to interpretatively associate the measurement of the macrostates with the microscopic resolution of the agents, so that two different microstates in the same macrostate-region are observationally indistinguishable for these agents (Robertson 2020, Section 4). Hemmo and Shenker (2012, Chapter 5) detailed two conceptually different strategies of justifying Boltzmannian coarse-graining: one based on the subjective 'microstatistical indistinguishability' and another based on capturing TD-regularities. These authors defended the ability of the second strategy legitimizing macrostatistical partitions to conceptually account for certain macroscopic TD-regularities that emerge from this coarse-graining, enabling an inter-subjectively objective (although counterfactual) characterization of Boltzmann entropy. Their argument run as follows: that the measure of a macrostate can be satisfactorily interpreted as a certain degree of ignorance à la Jaynes (1965) does not explain why a gas will eventually expand over the entire volume of its container. However, Hemmo and Shenker (2012, p.96) argued that this subjective-epistemic interpretation of macrostates could also make an explanatory contribution in accounting for why macroscopic observers (i.e., scientists in SM-practices) are able to grasp certain macrostatistically-describable thermodynamic regularities and not others[10]. Then, Hemmo and Shenker defended that the way in which these partitions of the phase space are interpreted can be decisive (i) to understand their capacity to specify macrostatistical regularities, and therefore (ii) to objectively interpret Boltzmann's entropy. Illustratively, partitioning $\Gamma$ into Boltzmannian macrostates according to the

[10] "Since macrostates are sets of microstates that are indistinguishable by an observer, and since the actual state of the universe at any moment is a single microstate, it turns out that the statement that a system is in some macrostate must involve some sort of ignorance on the part of the observer, as to which of the microstates in the macrostate is actual. (…) this ignorance does explain why we happen to see these specific regularities – namely, the ones describable in terms of thermodynamic macrostates – rather than any other regularities that may exist in the world but cannot be perceived by human observers" (Hemmo and Shenker 2012, p.96).

mechanical meaning of volume V (i.e., dispersion of molecular positions), temperature T (i.e., average molecular kinetic energy) and pressure P (i.e., molecular collisions per area) would allow us to encode macroscopic TD-regularities like the ideal gas law' PV = $n$RT.

### 2.2.3. Gibbsian Statistical-Mechanical Entropy

The American Josiah Willard Gibbs carried out a theoretical proposal detached from Boltzmann's theoretical pretensions in explaining-reducing the content of TD to strictly classical-mechanical concepts (Frigg 2008, p.51) and based on the atomic hypothesis of matter. On the other hand, Gibbs' interests were directed fundamentally to provide theoretical coherence and systematicity to SM, whose work *Elementary Principles in Statistical Mechanics* (Gibbs 1902) is considered in the literature as the actual historical origin of this physical discipline (e.g., Uffink 2007, p.992).

#### *Gibbs's Fine-Grained Entropy*

One of the most notable theoretical differences between these two SM-frameworks is that, while BSM describes the behavior of individual n-molecular kinetic macroscopic systems, GSM depends descriptively on the so-called thermodynamical 'ensembles', usually conceived as a continuous (uncountably infinite) collection of virtual and independent copies of a single TD-physical system distributed over Γ at a given time t: "We may imagine a great number of systems of the same nature, but differing in the configurations and velocities which they have at a given instant, and differing not merely infinitesimally, but it may be so as to embrace every conceivable combination of configuration and velocities." (Gibbs 1902, p.v).

There are different types of ensembles depending on the parameters used to macrostatistically represent the physical system in question. Among the most important are the 'microcanonical ensemble' or 'NVE', used for collections of isolated systems that do not exchange any type of energy or matter. We also find the 'canonical ensemble' or 'NVT', which represents closed physical systems that interact by exchanging heat with their external environment, but do not exchange matter or other forms of energy with it. The particular state of each ensemble at each moment t is formally represented within GSM by a probability density function ρ over the phase space Γ of the system; wherein ρ (i.e., $\rho(x)$: $x \rightarrow$ r; r = 1 > r > 0) would encode either the probability that the actual microstate of the system is within the phase region $\Gamma_\rho$ (ontological interpretation of ρ) or even our belief about a microstate $x$ being the actual one (epistemic interpretation of ρ). As we see in Section 2.3.2, not only the density function ρ would depend on the particular ensemble we choose, but also our interpretation of these probabilities will be pivotal to specifying what ρ properly means in an ensemble description. It should be remarked that the microcanonical ensemble is represented by a uniform probability measure $\rho_\mu$. Also, according to GSM, the values of each particular observable variable $A_i$ of the ensemble are encoded by a functional $f$ [ρ], whose expected value 〈A〉 can be computed by averaging over the density ρ:

$$\langle A \rangle = \int_{\Gamma} f(x)(x,t)dx \tag{8}$$

This phase averaging is a key procedure within the GSM theoretical apparatus not just because its highly effective capability to compute macro-values (see Frigg and Werndl 2018), but also precisely because of its ability to conceptually connect the ensemble-based 'level of description' (in Lavis' [2005] terms) with individual-systems descriptions. One of the main assumptions of the GSM framework is that the 'statistical' equilibrium state (analogous to the TD-equilibrium state) of an ensemble corresponds to that distribution ρ which is stationary, because the macroscopic average values 〈$A_i$〉 do not change. On these technical and theoretical grounds, Gibbs (1902) defined his notion of SM-entropy as

$$S_G(\rho) = -k_B \int_{\Gamma} \rho \ \log(\rho) d \tag{9}$$

i.e., the integral over the volume of the phase region of phase space occupied by ρ, where ρ stands for the density representing the ensemble. Its description of thermodynamically irreversible processes such as equilibration presents certain significant theoretical problems. The source of the issue lies in the statistical mechanical fact that the total derivative of the density ρ with respect to time is zero, which implies that, following Liouville's theorem, that the phase volume of ρ will remain invariant during the evolution of the system Then, according to (8) the entropy of the ensemble of a freely expanding gas within a vessel will not increase during this irreversible process, but otherwise it will be conserved, contrary to TD-predictions. This is the so-called 'paradox of Gibbsian fine-grained entropy' (Callender 2011).

*Gibbs's Coarse-Grained Entropy*

As is well known, Gibbs (1902, Chapter 2) sought to solve this lack of predictive match between Gibbsian entropy and equilibration TD-processes by (i) partitioning phase space Γ into sets of microstates or 'cells' $\omega_i$ of equal Lebesgue-measured μ phase volume δω, and (i) introducing new descriptive resources by defining coarse-grained densities $\bar{\rho}$ from fine-grained densities ρ

$$\bar{\rho} = \frac{1}{\omega} \int_{\Gamma} \rho \, d \tag{10}$$

which are defined from a projection operator P*: ρ → $\bar{\rho}$, phase averaging the values of the fine-grained density ρ over the cells $\omega_i$ in which phase space Γ become partitioned. Recently Robertson (2020) employed Gibbsian coarse-graining within the Zwanzig-Zeh-Wallace framework to show how macroscopic TD-like time-irreversibility could be technically derived from microstates time-reversibility. Authors such as Ridderbos (2002, p.69) argue that the introduction of coarse-grained densities implies the generation of a concept different from that definable by fine-grained densities (or 'fine-grained GSM-equilibrium'), namely, a coarse-grained density will be in 'coarse-grained' equilibrium if it is homogeneously distributed over the cells as a result of the phase-average procedure. Having introduced these new SM-descriptive and technical resources, we can now define the concept of coarse-grained Gibbs entropy by the following expression

$$\mathrm{S}_G(\bar{\rho}) = -k_B \int_{\Gamma} \bar{\rho} \log (\bar{\rho}) d \tag{11}$$

Since it can be demonstrated that $\mathrm{S}_G(\rho) \leq \mathrm{S}_G(\bar{\rho})$ for any time (wherein $\mathrm{S}_G(\rho) = \mathrm{S}_G(\bar{\rho})$ holds when $\rho$ is already homogeneously distributed over the cells $\omega_i$, see Frigg [2008, p.75]), then we can employ this Gibbsian coarse-graining procedure to account for thermal behaviors such as equilibration so that the numerical results match the empirical TD-predictions.

### *Subjective vs. Objective Interpretations of Gibbs's Coarse-Grained Entropy*

As was the case with the Boltzmannian entropy concept, one of the most frequent interpretative coordinates with respect to specifying the meaning of the Gibbs entropy (in particular in its coarse-grained version) is the characterization as an agent-dependent or even artificial subjective property. Unlike its BSM alternative, the main criticisms of this GSM concept are directed at the artificial role it plays in the correction of the fine-grained paradox and the consequent prediction of processes with TD-like behavior. In this direction, Uffink (2010) points out that "the increase of entropy and the approach to equilibrium would thus apparently be a consequence of the fact that we shake up the probability density repeatedly in order to wash away all information about the past, while *refusing* a dynamical *explanation for this procedure*" (Uffink 2010, p.196. Italics are mine). For other authors (e.g., Sklar 1993), the concept of coarse-grained entropy lacks physical meaning precisely because the coarse-graining procedure constitutes an ad hoc modification of descriptive resources in the middle of the description procedure "Since we can't get from one empirically correct distribution to another via the dynamics, it seems the standard procedure is not justifiable from a mechanistic perspective" (Callender 2011, p.94). Also, as in the context of the Boltzmannian Γ-partitioning, there are no objective-universal elements involved in the choice of cell size $\omega_i$, hence the Gibbsian coarse-grained "depends essentially on our human choice of the size of the finite equal cells of boxes we partition [Γ]" (Grunbaum [1973], p.647).

In this agent-centric direction, Gibbsian cell sizes $\mu(\omega_i)$ have been usually justified (e.g., Tolman 1938, p.167) by interpretively associating them with the microscopic resolution of agents. The main conceptual problem with these agent-centric justifications of Gibbsian coarse-graining is that the GSM-descriptions of TD-like equilibration processes emerge as a result of a decrease in the microscopic resolution of agents[11]. Robertson (2020) recently proposed a novel justification of Gibbsian coarse-graining which does not explicitly rely on any observational limit nor ergodic considerations[12], but on what she calls the Zwanzig-Zeh-Wallace framework, which includes substantial modifications with respect to the standard Gibbsian procedure. First, a projection operator P* acts on the fine-grained density ρ by separating it into relevant[13] degrees of freedom and in irrelevant degrees of freedom, each set encoded in density $\rho_r$ and $\rho_{ir}$, respectively. Up to this point, one can define a pre-master equation for both densities $\rho_r$ and $\rho_{ir}$ that guide their respective Liouvillean dynamics. Based on such a justification strategy, Robertson (2020, p.565-568) justify the usage of a particular Gibbsian partitioning of Γ within the Zwanzig-Zeh-Wallace framework because of its role in deriving an autonomous irreversible dynamical pattern from a relevant forward-compatible density $\rho_r$ wherein the projection operation represents an abstract-descriptions generator that only rejects irrelevant details encoded in ρ. Based on this 'dynamical' (in Wallace's [2017] terms) justification of Gibbsian coarse-graining, the concept of Gibb's coarse-grained entropy could be interpretatively characterized as a non-agent-centric or subjective notion, since it is not conceptually based on the observational capabilities of agents.

### 2.2.4. Interpreting Thermophysical Entropy via Ordinary Disorder Concept

Apart from the agent-dependent vs agent-independence interpretative axis, one of the most decisive interpretative strategies for non-professionals and even within the scientific community has been to understand the meaning of the 'thermophysical' (i.e., either 'thermodynamic' or 'statistical mechanical') concepts of entropy through some non-technical metaphor-driven interpretation of 'molecular disorder'. This interpretative strategy goes back implicitly to the very origins of the TD-concept by Clausius (1865), who proposed to guide the intuitions about this notion (even before it was named 'entropy') through the degree of molecular disintegration represented in the dispersion of the components of the system. Explicitly, it was von Helmholtz (1882) who developed a systematic interpretation of entropy as a measure of molecular disorder, wherein the similarity of the movement of groups of molecules would intuitively suggest that the system has a low degree of entropy:

[11] Concerning this agent-based epistemic justification of Gibbsian coarse-graining, Frigg claimed that "The main justification is based on the finite accuracy of observations, which can never reveal the precise location of a system's micro-state in its γ-space. As the approach to equilibrium only takes place on the coarse-grained level, we have to conclude that the emergence of thermodynamic behaviour depends on there being limits to the observer's measurement resolution. This, so the objection continues, is misguided because thermodynamics does not appeal to observers of any sort and thermodynamic systems approach equilibrium irrespective of what those witnessing this process can know about the system's micro-state" (Frigg 2008, p.77).

[12] Malament and Zabell (1980) legitimized the introduction of Gibbsian partitions by postulating ergodic properties (mainly, a 'mixing' dynamics) in the described systems. Due to its later limited repercussion and for reasons of length, I will omit this strategy of legitimization of Gibbsian coarse-graining.

[13] The term 'relevant' is employed by Robertson (2020, p.552) to refer to those degrees of freedom from density ρ that contribute to obtaining a new 'autonomous' density $\rho_r$ by means of the coarse-graining operator P*.

> "Unordered motion, in contrast, would be such that the motion of each individual particle has no similarity to that of its neighbors. We have ample ground to believe that heat-motion is of the latter kind, and one may in this sense characterize the magnitude of the entropy as the measure of disorder" (Helmholtz 1882).

This Helmholtzian molecular-disorder interpretational strategy (prima facie aimed to understand Clausius TD-entropy [3]) was also adopted in Boltzmann's development of his BSM-concept (6), who not only linked the degree of spatial disorder of the constituent molecules of thermal systems with their entropy value, but also introduced central probabilistic considerations in his theoretical proposal "the system formed by these parts is also initially in an ordered state, and when left to itself it rapidly proceeds to the disordered most probable state" (Boltzmann 1896, p.443). That is, the level of order encoded in a high regularity in the configuration of the constituent molecules (e.g., a solid with crystalline structure) would be not only inversely proportional to its value of statistical mechanical entropy (2), but also highly improbable with respect to the space of possible molecular configurations. Thus, in the Boltzmannian framework the vast majority of the phase space $\Gamma$ of a system would correspond to those configurations of the system with less regularities or spatial patterns in the position and velocity of its molecular components (e.g., a gas in which its particles oscillate arbitrarily), being these also the most probable. Therefore, through the Second Law of thermodynamics expressed in this interpretative line, it could be predicted that all substances subject to this empirical law will tend to adopt their internal configuration (arrangement of molecular components) less spatially ordered via regular patterns.

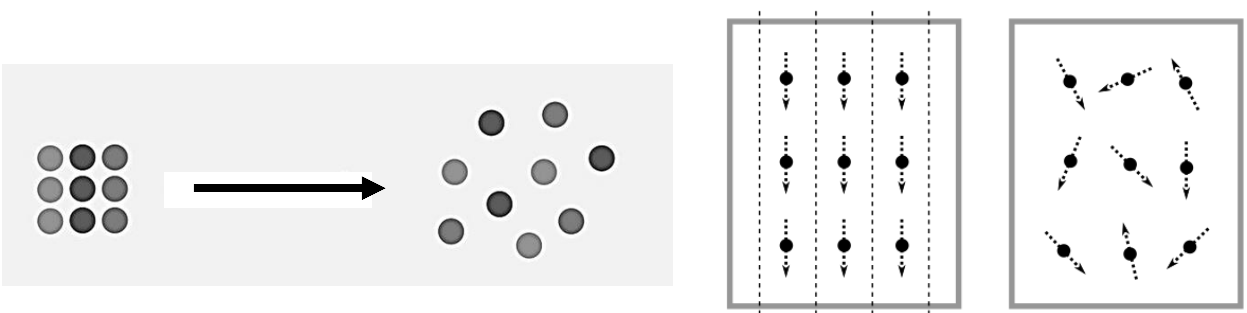

**Figure 1**. Entropy as measure of Disorder in SM and Adiabatic Demagnetization of a Gas.

One of the main criticisms, such as the one made by Burgers (1954), consists in pointing out the notion of 'disorder' as a vague and conceptually confusing, and even subjective and anthropomorphic term (Ben-Naim 2008). In other words, there is no single way agreed upon by the scientific community to qualitatively and precisely understand molecular disorder in the thermophysical context, grouping together different meanings: (i) order as similarity in relative molecular motion (Helmholtz 1882), (ii) order as pattern in molecular positions (Boltzmann 1896), or recently (iii) order as statistical correlation between position and molecular velocities (Haglund 2017), among others. However, it could be argued that the conceptual problem of this interpretation does not respond mainly to the vagueness of the everyday notion of 'disorder' (e.g., the disorder of a room), and the problem would persist even if we move to a technical-quantitative concept of disorder such as the one found in Landau theory with the 'order parameter'. In the expansion of a gas from a corner to the entire volume of a container by increasing its entropy, this increase could not be satisfactorily understood either by means of Landau's (hardly understandable for the layman) technical concept 'order parameter' precisely because the degree of order is the same at the beginning and at the end.

Finally, another of the main objections to disorder-driven interpretations of SM-entropies is to provide real cases that contradict the predictions derived from this interpretation. One of the most paradigmatic cases (Leff, 2007; Haglund, 2017) is in which an increase in entropy in certain liquid crystals causes these substances to pass from a state of molecular disorder to a state called 'smectic', where the molecules of this crystal are regularly configured-orientated in space by means of a sort of plane. Another example is that of a gas in a magnetic field that passes from a state of high regularity (visually explicit) in the molecular positions-velocities to another explicitly more irregular state without increasing the entropy of that substance. However, the criticism that there are counterexamples to this interpretation can be misleading, since it can lead to this type of answers "Disorder is a metaphor for entropy, not a definition for entropy. Metaphors are valuable only when they are not identical in all respects to their targets (...) The metaphor of disorder for entropy is valuable and thus imperfect" (Styer 2008, p.1031).

If from BSM one can interpretively exploit the spatial-disorder metaphor to grasp the meaning of entropy, from a Gibbsian framework entropy can also be visualized by the interpretation of entropy as 'mixed-up-ness'. The paradigmatic case in this strategic direction can be found in two gaseous substances whose component particles are mixed in the same volume (high-entropy state), while the separation of their corresponding particles in two contiguous but spatially separated volumes will indicate a low-entropy state in this system. To solve its interpretative dependence on spatial parameters (as with Helmholtzian disorder), other authors like Denbigh will seek instead to appeal to directly energy parameters[14] "it is not so much a question of mixing of the particles in space as of a mixing or sharing of their total energy" (Denbigh 1989).

[14] In the last decades, a substantial number of authors have proposed alternative interpretative frameworks based on the notion of 'energy spreading': systems far from thermal equilibrium tend to spatially spread their energy by the greatest number of dynamically accessible states, thus reaching equilibrium (i.e., maximum spread). In particular, Leff (2007) defended that the capacity of this metaphor to explicitly appeal to spatial, temporal and energy parameters gives a dynamic character to the way in which agents must understand the meaning of entropy, based on how the system travels through the space of possible states during its dynamic evolution. Close to this space-centric spectrum, we also find the interpretation of entropy as the degree of freedom of a system, linked to the variety or the number-volume of microstates $\mu(\Gamma_M)$ (which is why it is also called a 'multiplicist' interpretation, Leff [2007, p.1747]) contained in the macrostate-region $\Gamma_M$ in which the system is located. According to Styler (2008), this would allow (by combining interpretative perspectives) to solve that disorder-driven interpretations are exclusively focused on a single microscopic configuration $x_1$ of the system, also incorporating the degree of disorder that we would find in other macroscopically compatible configurations $x_i \in \Gamma_M$.

## 2.3. Landscape of Probability Concepts

As we have seen so far, the concept (or properly 'the concepts') of probability is indispensable to satisfactorily account for the main notions of entropy and information. However, the development of well-defined and theoretically robust concepts of probability was not only technically decisive for the historical development of SM (both BSM and GSM formalism) and information theory (IT), but also played a central role in the development of the intellectual trend that concerns us here, i.e., information physics. For this reason, it would be impossible to understand the conceptual foundations of this field without minimally detailing the main technical notions of probability that we find in this field. Let us now turn to detail the different concepts of probability in BSM, GSM and IT, and then explore the main interpretative strategies that can be found in the literature.

### 2.3.1. Framing Probability Concepts and Interpretations

Probabilistic concepts permeate all intellectual history since antiquity, although they begin to be technically defined from the 16th century onwards (Franklin, 2001; Hajek, 2019). Unlike the conceptual plurality underlying 'entropy' and 'information', all technical probability concepts have a formal-axiomatic definition generally accepted within the scientific community. This was provided by Kolmogorov (see Chapter 2.3.2.) in his 1933 *Foundations of the Theory of Probability*, characterizing probabilities P as (i) non-negative real numbers r between 0 and 1, (ii) for the set of possible events O, then P(O) = 1 (P is normalized), and (iii) is additive, i.e., P(A) + P(B) = P(A∪B). Despite having this axiomatically satisfactory Kolmogorov definition, the meaning of the concept (or concepts) of probability within contemporary physics is extremely difficult to specify[15].

For this reason, the task of minimally understanding its meaning cannot be performed independently of any interpretation of P that we may adopt (as happens to some extent with particular concepts of entropy and information), but depends closely on the particular interpretation that we adopt with respect to the concept of probability P. As we see below, each of these interpretations does not consist merely in assigning semantic values to the common probabilistic elements, but also incorporates a supplementary (or in certain cases also complementary) formal apparatus to the Kolmogorovian formal architecture. Despite the vast number of interpretations of probability that we can find in the literature, these can be classified (at least as far as the intellectual context of information physics is concerned) among 'objective interpretations of P' or simply 'objective probabilities' "in which probability is a feature of the material world independent of our knowledge or beliefs about that world." (Lavis 2011, p.51) and

[15] "Probability is the most important concept in modern science, especially as nobody has the slightest notion what it means" (Russell, 1929 Lecture [quoted in Bell 1945, p.587]).

'epistemic interpretations of P' or 'epistemic probabilities' in which probability is a measure of the degree of belief (rational, or otherwise) of an individual or group" (Lavis 2011, p.51).

Among the objective probabilities, the most conceptually simple is the so-called 'frequentist interpretation', where the probability of an event e is equivalent to the frequency (relative to a finite or infinite sequence of trials) of occurrence of this event within a random trial. It is therefore determined by the actual number of occurrences of the event relative to the total number of possible occurrences within a sequence. This interpretative architecture was famously defended by Reichenbach (1949) and formalized by von Mises (1957) in the case of limiting relative frequencies in 'collectives', the latter incorporating the axioms of (iv) convergence (i.e., the frequency of the event exists in the limit) and (v) random place selection, where the limiting relative frequency of an event in the collective is independent of the place where it occurs.

However, the most relevant objective probability within the SM conceptual landscape is that known as 'time averages', where the probability of occurrence of an event (i.e., the actual microstate *x* of the system) within a particular region R of is equivalent to the average time that the actual microstate *x* spends within this region R during a certain interval. As Frigg (2008, p.34) suggests, this interpretation presents certain conceptual problems such as defining what the relevant time interval is (e.g., time interval of a SM-measurement) or whether one has strong theoretical reasons that this average actually exists. Regardless of these issues and as I will detail below, the feasibility of the time average interpretation in the SM or IT domains ultimately depends on the appropriateness of the particular dynamics of the system.

After frequentism and time average, a third candidate of objective probabilities is found in 'propensity interpretations', where probabilities are ontic properties (conceptualized as 'tendencies' or 'propensities') of physical reality. Famously, one of the main proponents of the propensity interpretation was Popper (1959), motivated mainly by certain physical behaviors in the field of quantum physics. Finally, objective probabilities have also sought to be interpreted in the literature as chances via 'best systematizations' à la Lewis (1994b) of events contained in the Humean mosaic constituting the physical reality, seeking a balance between simplicity, strength and fit with which they systematize such events. These were used by Loewer (2004) to defend the compatibility between real chances and deterministic (non-chancy) dynamics, assessing the best ways of systematizing the distribution of initial conditions of a system with the maximum possible strength and simplicity, which is key in the case of SM.

As for the broad category of epistemic (or also 'subjective') interpretations of probabilities, their technical development dates back to the 18th and early 19th century with the configuration of Bayesianism and Laplace's classical perspective based on the principle of Sufficient Reason, i.e., in the face of uncertainty about the occurrence of events, we must assign the same probability to every possible event. However, it will be during the 1920s at the hands of Ramsey and Jeffrey (Uffink 2011, Section 3) that epistemic probabilities acquire a well-defined technical and conceptual architecture. According to this broad interpretative direction, P probabilities represent degrees of beliefs of rational agents. For epistemic interpretations such as de Finetti's 'personalist' one, any kind of belief is legitimate to be represented by probabilities as long as it fulfills certain coherence conditions delimited by the famous 'Dutch Book Theorem', i.e., that beliefs satisfy the above-mentioned Kolmogorov axioms (i-iii).

Undoubtedly, the most relevant epistemic interpretation in the intellectual domain of information physics is the neoclassical or 'objective epistemic' (as opposed to Ramsey's and Finetti's 'subjective epistemics') proposal of Edwin Jaynes (1957a) regarding probabilities in SM and IT. This interpretative proposal is based on the famous 'Maximum Entropy Principle' (MEP), which states roughly (more on this later) that the probability distribution P that best fits both (i) the agent's observable knowledge about the system, and (ii) an agent's uncertainty about the actual microstate of a system is the distribution that maximizes the IT-entropy value. In this sense, MEP (characterized as a rationality condition for agents) constitutes a generalization of Laplace's principle of Sufficient Reason. Therefore, Jaynes' (1957a) objective epistemic probabilities would not merely represent agents' beliefs, but their epistemic judgment regarding their ability to infer the actual microstate of the system. Having detailed the main interpretations of general probabilities, let us now turn to the particular probability concepts found in SM and IT.

### 2.3.2. Static Statistical-Mechanical Probability

The main historical attempts to reduce or explain macroscopic thermal phenomena to molecular dynamics (paradigmatically Boltzmann [1872], as we have seen) succumbed to the use of probabilistic modeling, either because of conceptual problems such as Loschmidt's reversibility objection or directly because of the pragmatic impossibility of calculating the vast amount of microstates. Although the first probabilistic considerations for modeling molecular quantities date back at least to the 'Hydrodynamica' written by Danniel Bernoulli in 1738, the first systematic uses of probability in gas kinetics are found in the work of Clausius and Maxwell during the 1850-1860s (see Uffink, 2007; Callender, 2011). In the particular case of Maxwell (1860), his derivation in 1860 of the molecular velocity distributions for a gas in thermal equilibrium (what is known as the 'Maxwell distribution') assuming independence of the i, j and k components of the velocities v of individual molecules, i.e., $f(v_i, v_j, v_k) = f(v_i)\, f(v_j)\, f(v_k)$. This author reformulated the derivation of this distribution in 1867 by assuming (similar to Boltzmann [1872], as we saw in Chapter 2.2.2) the statistical independence of the velocity of pairs of molecules before collision i.e., $f(v_1)\, f(v_2) = f(v'_1)\, f(v'_2)$. These types of probabilistic assumptions in gas kinetics have been referred to in the literature as 'static probabilities' (e.g., Callender 2011, p.91), because they were not constructed from molecular dynamics.

The rudimentary use of probabilistic tools in first-period gas kinetics (1860-1872) gave rise to what are now considered the standard SM-probability concepts for both BSM and GSM users. Let us first consider the Boltzmannian framework. The notion of probability in this context depends on the particular measure μ chosen to bound the size of uncountably infinite sets of numbers, precisely because in classical phase space $\Gamma$ each particular region $\Gamma_R$ is composed of an infinite number of microstates. Due to its many formal virtues (mainly, it is conserved via Liouville's theorem, where $\mu(\Gamma_R) = \mu(\Gamma_R)$ for all time t), the uniform Lebesgue measure is usually considered in the literature as the natural measure with respect to $\Gamma$. Probability measures $\mu$ as technical devices allow us to specify the size of distinct phase regions $\Gamma_R$ with infinite microstates, which is particularly useful in the Boltzmannian task of delimiting the size of the phase regions m occupying the macrostates. With these elements, it is assumed as a postulate in BSM that the probability P (wherein P = r, being r a real number between 0 and 1) of the actual microstate *x* being within the set $\Gamma_M$ of microstates associated with its actual macrostate M is defined by

$$P(M) = \mu(\Gamma_M) \tag{12}$$

wherein, the probability assignment to a macrostate corresponds to the uniform measure of the phase volume (or number of microstates) contained in the region associated with that macrostate. It should be mentioned that $\mu$ is normalized, so that the sum of measures of disjoint macrostates adds up to one. Because this 'Boltzmannian static probability' ($SP_B$) is defined over macrostates, it is also known in the literature as 'macroprobability' (Frigg 2008, p.12). Since in BSM a uniform probability measure $\mu$ is used, the probability P for an individual microstate of the actual macrostate-region $\Gamma_m$ being the actual one *x* is identical for all microstates contained in this macrostate-region $\Gamma_m$.

On the other hand, in GSM the selection of the particular probability measure describing the ensemble is carried out by different considerations than with BSM. Illustratively, if one wants to select the measure that maximizes the fine-grained Gibbs entropy (8) for ensembles in equilibrium with constant energy E, volume V and number of particles, then this will correspond to the uniform measure $\mu$ (or equivalently, the microcanonical density $\rho_\mu$) defined over the system's phase space $\Gamma$. However, unlike in BSM with (12), the uniform measure is not necessarily the appropriate density to represent canonical or microcanonical ensembles. In this SM-context, the probability P for a random microstate *x* being the actual one within the set of microstates $\Gamma_\rho$ on which the density $\rho$ of the chosen ensemble (whatever it is) is defined can be specified by

$$P(\Gamma_\rho) = \int_\Gamma \rho(\Gamma_\rho)d\Gamma \tag{13}$$

which reduces to $P(\Gamma_\rho) = \mu(\Gamma_\rho)$ in the case of microcanonical ensembles. This expression (13) is precisely the definition of the concept of 'Gibbsian static probability' ($SP_G$), homologous to $SP_B$ but only well-specified in equilibrium scenarios. Once I have technically defined the central SM-probabilistic concepts of $SP_B$ and $SP_G$, I should now specify and make explicit their meaning: What does it mean that the probability that the actual microstate *x* can be characterized by M is approximately $\mu(\Gamma_M)$? To do so, I must analyse which interpretation of the probabilities is the most appropriate to delimit the semantic content of $SP_B$ and $SP_G$.

Let us start with the objective probabilities. In particular, from a frequentist perspective one could claim that the probability of *x* actually being located at time *t* within the region $\Gamma_M$ or $\Gamma_\rho$ is P means precisely that P encodes the frequency with which we would find *x* in $\Gamma_M$ or $\Gamma_\rho$. For these SM-probabilities to be conceived frequentist-driven one should reconceptualize the uniform measure of macrostates in BSM and densities in GSM as if they were 'urns' from which we randomly draw not colored balls of different colors but individual microstates. In the first case we find Ehrenfests (1912, p.32-34), who popularized the probabilistic urn model to account for statistical statements based on BSM. In the second case, it was Gibbs (1902, p.163) himself who originally proposed to interpret the ρ-densities representing ensembles as if they were urns[16]. In addition to this urn-like reconceptualization of $\Gamma_M$ and $\Gamma_\rho$, for the $SP_B$ and $SP_G$ probabilities to constitute well-defined frequencies it is technically necessary for individual trajectories $\Gamma\gamma$ in $\Gamma_M$ and $\Gamma_\rho$ to correspond to 'collectives', satisfying the axioms of Von Mises (1931, p. 519) seen above. However, authors such as van Lith (2001, p.587) have argued that the dynamical evolution of the actual microstate *x* in the regions $\Gamma_M$ and $\Gamma_\rho$ does not satisfy condition (v) of random place selection, precisely because the deterministic dynamics of the microstates makes the selection of the microstate independent of its location in M and p, but dependent on its position within its trajectory. Indeed, that systems possess deterministic dynamics directly implies that $SP_B$ and $SP_G$ cannot be ontically understood as a tendency or disposition à la Popper (1959) of systems (Clark, 2001; Hoefer, 2007, p.558-559).

One of the most relevant ontological conceptions of $SP_B$ and $SP_G$ in the literature is to conceive the probability P of *x* being located within the region $\Gamma_M$ or $\Gamma_\rho$ be P as a time average of the trajectory of *x* in $\Gamma_M$ or $\Gamma_\rho$ during a certain time interval $\Delta t$. There is a consensus within the philosophical community (Lavis, 2011) that the theoretical feasibility of time average interpretations of these SM-probabilities ultimately depends on the dynamics of the system being 'ergodic' (or certain forms of 'ergodicity'), namely: for infinite time intervals, the time average that the microstate spends in $\Gamma_M$ or $\Gamma_\rho$ corresponds to the phase average of $\Gamma_M$ or $\Gamma_\rho$. As is well known, the ergodic program in SM suffers from multiple technical shortcomings[17], such as the 'zero measure problem' (i.e., there exists a subregion $\Gamma_S$ of the region of $\Gamma_R$ in which ergodicity is not satisfied), and notable conceptual problems, such as ensuring that physically meaningful systems are indeed ergodic (see Frigg, 2008, Section 2.4.3.). Recently Lavis (2011) advocated an objective time average interpretation of $SP_B$ (extendable to $SP_G$) based on what von Plato (1989) called 'ergodic decomposition', i.e., partitioning the phase space $\Gamma$ of the system into distinct non-overlapping regions $\Gamma_\mu$ in which the measure μ was ergodic[18].

[16] "On the frequency interpretation one thinks about an ensemble as analogous to an urn, but rather than containing balls of different colours the ensemble contains systems in different microstates (Gibbs 1902 [1981, p. 163]). The density [ρ] specifies the frequency with which we draw systems in a certain microstate." (See Frigg 2008, p. 153).

[17] "On the time average interpretation, [ρ] reflects the fraction of time that the system would spend, in the long run, in a certain region of the phase space if it was left to its own. Although plausible at first blush, both interpretations face serious difficulties and it is unclear whether these can be met" (see Frigg 2008, p. 153, 155).

[18] "Following von Plato 1989b we propose the use of the time-average definition of probability within the members of the ergodic decomposition. (…). It is now augmented by the second part, in which a version of Cartwright's (1999)

Let us now explore the feasibility of specifying the meaning of $SP_B$ and $SP_G$ by means of some epistemic interpretation, these being especially fruitful in the theoretical field of the latter concept. It was precisely Tolman (1938, p.38) with his 'fundamental hypothesis of equal a priori probabilities' who first explicitly advocated interpreting the densities in GSM as representations of the agent's knowledge. Tolman argued (applying the principle of Insufficient Reason to the theoretical scope of GSM) that the agent's ignorance about the microstate of the system should be represented by a uniform or microcanonical measure on the energy hypersurface of the system, a thesis later refined with the proposal of Jaynes (1957a) mentioned at the beginning of this section. Recently Uffink (2011) defended an epistemic conception of $SP_G$ based on de Finetti's personalist interpretation, in which "All that is required is that this probability distribution adequately represents what an agent actually believes about the location of the microstate of the system in that space. And since agents are entitled to believe whatever they want, under some rational consistency requirements, the problem of justifying a choice of this distribution simply does not occur." (Uffink 2011, p.48). The lack of empirical considerations in the choice of the particular density is taken by Uffink as a virtue of epistemic SM-probabilities[19], precisely because it does not require that the dynamics of the system be necessarily ergodic, as we have seen to be the case with certain objective interpretations of BSP. However, this jeopardizes the desirable characterization of SM as a theory dealing directly with physical systems (as in the case of Jaynes, 1957a). I will return to the analysis of this argumentative strategy later on.

### 2.3.3. Dynamic Statistical-Mechanical Probability

The concepts of SM-probabilities $SP_B$ and $SP_G$ are not the only probabilistic notions that play a central role in this area of physics. The main reason for introducing additional probabilistic concepts in both $SP_B$ and $SP_G$ is because statistical probabilities are not sensitive to the dynamics of systems, being especially problematic in accounting for TD-like processes of approaching equilibrium (Callender, 2011; Hemmo and Shenker, 2012). Illustratively, Boltzmann (1877, p.165) historically employed $SP_B$ to SM explain how the actual microstate $x$ of system away from equilibrium will eventually approach this macrostate-region, so that microstates always evolve in the future to macrostates with higher probability (Frigg, 2008, p.23). However, whether this is the case depends fundamentally on the dynamical behavior of the microstates, since neither $SP_B$ nor $SP_G$ eliminate the theoretical possibility that the actual microstate of system $x$ at $t_0$ (i) transits in

nomological machine is used to assign probabilities over the members of the decomposition. It is shown that this probability scheme is particularly well adapted to a version of the [BSM]" (Lavis 2011, p.52).

[19] "a subjective view on probability in statistical physics is clearly a viable view to take, when dealing with ensembles. But substantial issues within this approach remain: whether one takes a Bayesian, a de Finettian, or a Jaynesian approach on how to model the impact of empirical evidence. To me, de Finetti's approach seems by far the most promising amongst such approaches. This view, of course, brings along that one does not need any physical or empirical assumption for adopting a particular probability distribution over phase space. (Uffink 2011, p.48).

the future to lower entropy macrostates-regions $\Gamma_{M0}$ or (ii) that the past has already transited through the thermal equilibrium macrostates-region $\Gamma_{Meq}$.

To solve this theoretical deficiency, Albert (2000, Chap. 4) in the theoretical context of BSM proposed to assume that the initial macrostates-region $\Gamma_{M0}$ of a system (the author considers the universe itself, but it could be applied in other isolated systems) in an equilibration process has a uniform probability with respect to the Lebesgue measure $\mu$. This is what is known in the literature as 'Past Hypothesis', which encompasses: (i) postulating that the initial state of the system/universe was in fact in a low-entropy region $\Gamma_{M0}$, and (ii) a uniform measure $\mu$ on $\Gamma_{M0}$ representing the chance that the actual initial state was in a given subregion of $\Gamma_{M0}$. A proposal analogous to Albert's but applied to the Gibbsian framework has been developed by Wallace (2011), where the initial ensemble is represented by a uniform microcanonical density $\rho_{\mu}$. Focusing on the disciplinary scope of BSM, as the system evolves dynamically during a certain time interval $\Delta$t we consider how the uniformly distributed microstates from $\Gamma_{M0}$ transit through phase space $\Gamma$ as 'dynamical blobs' $\Gamma_{\rho}$, defined as the set of microstates resulting from the dynamical trajectories originated from the microstates of $\Gamma_{M0}$ during this time interval (Hemmo and Shenker 2012, Chap. 5). I should remark that these microstates from $\Gamma_{M0}$ evolve dynamically following the Liouville equation in the form of an incompressible blob $\Gamma_{\rho}$, keeping its $\mu$-measure invariant but not necessarily its shape (satisfying certain topological constraints), i.e., $\mu(\Gamma_{M0}) = \mu(\Gamma_{\rho t})$ for all t. The Past Hypothesis assumption would allow us to define the degree of probability for the actual microstate *x* of the system contained in the dynamical blob $\Gamma_{\rho}$ is within a particular macrostate-region $\Gamma_{M0}$ at time t, characterized by the macrovariables associated with it at t. Particularly the probability P of system's microstate at $t_1$ being within the macrostate-region $\Gamma_M$ is proportional to the size of the intersection between the dynamical blob and the macrostate at this particular time:

$$P(\Gamma_M) = \mu(\Gamma_{\rho} \cap \Gamma_M) / \mu(\Gamma_{\rho}) \tag{14}$$

With this we define the concept of SM-probability known as 'micro-probabilities' (Frigg, 2008, p.43-46)'conditional', 'transition' (Hemmo and Shenker 2012, p.131) or even 'dynamical probabilities' (DP) (Callender, 2011), whose pivotal role in clarifying the conceptual foundation of non-equilibrium-SM has been long debated in recent years. Unlike what happens with $SP_B$ and $SP_G$, the content of the SM-probabilistic concept of DP is not simply assumed to be theoretically primitive, but is derived from the particular dynamics (i.e., time-reversible, deterministic) of the Hamiltonian system. Of course, the SM-system's dynamics cannot necessarily be deterministic, as it is the case of Bernoulli-like processes (i.e., discrete-time stochastic), but won't consider those peripheral cases for the sake of simplicity. Moreover, conceptual compatibility between $SP_B$- $SP_G$ and DP is problematic, since not only may both assign different numerical values to certain set of microstates, but also (and this is key) their particular meaning may be substantially different. In order to specify this meaning comparatively, it is worthwhile at this point to evaluate the most relevant interpretations of DP that can be found.

On the one hand, one could understand the meaning of DP from an ontological interpretive strategy. Within a BSM framework, Hemmo and Shenker (2012, Chapter 6) proposed to select the probability measure upon which DP is articulated by strictly empirical considerations based on the frequency with which systems with thermodynamic behavior (e.g., the paradigmatic example of an expanding gas) are characterized by one set of macrovariables and another. The technical way to accomplish this task is to calculate the frequency with which the actual microstate *x* contained with the initial condition macrostate-region $\Gamma_{M0}$ of the system (i.e., the gas in a subvolume V of the whole container) at each experimental setting end up at the end of a certain interval Δt within the part of the dynamical blob p that overlaps a macrostate-region $\Gamma_{M1}$ or within another $\Gamma_{M0}$. Once this procedure has been repeated a considerable number of times, we select the probability measure that best fits the region corresponding to the overlap of $\Gamma_{M1}$ and $\Gamma_{\rho}$, as defined by DP. As Shenker (2020) remarks, this frequentist strategy to justify in an empirically-objective and conceptually consistent way is rooted on a doubly (pragmatically) impossible computational procedure[20] due to the uncountably infinite number of microstates of $\Gamma_{M0}$ and of microstatistical trajectories emerging from $\Gamma_{M0}$. This makes it possible to understand the meaning of DP as highly efficient computational shortcuts for performing predictions (Hemmo and Shenker, 2012, Section 6.5.).

Recently, Frigg and Hoefer (2015) developed a refinement of this frequentist objective perspective by developing a best-system interpretation of DP as 'Humean Objectives Chances', where the meaning of this SM-probabilities comes directly from the structure of phase space Γ (interpreted here as a Humean mosaic). While for the frequentist proposal of Hemmo and Shenker we consider fitness regarding $\mu(\Gamma_{\rho})/\mu(\Gamma_{M1})$ as the key parameter in the choice of measure u, for the best system proposal of Frigg and Hoefer we should not only consider its 'fit' but also include its 'simplicity' and 'strength' for this measure to constitute the best system. However "listing all points individually and checking whether they lie in [$\Gamma_{M1}$] is extremely cumbersome and won't make for simple system. So we have to reduce the complexity of the system by giving a simple summary of the distribution of points" (Frigg and Hoefer 2015, p.563). To solve this lack of simplicity the authors propose that DP depends not on individual microstates *x* and their Hamiltonian trajectory H(x) but on blob-like continuous distributions ρ and their Liouvillean trajectory, obtaining the latter by means of certain technical fitting procedures. Moreover, Frigg and Hoefer (2015, p.572) defended that the best Humean system (i.e., simplest/strongest) underlying DP would plausibly be a uniform

[20] This idea is nicely expressed in the following quotation "How, then, can we make predictions about the system's evolution? Suppose, by idealization, that we solve the equations of motion of the system: every microstate in $M_0$ is a starting point of an evolution, and we calculate the trajectory segment that starts at that microstate at time $t_0$ and follow their end points at times $t_1$, $t_2$, and $t_3$ (…). We do that for all the microstates in $M_0$, and examine first the set of end points at $t_1$ (…) We then carry on this calculation for the later moments $t_2$ and $t_3$ (…). (This is an idealization, and such a calculation is doubly impossible: the system is too complex, and the number of microstates, and hence of trajectory segments, is a continuous infinity. However, this idealization is useful in order to understand the nature and status of probability in statistical mechanics. There are shortcuts that enable practical predictions, which are actually carried out in statistical mechanics.)" (Shenker 2020, p.11).

Lebesgue measure $\rho_\mu$, since Dirac-delta peaked distributions $\rho_\delta$ over the initial condition macrostate $\Gamma_{M0}$ (providing a better fit than $\rho_\mu$) are computationally impossible to specify.

Let us now move to the domain of epistemic interpretations of SM-probabilities. In the case of Jaynes' (1957a) MEP, these allow one to assign prior (or 'static') SM-probabilities $\rho^{prior}$ as a function of the agent's macroscopic knowledge and uncertainty about the system's microstate. However, as Uffink (2011, p.44) pointed out, the theoretical apparatus linked to the MEP is not technically capable of representing how the epistemic states (beliefs, expectations, judgments, etc.) of agents are updated as the system begins to evolve dynamically from its initial macro-condition. At best, it can be possible to artificially solve this technical problem by an ad hoc à la Bayes conditionalization[21] of the initial distribution with respect to new microstate-based events e (see Jaynes 1983, p.250). However, this does not imply that 'new initial distribution' $\rho^{prior*}$ is actually able to encode (satisfactorily or not) the degree of knowledge that the agent possesses with respect to the system's dynamics, which (this is key) would be indispensable to justify DP from an epistemic interpretation of its meaning. If we now consider the role of epistemic interpretations of $SP_G$ à la Tolman (1938) to epistemically account for DP, it has been the object of harsh criticism by the Neo-Boltzmannian community in the past decades, as this famous quote by Albert shows:

> "Suppose that there were some unique and natural and well-defined way of expressing, by means of a distribution-function, the fact that "nothing in our epistemic situation favors any particular one of the microconditions compatible with X over any other particular one of them" (…) Can anybody seriously think that our merely being *ignorant* of the exact microconditions of thermodynamic systems plays some part in (…) in *making it the case*, that (say) *milk dissolves in coffee?*" (Albert 2000, p.64).

As Frigg and Werndl (2011a) or Myrvold (2016) suggest, the most philosophically interesting way to understand this objection of Albert's is that epistemic SM-probabilities are not powerful enough to account (descriptively or explanatorily) for DP, in particular, and for the microstate-based dynamics of thermal systems approaching equilibrium, in general. Within this debate on the plausibility of the epistemic meaning of DP, Uffink (2011) argues that epistemic SM-probabilities are unable to shed light on TD-like microstate dynamics, as well as that this interpretive strategy can illuminate the beliefs of agents in non-equilibrium SM scenarios[22]. This charitable position was replied by Callender: "why should we *believe* the gas likely moves toward equilibrium unless its *objectively* likely to move toward equilibrium? The thought is that the subjective probabilities

[21] A Bayesian updating would consist of using Bayes theorem to include new macrostatistical data into the model. This 'inadequacy' should be understood by the fact that the posterior distribution obtained from Bayes theorem is independent from the system's dynamics.

[22] "Instead, if we use subjective probability in statistical physics, it will represent our beliefs about what is the case, or expectations about what will be the case. And the results of such considerations may very well be that we ought to expect gases to disperse, ice cubes to melt, or coffee and milk to mix. They do not cause these events, but they do explain why or when it is reasonable to expect them." (Uffink, 2011, p.45). We could insert this position of Uffink within the tradition of explanations of phenomena by appealing to the 'nomic expectation' of certain physical laws, so that the dynamics of the system contribute to the fact that the beliefs of a subject will evolve until the expectation of a given phenomenon is extremely high. In Section 5.3. I will return to this issue in more detail.

should bottom out in something objectively probable. (…). Inasmuch as we want an interpretation that shares the goal of providing the mechanical underpinnings of thermodynamics, subjectivism won't help." (Callender 2011, p.105).

### 2.3.4. Information-Theoretical Probability

The concepts of probability not only play a role in the physical domain of SM but are also key in the formulation of the concepts of information in IT, and therefore also in a theoretically derived way (in the sense that to introduce the latter requires the former) in AIT and in the multiple semantic-IT proposals. In the case of IT-information, Shannon introduces the probabilities on the alphabet of symbols that constitute the source S (although because of the asymmetry of the communicative model, he could also consider the probabilities of the destination alphabet, usually coextensive with that of S). Beyond mentioning the 'frequency' of the English letters occur in his illustrative example, the meaning of the IT-probabilities is not (constitutively or interpretatively) specified in Shannon's (1948) theoretical proposal since: "nothing is said about how the probabilities are determined nor about their interpretation: they may be conceived as propensities theoretically computed, or as frequencies previously measured" (Lombardi et al. 2016b, p.2000).

Following Frigg and Werndl (2011, p.119), one of the main interpretative strategies we can pursue in this respect depends directly on the particular characterization of source S. On the one hand, we can assume that IT-probabilities are ontological properties of the source[23]. If we take this in a directly ontic sense (where the S-driven generation of symbol sequences must be taken as a completely stochastic process), then IT-probabilities end up being dispositions of source S to generate certain messages and not others. This interpretative strategy is only technically feasible when the source is stochastic and the system is ergodic, however, it is prima facie possible to conceive of non-stochastic sources (e.g., the generation of messages via everyday language). Relaxing its ontic nature and that of the source, we could also interpret IT-probabilities à la Shannon (1948) as frequencies of occurrence of symbols in messages.

On the other hand, one could assume interpretatively along the lines of Jaynes (1957a) that the IT-probabilities represent the degrees of beliefs of the agents regarding the occurrence of symbols in messages, which is conceptually favored by characterizing the source as non-probabilistic[24]. At this point we could ask ourselves which elements of the communicative model constitute the agent that possesses such degrees of belief, i.e., the source, the destination, etc. Even assuming that we can theoretically characterize the destination D as an epistemic agent in IT, the main problem

---

[23] "Alternatively, the source itself can be probabilistic. The probabilities associated with the source have to be ontic probabilities of one kind or other (frequencies, propensities, etc.). (…). In Shannon's setting this means that the $p(m_i)$ have to be equal to the source's objective probability of producing the message $m_i$." (Frigg and Werndl 2011, p.119).
[24] "If the source itself is not probabilistic, then the $p(m_i)$ express the beliefs and nothing but the beliefs – of receivers. For proponents of subjective probabilities these probabilities express the individual beliefs of an agent, and beliefs may vary between different receivers. Objectivist insists that all rational agents must come to the same value assignment. This can be achieved, for instance, by requiring that $S_S(P)$ be maximal, which singles out a unique distribution." (Frigg and Werndl 2011, p.119).

(apart from defining how these IT-probabilities are assigned, e.g., via MEP) lies in the fact that either it is not possible to represent the agent's degree of belief with respect to a single message m, or this degree of belief results from a 'statistical average' of the degree of belief with respect to the whole source. Note that this interpretative dilemma is the same one that arose in the analysis of IT-information, and therefore it can be somehow technically dissolved by reconceptualizing them as AIT-probabilities or SIT-probabilities.

## 2.4. Landscape of Information Concepts

Together with entropy, the most relevant conceptual manifold in the field of information physics is, evidently, that of information. As with the notion of entropy, within this intellectual current the term 'information' does not univocally express a single concept of information, and sometimes even certain informational concepts will bear a different name. As Shannon pointed out: “The word ‘information’ has been given different meanings by various writers in the general field of information theory. […] It is hardly to be expected that a single concept of information would satisfactorily account for the numerous possible applications of this general field” (Shannon 1993, p. 180). The conceptual landscape of information deployed by the physics of information ranges from the ordinary notion of information to technical and quantitative informational concepts, semantic notions and others without semantic dimension, concepts linked to the knowledge of agents or others that are not. In order to clarify these subtle conceptual differences, we now proceed to explore the main notions of information, their underlying theoretical motivations and the possible strategies for interpretatively specifying their meaning.

### 2.4.1. Ordinary Information

As I will see in the following chapters, one of the most important concepts of information for information physics, both for its historical shaping and for the consistency of its main arguments, is its everyday or ordinary meaning (OI). The term 'information' that we use in everyday language (i.e., this includes contexts of real scientific practices that I will focus on) does not imply a univocal concept with well-defined meaning; on the contrary, its various uses as such have journalistic, computational, or directly philosophical connotations, directly related to truth or meaning. This multiform everyday non-technical concept is usually characterized within the recent philosophical literature, both in the philosophy of information pioneered by Floridi (2011) and in the philosophy of physics (Timpson 2013, p.10-16; Lombardi et al. 2016a, 2016b), as a notion with explicit semantic character[25], remarked in its directed intentionality or 'aboutness' manifested in everyday

[25] “the pervasiveness of the notion of information both in our everyday life and in our scientific practice does not imply the agreement about the content of the concept. (…) it is a polysemantic concept associated with different phenomena, such as communication, computation, knowledge, reference, meaning, truth, etc.” (Lombardi et al. 2016b, p.1984).

expressions, e.g., 'information *about* something'. In those philosophical assessments of OI, this non-technical concept is also specified as an agent-centric notion, since it is usually (mentally or not) possessed by individuals or is employed by them e.g., ‘Bob *has* information about something’. Its intentional character means that it is an asymmetric notion wherein ‘Bob *has* information about P’ does not mean that ‘P *has* information about Bob’. And finally, the ordinary concept of information is a notion closely related to that of ‘knowledge’ (usually, they are treated as synonymous), therefore it should be considered as an epistemic notion, due to the fact that P is 'informational' for an epistemic agent B if P somehow provides knowledge to that epistemic agent[26]. This everyday notion of information (as we saw with the ordinary notion of 'order') is a non-technical concept, which simply means that its multiple meanings are not theoretically determined in some formal specification but also that there are no well-defined guidelines that allow standardized uses of this concept. This everyday notion of information will serve as an inspiration throughout history (e.g., Carnap and Bar-Hillel, 1952; Floridi, 2011) to develop technical concepts that capture some of its central properties (i.e., semantics, epistemics, etc.).

### 2.4.2. Information-Theoretical Information

*Hartley Information*

The historical origins of the technical concepts of information can be traced back to engineering studies of telegraph signal transmission during the 1920s. In particular, the symbolic origin is often attributed to a 1924 paper entitled ‘Certain Factors Affecting Telegraph Speed’. Its author, Bell Laboratories engineer Harry Nyquist, proposed to measure the transmission speed of "intelligence" using certain signals by implementing a method of digitizing continuous waves that would improve the efficiency of such transmission. Following Nyquist's proposal, the first quantitative-technical concept of "information" was developed in a paper by Bell Laboratories researcher Ralph Hartley (1928), whose quantity was intended to measure the amount of information obtained by the occurrence of certain equally probable elements among a set of n possible excluding alternatives. As this author pointed out, with the choice of the term 'information' (whose use goes back several centuries, see Adriaans [2020]) as opposed to Nyquist's notion of "intelligence", the aim was to eliminate the psychologistic connotations of the latter. Without the need for probabilistic assumptions or assuming equal probability of his symbols (Frigg and Werndl 2011, p.121) , Hartley's measure of information (loosely associated with the "meaning" conveyed by his symbols (ibid, p.535) can be represented mathematically by a logarithmic function (also used by Nyquist [1924]) that has certain nice mathematical properties, namely: (i)monotonic increase with respect

[26] “The everyday concept of information is closely associated with the concept of knowledge, language, and meaning; and it seems, furthermore, to be reliant in its central application on the prior concept of a person (or, more broadly, language user) who might, for example, read and understand the information; who might use it” Timpson 2013, p.11).

to n, (ii) 'branching' or independence of the grouping of alternatives (i.e., $H_H$ (n.m) = $H_H$(n) + $H_H$(m)), and (iii) normalization. Thus, its proper definition is:

$$H_H(n) = \log n \tag{15}$$

*Shannon Information*

However, Hartey and Nyquist's conceptual proposals will hardly have any immediate influence (Gleick, 2010, p.202), until their recovery two decades later by Claude Shannon in his theory. He developed during the 1940s the pivotal concept of 'information-theoretical entropy' as a quantitative measure of information, originally presented in his groundbreaking paper ‘A Mathematical Theory of Communication’ (1948). As an inquiry field with a domain-restricted object of study, Shannon's (1948) communication theory or (as popularized in the early-1950s) information theory emerged to statistically evaluate the efficiency with which messages can be encoded and transmitted in different signal-transmission or communication models, based on this author's early work (Shannon 1945) on enhancing the efficiency of message-encryption and message-decryption procedures during World War II (see Kline 2015, Section 2). To define its proper domain of application, Shannon displayed on the second page of his article a diagram illustrating all the basic elements of his communicative model and their operation using arrows.

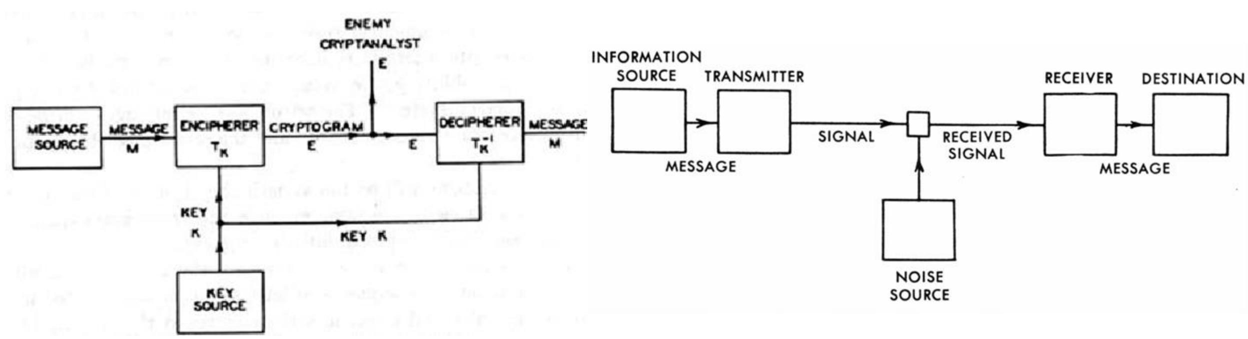


**Figure 2.** Original diagrams comparison between Shannon’s 1945 Enciphering-Deciphering Model (Left) and his 1948 Communication Model (Right) displaying conceptual evolution in IT.

First, we encounter the information source S, which is the element that selects a particular message m or sequence of IT-states or ‘symbols’ (m = $s_1$, $s_2$, ..., $s_i$) produced by the source from a set of possible symbols or alphabet. Next, this message s is encoded via signals by a transmitter (or ‘encoder’) T, who sends it over a particular communication channel C. Such a communication channel may be 'noiseless' or 'noisy', depending on whether C possesses a certain amount of noise that may alter the signals transmitting the message. Subsequently, the signals reach the receiver (or ‘decoder’) R, who decodes the signals and reproduces the message m, finally ending up in the destination D. So much for the core of Shannon's (1948) model. Nothing else is specified about the nature of each of these elements, for example, whether the source-recipient or the transmitter-receiver have an agent-centered character, presupposing a sort of epistemic agent. Assuming this background model, Shannon developed the technical concept of the entropy H (also '$H_{IT}$' or 'Shannon information') of the source H(S) to statistically characterize the capacity required by the communication channel CH to successfully transmit the sequence of symbols $s_i$ by the information source S, defining on the probability of occurrence $p(s_i)$ of each of the individual symbols and employing a logarithm function in base 2:

$$\mathrm{H}(S) = -\sum_{j=1}^{m} \mathrm{p}(s_i)\log(\mathrm{p}(s_i)) \tag{16}$$

That is, the more improbable the sequence of symbols selected m, the higher the value of H(S). For example, if we consider the English language as source S, then H("sky") < H("xzy") due to the higher frequency of occurrence of 'sky' than of 'xzy'. $H_{IT}$ was supposed to be the only function satisfying the following axioms: (i) H is continuous over the probability distribution ρ defined over the set of possible symbols, (ii) H is additive for any two independent messages, i.e., $H(m_1 + m_2) = H(m_1) + H(m_2)$, (iii) H increases monotonically for n alternatives, (iv) H is independent on how messages are grouped, and (v) H is bit-normalized, i.e., $H(\frac{1}{2}, \frac{1}{2}) = 1$ (Frigg and Werndl 2011, p.118; Klir, 2006, Section 3.2.2). Following Hartley's (1928) proposal, the use of the logarithm function in (12) (via Stirling's approximation of the number of possible messages, i.e., $2^{NH(S)}$) is pragmatically convenient in the development of this concept because of its ability to mathematically compress the size of the possible messages that can be generated by the source, where certain important technical parameters vary linearly with respect to this function. (Shannon, 1948, p.349; Lombardi et al. 2016b, p.1986). The base 2 of the logarithm in (16) is precisely what allows this quantity to be expressed in binary units or 'bits' (a term attributed by Shannon [1948] to John Tukey in 1947, although it is historically debatable, see Gleick 2011).

Interestingly the technical concept of Shannon information has certain particular properties. Firstly, note that Shannon entropy of the source H(S) does not straightforwardly quantify the amount of information contained (if this 'contained' can make sense in this context) in an individual message m, but the amount of technical information generated in the selection of the symbols of the message m regarding the entire source S. Although this should be interpretatively specified, the main reason underlying its 'ensemble-based' constitution might be that "Shannon didn't focus on individual amount of information, because he was interested in solving problems related with the transmission of any message in technological situations" (Lombardi et al. 2016b, p.1991). As Lombardi et al. (ibid.) point out, although $H_{IT}$ does not quantify m-based quantities of information, it would in principle be possible to derive the amount of $H_{IT}$ generated in the occurrence of individual symbols $s_i$:

$$\mathrm{I}_{IT}(s_i) = -\log(\mathrm{p}(s_i)) \tag{17}$$

The second conceptually interesting feature of H(S) resides precisely in the fact that (16) it does not depend on the meaning of individual symbols $s_i$ or sequences of symbols m, as Shannon himself defended "Frequently, the messages have meaning; that is they refer to or are correlated according to some with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem" (Shannon 1948, p.3). Third, it should be noted that information-IT is constitutively independent of the way in which the information units are physically implemented (mainly, letters from the source and bits) and therefore conceptually detached from any physical theoretical domain. In this sense, as suggested by Lombardi et al. (2016b). It would not be technically correct to speak of Shannon's theoretical proposal as a 'classical information theory' analogous to the quantum information theory articulated by Schumacher (1995) almost half a century after Shannon (1948). Shannon's IT is not a physically 'classical' theory precisely because it is theory-neutral (Lombardi 2016b), since their concepts like H can be equally applied to both classical and quantum physical systems.

Apart from the source-based $H_{IT}(S)$ concept, IT also allows to define an analogous notion for the information received by the destination D. In the same way as with H(S), we can calculate the amount of entropy H(D) generated during the reproduction of the message m by the destination D, if we consider the probability $p(d_i)$ of each individual symbol $d_i$ received by the destination:

$$\mathrm{H(D)} = -\sum_{j=1}^{m} \mathrm{p}(d_i)\log(\mathrm{p}(d_i)) \tag{18}$$

Once both entropy quantities H(S) and H(D) have been defined, we can analyse the statistical dependence between them by means of the concepts of (i) mutual information H(S;D), amount of information generated by source S and received by D; (ii) equivocation E, or information produced by S but not received by D; and (iii) noise N, amount of information received by D but not produced by S (see Lombardi et al. 2016). Departing from these elements, Shannon (1948) defined the channel capacity C as the maximum amount of mutual information H(S;D) that can be communicatively transmitted through this channel: this is the so-called ‘Noisy-Channel Coding Theorem’. Apart from the Noisy-Channel Coding Theorem, the most important theoretical result of IT is the Noiseless-Channel Coding Theorem, which states that H(S) measures the optimal compression of messages *m* (i.e., into sequences of binary units) to be transmitted over noiseless channels. In this direction “Shannon information is, then, appropriately called a measure of information because it represents the maximum amount of messages consisting of letters produced by a source [S] can be compressed” (Timpson 2013, p.22).

### *Interpreting Shannon Information*

Following Lombardi et al. (2016b), despite the technical precision with which its meaning is defined, the technical-quantitative concept of $H_{IT}$ constitutes an enormously fruitful interpretative landscape. These authors defended the existence of three interpretative strategies with respect to this concept, namely the formalist, the epistemic and the physical. As I will argue below, this classification of interpretative lines has no satisfactory applicability within the fine-grained analysis of the multiple descriptive tactics underlying information physics precisely because of the extremely complex interweaving of interpretative positions. In fact, a myriad of authors such as

Brillouin (1956), Tribus and McIrvine (1971) or Caves (1990) interpreted information in a simultaneously epistemic and physical way, but all of them defended substantially different ways of understanding $H_{IT}$. Beyond this point, what Lombardi et al. (2016b) call (and even defend) 'formal interpretation', historically developed by Khinchin (1957) and subscribed by authors such as Reza (1961) or Cover and Thomas (1991), is an interesting way of specifying this concept of information as a mere syntactic tool without specified semantic content: “the formal view endows the concept of information with a generality that makes it a powerful formal tool for science. (…). It is not only that messages have no semantic content, but that the concept of information is a purely mathematical concept, whose “meaning” has only a syntactic dimension. It is precisely from its syntactic nature that the generality of the concept derives” (Lombardi et al. 2016b, 2007).

Beyond these coarse-grained interpretative strategies, we can find conceptual regions of IT in which the interpretative exercise is decisive in order to give a satisfactory account of its theoretical content. Contrary to what is widely assumed (e.g., Dretske 1981, Chapter 2; Kolmogorov, 1983, p.29) and theoretically similar to what happens in GSM, one could be prima facie interpretatively assume that the quantity H(S) semantically refers to individual quantities of information (formalized as $I(s_i)$) if the latter concept is consistently interpreted as quantifying averages of information, as was suggested by Lombardi et al: “it is essential to the interpretation of the concept of information: in this context it is important to decide whether $I(s_i)$ measures information and H(S) is a weighted average, or the concept of information is directly embodied in H(S) and the talk about the information carried by individual messages makes no sense. The difference between these two positions is strongly related with the way in which information is conceived: if primarily as what is transmitted in a situation of communication, or as a measure of the optimal compression of the source’s messages” (Lombardi et al. 2016b, 1992). That is, these authors locate an interpretative dilemma at the theoretical heart of IT: either we orthodoxly consider that Shannon's meaning information is based on source-dependent 'informational ensembles' (analogous to GSM-ensembles), or we develop an IT-concept of 'individual-state information' $I(s_i)$ derived via averaging the values of $H_{IT}(S)$ per individual states $s_i$. The latter averaging procedure (AP) technically depends on the system (communicative or otherwise) being ergodic, which is guaranteed for the Marvokian processes considered by Shannon. In fact, the following technical concept of information that I will explore can be considered as an 'external' way (since it depends on developing an additional theoretical apparatus to IT) of solving this dilemma.

### 2.4.3. Algorithmic-Complexity Information

Although the notion of Shannon entropy constitutes de facto the most decisive concept of technical and quantitative information, some authors tried to overcome certain technical and theoretical deficiencies of this IT-concept by developing a new ‘computationally-enhanced’ concept of information. This was the case of what is known as algorithmic information theory (AIT), from which the informational concept of 'algorithmic complexity' or 'Kolmogorov complexity' was configured by integrating Shannon’s IT and certain technical elements of the theory of computation, such as the notion of 'Turing computability' (i.e., a sequence s is Turing-computable if their symbols can be generated by an abstract machine without halting). Interestingly, AIT was developed independently by three scientists to satisfy significantly different theoretical goals. On the one hand, Ray Solomonoff (1960) sought to develop an apparatus that would account for inductive inference (in a way analogous to what Carnap and Bar-Hillel intended, as we shall see below) by combining a Bayesian framework and assuming Occam's razor as a theoretical principle. Starting from certain results in computer theory, the young student Gregory Chaitin (1966) arrived at the same theoretical results as Solomonoff (then unknown) and the famous Russian mathematician Andrei Kolmogorov (1965). In his ‘Three Approaches to the Quantitative Definition of Information’ Kolmogorov (1965) distinguishes between the two technical concepts of information (i.e., combinatorial and probabilistic, information-IT being of the latter type) and proposes to develop a new concept that would make it possible to quantitatively and technically specify the information conveyed by a single object in terms of its complexity or randomness:

> “Our definition of the quantity of information has the advantage that it refers to individual objects and not to objects treated as members of a set of objects with a probability distribution given on it. The probabilistic definition can be convincingly applied to the information contained, for example, in a stream of congratulatory telegrams. But it would not be clear how to apply it, for example, to an estimate of the quantity of information contained in a novel or in the translation of a novel into another language relative to the original. I think that the new definition is capable of introducing in similar applications of the theory at least clarity of principle” (Kolmogorov 1983, p.29).

In the theoretical context of AIT, to determine the information of a particular sequence of symbols or message $s_i$ (in classical terms of Shannon's theory, although one could also consider non-sequential data structures) it would be necessary to measure its algorithmic complexity or Kolmogorov complexity K, defined as the minimum number of symbols $s_i$ that is necessary to reproduce that exact sequence of symbols *m* by means of a Turing machine P (or universal computational program); or in other words, the algorithmic information K(*m*) of a particular sequence *m* corresponds to the length of the shortest program that can replicate that sequence:

$$K(m) = |P_{min}(m)| \tag{19}$$

That is, the algorithmic information of a sequence s is inversely proportional to the ability of some conceivable program P to compress the original sequence *m* into a new shorter sequence *m'*. Of course, whether a sequence can be effectively compressed depends on the regular patterns we find among its encompassing symbols $s_i$. In order to detail the operation of this informational concept, let us illustrate its applicability in a case by comparing the sequence *m* with the sequence *t*.

$$m = 19191919191919191919191919191919191919191919191919191919191919$$

$$t = 43138998924098346809238$$

What is interesting is that, although s contains a significantly larger number of numeric symbols than t (in technical notation $|t| < |m|$, wherein $|t| = 23$ and $|m| = 62$), this sequence s can be effectively compressed by an algorithm or program $P^m$ into a much shorter sequence such as

$$m' = \mathrm{W19X31}$$

wherein ‘W’ stands for ‘write’ and ‘X…’ represents ‘n-times’. In this case, the fact that '19' is repeated 31 times in s (this can be characterized as a 'pattern' in s) is exploited by the $P^m$ program to compress s by means of t. Therefore, according to definition (14) the algorithmic information contained in this single sequence will be $K(m) = |P^m{}_{min}(m)| = 6$. However, prima facie there is no $P^t$ algorithm that allows us to compress the sequence t into a new shorter sequence, therefore, *t* would constitute an incompressible sequence (i.e., the shortest possible sequence that allows us to reconstruct *t* is t itself). According to this computationally-enhanced informational concept, the greater the randomness in the considered sequence, the greater its algorithmic information. The fact that *t* is incompressible (with respect to the technical precepts of AIT) due to the apparent randomness with which the symbols appear in t means in this theoretical context that *t* possesses a high amount of algorithmic information, which is quantitatively determined by $K(t) = |P^s{}_{min}(t)| = 23$. Thus, although s has almost three times as many symbols as *t*, the individual sequence t contains a higher degree of algorithmic information than s because it is algorithmically incompressible, i.e., $|t| < |m|$ but $K(m) < K(t)$.

### 2.4.4. Theoretically-defined Semantic Information

If the AIT concept of algorithmic information arose to solve the theoretical problem that IT could not define with technical precision the amount of information possessed by individual objects (Kolmogorov 1983, p.29), the lack of semantic character of IT-entropy motivated the development of new technical concepts of information that incorporated a certain sensitivity to the semantic properties of sequences or other types of structures. I should remark that here we must distinguish between the main concepts of semantic information of a strictly technical (though not necessarily quantitative and/or statistical) nature that we find in the literature and the ordinary non-technical concept of 'information'[27].

One of the first semantic information theories (SIT) was developed by the philosophers Rudolf Carnap and Yehoshua Bar-Hillel (1952) a few years after Shannon's proposal was published. In their 'An Outline of a Theory of Semantic Information' the authors suggested that their main motivation was that the “Prevailing theory of communication (or transmission of information) deliberately neglect the semantic aspects of communication, i.e., the meaning of the messages” (Carnap and Bar-Hillel 1952, p.1). Based on the difference between 'information content' and 'amount of information', Carnap and Bar-Hillel developed a semantic-IT proposal in connection

[27] “The everyday concept of information is closely associated with the concept of knowledge, language, and meaning; and it seems, furthermore, to be reliant in its central application on the prior concept of a person (or, more broadly, language user) who might, for example, read and understand the information; who might use it” (Timpson 2013, p.11)

with their general framework of inductive logic as a foundation of probability, where the object of application of their concept of 'information content' would not be the sequences of symbols or the wave signals of IT but properly the propositions expressed by declarative statements as vehicles of semantic information. The authors argued that this restriction was not problematic, since any other possible informational vehicle could be characterized in propositional terms[28].

It should be noted that their construction of the concept of 'information content' does not depend on formally modifying the notion of IT-entropy, but on technically specifying the ordinary notion of information mentioned above: "We attempt to explain the pre-systematic concept of information, even as it is applied to sentences or propositions" (ibid., p.3). Their conceptual proposal is as follows. Each proposition *i* asserting that state Z is the case is associated with state descriptions compatible with Z being the case. For any proposition *i* (subject to the Boolean algebra of propositional logic) involving state Z, its set of state descriptions or 'range' is denoted by R(*i*), where (illustratively) R(*i*) and R(¬*i*) contain no state in common. Thus, Carnap and Bar-Hillel (1952, p.14) postulated the existence of a measure called 'proper m-function' m(Z) which encodes the degree of probability that state Z is the case, assuming certain basic conditions such as (a) unity, $m(i) + m(\neg i) = 1$, or (b) additivity of m under the conjunction of independent propositions. From this m-function the authors defined their concept of 'information content' as

$$\mathrm{cont}(i) = m(\neg i) = 1 - m(1) \tag{20}$$

where the semantic information of a particular proposition *i* is determined by the degree of probability that the state Z implied by *i* is not the case. That is, the more improbable the state Z implied by *i* is, the more semantic information content *i* will possess. As regards their quantitative concept of 'amount of information', this was defined by the logarithm function of the ratio between the total number of state descriptions N and the number $N_R$ of state descriptions associated with the rank R(i) of a particular proposition *i*,

$$\inf(i) = -\log_2(N_R/N) = \log_2(N_R/N) \tag{21}$$

Unlike Carnap-Bar-Hillel's (1952) proposal, technically specifying the ordinary notion of information by means of the technical tools of inductive logic, we can also find in the literature other SITs that defend a direct reconceptualization of entropy-IT, seeking to give it a semantic dimension. In this strategic direction, the physicist Donald MacKay (1969) sought to include in his ambitious information theory the concept of Shannon's IT-entropy, redefining it as "selective information", precisely because its value depends on how the units-symbols with which the messages to be transmitted are 'selected' from the source S. Recognizing this informational measure as merely syntactic, this author will also argue in the Chapter entitled 'Information, Mechanism, and Meaning' (MacKay 1969, Chapter 3) that the so-called descriptive-information concepts of 'metric information' (i.e., Fisher's [1935] measure of statistical information defined

[28] "The restriction of the range of application of the concepts to be explicated to sentences (or propositions) is probably not serious, since other applications seem to be reducible to this one". (Carnap and Bar-Hillel 1952, p.1).

as the reciprocity of a parameter O about a variable X) and structural information (i.e,. Gabor [1946] measure of physical information based on minimum units of phase-frequency-time volume) possess a semantic character derived from the intentionality of these agent-observers "It appears from our investigations that the theory of information has a natural, precise, and objectively definable place for the concept of meaning" (MacKay 1969, p. 93). However, MacKay seems to suggest (although it is not very clear) that the semantic character of these concepts and measures of information would depend mainly on the capacity of the agent-observers to employ them as representational tools of the objective phenomena.

Undoubtedly, one of the most influential semantic-IT within the philosophical literature is that developed by Fred Dretske (1981) in *Knowledge and the Flow of Information*, wherein he defended a naturalistic theory of human perception based closely on Shannon's theory of communication. Although this author recognizes from the beginning the Shannonian dictum that its entropy measure does not capture the plausible meaning of a sequence of symbols, it is also part of Weaver's suggestion that certain engineering considerations might be relevant to measuring certain semantic relationships (Shannon and Weaver 1949, p.1). To this end, communication theory should be technically and theoretically supplemented to derive semantic relations from the statistical correlations between the occurrence of symbols and the occurrence of events: "the underlying structure of this [communication] theory, when suitably supplemented, can be adapted to formulate a genuinely semantic theory of information" (Dretske 1981, p.x). As we pointed out at the beginning, one of the main reasons why Shannon's communication theory is insensitive to the semantic content of messages is precisely because its entropy measure not a measure of the amount of information of a particular sequence of signs (as one would expect from a measure of semantic information) but a statistical ensemble of possible messages. Dretske (1981) proposes to develop a measure $I_S$ of the information transmitted by an individual event $y_i$ (although the author prefers to speak of 'signals' and 'states of affairs' to refer to the events they refer to and to the events referred to, respectively) to solve this theoretical problem, representing the amount of information that a sequence of particular events $y_j$ transmits about another event or particular state of affairs $x_i$, thus assimilating the intentional character or 'aboutness' that must characterize any robust semantic information concept. This measure could be formulated by the following mathematical expression

$$\mathrm{I}_S(s) = -\log p\,(s_i) - \mathrm{H}(p(t_i s_i)) \tag{22}$$

This measure would represent the amount of information generated by the existence of certain signal-event $s_i$ about the occurrence of other event $t_j$, which is determined by the logarithm of the improbability of occurrence of the signal-event $s_i$ minus the $H_{IT}$-encoded unexpectedness of the conditional probability distribution between the event referring to $s_i$ and the referred event $t_i$ (Dretske 1981, Section 3; Timpson 2013, p.39). Therefore, his measure of semantic information (17) would strongly depend on the non-semantic information measure of Shannon’s IT-entropy, deriving semantic relations from the statistical correlations between the two events. With this measure of the amount of information contained in a single event, Dretske would have the necessary resources to define the semantic information that is contained in a particular signal S about or about a certain event *q*:

$$\text{A signal } s_i \text{ contains the information that } q =_{def} p\,(q \mid s_i) = 1 \tag{23}$$

That is, an event that acts as a $s_i$ signal (that is, the referring event) would contain information about another *q* event (or target event) if there is a complete correlation between the signal $s_i$ and the referenced event *q*. This would capture the modal amount of semantic information; namely, that each time the signal $s_i$ occurs, it necessarily depends on the fact that the event *p* to which it refers has also occurred. Dretske (1981) explains this modal feature by claiming that the event referred q acts properly as the cause of the signal $s_i$, wherein $s_i$ in turn constitutes the effect of that causal connection. The fact that a signal $s_i$ has semantic information about *q* is based on the existence of an asymmetric causal relationship between *q* and $s_i$, which for reasons of extension I will not go into further. Then, according to this Dretskean proposal, the semantic information would thus constitute a statistical (and relative) measure derived from the statistical correlation between events, where its quantity is statistically determined by the conditional probability in (23): while the value 1 would represent that the signal $s_i$ has the maximum semantic information that q, the value 0 would represent that the signal $s_i$ has no information that *q*. Finally, it should be mentioned (as I will detail in Section 5.2.2) that there are authors such as Manolo Martínez (2019) or Alastair Isaac (2019) who have recently vindicated the possibility of developing technically-robust semantic information concepts from Shannon-IT beyond the Dretskean program.

## 2.5. Connection Atlases on Entropy, Information and Probability Concepts

Previously, I have explored the conceptual landscapes around the main notions of entropy, probability and information in various theoretical scenarios, focusing on classical thermal physics and obviating their appearance in fields such as computation or dynamical systems theory. Having completed this initial task, I should now systematically trace the main connections that we find in these local conceptual and interpretative domains that I have just delimited, emphasizing those that will have a certain relevance in the theoretical architecture of information physics. I will now focus on neutrally assessing the unfolding of 'positive' conceptual relations (i.e., how two concepts connect) rather than 'negative' ones (i.e., in what ways two concepts do not relate), since this task will be properly carried out in Chapter V, after having analysed the historical development of this intellectual trend during Chapters III, and IV. In this sense, the assessment carried out in this Section 2.5 won’t be critical but just expositive, since my properly critical task is going to be fully developed in Chapter V. As in the previous sections, the construction of 'atlases', understood as sets of conceptual maps making explicit their relations, will also depart from the main theoretical results and consensus generated in the recent philosophy of physics on those essential topics.

### 2.5.1. Framing Conceptual Relationships

Before starting this analysis, I should preliminarily differentiate between different types of relationships between technical concepts (which would not include prima facie 'ordinary order' or 'ordinary information') according to the conceptual dimension they encompass. The conceptual connections most relevant to this task are (i) 'superficial' (terminological or formulaic) relations between their representational formats, (ii) numerical relations between their derived results, (iii) semantic relations between their meaning, and (iv) interpretative relations between the ways of understanding their content. Let us delve minimally into these.

First, two technical concepts C and C' possess a terminological and/or formulaic relationship if both are expressed by the same technical term and/or by the same mathematical formula, or at least by sufficiently similar terms and/or formulas. Illustratively, the concepts of 'Gibbs entropy' and 'Shannon entropy' possess both a terminological connection (both contingently share the term 'entropy') and a formulaic connection (both share a similar mathematical expression). Although this kind of conceptual connection might somehow seem trivial or philosophically uninteresting, it has had a pivotal role in the historical unfolding of the physics of information. Second, two technical and quantitative concepts C and C' possess an arithmetic relationship if by algebraic manipulation of both it is possible to computationally derive the same numerical values. For example, the concepts '$S_G$' and '$S_B$' have an arithmetic connection if by using both it is possible to derive (by means of phase-averaging or discretization procedures, for instances) the same numerical results as Joules/Kelvin degrees with respect to the same SM-model. In their 'A Guide for the Perplexes' Frigg and Werndl (2011) exhaustively analysed the existing numerical relationships within the landscape of entropy concepts.

Thirdly, two (usually technical-quantitative) concepts C and C' possess a semantic relation if (a) both employ the same representational resources to refer (co-intensional) or if (b) both refer to the same magnitude or property (co-referential). For example, the concepts Boltzmann combinatorial entropy and Boltzmann entropy possess a semantic connection, because (b) both refer to the same possible values of position and velocity of kinetic systems (they are co-referential), but (a) they do so by means of different representational resources (then, they are not co-intentional): namely, arrangements D in μ-space and microstates in Γ, respectively. Of course, one could find many different kinds of (partial) semantic relationships between scientific concepts and their representational resources, mainly (i) idealizations (i.e., deliberate distortion of certain descriptive resources of C by C'), (ii) abstraction (i.e., omission of properties of C by C') (see Frigg and Hartmann 2006), and (iii) approximations (inexact use of representational resources of C by C') (see Norton 2012). In philosophy of physics (unlike in physics itself) two notions C and C' have been usually considered as conceptually identical not if they are numerically related, but mainly if they possess a 'strong' (or complete) identity-like semantic relation, either via (i) or (ii) (e.g., Carnap 1977). This is the case of Hemmo and Shenker, who argued that the concepts '$S_{TD}$' and 'von Neumann entropy' (i.e., the quantum analog of Gibbs entropy) were numerically related, but not conceptually identical precisely because of their lack of ‘strong’ (or ‘complete’) co-referential semantic relationship: “Identities of physical properties mean that the two quantities refer to the same magnitude in the world. We have shown that for finite gases this identity of reference does not hold, and in the infinite case where the argument hold (arithmetically) the system in question is not physical” (Hemmo and Shenker 2006, p.172).

Finally, due to the enormous complexity of this field of study and the central role of interpretation in contemporary physics, we must also consider certain fine-grained conceptual interrelationships associated with the possible interpretative strategies of concepts. In this sense, we would say that a concept C is fine-grained conceptually connected with the concept C' by means of an interpretative strategy I if the specification of the representational resources and reference of C by means of strategy I allows us to fix the representational resources and reference C'. Illustratively, the concept of 'combinatorial Boltzmann entropy' is fine-grained conceptually connected with 'Boltzmann entropy' via an agent-based interpretation I precisely because that interpretative specification I of the latter concept as a measure of the agent's inability to infer the actual μ-arrangement D of the system would allow us to further specify the meaning of the former concept as quantifying the agent's inability to infer the actual microstate *x* of the system within a macrostate-region $\Gamma_M$. Once these relationships have been minimally detailed, let us move on to explore the main existing connections of the first conceptual domain.

### 2.5.2. Connection Atlas over Entropy Concepts

If, as mentioned above, entropy is one of the most difficult scientific concepts, the clarification of the connections between the various concepts of entropy is one of the most important tasks within the philosophy of thermal physics. The importance of performing the analysis of conceptual relations between different classical thermophysical entropy notions lies precisely in the fact that, historically, the theoretical validity of SM-entropies has depended on their possible connections with the empirically meaningful notion of $S_{TD}$. However, different schools (without distinguishing between scientific or philosophical) of physical thought have prioritized certain types of connections over others, subordinating them to different theoretical claims such as the reduction of TD in SM, as I proceed to analyse.

On the one hand, proponents of the GSM-apparatus have argued that the theoretical validity of $S_G$ depends on its numerical connection with the $S_{TD}$, because by means of Gibbsian coarse-graining we can adjust the numerical values so that the ensemble entropy increases without exception during a process of approaching equilibrium. This is the case of Jaynes (1965), who not only criticized the Ehrenfests (1912) defense of $S_B$ as the 'correct' entropy, but also argued that $S_G$ coarse-grained was the correct entropy because of its numerical connection with the $S_{TD}$ “Gibbs formula gives the correct entropy (…) objections to the Gibbs [coarse-grained entropy] are immediately refuted by the fact that the Gibbs canonical ensemble yields correct thermodynamic predictions” (Jaynes 1965, p.391). However, such arithmetic-driven arguments will be criticized from the philosophy of physics (e.g., Sklar 1993), in general, and from the Neo-Boltzmannian trend at the end of the $20^{th}$ century, in particular. In the latter direction, pro-BSM authors such as Callender will argue that the mere numerical-syntactic relation between the predictions derived from $S_G$ and $S_{TD}$ do not imply a meaningful conceptual connection between the two "SM need not and should not search for a mechanical counterpart for entropy that exhibits monotonic behavior with time. For when we find such an entropy, we immediately know that it is irrelevant to the behavior of real individual systems in the world" (Callender 1999, p.367).

On the other hand, philosophers of physics and Neo-Boltzmannians have claimed the theoretical primacy of $S_B$ over $S_G$ due to the semantic connections of the former with $S_{TD}$, because $S_B$ is a quantity referring directly to individual systems (likewise TD) via individual microstates *x*, while $S_G$ refers to properties not of individual physical systems but of ρ-encoded fictitious collections or ensembles (differently from TD). Similarly, while the equilibrium macrostate in BSM refers to a property of individual physical systems, the statistical equilibrium state (both fine and coarse-grained) is a concept referring to a statistical property of ensembles (Frigg 2008). However, the concept of $S_B$ has problems when numerically replicating the results of certain TD-predictions: namely, for systems approaching equilibrium, $S_{TD}$ would necessarily increase until it reaches its maximum numerical value in the equilibrium state (thus following the Second Law), i.e., $S_{TD}(t_1) \geq^* S_{TD}(t_0)$; while in the case of $S_B$ there is a tiny probability that its value will fluctuate below its maximum value once the system reaches equilibrium, i.e., $S_B(t_1) \geq^* S_B(t_0)$ wherein $\geq^*$ means 'more or equal' with certain exceptions. That SM-entropies trade-off between having a robust semantic connection or a correct numerical connection to the $S_{TD}$ concept is what is known in the literature as 'Sklar's dilemma' (e.g., McCay 2020), as it was Lawrence Sklar (1993) who first pointed it out in his foundational work *Physics and Chance*. This balance is crystal clear in this quote:

> "Thermodynamics states that once an isolated system achieves equilibrium, it stays in equilibrium forever (until disturbed by external in- fluences). Gibbsian SM agrees. Gibbsians suppose that the probability distribution characterizing equilibrium is time independent. Boltzmannian SM disagrees, however, as it abandons the idea that equilibrium is stationary in time. The Boltzmann approach balances this affront to thermodynamics by retaining the idea that equilibrium and entropy are properties of individual systems. The Gibbs approach pays for its strict agreement with the thermodynamic laws by relinquishing the idea that entropy and equilibrium are properties of individual system" (Callender 1999, p.357)

Beyond the conceptual connections in Sklar's dilemma in equilibrium approximation contexts, we could also find relations between the thermophysical concepts of entropy in equilibrium scenarios. As Frigg and Werndl (2011, p.130) pointed out, the multiple numerical connections between entropic measures must be evaluated on a case-by-case basis[29]. In particular, from the vast amount

[29] "A [numerical] relation between $S_{G,f}(\rho)$ and the TD entropy can be established only case by case. $S_{G,f}(\rho)$ coincides with $S_{TD}$ in relevant cases arising in practice. For instance, the calculation of the entropy of an ideal gas from the microcanonical ensemble yields [Sackur-Tetrode equation] – up to an additive constant (Kittel 1958, p. 39)". (Frigg and Werndl 2011, p.130).

of results in SM and TD (Kittel, 1958; Reiss, 1965; Emch & Liu, 2002) we can derive the existence of a 'sweet spot' between the main three thermophysical entropies, where the numerical values of these three measures can come to coincide simultaneously (of course, this also applies for the numerical binary relationships between $S_{TD}$ and $S_B$, $S_{TD}$ and $S_G$, $S_B$ and $S_G$). This 'Golden Entropy Triangle' is determined by the following conditions: (i) the application-target (or reference) should be an ideal monoatomic gas of non-interacting identical hard-spheres molecules in thermal equilibrium, (ii) for $S_G$ a microcanonical ensemble is employed, (iii) the macrovariables $V_0$ and $T_0$ of the gas are fixed with respect to an energy value E (see Frigg and Werndl 2011, p.130). Under these particular conditions, both $S_B$ and $S_G$ are numerically equivalent to the Sackur-Tetrode formulation of $S_{TD}$ for the system under consideration:

$$S_{TD} = nk \log\left(\left(\frac{T}{T_0}\right)^{\frac{3}{2}}\frac{V}{V_0}\right) \tag{24}$$

wherein $T_0$ and $V_0$ encode the temperature and volume of the system at $t_0$ in a quasi-static process. However, I should finally remark that the domain of conceptual application delimited by this 'Golden Entropy Triangle' is, independently of the degree of idealization it implies (i.e., zero interactions or perfectly elastic), exclusively restricted to certain noble (e.g., Helium, Neon, Argon, etc.) or non-noble (e.g., oxygen, nitrogen) gases under ideal conditions such as high ambient temperature and low pressure. Then, one should conclude from the above argument that the only possibility to connect in a non-interpretative way (arithmetically) the three main thermophysical entropy concepts would be strongly restricted to the physically insignificant field of a tiny set of real gaseous substances under extremely particular experimental conditions.

### 2.5.3. Connection Atlas over Information Concepts

After charting the core conceptual connections between entropy notions, the next task is to do the same within the landscape of information concepts. The importance of this task for the philosophy of physics/information and with respect to the goal of delimiting the conceptual architecture of information physics lies, on the one hand, in the ability of technical measures of information to refine certain properties (i.e., intentionality, agentiality, and semantic-epistemic character) of OI, and on the other hand, in how the notions connect with the interdisciplinarily successful and popular concept of $H_{IT}$.

First, let us explore the central conceptual connection between OI and information-IT, which will play a central role in my historical-philosophical analysis. Beyond the contingent and conventional choice of the common term 'information' and because the former (OI) is not a technical-quantitative concept, there are no possible arithmetical connections between the two, although there are (again, contingently) terminological ones. Due to (i) the lack of semantic character of $H_{IT}$ (Shannon 1948), (ii) the independence of agents and (iii) the direct reference of $H_{IT}$ to the probability distribution of the source[30] S (at least in its original definition, ibid.), this concept of information-IT would not be semantically connected neither in reference nor representational resources (intension) with OI, having (i) a strong semantic content, (ii) based on agents, and (iii) refers to the content of proposition-like entities (Floridi 2011). If we follow the interpretative strategy of Lombardi et al. (2016a) to consider the amount of individual information I(pi) (we recall that this is derived by averaging $H_{IT}$) as an informational concept, then this would be semantically connected with OI in that (iii) both denote individual objects (messages and propositions, respectively). However, insofar as individual information $I_{IT}(p_i)$ is conceptually derived from $H_{IT}$, the former remains prima facie a concept without intrinsic semantic dimension (let alone without the robust semantic character that characterizes OI), so that its conceptual connection with OI is only partial.

However, some recent authors have argued that it is possible to trace a fine-grained conceptual connection between OI and $H_{IT}$ by means of an interpretative specification of the meaning of the former and the latter. This is the case of Timpson (2013, p.30), who argued that understanding OI as an everyday form of agent uncertainty regarding the probabilistic outcome of an experiment, then $H_{IT}$ meaning could be specified as a measure of ordinary information (proto-quantitative, as we shall see) in this first sense. However, this interpretative conceptual connection would only occur in a limited scenario in which a key condition is met: namely, that any possibility of an agent acquiring knowledge about (agency and semantic-epistemic character of OI) the probabilistic outcome of the source S is restricted entirely and exclusively to how the source S generates the outcomes[31]. An example is a scenario in which an agent B cannot employ his background knowledge to infer what is the missing letter of the sequence "hom_" beyond his knowledge of 'p(e)' concerning the source S. If we include in this interpretative strategy of Timpson's the concept of individual information $I_{IT}(p_i)$ derived by averaging (AP) from $H_{IT}$, then we obtain a 'sweet spot' conceptually connecting these two notions of IT and OI.

In this 'Golden Information Triangle' the meaning of the concepts $H_{IT}$, $I_{IT}$ and OI is interpretively specified as a proto-quantitative measure of the amount of ordinary information that an agent can obtain exclusively about an individual message ($I_{IT}$) or about the generation of possible messages

---

[30] "It is worth emphasising that [$H_{IT}$] is a technical conception of information which should not be taken as an analysis of the various senses 'information' has in ordinary discourse. In ordinary discourse information is often equated with knowledge, propositional content, or meaning. Hence 'information' is a property of a single message. Information as understood in information theory is not concerned with individual messages and their content; its focus is on all messages a source could possibly send. What makes a single message informative is not its meaning but the fact that it has been selected from a set of possible messages." (Frigg and Werndl 2011, p.119).

[31] "H(X) (etc.) *could* be understood in the advertised way in terms of the uncertainty/information link; and we *could* accept that quantity as sometimes measuring an amount of information in the everyday sense, but only *sometimes*. That is, when the scenario is pared down to the extent that one's only interest is the given experiment and what the probabilities of its various outcomes are: a situation when none of the other standard features involving what one learns, or what one might infer, or what one's various interests might be, are in play. In this little corner, it perhaps does no harm to grant H(X) (etc.) as measuring information in the everyday sense, as applied to this restricted situation. But one will note that this is far from *a representative epistemic scenario*." (Timpson 2013, p.30. Italics are mine).

($H_{IT}$). However, extending Timpson's (2013, p.30) of the OI-$H_{IT}$ vertex[32] to the whole area of the OI-$H_{IT}$-$I_{IT}$ 'Golden Triangle', the truth is that within this extremely limited conceptual region where this connection occurs: (i) OI ends up being an epistemically insignificant concept (i.e., the agent's OI is limited to the outcome of symbols) and (ii) both $H_{IT}$ and $I_{IT}$ end up being technically irrelevant measures (i.e., their function ends up being limited to measuring OI in IT scenarios). Consequently, the claim of Carnap-Bar-Hillel's (1952) or Floridi's (2011) SITs to conceptually refine the notion of OI immediately becomes meaningless, since their philosophical applicability is limited to certain unrealistic or epistemically insignificant cases.

Second, another major conceptual connection in the informational domain is that between $H_{IT}$ and algorithmic information K. Similarly to Hartley's $H_H$ being both formal and arithmetically a particular case of Shannon's $H_{IT}$ for uniform distributions[33] $\rho_\mu$ (analogously to the formal connection between $S_G$ and $S_B$ in SM), the notion $H_{IT}$ could be conceptually characterized as a formal-arithmetically particular case of K-averages for random distributions of source S. Based on this formal connection, some computationally-oriented authors (Zurek, 1990; Cover and Thomas, 1991, p.3) regard K as a more conceptually fundamental than $H_{IT}$. The arithmetical conceptual relation between K and $H_{IT}$ was historically popularized by Bennett (1982) and has been recently pointed out within the philosophical domain by Lombardi et al. (2016b, p.1997), who also assessed the semantic and interpretative dimension of this connection. Concerning the former kind of conceptual connection, H(S) would be equivalent to the K-average per symbol-state of the (randomly-selected) message $m_i$ and source-symbols $s_i$ under certain $\rho$ over $s_i \in S$:

$$\langle K(m_i)\rangle = \Sigma\, p(s_i)\, K(m_i) \approx H(S) \tag{25}$$

Apart from this arithmetic connection (see Lombardi et al. 2016b, p.1997), the way in which the meaning of K is established can help to fix the meaning of $H_{IT}$ as a measure not of the improbability of message transmission but properly of the degree of optimal message compression. From the latter authors' analysis one can infer the existence of a bifurcation of interpretative strategies on $H_{IT}$ depending on whether one intends to connect it conceptually with algorithmic information K or individual information $I_{IT}$, both defined over individual objects. It should be noticed that these two interpretative strategies are based on the same averaging procedure (AP), except that in the

[32] "Furthermore, one will note that as it stands, the everyday notion of information does not really admit of the kind of precise degree of weighting of amount that a quantitative measure (such as *H*(*X*)) would import. Judgements of amount supported by the everyday concept will generally be qualitative, partial and interest relative. Thus if employing a measure such as *H*(*X*) to measure uncertainty/information (in its little [Golden Information Triangle] where we may happily let it run free), we should bear in mind that we are *creating* an answer to a quantitative 'how much' question —an answer which did not exist before— rather than bringing to bear on the situation a finely tuned instrument which is finally able to reveal some pre-existing, but previously hard to measure, facts." (Timpson 2013, p.30).

[33] "On the face of it this entropy is based solely on the concept of mutually exclusive alternatives, and it does not invoke probabilistic assumptions. However, views diverge on whether this is the full story. Those who deny this argue that the Hartley entropy implicitly assumes that all alternatives have equal weight. This amounts to assuming that they have equal probability, and hence the Hartley entropy is the special case of the Shannon entropy, namely the Shannon entropy for the uniform distribution." (Frigg and Werndl 2011, p.121).

first case ($H_{IT}$ as message compression) it is performed on K-quantities and in the second case ($H_{IT}$ as message transmission averaging) it is performed on $H_{IT}$-quantities.

### 2.5.4. Connection Atlas between Entropy and Information Concepts

Having detailed the main connections between concepts of entropy and information, we are now in a position to state that the whole theoretical architecture of information physics rests fundamentally on the possibility of drawing robust conceptual connections between the notions of entropy and information. As I suggested earlier, this section is focused on exposing the main and most recent philosophical results on the existence of these intricate relationships, which will be indispensable both in my analysis of the historical unfolding (Chapter III and IV) and the conceptual foundations (Chapter V) of information physics.

First, one of the most important conceptual connections in this domain are those between OI and SM-entropies. Prima facie, this relationship is not terminological, formulaic or arithmetical, the latter two due to the non-technical non-quantitative character of OI. Nor is it semantic, since OI refers directly to agential, semantic and epistemic elements via non-technically defined representational resources (e.g., via ordinary language or mental imagery), whereas SM-entropies do not constitutively refer to such elements with their mathematically defined representational resources. However, we could develop an interpretive strategy that fixes the meaning of OI by means of representational resources of SM-entropies. This is the paradigmatic historical case of Brillouin, who argued that "[Boltzmann] entropy measures the lack of [ordinary] information about the actual structure of the system. This lack of [ordinary] information induces the possibility of a great variety of microscopically distinct structures, which we are in practice unable to distinguish from one another" (Brillouin 1962, p.160). Notice that Brillouin reconceptualized (implicitly, as I will defend below) OI by means of representational resources proper to BSM, namely: (i) it belongs to an epistemic agent (i.e., observer in SM) is presupposed, (ii) it possesses semantic and intentional character (i.e., information about the actual structure), and (iii) it constitutes an epistemic resource (i.e., it allows the agent to distinguish between different microstates). Thus, this BSM-based interpretative specification of the meaning of OI would allow to semantically connect this notion with an epistemic and agent-centric interpretation of $S_B$.

Second, the most explicit entropy-information conceptual connection that we can find once this conceptual landscape is delimited is that between $S_G$ and $H_{IT}$. As already mentioned, their relationship is initially terminological (both are more or less aptly named 'entropy') and subsequently formulaic, due to the fact that the mathematical formulation of $S_G$ and the continuous version of $H_{IT}$ is identical except for the logarithm base and the multiplicative constant k (See Frigg-Werndl 2011, P.129). In addition to this superficial connection between the two, we can consider $S_G$ and $H_{IT}$ to also possess an 'intensional' semantic relationship in the sense that they both rely on probability densities $\rho$ as essential representational resources (which also applies for $S_B$ in the case of uniform distributions). However, this is not constitutively an 'extensional' semantic relation, since SM-densities defined over $\Gamma$ refer (independently of their particular interpretation) to microstates of molecular systems and IT-densities defined over the space of possible messages refer to sequences of symbols to be communicatively transmitted. By means of an interpretative strategy such as the one carried out by Jaynes (1957a), we could interpretatively specify the meaning of the IT-densities so that, once redefined over the $\Gamma$ of a molecular system (instead of the state space of possible messages), these $\rho$ now refer to the system's microstates. In the same way, we can also interpret à la Frigg (2004) the system's $\Gamma$ as a state space of possible messages by introducing a Gibbsian partition, specifying each cell $\omega_i$ as a possible symbol $s_i$ and each trajectory as a message. By means of this kind of meaning-specification strategies, $S_G$ and $H_{IT}$ would end up interpretatively unfolding a semantic (and plausibly also arithmetic) connection not only between their $\rho$-based representational resources, but also between the reference of these.

Regarding the prima facie only terminological connection between $S_B$ and $H_{IT}$, the conceptual distance in terms of semantic relations (not only referential but also in terms of their representational resources) between the two is even greater than with $S_G$, precisely because the semantic capacity of $S_B$ depends directly not on distributions but properly on 'macrostates' on which a distribution is defined. In this sense, the possibility of consistently translating the BSM representational resources underlying $S_B$ into IT-resources would further entail interpretively justifying the partitioning of the state space of possible messages into macrostates by conceptual-interpretative IT-resources. Recently, Shenker (2020) developed an interpretive proposal that hints at a plausible semantic connection between $S_B$ and $H_{IT}$ representational resources and their references[34]. Shenker proposed to define $H_{IT}$ over a physical system's $\Gamma$, partitioning $\Gamma$ into different macrostate-regions $\Gamma_{Mi}$ whose microstates are associated with a particular symbol $s_i$ and whose probability measure $\mu(\Gamma_{Mi})$ is proportional to the probability of transmitting the symbol, so that the actual trajectory $\gamma$ of the system could be interpreted as the 'source' S generating the messages as it transits through different 'macrostate-regions' $\Gamma_{Mi}$ of the systems.

Having detailed some of the fine-grained interpretative connections that can be drawn between entropic concepts and informational concepts, I can now evaluate how the conceptual constraints that we found in a particular domain occur when we attempt to translate these relationships to another domain. It is at this point wherein those relationships emerge not between concepts but between conceptual relationships. One of the most illustrative examples of this sophisticated relational typology is found in the orthogonal connection between the semantic capacity of $S_B$ and

[34] "Consider a given physical system, with a given dynamics. This dynamics fixes (or just is) the set of trajectories in the state space. Suppose also that the microstates in the state space are partitioned to sets, such that the microstates in each set share a certain macrovariable; these sets are macrostates. If we take each macrostate (under the given partition) to stand for (or to be) an alphabetic symbol, then the physical system can be seen as a "source" that "sends symbols", in the sense of Shannon's theory. Each microtrajectory would be a particular message sent by the source along an ideally noiseless channel. And conversely, if we wish to send chains of symbols, we can build a system with suitable dynamics, such that, together with a suitable partition of the microstates to sets according to certain suitable macrovariables, each of the micro-trajectories of that system would generate a possible sequence of symbols" (Shenker 2020, p.16).

$I_{IT}$ or K to refer to individual objects (micro-trajectories of systems $\Gamma\gamma$ and individual sequences of symbols $s_i$, respectively) and the semantic capacity of $S_G$ and $H_{IT}$ to refer to statistical collections of objects (ensembles of possible microstate trajectories and ensembles of possible sequences, respectively). In this conceptual architecture, the interpretative possibility of both $S_G$ in GSM and $H_{IT}$ in IT to refer by means of these notions to individual objects (mimicking this referential property of $S_B$ and $I_{IT}$ or K, respectively) depends on the same theoretical mechanism, namely, the realization of phase averages over the statistical collections of objects to which they refer.

Finally, one might wonder about the possible relationships between the more conceptually distant notions, namely: $S_{TD}$ on the one hand, and $H_{IT}$ or even more radically, OI. As far as the first pair of concepts is concerned, we subscribe to the idea that "At first sight the information theoretic and the thermodynamic entropy have nothing to do with each other." (Frigg & Werndl 2011, p.3-4) except for the superficial terminological connection between the two. Even if we evaluate the concepts $S_{TD}$ and $H_{IT}$ in greater depth, we can conclude the inexistence of numerical or semantic connections (they do not share reference or representational resources) between them. Any possibility of connecting these two notions will be in an 'indirect' or 'conceptually mediated' way, for example, by arithmetically relating $S_{TD}$ and $S_G$ in the particular context ('Golden Entropy Triangle') delimited in Section 2.5.2, and then semantically connecting the representational resources of $S_G$ and $H_{IT}$ also in the particular context specified above. Insofar as we must situate ourselves in the Golden Entropy Triangle, this indirect conceptual connection between $S_{TD}$ and $H_{IT}$ will possess a remarkable degree of physical insignificance. Even more radically, any possible indirect conceptual connection between $S_{TD}$ and OI will eventually depend on finding a domain of conceptual application at the intersection between the Golden Entropy Triangle and the Golden Information Triangle. Therefore, in the remote case where such an intersection contains any non-trivial domain of conceptual application, it will simultaneously possess a remarkable degree of both physical insignificance (Section 2.5.2.) and epistemic irrelevance (Section 2.5.3).

### 2.5.5. Connection Atlas between Probability and Entropy-Information Concepts

Although the above manifold of connections between concepts of entropy and information is the master pillar of information physics, in turn the conceptual relationships and interpretative strategies concerning different notions of probability in SM and IT will be decisive in shaping this intellectual trend. Unlike the general type of conceptual relationships I have analysed so far, the connections involving probabilistic concepts are mostly interpretatively-driven for the reasons previously discussed in Section 2.3, which roughly implies that the semantic entanglement among the representational resources of concepts increases within this probabilistically-extended domain.

Initially, I should clarify that probabilities are not only conceptually connected with other probabilistic notions as well, such as SM-entropies or IT-information concepts. This is the case of OI, whose meaning can be connected once interpretatively specified with the main concepts of probability in SM and IT. First, probabilities refer by means of different representational resources to somehow objective properties (both individual $x$ and collective $\rho$) or to certain (particular or group) epistemic states $\varphi$ of scientific agents, depending on whether we specify the meaning of these probabilities by means of some objective or epistemic interpretation, respectively. Following Callender's (2011) argument[35], because the proposition-like content of epistemic states $\varphi$ intended to refer are ontological properties, in the end the concepts of probabilities refer directly (ontologically) or indirectly (epistemically). For example, $P(x) = 0.7$ can interpretatively represent either (a) the objective probability of $x$ being the actual microstate is 0.7; or (b) and epistemic state like $\varphi$ = 'my degree of belief in $x$ being the actual system's microstate is 0.7', whose content also refers to microstate $x$. Second, insofar as by means of epistemic SM-probabilities we can interpretatively specify their fine-grained meaning in such a way that certain probabilistic descriptions (i) presuppose an agent entertaining $\varphi$, (ii) possess the semantic-intentional character of mental states $\varphi$, and (iii) $\varphi$ provides knowledge to the agent; then this epistemically interpreted probabilistic concept will thereby semantically connected to the ordinary information concept OI.

Secondly, it is worth noting the conceptual distinction between the SM-probability notions GSM and dynamical probabilities (DP) and their SM-associated entropy notions $S_B$ and $S_G$. Their semantic disconnection was recently defended by Shenker (2020, p.16) in the case of $S_B$ and DP, who pointed not only to the difference in representational resources between the two (i.e., size of macrostates $\mu(\Gamma_M)$ and size of overlaps $\mu(\Gamma_\rho \cap \Gamma_M)$ but also to the empirical contingency of their numerical and referential connection through an objective interpretation of DP. That is, whether the measure $\mu(\Gamma_\rho \cap \Gamma_M)$ increases as the measure of the macrostate-regions $\mu(\Gamma_{Mi})$ in which the system's microstate $x$ might transit does not depend on the latter fact (on which $SP_B$ are defined), but on whether the ontically-interpreted frequency of individual trajectories $\gamma \in \Gamma_\rho$ ending in $\Gamma_M$ increases during the evolution of the systema[36]. If we move to the case of GSM, the previous 'intensional disconnection' pointed out by Shenker (2020, p.16) between $DP_B$ and $S_B$ does not occur at the same extend in the case of GSM (and its blob-like dynamic evolution) and $S_G$. This is because GSM and $S_G$ share $\Gamma$-defined densities $\rho$ (likewise BSM and $S_B$ share $\Gamma_M$) as their fundamental representational resources while $DP_B$ and $S_B$ don't have dynamical blobs $\Gamma_\rho$ (which is only defined for $DP_B$) as a common representational element.

Now, let's consider the interrelation between IT and SM via probability concepts. I should point out that in any fine-grained conceptual interrelation between BSM and IT mediated by an

[35] "The difference between objective and subjective accounts is that the former tells us why gases go to equilibrium, whereas the latter tell us why *we* ought to *believe* that gases move towards equilibrium (…). At bottom, then, the complaint is really this: why should we *believe* the gas likely moves toward equilibrium unless its *objectively* likely to move toward equilibrium? The thought is that the subjective probabilities should bottom out in something objectively probable." (Callender 2011, p.105)

[36] This conceptual disconnection is grounded on the fact that "increasingly large overlaps (i.e., probabilities, according to the probability measure) will be associated with increasingly large macrostates (i.e., entropy, according to the entropy measure) is based on contingent assumptions concerning the dynamics of our world. This is not a conceptual identification, but a contingent empirical correlation found in experience (…) The measure for entropy and the measure for probability are derived in two different ways, and result from two distinct empirical findings, and therefore are two distinct theoretical concepts" (Shenker 2020, p.16).

interpretation of their probabilistic concepts we find a significant incompatibility of their representational resources. On the one hand, SM deals diachronically with individual (individual trajectories $\gamma$) or group (sets of trajectories $\rho$) continuous entities generated through Hamiltonian and Liouvillean evolution of these entities, respectively. On the other hand, IT deals diachronically with the generation of individual sequences of discrete symbols $s_i$ (individual messages *m*) or in groups (sets of messages). As noted above, this incompatibility of representational resources of SM and IT (assessed in Section 5.2) could technically be technically solved by discretizing $\Gamma_{IT}$ into (i) Boltzmannian macrostates-symbols $\Gamma_{Mi}$ on which to reinterpret $SP_B$ from IT and derive $DP_B$ following Shenker (2020), or into (ii) Gibbsian symbol-cells $\omega_i$ of equal probabilistic measure $\mu(\omega_i)$ and derive $DP_B$ following Frigg (2004). However, in the interpretive specification of IT from $DP_B$ this second type of partitions is restricted to the particular case where the probability of all symbols is identical. In any case the hybrid SM-IT partition of the $\Gamma$ of possible messages would allow defining common probabilities for both conceptual domains.

Shenker (2020) also defended the possibility of conceptually deriving IP (i.e., IT-probabilities, Section 2.3.4) from $DP_B$ and $SP_B$ (i) by means of a frequentist interpretation of the latter that would allow to endow the former with microstate-driven dynamic content, and (ii) from her framework to reconceptualize IT-concepts by means of BSM-representational resources (see Section 2.5.4.). His proposal is as follows: "The total probability $p_i$ for producing the symbol $s_i$ can be the combination of all the probabilities of producing it in all possible circumstances; and the physical realization of this might be that the total probability $p_i$ of macrostate [$\Gamma_M$] should be given by the sum of the probabilities of arriving in that macrostate from all the other macrostates, within a given time interval" (Shenker 2020, p.16). That is, the probability of generating a particular symbol $s_i$ (associated with a macrostate) by means of the actual trajectory $\gamma$ of the system is proportional to $\mu(\Gamma_{\rho} \cap \Gamma_M)$, namely, the measure of the intersection between the dynamical blob $\rho$ of all possible trajectories and the macrostate-region $\Gamma_M$ associated with that symbol. Interestingly, Frigg (2004) arrived to that very conceptual connection from a same-measure cells-symbols partitioning of $\Gamma$ and an epistemic interpretation of $DP_B$ as a measure of the degree of unpredictability of individual trajectories $\gamma$.

Thus, the interpretive specification of $SP_B$ shows us that the possibility that the SM-trajectories generate a particular IT-message depends both on (a) the probability measure of each macrostate-symbol (grounding $SP_B$) and (b) the particular system's dynamics. But because Shannon (1948) considered message generation as Markovian processes, the particular Hamiltonian dynamics that could implement this sequence should be that of Bernoulli systems (Berkovitz et al. 2006), i.e., those Hamiltonian dynamical systems in which the predictions of the future microstate are independent on the location of the actual microstate. Therefore, the possibility of connecting IT and BSM by interpretive specification of their probabilistic representational resources would depend on applying such a conceptual connection on Hamiltonian systems whose Bernoulli-driven dynamics are maximally unpredictable.

## 2.6. Evolutive Conceptual Landscapes

During this section I have explored the conceptual foundations on which information physics historically unfolded during the mid-twentieth and early twenty-first centuries. As I have tried to show, these conceptual elements are distributed along different logical geographies (Ryle, 1948; Frigg, 2008) or 'conceptual landscapes' associated with theoretical domains, among which we include thermodynamics, statistical mechanics (in its Boltzmannian and Gibbsian version), information theory, algorithmic information theory or even semantic information theories. Each of the entropic, informational and probabilistic concepts analysed in this section should not be understood as definitive closed semantic vehicles, but as semantically-convoluted representational resources historically developed by a community of epistemic agents with different theoretical interests and in different fields of application (Chang, 2004; Uffink, 2007). While $S_{TD}$ emerged during the nineteenth century as a computational-differential tool to numerically compute the effectiveness of thermal machines, $H_{IT}$ emerged during and immediately after WW2 from the need to optimize the transmission of messages between noisy channels. Due to the semantically convoluted character of the concepts analysed here, the task of interpretively reorganizing their meaning (Lombardi et al. 2016b) was not only required within the philosophical analyses of these disciplines, but also as a necessary part of the development of these scientific fields, as the Ehrenfests (1912) or Jaynes (1965) advocated in the case of SM long before the disciplinary consolidation of SM philosophy during the 1990s (e.g., Sklar 1993).

As these conceptual elements, the display of a wide variety of conceptual connections over these rich conceptual landscapes (assessed during the previous section by means of relational 'atlases') constitutes a convoluted diachronic process involving the multiple theoretical claims and disciplinary interests of various agents. As I argue below, the fact that information physics as a stream of scientific thought rests intellectually on a web of conceptual connections deployed over the last seven decades justifies the thesis that there is no satisfactory way to assess the exploitation of this web other than to understand its historical unfolding. Similarly, also my own philosophical analysis (including the making of relational atlases on which to develop such analysis) should be recognized as part of a historically unfolded process of conceptual evaluation that goes back at least to the critical work of Carnap (1977) in 1953 at the very historical origins of this intellectual trend. In short, the flourishing of new conceptual landscapes on this convoluted relational fabric implies that these emergent logical geographies are entities subject to a constant and subtle evolution that in turn require to be philosophically evaluated based on their historical unfolding. This is my next task.

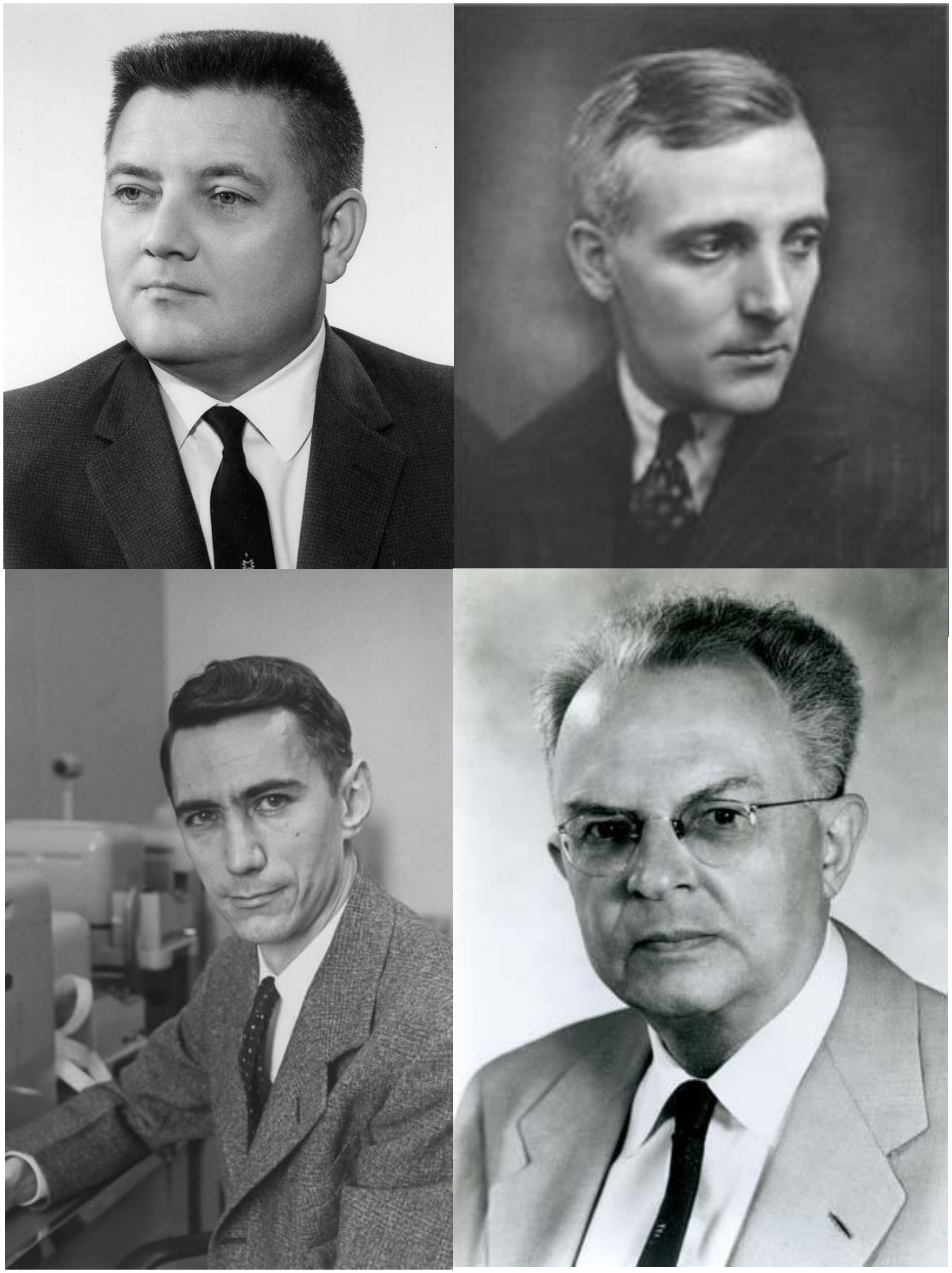

**Image 1**. Main authors in Chapter III. (Top-left) Edwin Thompson Jaynes (1922-1998). (Top-right) Léon Nicolas Brillouin (1889–1969). (Bottom-left) Claude Elwood Shannon (1916–2001). (Bottom-right) Rudolf Carnap (1891–1970).

# Chapter III

## *The Informationalization of Thermal Physics*

> "A new territory was conquered for science when the theory of information was recently developed. This discovery opened a new field for investigation and inmediately attracted pioneers and explorers. It is an interesting phenomenon to watch, in the history of science, and such a sudden expansion of the domain of scientific research deserves closer considerations" (Brillouin 1956, p.ix)

If there is an origin in the history of information physics as a scientific research program, this should be in the 1950s. However, the emergence of this phenomenon within the post-war scientific community cannot be understood as a homogeneous intellectual process. On the contrary, a significant multiplicity of disciplinary interests, theoretical assumptions, intellectual aims and different interpretative strategies were initially intertwined in this convoluted historical process. Progressively, these diverse factors crystallized during the 1950s into two major theoretical agendas or informational traditions in classical statistical thermophysics, namely, the Brillouinian and the Jaynesian. The historical phenomenon that indisputably articulated the development of this convoluted current of thought will be the exponential intellectual capitalization on Shannon's IT, what the author called a 'bandwagon':

> "Information theory has, in the last few years, become something of a scientific bandwagon (…) Our fellow scientists in many different fields, attracted by the fanfare and by the new avenues opened to scientific analysis are using these ideas in their own problems (…) Although this wave of popularity is certainly pleasant and exciting for those of us working in the field, it carries at the same time an element of danger (…) It will be all too easy for our somewhat artificial prosperity to collapse overnight when it is realized that the use of a few exciting words like *information*, *entropy*, *redundancy*, do not solve all our problems" (Shannon 1956, p.3)

### 3.1. Reconstructing the Thermal Engine of Shannon's Bandwagon

Although Shannon's IT bandwagon as an intellectual-historical phenomenon affected during the 1950s a myriad of areas of scientific research (e.g., linguistics, sociology, anthropology, neurology, etc.), it was especially prolific disciplinary fields of the natural sciences such as molecular biology (Yockey 1992) and, as far as this topic is concerned, classical statistical physics. In this section I will attempt to conceptually assess the origin of this convoluted historical-intellectual process in the field of thermophysics, namely: the emergence of new interpretative strategies, the development of new technical concepts, the emergence of unknown theoretical niches, or the novel interrelation between pre-existing theories. In order to perform this historical-philosophical task, I will try to specify (among the vast amount of elements that one could find in this historical domain) the minimum set of factors that explain at most how the bandwagon was possible in this scientific domain, baptized as its 'thermal engine' in the title of this section[37]. The following quote from the philosopher Bar-Hillel is quite illustrative of how Shannon's bandwagon heat engine was critically perceived by the authors of the time:

> “The words ‘cybernetics’ (and its derivatives) and ‘information’ were surely two of the most used, and misused, members of the scientific vocabulary of the fifties (…) Certain curious expressions of Wiener, von Neumann, Weaver and others did nothing to disperse these confusions or the halo of mystery which sometimes surrounded this newly discovered “commodity” whose importance was put on a par with that of “energy”, for instance. True enough, Shannon explicitly dissociated himself from those who interpreted his measures of information as measures of meaning or semantic content, but few were thereby discouraged from making just this interpretation, though they occasionally paid lip-service to Shannon’s disclaimer” (Bar-Hillel 1964, p.11)

### 3.1.1. Norbert Wiener: Complementary Connection between Information and Entropy

One of the decisive factors in the process of informationalization of thermal physics was the development of an alternative information theory to Shannon's by the multidisciplinary author Norbert Wiener (1894-1964), a child prodigy who received his PhD in philosophy of mathematics from Harvard at the age of 18. Just as Shannon's (1948) theoretical proposal was conceptually rooted in his work done during WW2 in the field of message encryption, Wiener's (1948) main proposal set forth in his celebrated book *Cybernetics* emerged from his research on weapon automation and anti-aircraft systems developed at MIT together with engineer Julian Bigelow (Kline 2015, p.19). In this field of study, Wiener developed a filter (now known as the 'Wiener filter') with which to significantly reduce the noise contained in certain radar signals by non-noisy reconstruction of these radar signals, representing these continuous signals as 'time-series'. This greatly increased the ability to predict the position-trajectory (interpreted as a sort of 'time-series-encoded' message) of German bombers, as described by Mindell (2002). Precisely because of this epistemic virtue of his proposal, Wiener referred to his theoretical apparatus as 'predictive theory'.

Interestingly, as early as 1944 (Kline 2015, p.22) Wiener developed a technical concept of 'quantity of information' whose content captured the ability to infer data through his time-series analysis.

[37] The critical analysis of these intellectual factors will be carried out both in Section 3.4. by Carnap and in Chapter V from my conceptual-evaluation proposal.

As noted above, the culmination of this information theory project resulted in the publication of his *Cybernetics. Or control and communication in the animal and the machine* (Wiener 1948). Unlike Shannon's (1948) foundational article, this was not a proposal restricted exclusively to the field of signal transmission, but intended to encompass several fields of application: namely, SM, animal neurology, radar signals, or sociology. Also, unlike Shannon, Wiener combined highly speculative insights (e.g., on the notion of 'time' in Bergson [Wiener, 1948, Chap, 1] with technically precise mathematical derivations. Cybernetics as a multidisciplinary theory of information did not restrict its focus on communicative processes of information transmission (in its case, by means of continuous time-series-encoded signals), but ambitiously extended it to any informational feedback mechanism, regardless of its ontological nature or scope of application.

The concept of information that Wiener intended to develop in this work was found to be directly associated with the capacity of an agent (either human subject or an automaton) to control these complex feedback mechanisms in its favor. In this sense, the term 'cybernetics' (which comes directly from the Greek term κυβερνητική [*kybernētikḗ*]', meaning ‘governor’ or 'pilot') denotes the properly epistemic capacity of an agent to exploit complex informational feedback systems to obtain predictive knowledge about the temporal evolution of those systems. This suggests that Wiener (1948) intended to construct a technically-defined and statistical concept of information by refining the ordinary notion OI (i.e., agent-centric, semantic and epistemic) applicable to these domains, roughly definable as that which provides an agent with the ability to predict the evolution of a system, see Section 2.4. Thus, Wiener’s concept of information ($I_W$) was characterized by the author as a measure of the degree of 'order' (in its ordinary sense OO, Section 2.2.4) of the target system, so that the more ordered it is, the easier it will be for the agent to predictively obtain information about the system. Since Wiener's proposal deals not only with inorganic but also with organic systems, this author uses indistinctly the ordinary concepts of 'order' and 'organization' to characterize the degree of structural and functional regularity (respectively) in the system. It was precisely in the degree of order or organization of the sequences of symbols that encodes their 'amount of meaning', as Wiener (1950) defends in his later work *The Human Use of Human Beings*, where the author interpretatively developed the semantic dimension of his concept of information "It is quite clear that a haphazard sequence of symbols or a pattern which is purely haphazard can convey no information" (Wiener 1950, p.6).

Also, Wiener (unlike Shannon) intended to closely connect his information concept $I_W$ with the physically meaningful concept of $S_{SM}$. However, before detailing this connection we must specify how it depends significantly on combining two distinct interpretative strategies and of which he presents himself as the intellectual heir. On the one hand, Wiener takes up the famous interpretation of negative entropy as a measure of the degree of order (and organization) of a system famously developed by Schrödinger (1944) in *What is Life?*. Although this interpretation is by no means originally from this author (see Section 2.2.4), the main conceptual novelty introduced by Schrödinger in this text is to point out the mechanism that living organisms possess (as opposed to inorganic matter) to remain TD-alike in a low entropy macrostate-region $\Gamma_{M_0}$ in exchange for TD-like increase in the entropy of their environment: “a stream of negative entropy upon itself, to compensate the entropy increase it produces by living and thus to maintain itself on a stationary and fairly low entropy level. If D is a measure of disorder, its reciprocal, l/D, can be regarded as a direct measure of order. Since the logarithm of l/D is just minus the logarithm of D, we can write Boltzmann's equation thus: $-(S_B) = k \log (l/D)$. Hence the awkward expression 'negative entropy' can be he replaced by a better one: entropy, taken with the negative sign, is itself a measure of order. Thus, the device by which an organism maintains itself stationary at a fairly high level of the orderliness (= fairly low level of entropy) really consists on continually sucking orderliness from its environment” (Schrödinger 1944, Chap. 6). Wiener relied on this order-driven interpretative strategy of Schrödinger (1944) (which was free of information concepts) to defend those cybernetic agents also possess the ability to exploit these ‘orderliness-conceived’ entropic gradients to not decrease their own $S_B$ value.

On the other hand, Wiener relied on certain epistemic interpretations of entropy to develop his conceptual connection between entropy and information. Interestingly, there are epistemic interpretations that made use of the ordinary concept of information OI on which Wiener (1948) developed his theoretical proposal. The paradigmatic example of this can be found in the following quote from the chemist G. N. Lewis “Gain in entropy always means loss of information, and nothing more. It is a subjective concept, but we can express it in a least subjective form, as follows. If, on a page, we read the description of a physical-chemical system, together with certain data which help to specify the system, the entropy of the system is determined by these specifications. If any of the external data are erased, the entropy becomes greater; if any essential data are added, the entropy becomes less.” (G. N. Lewis 1930, p.573). That is, the more information (OI) an agent has about the microstate of a system, the smaller will be the measure of its macrostate-region and therefore of its entropy. Combining this epistemic-informational interpretative line of Lewis (A) with Schrödinger's disorder-based (B), Wiener (1948, p.18) defined his statistical concept of information as a measure of the degree of order-organization (or equivalently, of predictiveness) of the target system (B) and therefore also of its amount of negative entropy: "In Wiener's view, information was not just a string of bits to be transmitted or a succession of signals with or without meaning, but a measure of the degree of organization in a system” (Conway and Siegelman 2005, p.132). Illustratively, an agent can predictively extract more information about a crystalline substance whose components are well spatially arranged (i.e., in high negative entropy state) than about an apparently spatially-disordered gas in a vessel (i.e., low negative entropy or high entropy).

Thus, unlike Shannon's theoretical proposal with $H_{IT}$, Wiener developed a technical concept of $I_W$ by combining the above interpretative strategies whose meaning was intrinsically statistical mechanical: "The notion of the amount of information attaches itself very naturally to a classical notion in statistical mechanics: that of entropy. Just as the amount of information in a system is a measure of its degree of organization, so the entropy of a system is a measure of its degree of disorganization; and the one is simply the negative of the other (…) We have said that amount of information, being the negative logarithm of a quantity [$\mu(\Gamma_M)$] which we may consider as a probability, is essentially a negative entropy" (Wiener 1948, p.18). This complementary conceptual connection between $S_B$ and $I_W$ underlying Wiener's proposal involves a technical refinement of the interrelationships presupposed de facto in Lewis' (1930) epistemic interpretation of $S_B$, encoding the amount of epistemic (i.e., it provides predictive knowledge) and semantic (i.e., it refers to the microstate of the system) information that agents can extract from the system into the statistical entropy of the macrostate-region corresponding to their macroscopic knowledge. Thus, $S_B$ and $I_W$ are but two complementary measures describing two different facets of a single property common to agents and systems: in the same way that the entropy of a closed system tends to increase according to the law, the information that an agent can predict about the microstate of this system would also tend to decrease[38]. This conceptual proposal also allowed Wiener to reconceptualize interpretatively part of the disciplinary edifice of thermal physics, as shown by this attempt to derive the Second Law of TD in a communicative theoretical context "It will be seen that the processes which lose information are, as we should expect, closely analogous to the processes which gain entropy. (…) No operation on a message can gain information on the average. Here we have a precise application of the second law of thermodynamics in communication engineering. Conversely, the greater specification of an ambiguous situation, will (…) generally gain information, and never lose it." (Wiener 1948, p. 78-79).

It should be noted that the intellectual development of this cybernetic theory of information was substantially influenced by the results that led to Shannon's theory of communication, and vice versa, as Weaver (co-author and populariser of the latter) recognized[39] (Shannon and Weaver 1949, footnote 1). One of the most striking points resulting from a comparative analysis of both proposals can be found in the fact that both derived a mathematically identical concept of information, $H_{IT}$ and $I_W$, except for the difference of sign in their formulations. Some authors later interpreted this formula-based difference as implying different ways of relating their respective concepts of entropy and information[40]. However, Shannon and Wiener concluded in October 1948 through private correspondence that this sign difference was a mere conventional accident and both notions

[38] "A measure of information is a measure of order. Its negative will be a measure of disorder (…) This measure of disorder is known to the statistical mechanist as entropy (…) [which concept] is associated with that of pattern, and represents the disorder in a class of patterns. Amount of information is a measure of the degree of order (…) [The] amount of information is a quantity which differs from entropy merely by its algebraic sign and a possible numerical factor. Just as entropy tends to increase spontaneously in a closed system, so information tends to decrease; just as entropy is a measure of disorder, so information is a measure of order. (Wiener 1950, p. 18-21, 129).

[39] "Professor Wiener, on the other hand, points out that Shannon's early work on switching and mathematical logic antedated his own interest in the field; and generously adds that Shannon certainty deserves credit for independent development of such fundamental aspect of the theory as the introduction of entropic ideas. Shannon has naturally been specially concerned to push the applications to engineering communication, while Wiener has been more concerned with biological application (central nervous system phenomena, etc.)" (Shannon and Weaver 1949, fn.1).

[40] For instance: "although [Wiener and Shannon] conceived of information in similar ways, Wiener was more inclined to see information and entropy as opposites (…) Like Brillouin and many others of his generation, Wiener accepted the idea that entropy was the opposite of information. The inverse relation made sense to him because he thought of information as allied with structure and viewed entropy as associated with randomness, dissipation and death (…) Claude Shannon took the opposite view and identified information and entropy rather than opposing them. Heuristically, Shannon's choice was explained by saying that the more unexpected (or random) a message is, the more information it conveys" (Hayles 1990, p.102, 305).

were identical concepts on the basis of their arithmetic connection “It was interesting to note how closely your work has been paralleling mine in a number of directions. I was somewhat puzzled by one or two points. You consider information to be negative entropy while I use the regular entropy formula (…). with no change in sign. I do not believe this difference has any real significance but is due to our taking somewhat complementary views of information (...) We would obtain the same numerical answers in any particular question.” (Shannon [1948] to Wiener, quoted on Kline 2015, p.15). In any case, although both authors defended the insignificance of this sign difference between both quantitative concepts, this superficial distinction had a substantial impact on the reception of both theoretical proposals during the 1950s, as I will detail below

### 3.1.2. John von Neumann: Information Theory as a Naturalized Epistemology for Science

The most decisive author in the historical-intellectual process of informationalization of thermal physics is the Hungarian mathematician John von Neumann (1903-1957), pioneer of the computer sciences and main developer of the formalism of quantum mechanics. His relevance lies precisely in the radical impact generated by his theoretical proposals and the vast immediate influence that his ideas had on several generations of scientists for more than three decades. To understand the dimension of von Neumann's intellectual contribution to this current of scientific thought, one must go back to the formulation of Maxwell's demon problem, whose different modes of influence on the configuration of information physics we will deal with in more detail later on.

As Earman and Norton (1998) discussed in detail, in a letter from Maxwell to Tait in 1867 Maxwell postulated the existence of a finite being whose knowledge of the positions and velocities of the molecules of a gas would eventually allow him to modify its entropy values contrary to the Second Law of TD. According to what Maxwell himself defends in his 1878 article on 'Diffusion' in the Encyclopedia Britannica, by means of an epistemic interpretation of the molecular-mechanical meaning of entropy we could understand the action of this 'finite being' as a theoretical possibility: “the idea of dissipation of energy depends on the extent of our knowledge (…) the notion of dissipated energy would not occur to a being who could not turn any of the energies of nature to his own account, or to one who could trace the motion of every molecule and seize it at the right moment” (Maxwell 1878 [1953, p.646]). This interpretative strategy allowed us to assume that certain concrete epistemic states (in particular, those possessed by this being or 'demon', as Lord Kelvin later called it (Leff and Rex, 1990) could be employed to cause certain thermal effects in macroscopic systems contrary to the phenomenological laws that theoretically govern their behavior. More than four decades after the letter to Tait, Marian Smoluchowski (1912) posed a dilemma concerning these epistemic states of demonic-like intelligent (or ‘epistemic’) agents: either these states are extra-physical and cannot be described by physical representational resources (i.e., maintaining a dualistic position), or else the causal interaction of these epistemic states with physical properties implies that they can also be (i.e., from a reductive perspective) physically characterized: “there is certainly no doubt that an intelligent being to whom physical phenomena are transparent could bring about processes that contradict the Second Law. Indeed Maxwell has already proven this with his Demon. However, intelligence extends beyond the boundaries of physics. On the other hand, it is not to be excluded that the activity of intelligence, the mechanical operation of the latter, is connected with the expenditure of work and the dissipation of energy." (Smoluchowski 1912, p.1080).

The Hungarian physicist Leo Szilard's (1929) historically assumed this second horn of Smoluchowski's dilemma in his famous paper 'On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings', where he naturalized the epistemic-driven action of the demonic agent by means of a physical process of macroscopic measurement on equally probable alternatives. In particular, a single molecule was considered in a cylinder (initial stage), to which a partition was introduced dividing it into two equal volumes $V_1$ and $V_2$, subsequently a detection is performed on which of the two volumes the molecule is located (generating an epistemic state $\varphi_1$ and $\varphi_2$), then a piston is inserted in the opposite volume, and finally the partition is eliminated. From the operation of this one-molecule engine, Szilard derived (in an obscure way, see Earman and Norton [1998, p.456]) the idea that obtaining the epistemic state φ containing the macroscopic location of the particle would entail a certain average entropy production of $k \ln 2$, ultimately compensating the entropic reduction that could be generated in the cycle. As Köhler (2001, p.103) pointed out, this was the first time in the history of physics in which an author claimed to physically naturalize an epistemic state φ (i.e., knowledge of the macroscopic location of the molecule either in $V_1$ or $V_2$) and its logical structure with respect to other possible epistemic states $\varphi_i$ relative to that particular scenario[41]. Other authors such as Norton (2013) defended that this choice of the 'naturalistic' horn of Smoluchowski's (1912) dilemma by Szilard (1929) constituted a historical source of distraction concerning the proper micromolecular solution (analysis of thermal fluctuations) with respect to Maxwell's demon problem. In any case, the intellectual impact of this paper of Szilard was decisive, especially on von Neumann.

In his famous ground-breaking book *Mathematical Foundations of Quantum Mechanics* (where for the first time a common mathematical formalism for this physical discipline was developed), von Neumann (1932) carefully analysed and took to its ultimate interpretative consequences Szilard's (1929) physical-naturalistic position regarding epistemic states. In this evaluation, this mathematician defended the interchangeability of certain forms of knowledge (i.e., knowing alternative locations of a molecule) with thermophysically significant quantities of $S_{TD}$ by means of certain controlled physical operations[42]: "L. Szilard [1929] has shown that one cannot obtain

---

[41] "[Szilard] was the first to explicate more precisely than anyone had done before how cognitive states and processes -in particular memory and discrimination- could be treated within physics. It was crucial for understanding the physical workings of an intelligent agent to represent its observational and logical operations within the physical theory itself. Boltzmann (1877) had already conjectured on the widely-recognized relation between thermodynamic entropy and states of knowledge. But for many physicists such as John von Neumann and Wolfgang Pauli, Szilard (1929) was a revelation they would never forget" (Köhler 2001, p.103).

[42] "at the end of the [Szilard's cycle], the molecule is back in volume V; but we now no longer know whether it is located on the left or right, although there is a compensating entropy decrease of x In 2 (in the reservoir). That is, we

this "knowledge" without at least a compensating entropy increase of $k \ln 2$ - in general, $k \ln 2$ is the "thermodynamic value" of knowing which of two cases of an alternative obtains. - All attempts to carry out the process described above not knowing in which half of the container the molecule is located may be shown to be invalid, although they involve, under certain circumstances, quite complicated automata-mechanisms.” (von Neumann 1932, p. 213). Thus, the interchangeability between (i) the attainment of epistemic states by macroscopic measurements and (ii) a minimal production of entropy shown by Szilard (1929) led von Neumann to consider that such scientific forms of knowledge and observation-like epistemic processes (also decisive in the then-novel quantum domain, i.e., the uncertainty principle) could be integrated as ‘physical states’ and ‘physical processes’ in a technically-mathematically rigorous and conceptually consistent way within the theoretical apparatus of physics[43]. However, this early von Neumann's intellectual goal of naturalizing the epistemic elements emerging in the physical sciences would drag on for the next few decades, due in part to the lack of theoretical resources in the early 1930s to carry out this monumental task.

These ideas have generally been recounted since the 1950s in the scientific (and later on also in the philosophical) literature from an 'informational' point of view, e.g., "Demon's information processing [in Szilard's work]" (Earman and Norton 1998, p.437). The fact is that, until this historical moment (circa 1932) no author had explicitly employed the everyday notion of 'information' to describe epistemic states and processes in this context, let alone any technical concept of 'information' (Hartley's 1928 notion $H_H$ was historically popularized to the scientific community via Shannon [1948]). Instead, Maxwell (1878), Smoluchowski (1912) or Szilard (1929) preferred psychologically-flavored terms such as 'intelligence' (also employed by Nyquist [1924] in the field of telecommunications) or more suggestive terms such as 'knowledge' à la von Neumann (1932), either in English in the case of Maxwell or originally in German for the rest. Of course, one can assume that all these authors implicitly used OI, but this does not change the historical fact that the uncontrollable proliferation of informational concepts in physics occurred de facto only from the late 1940s and not before. The history of this stream of scientific thought has shown that the choice of terminology, although prima facie a mere conventional exercise, has turned out to be far from trivial in terms of its impact (see Wicken 1987).

The relevance of naming choices is famously the case with the well-known popular accounts that place von Neumann as a decisive intellectual factor in Shannon's choice of the name 'entropy' for his concept, the central source of which (see Thims 2012, p.52) is attributed to an in-person

have exchanged our knowledge for the entropy decrease of x In 2. That is: in volume V the entropy is the same as that in volume VI2 under the assumption one knows in which half of the container the molecule is located” (von Neumann 1932, note 202).

[43] “We here see that von Neumann quite early confidently regarded thermodynamic entropy decrease as exchangeable with - hence equivalent to - information increase, i.e. increase of knowledge. Thus it appears that epistemological states and processes may be understood as physical states and processes.” (Köhler 2001, p.103).

interview conducted by Myron Tribus with Shannon in 1961. One of the versions[44] (Tribus 1963, 1978) most frequently cited is that reported by Tribus in his 1971 paper 'Energy and Information' (co-authored with E. McIrvine), where he posits a hypothetical meeting between von Neumann and Shannon at the Institute of Advance Studies at Princeton during the 1940/1941 academic year "[Shannon said:] 'My greatest concern was what to call it. I thought of calling it 'information', but the word was overly used, so I decided to call it 'uncertainty'. When I discussed it with John von Neumann, he had a better idea. Von Neumann told me, 'You should call it entropy, for two reasons. In the first place you uncertainty function has been used in statistical mechanics under that name. In the second place, and more importantly, no one knows what entropy really is, so in a debate you will always have the advantage." (Tribus and McIrvine 1971). Authors such as Denbigh (1982), Wicken (1987) or Ben-Naim (2008, p.xvii) have described this terminological suggestion by von Neumann as a disservice to science[45]. However, the historical truth of this quasi-legendary episode has neither been confirmed (nor completely denied) by Shannon himself, as shown by one of his last interviews (already suffering from Alzheimer's) with engineer Robert Price in 1982[46].

Regardless of whether this was the case or not, the fact is that von Neumann followed with enormous interest the progressive conceptual and technical evolution of the information theories of both Wiener (1948) (the two worked together in 1945 on an interdisciplinary computing project at MIT) and Shannon (1948), with whom he maintained a constant exchange of ideas throughout the 1940s (see Kline 2015, p.10, 29, 36). One can have good reason to believe (see quote below) that von Neumann was able to see in the gradual development of these sophisticated theoretical tools an excellent opportunity to advance his goal of describing in a technically rigorous way the epistemic states of thermal physics. Just a few months after the publication of the successful works of Shannon (and Weaver 1948) and Wiener (1948), the Hungarian mathematician explicitly laid the conceptual foundations for the informationalization of thermal physics in part of a talk given at the University of Illinois entitled 'Theory and Organization of Complicated Automata':

> "I have been trying to justify the suspicion that a theory of information is needed and that very little of what is needed exists yet. Such small traces of it which do exist (...) it is likely to be similar to

[44] A less frequent version of this story is the one of 1961: "In the 1961 interview with Shannon, to which I referred, I obtained one anecdote which seems to me worth recording on this occasion. I had asked Dr. Shannon what his personal reaction had been when he realized he had identified a measure of uncertainty. Shannon said that he had been puzzled and wondered what to call his function. 'Information' seemed to him a good candidate as a name, but 'Information' was already badly overworked. Shannon said he sought the advice of John von Neumann, whose response was direct, "You should call it 'entropy' and for two reasons: first, the function is already in use in thermodynamics under that name; second, and more importantly, most people don't know what entropy really is, and if you use the word 'entropy' in an argument you will win every time!" (Tribus 1963 [1979 p.2f.]).

[45] Illustratively "In my view von Neumann did science a disservice!" (Denbigh 1982) or "My reason for embracing Denbigh's statement is that information (or choice or uncertainty) is a simple, meaningful and well-defined concept. Renaming it entropy merely corrupt the term information" (Ben-Naim 2008, p.xviii).

[46] "Price: Well, there are a couple of stories. There's the one that Myron Tribus says that von Neumann gave you the word entropy, saying to use it because nobody, you'd win every time because nobody would understand what it was. (…) / Shannon: von Neumann told that to me? / Price: That's what you told Tribus that von Neumann told that to you. (…) / Shannon: No, I don't think he did." (See https://ethw.org/Oral-History:Claude_E._Shannon).

> two of our existing theories: formal logics and thermodynamics. It is not surprising that this new theory of information should be like formal logics, but it is surprising that it is likely to have a lot in common with thermodynamics (…) Thermodynamical concepts will probably enter into this new theory of information. There are strong indications that information is similar to entropy and that degenerative processes of entropy are paralleled by degenerative processes of information. It is likely that you cannot define the function of an automaton, or its efficiency, without characterizing the milieu in which it works by means of statistical traits like the ones used to characterize a milieu in thermodynamics (…) An automaton in which one part is too fast for another part, or where the memory is too small, or where the speed ratio of two memory stages is too large for the size of one, looks very much like a heat engine which doesn't run properly because excessively high temperature differences exist." (Von Neumann, 1949 [1966], p. 62f.)

Therefore, for von Neumann, information theory should be precisely the necessary tool that would allow us to systematically connect formal logic (i.e., modelling knowledge by means of propositions) with the empirical scientific domain of thermodynamics (necessarily including also statistical thermodynamics or SM). As he notes at the end of the quote, this threefold TD-information-theoretical-logical connection would allow him to carry out his Szilardian task of naturalizing epistemic states of scientists, such as memory units, so that these can be described as physical systems with thermal behavior. In fact, von Neumann (1949) suggests on the basis of "strong indications" that IT would in turn allow to rewrite the Second Law (i.e., "degenerative processes of entropy") in strictly informational terms (i.e., "degenerative processes of information"). This was justified by the author on the basis not so much of the terminological connection of their notions (whose role I have discussed) but on the formal similarity of the mathematical expressions of $H_{IT}$ and $S_G$ (see Section 2.5.4), to which is also added the formal similarity of both with a quantitative inductive-logic concept developed by Carnap (1945) at the time (Köhler 2011, p.106). Note that, according to this presumable connection, as long as the meaning of $H_{IT}$ and $S_G$ (interpreted as $S_{TD}$) can be connected with the epistemic states (propositionally encoded) of scientific agents whose content refers to physical states, then both concepts should be interpretively connected implicitly also with OI. All this complex systematic interrelation that von Neumann deploys in this 1949 talk evolved in the early 1950s as a deep conceptual TD-ID-logic identification, as reported by Bar-Hillel (1964, p.1) on his visit to Princeton in 1952: "During one of my visits to him in Princeton, in 1952, von Neumann also came to see him, and we started discussing the talk I had heard von Neumann deliver shortly before at an AAAS meeting in St. Louis, in which he had proclaimed, among other things, a triple identity between logic, information theory and thermodynamics." (Bar-Hillel 1964, p.1).

In short, Shannon's (1948) IT came to mean for von Neumann in the mid-1950s not a mere statistical analysis of the transmission of communicative signals as his own author advocated, but above all an opportunity to describe naturalistically (i.e., in a scientific and mathematically rigorous way) how epistemic agents obtain knowledge via the informational-physical behavior of their neural elements "Another thing about which we can't tell today as much as we would like, but about which we know a good deal, is that it might have been quite reasonable to expect a vicious cycle when one tries to analyse the substratum which produces science, the function of human intelligence. The whole evidence of exploration in this area is that the system which occurs in intellectual performance, in other words in the human nervous system, can be investigated with physical [TD] and mathematical [logical] methods" (von Neumann 1958, p.486f.). Unlike Wiener (1948) and following Köhler (2001, p.114), von Neumann's intellectual profile was not inclined to deal in a philosophically speculative way with the main epistemic problems emerging from physical practices, but to physicalize such problems à la Szilard (1929) and then deal with them satisfactorily with our best scientific tools. In any case, the influence of his enormously suggestive ideas was essential for the generation of this whole intellectual trend[47].

### 3.1.3. Warren Weaver: Shannon's Information Theory beyond Shannon

Along with Wiener and von Neumann, the work of Warren Weaver (1894-1978), director of the Division of Natural Sciences at the Rockefeller Foundation from 1932 to 1955, was one of the key pieces in the informatization of thermal physics. However, his task was not to develop his own IT proposal (as was the case with Wiener) or to promote the construction of a theoretical apparatus with an explicit function in the physical field (along the lines of von Neumann), but precisely to convert Shannon's (1948) original theoretical proposal (Section 2.4.2) into a product (i) accessible[48] and (ii) intellectually attractive to the scientific community in general and not exclusively to the particular field of telecommunications engineering.

To account for Weaver's role, one should first understand a key difference between the disciplinary scope of Shannon's (1948) and Wiener's (1948) proposal in the late 1940s, beyond the technical difficulty of both. While Wiener intended to actively defend and popularize his cybernetics as a new theoretical discipline whose scope of application encompassed a myriad of domains, Shannon wanted to clarify that the application of his results was focused on the analysis of signal transmission, thus "Shannon did not popularize information theory nearly as much as Wiener popularized cybernetics" (Kline 2015, p.123). The contrast between the intellectual immodesty of the former and the intellectual modesty of the latter had a direct impact on the immediate publicity of both works, as Kline points out: "The media did not publicize Shannon nearly as much as they did Wiener. (…) The press portrayed the two men as opposites. Wiener was celebrated as a mathematical genius who spoke several languages, cared about the social implications of science,

[47] "Von Neumann consistently avoided "philosophical" discussions of epistemological issues. He was all the more engaged in implementing logic and methodology in concrete engineering applications, as the standard history of computers by Goldstine (1972) shows. It is obvious that his efforts to explicate knowledge acquisition using automata and nerve nets ultimately went back to Szilard's "physicalizing" thermodynamic knowledge, as von Neumann (1966, p. 62f.), himself indicated." (Kohler 2001, p.114).

[48] "If a mathematically rigorous theory of information was almost unreadable to engineers, even a mathematically light version was tough sledding for lay audiences. Nevertheless, journalists, popular science writers, scientists, and engineers introduced Shannon's theory to general readers soon after it appeared in the Bell System Technical Journal in 1948" (Kline 2015, p.121).

and made “sensational” claims about cybernetics. Shannon was seen as a modest genius, a wiry figure (…) who quietly did important research for the corporate giant AT&T” (Kline 2015, p.125).

Everything suggested that Shannon's original article (1948) in the Bell System Technical Journal was going to have the same fate as those of Nyquist (1924) and Hartley (1928), i.e., not to be known beyond the technical field in question. As carefully analysed by Kline (2015, p.121), Weaver came across the article in the summer of 1948, and after exploring with Shannon the possibility of applying IT to machine translation processes (his field of research at that time), the latter proposed to the head of the Rockefeller Foundation the possibility of writing a text that would present Shannon's article in a way accessible to the non-technical public. This text finally became the 'Recent Contributions to the Mathematical Theory of Communication', and was incorporated as a presentation of Shannon's proposal in the famous book *The Mathematical Theory of Communication* (Shannon and Weaver 1949). Already the change from the indefinite article 'A' of the original title (Shannon 1948) to the definite article 'The' of their joint book indicates a desire on Weaver's part to popularize Shannon's IT as 'the' information-theoretic proposal as opposed to the then enormously popular cybernetics of Wiener (1948). To accomplish this task in a satisfactory way, Weaver sought to present Shannon IT, and in particular his $H_{IT}$ concept, in a way closer to Wiener's conceptual proposal $I_W$ than to Shannon's (1948) original constraints (note that Weaver was well acquainted with Wiener's intellectual work, who was his boss during the radar signals project during WW2 [Kline 2015, p.126-127]). Namely, in his famous introduction (Shannon and Weaver 1949, p.93-117), Weaver intended to popularize the $H_{IT}$ concept as a notion (i) with semantic potential and (i) conceptually connected to thermophysical notions of entropy.

Regarding the originally-defined nonsemantic character of Shannon’s $H_{IT}$, Weaver was well aware that this concept "at first seems disappointing and bizarre - disappointing because it has nothing to do with meaning," (ibid, p.114). For this reason, Weaver (unlike Shannon) did not want to close off the possibility of theoretically exploring the 'attractive' or 'non-disappointing' non-explicit semantic dimension of $H_{IT}$ from the very beginning. This author distinguished between the plausible conceptual dimensions of $H_{IT}$ in communicative contexts, namely: (i) the syntactic "How accurately can the symbols of communication be transmitted?" (ibid), (ii) the semantic "How precisely do the transmitted symbols convey the desired meaning?" (ibid) and (iii) the pragmatics "How effectively does the received meaning affect conduct in the desired way?" (ibid). Whereas Shannon focused exclusively on the syntactic dimension of $H_{IT}$, Weaver literally suggested that this concept "actually is helpful and suggestive for the level [ii, semantic] and [iii, pragmatic] problems" (ibid., p.114). This (by no means innocent) 'promise of semanticity' had a decisive later impact not only on the popular assimilation of IT and $H_{IT}$, but also on certain theoretical programs[49]; for example, constituting the source of Drestke's (1981) legitimacy for erecting a concept of semantic information about $H_{IT}$, as we saw in Section 2.4.4.

[49] “Weaver thus did more than simplify information theory. His mathematical account was simpler than Shannon’s, but he went beyond Shannon to point out the semantic and pragmatic implications of the theory, to which social scientists responded” (Kline 2015, p.122).

Considering the nonphysical content of $H_{IT}$, Weaver was able to take advantage of the suggestive inertia generated both by the term 'entropy' and by its formal similarity with the thermophysical concepts of entropy to encourage among the scientific community the 'physicalization' of $H_{IT}$: "when one meets the concept of entropy in communication theory, he has a right to be rather excited-a right to suspect that one has hold of something that may turn out to be basic and important." (Weaver on Shannon and Weaver 1949, p.103). Weaver exploited the TD-IT conceptual connections already advocated by von Neumann (1949) (who further popularized Szilard's [1929] results) to present $H_{IT}$ in connection not only with Hartley's $H_H$, as claimed by Shannon (1948), but also with the thermophysical entropy concepts:

> "Dr. Shannon's work roots back, as von Neumann has pointed out, to Boltzmann's observation, in some of his work on statistical physics (…) that entropy is related to "missing information", inasmuch as it is related to the number of alternatives which remain possible to a physical system after all the macroscopically observable information concerning it has been recorded. L. Szilard […] extended this idea to a general discussion of information in physics, and von Neumann […] treated information in quantum mechanics and particle physics. Dr. Shannon's work connects more directly with certain ideas developed some twenty years ago by H. Nyquist and R. V. L. Hartley, both of the Bell Laboratories; and Dr. Shannon has himself emphasized that communication theory owes a great debt to Professor Norbert Wiener for much of its basic philosophy." (Weaver, on Shannon and Weaver 1949, Note 1)

As is clear from this key paragraph, Weaver non-explicitly employed the concept of ordinary information (and not $H_{IT}$) à la Wiener's (1948) $I_W$ when it refers to the set of microstates compatible with the agent's macroscopic knowledge of the system, although he claims to connect it (in a tentative and superficial way) with the notion of $H_{IT}$ by mentioning 'number of alternatives'. This interchangeable use of OI and $H_{IT}$ against Shannon (1948) is shown to be in tune (willingly or unwillingly) with Weaver's claim to semanticize $H_{IT}$. Assuming this, Wiener also implicitly subscribed a 'complementarist' or 'missing-information' conceptual connection between $S_B$ and OI (i.e., the more entropy the less information) in a manner similar to the interpretive strategy advocated technically by Wiener (1948) and non-technically by Lewis (1930) (see Section 3.1.2). Regardless of what vague connections underpin Weaver's $H_{IT}$-physicalization, it depends significantly on placing $H_{IT}$ within a historical narrative in which it is presented as a natural outcome of the evolution of the concepts of thermophysical entropy. This narrative or historiographic exercise have had a significant impact not only within the general popularization of science (e.g., Seife 2007) but also within the academic historiography of IT: "Entropy has a long history in physics, and during the twentieth century had already become closely associated with the amount of information in a physical system. Weaver carefully credited these roots of Shannon's work" (Aspray 1985, p.124).

In short, Weaver not only intended to present Shannon's (1948) overly technical proposal in a simplified way for the general public or for scientists from other disciplines. On the contrary, he sought to transform communication theory into something more than a statistical tool for signal transmission, namely: into a highly attractive and intellectually suggestive scientific product for countless research domains, including (as is our case) thermal physics. As one could in the following years of the twentieth century, reality surpassed any expectations Weaver might have had in 1948, placing not only Shannon in the Olympus of American science "[Weaver] was already telling his president [in 1948] that Shannon had done for communication theory "'what Gibbs did for physical chemistry'" (Gleick 2011, p. 221) but also turning $H_{IT}$ into one of the most important concepts for modern physics. Of course, this would not have been possible without the success of Weaver's popularization-repackaging of Shannon's proposal.

### 3.1.4. Shannon's Bandwagon Thermal Engine Reconstructed

Up to this point, let summarize the main tenets from this section. The intellectual origin of this Shannon's bandwagon within classical thermal statistical physics in the late-1940s/early-1950s can be satisfactorily explained by three conceptually-driven historically unfolded explananda or (explanatory factors): (a) Wiener's (1948) development during the 1940s of a cybernetic information theory (alternative to Shannon's IT) based on informational concepts with physical and semantic content; (b) von Neumann's (1933, 1954 [1963]) search for a theoretical tool to deal scientifically and technically with certain epistemic elements emerging in thermo-statistical physics; and (c) Weaver's aim to turn Shannon's IT into a scientifically attractive product by popularizing the meaning of $H_{IT}$ as having a plausible semantic and physical content. Although these three elements may seem to derive from the particular interests and objectives of each of these three historical actors, the truth is that if we pay attention to their conceptual bases we could extract certain common patterns, namely: that the potential of any technical informational concept C to be assimilated theoretically within statistical physics depends on it possessing certain intuitively attractive properties (already non-technically encoded in OI), i.e., that it attributes knowledge to scientific agents and that it refers to physical properties. That is, that this concept C possesses (i) semantic and (ii) physical content.

Each of these authors aimed to achieve this very purpose by means of different strategies and departing from different intellectual motivations. Initially, Wiener (1948) developed on his own a technical information concept $I_W$ whose content is constitutively defined as having both (i) thermophysical (since it is interpreted via OO as negative order-interpreted quantities of $S_B$), and (ii) semantic, since it is also interpreted as a measure of meaning by somehow technically refining OI (Wiener 1950). Second, Von Neumann (1949) suggested that IT could be technically employed to conceptually connect (i) thermophysical phenomena $\theta$ (from the formal-formulaic relation between $H_{IT}$ and $S_G$), and (ii) the epistemic states $\varphi$ (encoded as propositions within inductive logic) of scientific agents referring to such phenomena, thus technically refining OI. Thirdly, Weaver (1949) popularized the $H_{IT}$ concept (originally defined by Shannon [1948] as constitutively non-semantic and non-physical) as a technical notion (i) historiographically and interpretatively connected with physical entropy concepts such as $S_B$, and (ii) with a plausible semantic dimension of OI to be explored.

Finally, I should note that this initial stage in the informationalization process of classical thermal physics that I have just analysed (circa 1948-1951) is characterized by the inexistence of systematic and well-defined proposals on how informational concepts are integrated in a theoretically consistent and technically correct way within this physical domain. As has been shown, all that we find in Wiener (1948), von Neumann (1949) or Weaver (1949) (as well as in their closest intellectual entourage) are mere suggestive (but superficial) indications of how to assimilate conceptual meanings, future promises of the interdisciplinary scope of certain theoretical tools or mere attempts at conceptual application. For a significant disciplinary consolidation of the bandwagon in thermo-statistical physics we had to wait until 1956-1957, when the first thermo-informational architectures ('classics' in terms of their impact to this day) of Brillouin (1956) and Jaynes (1957a, 1957b) emerged after a complex historical process of conceptual consolidation and standardization of the interpretative resources available in the scientific community of the time. Due to the enormous difficulty of this task, I will devote the next two sections to an in-depth historical-philosophical analysis of the intertwined conceptual railways underlying the intellectual genesis of these two informational traditions.

## 3.2. Brillouinian Informationalism

Undoubtedly, Léon Brillouin[50] (1889-1969) is one of intellectual leaders of information physics as a current of scientific thought. One of the most interesting aspects of this author is that his own intellectual evolution from the late 1940s to the 1960s constituted a significant thermometer (or properly 'epistemometer') of the main trends of scientific thought within the classical-physics community of the time, as well as of how they were evolving. From an initial affinity (circa 1949) for the theoretical interests and suggestive concepts of Wiener's cybernetics, Brillouin developed in the early 1950s his own concept of information (based on his notion of 'negentropy') that could be applied to the problem of scientific measurement, on which he intended to erect (circa 1956) a complete reformulation of the theoretical edifice of classical thermal physics. As far as we are concerned, Brillouin consolidated a certain 'school of thought' (or 'research program', in Lakatosian

[50] Born in Sèvres, Léon belonged to a long line of French physicists, including his father, Marcel Brillouin, introducing SM in France; and his grandfather, Éleuthère Mascart. Trained under the supervision of Paul Langevin in quantum solid physics, Brillouin developed part of his career in French scientific institutions such as the Collège de France or the Institut Henri Poincaré. However, due to the war, he began to work in the United States (from the University of Wisconsin-Madison in 1942 to Harvard University during 1947-1949), where he developed the rest of his scientific career, even becoming an American citizen in 1949. Brillouin was recognized in the history of modern physics for his theoretical advancements in solid state physics, in wave physics and for the systematic introduction of information concepts into thermal physics. Primary material about this author, including some of his correspondence and transcripts of oral conversations, can be found physically at Niels Bohr Library & Archives, American Institute of Physics. One Physics Ellipse, College Park, MD 20740, USA, see https://history.aip.org/phn/11412022.html.

terms) within informational physics itself, whose idiosyncrasy lies mainly in (i) taking advantage of the historical-intellectual inertia of Shannon's IT bandwagon without relying on the $H_{IT}$ concept, (ii) interpretatively assuming a complementarist connection between (epistemic, semantic) information and (physical) entropy, and (iii) incorporating the agent's epistemic states in observational processes in the thermophysical description of the observed elements. For this reason, in order to satisfactorily understand Brillouin's proposal (as well as later 'Brillouinian tradition', e.g., Earman and Norton [1999, p.14]) we have no choice but to delimit the conceptual and interpretative sources from which he draws at each stage of his intellectual evolution.

### 3.2.1. The European School of Information Theory

> "Future historians of science may find it curious that despite the volume of activity sparked off by Wiener's and Shannon's classic publications in the United States, it was in London that the First International Symposium on Information Theory was held in the summer of 1950" (MacKay 1969, p. 9)

In order to understand the intellectual soil from which not only Brillouinism but also other related proposals (e.g., Raymond, 1950; Rothstein, 1951) emerge, it is interesting to evaluate the main interpretative concerns and theoretical problems of what has become known as the 'European School of IT'. Before the well-known radical popularization of Shannon's communication theory of signal transmission in the early fifties, several scientists in England such as Fisher (1934) and Gabor (1946) had already independently developed what can be called as 'information theories'. All these theoretical proposals focused grosso modo on interdisciplinary assessing of how scientists acquired knowledge through practical procedures like measurements or observation, in contrast to Shannon's IT (and somehow also to Wiener's IT) which focused technically on statistically assessing signal transmission processes. In the words of Colin Chery: "This wider field, which has been studied in particular by MacKay, Gabor, and Brillouin, as an aspect of scientific method, is referred to, at least in Britain, as information theory, a term which is unfortunately used elsewhere synonymously with communication theory." (Cherry 1957, p.176). This intellectual tendency shaped what McCulloch called the 'English School' or what Bar-Hillel called the 'European School' of IT, a trend of scientific thought long neglected in the history of science (with some exceptions, see Kline [2015]) and deeply forgotten by the recent philosophy of physics, despite the pivotal role it played in shaping many trends of scientific thought in the fifties. If we consider that the pragmatic criterion for classifying authors into different schools should be based essentially on their common intellectual patterns (theoretical objectives and interests, interpretative strategies, etc.) and not on their contingent nationality, then Brillouin's (1956) proposal constituted one of the most conceptually consistent and technically sophisticated theoretical manifestations of this 'European School'.

One of the oldest was developed by the British statistician Ronald Fisher, who originally proposed in his *Probability, Likelihood and Quantity of Information in the Logic of Uncertain Inference* (Fisher 1934) the concept of the 'amount of information' that a random observable variable X possesses *about* a parameter O as the reciprocal value of the variance of X with respect to O. As this definition makes explicit, Fisher's technical-statistical conceptual proposal (later popularized as 'Fisher information' $I_F$, e.g., Frieden [1998]) constitutes a mere interpretative transposition of certain conceptual properties of OI to the domain of statistical practices, mainly (i) its semantic-intentional character (i.e., the variable X refers or is somehow directed to the parameter O, and not the other way around) and (ii) its epistemic connotations (i.e., the reciprocal variance of X codifies the scientist's amount of knowledge about O when he receives a value of X).

Apart from Fisher, the Hungarian-British physicist and engineer Dennis Gabor, then associate member of the London College and Nobel Prized in 1971 for his invention of the holography, independently developed an information-theoretical proposal in his 1946 paper entitled *Theory of Communication*, anticipating by a couple of years Shannon's (1948) paper. According to the author, this work (in which the work of Nyquist [1924] and Hartley [1928] is recognized as antecedents) constitutes "an inquiry into the essence of the "information" conveyed by the channels of communication" (Gabor 1946, p.429). In order to technically define this particular concept of 'information', Gabor focused on evaluating the representational resources used to physically describe continuous communication signals in terms of their frequency and time, mainly the waveform analysis of Fourier-transform theory. In this sense, the meaning of its informational concept is not theoretically derived from the statistical properties derived from the transmission of signals (discrete or continuous) in communicative channels à la Shannon (1948) or Wiener (1948), but is constructed from the physical properties of the waves that make up these signals, i.e., essentially the relationship between their frequency and their time. In this sense, the communicative signals are represented by Gabor (1946) two-dimensionally using frequency and time as coordinates, calling this representation 'information diagrams'. Once with these elements, Gabor's concept of 'information' $I_G$ is defined by the area of the information diagram associated with a signal, encoding the number of independent data transmissible by this signal considering (i) the minimum area into which this diagram can be partitioned, and (ii) the impossibility of simultaneously extracting data with high precision on the frequency and time of the signal. This last 'uncertainty relation' of Gabor (1946) shows how his technical concept of $I_G$ information (like Fisher's $I_F$ and unlike Shannon's $H_{IT}$) intends to mimic certain properties of OI, mainly (i) its semantic character (i.e., information diagrams refer to signal properties) and (ii) epistemic, in that the area of these diagrams provides precise knowledge about the frequency or time dimension (but not both) of the signals. Gabor's IT proposal is mainly marked by two factors, his largely engineering profile and his interest in evaluating the epistemic limits of scientific observational processes. This was clear in his 1951 proposal, wherein he developed an optical mechanism to detect the classical (i.e., non-quantum) position of single molecules in Szilard's one-molecules engines with an entropy cost lower than that specified by Szilard (1929) and (even generalized to other processes) by von Neumann (1933, 1949), see Section 3.1.2. Gabor's main intellectual contributions to the informationalization of thermal physics (i.e., the importance of uncertainty relations and the claim to refute the Szilard-von-Neumann thesis) had a limited impact in general, although significant in the evolution of Brillouin's (e.g., 1956, Chap, 8) proposal, due to the fact that the latter considered the former as one of his main interlocutors.

These conceptual proposals of Fisher (1934) and Gabor (1946, 1954) were assimilated by the young King College researcher Donald MacKay in the late forties, who later in the fifties would popularize the later-known as the European (or ‘English’ in McCulloch’s terms) School of Information Theory. From his early interest (circa 1949) in the problem of the limits of scientific measurement "the art of physical measurement seemed to be ultimately a matter of compromise, of choosing between reciprocally unrelated uncertainty" (MacKay 1969, p.1), MacKay became interested in the question of how we should quantify and understand ‘information’ as it appears within the context of scientific measurement. It should be noted that, MacKay's IT wasn’t independent from particular physical theories as was the case of Shannon's IT, but it rather derived (following Gabor [1946]) from assessing the physical theories descriptively employed by scientific agents to obtain knowledge. Originally presented at the first London Symposium on Information Theory in 1950, MacKay aimed to integrate in his ambitious proposal both (i) Fisher's information concept $I_F$ (renamed as 'metric information' or simply 'metron') quantifying the OI from a variable X about a parameter O, and (ii) Gabor's information concept $I_G$ (renamed as 'structural information' or simply 'logon'), quantifying the uncertainty-interpreted OI between frequency-time parameter. Both $I_F$ and $I_G$ belong to what this author calls 'descriptive information', due to their semantic role in being used by observers or scientific agents to properly describe the properties-parameter observed. As we briefly mentioned in Section 2.4.4, MacKay defended that these informational concepts possess, unlike $H_{IT}$, an explicit semantic character derived (similarly as OI) from the intentionality of these agent-observers "It appears from our investigations that the theory of information has a natural, precise and objectively definable place for the concept of meaning" (MacKay 1950 [1969, p.93]). MacKay not only ambitiously aimed to theoretically integrate (more or less successfully) the ‘descriptive information’ conceptual proposals of (i) Fisher’s $I_F$ and (ii) Gabor’s $I_G$, but also aimed to incorporate (iii) Shannon’s $H_{IT}$, reinterpreted as an (agent-centric) measure of 'selective information’. Due to emergence of Shannon’s bandwagon, his theoretical proposal was at first eclipsed and later forgotten by the scientific community during the fifties, being finally published two decades after in his 1969 book '*Information, Mechanism and Meaning*'.

### 3.2.2. Raymond, Rothstein and Brillouin’s Early Intellectual Stage

During his time at Harvard in the late forties, Brillouin began to give theoretical consistency to his concerns about the processes of obtaining knowledge and his intuitions (conceptually dependent on OO and OI) about the meaning of the concepts of entropy beyond the confinements of physics. Interestingly, these intellectual concerns and intuitions was a constant in the intellectual evolution of this author, whose different stages could be characterized by the different interpretative strategies and conceptual resources through which they are theoretically specified. The main apparatus from which to coherently channel this formless amalgam of concerns, interests and intuitions that was available to Brillouin in that particular circumstance (i.e., Harvard Cruft Laboratory and academic year 1948/1949) was Wiener's (1948) cybernetics (Section 3.1.1) along with its entire conceptual scaffolding. The direct result of this initial theoretical pretension is found in his 1949 paper *Life, Thermodynamics, and Cybernetics*, whose content explores themes close to Schrodinger's (1944) *What is Life?* from an explicitly Wienerian orbit. The following is an illustrative paragraph: “Scientific information represents certainly a sort of negative entropy for Wiener, who knows how to use it for prediction, and may be of no value whatsoever to a non-scientist. (…) In addition to the old and classical concept of physical entropy, some bold new extensions and broad generalizations are needed before we can reliably apply similar notions to the fundamental problem of life and intelligence” (Brillouin 1949, p.86-87).

In the above quote we see symptoms of being at the beginning of a process of conceptual development, such as the interchangeable use of the term 'intelligence' and that of 'information' to express the same notion (as occurred historically in IT with the transition from Nyquist’s [1924] intelligence to Hartley’s [1928] information, or in Maxwell's demon problem with the transition from Smoluchowski’s [1912] intelligence  to Wiener’s [1948] information). As was the case with later work, here the term 'information' is (i) closely connected (if not directly identified) with OI, and (ii) complementarily interrelated with concepts of physical entropy; thus using $I_W$ as a technical reference concept that fulfils (i) and (ii). This last conceptual platform served provisionally Brillouin in 1949 to express technically his intuitive interpretation of the concepts of physical entropy as a hybrid measure of the ontological disorder of the observed systems and of the epistemic unpredictability of the observing agents. However, as noted in the above quote, the Wiener information $I_W$ as a (complementary) conceptual extension of $S_B$ is insufficient to delimit in a technically rigorous and conceptually consistent way the degree of inability of agents to gain knowledge about the microscopic configuration of systems. This conceptual deficiency of $I_W$ was one of the main theoretical motivations that articulated Brillouin's intellectual evolution during the following years.

A significant advance in this direction is found in the conceptual refinement work of Richard Raymond (1950), who also started from a complementary interpretation of the physical entropy and information connection à la Wiener [1949] or Brillouin [1949]) and specified that the thermophysical content of $I_W$ lies precisely in its capacity to move systems away from their equilibrium state towards non-equilibrium metastable states. This allowed Raymond to technically define a broad concept of entropy that, generalizing notions such as $S_B$, is quantitatively sensitive to the thermophysical content of information: "the entropy of a system may be defined quite generally as the sum of the positive thermodynamic entropy which the constituents of the system would have at the thermodynamic equilibrium and a negative term proportional to the information necessary to build the actual system from its equilibrium state" (Raymond 1950, p.273). With these new conceptual resources (in a way improving the Wienerian notion $I_W$ in terms of theoretical definition) Raymond (1951) redescribes the action of the Maxwellian demon in the original model of two vessels connected by a trap door, so that its obtaining information about the molecules generates a pressure differential between the two vessels. That is, the increase of OI-based information (semantic and epistemic, i.e., it provides knowledge about the gas) by the demon-agent causes the gas to transit to a metastable state of non-equilibrium in which it would 'store' or 'encode' that amount of information.

Raymond's (1950) conceptual proposal provided him with the representational tools that Brillouin (then an IBM consultant) needed to advance his technically and thermophysically-defined notion of information. This is clearly seen in his celebrated 1951 paper ‘Maxwell's Demon Cannot Operate: Information and Entropy’, where he passed from Wiener information $I_W$ to Raymond's notion as the technically-defined information concept of reference, to whom he acknowledged his profound influence "R. C. Raymond recently published some very interesting remarks on similar subjects (...) He succeeded in applying thermodynamic definitions to the entropy of information and discussed the application of the second principle to such problems" (Brillouin 1951, p.123). Thus, Brillouin considered that the complementary relationship between entropy and information could not only be specified by an intuitive interpretation (e.g., Lewis 1930) of the content of both à la Wiener (1948), but should also be specified technically and quantitatively within the theoretical apparatus of TD and SM à la Raymond (1950). This was the key to the development of one of the most important concepts in the whole history of information physics, whose famous name is introduced in this paper for the first time: “We coined the abbreviation “negentropy” to characterize entropy with the opposite sign. This quantity is very useful to consider and already has been introduced by some authors, especially Schrödinger. Entropy must always increase, and negentropy always decrease. Negentropy corresponds to “grade” of energy in Kelvin’s discussion of “degradation of energy.” (Brillouin 1951, p.120).

Thus, Brillouin (1951) tried to define his technical concept of $I_B$ information through the notion of negentropy as a thermophysically significant quantity (i.e., associated with degrees of utility of energy) interchangeable with negative quantities of entropy by means of certain processes, mainly observation "The physicist making an observation (...) transforms negative entropy into information" (ibid. p.123). Thus, the information $I_B$ that scientific agents possess about the microscopic configuration of systems would be theoretically-embedded within the SM-descriptions of these systems by corresponding numerically to negative quantities (differences) of $S_B$. In an implicit analogy with Clausius' (1865) use of TD-cycles to define $S_{TD}$ (see Section 2.2.1), Brillouin used a variety of 'thermoinformational cycles’ (generalized to non-equilibrium states) to define his informational concept of negentropy. Also, to illustrate the applicability of this concept, Brillouin rewrote the Maxwellian demon problem in such a way that this agent can recover part of the $S_B$-entropy generated during an equilibrium process by using $I_B$-information "to decrease the degradation of energy" (Brillouin 1951, p.121). Like von Neumann (1949) and without going into technical details, Brillouin (1951) concludes (circularly, since he presupposes from the beginning that the Second Law applies, see Earman and Norton [1999, p.9]) with the failure of the demon to violate the Second Law.

One of the most significant aspects of this intellectual stage is his encounter with the then accelerated bandwagon of Shannon's IT, in whose exposition his interpretative reformulation is manifest via Weaver (1949) and/or von Neumann (1949) "Shannon, however, compares information with positive entropy, a procedure which seems difficult to justify since information is lost during the process of transmission, while entropy is increased. Norbert Wiener recognized this particular feature and emphasized the similarity between information and negentropy" (Brillouin 1951, p.120). As the reader can see in the quote, Brillouin not only points out the incompatibility between 'complementarist' à la Wiener (1948) and 'identificationist' à la von Neumann (1949) interpretations of the relationship between entropy and information, but also projects his own information concept $I_B$ (OI-based, semantic and epistemic) onto a $H_{IT}$-driven description (non-OI-based and nonsemantic) of a transmission process. The inability to satisfactorily account for the conceptual architecture and interpretative strategies derived from Shannon's IT (increasingly present within the scientific community in the 1950s) by means of the elements of his 1951 proposal was one of Brillouin's main theoretical motivations to further develop and reconfigure the informational framework. But to achieve this theoretical objective, the systematic analysis of the notions of entropy and information developed by Rothstein in 1951 was conceptually indispensable.

The physicist Jerome Rothstein (1918-2002) is undoubtedly the most decisive author within the Brillouinian tradition (apart from Brillouin himself) of informational thinking in thermophysics, in terms of the creativity of his conceptual proposals and his prolific intellectual dedication in this field. His analysis deployed in 'Information, Measurement, and Quantum Mechanics' (Rothstein 1951) was the Rosetta stone that allows Brillouin to solvently erect a systematic informational proposal on a non-Shannonian information concept such as his $I_B$ in an intellectual environment articulated around $H_{IT}$. In the line of Gabor (1946) or Brillouin (1951), Rothstein (1951) employed a notion of information conceptually dependent on $I_B$ and constitutively defined on the scientific practices of observation: "Observation (measurement, experiment) is the only admissible means for obtaining valid information about the world" (ibid, p.89). Also following Wiener (1948), Raymond (1950) or Brillouin, this author defended here a complementarist interpretation of the relationship between entropy and information. However, unlike Brillouin's (1951) defensive position in the face of the prima facie incompatibility of complementarism with the conceptual architecture derived from $H_{IT}$, Rothstein (1951) (aware of this interpretative incompatibility[51]) sought to integrate the two by connecting IT-based and SM-based representational resources "In the physical case the message is not sent (…) so that physical entropy measures how much physical information is missing. (…) The more information we lose, the greater the entropy, which

[51] Rothstein (1951) expressed this interpretative incompatibility as follows "It may seem confusing that a term connoting lack of information in physics is used as a measure of amount of information in communication, but the situation is easily clarified" (ibid. p.89).

statistical equilibrium corresponding to minimal information (…) We can thus equate physical information and negative entropy (or negentropy, a term proposed by Brillouin)" (ibid, p.89). Note that in this quote, the author uses 'physical information' to refer directly to the $I_B$ concept, since for $H_{IT}$ he used the expression 'entropy of information theory' (p.90).

The Rothsteinian connection of representational resources underlying $H_{IT}$ and $S_B$ consists in making a systematic analogy between the conceptual elements of Shannon's IT communicative models and the measurement models in SM, from which he will finally draw an equivalence between the two "progress either in theory of measurement or in theory of communication will help the other. Their logical equivalence permits immediate translation of results in one field to the other" (ibid, p.90). Interestingly, this conceptual equivalence is represented visually by Rothstein by superimposing the measurement process diagram and the classic Shannon (1948) diagram of a communicative process. The establishment of this conceptual equivalence between IT-communication and SM-measurement serves not only to fit $H_{IT}$ interpretatively within an explicitly complementarist program ("the entropy of information theory is, except for a constant depending on the choice of units, a straightforward generalization of the entropy concept of statistical mechanics" [Rothstein 1951, p.90]) but also to attempt to connect $H_{IT}$ interpretatively with the OI implicit in the measurement model. Insofar as the communicative destination D is equivalent to the scientific observer, endowing D with an agent-based character alien to its original definition by Shannon [1948], the message *m* as equivalent to the value *v* of the measurement performed can be interpreted (i.e., agent-based, semantically and epistemically à la OI) as providing knowledge to the destination D about the information source S. In this way, the message *m* would convey OI to the destination about S just as the measured value *v* conveys OI to the observer about the measured system.

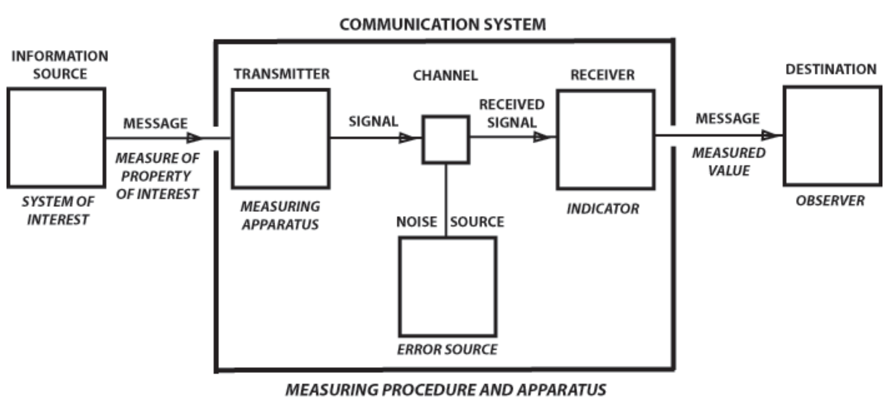


**Figure 3**. Rothstein (1951)'s Identity between Communication and Measuring Problem.

In this way, and without going into the semantic consistency and syntactic correctness of this systematic analogy[52], Rothstein (1951) shows how the conceptual architecture of $H_{IT}$ and IT is not only interpretatively unproblematic concerning a plausible complementarist informational framework (as with Brillouin [1951]) but can also be used to technically refine an informational interpretation of the measurement-SM processes. The latter is particularly important when it comes to an informational understanding of the Maxwellian demon (as well as its Szilardian treatment) as a problem about the observer-SM: "If, like physicists, the demon gets his information by means of measuring apparatus, then the price is paid in full" (Rothstein 1951, p.89). In this direction, another of the characteristics of Rothstein's (1951, 1952) theoretical apparatus that was decisive in Brillouin's intellectual evolution during 1951-1956 was the combination of technical precision in the definition of concepts (and conceptual relations) and a marked vocation for systematicity, pretending to apply his proposal to the very principles of TD: "we can formulate the statistical expression of the Second Law of thermodynamics rather simply in terms of information: our information about an isolated system can never decrease (only by measurement can new information be obtained) (...)" (Ibid, p. 90)[53]. In short, Rothstein's (1951) fine-grained interpretative proposal provided the theoretical resources for Brillouin (as he recognized[54] in 1956) to progressively advance from the creation of his 1951 Wienerian-prone concept of negentropy to the development of a Shannon's IT-compatible informational framework systematically applicable to thermal physics, which I proceed to analyse below.

### 3.2.3. Brillouin's Negentropy Principle of Information

The publication of the book *Science and Information Theory* (SIT) in 1956 constitutes not only the culmination of Brillouin's intellectual evolution in the field of the informationalization of thermal physics, but also represents the first truly systematic theoretical proposal in the (then recent) history of information physics. Interestingly, in the introduction the author speaks of the impact of IT on the science of the 1950s in enormously optimistic terms, pointing to it as an opportunity to explore the fruitfulness of new conceptual landscapes[55]. Already by 1956 intellectual trends within the physics community had de facto changed from those of the late forties or early fifties: while Wiener's (1948) cybernetic was initially presented as a suggestive and promising theoretical platform for the introduction of informational concepts in classical

---

[52] For instance, Mari (1999) pointed out that this analogy between communication and measurement becomes inconsistent because every message can be known independently of its transmission, while the state of the measured system could not be known independently of its measurement.
[53] Since information is interpreted as negentropy, the idea that "our information about an isolated system can never decrease" can lead to interpretatively incoherent theses such as the one stating that the entropy of an isolated system can never increase. These conceptual deficiencies of Brillouin's proposal will be thoroughly assessed in Chapter V.
[54] For instance, after defining his negentropy concept in 1956, Brillouin defended that "The connection between entropy and information has been clearly discussed in some recent papers by Rothstein in complete agreement with the point presented in this chapter" (Brillouin 1956, p.161).
[55] "A new territory was conquered for science when the theory of information was recently developed. This discovery opened a new field for investigation and immediately attracted pioneers and explorers." (Brillouin 1956, p.ix).

thermophysics, by the mid-fifties this was significantly relegated (see Kline 2015, Chapt. 4) (in part, chastising its high speculative content) by the then ubiquitous Shannon's IT "A new scientific theory has been born during the last few years, the theory of information. It immediately attracted a great deal of interest and has expanded very rapidly" (SIT, vii). Wiener's cybernetic proposal went from being 'the' reference (Brillouin 1949, 1951) to receiving no mention[56] in the 350 pages of his 1956 work in only half a decade. Shannon's IT bandwagon had completely eroded the intellectual environment of the time, which is clearly reflected in the intricate interpretive strategies underlying Brillouin's thermo-informational proposal in SIT.

First, Brillouin devoted seven chapters in SIT (from Chapt. 1 'The Definition of Information' to Chapter 7 'Applications to Some Special Problems') to expound Shannon's IT in detail as a strict statistical analysis of signal transmission processes, without pretending to establish (or suggesting) theoretical relations with SM or TD. In Chapter 8 entitled 'The Analysis of Signals: Fourier Method and Sampling Procedures', Brillouin used Gabor's (1946) IT proposal as a conceptual bridge between mere signal analysis and a strictly physical study of signals. Beyond this very extensive exposition of communication theory, Brillouin went from criticizing Shannon's IT-dependent interpretative strategies of connecting information and entropy in his 1951 paper (see previous section) to extolling him as the 'rediscoverer' of such a conceptual relationship in SIT, being highly charitable about the compatibility of Shannonian-like identificationist interpretations with his own complementarist position[57]: “The connection between entropy and information was rediscovered by Shannon, but he defined entropy with a sign opposite to that of the standard thermodynamical definition. Hence what Shannon calls entropy of information actually represents negentropy. This can be seen clearly in two examples (…) where Shannon proves that in some irreversible processes (an irreversible transducer or a filter) his entropy of information is decreased. To obtain agreement with our conventions, reverse the sign and read negentropy” (SIT, p.161).

Another point where this pro-Shannonian attitude can be appreciated is in defending that his conceptual proposal of information $I_B$ (already partially defined in 1951) is independent of OI, because (i) $I_B$ lacks the semantic content and epistemic amount of OI, and (ii) OI does not have a quantitative dimension "we define "information" as distinct from "knowledge", for which we have no numerical measure" (SIT, p.9). Thus, Brillouin interpretatively characterizes his quantitative and statistical notion of $I_B$ information as a technical concept (counter-intuitive with respect to OI) that does not depend on the particularities of the agents since its value is absolutely objective, dependent only on $\Gamma_M$: “The restriction that we have introduced enable us to give a quantitative definition of information and to treat information as a physical measurable quantity. (…) This definition may look artificial at first sight, but it is practical and scientific. It is based on a collection of statistical data on each problem to be discussed, and these data, once available, are the same for all the observers. Hence our definition is an absolute objective definition, independent of the

[56] The only reference to Wiener in SIT is the expression ‘Wiener-Khintchine formula’ (wherein ‘Khinchin’ is misspelled) and ‘Wiener’s autocorrelation function’ in p.101 and p.112, respectively.
[57]As the reader might have noticed, this is an interpretative tactic already followed by von Neumann (see Section 3.1).

observer” (SIT, p.x-xi). Beyond this conceptual distinction, Brillouin showed in SIT a significant proto-philosophical or at least meta-scientific interest (he went so far as to expound the operationalism of P. W. Bridgman [1927]) in scientific concepts, e.g., "new definitions amount to a verbal translation of formulas given symbolically and based on postulates" (SIT, p.x). This concern was further developed in his 1964 book *Scientific Uncertainty and Information*. In line with his philosophical concern, this author also showed his awareness of the SIT proposal of Carnap and Bar-Hillel (1952), recognizing the formal connection of the notion of semantic information of these two authors (see Section 2.4.4) with his own $I_B$ concept (SIT, Section 20.3).

As the reader might have noticed, the whole theoretical architecture of Brillouin’s proposal in SIT rests on the conceptual foundations of BSM (Section 2.1.2), i.e., partition of $\Gamma$ into macrostate-regions, $S_B$, $SP_B$, etc. As a depart from his previous developments, Brillouin differentiated in Chapter 12 of SIT between two concepts of information $I_B$ according to the representational resources on which they are semantically defined, namely: (i) free information $I_{Bf}$, where the possible states referred to have no physical content, and (ii) bound information $I_{Bb}$, where the states referred to are possible microstates (or ‘complexions’ in terms of Planck) of a physical system. Although $I_{Bb}$ are simply particular cases of $I_{Bf}$, the conceptual difference between the two is key to his proposal, since the former notion is 'physically meaningful' (in Brillouin's terms, SIT, p.152) but the latter is not. In fact, (and this is essential) the physical meaningfulness of bound information $I_{Bb}$ is what determines the validity of its semantic connection with negentropy N: “Thus only information connected with certain specific physical problems, that is bound information, will be thought as related to entropy (…) In the case of free information, we prefer not to think of a connection between information and entropy, since the relation between entropy and the number of cases is defined only if the cases are complexions of a physical system” (SIT, p.152-153).

In this sense, Brillouin (SIT, Chapter 1) also suggests that the $H_{IT}$ of a set of equiprobable states is formally, arithmetically and semantically identical to its $I_{Bf}$, which makes possible a conceptual identity between both notions. However, the conceptual identification between $H_{IT}$ and $I_{Bf}$ is not immediate for this author, since it depends on the states on which the former is defined being physically meaningful. As for the criterion of 'physical meaningfulness', Brillouin (SIT, Chapter 12) hints that this partially depends on the use of empirically TD-meaningful macrovariables (temperature, pressure, etc.) to define the possible states. However, this charitable interpretation of his (non-explicit) criterion of meaningfulness is not consistently followed throughout his proposal precisely because "Sometimes it is difficult to distinguish between free and bound information, particularly in cases involving a human observer" (SIT, p.155), which will even become a manifest conceptual deficiency of his interpretative strategy "Here, again, we see the impossibility of always distinguishing clearly between free and bound information" (p.157, fn. 4).

Once the conceptual connection between bound information $I_{Bf}$ and negentropy N (or $-S_B$) was assumed, Brillouin further developed a 'hybrid' interpretative specification (closely similar to Wiener’s interpretation [1948]) of the thermophysical content of $S_B$, turning it simultaneously into (i) a measure of the ontological disorder of the system to which the (micro)states to which it refers belong, and (ii) a measure of the lack of epistemic information possessed by the observer about those microstates, à la Lewis (1930). "Entropy is usually described as measuring the amount of disorder in a physical system. A more precise statement is that entropy measures the lack of information about the actual structure of the system. This lack of information introduces the possibility of a great variety of microscopically distinct structures, which we are, in practice, unable to distinguish from one another. Since any of these different microstructures can actually be realized at any given time, the lack of information corresponds to actual disorder in the hidden degrees of freedom" (ibid, p.160). Apart from implicitly assuming a conceptual connection between OO and OI (i.e., order in the system is equivalent to information about the system), the strongly epistemicist interpretation (ii) of $S_B$ was for Brillouin conceptually prior to its disorder interpretation[58] (i), since, according to the above quotation, the degree of inability of the observer to distinguish macroscopically the actual microstate among a set of possible microstates somehow implies theoretically that the number (i.e., probability measure) of macroscopically compatible microstates $\Gamma_M$ is proportional to the degree of disorder of these microstates $\Gamma_M$.

Assuming the consistency of his 'bound information' concept $I_{Bb}$ and the coherence of his epistemic interpretation of $S_B$, Brillouin sought to technically specify the complementarist interrelation between $S_B$ entropy and $I_{Bb}$-information (originally attributed to the work of Szilard[59], following the historiographical narrative of von Neumann [1949] and Weaver [1949]) simply by defining $I_{Bb}$ as differences of $S_B$ quantities. Through this close definitional nexus, the content of $I_{Bb}$ would end up being conceptually related (and even identifiable) with the notion OI through the epistemic interpretation of $S_B$. That is, if $S_B$ can be understood epistemically as the (negative) amount of OI that the observer possesses over the actual microstate of the observed system, also by definition, $I_{Bb}$ could be assimilated as the (positive) amount of OI that the observer lacks over the actual microstate of the observed system. As Earman and Norton argued: "Brillouin's labeling of the quantity as 'information' is intended, of course, to suggest our everyday notion of information as knowledge of a system. (Earman and Norton 1999, p.8). In this sense, although Brillouin distinguished $I_{Bb}$ and OI as conceptually distinct notions (SIT, Chapter 1), they would end up interpretatively connected through the definitional relation of $I_{Bb}$ and $S_B$, and the epistemic conception of $S_B$. Although I will expand this idea later on, I argue that the very impossibility to conceptually differentiate between free and bound information is grounded precisely on this interpretative 'tension' of Brillouin's theoretical proposal.

In fact, upon this Brillouinian definition of the interrelation between entropy $S_B$ and bound information $I_{Bb}$ is articulated the conceptual core of his informational thermophysics, which he

[58] "Every physical system is incompletely defined. We only know the values of some macroscopic variables, and we are unable to specify the exact position and velocities of all the molecules contained in a system. We have only scanty, partial information on the system, and most of the information on the detailed structure is missing. Entropy measures the lack of information; it gives us the total amount of missing information on the ultramicroscopic structure of the system" (ibid, p.xii).
[59] "The origin of our modern ideas about entropy and information can be found in an old paper by Szilard, who did the pioneer work but was not [informationally] well understood at the time." (SIT, p.161).

called the 'Negentropy Principle of Information' (NPI) and detailed in Chapter 12 of SIT: "This point of view is defined as the negentropy principle of information, and it leads directly to a generalization of the second principle of thermodynamics, since entropy and information must be discussed together and cannot be treated separately" (SIT, p.xii). As Brillouin suggests in this quote, the NPI intends to reformulate the Second Law of TD not on the basis of the $S_{TD}$ concept, like Clausius (1865) (i.e., the entropy of a closed thermal system cannot decrease during its evolution), but by means of his informational notion of $I_B$ negentropy (i.e., the entropy of a closed thermal-informational system cannot increase during its evolution). However, due to its strictly statistical content, NPI is properly an informational reformulation not (at least directly) of the Carnot principle but of the Boltzmann theorem of BSM, where the evolution of systems to their equilibrium state can only be guaranteed statistically on average and considering the possibility of fluctuations: "The sum of negentropy and information may remain constant in a reversible transformation and will decrease otherwise. These inequalities apply only to average values. In all applications of Carnot's principle, the possibility of the occurrence of random fluctuations is presents" (SIT, p.155). Interestingly, Brillouin's informational proposal is not only based on the conceptual framework of BSM but also replicates in SIT some of the explanations of Boltzmann himself (1896). Illustratively, Brillouin used $SP_B$ (without further justification) to BSM-explain the equilibrium approach of thermal systems according to TD (see Section 2.3.3): "When the system is isolated and abandoned to itself, it naturally evolves toward the average structure of greatest probability" (p.153).

To illustrate the NPI content in concrete cases, Brillouin used as a case study a monoatomic ideal gas uniformly contained (i.e., in equilibrium state) in a thermally isolated vessel of volume V. This particular domain of application corresponds to what I previously called (see Section 2.5.2) the 'Golden Entropy Triangle', where it is possible to connect arithmetically the concept of entropy $S_{TD}$ with that of entropy $S_B$ by means of the Sackur-Tetrode formula, used by Brillouin (SIT, p.156) to compute its entropy value. However, since to define $I_{Bb}$ information quantities requires $S_B$ entropy differences and in this case study no macroscopic entropy changes occur (i.e., the gas is already in equilibrium), Brillouin proposes a dynamic scenario in which the gas expands from a subvolume V' of the container until it occupies all of V. During this process of free expansion of an ideal gas, the positive entropy difference $S(V) - S(V') = k (\ln V - \ln V')$ corresponds via NPI to an amount of information $I_{Bb}$ that, according to Brillouin's analysis, the gas loses during this process "Increase of entropy and loss of information proceed together. One may say that the gas progressively "forgets" the information" (p.157). Note that in this case of conceptual application the interpretative tension noted above emerges. While Brillouin states that as the gas microstate evolves toward the equilibrium macrostate-region it is the gas itself that loses this amount of information $I_{Bb}$ (i.e., it logarithmically increases the size of the final macrostate-region relative to that of the initial macrostate-region), this scenario can also be interpreted so that it is the observer itself that loses $I_{Bb}$ (or equivalently OI) over the actual microstate of the gas as it expands in V. This interpretative tension is recognized by Brillouin as the gas progressively "forgets" the information" (p.157), in reference to the (strictly epistemic and then non-informational, see Section

3.1.2) analysis by von Neumann of an identical scenario in 1932, expanding Szilard's (1929) results "J. von Neumann also considers the case of the observer forgetting the information, and states that this process means an increase of entropy" (SIT, p.157, fn, 4).

This is an example of 'negentropy $\rightarrow$ information' transformation, the first part of the 'info-thermal' cycle already mentioned in 1951 and developed in depth in SIT (p.164). To illustrate the second part of the cycle (i.e., 'information $\rightarrow$ negentropy' processes), Brillouin (SIT, Chapter 13) applied NPI on Maxwell's demon problem, first by rewriting Szilard's (1929) analysis using his own conceptual tools (i.e., $S_B$ or N, and $I_{Bb}$) and then by participating in Gabor's (1954) discussion of the possibility of measurements with a lower entropic cost than Szilard pointed out. Interpreting that measurement processes generate an amount of $I_{Bb}$ information equivalent to the increment of the macrostate-region measurement during that process (i.e., $I_{Bb} = \Delta S_B(\Gamma_M) + \Delta S_B(\Gamma_{M0}) = k \log (\Gamma_M/\Gamma_{M0}) > 0$), Brillouin (SIT, p.168) concluded that “every physical measurement requires a corresponding entropy increase, and there is a lower limit, below which the measurement becomes impossible (…) A more accurate discussion proved later that the exact limit is $k$ ln 2, or approximately 0.7 $k$ for one bit of information obtained”. This is what has been referred to in the literature as 'Szilard's principle' (see Earman and Norton 1999, p.5), reformulating in informational terms the epistemic thesis of Szilard (1929). In short, Brillouin (SIT, Chapter 13) pretends to explain informationally the Maxwellian demon as a sort of 'info-thermal' machine (à la Rothstein 1951) capable of completing a cycle 'negentropy $\rightarrow$ information $\rightarrow$ negentropy', whose 'exorcism' (i.e., demonstration of why the demonic operation fails to violate the Second Law) resides in the fact that the average entropy generated during the measurement process (ii) information $\rightarrow$ negentropy is always greater than or equal to the amount of information obtained during the process (i) negentropy $\rightarrow$ information (SIT, p.184).

### 3.2.4. Brillouinian Tradition in Informational Thermal Physics

The informational proposal of Brillouin’s (1956) SIT had a considerable impact within the thermophysical community of the late fifties and early sixties, reaching a highly popular second edition in 1962. This informational proposal constitutes a platform from which to systematically reconceptualize the theoretical apparatus of BSM by (i) certain informational notions (i.e., free information $I_{Bf}$ and bound information $I_{Bb}$), (ii) a complementary and technically-quantitatively defined connection between concepts $I_{Bb}$ and $S_B$ (generalizing the Second Law by means of NPI), and (iii) interpretatively specifying the content of $S_B$ by means of an epistemic (or OI-based) strategy. Among the main theoretical virtues with which Brillouin's (1956, 1962) NPI-based informational thermophysics was popularized within the scientific community (e.g., Leff and Rex, 1990; 2003) was (a) his suggestive informational treatment of the scientific measurement problem, and (b) the informational reformulation of the Szilardian (1929) solution to Maxwell's demon problem (understanding this problem not as a question about the foundations of TD-SM but as an opportunity to conceptually relate entropy and information).

Building on these features, the adherence of certain authors to Brillouin's proposal conformed what I would call the 'Brillouinian tradition' in informational thermophysics, understanding this as a current of physical thought articulated around the conceptual architecture and interpretative strategies delimited in Brillouin's work. Retroactively, one should integrate the conceptual proposal of Raymond (1950) and Rothstein (1951) within the Brillouinian tradition precisely because of their indispensable contribution to the development of SIT, see Section 3.2.2. An illustrative example of the immediate impact of Brillouinism in the early sixties can be found in Philip Rodd's 1964 paper entitled 'Some Comments on Entropy and Information'. Rodd (1964, p.145) defended from a strongly Brillouinian framework (i) a definite complementarist conception of the relation between entropy $S_B$ and information $I_B$, and (ii) an epistemic interpretation of $S_B$ as a measure of the microscopic uncertainty U of the observer, so that $I_B = U_0 - U_1$ and $S_0 - S_1 = I_B$. On the basis of Brillouin's exorcism in SIT, Rodd derived (see Earman and Norton 1999, p.11) an informational reformulation of the Second Law, i.e., $\Delta(S - U) \geq 0$.

Due to multiple reasons, such as Brillouin's short-lived advocacy and active popularization of his informational program (he published SIT at the age of 67 and died thirteen years later in 1969) and the lack of interpretative closure of his proposal, the Brillouinian tradition did not consolidate socio-institutionally into a school of physical thought with a clearly definable identity, unlike what happened with the Jaynesian tradition. In this sense, Brillouinism as an informationalist current constantly mutated during the second half of the twentieth century, incorporating interpretative lines in new conceptual architectures and diverging its interests and theoretical objectives among emerging intellectual trends, as I will show in detail in Chapter IV.

## 3.3. Jaynesian Informationalism

Along with Brillouin, the American physicist E. T. Jaynes (1922-1998) was one of the main actors responsible for the rapid informationalization of statistical physics in the fifties and afterwards. Jaynes developed during the first half of the 1950s and presented in a couple of papers (Jaynes 1957a, 1957b) a systematic reformulation of the theoretical architecture of SM from the technical and conceptual resources of Shannon's IT. This constituted what is now known as the 'Maximum Entropy' formalism or approach of SM, progressively generating an informational research program within modern statistical physics[60] radically different from the Brillouinism of his time. This separation of 'informationalist traditions' in the late 1950s can be characterized by the significant difference in their disciplinary pretensions, their conceptual architectures, and their

[60] Although Jaynes proposal also applies to quantum mechanics and other statistical physics domains, I will restrict my argument to his application in classical statistical physics.

interpretative strategies. In this sense, Jaynes' informational SM-apparatus based on his ‘Maximum Entropy Principle’ (MEP) consisted essentially of (i) grounding SM as a theory of statistical inference applicable to certain physical domains, and (ii) connecting the conceptual architecture of IT and GSM through an epistemic interpretation of $SP_G$ and IP (and derivatively also of $S_G$ and $H_{IT}$). Following the dialectical strategy that I traced in the case of the Brillouinian tradition, before evaluating the conceptual and interpretative foundations of Jaynes' proposal, it would be worthwhile to analyse the main lines of thought that converge in his work.

### 3.3.1. Intellectual Origins of Jaynesianism

> “it was his vagueness on these conceptual questions (allowing every reader to interpret the work in his own way) that made Shannon’s writings, like those of Niels Bohr, so eminently suited to become the Scriptures of a new Religion, as they so quickly did in both cases” (Jaynes 1978, p.24)

At the end of the 1970s, Jaynesianism as an informationalist program was already significantly consolidated and rooted within the physics community of the time. An example of this is the celebration of the Maximum Entropy Formalism Conference at MIT on May 2 and 4, 1978, in which the main adherents of this theoretical framework met to share results obtained, unpublished applications and speculations on possible lines of work. At this meeting, in a talk entitled 'Where do we stand on Maximum Entropy?' Jaynes (1978) himself reconstructed the main intellectual influences that converged in his proposal and the process of conceptual development of his framework during the forties and fifties (which somehow manifests the assimilation of MEP in 1978 as an established school). In this talk Jaynes poses his theoretical framework as the result of two intellectual trends: namely, on the one hand classical probability theory (Bernoulli, Laplace) and modern epistemic conceptions of statistical inference (Jeffrey, Cox), and on the other hand the evolution of SM up to the appearance of Shannon's IT, accepting the bandwagon-driven historiographical narratives promoted by Weaver (1949) and von Neumann (1949).

First, the classical probability theory of Bernoulli and Laplace played a central role in shaping MEP and Jaynesian thought, sometimes referred to as 'neoclassicism' in this sense. This is not simply because the MEP (as a probability assignment procedure) constitutes an explicit generalization of the 'principle of Insufficient Reason' (Uffink 1995, p.225-229), where possible states are assigned the same probability; but precisely because of the recognition that this probability assignment is intended to represent not a state of reality but an epistemic state of the agent. However, this mechanism limited the assignment of probabilities only to applicative domains in which equally probable cases could be enumerated. The introduction of Bayes' theorem (whose authorship belongs to Laplace, according to Jaynes [1978, p.6]) made it possible to introduce probabilities incorporating new data, thus being able to represent also dynamics of reasoning in fields such as scientific practices: "probability theory interpreted and used as the logic of human inference does rather well in dealing with problems of scientific reasoning -just as James Bernoulli and Laplace thought it would, back in the 18th Century" (Jaynes 1990, p.390).

In the work of Harold Jeffreys (1939), strongly positioned against Fisher's (1934) 'informational' anti-classicism (see Section 3.2.1), Jaynes found the tools to assume the usefulness of a Laplacian framework of epistemic probabilities in a largely hostile (i. e. mainly frequentist) intellectual environment. e. mainly frequentist) "As a student in the mid-1940's, I discovered the book of Jeffreys (1939) and was enormously impressed by the smooth, effortless way he was able to derive the useful results of the theory, as the sensible philosophy he expressed" (Jaynes 1978, p.9 ). In spite of its technical refinement and its vast range of scientific applications, Jeffreys' proposal did not contain any solid justification that would legitimize the recovery of a methodology (i.e., classical Laplacian probability) already considered within the statistical community as outdated. This legitimization was found in a short article by Cox (1946), where the justification task was replaced by the theoretical task of constructively developing a set of well-defined procedures so that it should be possible to represent degrees of belief in a mathematically consistent way "From the day I first read cox's article I have never for a moment doubted the basic soundness and inevitability of the Laplace-Jeffreys methods, while recognizing that the theory needs further development to extend its range of applicability" (Jaynes 1978, p.10). This new landscape of conceptual application to which Jaynes refers above was provided precisely by the appearance of Shannon's IT in 1948, which together with Cox's proposal (1946) would form the theoretical foundations of MEP, as already pointed out in the seventies by Jaynesian authors Costa and Tribus: "Jaynes' principle enters as a synthesis of Cox's results (...) and Shannon's results in communication theory" (Costa and Tribus 1974, p.240).

Jaynes (1978, p.23-27) recounts how the publication of Shannon's IT had a tremendous impact on his intellectual formation. Initially, he had significant problems in interpretatively specifying the content of the conceptual elements of this theory, for example: (i) interpreting sender and receiver in strongly agential terms, conceiving $H_{IT}$ as a measure of sender information and receiver uncertainty; or (ii) trying to understand IP epistemically from Shannon's (1948) originally frequentist presentation. As I argue below, this interpretative difficulty will be key in the conceptual evaluation of MEP in Chapter V. However, by algebraically manipulating the mathematical instrumentation of IT Jaynes quickly began to glimpse the possibility of establishing certain conceptual connections with GSM beyond mere formal analogies and of building a probability selection criterion on these promising tools[61]:

> "if you start to solve this problem of maximizing the entropy subject to certain constraints, you will soon discover that you are writing down very familiar equations. The probability distribution over messages is just the Gibbs canonical distribution with certain parameters. (…) Here was a problem of statistical inference (…) in which we are to decide on the best way of encoding a message, making use of certain partial [OI] information about the message. The solution turns out to be

[61] "In [$H_{IT}$], Jaynes saw the potential basis for a fully general criterion for Bayesian prior selection." (Vasudevan 2020, p.1119)

> mathematically identical with [GSM], which physicists had been trying, long and unsuccessfully, to justify in an entirely different way" (Jaynes 1978, p.26)

The new conceptual landscapes behind the technical fabric of IT presented themselves to Jaynes between 1949 and 1950 as an extremely fertile field (with application domains extending beyond BSM and GSM) on which to lay a new epistemic-probabilistic informational foundations for SM, far removed from the then promising frequentist-centric ergodic programs. As Jaynes remarked, this promising theoretical project intended to join the bandwagon was originally conceived in 1951 "All this was clear to me by 1951; nevertheless, no attempt at publication was made for another five years." (Jaynes 1978, p.27). The reasons for this time extension were both conceptual and of justificatory nature. First, Jaynes faced the theoretical problem of consistently and correctly reformulating the (discrete) conceptual elements of IT as $H_{IT}$ and IP by means of the representational resources of GSM as Γ-continuous densities ρ (see Section 2.5.4, 2.5.5 & 5.2.2). Second, Jaynes wanted to gauge the susceptibility of his intellectual environment by expressing his theoretical intentions to George Uhlenbeck, who in the summer of 1951 was teaching a course on SM at Stanford: "I had expected, naively, that he [Uhlenbeck] would be enthusiastic about Shannon's work, and as eager as I to exploit these ideas for [SM]. Instead, he (...) adamantly rejected all suggestions that there is any connection between entropy and information" (Jaynes 1978, p.27). In this encounter, Uhlenbeck posed to Jaynes the interpretative maze to which Jaynesianism was condemned: while information (recall that underlying the epistemic interpretation of probabilities lies OI) demands an agent, the content of $S_{TD}$ cannot be expressed agentially precisely because it is an experimentally meaningful physical measure independently of the agent specifying its content (ibid., p.27-28).

The way out of this conceptual maze (in which he was imprisoned between 1951-1956) was found by Jaynes in the development of an interpretation specifying both the thermophysically meaningful content of $S_G$ (which depends on its connection with $S_{TD}$) and its dependence on the agent's epistemic states. Among the intellectual sources that enabled such a strategy one could find, first, Tolman's (1938) defense of the GSM framework from an epistemicist perspective akin to Jaynes' theoretical interests "The science of [SM] has the special function of providing reasonable methods for treating the behavior of mechanical systems under circumstances such that our knowledge of the condition of the system is less than the maximal knowledge which would be theoretically possible" (Tolman 1938, p.1). Apart from recognizing his affinity with Tolman's interpretative strategy[62], Jaynes found in the epistemic-informational conception of SM-entropy advocated by Ter Haar in 1954 (quoted in Jaynes [1957a, fn.1]) the key piece he needed "We can express this slightly differently by stating that the entropy measures our lack of knowledge or lack of detailed [OI] information, since [μ(Γρ)] (...) gives us a measure for the volume in Γ-space in which the representative point can be found, once the energy is stated to lie within a given interval" (Ter

[62] "A notable exception [of mentioning SG] is the monumental work of R. C. Tolman (1938) (…) [who] repeatedly stresses the superiority of Gibbs' approach, although he still attempts to base the second law on coarse-graining" (Jaynes 1965, fn.3).

Haar 1954, p. 232 ). Once with these related interpretative resources, *l'esprit d'escalier* whispered in the mid-fifties to Jaynes the solution to the problem that Uhlenbeck posed to him several years earlier[63], feeling then legitimized to bring to light his informational proposal to reconceptualize statistical physics.

### 3.3.2. Jayne's Maximum Entropy Principle

Jaynes' SM-proposal first appeared in two papers published in 1957 in The Physical Review under the title 'Information Theory and Statistical Mechanics' and 'Information Theory and Statistical Mechanics II', after receiving both negative (directed at its lack of applicability and its epistemicist views, see Jaynes [1978, p.28]) and positive comments from the reviewers. In the first, Jaynes (1957a) laid out the conceptual and interpretative foundations of his informational framework, defining MEP and applying it to equilibrium-SM, while in the second (1957b) he delimited its possible application to non-equilibrium-SM and extension to the matrix-formalism of quantum mechanics. As he notes already in the introduction, his main theoretical goal was intended to be not only to incorporate in SM the new statistical tools developed in the disciplinary context of Shannon's IT, but to employ these to recast SM as a theory of inference on a conceptual ground free of mechanical assumptions (i.e., ergodicity, Birkhoff-von-Neumann metric transitivity, etc.).

To achieve this goal, Jaynes establishes MEP as the theoretical core that articulates his conceptual architecture and his interpretative strategies is MEP, which is defined as an algorithmic procedure (à la Gibbs [1902]) or constructive criterion (à la Cox [1946]) for assigning probability distributions over the $\Gamma$ of a physical system considering: (i) that the information (OI) that the agent possesses about the representative microstate of the target system is not complete but partial, and (ii) starting from a certain knowledge (also OI) macroscopic of the system. On the one hand, this macroscopic knowledge that the agent possesses (linked to a set of observational values Ai) is represented in the context of MEP as a constraint C over the space of possible distributions $\rho \in P$. On the other hand, the probability distribution $\rho$ that is selected within P must be the least biased concerning the lack of information (OI) of an agent about the actual microstate, which according to Jaynes is precisely the one that maximizes $H_{IT}$. The conceptual justification provided by the author to legitimize the latter idea lies precisely in interpreting $H_{IT}$ as the only quantity that meets certain consistency criteria and that also represents the 'intuitive' degree of uncertainty that the agent possesses *about* the actual microstate (therefore, connecting interpretatively $H_{IT}$ with OI): "The great advance provided by information theory lies in the discovery that there is a unique, unambiguous criterion for the "amount of uncertainty" represented by a discrete probability

[63] "The entropy of a thermodynamic system is a measure of the degree of ignorance of a person whose sole knowledge about its microstate consist of the values of the macroscopic quantities [Ai] which define its thermodynamic state" (Jaynes 1978, p.28).

distribution, which agrees with our *intuitive notions* that a broad distributions represents more uncertainty than does a sharply peaked one" (Jaynes 1957a, p.622. Italics are mine).

As Jaynes (1957a, p.621) himself acknowledged, MEP involves a radical reversal of the conceptual architecture of SM and a rejection of the original theoretical claim (mostly associated with BSM, though also GSM [e.g., Robertson 2020]) to derive the macroscopic properties of molecular systems by means of the dynamical properties of their components: "Previously, one constructed a theory based on the equation of motion, supplemented by the additional hypothesis of ergodicity, metric transitivity or equal a priory probabilities, and the identification of entropy was made only at the end, by comparison of the resulting equations with the laws of phenomenological thermodynamics. Now, however, we can take entropy as our starting concept" (Jaynes 1957a, p.621). However, the theoretical weight to be borne by $H_{IT}$ as a primitive notion of a 'physical' (I will clarify this point later) domain such as SM is intended to be alleviated by Jaynes through a robust conceptual connection with the thermophysically meaningful notion entropy $S_G$. In contrast to what von Neumann (1949) promulgated in the origins of the bandwagon (and the role it played in its intellectual genesis [Jaynes 1978, p.24-16]), Jaynes (1957a) argued that the conceptual connection between $H_{IT}$ and $S_G$ could not rest on the merely formulaic relationship between their mathematical expressions, but on a coherent interpretative identification (see Section 2.5.1) "The mere fact that the same mathematical expression $-\Sigma\, p_i \log p_i$ occurs both in [SM] and in [IT] does not in itself establish any connection between these fields. This can be done only by finding new viewpoints from which [$S_{TD}$] and [$H_{IT}$] appears as the same concept." (1957a, p.621). This intellectual caution of Jaynes against formal analogies was permanent throughout his career[64].

I must open here a brief aside to highlight that Jaynes further defended that $S_G$ (and not $S_B$) was closely connected conceptually with $S_{TD}$ (also conceived as an 'anthropomorphic' or agent-based notion[65]) precisely because of the equal numerical relations derivable from their predictions: "Gibbs formula gives the correct entropy, as defined in phenomenological thermodynamics (...) objections to the Gibbs [$S_G$] are immediately refuted by the fact that Gibbs canonical ensemble does yield correct thermodynamic predictions" (Jaynes 1965, p.391). Thus, Jaynes seems to accept arithmetical connections and at the same time to reject merely formulaic-syntactic connections as a criterion for conceptual identification, as he defends in the cases of $S_{TD}$-$S_G$ and $S_G$-$H_{IT}$, respectively. Returning to the relation between the latter conceptual pair, Jaynes dissociates himself from the 'complementarist' tradition (Wiener, 1948; Brillouin, 1956) and defends an

[64] "It perhaps takes a moment of thought to see that the mere fact that a mathematical expression like $-\Sigma\, p_i \log p_i$ shows up in two different fields, and that the same inequalities are used in two different fields does not in itself establish any connection at all between the fields. Because, after all $e^x$, $\cos\theta$, $J_0(z)$ are expressions that show up in every part of physics and engineering. Every place they show up (…) nobody interprets this as showing that there is some deep profound connection between, say, bridge building and meson theory." (Jaynes 1983, quoted on Ben-Naim 2008, p.25-26).

[65] "entropy is an anthropomorphic concept, not only in the well-known statistical sense that it measures the extent of human ignorance as to the microstate. Even at the purely phenomenological level, entropy is an anthropomorphic concept. For it is a property, not of the physical system, but of the particular experiments you or I choose to perform on it" (Jaynes 1965, p.398).

'identificationist' interpretative strategy of $H_{IT}$ and $S_G$ based theoretically on an epistemic understanding of the probabilistic representational resources (densitie of both. As mentioned in Section 2.3.1, in this epistemic interpretation of $SP_G$ (key in the conceptual architecture of MEP) Jaynes sought to dissociate himself from the subjectivist-personalist positions of authors such as Ramsey and Finnetti (Uffink 2011) to develop a conception based on the epistemic states of scientific agents, but without renouncing the objective content of these states: “Our probabilities *and the entropies based on them* are indeed “subjective” in the sense that they represent human [OI] information. But they are completely “objective” in the sense that they are determined by the [OI] information specified, independently of anyone’s personality, opinions, or hopes. It is “objectively” in this sense that we need if information is ever to be a sound basis for new theoretical developments in science” (Jaynes 1990, p.390. Italics are mine).

Although not made explicit in detail in his 1957a and 1957b papers, this line of interpretation is equally applicable to IP under the implicit assumption of an agent possessing $\varphi$ (see Jaynes 1978, p.24). As displayed in the last quote of the previous paragraph, it is precisely on this epistemic interpretation of $SP_G$ and IP that the theoretical weight of the meaning of their respective concepts $S_G$ and $H_{IT}$ rests, which is also specified by Jaynes in an explicitly epistemic (but also objective[66]) way: "The essential point in the argument above is that we accept the von-Neumann-Shannon expression for entropy, very literally, as a measure of the amount of uncertainty represented by a probability distribution; thus entropy becomes the primitive concept with which to work, more fundamental even than energy" (Jaynes 1957a, p. 629). Thus, although the concept (once identified) $S_G$-$H_{IT}$ is conceptually primitive within the Jaynesian SM-formalism, the interpretive specification of its content is later or derivative with respect to the interpretive specification of (once identified) $SP_G$-IP both in the architecture of MEP and in Jaynes' own intellectual evolution. That is, whether $S_G$-$H_{IT}$ can be coherently understood as a measure of the agent's uncertainty about the actual microstate depends (within this convoluted interpretative strategy of Jaynes) on $SP_G$-IP representing the agent’s degrees of belief about the occurrence of individual microstates.

This 'integral' epistemic interpretation of Jaynes is not only directed towards the particular concepts of probability ($SP_G$-IP), information ($H_{IT}$ and OI) and entropy ($S_G$) included in MEP, but also aimed to provide an informational way of understanding SM not as a theory about a physical domain (such as TD, BSM or GSM) but as a theory about the operation of statistical predictions or inference in SM: "If one considers statistical mechanics as a form of statistical inference rather than as physical theory, it is found that the usual computational rules, starting with the determination of the partition function, are an immediate consequence of the maximum-entropy principle" (1957a, p. 620 ). In this sense, the theoretical goal of Jaynesian-SM (also 'predictive-SM' [Parker 2011, p.837] or simply MEP) would not be to generate statistical descriptions of the molecular behavior of macroscopic systems (as is the case for BSM or GSM) but to exploit statistical techniques that allow to increase the effectiveness of SM-predictions. The MEP

[66] “we can also say on the “objective” side, that [$\Gamma\rho$] measures the degree of control of the experimenter over the microstate, when the only parameters he can manipulate are the usual macroscopic ones” (Jaynes 1965, p.396).

procedure to generate SM-predictions[67] is as follows: (i) we measure certain macrovariables $A_i$ of the system at $t_0$, (ii) considering the value of $A_i$ as constraint C we assign the density ρ that maximizes $H_{IT}$, (iii) we let ρ evolve dynamically over Γ during $t_0$-t, obtaining predictively the values of $A_i$ at t, and finally (iv) we introduce (via MEP-averages, in contrast to GSM-averages) a new density ρ* that maximizes $H_{IT}$ considering the new constraints $A_i(t)$. This same procedure was used by Jaynes (1983) to develop his 'incredibly short' proof of the Second Law (see Parker, 2011, p.837), showing that $S_G(t_0) \leq S_G(t)$ for any ρ* maximizing $H_{IT}$ under $A_i(t)$.

Although Jaynesian-SM is intended to be legitimized (almost exclusively) as a theory of SM-prediction, its author argues that this conceptual framework also may offer interpretative keys to an informational understanding of certain physical principles. Illustratively, one of the dynamic consequences of the epistemic interpretation of $SP_G$-IP is that the meaning of DP and Liouville's theorem could also be specified epistemically (see Section 2.3.3), which according to Jaynes would illuminate the content of (a particular GSM reformulated version, à la Tolman [1938]) of the Second Law: "we finally see the real reason for the second law: since phase volume is conserved in the dynamical evolution, it is a fundamental requirement on any reproducible process that the phase volume compatible with the final state cannot be less than the phase volume which describes our ability to reproduce the initial state" (Jaynes 1965, p.396). Later on, neo-Jaynesian authors (e.g., Parker 2011, fn, 11) deepened in the dynamic dimension of Jaynes interpretative line, problematizing whether the ability of agents to preserve their degree of uncertainty during the evolution of ρ might correspond to the fact that density ρ follows a Liouvillean dynamic over Γ. This point is especially decisive in the theoretical legitimization of MEP in the context of nonequilibrium-SM precisely because (analogously to coarse-graining GSM, [Tolman 1938]) the plausible numerical validity of its predictions is constitutively independent of the dynamics of the system due to the indispensable use of MEP-averages. Against this type of objection, Jaynes argued years later that "The point is that we are not ignoring the dynamics, and we are not getting something for nothing, because we are asking (…) only for predictions of experimentally reproducible things, and for these all circumstances that are not under the experimenter control must, of necessity, be irrelevant[68]" (Jaynes 1978, p.29).

Finally, Jaynes intended by means of his MEP proposal to rest all the interpretative weight of SM on a particular way of understanding the use of probabilities in the natural sciences based on Laplace's dictum "probability is nothing but common-sense reduced to calculation" (Jaynes 1957a, p.626). As is clear from his posthumous work *Probability Theory: The Logic of Science*, Jaynes (2003) ambitiously attempted throughout his academic career to reconceptualize the disciplinary function of probability theory and statistical inference in a way parallel to what von Neumann (1949) directly attempted with Shannon's IT (Section 3.1.2): namely, as a technically sophisticated

[67] "In the problem of prediction, the maximization of entropy is not an application of a law of physics, but merely a method of reasoning which ensures that no unconscious arbitrary assumptions" (Jaynes 1957a, p.630).
[68] It should be noticed that this argument is almost identical to Robertson (2020) justification of GSM coarse-graining procedure.

platform from which to develop a naturalized epistemology of the natural sciences: “We suggest that the proper tool for incorporating human [OI] information into science is simply probability theory -not the currently taught “random variable” kind, but the original “logical inferential” kind of James Bernoulli and Laplace. (…) When supplemented by the notion of [$H_{IT}$] this becomes a mathematical tool for scientific reasoning of such power and versatility that we think it will require a century to explore its capabilities” (Jaynes 1990, p.387).

### 3.3.3. Jaynesian Tradition of Informational Thermal Physics

The appearance in 1957 of the Jaynesian framework based on MEP generated enormous interest within the physics community at the time. More than 2000 requests for reprints of his 1957a paper were solicited during the first year (Jaynes 1978, p.28), which is a significant figure for the time. Even so, the intellectual environment of the time was relatively cautious about the many theoretical promises with which Jaynes popularized his MEP in his foundational papers, posing to solve a vast number of difficult physical problems (including the derivation of TD-irreversibility by means of SM) by a simple statistical procedure. For this reason, Jaynes devoted himself from 1956 until 1962 to refining his theoretical apparatus and to trying to pinpoint the precise conceptual origin (on which to base his theoretical legitimacy) of the many derivable predictive successes of MEP. Regarding the latter, Jaynes (1978, p.29-30) concluded that MEP works for precisely the same reason that $SP_B$ is "the real reason why all [SM] works" in both equlibrium-SM and non-equilibrium-SM (which would be difficult to defend now, see Section 2.3.3). In any case, in the sixties the group of adherents to Jaynes MEP proposal grew considerably, stabilizing in the seventies as a 'school of physical thought' (symbolized by the celebration of the Maximum Entropy Formalism Conference at MIT on 1978, see the beginning of Section 3.3.1). Another sign of the socio-institutional consolidation of the Jaynesian tradition (in contrast to the Brillouinian) in the 1970s as a physical program are the first critical attacks directed not to its technical procedures and interpretative strategies but to the whole MEP framework (Friedman and Shimony, 1971; Friedman, 1973; Shimony, 1973). Beyond this criticism, which I analyse below, Jaynes actively (and sometimes harshly [Jaynes 1985]) defended the interpretative architecture of his proposal during the more than four decades that separated his foundational publications in 1957 (when he was only 35 years old) until his death in 1998. As we will see in Chapters IV, Jaynes' intellectual figure crosses in multiple ways, either as a protagonist or as a reference scientist, practically the entire historical unfolding of information physics.

…

In short, the Brillouinian and Jaynesian traditions emerged during the late-1950s as the two most sophisticated programs exploiting information concepts within the domain of thermal physics. If it were legitimate to speak of ‘informational thermophysics’ as a proper domain of physical research (a question I will return to later), Brillouin and Jaynes would have a historical role analogous to that of Boltzmann and Gibbs for SM, respectively. Both emerge after complex processes of intellectual evolution linked to the effects of Shannon's bandwagon in the thermophysics of the fifties (via Wiener, von Neumann and Weaver, see Section 3.1), either by capitalizing on its popularity (à la Brillouin) or its conceptual resources (à la Jaynes). Brillouinism and Jaynesianism constituted two completely different strategies of interweaving interpretatively (complementarily or identificatively, respectively) previously existing notions of entropy ($S_B$ and $S_G$), information ($H_{IT}$ and OI) and probability ($SP_B$ and $SP_G$), or also developing new concepts (e.g., $I_{Bb}$) that can satisfy the new theoretical needs that emerged in this intellectual context. This difference also occurs at the level of their main intellectual aims: while Brillouin (1956) intended to explain how observer information in BSM has a certain thermophysical significance in certain scientific processes such as measurement, Jaynes (1957a) intended to exploit information theory interpretatively to increase technically the predictive effectiveness in SM. In any case, this cold war between the Brillouinian and Jaynesian traditions did not last long within the history of information physics, as I am going to defend in Chapter IV.

| **Informational Traditions** | **Brillouinism** | **Jaynesianism** |
|---|---|---|
| Foundational Works | Brillouin (1956) | Jaynes (1957a, 1957b) |
| Main Intellectual Influences | Wiener, Gabor, Rothstein | Jeffrey, Cox, Shannon |
| Core Theoretical Principle | Negentropy P. Information (NPI) | Maximum Entropy P. (MEP) |
| Function of the Principle | Enable Info-Explanations in BSM | Enhance Prediction in SM |
| Background SM-Formalism | BSM (non-committal) | GSM (committal) |
| Background Purpose | Describe φ via TD (BSM) | Reformulate SM via φ-Prob |
| Interpreting Theory | Generalize TD into IT-TD | Generalize SM into Prob |
| Core Entropy Concept | $S_B$ (identified with $S_{TD}$) | $S_G$ (identified with $S_{TD}$) |
| Core Information Concept | $I_{Bb}$ (distinguished from OI) | $H_{IT}$ (identified with OI) |
| Core Conceptual Connection | Complementarist ($-S_B = I_{Bb}$) | Identificationist ($S_G = H_{IT}$) |
| Probability Concept | $SP_B$ (explains non-equilibrium-SM) | $SP_G$, IP |
| Probability Interpretation | - | Epistemic (OI) - Subjective |
| Entropy Interpretation | Epistemic (OI) / Disorder (OO) | Epistemic (OI) / Objective |
| Information Interpretation | Agent (OI) / Physical ($S_B$) | Epistemic (OI) / Objective |

**Table 1**. Comparative Analysis between Brillouinian and Jaynesian Traditions

## 3.4. Carnapian Criticism of Shannon's Bandwagon

The informationalization of thermal physics since the 1950s was not only a pivotal process in the history of this scientific domain and this school of physical thought, but also (as I intend to defend here) in the history of the philosophy of the concepts of entropy, information and probability. The historical role played by the philosophical critique of the conceptual and interpretative foundations of informationalism has been systematically disregarded by part of the philosophical literature (with anecdotal exceptions, e.g., Köhler [2001]), mainly because the results obtained in this direction were historically buried due to the low tolerance to criticism on the part of certain informationalist authors of the time. In particular, it was the philosopher Rudolf Carnap (1891-1970), together with his disciples Bar-Hillel and Shimony, who developed not only a systematic evaluation of the main conceptual landscapes (entropy, information, and probability) of this current, but also a critical analysis of the interpretative incoherencies and conceptual inconsistencies underlying certain uses of these notions in the early 1950s and by the Brillouinian and Jaynesian traditions.

### 3.4.1. Early Criticism of Shannon's Bandwagon

The fact that the popularization of Shannon's IT immediately generated a tsunami of application interest in areas of scientific research does not mean that there were no voices within the scientific community that warned of the risks that this could entail. A particular illustrative (and somewhat prophetic) example is the psychologist and computer scientist Joseph Licklider, who warned that "It is probably dangerous to use this theory of information in fields for which it was not designed, but I think the danger will not keep people from using it" (Licklider 1950). This kind of cautious and intellectually conservative attitude towards the possible theoretical applications of IT rooted especially (as one would expect) among the information theorists themselves in the early fifties, such as Elias, Fano[69] or Shannon himself. The latter was one of those who most actively warned of the dangers of jumping on the bandwagon of his own theory "Although this wave of popularity is certainly pleasant and exciting for those of us working in the field, it carries at the same time an element of danger" (Shannon 1956). However, as the number of scientific productions (publications or presentations at conferences) applying IT in scientific fields surpassed that of communication engineering (Kline 2015, p.115), caution turned into more or less explicit criticism. This is the case of Peter Elias, who in a 1958 text entitled 'Information Theory, Photosynthesis and Religion' parodied (similarly to what Alan Sokal [1996] did in the 1990s

[69] "I believe, for instance, that expressions for the amount of information such as those which appear in Shannon's paper can only be used in problems where the transmitter and receiver are well-identifiable, and where one can assume the existence of an ensemble of messages with known statistical characteristics." (Robert Fano to Colin Cherry, February 12, 1953, CC, 12/3/1.).

regarding the exploitation of modern physics concepts by continental philosophy) the intellectual tendency to exploit the technical and conceptual resources of IT in fruitless applicative domains.

In addition to these famous general criticisms of IT misuses[70], other authors also put their critical focus on key intellectual trends in the informationalization of thermal physics. In private conversation with Colin Cherry, Robert Fano, a member of what has become known as the 'Americal School of IT' (Kline 2015, p.112) along with Shannon and Elias, differentiated in 1952 between the informationalist positions he encountered at the second London symposium organized by Cherry, namely: (i) communication theory (Shannon's IT), (ii) wave form analysis (Gabor's IT), (iii) classical statistics (Fisher's proposal) and (iv) miscellaneous philosophical speculations (e.g., Wiener's IT). Thus, Fano separated the informational trends according to their theoretical pretensions, separating Shannon's (1948) modest aims from MacKay's broad 'English-School' proposal (integrating (ii) and (iii), see Section 3.2.1) and from Wiener's ambitious and speculative cybernetics (i.e., (iv) above, see Section 3.1.1). As far as the latter line of research is concerned, much of the criticism received by the scientific community is directed not so much at Wiener's (1948) immeasurable degree of speculation (not to mention his interpretative vagueness and conceptual indefiniteness) but at his attempt to converge on Shannon's bandwagon. In this sense, the Russian translator of Shannon's original article intended to 'clean' the formal apparatus of IT from any kind of interpretative connotations alien (being especially hard on cybernetics) to the original proposal[71]. Interestingly, this interpretative cleansing of Shannon's IT in the Soviet intellectual milieu of the 1950s would influence the emergence of the 'Formalist School in IT' (Lombardi et al. 2016a) led by the mathematician Aleksandr Khinchin (1957), who provided solid mathematical foundations to IT and conceived it as a mere statistical tool (i.e., a branch of probability theory) without any content.

As it was analysed, the successful popularization of Shannon's IT by Weaver (1949) and the pretensions of von Neumann (1949) and Wiener (1948, 1950) generated a progressive climate of intellectual excitement within statistical thermal physics during the fifties. This, added to the enormous impact of the systematic theoretical proposals of Brillouin (1956) and Jaynes (1957a,

---

[70] "There is something frustrating and elusive about information theory. At first glance, it seems to be the answer to one's problems whatever these problems may be. At second glance it turns out that it doesn't work out as smoothly or as easily as anticipated (…) So nowadays one is not safe in using information theory without loudly proclaiming that he knows what he is doing and that he is quite aware that this method is not going to alleviate all worries." (Quastler 1955, p.2).

[71] "The terminology of the statistical theory of electrical signal transmission and a number of its concepts are utilized by some foreign mathematicians and engineers in their speculations related to the notorious 'cybernetics'. For example, building upon superficial, surface analogies and vague, ambiguous terms and concepts, Wiener, Goldman, and others attempted to transfer the rules of radio communication to biological and psychological phenomena, to speak of the 'channel capacity' of the human brain, and so on. Naturally, such attempts to give cybernetics a scientistic look with the help of terms and concepts borrowed from another field [TD-SM] do not make cybernetics a science; it remains a pseudo-science, produced by science reactionaries and philosophizing ignoramuses, the prisoners of idealism and metaphysics. At the same time, the notorious exercises of philosophizing pseudoscientists cast a shadow on the statistical theory of electrical signal transmission with noise—a theory whose results and conclusions have great scientific and practical importance." (Russian translation [1953] of Shannon 1948).

1957b), caused a low tolerance to the criticism of informationalist positions from the late 1950s onwards, even becoming assimilated within the pedagogical circuits of this discipline (e.g., Amnon Katz's [1967] well-known textbook *Principles of Statistical Mechanics*). But before this unstoppable informational tsunami completely flooded all the nooks and crannies of thermophysics, some critical voices were able to breathe briefly. This is particularly the case of physicist Dirk ter Haar, who in his popular textbook *Elements of Statistical Mechanics* of 1954 criticized the physical interpretations of the content of $H_{IT}$ and its complementary relationship with entropy concepts “The relationship between entropy and lack of information had led many authors, notably Shannon, to introduce “entropy” as a measure for the information transmitted by telephone lines, and so on, and in this way entropy has figured largely in discussions in information theory. It must be stressed here that the entropy introduced in information theory is not a thermodynamical quantity and that the use of the same term is misleading. It was probably introduced because of the rather loose use of the term “information”” (ter Haar 1954, p.232). As can be seen in this quote, the complementarist interpretations of the information-entropy connection had already taken root to some extent within the imaginary of the scientific community. Beyond these punctual critical comments by ter Haar, one must dive deep into the intellectual history of this period to find (or rather 'reconstruct') an argumentatively robust and critical attack on the conceptual consistency and interpretative coherence underlying this convoluted informational trend.

### 3.4.2. Carnapian Explication of Probability and Information Conceptual Landscapes

To satisfactorily understand the dimensions of the Carnapian critique of the informationalization of thermal physics, we must go back to the development of his late (from the mid-1940s) methodology of conceptual analysis, whose current validity has been greatly vindicated in the last two decades (e.g., Maher 2007). As I proceed to detail below, Carnap systematically applied his method of conceptual evaluation on each of the main conceptual landscapes explored in Chapter II and on which the currents of thought specified in this chapter took root. This method is known as 'explication', the classic exposition of which can be found in his 1962 'Logical Foundations of Probability'. Long before this text, Carnap first displayed the functioning of conceptual explication in his famous 1945 article 'The Two Concepts of Probability', stating that: “in each case of an explication, we call the old concept, used in more or less vague way either in every-day language or in an earlier stage of scientific language, the explicandum; the new, more exact concept which is proposed to take the place of the old one the explicatum” (p.513). Therefore, explication should not be understood as a merely descriptive methodology (i.e., characterizing previously existing concepts) but above all as a philosophical tool with an important normative or even constructive content (i.e., creating new or redefining old concepts), along the lines of what is currently known as 'conceptual engineering' (Floridi 2019).

As the title of the paper anticipates, the first concept on which this explication-based analytical methodology was applied was that of probability, whose main theoretical problem it poses is that of "finding an adequate definition of the concept of probability that can provide a basis for a theory of probability." (ibid). That is to say, in the case of this conceptual task, explication does not properly consist in developing a new concept but in technically defining an already existing conceptual landscape, as an intuitive notion since antiquity and minimally mathematically defined since the time of the Pascal-Fermat correspondence (see Section 2.3). This, according to Carnap, is an important conceptual issue within the scientific area "Thus we have here an instance of that kind of problem -often important in the development of science and mathematics- where a concept already in use is to be made more exact or, rather, is to be replaced by a more exact new concept" (ibid). In his exposition of this methodology, this author distinguished between the conceptual landscape object of the explication or 'explicatum' (plural 'explicata') and the conceptual landscape resulting from the explication or 'explicanda'.

In the case of the explicatum required by a technically-robust theory of probability, this would correspond to the underdefined semantic content associated with the pre-scientific or pre-systemically defined probability concept[72], which should be traced in the multiplicity of non-regularized uses and applications that emerge from this wild intuition-driven conceptual region: "When we look at the formulations which the authors themselves offer in order to make clear which meanings of 'probability' they intend to take as their explicanda, we find phrases as different as "degree of belief", "degree of reasonable expectation", "degree of possibility", "degree of proximity to certainty", "degree of partial truth", "relative frequency" (ibid). Note that although the concept of probability has an axiomatically structured body around Kolmogorov's axioms, its enormous interpretative potential causes its possible meanings to unfold irregularly and disorganized beyond this axiomatic definition. In this sense, Carnapian explication as a conceptual-analytic methodology can be understood (along the lines of my proposal in Chapter II) as a meticulous evaluative-normative process of regularizing the manifold of different uses and meta-conceptually organizing all of the plausible interpretative strategies of specifying the semantic content of a (somehow underdefined) conceptual landscape.

The result of the Carnapian explication of probability was the technical refinement of the previous landscape-explicata into two distinct interpretative modes of specifying its meaning, delimiting two broad conceptual regions (prima facie non-overlapping) grouped under certain common semantic patterns: "we are, of course, distinguishing between two explicanda and not between the various explicata offered by these theories (...) The two concepts are: (i) probability$_1$ = degree of confirmation, (ii) probability$_2$ = relative frequency in the long run" (ibid). As Carnap (ibid, p.523) recognized, the conceptual distinction underlying this explication is not orthogonal to other interpretative coordinates that one usually find in the literature on probabilistic concepts, such as the mathematical-philosophical or subjective-objective distinctions. This explication of the probability landscape allowed Carnap to develop analytical tools to evaluate forms of 'conceptual confusion' between, on the one hand, the interpretative strategies that conceptual users claim to

[72] "the explicatum for every theory of probability is the pre-scientific concept of probability, i.e., the meaning in which the world 'probability' is used in the pre-scientific language" (p.517).

defend, and on the other hand, the methodologies that these users actually employ to specify conceptual meaning. Carnap gives as a paradigmatic example of conceptual confusion within the probability landscape (particularly in the realm of inductive logic) what he calls 'psychologism' (expanded on Carnap [1963]), where “Some authors, from Laplace and other representatives of the classical theory of probability down to contemporary authors like Keynes and Jeffreys, use subjectivistic formulations [i.e., epistemic interpretative strategies] when trying to explain what they take as their explicandum (…) However, an analysis of the work of these authors comes to quite different result if we pay more attention to the methods the authors actually use (…) most and perhaps all these authors use objectivistic rather that subjectivistic method” (ibid, p.525).

During the second half of the forties and early fifties, Carnap continued to develop his method of conceptual analysis as a support for his program of founding inductive logic. Due to the already discussed intellectual tsunami that occurred in the scientific community during this historical period, the next conceptual landscape on which Carnap applied his explanatory method was the one linked to informational notions, particularly with those not theoretically defined (hence, in the semantic environment of OI). As argued in Section 2.4.4, this explication of the pre-theoretical concepts of information was theoretically indispensable for the co-development, together with the also philosopher Yehoshua Bar-Hillel, of his SIT proposal, deployed in ‘An Outline of a Theory of Semantic Information’ (1952) and ‘Semantic Information’ (Bar-Hillel y Carnap 1953). As they stated: “We intend to explicate the pre-systematic concept of information [OI], insofar as it is applied to sentences or propositions and inasmuch as it is abstracted from the pragmatic conditions of its use. We shall then define, on the basis of this systematic concept of semantic information, various explicata for the presystematic concept (or concepts) of amount of semantic information” (Carnap and Bar-Hillel 1952, p.3).

The key here is that this explication of the informational landscape not only resulted in the development of its technical and quantitative concept of SIT as explanandum (Section 2.4.4), but also (similar to what happened with the explication of probabilities) in obtaining new analytical tools with which to evaluate conceptual misuses within this landscape. In the case of OI, being a pretheoretical and not technically defined concept, its uses or possible applications are not disciplinarily regulated, so one could not strictly speak of 'misuses' of OI (beyond those derived from the domain to which it is applied). However, the intentional or unintentional exploitation of the conceptual properties of OI (agential, semantic, and epistemic) from a technical-theoretical information concept that does not explain these conceptual properties constitutes a 'conceptual confusion'. This is the paradigmatic case of the explicit uses (encouraged by its Weaverian popularization, Section 3.1.3) of $H_{IT}$ through the conceptual resources of OI: “Impatient scientists in various fields applied the terminology and the theorems of statistical information theory to fields in which the term ‘information’ was used, presystematically, in a semantic sense, i.e., one involving contents or designata of symbols, or even in a pragmatic sense, i.e., one involving the users of these symbols” (Bar-Hillel and Carnap 1952, p.147).

In this sense, the fact that $H_{IT}$ was constitutively defined by Shannon (1948) (and in general within IT) as a technical concept insensitive to semantic content constitutes for Carnap and Bar-Hillel (1952) a criterion for evaluating the semanticizing interpretative uses of $H_{IT}$ as 'misinterpretations' or 'misconceptions' “It has, however, often been noticed that this [semantic] asceticism is not always adhered to in practice and that sometimes semantically important conclusions are drawn from officially semantics-free assumptions. In addition, it seems that at least some of the proponents of communication theory have tried to establish (or to reestablish) the semantic connections which have been deliberately dissevered by others.” (Carnap and Bar-Hillel 1952, p.2). However, as years later Bar-Hillel rightly pointed out in his 1955 ‘Examinations of Information Theory', we must differentiate here between the explicit-intentional misinterpretation of $H_{IT}$ (interpreting it as explication of OI) and the natural tendency to cognitively exploit OI to understand the meaning of $H_{IT}$: “However, it is psychologically almost impossible not to make the shift from the one sense of information, for which this argument is indeed plausible, i.e., information = signal sequence, to the other sense, information = what is expressed by the signal sequence, for which the argument loses all its persuasiveness.” (Bar-Hillel 1955, p.94).

### 3.4.3. Explicating Entropy and Carnapian Arguments against Brillouinism

Finally, during his fellowship at Princeton between 1952 and 1954 Carnap intended to systematically explicating the main existing notions in the conceptual landscape of entropy. This monumental work would complete the Carnapian trilogy of critical conceptual evaluation of three of the main concepts in post-war science, i.e., probability, information and entropy, precisely the three conceptual landscapes on which informational thermophysics emerged at this historical moment (Section 2). Unfortunately, the history of this philosophical project did not have a happy ending, as we shall see below, so its publication was postponed (posthumously) until 1977. Like his 'first two parts' (Carnap 1945; Carnap and Bar-Hillel 1952), this work also embarked on his macro-program of grounding inductive logic, so that one of the results Carnap intended to obtain from his explication of entropy was the development of an abstract technical-quantitative concept of entropy without physical content that would fulfil certain functions within his theory of logic. If Bar-Hillel accompanied Carnap in the lush landscapes of informational concepts, the latter's exploits in the conceptual domains of entropy were collected for posterity by the also philosopher of physics and physicist Abner Shimony (1928-2015), undoubtedly the main critic of the informationalization of physics (especially at the hands of Jaynesianism) in the seventies and eighties. After facing several obstacles, Shimony managed to edit the never-finished *Two Essays on Entropy* seven years after Carnap's death (namely, 1977) and in an intellectual environment significantly different from the one in which its author developed this work. In his 'Introduction', Shimony faithfully reconstructed the results of this work and incorporated them into the main debates of the 1960s on this landscape.

Unlike his explication of probability in 1945, the explication does not occur on an ordinary pre-theoretical notion of entropy (because all concepts originated already in a theoretical medium) but is carried out on the multiple classical concepts of entropy that one finds in the domains of TD and SM (see Section 2.2). Beyond the various classification strategies that might arise in the physics literature (e.g., Shimony [in Carnap 1977, p.viii] presents seven different ones[73]), Carnap articulated his explication of entropy concepts around the characterization of these notions as 'logical' or properly 'physical'. Carnap (ibid., p.4) establishes $S_{TD}$ as a quantitative concept referring to an objective property of the state of the system at the same descriptive level as other 'macro-magnitudes' (in Carnapian terms) such as temperature or pressure. The fixation of $S_{TD}$ as a 'fixed conceptual point' in the entropy landscape allowed Shimony (in Carnap [1977, p.viii]) to reconstruct Carnapianly a spectrum of entropic concepts that would organize their meaning according to their degree of 'physical objectivity'. To analyse the conceptual interrelations of this domain, Carnap (ibid., p.33-42) induces a mode of partial explication controlled by his 'principle of physical magnitudes': "if a physical quantity is ascribed a definite value (which a certain precision) at a given time by a coarse description, then a finer description of the system at the same time must ascribe a compatible value to the quantity". This numerical criterion of conceptual identification (Section 2.5.1) allows Carnap to defend that the combinatorial concept $S_B$ possesses objective physical content precisely because the $\Gamma_M$-based $S_B$-descriptions of a monoatomic ideal gas in equilibrium (i.e., recall the Golden Entropy Triangle' of Section 2.5.2) coincide numerically with the $A_i$-based macroscopic $S_{TD}$-descriptions of this very scenario (Carnap 1977, p.20-22).

However, as Shimony (ibid., xi) pointed out, the Carnapian principle of physical magnitudes is a necessary but not sufficient criterion to identify the objective physical content of SB. Moreover, it would be necessary to carry out an 'objectivist' interpretative specification of the content of $S_B$, what Carnap called 'Method I', which would further imply assuming (via ergodicity and a time-averaged conception of $SP_B$ and DP, or what in 1945 he called 'probabilities$_2$') that the microstate-dynamic description $x$ of the system will conform to $\Gamma_M$-descriptions of the macromagnitudes $A_i$. This Method I satisfy both (a) the principle of physical magnitude and (b) a new necessary condition called ‘principle of disjunction’, i.e., a physical magnitude $A_i$ have the same value on the basis of the disjunction of microstatistical descriptions (ibid, p.36). On the other hand, Carnap proposes the 'non-objectivist' interpretative specification of entropic concepts such as $S_B$ by means of 'Method II', which consists in explicating $S_B$-descriptions by means of any descriptions $D_i$ whose value is independent of macromagnitudes $A_i$. A paradigmatic example (ibid., p.xi) of this interpretation in the context of the classical combinatorial argument under $S_B$ (Section 2.2.2) is not specifying the number $n$ of particles in each cell $\omega_i$ but the particular cell in which each individual particle is located. From this, Shimony remarked that "Since the entropy defined by Method II depends upon the specificity of the description [$S_B$] it is a logical or epistemological rather than a

[73] “A simple classification of conceptions of entropy is certain to be inadequate because of the many considerations involved: various levels of descriptions, large number of components in a thermodynamic system, permutability of identical components, dynamics of an isolated system, interaction of a system with different types of environment, uncertainty and information, and experimental control” (ibid, p.viii).

physical concept. Those statistical mechanicians who conceive of entropy as a measure of lack of information are committed to something like Method II" (ibid.).

At this point of his argument, Carnap critically charges against the use of Method II (closed to what I call epistemicist interpretations) to specify the proper thermophysical content of entropy concepts in SM[74]. As Shimony pointed out in the previous quote, one of the main ways to tactically exploit this interpretative method are the complementarist conceptions à la Wiener (1948) or Brillouin (1949, 1951) (Section 3.1.1. and 3.2) of the relation between entropy and information. Due to the intellectual environment in which Carnap was immersed circa 1952 at Princeton, he was fully aware of the entrenchment of the complementarist conception within the scientific community. Note that Carnap does not criticize the theoretical legitimacy of Method II, since he himself employs this path to develop a concept of entropy for inductive logic. However, from his explication of entropy one can draw a critique of the 'conceptual confusion' implicit in simultaneously using (i) an epistemicist (Method II) interpretation of $S_B$, and (ii) defending the thermophysical and empirically meaningful content of $S_B$ by this interpretative line: “The main result of our discussions is that the general statement of equality of entropy and negative amount of information can be maintained only if Method II [epistemic $S_B$] is chosen. However, in this case the resulting concept $S_B^{II}$ (in any of its versions) is not a physical but a logical concept. The customary use of the term ‘entropy’ for this concept is apt to lead to confusion.” (Carnap 1952 [1977], p.71-72). Beyond this general criticism of the interpretative incoherence underlying the information-entropy complementarity, Carnap employed his methodology of conceptual evaluation on Brillouin's well-known particular theoretical proposal in 1951:

> “In particular L. Brillouin in several articles has investigated the relation between negentropy and amount of information. (…) He shows that the increase ΔS in the total entropy of the physical system to be investigated together with the surrounding experimental arrangement can never be less than the increase Δinf in the amount of information (…) However, when Brillouin proceeds to identify negentropy with amount of information, I cannot follow him any longer. He defines entropy by $k$ ln P, where P is the number of “possible structures” or “possible states” or “complexions”; thus he means something like $S_B^{II}$ (omitting the multiplicative constant, as is frequently done). At any rate, it seems from his discussions that he implicitly uses what we have called Method II [epistemic interpretation] He does not seem to be aware that the definition of [$\Gamma_M$] which he uses (and which he ascribes to Boltzmann and Planck) makes [$S_B$] a logical rather that physical concept” (ibid, p.72-73)

This Carnapian critique of Brillouin's use of the epistemic interpretation (Method II) of $S_B$ as a thermophysically meaningful concept (which even went against the theoretical aim of Boltzmann, according to Carnap[75]) would also be key as a plausible critique of the general Brillouinian

[74] “If the extension of $S_B$ is intended to maintain the character as a physical concept in [SM] which is related to [$S_{TD}$], their Method II must be rejected and Method I must be applied” (Carnap 1977, p.44).
[75] “Therefore [a Method II interpretation of $S_B$] it is entirely out of line with Boltzmann’s aim to construct a physical, statistical concept corresponding to [$S_{TD}$]” (Carnap 1977, p.33).

program based on NPI. A reconstruction of this plausible Carnapian argument against the NPI is the following. As he stated in his explication of probabilities, descriptions derived from interpretatively specified concepts via Method II "is to be established by logical analysis alone (...) It is independent of the contingency of facts because it does not say anything about facts (although the two arguments do in general refer to facts)" (Carnap 1945, p.522). The famous analytic-synthetic distinction (without entering into any debate) would apply in a plausible Carnapian explication of the Brillouinian concept of $I_{Bb}$ as a mere 'analytic' (or by mere definition) reformulation of differences of the Method II (or epistemically) interpreted concepts $S_B(\Gamma_{M1})$ - $S_B(\Gamma_{M0})$. That is, because according to NPI the content of $\Gamma_M$-descriptions (and hence $S_B$) is specified by Method II as an argument of $I_{Bb}$ and not by their correspondence (via the Carnapian principle of physical magnitudes) with physical macromagnitudes Ai. Insofar as the content of $S_B$ and by definition also the content of $I_{Bb}$ are specified via NPI independently thermophysically meaningful properties $A_i$, both Brillouin's $S_B$ an $I_{Bb}$ are strictly epistemological notions defined over sets of microstates $\Gamma_{Mi}$ (independently of their correspondence with $A_i$). Thus, according to this reconstructed Carnapian explication, NPI would define an analytic interrelation between epistemic concepts $I_{Bb}$ and $-S_B$ (i.e., being $I_{Bb}$ just a 'terminological shortcut' for $-S_B$), where $I_{Bb}$ values will trivially coincide with the numerical values of $S_{TD}$ in those cases wherein $-S_B$ values coincide with those of $S_{TD}$ (see Section 2.5.3.).

Had he expanded his explication-based critique to this point, Carnap would have been almost three decades ahead of the critique that Kenneth Denbigh (1981) developed (of course, from a methodology that is neither Carnapian, in particular, nor properly philosophical, in general) against Brillouin's epistemic interpretation of $S_B$ and its merely definitional (i.e., analytic) connection with $I_{Bb}$:"[Brillouin's NPI] has encouraged the idea that almost any form of 'information' [$I_{Bb}$] is freely interconvertible (after appropriate change of sign) into [$S_{TD}$], or vice versa. My own view is that the interconvertibility thesis can be maintained only in certain special cases where it becomes trivially true, due to the reference to 'information' being unnecessary (…) ''bound information'' [$I_{Bb}$] is really nothing more than a name given by Brillouin to [$S_B(\Gamma_{M1})$ - $S_B(\Gamma_{M0})$]" (Denbigh 1981, p. 115). Other authors in the context of the philosophy of physics also reached the same conclusion in the late 1990s as we did with our reconstruction of the Carnapian explication (1952 [1977]) and Denbight (1981). This is the case of Earman and Norton (1999), in whose analysis of Brillouin's informational exorcism of the Maxwellian demon of Brillouin (1962, Chapter 13) argued that $I_{Bb}$ is a mere 'label' or 'terminological reformulation' and that the (physically meaningful) interrelation between $I_{Bb}$ and $S_B$ defined by NPI would depend on a precise conceptual distinction between $I_{Bb}$ and $I_{Bf}$: "the quantity [$I_{Bb}$] is an oddly labelled quantity of entropy (…) For the generalised principle [NPI] to have as precise a meaning as the original principle, one would want a precise sense in which the negentropy of [$I_{Bf}$] is interconvertible with that of [$I_{Bb}$] (…) Brillouin's examples do not provide such precision" (Earman and Norton 1999, p.8).

### 3.4.4. Shimony's Carnapian Objections against Jaynesianism

As a conclusion of the Carnapian conceptual evaluation of $S_B$ we could state that the physically meaningful ('synthetic') or epistemic-logical ('analytic') content of this concept depends on the particular method (I or II, respectively) wherein it is specified its content interpretatively, i.e., in relation (i) to microstate-based dynamic descriptions of the macromagnitudes $A_i$ or (ii) to other types of descriptions, respectively. As for the $S_G$ concept, Carnap (1977, Chapter 8) discusses at length the specification of its content during the coarse-graining procedure à la Tolman (1938), and whether in this scenario $S_G$ constitutes an explication of $S_{TD}$. His argument was the following: since the value of $S_G$ in coarse-graining $\bar{\rho}$-descriptions of equilibration processes depends on the 'arbitrary' choice of the size of cells $\omega_i$ and not directly on how physical macromagnitudes $A_i$ are $\rho$-described, then, (i) the content of coarse-grained $\bar{\rho}$-descriptions need to be specified via Method II, and (as a consequence of (i)) (ii) coarse-grained $S_G$ does not satisfy the Carnapian principle of physical magnitudes and thus does not constitute an explication of $S_{TD}$. On the basis of this Carnapian conceptual assessment of coarse-grained $S_G$, Shimony (ibid, viii) places $S_B$ closer to the 'physically objective' side of the entropic landscape (fixed by $S_{TD}$) than $S_G$, precisely for the following reason: in the case of nonequilibrium-SM descriptions, the content of $S_B$ can either be specified via Method I or II, the content of coarse-grained $S_G$ must necessarily be specified 'epistemically' (though not necessarily 'subjectively', as Robertson [2020], Section 2.2.3] notes) via Method II to match numerically with $S_{TD}$-derived descriptions. Next, Shimony (ibid., p.ix) pointed out how the landscape of non-physically meaningful entropies extended after the early 1950s beyond Brillouin's (1951) epistemic interpretations of $S_B$ and Tolman's (1938) of $S_G$:

> “Unfortunately, a full exposition of the other extreme of the spectrum [non-physically objective entropies] did not exist at the time that Carnap wrote this essay [1952-1954], even though the nonphysical conception was “in the air”. Had Jaynes’s information-theoretical formulation of statistical mechanics [1957] been available to Carnap, his essay might have gained both in sharpness of focus and in historical accuracy, by taken the formulation of Jaynes rather than of Gibbs as the target in his analysis of the relation between entropy and uncertainty” (Shimony in Carnap 1977, p.ix)

In his 'Introduction' to *Two Essays on Entropy*, Shimony (ibid., p.xvii-xviii) redirected (or used from his historical vantage point) the Carnapian critique of $S_G$ as an explication of $S_{TD}$ to attack the conceptual foundations on which the Jaynesian theoretical apparatus stands. This attack on MEP is far from an isolated case, but is framed as part of a larger critical project that made Shimony (along with fellow philosophers Kenneth Friedman and Maria Dias) the nemesis of the Jaynesian program in SM during the 1970s and 1980s (Friedman and Shimony, 1971; Friedman, 1973; Shimony, 1973, 1985; Dias and Shimony 1981). First, Shimony legitimized his use of Carnapian arguments against the epistemic interpretation of $S_G$ (particularly its coarse-grained version) on the fact that the Jaynesian conceptual architecture theoretically depends on the representational resources of GSM, as argued in Section 3.3. In this vein, Shimony argued that the epistemic interpretation of $SP_G$ and $S_G$ defended by Jaynes (1957a) was (regardless of its historiographical validity) conceptually incoherent precisely because of the numerical discrepancies between $S_G$ and

$S_{TD}$ that Carnap showed in different applicative contexts: "Carnap's exhibition of the discrepancy between $S_G$ and [$S_{TD}$] (…) constitute a serious objection to Jaynes's program" (Shimony 1977, p. xvii). That is, insofar as epistemic $SP_G$ purports to encode OI of the agent on physical macrovalues $A_i$ on which $S_{TD}$ is defined, any numerical discrepancy (i.e., assuming a numerical criterion of conceptual identity) between $S_G(\rho)$ and $S_{TD}$ at a time t would delegitimize interpreting ρ as representing the agent's macroscopic knowledge about $A_i$ at t.

In his Chapter 8 'Gibbs's Definition of Entropy for an Ensemble' Carnap analysed six cases where this numerical discrepancy occurred. For example, Case 5 explores the (unlikely but microphysically possible) case where the gas (i) TD-description: is in two quasi-equilibrium states at $t_0$ and $t_1$, or equivalently (ii) BSM-description: the gas microstate *x* transits from one macrostate-region $\Gamma_{M0}$ at $t_0$ to another macrostate-region $\Gamma_{M1}$ at $t_1$ (wherein $\Gamma_{M0} > \Gamma_{M1}$). In this scenario, while the $S_{TD}$ and $S_B$ values would decrease, the coarse-grained $S_G$ value would increase due to (iii) GSM-description: the ρ-encoded ensemble is dynamically dispersed by new cells $\omega_i$. In Carnap's words "This shows again that $S^{II}_G$ [epistemic $S_G$] has nothing to do with $S_{TD}$. In this case most of the new cells into which the ensemble has spread by [$t_1$] represent states with lower values of $S_B$ and of $S_{TD}$. $S^{II}_G$ in no way reflect this fact; it merely reflects the fact of spreading" (ibid, p.56). Therefore, if according to Carnap epistemic $S_G$ does not correspond (numerically and semantically) to $S_{TD}$, then Shimony (p.xviii) argues against Jaynes that $H_{IT}$ (i.e., assuming in his favor the identity between $S_G$ and $H_{IT}$) will not conceptually correspond to $S_{TD}$ either.

In addition to this Carnapian attack on the interpretative foundations of the Jaynesian, Shimony immediately recovers the critical line developed in Friedman and Shimony (1971) on certain technical problems of MEP. In particular, the latter two authors pointed out that the MEP-averaging mechanism does not properly constitute a procedure for updating prior densities with respect to new data (as is Bayes' theorem as a procedure for updating prior densities), but a mechanism for reintroducing prior probabilities. On the basis of this idea, Friedman and Shimony (1971) argued (as Cyranski 1979, p.298 later did) that MEP is a procedure inconsistent with the theoretical framework underlying Bayesian probabilities, the debate on which would continue for several decades (see also Skyrms 1981). Interestingly, a decade later Shimony together with Dias (Shimony and Diaz 1981) employed the Carnapian concept of 'lambda comtinuum' (Carnap 1952) to demonstrate how a statistical statement could be employed interchangeably as a constraint in MEP and as evidence in Bayesian conditionalization.

Finally, the third argument of Shimony (in Carnap 1977, p.xviii) against Jaynesianism is directed at the conceptual inconsistency implied by the identification of $S_{TD}$ and $H_{IT}$, where the latter is also epistemically interpreted as a measure of uncertainty. To support his thesis, this philosopher relied on the results obtained by Lazlo Tisza and Paul Quay (1963) in the field of equilibrium-SM, where they showed that, while the value of the $H_{IT}$ of a substance composed of a single type of molecule depends numerically on the number of isomers (well-defined variant of the same type) of that molecule, the value of the $S_{TD}$ of this substance is completely independent of the number of isomers that this substance may have. Thus, this illustrative example serves to support "a clear conceptual distinction between [$S_{TD}$] and [$H_{IT}$]" (Shimony in Carnap 1977, p.xviii), thus collapsing the thermophysical significance of the predictive results derivable from MEP. As early as the 1980s, Jaynes was fully aware that the Neo-Carnapian arguments of Shimony (et al.) posed a major threat to the conceptual foundations of Jaynesian treason, even if this in no way implied (as history has shown) a questioning of his theoretical program by his community of adherents. Even while aware of the minimal impact of Shimony's critique within (and even outside) the Jaynesian community, Jaynes (apparently unlike Shimony) seemed to assume ad hominen attacks as legitimate argumentative resources in intellectual debates:

> "[Shimony] seems to have made it his lifelong career to misconstrue everything I wrote many years ago, and then compose long pedantic commentaries, full of technical errors and misstatements of documentable facts, showing no awareness of anything done in this field since then -and which, to cep it all off, attack not my statements, but only his own misunderstandings of them. The conflict is not between Shimony and me, but between Shimony and the English language"

This (maybe right, but certainty unfortunate) comment by Jaynes (1985, p.135) can be found in an article published in Synthese, a leading journal in the philosophy of science domain. Interestingly, Jaynes dedicated some words in this very same text to the community of philosophers of science:

> "Of course, if philosophers wish to discuss the rationale of science among themselves, in their own journals, without pretending that they are making new contributions to science, they have every right to do so. We physicists also gossip among ourselves about work in other fields" (ibid, p.134)

### 3.4.5. Conclusion: Could Philosophical Criticism Have Stopped the Bandwagon?

This development of the systematic Carnapian evaluation of the conceptual landscapes of probability, information and entropy during the period 1945-1954 constitutes (as I have tried to posit in this Section 3.4.) the germ of an analytically robust critique of the intellectual embryo of what are retrospectively the conceptual foundations of informational thermophysics in the Brillouinian and Jaynesian tradition. This is particularly evident in Carnap's *Two Essays on Entropy*, from which (i) Shimony draw key critical elements for an attack on the Jaynesian program, and (ii) Carnap himself robustly undermined attempts to defend the thermophysical content of entropic concepts by epistemicist or Method II interpretations, upon which rested (in particular) the Brillouinian program and (in general) complementarist misconceptions of the entropy-information connection:

> "It seemed to me that the customary way in which the statistical concept of entropy is defined or interpreted makes it, perhaps against the intention of the physicists, a purely logical instead of a physical concept; if so, it can no longer be, as it was intended to be, a counterpart to the classical macro-concept of entropy introduced by Clausius, which is obviously a physical and not a logical concept" (Carnap 1963, p.36f).

This *Two Essays on Entropy* as a cornerstone of Carnapian criticism was put together during Carnap's fellowship at the Institute for Advance Study at Princeton during 1952-1954. However, Carnap found at Princeton an intellectual environment in which the informationalist ideas consolidated in the early 1950s (Section 3.1) were deeply rooted. This prima facie would not have been an impediment for the German philosopher's project to come to light at that time. However, the enormous influence that some of the key scientific figures of the twentieth century exerted on Carnap was decisive in dissuading him from the idea of publishing the results of his studies on the uses of entropy concepts (see Köhler 2001, p.100-101). The series of episodes in which this dissuasion took place consisted of several conversations between, on the one hand, the German physicist Wolfgang Pauli and von Neumann (remember, the main promoter of the informationalization of thermal physics), and on the other hand, Carnap in the company of his collaborator Bar-Hillel. This conversation was remembered by Carnap himself (compiled in the Autobiography in the Library of Living Philosophers [Carnap 1963]) as a failed attempt to show argumentatively the existence of certain interpretative incoherencies and conceptual inconsistencies in certain uses of the concepts of entropy, information and probability:

> "I had some talks separately with John von Neumann, Wolfgang Pauli, and some specialists in statistical mechanics on some questions of theoretical physics with which I was concerned. I certainly learned very much from these conversations; but for my problems in the logical and methodological analysis of physics, I gained less help than I had hoped for. (…) I had expected that, in the conversations with the physicists on these problems, we would reach, if not an agreement, then at least a clear mutual understanding. In this, however, we did not succeed, in spite of our serious efforts, chiefly, it seems, because of great differences in point of view and language" (Carnap 1963, p.36f)

Such a disagreement that Carnap reported between these physicists and his own philosophical assessments of Carnap is especially evident when we move to the views of the other participants. Pauli, who perused the preliminary manuscript of *Two Essays on Entropy* and prepared a detailed response, not only defended the epistemicist (Method II) interpretation of thermophysical entropy concepts[76] but also argued that this manuscript should not be published because its philosophical assessments would contribute to conceptual confusion within the physics community:

> "Dear Mr. Carnap! I have studied your manuscript a bit; however I must unfortunately report that I am quite opposed to the position you take. Rather, I would throughout take as physically most transparent what you call "Method II". In this connection I am not all influenced by recent information theory (…) Since I am indeed concerned that the confusion in the area of foundations of statistical mechanics not grow further (and I fear very much that a publication of your work in this present form would have this effect)" (Pauli 1954 [1999, p.541])

[76] In fact, Pauli confessed that "In this way one comes to what you have named "Method II", which appears to me to be the only satisfactory one: the amount of entropy describes logically our subjective knowledge of the system, the thermodynamic entropy corresponds in particular to that information about systems (this appeared to the last century as "objective reality") obtainable by infinitely slow diffusion processes" (Pauli 1999, p.544).

As for the conversation with von Neumann, where they tried to show argumentatively the conceptual inconsistency underlying his identification of $S_G$ and $H_{IT}$ solely on the basis of the formal similarity of their mathematical expressions (see Section 3.1.2). Bar-Hillel, who was present at the conversation, reported von Neumann's refusal to change his rigid intellectual stance:

> "During one of my visits to him in Princeton, in 1952, von Neumann also came to see him [Carnap], and we started discussing the talk I had heard von Neumann deliver (…) in which he had proclaimed, among other things, a triple identity between logic, information theory and thermodynamics (…) We were quite ready to agree that there existed a certain formal analogy (…) and had indeed ourselves shown (…) how entropy-like expressions, such as the famous $-\Sigma\, p_i \log p_i$, occurred in (inductive) logic, but could not see how stronger relationships could possibly be supposed to exist. We tried to convince von Neumann that this way of presenting the analogy as an identity must lead to confusion. His calm reply was that he could see how logicians and methodologists might be worried by his statements but that no physicist would misunderstand them." (Bar-Hillel 1964, p.11, quoted in Köhler 2001, p.100-101)

Our whole paper could be understood as an attempt to show historically and philosophically how von Neumann could not have been more wrong with his response to the critique of Carnap and Bar-Hillel (last strawberry in the quote above). Carnap assumed at the time (i.e., circa 1954) that, if scientific geniuses like Pauli or von Neumann were so reluctant to accept his conceptual assessments, then the rest of the scientific community would hardly adopt a more open attitude to these critical lines. The intellectual tsunami of informationalism was already an unstoppable current of scientific thought. Faced with this bleak scenario, Carnap had no choice but to abandon his critical project indefinitely in the sands of time, being unearthed posthumously by Shimony (1977) in a world in which the conceptual landscapes and the interests of the thermophysical community were no longer the same as in the 1950s. Regardless of the validity or not of these Carnapian analyses (to which I will return later), in that talk von Neumann could have done a significant disservice to science as a critical domain of inquiry.

> "I know that Carnap decided, after this talk, to postpone sine diem the publication of a long paper he had almost finished on the role of inductive logic in statistical mechanics. In this paper he had made some forceful and, to my mind, very illuminating criticisms of certain standard formulations in this theory. I myself omitted some of my remarks on physics in the printed version of my Princeton talk" (Bar-Hillel 1964, p.11)

# Chapter IV

## The Golden Age of Information Physics

> "in the last several decades there has been significant growth in a field called "The Physics of Information", whose motivating conception is precisely that information processing must be instantiated through physical systems, and that the nature of these system matters. The key slogan is that 'Information is Physical'" (Maroney and Timpson 2018, p.103)

As I have argued in Chapter III, information physics emerged in the field of thermal physics during the 1950s as an intellectual current characterized mainly by the use of informational concepts to reinterpret physical phenomena and rewrite problems such as measurement or Maxwell's demon. This intellectual current (led by the Brillouinian and Jaynesian traditions) evolved until the 1980s, when the emergence of computational thermophysics and other factors boosted its popularity within the physics community. It was precisely in the 1990s that physical informationalism went from being a mere 'alternative' current of thought to become (i) a hegemonic intellectual trend not only in thermophysics, but also in other fundamental domains such as quantum and relativistic physics; and (i) an institutionalized subdiscipline of physics since 1990. This caused that at the end of the twentieth century informational concepts began to flood any kind of scientific production in physics "Physicists use the concept of information in various senses: some notions of information [e.g., $H_{IT}$] are mathematical and precise, while others [e.g., OI] are qualitative and heuristic. As the history of physics and the technology of science has progressed, information as a concept and tool has become more central to the discipline" (Harshman 2016, p.7). As I will try to evaluate in this chapter, this intellectual hypertrophy of physical informationalism in the nineties had as its main consequence the first explicit proposals of informational metaphysics at the hands of the leaders of this movement, such as the famous 'It from the Bit' by Wheeler (1990) or 'Information is physical' by Landauer (1991). This ontology-centred turn of physical informationalism[77] was received with interest by part of the philosophical community since the

[77] "perspectives on information within physics can tend toward more epistemological or more ontological. On the epistemological side, an implicit assumption is that information is a property of the interaction of the system and the observer (…) In contrast, the ontological perspective is that information is a real thing" (Harshman 2016, p.7).

mid-2000s, both by those authors with the pretension of developing an informational metaphysics (Ladymanand Ross 2007) and by those authors who sought to criticize these inflationary ontologies (Timpson, 2013; Lombardi et al. 2016a). It is in this precise context that my assessment of the conceptual and interpretative foundations underlying the theoretical and ontological inflation of information physics in the 1990s is inserted.

## 4. 1. Intellectual Evolution of Informational Thermophysics

As I have analysed extensively in Chapter III, thermal physics was heavily informationalized during the fifties through Shannon's IT bandwagon, resulting in the Brillouinian and Jaynesian traditions. I aim to show in this Chapter IV that informationalism as a current of physical thought reached its intellectual peak, in terms of popularity and radicality in its theoretical positions, during the nineties. But, how was it possible to reach this latter historical stage from the former one? In the following, I will attempt to answer this question. I will argue that this convoluted intellectual process (initiated in the early-1960s and completed in the late-1980s) can be satisfactorily understood through the combination of three historical factors, namely: (a) the multiple confluences of the theoretical goals, conceptual resources and interpretative strategies of Brillounianism and Jaynesianism, (b) the insertion of the concept of algorithmic information K into the physical realm and the computational turn of informationalism in the 1980s, and finally (c) Bekenstein's derivation of black hole entropy from informationalist presuppositions.

### 4.1.1. Informational Syncretism: Tribus and Brillouinian-Jaynesian Synthesis

One of the main factors that caused the radicalization of the thermophysical informationalism during the eighties-nineties was the intellectual confluence of informationalist schools with a certain intellectual standing within the scientific community. As I have already analysed in Chapter III, from the late 1950s onwards two systematic thermo-informationalist traditions were consolidated (within a vast amalgam of theoretical proposals) on different families of interpretative strategies and different principles of conceptual interconnection, namely: Brillouinism and Jaynesianism. Beyond the particular historical drift of each of these intellectual lines, both converged (more superficially or more systematically) in the 1960-1970s within certain theoretical proposals of great impact within the literature of the following decades.

This is the case of engineer Myron Tribus (1921-2016), undoubtedly one of the main figures of informationalism during the transition period between the 1950s (Chapter III) and the golden era of the 1990s (Chapter IV). As a professor of thermodynamics at UCLA during the fifties, Tribus was first an enthusiast of Brillouin's theoretical proposals (1951, 1956) and later an adept of the way in which Jaynes (1957a, 1957b) conceptually connected IT and thermal physics. Recall that it was precisely Tribus who in an interview with Shannon in 1961 asked him about the origin of his terminology 'entropy' (Section 3.1.2). As an intellectual culmination of the informational frenzy of the fifties, this author developed in his 1961 book *Thermodynamics and Thermostatics. An Introduction to Energy, Information and States of Matter, with Engineering Applications* a systematic reformulation of TD and SM along clearly Brillouinian and Jaynesian conceptual-interpretative lines. His intellectual profile was strongly interdisciplinary, seeking to connect different research domains with thermophysics (interestingly, Tribus is the founding father of what is currently known as 'thermoeconomics'). After maturing his theoretical position during the sixties, Tribus presented in the famous 1971 article 'Energy and Information' (co-authored with E. McIrvine) his most conceptually sophisticated Brillouinian-Jaynesian informationalist proposal, which had the greatest impact within the scientific community. Tribus' theoretical proposal could not be understood historically as a mere reformulation of the predominant thermo-informational traditions in the seventies (i.e., most of the proposals in the literature), but properly as a reconstruction of the conceptual architecture and a radicalization of the theoretical assertions.

One of the main points of conceptual demarcation of Tribus' (and henceforth also McIrvine's) proposal from the informationalism inherited from the 1950s is not the plausible conceptual connection (complementarist or identificatory) between entropy and information but properly between 'energy' and information. As a central thesis, Tribus articulated his informational proposal on a meaningful conceptual relation between the TD-significant notion of 'energy' and a technical notion of information (conceptually interchanging $H_{IT}$ and a OI-like epistemic information, as we shall see):"Today we know that it takes energy to obtain knowledge and that it takes information [OI] to harness energy (…) whereas the theory of instrumentation shows why energy is needed to obtain information, recently advances in information theory show why information is needed for transformation of energy" (Tribus and McIrvine 1971, p.179). Tribus' use in this quote of a concept of information semantically interchangeable with that of knowledge is conceptually based on an explicit epistemic interpretation of probabilities SP à la Jaynesian, such that these probabilistic concepts allow the epistemic states of agents to be mathematically modelled: "Ideas about probability play a central role in any theory of knowledge. In modern information theory probabilities are treated as a numerical encoding of a state of knowledge" (Ibid, p.179). As a fervent Jaynesian (Section 3.3), Tribus strategically capitalized on this line of epistemic interpretation of probabilities to epistemically interpret the technical-statistical information concept of $H_{IT}$: "By observing that knowledge can thus be encoded in a probability distribution (...), we can define information as anything that causes an adjustment in a probability assignment" (Ibid). Of course, this interpretative stance straightforwardly led to epistemically conceiving $H_{IT}$ as an uncertainty-like notion "Numerous workers have demonstrated that Shannon's measure [$H_{IT}$] of uncertainty which he called entropy, measures how much is expected to be learned about a question when all that is known is a set of probabilities" (Ibid).

Going a step beyond the interpretative strategies of the Jaynesianism tradition, Tribus argued for an identification not between the concepts of $S_G$ and $H_{IT}$ (raised by von Neumann [1949] and technically defined by Jaynes [1957a]) but properly between $S_{TD}$ and $H_{IT}$ "It can be shown that

Shannon's function [HIT] and Clausius' function [$S_{TD}$] are the same (...) The proof that they are indeed the same (and not merely analogues) has been dealt with extensively elsewhere" (Ibid, p. 179, 181). As can be extracted from this quote, the authors assumed theoretically both the conceptual identity between $S_{TD}$ and $S_G/S_B$ and that the validity of the conceptual identity (interpretively mediated by epistemic probabilities) between $S_{TD}$ and $H_{IT}$ had already been demonstrated by the previous informationalist tradition. Regardless of whether this conceptual identity had been 'proven', Tribus contributed to popularizing in the scientific community of the 1970s the belief that $S_{TD}$ and $H_{IT}$ were the same concept. Because this identification was ultimately based on an epistemic conception of probabilities, this conceptual entropy-information identification ends up implicitly extending interpretively also to OI (as in Jaynes' proposal, see Section 3.3) "The concept of an inherent connection between the entropy Clausius and the intuitive notion of information preceded Shannon's work by many years" (Ibid, p.182).

Of course, to theoretically justify the strong conceptual identification $S_{TD}$-OI Tribus (and McIrvine 1971, p.181) focused on the historical origins of Maxwell's demon problem, defending his approach as a sort of prehistory of informationalism in thermophysics "From Maxwell's time on many leading investigators empowered the relation between observation and information on the one hand and the Second Law of thermodynamics on the other" (Ibid, p.182). Within the line of thermophysical research articulated on Maxwell's demon, Tribus argued that the Brillouinian proposal constitutes the definitive theoretical resolution of this fundamental problem: “The demon was finally “exorcised” in 1951 by Léon Brillouin, who pointed out that if the demon were to identify the molecules, he would have to illuminate them in some way, causing an increase of entropy that would be more than compensate for any decrease in entropy such a being could effect” (Ibid, p.181). In this line, Tribus adopts a properly Brillouinian interpretative strategy (Section 3.2) to defend that the $S_{TD}$-$H_{IT}$ connection is only thermophysically valid if $H_{IT}$ is technically defined by observationally significant quantities $A_i$ (which includes different types of energy, mainly heat Q) via Brillouin's bound information $I_{Bb}$ "There is no conflict between abstract Shannon information and thermodynamic information, as long as the questions we ask are physically real questions" (Ibid, p.183). It is precisely the interpretative use of bound information that allows Tribus to develop the conceptual infrastructure to defend argumentatively his central thesis in this 1971 paper, namely, the existence of a robust connection between energy and information: “Whatever the short-term outcome in educational circles, it is certain that the conceptual connection between information and the Second Law of thermodynamics is now firmly established. The use of “bound information” in the Brillouin sense of necessity involves energy. The use of energy, based on considerations of thermodynamic availability, of necessity involves information. Thus information and energy are intrinsically interwoven” (Ibid, p.188).

This radical syncretism of the conceptual architecture of Brillouinism and Jaynesianism in the theoretical proposal of Tribus (and McIrvine) (1971) made him the major intellectual advocate of informationalism in thermophysics during the seventies, in general, and of both lines of thought, in particular. Interestingly, Jauch and Baron (1972, p.220) highlighted the informationalist drift at this time "There is a widespread belief that thermodynamics is more or less closely related to the concept of information as used in communication theory". These authors concentrated their critical attack on the Brillouinian strategies of theoretically justifying the complementarist connection between information and entropy (via NPI) on the basis of the formal similarity between the mathematical expressions of both concepts “The similarity of the formal expressions in the two cases has misled many authors to identify entropy of information (as measured by the formula of Shannon) with negative physical entropy.” (Ibid). Jauch and Baron's argument focused on criticizing the conceptual inconsistencies of Szilard's model on which much of Brillouinism depends "Szilard's experiment is based on an inadmissible idealization [i. e. the insertion of the piston violates the Gay-Lussac law]; therefore it cannot be used for examining the principles of thermodynamics (...). e. the insertion of a piston violates the Gay-Lussac law]; therefore it cannot be used for examining the principles of thermodynamics (...) Thus (...) it is no paradox; and has nothing to do with information" (Jauch and Baron 1972, p.232).

The Jaynesian physicist Olivier Costa de Beauregard and Tribus himself responded to Jauch and Baron's criticisms in their 1974 article 'Information Theory and Thermodynamics'. This critical strategy was a symptom of the informational syncretism that began to proliferate at the time. While Jauch and Baron criticized the thermophysical insignificance of information from Brillouinism, Costa de Beauregard and Tribus (1974, p.246) exploited the Jaynesian strategy of epistemically interpreting probabilities to defend that "in order to have the entropy balance [in a Szilard’s cycle] right, we must include in it the information [OI] gained by the observer." Thus, according to these authors, such a thesis would somehow justify not just a conceptual identification between entropy and information, but an interpretation of IT as a straightforward theoretical generalization of TD (but not the other way around): “we believe that, while it is possible, and very informative, to deduce Thermodynamics from a general theory of Information, the converse is not possible. Thermodynamics is too rooted in specifics of physics to produce a general theory of information” (Costa de Beauregard and Tribus 1974, p.246). Therefore, the progressive hybridization of Brillouinian-Jaynesian interpretative strategies by authors such as Tribus from the seventies onwards allowed the strengthening of argumentative tactics through which to defend the disciplinary pretensions of informationalism against possible objections of theoretical misuse, conceptual inconsistency and interpretative incoherence. As a consequence of this progressive 'intellectual immunization', informationalism in thermophysics began to proliferate uncontrollably also outside the frameworks delimited by the Brillouinian and Jaynesian tradition, generating new informational tendencies in physics with a disciplinary scope, theoretical ambition and interpretative radicality never seen before in the history of this intellectual current.

### 4.1.2. Algorithmic Information and the Computational Turn of Informationalism

Apart from the progressive syncretism between different traditions, another factor that promoted the rise of informationalism within the physics community of the nineties was the incorporation of new conceptual and technical resources by the algorithmic information theory AIT developed by Kolmogorov (1965) and Chaitin (1966) during the mid-sixties (Section 2.4). Unlike the immediate thermophysical bandwagon originating in the 1950s from Shannon's IT (Section 3.1), the assimilation of Kolmogorov and Chaitin's AIT theoretical proposal (or 'AIT bandwagon') into the physical realm was a subtle and historically extended intellectual process. Just as Shannon conceived IT as a statistical tool without physical content, Kolmogorov-Chaitin's AIT was originally conceived as a formal tool (i.e., without physical content) aimed at the domain of probability theory and computation, respectively. Just as I analysed in the case of Shannon's IT, certain popular narratives tried to emphasize that AIT physical applicability was intrinsic to that theory: “For Kolmogorov, these ideas belonged not only to probability theory but also to physics. To measure the complexity of an orderly crystal or a helter-skelter box of a gas, one could measure the shortest algorithm needed to describe the state of the crystal or gas.” (Gleick 2010, p.338). Interestingly, the AIT apparatus was historically introduced into the field of thermophysics through two distinct intellectual pathways. On the one hand, the concept of complexity or algorithmic information K was first systematically applied in the field of dynamical systems in 1978 through what is known in the literature as 'Brudno's theorem' (Brudno, 1978; Frigg 2004, p.430). This theorem conceptually identifies the notion of information AIT with the chaos-theoretic notion of Kolmogorov-Sinai entropy (see Frigg 2004), allowing to interpret the content of this concept K of information as a measure of the dynamical randomness of microstate trajectories.

On the other hand, and as far as my task here is directly concerned, AIT theory and the K-concept began to gain predominance within thermophysical informationalism (particularly within computational thermophysics) in the eighties thanks to the close intellectual collaboration between Gregory Chaitin and Bennett since the second half of the seventies. Ten years after his founding paper in 1965, Chaitin started working in 1975 at the IBM Watson Lab together with key figures such as Bennett himself or Landauer, who later formed what Feynman (1996, p.148) called the 'IBM group'. For Bennett, who in the late 1970s was developing the theoretical foundations of computational thermophysics, Shannon's IT posed certain interpretative problems in the informational characterization of thermophysical processes: "Ordinary information theory offers no solid grounds for calling one N-bit string "more random" than another, since all are equally likely to be produced by a random process such as coin tossing" (Bennett 1982, p.936). Based on this interpretative difficulty, the author points out the conceptual problem that both $H_{IT}$ and $S_G$ are "inherently statistical concept" (Ibid) because they can only be defined for probabilistic distributions of events and not for individual events. Bennett found in the conceptual tools developed in Chaitin's AIT (1975a) a fantastic theoretical opportunity to solve this problem, wherein this intellectual influence is explicitly acknowledged in his foundational 1982 article[78]. In this direction, Bennett defended that “the notion of algorithmic entropy resolves the conceptual problem [regarding individual microstates as having entropy] by providing a microstate function,

[78] “I wish to thank Rolf Landauer and Gregory Chaitin for years of stimulating discussions of reversibility and entropy” (Bennett 1982, p.939)

H(x), whose average is very nearly equal to the macroscopic entropy S(p), not for all distributions p (which would be impossible) but rather for a large class of distributions including most of those relevant to [SM]” (Bennett 1982, p.936). This conceptual connection between $H_{IT}$-$S_G$ and K was concisely described in Section 2.5.

To give physical content to the informational notion K through this conceptual connection with $H_{IT}$-$S_G$, Bennett (who used the term "algorithmic entropy") interpretively relied on Chaitin's (1975b) theoretical proposal to define K quantities over data structures other than binary sequences à la TI, such as coarse-grained cells in GSM. Thus, the algorithmic entropy K of an individual microstate *x* of the system is not properly defined as an observable physical quantity $A_i$ but as the number of bits $B_i$ (i.e., binary partitions of Γ) needed to describe *x* by a given universal program. Because this (still underdeveloped) notion of 'algorithmic entropy' is directly applicable to a microscopic configuration of the system: “Algorithmic entropy is a microscopic analog of ordinary statistical entropy [$S_G$] in the following sense: if a macrostate p is concisely describable, e.g., if it is determined by equations of motion and boundary conditions describable in a small number of bits, then its statistical entropy is nearly equal to the ensemble average of the microstate’s algorithmic entropy” (Bennett 1982, p.938). Note that prima facie the specification of a microstate *x* within a continuous density ρ constitutes either a computationally intractable or extremely difficult process (on the order of $10^{24}$ bits). To justify theoretically the conceptual connection underlying the algorithmic entropy of microstates in the face of these feasibility problems, Bennett appealed directly to Chaitin's (1975b, Theorem 3.2) use of Monte Carlo techniques to deal with the algorithmic complexity of continuous objects such as ensembles densities ρ. Undoubtedly, this Bennettian proposal in 1982 was precisely the historical origin of what it could be called the 'algorithmization' of thermophysical informationalism, as Caves (1990, p.93) pointed out eight years later “This article [Bennett 1987] led me to Bennett’s review of reversible computation, the last section of which makes the seminal connection between entropy and algorithmic information”.

Undoubtedly, this progressive algorithmization (or similarly ‘complexification’) in the eighties of previous thermo-informationalist currents was an entangled intellectual process encouraged by the historically simultaneous rise of Bennett-Landauer’s computational thermophysics, encouraging a significant computational turn in the informational thinking trends of the time. In this direction, Bennett (1988, Section 5) also raised the conceptual plausibility of physically interpreting the content of Chaitin's (1987) AIT-based notion of 'logical depth', defined roughly as the amount of computational resources required to reproduce a given object by means of a computer program. The Bennettian strategy of epistemically interpreting Chaitin’s concept of logical depth as a measure of the physical complexity of a given system or process depends conceptually on the validity of the so-called 'Turing-Church physics thesis' (put forward by David Deutsch [1985], see Piccinini and Maley 2021), which states that certain physical processes can be computed.

In this computational turn of informationalism, the concept of algorithmic information K was assumed by the scientific community of the eighties as advantageous over Shannon's $H_{IT}$ notion for its application within physical contexts, even by the most 'info-skeptical' scientists. This was the illustrative case of Jeffrey Wicken, who in his 1987 paper 'Entropy and Information: Suggestions for Common Language' systematically criticized that Shannon's IT conceptual resources could be satisfactorily employed to model thermophysical properties of systems, mainly through the (complementary or identifying) connection between $H_{IT}$ and $S_{TD}$ (as was discussed in Chapter III). In spite of this critical position against traditional informationalism (i.e., Brillouinian or Jaynesian), Wicken defends the value of algorithmic entropy as a measure of the ontological complexity in the structural relations existing between the elements of a system, whose physically interpretable content resides precisely in the possibility of specifying properties of individual objects (microstates or messages): "Shannon could have avoided these difficulties in part had he confined entropy to a property of the ensemble rather that extending it to the message itself (…) There is, in fact, *a completely appropriate alternative* to "entropy" in information theory (…) This has been discussed by Chaitin (1965, 1975) in his algorithmic approach to information theory. Here, the "entropy" of the symbol ensemble is replaced by the program of information required to specify the sequence" (Wicken 1987, p.184. Italics are mine)

### 4.1.3. Bekenstein and Black Hole Thermodynamics

Together with the syncretism of informational traditions in thermophysics and the incorporation of algorithmic information concepts in these intellectual trends, another of the historically decisive factors for the rise of informationalism in physics in the nineties was Bekenstein's derivation of black hole entropy in 1972 from the conceptual resources and interpretative strategies of Brillouinism and Jaynesianism. Let us briefly approach the intellectual origins of this theoretical proposal. The Mexican-Israeli physicist Jacob Bekenstein (1947-2015) was encouraged by John Wheeler (then his doctoral supervisor at Princeton and future protagonist of the informationalism of the nineties) to study certain singularities in certain solutions of Einstein field equations, the latter baptizing them as 'black holes' in a 1967 talk on pulsars (Nolte 2018, p.168). Such singularities were first discovered by Schwarzschild in 1916 (while participating in World War I) the year after the publication of Einstein's equations of general relativity. It was precisely during the 1960s that the idea became popular that, according to Oppenheimer-Snyder's calculations in the 1930s (ibid, p.168) these mathematical singularities (à la Schwarzschild) should have physical content, implying the existence of certain gravitational singularities in the structure of spacetime. According to such solutions to the Maxwell and Einstein field equations (i.e., encoding electromagnetic and gravitational properties, respectively), these 'black holes' could be fully characterized by three classical observable parameters, namely mass, electric charge and angular momentum. The rest of the physical properties of black holes ceased to be predictable in a certain region where their gravity reached a critical value, called their 'event horizon'. This theoretical principle was formulated by Bekenstein (although popularized by Wheeler[79]) with the expression "black holes have no hair", from which arises what is now known as the 'no-hair theorem'.

These physical anomalies that constitute black holes gave rise to the idea those spacetime objects could also be theoretically problematic in the field of thermophysics. Bekenstein encountered in the late 1960s during his doctoral stage a thermophysical context strongly marked by the Brillouinian and Jaynesian informationalist traditions. As the historian Israel Belfer argued on Bekenstein's intellectual evolution: “[Bekenstein] immersed himself in recent developments and the classic literature of thermodynamics and statistical mechanics, including the attempts by Edwin T. Jaynes and Léon Brillouin to incorporate tools of information theory in physics” (Belfer 2014, p.71). Like Tribus (and McIrvine 1971), Bekenstein's theoretical proposal constitutes another paradigmatic example of this intellectual tendency in the 1970s to employ the Brillouinian and Jaynesian traditions to develop a syncretic informational theory. Because of this informationalist influence during his scientific training, Bekenstein employed these conceptual resources and interpretative strategies (Chapter III) to clarify certain theoretical problems in his field of black hole research (initially a classical physics domain, but from the mid-1970s on this became an intrinsic quantum phenomenon). An illustrative example of these cognitive tactics can be found in what this author called 'Wheeler's demon' (Bekenstein 1972), as a mental experiment analogous in the thermo-gravitational domain to 'Maxwell's demon' on which traditional informationalism was centered. This was based on a hypothetical situation posed by Wheeler: suppose we throw an object with high entropy $S_{TD}$ (e.g., the mixing of hot tea and cold milk) beyond the event horizon of a black hole until it eventually disappears, then the overall $S_{TD}$ entropy of the universe will decrease accordingly, thus violating the Second Law of TD:

> “Suppose that a body containing some common entropy goes down a black hole. The entropy of the visible universe decreases in the process. It would seem that the Second Law of thermodynamics is transcended here in the sense that an exterior observer can never verify by direct measurement that the total entropy of the whole universe does not decrease in the process. However, we know that the black-hole area “compensates” for the disappearance of the body by increasing irreversibly. It is thus natural to conjecture that the second law is not really transcended provided that it is expressed in a generalized form: The common entropy in the black-hole exterior plus the black-hole entropy never decreases. This statement means that we must regard black-hole entropy as a genuine contribution to the entropy content of the universe” (Bekenstein 1973, p.2336)

This thought experiment led Bekenstein to propose an interpretative strategy whereby, just as Brillouin generalized the concept of entropy via NPI to informational processes to exorcise Maxwell's demon, the concept of entropy $S_{TD}$ could also be generalized thermo-gravitationally to black holes to exorcise Wheeler's demon. Just as for Brillouin's (1956) NPI postulates that the sum of negentropy N and information $I_{Bb}$ cannot decrease during evolution (Section 3.2), for Bekenstein the sum of $S_{TD}$ and black hole (also known as ‘Bekenstein-Hawking’) entropy $S_{BH}$

[79] John Archibald Wheeler, “Feynman and Jacob Bekenstein,” accessed June 21, 2021, http://www.webofstories.com/play/9549. https://www.youtube.com/watch?v=BIHPWKXvGkE&t=6m

could not theoretically decrease (Bekenstein 1973, p.2336). Thus, by throwing a highly entropic object into the event horizon of a black hole, this Wheelerian demon action would violate this thermo-gravitational generalization of the Second Law of TD, precisely because it presupposes that the entropy of the black hole $S_{BH}$ should proportionally increase in compensation.

Bekenstein originally postulated that the $S_{BH}$ black hole entropy should be defined on the basis of an observable property of the black hole that never decreases in its evolution, by analogical reasoning with respect to $S_{TD}$, see Section 2.2. From certain results obtained by the British physicist Stephen Hawking (1942-2018) in 1971, this author assumed that due to the non-conservation of the baryon number this 'fundamental' property should be the area A of the event horizon of the black hole (Bekenstein 1972), whose increase was proportional to the increase of its mass M (i.e., one of its observables according to the no-hair theorem). Assuming that this area A would play (via Einsteinian energy-mass equivalence) an analogous role as a quantity of energy to heat Q in the $S_{TD}$ formula, Bekenstein (1973) conjectured that the constant of proportionality between $S_{BH}$ and A would have an approximate value of 0.276 encoded by the ratio between the Boltzmann constant $k_B$ (giving thermophysical significance) and the formula for the area of a sphere 8π. Months later, Hawking, who independently was also trying to describe the thermophysical properties of these gravitational anomalies, determined that the proportionality constant was precisely ¼. The next step in the construction of the $S_{BH}$ concept consisted in partitioning the area A of an event horizon into minimal regions in a physically robust way for its quantification. As collected by Belfer (2014), Wheeler proposed to Bekenstein to employ as a universal physical measure the fundamental Planck length $l_P$ to quantize geometrically the total area of the surface A by minimal Planck areas or triangles of side $l_P^2$. Because of the absolute values of Planck areas with edges $l_P^2$, they would play an analogous conceptual role in $S_{BH}$ to the absolute Kelvin temperature T in $S_{TD}$. Once with these representational resources, Bekenstein defined (analogously to $S_{TD} = \int dQ/T$) the black-hole entropy concept $S_{BH}$ by the following expression:

$$S_{BH} = \frac{k_B A}{4\, l_P^2} \tag{1}$$

Following the Brillouinian tradition, Bekenstein clearly adopted in his 1973 paper '*Black holes and entropy'* an epistemic interpretative stance with respect to entropic concepts and a complementarist perspective of the relationship between entropy and information "The entropy of a system measures one's uncertainty or lack of information about the actual internal configuration of the system" (Bekenstein 1973, p.2335). To employ this interpretative strategy as a theoretical legitimation of his own thermo-gravitational concept of entropy $S_{BH}$, Bekenstein (ibid.) reconceptualized Planck-triangles quantifying area A as binary units of $H_{IT}$ information (also epistemically conceived as a measure of agent uncertainty) or 'Bekensteinian bits'. Then, the number of Bekensteinian bits possessed by the epistemically accessible area A of a black hole's event horizon is directly proportional to the number of internal configurations (i.e., microstates of the black hole) that are epistemically inaccessible to an external observer. Consequently, Bekenstein interpreted complementarily (à la Brillouin) the entropy $S_{BH}$ as a thermophysical measure of the inability of the external agent to distinguish configurations of the black hole beyond its event horizon: "It is then natural to introduce the concept of black-hole entropy as the measure of the inaccessibility of information (to the external observer) as to which particular internal configuration of the black hole is actually realized in a given case" (Bekenstein 1973, p.2335).

As it was remarked before, Bekenstein's argumentative strategy in 1973 is intellectually nourished not only by the Brillouinian tradition but also by Jaynesianism. This is precisely the case of the theoretical dependence of his proposal on the conceptual identification à la Jaynes (1957a) between $H_{IT}$ (defined on Planck-units Bekensteinian bits) and $S_{TD}$, as clearly seen in the reconstruction of the Bekensteinian strategy in 1973 by Christian Wuthrich (2017, p.9). On the one hand, the horizon area A of a black hole is conceptually identical to its epistemic $H_{IT}$ (measure of uncertainty); on the other hand, $H_{IT}$ is conceptually identical to $S_{TD}$ (i.e., Jaynesian conceptual identification); therefore, the horizon area A of a black hole should be conceptually identical to its $S_{TD}$. Strategically employing the informational concepts of the Brillouinian and Jaynesian traditions, Bekenstein (1972, 1973) was able to descriptively characterize (and theoretically legitimize) black holes as thermophysical phenomena by means of his notion of black hole entropy $S_{BH}$. In this sense, Bekenstein employed the Jaynesian device of reversing the standard theoretical procedure of non-informational thermophysics, employing entropic concepts to derive thermal properties of systems and not vice versa (see Section 3.3). Because of the presumable thermophysical content of this $S_{BH}$ concept, it would eventually contribute to the entropy $S_{TD}$ of the universe by saving the (generalized) Second Law of TD from being violated by a plausible Wheeler demon.

As can be deduced from its current name, the concept of black hole entropy was developed simultaneously by Bekenstein and Hawking with no greater intellectual cooperation than the common acceptance of the surface area of the event horizon as the fundamental concept that allows black holes to be characterized as thermal phenomena (Belfer 2014). Both arrived at the same conceptual result from substantially different interpretative strategies and theoretical goals: while Bekenstein relied on Brillouinian and Jaynesian resources to interpretatively legitimize the thermophysical content of his informational derivation of $S_{BH}$, Hawking initially argued that his non-informational interpretation of $S_{BH}$ was not to be understood in a thermophysically meaningful way[80]. Note that while Bekenstein (1973) explicitly spoke of "black hole thermodynamics", Hawking preferred to use non-thermal expressions such as "four laws of black hole mechanics". Paradoxically, Hawking's discovery in 1974 that black holes possessed a certain quantum radiation (now known as 'Hawking radiation') suggested the theoretical possibility of defining a concept of

[80] "Though working on the same problem apparently in parallel, at least in their initial theoretical results, Bekenstein and Hawking did not collaborate on the development of black hole thermodynamics. The entropy equation is commonly named after both of them, but they perceived it very differently. Bekenstein found a description for entropy proper, the actual thermodynamics of singularities, adopting the information-theoretical oriented mindset of Brillouin and Jaynes. [$H_{IT}$] was hence a physical attribute of the system in a generalized form of the quantified unknown information [OI] (...) On the other hand, Hawking, together with Bardeen and Carter, formulated "four laws of black hole mechanics" which are similar to, but distinct from, the four laws of thermodynamics. They capitalized on the correlation between the event horizon and the surface gravity, achieving a mathematical thermodynamic formulation, but took care to not treat it as an actual physical attribute of black holes." (Belfer 2014, p.83).

temperature applicable to these gravitational phenomena, theoretically supporting Bekenstein's position over Hawking's own. This was subsequently recognized by the British physicist, as his biographers pointed out: “Bekenstein, had been on the right track after all. What Hawking discovered was that black holes do radiate (in the form of particles) which means that they do have a temperature, and that the four laws of black hole mechanics which he, Bardeen and Carter had been so careful to stress as only analogous to thermodynamics, were actually laws of black hole thermodynamics in disguise” (Larsen 2005, p.40). As I am going to assess in the following Section, this apparent 'victory' of the Bekensteinian characterization of black holes as properly thermo-gravitational phenomena constituted historically a springboard of informationalism to the intellectual front lines of the physics community of the time.

## 4. 2. Intellectual Explosion of the Physics of Information in the 1990s

> “The specter of information is haunting sciences. Thermodynamics, much of the foundation of statistical mechanics, the quantum theory of measurement, the physics of computation, and many of the issues of the theory of dynamical systems, molecular biology, genetics and computer science share information as a common theme.” (Zurek 1990, p.vii)

As already explored in Chapter III, informationalism in thermophysics experienced a period of intellectual splendor during the fifties, starting with Shannon's bandwagon, from which the Brillouinian and Jaynesian traditions emerged. This first thermo-informationalist generation of the fifties had a decisive impact during the sixties and seventies, forming the conceptual ground from which a second generation of authors such as Tribus (and McIrvine 1971) or Bekenstein (1973) emerged (Section 4.1). The gradual assimilation of algorithmic information in the field of physics and the development of computational thermophysics, both promoted by Bennett (1982), conceptually fertilized the intellectual soil of the eighties from where the greatest stage of physical-informationalist thought in history flourished. This third generation, which emerged in the late eighties and nineties, was intellectually led by physics geniuses such as John Wheeler (1990) or Richard Feynman (1996), where multiple trends of thought and new physical-informational disciplines such as quantum computing or quantum information theory converged. It is precisely from this intellectual hegemony of informationalism within the scientific community of the nineties that 'information physics' emerges not as an alternative current of scientific thought (as occurred in the fifties, Section 3.1) but as a promising discipline from which to unify the fundamental domains of contemporary physics.

4.2.1. The Manifesto of Information Physics (1990)

If one must fix a historical moment in which physical informationalism acquired a greater awareness of its impact on the scientific community and an explicit will to consolidate itself as a disciplinary field of its own, this was the workshop on ‘Complexity, Entropy, and the Physics of Information’ held in May-June 1989 at the Santa Fe Institute (Santa Fe, New Mexico), one of the main institutions in the promotion of interdisciplinarity in physics in the 1980s. In this event, the theoretical proposals of the major intellectual figures in the history of physical informationalism, including John Wheeler, Wojciech Zurek, Charles Bennett, Edwin Jaynes, Carlton Caves, Seth Lloyd, Asher Peres, Paul Davies or Benjamin Schumacher, among many others, were displayed. Here not only did the main thermo-informational traditions (i.e., Brillouinism, Jaynesian and Bennettian or computational thermophysics) converge historically, but they also coordinated their particular intellectual pretensions under the same theoretical goal: the creation of information physics as a fundamental research domain with respect to a wide range of scientific areas. The main results of this meeting were collected as proceedings in the homonymous volume *Complexity, Entropy, and the Physics of Information* (ed. Zurek 1990), constituting the foundational work in which information physics is born as an institutionalized discipline (in spite of its more than four decades of previous intellectual history. As a foreword to this volume, the physicist Wojciech Zurek (born 1952) (as editor and spokesman of this intellectual movement) deployed a manifesto for information physics, highlighting the growing centrality of informational concepts in all areas of fundamental physics, starting with thermophysics (TD-SM): "The specter of information is haunting sciences. Thermodynamics, much of the foundation of statistical mechanics, the quantum theory of measurement, (…) Black hole thermodynamics (…)" (Zurek 1990, p.vii). Note that each of these pillars of informational physics developed during the second half of the twentieth century corresponds to each fundamental physical domain (i.e., thermophysics, quantum, and relativity), suggesting the claim to constitute information physics as a vehicle for the theoretical unification of these domains, as was clear in Wheeler's proposal.

In this manifesto, Zurek recognized that the main intellectual engine of information physics comes directly from the historical deployment of informational thermophysics seen in Chapter III (Shannon's bandwagon, Brillouinism and Jaynesianism), relying on the plurality of interpretative strategies to interrelate entropic and informational concepts. Of course, the informational (Brillouinian or Bennettian) tactics to exorcize Maxwell’s problem played a central role in popularizing thermal informationalism as an epistemically successful physical trend. Therefore, one of the foundational pillars of the physics of information was grounded on “A deep analogy between thermodynamic entropy [$S_{TD}$] and Shannon’s information-theoretic entropy [$H_{IT}$]. Since the introduction of Maxwell’s Demon and, particularly, since the celebrated paper of Szilard and even earlier discussion of Smoluchowski, the operational equivalence of the gain of information and the decrease of entropy has been widely appreciated. Yet, the notion that a subjective quantity such as information could influence “objective” thermodynamic properties of the system remains highly controversial” (Zurek 1990, p.vii).

The second pillar of information physics according to Zurek's manifesto corresponds to the problem of quantum (though also classical) measurement. As I have discussed in section 3.2, scientific measurement processes began to be systematically interpreted in informational terms (i.e., information acquisition processes) already since the Brillouinism of the early 1950s. This made it possible to informationally reconceptualize the results of Szilard (1929) and von Neumann (1933), who argued that measurement processes (classical and classical-quantum, respectively) had direct thermal consequences "It is, however, difficult to deny that the process of information acquisition can be directly linked to the ability to extract useful work. Thus, questions concerning thermodynamics, the second law, and the arrow of time have become intertwined with a half-century puzzle, that of the measurement problem in quantum physics" (Zurek 1990, p.vii). Undoubtedly, the problem of measurement is one of the main interpretative problems central to the field of quantum physics, and thus to all of physics. That the use of informational concepts can epistemically contribute to shed light on this question would mean that informationalism is indispensable for the progress of modern physics: "Only very few discussions of the measurement problem in quantum theory make an explicit effort to consider the crucial issue- the transfer of information. Yet obtaining knowledge is the very reason for making a measurement. Formulating quantum measurements and, more generally, quantum phenomena in terms of information should throw a new light on the problem of measurement, which has become difficult to ignore in light of new experiments on quantum behavior in macroscopic systems" (Zurek 1990, p.vii).

After the thermophysical and quantum domains, the other area of fundamental physics in which the informationalist current of thought was introduced was general relativity. As I evaluated in Section 4.1, this occurred precisely with Bekenstein's (1973) proposal to describe black holes informationally, encompassing not only the relativistic domain but also the thermophysical and later (after Hawking's [1974] results) the quantum domain: "Black hole thermodynamics has established a deep and still largely mysterious connection between general relativity, quantum, and statistical mechanics. Related questions about the information capacity of physical systems, fundamental limits on the capacity of communication channels, the origin of entropy in the Universe, etc., are a subject of much recent research." (Zurek 1990, p.viii).

After presenting the thermophysical, quantum and relativistic pillars of information physics, Zurek continues the manifesto of this current of thought by detailing its recent computational-centric (i.e., information processing) extension since the 1980s. In this sense we would generally speak of 'physics of computation' as a branch of informational physics, which encompasses not only the then underdeveloped quantum computation (interestingly, Bennett [1990] was one of its founders) but above all the thermophysics of computation developed by Bennett and Landauer in the 1980s: "physics of computation explores the limitations imposed by the law of physics on the processing of information. It is now established that both classical and quantum systems can be used to perform computations reversibly (…) These results are of fundamental importance, as they demonstrate that, at least in principle, processing of information can be accomplished at no thermodynamic cost. Moreover, such considerations lead one to recognize that is actually the erasure of the information which results in the increase of entropy" (Zurek 1990, p.viii).

Finally, and also in the line of the physics of computation, Zurek acknowledges the recent impact of the concept of algorithmic information within informationalist thought (see Section 4.1). As I aim to show below, several authors of this golden period of physical informationalism capitalized on this concept to generate baroque theoretical architectures that could satisfactorily cope with the new intellectual objectives: "Algorithmic randomness – an alternative definition of the information content of an object based on the theory of computation rather than on probabilities- was introduced more than two decades ago by Solomonov, Kolmogorov, and Chaitin (…) It is tempting to suggest that physical entropy (…) should take into account its algorithmic randomness (…) It provides a direct link between thermodynamics, measurements, and the theory of computation" (Zurek 1990, p.xi).

#### 4.2.2. Zurek's Themo-Informational Program and Physical Entropy

One of the intellectual summits of the golden period of information physics can be found in the conceptually ambitious and interpretatively sophisticated theoretical proposal of Zurek (then a member of the Santa Fe Institute, which marked his intellectual profile) first deployed in his 1989 paper 'Algorithmic randomness and physical entropy' and expanded in his contribution to 'Complexity, Entropy, and the Physics of Information' (Zurek 1990). On the one hand, it was a conceptually ambitious proposal in the sense of pretending to integrate under the same conceptual architecture the main notions of thermophysical ($S_{TD}$, $S_B$, and $S_G$) and informational ($H_{IT}$ and K) entropy, whose range of concepts is not comparable with any other proposal in the whole previous history of thermophysical informationalism (Chapter III). This claim of conceptual interrelatedness was termed by Ladyman and Ross (2007, p.213-214) as 'grand reification'. On the other hand, it was an interpretatively sophisticated proposal because of its claim to coherently unify Brillouinian and Jaynesian interpretative strategies (a characteristic feature of the post-1970 thermo-informationalism, as noted in Section 4.1) to theoretically legitimize its conceptual architecture: "Particularly influential works arguing in favor of the identification of entropy with missing information are due to Jaynes and Brillouin. Yet all but a few of the most fervent supporters of the information-theoretic interpretation will agree with the notion that individual states of physical systems can be either ordered or random regardless of our information about them" (Zurek 1989, p.4741). Then, Zurek (1989, 1990) intended to build his theoretical edifice on the (objectively interpreted) randomness of the individual microstates of physical systems, thus breaking with the previous informationalist tradition centered on the knowledge of agents about these microstates (Brillouinism) or about the ensembles that contain them (Jaynesian).

Instead of distinguishing between thermophysical entropic notions and informational notions, the Zurek proposal starts from a conceptual distinction between entropic-informational concepts defined over sets of objects or probabilistic ones ($S_B$, $S_G$, or $H_{IT}$) (which initially assumes their identification à la Jaynes [1957a]) and those defined over individual objects, i. e. algorithmic entropy K. On the one hand, this author followed an epistemic interpretation (in the Jaynesian tradition) of the unified $S_B$-$S_G$-$H_{IT}$ concept, so that the ensemble-like ρ set of objects on which it is defined represents epistemic states of the agents: "Boltzmann-Gibbs-Shannon entropy has an information-theoretic flavor and is a subjective property of the microstates (…) This entropy is an (objective) property of an ensemble rather than of a microstate. However, ensembles are usually defined subjectively in a manner that may depend on the state of the knowledge of the observer" (Zurek 1989, p.4745). As a straightforward consequence of this epistemic conception of $S_B$-$S_G$-$H_{IT}$, it would be complementarily connected with OI via a classical 'missing information' interpretation. On the other hand, Zurek defends an objective conception of the individual microstates of physical systems, which are interpreted according to the Brillouinian tradition from an ordinary concept of disorder (OO) or randomness (Section 2.2): "A broken glass (…) is more random than the unbroken one not because our ignorance about it has dramatically increased when it was shattered by the hammer, but because it has become more disordered" (Zurek 1989, p.4741).

The next step of his interpretative strategy is to technically specify the content of the ordinary concept of randomness-disorder (OO) of microstates by means of the concept of algorithmic information K, which is possible because both are defined on individual objects (microstates and sequences, respectively). Once the content of the algorithmic information K is thermophysically interpreted as a technical measure of the objective randomness of individual microstates, it would also be reconceptualized as a notion of 'algorithmic entropy': "The intuitive notion of randomness is, we believe, formally expressed by the algorithmic definition of entropy. We have demonstrated that equilibrium systems are expected to have the same equilibrium thermodynamic properties regardless of which of the two paradigms is employed to define their entropy. Therefore we have the "objective[81]" quantity -algorithmic randomness- which measures the disorder in the system" (Zurek 1989, p.4741). That is, the thermophysical content of the algorithmic entropy depends on its ability to represent the macroscopic observable values in a numerically identical way as they are represented by the unified entropic concept $S_B$-$S_G$-$H_{IT}$. Extending the then classic Bennettian (1982) connection[82] between $H_{IT}$ and K, Zurek demonstrates in his 1989 paper that the numerical connection between the probabilistic entropy $S_B$-$S_G$-$H_{IT}$ of an ensemble ρ and the algorithmic entropy K of a microstate $x$ holds precisely in the applicative domain of macroscopic thermal equilibrium. Another decisive result of his theoretical proposal (Zurek 1989, Appendix B) is the demonstration that the algorithmic entropy K of a microstate coincides numerically (and is therefore conceptually identified) with the $S_{TD}$-$S_B$ of the microstate of a monoatomic ideal gas, which in Section 2.5. I called the 'Golden Entropy Triangle'.

[81] As the reader might have noticed, that K can be defined for individual microstate does not entails that K can be straightforwardly considered as an objective quantity.

[82] Zurek(1990) pointed out that "More recently, Bennett, in an influential paper [1982] has pointed out that the average algorithmic entropy of a thermodynamic ensemble has the same value as its statistical (ensemble) entropy".

In line with Zurek's theoretical strategy, the very concepts of probabilistic entropy and algorithmic entropy are not conceptually identified but possess a complementary and interpretatively separate conceptual interrelationship. Combining (i) his epistemic Jaynesian interpretation of probabilities and (ii) the fact that K is defined independently of probabilities, then (iii) K could be interpreted as a measure of an ontic-objective property (disregarding its description-size dependence): “algorithmic randomness can be expressed as the number of bit in the smallest program for a universal computer can reproduce the state in question (…) In contrast to the traditional definitions of entropy, algorithmic randomness can be used to measure disorder without any recourse to probabilities (…) Algorithmic entropy is a property of an objective property of a microstate, except for the O(1) corrections which are related to the size of the description” (Zurek 1989, p.4731, 4745). It is then that Zurek (1989, p.4741) constructs his fundamental concept of 'physical entropy' defined by the sum of the probabilistic entropy-information of the ensemble and the entropy-information of the microstate. In this sense, Zurekian physical entropy is an entropic-informational concept that encompasses the complementary functions of (a) statistical entropy quantifying the agent's lack of information (agent-dependent epistemic property) and (b) algorithmic entropy quantifying the disorder-randomness of the microstate of the molecular system (ontic and objective property). This amphibolic character of physical entropy is illustrated by Zurek (1989, p.4741) with the motto 'Physical Entropy = Missing Information + Known Randomness'.

In his contribution to the book-manifesto of information physics entitled 'Algorithmic Information Conte, Church-Turing Thesis, Physical Entropy, and Maxwell's Demon', Zurek (1990) sought to show the epistemic potential of his articulated theoretical proposal on the concept of physical entropy through its systematic application to the problem of scientific measurement and to that of Maxwell's demon. In the case of the first application domain, Zurek informationally reconceptualized scientific measurement processes as transformations of statistical entropy $S_B$-$S_G$-$H_{IT}$ quantities (encoded the agent's lack of information about the microstate) into algorithmic entropy K quantities (encoded the molecular disorder of the microstate): “Measurements can only convert uncertainty (quantified by the statistical entropy) into randomness of the outcome (given by the algorithmic information content of the data). The ability to extract useful work is measured by physical entropy, which is equal to the sum of these two measures of disorder. So defined physical entropy is, on the average, constant in course of the measurements carried out by the observer on an equilibrium system” (Zurek 1990, p.73). Note that analogous to the NPI within Brillouinism (in scientific measurements the indistinguishability of the agent’s $I_{Bb}$ decreases while the entropy $S_B$ of the system increases, leaving the sum N + $I_{Bb}$ invariant, see Section 3.2), the sum of $S_B$-$S_G$-$H_{IT}$ and K remains invariant during measurements in equilibrium contexts, decreasing the 'lack of information' of the observer and increasing the 'randomness' of the system in compensation. However, Zurek's concept of physical entropy allows to describe in thermo-informational terms the measurement processes without thermophysical consequences (as it happens with Szilard's principle in the case of Brillouinism): “Indeed, physical entropy S has the great advantage of removing the “illusion” that entropy decreases in the course of a measurement.

Of course, the ignorance (measured by H) does indeed decrease, but only at the expense of the increase of the minimal record size K" (ibid, p.88).

On the other hand, Zurek (1990, p.87) suggests that, in non-equilibrium scenarios, the increase in ontic randomness of the system decreases at a slower rate than the decrease in uncertainty of the agent as the number of measurements increases. According to the author, this opportunity provided by non-equilibrium processes is exploited by various 'intelligent beings' (i.e., epistemic agents) to their advantage, which would include Maxwell's own demon. Zurek wondered "how can one analyze actions of an "intelligent being" within the realm of physics?" (...) a physically sensible definition of what is an "intelligent being" was not easy to come by, especially in the times of Szilard" (ibid., p.81). In order to give a "physically sensible definition" beyond the simplistic models that dominate the history of informationalism, Zurek proposed to characterize Maxwell's demon as an 'information gathering and using system" or "IGUS", which consists of an entity (whose actions are either automated or not) that is able to make measurements and change its strategy depending on the measurement output. Interestingly, Zurek (1989a) constructed an exorcism of the Maxwellian demon on the basis of the Szilardian model (along the lines of Brillouin [1956] or Bennett [1982]), naturalistically redescribing its ability to store the position of the molecule in the form of an algorithmic sequence of zeros or ones. Assuming the validity of Landauer's principle and criticizing the inadequacy of the Bennettian exorcism (Earman and Norton 1999, p.17-18), Zurek claimed that an (algorithmically) 'efficient demon[83]' could strategically reduce the LP-dependent entropic costs of bit erasure by compressing its sequence-memory $s_i$ into a sequence-memory $s_i$ of smaller size, i.e., $|s_i| > |s_j|$. In short, Zurek's (1989, 1990) systematic proposal constitutes one of the historical summits of thermo-informationalism in its golden age, whose interpretative coherence and conceptual complexity possess a degree of originality and sophistication only intellectually comparable to other systematic (pre-algorithmic and not-computational) frameworks such as those of Brillouin (1956) and Jaynes (1957a).

### 4.2.3. Caves' Thermo-Informational Program

A sample of the enormous immediate impact that Zurek's (1989, 1990) proposal had within the physics community of the nineties can be found in the thermo-informational framework of the pioneer of quantum information theory Carlton M. Caves (born in 1950). Interestingly, the intellectual origin of Caves' theoretical proposal displayed in his 1990 'Entropy and Information: How Much Information is Needed to Assign a Probability?' constitutes a significant example of the accumulation of conceptual layers on which the new informational currents in thermophysics of the end of the century emerge: "First, I have been decisively influenced by E.T. Jaynes, the

[83] "For it took only a modest concern for economy in programming a devise an alternative Demon whose efforts elude exorcism by the anaugment Bennett-Landauer analysis. Such clever Demons were investigated by Caves (1990). These Demons seek to reduce the Landauer erasure cost by focusing on favourable states and compressing the descriptions used to record information" (Earman and Norton 1999, p.18).

leading exponent of the Bayesian view of probabilities (…) I take the Bayesian view explicitly as my starting point. Second, my interest in algorithmic information was sparked by Charles Bennett's Scientific American article on Maxwell's demon (…) Finally, (…) I discovered that much of what I had thought and dreamed about had been anticipated and developed by Wojciech Zurek. (…) Indeed this paper should be read in conjunction with Zurek's two paper" (Caves 1990, p.93). On the one hand, the Jaynesian tradition provides Caves with (i) an epistemic (though objective) interpretation of the probabilities on which $H_{IT}$ and $S_G$ are defined, and consequently (ii) the conceptual identification between $H_{IT}$ and $S_G$ (Section 3.3). On the other hand, Bennett (1982) provides (iii) the numerical connection between $S_G$-$H_{IT}$ and K (Section 4.1). Finally, Caves follows Zurek in his claim to (iv) provide interpretatively thermophysical content to the existing connection between $S_G$-$H_{IT}$ and K. This separation of the conceptual layers underlying the main thermo-informationalist proposals of the 1990s (namely, Zurek's and Caves') shows the increased intellectual complexity of this movement during this historical stage.

But not all these proposals in the nineties were intellectually rooted on the same first (post-1950s Brillouinism or Jaynesianism) or second (post-1970s syncretism and/or thermo-computationalism) generation of theoretical frameworks. In this sense, while Zurek starts (implicitly) from a Brillouinian interpretative strategy of entropy as a measure of disorder-randomness OO of microstates, Caves explicitly employs the Jaynesian interpretation of probability densities (and thus also epistemically conceiving $S_G$ and $H_{IT}$) as the basis of his proposal: "The probability assignment [ρ] reflect what one knows about the system" (Caves 1990, p.93). It should be noted that the thermo-informational proposals of Zurek and Caves not only have differences in their intellectual origins, but also in their interpretative strategies (where the former affect the conceptual scope of the latter). This is particularly significant in the specification of the meaning of the concept of algorithmic information K, where "he [Zurek] sees algorithmic information as a system property -specifically, a property of microstates- which supplements or even replaces the usual statistical definition of entropy; I, starting from a Bayesian orientation, see algorithmic information as a memory property, helping us to understand how we describe physical systems" (ibid). That is, whereas in the Zurekian interpretation algorithmic information K is reconceptualized as a measure of the molecular (and hence ontic) disorder of physical microstates, from the Neo-Jaynesian position of Caves (1990) algorithmic information K is epistemically reconceptualized as a property of certain states of knowledge (i.e., memory states).

Similarly, while for Zurek (1989) the difference between the two types of entropic-informational concepts manifests itself in the distinction between an epistemic and an ontological property (namely $S_B$-$S_G$-$H_{IT}$ and K, respectively), for Caves the conceptual difference between the two notion lies in their particular function with respect to statistical descriptions of the system states. Namely: on the one hand, the concept $S_G$-$H_{IT}$ allows to quantify the statistical OI-like information (encoded in posterior probabilities) acquired after updating the memory during a measurement; on the other hand, the concept K allows to quantify the OI-like information (i.e., epistemic and semantic, since it is *about* something, and agent-based) necessary to assign prior probabilities: "A memory, based on the prior [OI] information that it has *about* a physical system, assigns probabilities to system states. The average information that the memory acquires, should it discover the state of the system, is measured by the Gibbs-Shannon [$S_G$-$H_{IT}$] statistical information. The information content of the prior information is quantified by the algorithmic information [K] to generate the probability assignment." (Caves 1990, p.92)

In his 1994 paper 'Information, Entropy, and Chaos', Caves distinguished between two types of entropy-information conceptual connections that underlie not only his proposal, but all of information physics as a discipline "How are information and entropy connected? (…) The first of these connections is a purely mathematical relation [$H_{IT}$-$S_G$] between information and entropy; physics is introduced by the second connection [LP], which relates information to free energy and work (…) Information having been granted a physical status, one can talk about a "physics of information"" (Caves 1994, p.47). With this distinction, the author fundamentally separates between those strictly formal entropy-information conceptual connections of the Brillouinian and Jaynesian traditions[84] (Chapter III) and those conceptual connections with physical content delimited by the Szilard principle (Section 3.1.2) and Landauer's principle. In this direction, Caves (1994, p.47) separates the two main interpretative strategies in information physics according to how the epistemic states (i.e., memory) analysed are conceived. On the one hand, from an 'outside' (or naturalistic) interpretation these forms of epistemic information [OI] are characterized as physical states, which according to Caves might trivialize the entropy-information connection. This, of course, can be conceive as an attempt to avoid agency and psychological characterization of information concepts: "To avoid excessive use of "we" and "one", I use the following construct a physical system is observed by a "memory," which gathers, stores, and manipulates information about the system." (ibid. 48). On the other hand, from an 'inside' (or non-naturalistic) interpretation of informational states or bits are characterized simply by their capacity to refer to something of the system, where the entropy-information connection would ultimately depend on an absolute measure of information such as K[85].

## 4.3. Informational Ontologies in the Golden Age of Informationalism

One of the most evident symptoms that a current of thought has become intellectually consolidated within a scientific community is the appearance of the first ontological proposals derived from the extended use of informational concepts and interpretations. In the case of thermophysical informationalism, this occurred precisely during its golden period in the 1990s, when theoretically

[84] Caves summarizes the history of thermo-physical informationalism as follows "The close connection between information and entropy has been appreciated for some time, notably by Szilard (1929) in a seminal paper, by Brillouin (1962) in his identification of information with negentropy, and by Jaynes (1983) in a lifetime of work elucidating the Bayesian view of probabilities" (Caves 1994, p.48).

[85] With this distinction between inside-outside strategies, Caves (1994) set a precedent for what six years later was used as the central idea underlying the 'Earman-Norton (1999) dilemma'.

explicit statements intentionally aimed at emphasizing the physical reality of information began to emerge[86]. I would argue that the enormous increase in the intellectual credibility of informationalism during the eighties (for the reasons seen in Section 4.2) allowed a generalized radicalization of its theoretical positions, giving rise to the emergence of multiple ontological theses in the early nineties such as the celebrated "It from the Bit" (Wheeler 1990, p.3), "Information is Physical" (Landauer 1991), "the world is algorithmically compressible" (Davies 1990, p.63), "information is as much a part of the physical universe as is matter and energy" (Stonier 1990, p.2). Let us now evaluate the conceptual foundations and interpretative strategies underlying each of the above proposals for informational ontologies.

### 5.3.1. Wheelerian Informational Metaphysics

John Wheeler is one of the most important theoretical physicists of the second half of the twentieth century. The key to his scientific output concerning the historical drift of the physics of information physics is not so much his particular achievements (among which we find, for example, the Wheeler-DeWitt equation or the description of the Breit-Wheeler process in quantum field theory) but his own intellectual evolution during the second half of the twentieth century. In his biography *Geons, Black Holes, and Quantum Foam: A Life in Physics*, Wheeler (1998) explores how his belief about what is the fundamental physical domain that could account for any physical phenomenon (what is known among the scientific community as the 'theory of everything') shifts from particle physics during the 1950s, to quantum and gravitational field theory in the 1960s and 1970s, to finally landing on the informational currents in physics during the 1980s and 1990s. Note that each of these intellectual stages (particle-centric, field-centric and information-centric) coincides historiographically with each of the generations of informationalist authors (i.e., fifties-sixties, seventies-eighties, late-eighties-nineties) discussed at the beginning of Section 4.2:

> "I think of my lifetime in physics as divided into three periods. In the first period, extending from the beginning of my career until the early 1950s, I was in the grip of the idea that Everything Is Particles. I was looking for ways to build all basic entities—neutrons, protons, mesons, and so on—out of the lightest, most fundamental particles, electrons, and photons… I call my second period Everything Is Fields. From the time I fell in love with general relativity and gravitation in 1952 until late in my career, I pursued the vision of a world made of fields, one in which the apparent particles are really manifestations of electric and magnetic fields, gravitational fields, and space-time itself. Now I am in the grip of a new vision, that Everything Is Information. The more I have pondered the mystery of the quantum and our strange ability to comprehend this world in which we live, the more I see possible fundamental roles for logic and information as the bedrock of physical theory." (Wheeler 1998, p.299)

[86] Note that this does not occur with the first generation in the fifties, where the concepts of information are assimilated fundamentally from an epistemic and not an ontological dimension (Chapter III), nor with the second generation with Tribus or Bekenstein in the seventies (Section 4.1) or the computational thermophysics of the eighties, where the ontic nature of information is still not explicitly defended.

Following Belfer's (2014, p.88) historiographical thesis, the Bekenstenian derivation of black hole entropy via information theory evaluated in Section 4.2. made an enormous impact on his teacher Wheeler, particularly from his new stint at the University of Austin from 1976 onwards. His expectation that Shannon's IT and informational concepts could play a decisive role within fundamental physics (i.e., quantum gravitation) was progressively increasing during the last decades of the twentieth century, as the informationalist current of thought increased its intellectual capital within the physics community. An example of this is the anecdote told by Bekenstein in his memoirs. He recounts that when he went to visit his teacher in 1986, the American was fully immersed in placing Shannon's information measure at the core of theoretical physics: "Those months in Austin were a pleasant return to Wheeler's scientific circle. He had more time for talking physics then. And many of these talks were about the role of information in physics" (Bekenstein 2006, p.111). As his disciples (including Zurek himself) mentioned in the speech they dedicated to Wheeler in the journal 'Physics Today' (Misner et al. 2009), it was precisely in this period when Wheeler laid, together with Bennett and Schumacher (1995), the conceptual foundations of what we now know as 'quantum information theory'. His position continued to become more radical during the late 1980s until he made his inaugural contribution to the volume 'Complexity, Entropy and the Physics of Information' analysed in Section 4.2, entitled ‘Information, Physics, Quantum: The Search for Links’ (Wheeler 1990, p.3-28). In this text we find his famous dictum:

> “every it -every particle, every field of force, even the spacetime continuum itself- derives its function, its meaning, its very existence entirely (…) from the apparatus-elicited answer to yes-or-not questions, binary choices, bits. It from bit symbolizes the idea that every item of the physical world has at bottom- at a very deep bottom, in most instances- an immaterial source and explanation; that which we call reality arises in the last analysis from the posing of yes-no questions and the registering of equipment-evoked responses; in short, that all things physical are information-theoretic in origin and this is a participatory universe” (Wheeler 1990, p.3)

As recent history has shown, these famous statements by Wheeler soon became the anthem of physical informationalism during this golden age, turning "It from the Bit" into the timeless slogan that popularized this intellectual movement within the scientific community. What differentiates this paragraph from any other historically previous physical-informationalist statement is not properly its overflowing intellectual ambition, but the evident will to defend an explicitly philosophical-ontological thesis and not a physical-theoretical thesis. As an ontological ("its very existence", "immaterial source") and epistemic ("explanation") thesis, Wheeler's (1990) informational metaphysics proposal is extremely vague in terms of interpretatively specifying the particular position he proposes. Thus, we agree completely with the analysis of Ladyman and Ross “when Wheeler suggests this he is clearly doing philosophy rather than physics, and we suspect the philosophy in question is naïve (…) A philosopher might be inclined to tremble at the boldness of this, but for the fact that it can be interpreted to suit any taste along the spectrum of metaphysical attitudes from conservative to radical” (Ladyman and Ross 2007, p.186, 212). Despite this ontological vagueness of the proposed Wheelerian pan-informational metaphysics, the 'all is information' thesis (whatever its meaning) is theoretically and unambiguously legitimized in the 'promising' ability of informational concepts to advance the fundamental unification of (i) thermophysics, (ii) general relativity and (iii) quantum mechanics in the direction of quantum gravity and black hole physics. As noted above, his disciple Bekenstein's (1973) use of informational concepts to quantize the discrete structure of space-time in the black hole domain (Section 4.1) constitutes the intellectual fuel underlying the Wheelerian program of developing a theory of everything on the basis of an ambitious physical-informational theoretical framework.

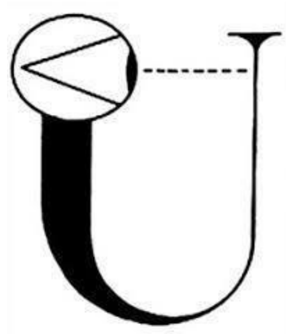

**Figure 4.** Wheeler’s Informational Uroboros (Zurek 1990, p.xiv)

Wheeler (1990) defended the metaphysical thesis that physical reality itself depended ontologically on how information [OI] about it was obtained by scientific measurements (i.e., by binary 'yes-no' question combinations), intellectually relying on the long-standing tradition of epistemizing (von Neumann, 1933) and informationalizing (e.g., Brillouin, 1956; Zurek, 1990) quantum measurements and their results obtained in the delayed-choice experiment. For example, an agent experimentally asks the question 'Does this system have molecular position values between $X_0$ and $X_2$? And we obtain as a result: 'yes'. Then, does this system have values between $X_0$ and X1? And so on. This is precisely what Wheeler calls the 'participatory universe', including both the agent-observer and the system-observed as part of that universe. This is illustrated by the way in which Wheeler represented the universe to the general public, as a capital 'U' surmounted by an eye (representing the agent-observer or receiver of information) that observes itself (representing the observed system or sender of information), see Figure 4 above. This image went down in history as the flagship of information physics (see Zurek 1990, p.xiv).

### 4.3.2. Landauerian Informational Naturalism

Apart from the Wheelerian informational metaphysics proposal, the famous slogan 'Information is Physical' contains the other great ontological thesis that became popular during this historical stage thanks to Rolf Landauer. Landauer underwent a pronounced intellectual evolution during the second half of the twentieth century. From modestly exploring certain practical limits in the development of physical computers in the sixties[87], Landauer took advantage of the exponential rise of computational thermophysics and information physics in the late 1980s to defend explicitly ontological proposals regarding informational concepts. This radical change of tone and proto-philosophical pretensions was deployed throughout several articles written during his last decade of life and academic production, mainly the well-known 'Information is Physical' (Landauer 1991) and his posthumously published 'Information is a Physical Entity' (Landauer 1999). With the intention of emphasizing the differential value of his proposal of informational metaphysics as opposed to Wheeler's, Landauer pointed out that “Wheeler’s suggestion, that the laws of physics are interlinked with the evolution of the universe and our observation of it, is not equivalent to my proposal” (Landauer 1991, p.29).

Unlike Wheelerian metaphysics, where it is postulated that 'informational principles' ontologically determine the fundamental physical regularities underlying the behavior of particles and fields, Landauer’s informational-naturalistic ontology is based on the other hand on defending that (i) physical laws constrain or even determine informational elements and (à la Wheeler) that (ii) informational principles constrain physical reality. In the line of Wheeler's proposal, this Landauerian ontological statement 'Information is Physical' is sufficiently interpretatively vague and philosophically imprecise (e.g., indefiniteness of the meaning of 'physical law') to imply a certain degree of argumentative circularity “Others have, in a variety of ways, suggested that space and time in the universe are not really described by a continuum and that there is some sort of discretization or some limit of information associated with a limited range of space and time (…) Information handling is limited by the laws of physics and the number of parts available in the universe; the laws of physics are, in turn, limited by the range of information processing available

[87] As I mentioned in Section 1.4, unfortunately the historical and conceptual analysis of Landauer's principle falls outside the scope of this work due to reasons of length and subject matter (i.e., non-computational informational physics). However, here I will give a brief summary of the main lines of this analysis. The so-called “Landauer’s principle” originates historically (Landauer 1961) as a modest thesis addressed to the practical field of physical construction of computers, being understood as a result of the intellectual confluence of various informational currents of the late 1950s and early 1960s within the context of IBM research. The rise of the thermophysics of computation as an informationalist tradition in the 1980s by Bennett (1982, 1987) generated an intellectual trend that turned Landauer's thesis into a universally valid principle of physics and one of the main ways to save the second law of TD. This peculiar historical evolution (driven by the radical informationalism of the 1990s, Chapter IV) provoked at the turn of the millennium an incipient philosophical interest in the conceptual foundations of this informational idea. Since the early 2000s, the philosophical community has been sharply divided between authors such as Ladyman (et al. 2007, 2013) who sought to conceptually reinforce its theoretical validity and critics such as Norton (2005, 2011) who warned of the conceptual inconsistencies on which computational thermophysics is built. Regardless of the constructive or destructive pretensions of these authors, the more than two decades of philosophical evaluation of this informational idea have proven to be highly fruitful. Fine-grained studies of the conceptual basis of Landauer's principle such as those of Hemmo and Shenker (2012, Chapter 11; 2019) have shown that this idea cannot justifiably be assumed as a universally valid principle due to its thermophysical and computational foundations, challenging the multiple layers of reconceptualization and reinterpretation that have been erected on Landauer's (1961) original thesis throughout its history “When Landauer (1961) introduced the principle, it was little more than a promising speculation, supported by a sketchy plausibility argument. Over half a century later, one might imagine that this would be sufficient time to place the principle on a more secure foundation. This has not happened. It is not for want of trying. However (…) the now burgeoning literature on Landauer’s principle persists in committing repeatedly a small set of interconnected errors in thermal analysis.” (Norton 2018, p.247).

(…) I do suggest that there is a strong two-way relationship between physics and information handling” (Landauer 1991, p.29). This apparent ontological co-dependence postulated by Landauer between 'physical principles' and 'informational principles' (I assume that the author has in mind his ‘Landauer’s principle’) ultimately depends on the validity of his central metaphysical ‘Information is Physical’ premise, wherein: "Information is not an abstract entity but exists only through a physical representation, thus tying it to all the restrictions and possibilities of our real physical universe" (Landauer 1999, p.63).

As in his seminal 1961 paper, part of the validity of Landauer's theoretical proposal (namely, his Landauer’s principle) resides in the fact that his concept of information and bits interpretatively acquire physical dimension simply by being defined on the possible microstates Γ of the physical system that computationally processes such information. That is, being philosophically charitable with this author's approach, information values only acquire existence through their 'representation' or 'implementation' via the phase space Γ (classical existence) or the Hilbert space (quantum existence) of a physical system. This idea was explicitly developed in this 1999 paper: “Information is inevitably inscribed in a physical medium. It is not an abstract entity. It can be denoted by a hole in a punched card, by the orientation of a nuclear spin, or by the pulses transmitted by a neuron. The quaint notion that information has an existence independent of its physical is still seriously advocated. This concept, very likely, has its roots in the fact that were aware of mental information [OI] long before we realized that it, too, utilized real physical degrees of freedom” (Landauer 1999, p.64). Without going into the conceptual consistency of his premise of 'physical representation' of information, Landauer intended to conceptually connect the technical-statistical notion of information to which he initially refers with the ordinary notion (OI) by means of a physicalist perspective on the mental states on which they are based. In short, this last stage of Landauer's intellectual production in the nineties is characterized by the pretension (more philosophical than physical-theoretical) of extracting the radical ontological-naturalistic consequences of the thermophysics of computation.

### 4.3.3. Davies Algorithmic-Informational Framework

After Wheeler and Landauer as ‘informational metaphysicians’ and founding fathers of the physics of information in the 1990s, many other authors followed in their path, profited from the rise of this intellectual current to defend straightforwardly ontological theses. This was the case of the English theoretical physicist and broadcaster Paul Davies (born in 1946), whose active militancy in the informationalist movement has continued during the last decades. In a contribution to the bible of physical informationalism (i.e., Zurek’s [1990] *Complexity, Entropy, and the Physics of Information* eds.), entitled 'Why is the Physical World so Comprehensible?', Davies (1990) benefited from the suggestive terminological connection between 'comprehensible' and 'compressible' to develop an informational metaphysical framework around the concepts of algorithmic information K and its physical applicability. In this proposal, Davies employed the recently-developed Bennettian (1987) concept of 'logical depth' (Section 4.2) to describe the non-random nature of the complexity (in an AIT-based sense) of much of physical reality, implying that "There exists a real external world which contains certain regularities. These regularities can be understood, at least in part, by a process of rational enquiry called the scientific method" (Davies 1990, p.62). That is to say, scientific analysis allows us to account for the non-random complexity with which certain physical phenomena are organized. It should be noted that Davies assumes not only a standard complementarist interpretation of entropy as a measure of the lack of information (OI) of an agent about the system, but also a reconceptualization of logical depth as a measure of the organization of the system which “refers instead to the quality of information [OI]” (ibid).

In this direction, Davies (1990) sought to exploit the concepts of Chaitin's AIT to describe the logic underlying physical practices, where scientists seek to develop or find regularities in data sets that allow them to be compressed by means of laws-algorithms. As an illustrative example, Davies characterizes Newton's laws as algorithms that allow scientists to exploit certain regularities they find in an encoded data set of planetary positions, where such Newtonian laws-algorithms allow them to algebraically-geometrically compress this positional data set (or any other similar set) in a way that such predictable positions became causally connected: “The existence of regularities may be expressed by saying that the world is algorithmically compressible (…) Viewed this way, the question “Why is the universe knowable?” reduces to “Why is the universe algorithmically compressible” and “Why are human beings so adept at discovering the compressions?” (…) the very fact that we exist as observers already constrains the universe to have the property of algorithmic compressibility (…) Secondly, there is a wide class of physical systems, the so-called chaotic ones, which are not algorithmically compressible” (Davies 1990, p.63). That is, for a dataset about a physical phenomenon to be algorithmically comprehensible by some scientific law according to this Davies framework means that the agent can epistemically 'comprehend' or obtain epistemic-semantic information (OI) about the phenomenon (i.e., allowing predictions, retrodictions, etc.), the level of quality of this information being proportional to the degree of compressibility of the dataset. In this sense, the fact that the data sets of chaotic systems are algorithmically incompressible by the main physical laws would indicate that these systems are 'incomprehensible', so that agents are unable to obtain predictive information (OI) about the behavior of these systems soon after their evolution.

### 4.3.4. Stonier’s Alternative Information Physics

One of the ontological proposals with the greatest subsequent impact and most conceptually radical within the informationalism of the nineties is undoubtedly the one developed by the German biologist Tom Stonier (1927-1999) during the nineties. Unlike the metaphysical theses developed by Wheeler (1990), Landauer (1991, 1999) or Davies (1990) (all representatives of 'orthodox' information physics), Stonier's proposal was systematically articulated in the eighties-nineties with a clear speculative flavour and significantly detached from the orthodox institutional-intellectual circuits of this discipline (Section 4.2). In fact, his conceptual architecture and interpretative strategy is more akin to the informationalism in the fifties than in the nineties (i.e., he disregards the algorithmic information K and computational turn in the eighties), typical of a Neo-Brillouinian proposal with the intellectual pretensions of this historical stage. Apart from his '*Beyond Information: The Natural History of Intelligence*' (1992) and '*Information and Meaning: An Evolutionary Perspective*' (1997), his key work is unquestionably the celebrated '*Information and the Internal Structure of the Universe*' (1990), whose dual purpose is to: "First, to examine the proposition that "information" is as much a part of the physical universe as is matter and energy, and (…) Second, to create a foundation on which to develop a general theory of information [and] explore the possibility of a new branch of physics -information physics" (Stonier 1990, p.2).

As for the pretension of developing a general theory of information (in the direction of MacKay, Section 3.2.1), Stonier concluded that a general IT providing 'information' with the same fundamental ontological level as 'energy' or 'matter' should constitute a branch of physics. In this sense, the author seems to ignore or disregard the disciplinary advances made in that same year i.e., 1990 by the physics community with the foundational manifesto of information physics (ed. Zurek 1990), Section 4.2. Interestingly, Stonier's (1990) information physics distanced from institutional or orthodox information physics in (i) its fundamentally thermo-classical framework (as opposed to the fundamentally gravitational-quantum one of Wheeler [1990] or Zurek [1990]), (ii) its explicit goal of integrating theoretically a semantic-epistemic information notion, and (iii) the disregard of algorithmic information K, putting the interpretive focus on $H_{IT}$. Interestingly, Stonier (1990, p.56) intends to erect his proposal of both a physical and semantic-epistemic 'agent-centric[88]' information concept by adopting a Neo-Brillouinian interpretative strategy centred on the notion of 'bound information' $I_{Bb}$ (Section 3.2): "Brillouin's concept [$I_{Bb}$] was designed to help overcome an anomaly in Shannon's theory: the more random is the arrangement of symbols – ie, the higher the entropy [$S_B$] – the greater the information [$H_{IT}$] content. Taken to its logical extreme this would mean that pure noise, which contain the greatest amount of entropy, would contain the greatest amount of information" (Stonier 1990, p.56). This 'anomaly' is precisely the so-called Carnap-Bar-Hillel Paradox (Floridi 2011, Chapt. 3), which naturally emerges from interpretatively combining[89] a non-semantic information $H_{IT}$ and a semantic information concept OI (as derivable from $I_{Bb}$): "The phenomenon we call "meaning" involves a gradient of relationship between physical information and mental interpretations" (p.21).

[88] Although agent-based, for Stonier the content of his information concept is independent from particular agents: "In contrast to physical information, there exists human information which includes the information created, interpreted, organized or transmitted by human beings (…) Patterns of information, at whatever level of complexity, need sensors to be perceived, and "intelligence" to be analysed and processed" (Stonier 1990, p.17-18).

[89] Equivalently, Carnap-Bar-Hillel Paradox which naturally emerges from interpretatively combining both an $H_{IT}$-based identificationist and an OI-based complementarist conception of the relation information-entropy (Chapter III) "However, Shannon's use of entropy as a metaphor is unfortunate. A physical relationship between information and entropy actually does exist. However, it is neither the direct relationship [identificationist] envisaged by Shannon, nor its negative [complementarist] as envisaged by Brillouin. (…) physical information relates to order, and as Schrödinger has shown, order varies inversely with Boltzmann's thermodynamic probability function [$S_B$]" (p.57).

Following a Brillouinian proposal, the first step of the interpretative strategy underlying Stonier's physics of information is to reconceptualize his broad notion of information $I_{Bb}$ as a measure of the degree of structural organization (exploiting Schrödinger's [1944] proposal) of a multi-component system. As a direct consequence, its complementary notion of entropy $S_B$ should be interpreted as a measure of the degree of structural disorganization. The main lines of such an interpretative strategy were described by the author into three basic principles (Stonier 1990, p.26):

> "1. All organized structures contain information, and as a corollary: No organized structure can exist without containing some form of information.
>
> 2. The addition of information to a system manifests itself by causing a system to become more organized, or reorganized.
>
> 3. An organized system has the capacity to release or convey information"

The next step after interpretatively characterizing information as a measure of organization is to define its conceptual relation to the notion of 'energy', which is developed by establishing parallels between the two. The skeleton of the argument is that if information is conceptually connected to the physically meaningful notion of energy, then it will also possess a physical dimension: "Energy is defined as the capacity to perform work. Information is defined as the capacity to organize a system -or to maintain it in an organized state (…) organization is the manifestation of information interacting with matter and energy (…) Information is a quantity which may be altered from one form to another. Information is a quantity which may be transferred from one system to another (…) Information, like energy, is an abstract quantity (…) information, like energy, also possesses a physical reality" (p.26-27). As can be seen in this quote, Stonier follows a Brillouinian interpretative line by articulating the conceptual information-energy connection on the basis of the concept 'bound information' $I_{Bb}$, whose content is both epistemic OI-like and meaningful TD-like. Thus the information about the system can degrade as its degree of disorganization (loss of structure among its components) increases, just as the energy possessed by the system can also degrade as the system's capacity to perform work decreases. It is precisely in this degradation of energy (generating Q in the context of TD) that Stoner (1990) introduces the third and last concept of his ontological triad: "Heat is the product of energy interacting with matter. Structure represents the product of information interacting with matter" (p.74).

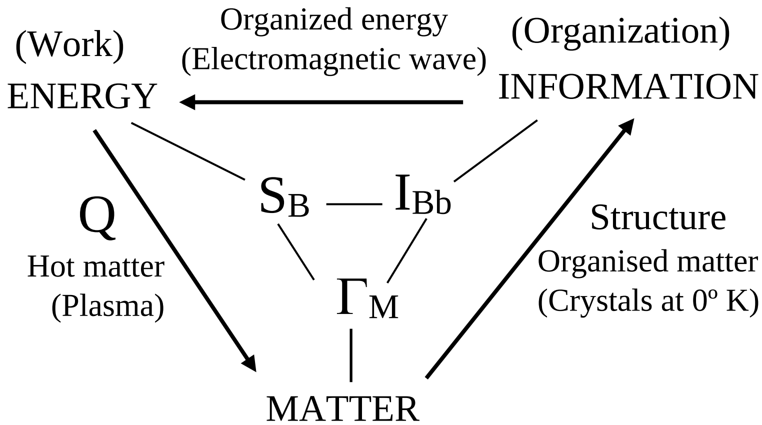


**Figure 5.** Stoner's (1990, Fig.6.1) Ontological Triad.

As this conceptual diagram in Figure 5. shows, the theoretical core of Stoner's proposal is based on the ontological triad 'information', 'energy' and 'matter' as fundamental pillars of physical reality: "INFORMATION and ENERGY must not be viewed as the opposites of a bipolar system, rather, they must be considered as the two angles of a triangle, with MATTER comprising the third. Such a conceptual model would define the boundaries of our physical universe" (p.75). To illustrate the main (binary) conceptual connections within the core of this model (Figure 5.), Stoner unfolded the following (visually suggestive) state-of-matter cases studies: "(1) A mixture of pure energy and matter, lacking information, would comprise a plasma of fundamental particles. (2) A mixture of matter and pure information, lacking energy, is exemplified by a crystal at 0 K. (3) A mixture of information and energy, lacking matter, would consist of massless particles such as photons travelling through space devoid of matter" (ibid).

Therefore, Stonier intended to develop (whether he succeeded is a different issue) in 1990 a conceptually systematic and interpretatively coherent information physics as an epistemically fruitful discipline, intellectually independent of that presented by Zurek (eds. 1990), Section 4.2 "Information physics will become an established and fruitful science when the material presented in this book has been properly quantified on the one hand, and experimentally or observationally verified on the other (…) Any general theory of information must begin by studying the physical properties of information as they manifest themselves in the universe" (p.113). Finally, Stonier (1990, p.114) unfolded the two axioms (and a central conclusion derived from them) of information physics as a general theory of information:

> "[A1] Information is a basic property of the universe.
>
> [A2] The information contained by a system is a function of the linkages binding simpler, into more complex units.
>
> [C] The universe is organized into a hierarchy of information levels"

### 4.3.5. Rise of Pan-Informational Metaphysics in the 2000s

Interestingly, these informational ontological proposals in the 1990s (led by Wheeler's [1990] 'If from the Bit' and Landauer's [1991] 'information is physical') had a decisive impact within the physicist community. One of the main intellectual consequences of this ontological turn of this current of scientific thought was the fertilization of the intellectual soil for the emergence of informational metaphysics among the next generation of informational physicists (e.g., Lloyd, 2006; Vedral, 2010) and popularisers (e.g., Seife, 2006; Gleick, 2010) at the turn of the millennium. This was precisely the case of Seth Lloyd, one of Wheeler's disciples and main promoters of this current in the nineties, who sought to develop theoretically the suggestive ontological proposal of his teacher (Section 4.3.1) in such a way that physical reality itself is characterized as a quantum computer processing on the order of $10^{90}$ bits (Lloyd 2006). This ambitious position was systematically stated in his 2006 *Programming the Universe*: "The universe is the biggest thing there is and the bit is the smallest possible chunk of information. The universe is made of bits. Every molecule, atom, and elementary particle register bits of information. Every interaction between those pieces of the universe processes that information by altering those bits (…) The universe is quantum computer." (Lloyd 2006, p.3). Another example in this direction was the proposal of Vlatko Vedral (2010), who influenced by Landauer's proposal[90] sought to place information as the ontological foundation of physical reality: "before we can even speak about why things are connected, we need to ask ourselves why things exist in the first place. I will argue (…) that the notion of 'information' gives us the answer to both questions. Curiously, this makes information a far more fundamental quantity in the Universe that matter or energy" (Vedral 2010, p.2). Both Lloyd's (2006) and Vedral's (2010) informational metaphysics share a strong computational and quantum character[91] (i.e., focused on discrete units of physical information), which is why they are often referred to in the literature as 'digital ontologies' (Floridi 2011, Section 13) as opposed to 'analog ontologies' based on continuous units.

…

> "Every particle in the universe, every electron, every atom, every particle not yet discovered, is packed with information […] that can be transferred, processed, and dissipated. Each star in the universe, each one of the countless galaxies in the heavens, is packed full of information, information that can escape and travel. That

[90] "The three worlds that I read back in autumn 1994, which changed my perspective so markedly, were 'Information is physical'. The three words, in this order (…) made me realize that indeed maybe information is the answer" (Vedral 2010, p.3).

[91] "The conventional view is that the universe is nothing but elementary particles. That is true, but it is equally true that the universe is nothing but bits – or rather, nothing but qubits. Mindful that if it walks like a duck and it quacks like a duck then it's a duck (…) since the universe registers and processes information like a quantum computer, and is observationally indistinguishable from a quantum computer, then it is a quantum computer." (Lloyd 2006, p. 154).

information is always flowing, moving from place to place, spreading throughout the cosmos." (Seife 2006, p.3)

## 4. 4. Philosophical Developments and Criticisms of Informational Ontologies

One of the main intellectual manifestations of the radicalization of physical informationalism in the nineties was the emergence of various informational ontologies, placing information in a central place in the picture of reality provided by fundamental physics. Despite the heterogeneity of these metaphysical proposals, the enormous impact of this intellectual current resulted in the popularization within (Wheeler, 1990; Stonier, 1990) and outside (e.g., Seife 2007) the scientific community of the widespread belief that physical reality is (in one way or another) composed of information as a sort of fundamental entity. Interestingly, these kinds of ontological conclusions historically derived from the informationalist tradition also gained interest from the philosophical community since the mid-2000s. Some authors, notably James Ladyman and Don Ross (2007), sought to develop a fine-grained informational ontological proposal. Another group of authors, such as Christopher Timpson (2004, 2013) and Olimpia Lombardi (et al. 2016a, 2016b), had criticized inflationary attitudes towards information. All of them, despite the difference between their positions, unanimously criticized that most ontological-informational theses (Section 4.3) are articulated on the basis of inconsistent conceptual resources and incoherent interpretations.

### 4.4.1. Ladyman and Ross Information-Structural Realism

James Ladyman has been not only instrumental in the philosophical defense of Landauer's principle and computational thermophysics (see Footnote 86), but also placed informational ontologies at the center of philosophical discussions. Together with Don Ross in his 2007 *Every Thing Must Go. Metaphysics Naturalized*, both authors aimed to develop a naturalistically sophisticated and philosophically-robust ontological proposal that exploited elements of physical informationalism while being critical of the coherence of the 'philosophically naïve' ontologies seen in Section 4.3: "we return to the claim that the world is made of information and distinguish between two interpretations: (i) The world is made of a new substance or substantive particular called information 'infostuff'. (ii) The world is not made of anything and information is a fundamental concept for understanding the objective modality of the world. We reject (i) as a further example of domesticating metaphysics" (Ladyman and Ross 2007, p.189).

By 'domesticated metaphysics', the authors refer to that type of ontological frameworks based on affirming that reality is composed of fundamental entities that play the role of 'building blocks' (i.e., from whose combination new ontological levels are generated). Undoubtedly, the informational ontology of Stonier (1990) (as well as the rest of the proposals in Section 4.3) constitutes a paradigmatic case of domesticated metaphysics, since physical reality can be derived from the combination of three fundamental building blocks, i.e., energy, matter and above all information. Against the strategy of theoretically domesticating metaphysics via entities, Ladyman and Ross (2007, Chapter 2) defend a perspective known as 'structural realism' in philosophy of science, where our most empirically satisfactory physical theories (i.e., quantum field theory) suggest to us that reality is fundamentally composed not of entities but of a certain kind of ontologically prior relations. Without entering into the debate on scientific realism, it should be mentioned that there are different types of structural realism depending on the type of structures considered. For the ontological proposal of Ladyman and Ross, the fundamental structures that would characterize a certain real domain are certain informational structures that allow us to grasp the regularity-driven or modal content[92] of the fundamental physical reality that underlies that domain (even when this corresponds to a special-science domain): “We may entertain the endorsement of information-theoretic fundamentalism in the following deflationary sense: ‘it is impossible to distinguish operationally in any way reality and information (…) the notion of the two being distinct should be abandoned” (ibid, p.189).

Despite this 'IT-fundamentalism' advocated by Ladyman and Ross, the concept of information on which they built their proposal was not (as might be expected) $H_{IT}$, but properly K because of its ability to directly quantify individual objects. In particular, the authors conceptually assimilated K through the notion of 'logical depth' (Bennett 1987) (Section 4.1.2) “To obtain an objective measure of informational content in the abstract (that is, [non-TD]) sense, one must appeal to facts about algorithmic compressibility [K] (…) The important measure for our purpose will be logical depth. This is a property of structural models of real patterns” (ibid, p.220). With the expression 'real patterns', Ladyman and Ross refer to the classic paper by Daniel Dennett (1991), where from Conway's Game of Life they describe the 'macro' or 'observable' emergence of patterns (structure of cells with relatively autonomous behavior) solely as a result of the interaction at the local 'micro' level of individual cells. Ladyman and Ross (2007, p.227-231) reconceptualized Dennett's (1991) philosophical proposal by means of the informational concept of logical depth such that 'real patterns' allow compressing a vast amount of redundant information in a way that "must be required if some counterfactual-generalization-supporting information is not lost" (Ladyman and Ross 2007, p.231). That is, the logical-depth information of a particular pattern somewhere in the cell-bits map (i.e., analogous to an equal-cell partitioned Γ) is equivalent to the information content of that same pattern anywhere else in the very cell-bits map, which allows compressing all the redundant information of its encompassing individual cells via some (algorithmically conceived) theoretical/model-based description. In this sense, as can be seen, the informational ontology

[92] Ladyman and Ross relied on Maudlin (2002a) to support the model character of information “Tim Maudlin shows that information is modal by arguing that information can be transmitted (2002a). The point is easily established by the following example: Bob tells Alice that if he gets spin up he will phone her at 6pm, and if he gets spin down the he won’t. If Bob doesn’t call then Alice gets information about the measurement without energy or matter being transmitted. As Maudlin says, in such a case ‘information is transmitted only because of the *model* structure of the sition (2002a, p.163, his emphasis). Alice only knows the result of the measurement because of what would have happened had the result been otherwise” (Ladyman and Ross 2007, p.188-189).

proposal of Ladyman and Ross in 2007 constitutes a conceptual refinement and interpretative reinforcement with respect to the informational ontology of Davies in 1990 (Section 4.3), both theoretically based on the AIT notion of 'logical depth'.

After the concept of information K and that of compressibility, the informational ontology of Ladyman and Ross is also built on a technical notion of 'projection'. A given pattern expressed by a structure x (e.g., a cell-ship in Conway's Game of Life) is projectable onto another structure y of lesser 'logical depth' (Section 5.2), if x → y, wherein the informational-modal content of x could be equivalently transferable to the content of y: “Projection is a particular kind of ‘information flow’ (…) we are interested instead in properties of information structures that warrant inference to real patterns. ‘Projectability’ is the concept of information-carrying possibility -applied now not to channels but to models of real patterns and ultimately to real patterns themselves” (Ladyman and Ross 2007, p.224). From a physicalist position, Ladyman and Ross (2007) propose that the most empirically successful fundamental physical theories (i.e., quantum field theory) allow us to model the structures of higher logical depth corresponding to 'real patterns', whose informational content is projected onto other structures of lower logical depth distributed by the special sciences. That is, the models of particular sciences are precisely those that contain the least amount of redundant information among all sciences, where the models of fundamental physics have a high degree of informational redundancy (whose patterns can also be characterized as ‘real patterns’ as long as they are non-reductively[93] projectable to those of the particular sciences).

Once these conceptual elements are in place, Ladyman and Ross (2007, p.233) unfolded the core of their physical-informational ontology called 'rainforest realism' through the following definition of real pattern “To be is to be a real pattern; and a pattern x → y is real iff (i) it is projectible; and (ii) it has a model that carries information about at least one pattern P in an encoding that has logical depth less than the bit-map encoding of P, and where P is not projectible by a physically possible device computing information about another real pattern of lower logical depth than x → y” (p.233). Thus, from the rainforest realism developed by Ladyman and Ross (2007, 2013), the physical reality should be ontologically identified with the set of non-redundant data[94] (statistical information) provided by our most epistemically successful fundamental physical theories.

[93] In this respect, Ladyman and Ross (2007, p.33) defended that “There are various possible ways in which real patterns at different scales can carry information about one another without inter-reducing”

[94] In this regard Ladyman, Ross and Kincaid (2013) further clarified this position in a later contribution entitled ‘The World in the Data’, wherein they defended that “The fundamental empirical structure of the world is not mathematical but statistical. And there is no such thing as purely formal statistics. The ‘principles’ of statistics are simply whatever dynamics emerge from our collective exploration of, and discovery of patterns in, data. Applied scientists are continually reminded of this in their day-to-day work, especially as our new technological capacity at last realizes the potential for limitlessly evolving of co-adaptive growth of new techniques for discerning real patterns and new demand for statistical innovation arising from the discovery of new possible patterns. (...) It is made more, not less, persuasive by the grandest discovery of the twentieth century, that fundamental physics, and therefore reality itself, are irreducibly statistical. What is the world? It is the endless weave of patterns to be extracted from noise, at an endless proliferation of mutually constraining scales, that we will go on uncovering forever as long as we have the collective and institutional courage that comes from love of objective knowledge, the great moral core of the Enlightenment.”

4.4.2. Timpson against Informational Reification

Apart from the informational ontology proposal of Ladyman and Ross (2007) within the field of philosophy of science, the problem of the ontological nature of information also aroused interest in other philosophical fields. This was the case of the information philosopher Luciano Floridi (2011, p.11), who not only defended the theoretical neutrality of information concepts (i.e., $H_{IT}$ is not a classical physical concept) but also called the question of the ontological nature of information as 'Wiener's problem', precisely because he was a pioneer author[95] (Section 3.1.1) in raising this issue "Information is information, not matter or energy." (Wiener 1948, p.132).

One of the authors who analysed more deeply Wiener's problem in fundamental physics was Christopher Timpson (2004, 2013), focusing not on the particular field of thermal physics (as is our case) but on quantum mechanics and quantum information theory. The philosophical aim of his 2013 *Quantum Information Theory and the Foundations of Quantum Mechanics* was to assess the conceptual consistency and interpretive coherence underlying certain widespread uses within the physics community of informational notions. Although in this thesis I focus on thermophysical informationalism, much of the philosophical assessment to be carried out by Timpson (2013) on what we might call 'quantum-physical informationalism' could be hugely relevant to our domain-restricted task as well. This quantum-physical informationalism (historically emerged in the 1990s, Section 4.2) consists in applying informational concepts (e.g., explicitly $H_{IT}$ or quantum-$H_{IT}$ and implicitly OI, among others) to interpret quantum formalism or to explain the functioning of quantum teleportation. Regarding the latter field of application, some authors (in the bandwagon of informational physics of the 1990s) such as Penrose developed an informational interpretation of the famous phenomenon of quantum teleportation between Alice and Bob as if it were a continuous transport of an informational substance: "How is it that the *continuous* 'information' of the spin direction of the state that she [Alice] wishes to transmit (...) can be transmitted to Bob when she actually sends him only two bits of discrete information?" (Penrose 1998, p.1928).

According to Timpson (2013, p.95), these types of quantum-informational interpretation of quantum teleportation processes such as Penrose's (1998) constitute inconsistent uses of the notion of $H_{IT}$ information: "Puzzles arise when we feel the need to tell a story about how something travels from Alice to Bob in teleportation. In particular, it might be felt that this something needs to travel in a spatial-temporally continuous fashion (...) But if the [$H_{IT}$] doesn't pick out a particular, then there is nothing to take a path, continuous or not, therefore the problem is not a

[95] According to Bynum (2008), Wiener was a historical pioneer of the informational metaphysics of the nineties "Near the end of the Second World War, Wiener had a crucial new insight: he realized that *information is physical* since it is governed by the second law of thermodynamics. Information, he saw, plays a significant role in every physical entity and process. Delighted by this discovery, he walked the halls of his building at MIT, in one of his famous "Wiener walks" (Conway and Siegelman, 2005), telling everyone he met that "information is entropy" (…). This was an important new addition to the foundations that he and others were building of the new age that he envisioned" (Bynum 2008, p.204).

genuine one, but an illusion." (Timpson 2013, p.95). Timpson's point here is precisely that underlying this line of interpretation is an IT-based reconceptualization of the quantum teleportation process such that $H_{IT}$-encoded information in a quantum state of a system controlled by 'Alice' is transported through a continuous spatiotemporal channel to the quantum state of another system (quantum entangled with the first) controlled by the recipient 'Bob'. However, this informational reconceptualization incurs a misinterpretation of the $H_{IT}$ concept for the following reasons. On the one hand, the attempt to physically characterize the communicative channel as a spatiotemporal channel between the quantum systems of Alice and Bob contradicts the lack of spatiality of the 'communicative channel' posited by Shannon (1948), Section 2.3. On the other hand, that a quantum state can be exploited to encode information $H_{IT}$ (e.g., Schumacher 1995, Feynman 1996) does not imply that there is an 'informational' physical entity that transports along a continuous path from Alice to Bob, so assuming the latter is interpretatively incoherent.

For Timpson (2013) this type of interpretative misuse of informational concepts in quantum mechanics is only a particular case of the inconsistency that underlies any ontologically inflationary position regarding information, such as those paradigmatically defended from the physical informationalism of the nineties (Wheeler, 1990; Stonier, 1990; Landauer, 1991, 1999), Section 4.3. According to this author, the conceptual inconsistency of these ontologically inflationary proposals could be satisfactorily explained as categorical errors: "[inflationary informational ontologies] simply involves a category mistake. Pieces of information, quantum or classical, are abstract types. They are not physical, it is rather their tokens which are" (Timpson 2013, p.69). In order to evaluate the logical grounds of this categorical mistake, Timpson (2013, p.16) introduces the distinction (somewhat external to Shannon's IT [see Lombardi et al. 2016b]) between (i) the 'bits' of classical or quantum-encoded $H_{IT}$ information, which are the amount of information $H_{IT}$ produced by the source S, and (ii) the 'pieces' of classical or quantum information, which correspond to the particular objects (e.g., sequences of symbols, quantum states, etc.) generated by the source S. It is precisely in the second case where the conceptual distinction (proper to the philosophical domain) between types and tokens would apply, stating that the pieces of information produced by the source could not be consistently characterized as 'tokens' but properly as 'types', and therefore it would be a categorical error to characterize types[96] as physical entities (a property applicable only to particular tokens) and not as mere general abstracta:

> "The abstract nature of pieces of [$H_{IT}$] and of quantity of information proved important when we came to consider the question of the flow of [$H_{IT}$]. [$H_{IT}$], whether piece or quantity, is not a kind of stuff that flows around, no matter how aethereal. This is the wrong picture; and one stemming from a logical mistake: thinking of abstract items as oddly nebulous concrete ones. Even with this mistake corrected, I argued that one can't interpret [$H_{IT}$] as telling us how [$H_{IT}$] as a quantity might flow, analogous to a flow of energy, for example; for [$H_{IT}$] as a quantity in the Shannon theory does not play a part in the occurrent characterization of the properties of systems (save of a source): how

[96] "types are abstracta. They are not themselves part of the contents of the material world, nor do they have a spatio-temporal location." (Timpson 2013, p.27).

> compressible the output of a source is, or how much a particular message can be compressed, are apparently not quantities it makes sense to think of as flowing around as energy or heat" (p.43)

It is precisely on the basis of this conceptual distinction that Timpson (2013, section 2) intended to explain the emergence of inflationary ontological attitudes about information (e.g., Wheeler's [1990] "It from the Bit" or Landauer's [1991]) through an entangled progressive historical-cognitive process of nominalization of the original verb 'to inform' i.e., to provide knowledge to an agent. Precisely from this historical-cognitive nominalization of the verb 'informational' would progressively derive the singular noun 'information', which would eventually allow to fall into the category error between types and tokens that underlies the proposals of ontological inflation of information. In other words, by nominalizing the verb 'to inform', not only does one obtain the singular name 'information' for use within a scientific community, but also its conceptual users would eventually cease to understand the content of this $H_{IT}$ notion as an (abstract) action, hypostatizing or 'reifying' (using Ladyman and Ross's terms [2007, p. 189]) its content and implicitly reconceptualizing information as a (concrete) fluid-like substance.

### 4.4.3. Lombardi against Informational Physicalism

This deflationary program of Timpson (2013) with respect to informational ontologies in physics has recently been criticized by quantum physics philosopher Olimpia Lombardi (et al. 2016a) in her 'Deflating the deflationary view on information' (Lombardi et al. 2016a). In addition to pointing out that Timpson's distinction between bit-quantities of information $H_{IT}$ and pieces of information is completely alien to Shannon's (1948) original theoretical framework IT (Section 2.3), Lombardi et al. (2016a) directed their critical charge at the argumentative methodology underlying the Timpsonian evaluation of inflationary ontologies[97]. According to these authors, Timpson's (2013) analytical argument depends on assessing the ontological implications derived from the grammatical category of the term 'information' employed by proponents of informational metaphysical frameworks that emerged in the 1990s (Section 4.3). However, Lombardi et al. (2016a) argue that proponents of informational ontologies such as Wheeler (1990) or Landauer (1991) in no way rely interpretively on the linguistic functions of informational terms (this is a strategy more typical of what Ladyman and Ross [2007] call 'domesticated metaphysics'), but on the function that these concepts play within scientific practices.

[97] The core of Lombardi et al. (2016a) criticisms was synthetized as follows "It is true that Timpson distinguishes between the everyday notion and the technical notion of information. Nevertheless, in both cases the strategy is the same: to analyse the grammatical role played by the word 'information' in the non-formal language, and to draw ontological conclusions from that analysis. However, physicists do not appeal to that strategy to decide what a physical item is when they say, as Rolf Landauer (1991, 1996), that information is physical. If one does not want to turn the structure of non-formal languages into the key witness about what exists and does not exist in the physical world, a more reasonable strategy seems to be admitting that physics supplies us the better tools to know what the physical world is. Therefore, in order to decide whether or not a certain item belongs to the physical world, we should see what role it plays in physical science" (Lombardi et al. 2016a, p.224).

Regarding the constructive part of their proposal, Lombardi et al. (2016a) argued that the evaluation of the content or ontological nature of the particular concept of $H_{IT}$ information should be analogous to the strategy (in no case based on the analysis of grammatical or linguistic categories à la Timpson [2013]) by which he evaluates other similar concepts in the history of physics, such as those of field and energy: "These discussions [on Wiener's problem] are philosophically interesting when one admits that it is physics and not grammar that is the best clue for discovering the content of the physical world. It does not matter what kinds of words are used to refer to properties, such as charge and mass, and to name items that tended to substantialization through the history of science, such as fields and energy" (Lombardi et al. 2016b, p.227). In the case of the concept of energy, various authors (including professional scientists, as is the case of Stonier) throughout history have made the mistake of referring to energy as a kind of fluid substance that bodies exchange, i.e., "the two balls exchanged energy" (Stonier 1990, p.81). Although this implicit reification of certain uses of the concept of energy is interpretatively incoherent or allows us to incur in categorical errors, the fact is that (against Timpson [2013] deflationary proposal) one cannot reject that energy should be ontologically included as part of our physical reality. It is precisely the role played by the concept of energy within certain scientific contexts what should ultimately constitute the criterion to determine its ontological nature or not. And likewise, it is the role played by the concept of information $H_{IT}$ in our physical theories that would properly fix its physical nature.

Lombardi et al. (2016a) finally conclude by arguing that, since certain fundamental physical theories postulate that $H_{IT}$ information is part of our physical reality (Section 4.3), then there is no reason to think that $H_{IT}$ information can have a physical nature: "What only matters is that all those items [energy, fields, and information] inhabit the world of physics, that is, according to physics they are part of the furniture of the world. And this implies that *contemporary physics offers no grounds to deny the possibility of a non-trivial and meaningful physical interpretation of the concept of information*" (ibid, p.227, Italics are mine). Therefore, if there is a relevant interpretation of a concept of information within a physical theory, then that to which this concept of information refers could belong to physical reality. Thus, Lombardi et al. (2016a, 2016b) finally defend a pluralistic position regarding the possible interpretations of informational concepts: an information concept will have physical content if there is a non-trivial physically meaningful interpretation of it.

# Chapter V

## Conceptual Foundations of Information Physics

Now that we have reached this point in this journey, let us look back at what we have covered so far. During the two previous chapters I have analysed the historical unfolding of the current of scientific thought that is called 'information physics' (in particular, 'informational thermal physics'), from the start of Shannon's bandwagon in the early 1950s to the radicalization of this intellectual trend in the 1990s. As I pointed out at the beginning, my historical evaluation has not been a historiographic task centred on the description of a succession of events. Such an evaluation has been properly conceptual, delimiting: (i) the theoretical motivations, (ii) the interpretative and epistemic strategies, (iii) their main criticisms in the literature, and (iv) and the intellectual pretensions underlying the historical evolution of the uses of informational notions within the domain of classical statistical thermophysics during the second half of the twentieth century. In this sense, the previous historical-conceptual analysis has been the diachronic part of my philosophical evaluation of this decisive current of physical thought. Now it is time to reap the philosophical fruits that we have been sowing in the previous chapters.

My task in this chapter is to develop the synchronic part of my philosophical assessment, devoted essentially to a critical analysis of the conceptual foundations underlying the use of informational notions in this physical domain and historical period. With this I intend to complement (avoiding argumentative redundancy) the philosophical results obtained from the previous historical-conceptual analysis, evaluating how the use of various informational concepts would allow the generation of knowledge about thermal phenomena or would make it possible to adopt realistic attitudes towards the physical nature of these informational properties. To this end, I will develop (on the one initially outlined in Chapter II) a 'minimalist' framework of conceptual evaluation, designed expressly for the arduous philosophical task that concerns us here and not to position ourselves in debates on semantics or epistemology of scientific concepts. Finally, this project is presented as a direct heir to the critical analyses of Rudolf Carnap (and his intellectual milieu) (Section 3.4), John D. Norton, James Ladyman or Christopher Timpson (Section 4.4) and many other philosophers who have contributed to the clarification of certain local issues in this field.

## 5.1. Assessing Epistemic Contributions of Informational Concepts

In this section I briefly delimit the methodological framework I will use in this chapter to analyse the epistemic contribution of informational concepts in thermophysics. To do so, I assume as background the ‘minimalist’ representational-inferential approach to concepts advocated in Chapter II (i.e., understanding concepts as representational resources used to infer about target phenomena), assessing how the exploitation of their (i) syntactic properties, (ii) semantic content and (iii) pragmatic dimension contributes to gaining knowledge about the domain of investigation. Additionally, I assume a Lakatosian perspective[98], but exclusively in the sense that I will adopt as a unit of epistemic analysis not an individual thermoinformational theoretical proposal $T_i$ (for instance, Brillouin 1951) but a diachronic unfolding of these proposals $P = T_1, T_2, ..., T_N$ (e.g., the Brillouinian or Jaynesian tradition, see Section 3.2) thus characterizing them as a 'research program' (Lakatos 1978) or even 'research tradition'[99] (Laudan 1977).

Thus, although necessary for the unfolding of this argument, my ultimate aim is not to evaluate the epistemic contribution of the use of informational concepts within a single synchronic proposal, for example within Brillouinism (Section 3.2.) or Jaynesianism (Section 3.4). My claim is precisely to analyse how the historical unfolding of uses of informational concepts within the thermo-informational program (Chapters III and IV) has evolved diachronically in terms of epistemic contribution to the field of thermophysics. That is, whether historical changes in descriptive tactics within this tradition (e.g., the use of algorithmic information since the 1980s) have made a real epistemic advance in the way in which the thermo-informational program aims to describe, explain, predict phenomena within this physical domain and vis-à-vis alternative programs such as non-informational thermophysics. As will be remarked in Chapter VI, by 'advance' or 'epistemic progress' I mean nothing other than the ability of a proposal $T_1$ to predict, explain and/or interpret known phenomena (e.g., the adiabatic demagnetization of an ideal gas) or solve conceptual new puzzles (e.g., framing the observer's loss of knowledge during a thermal process) more successfully than a predecessor $T_0$ proposal within the program.

[98] It should be noted at this point in this analysis that, although I provisionally adopt Lakatos' research programs as diachronic units of epistemic analysis, I will not subscribe during this chapter to his method of scientific research programs (Lakatos 1978, Section 1.3) and the positive-negative heuristics dynamics as a framework for describing the historical dynamics of theories. The main difference is that our focus is not on assessing the rationally diachronically unfolded by research programs progress (theoretically or empirically) or come to be refuted by other programs, but properly on analysing how the various conceptual usages underlying such programs contribute to interpreting, explaining and predicting phenomena within the research domain.

[99] Nikles (2017, Section 1.5) shows how the main difference between Lakatos' (1978) research programs and Laudan's (1977) research traditions is that the former somehow presuppose that scientific methodology remains fixed throughout its historical evolution, whereas for the latter this methodological fixity is a purely contingent fact. In this sense, my analysis of informational thermophysics suggests a certain incompatibility with the discontinuist proposals of the Kuhnian paradigms (Kuhn 1962) or the methodological anarchism of Feyerabend (1975), mainly due to the historical continuity of the uses of informational concepts that I have detailed in Chapter III, & IV.

Although one can speak of different thermo-informational traditions (i.e., Brillouinism, Jaynesianism, etc.) throughout history, it can be interesting to refer to a single thermoinformational research program constitutively characterized by the explicit and systematic use of informational concepts. Thus, what Lakatos (1978, p.48) called the 'hard core' of a research program[100] would correspond in the particular case of the thermoinformational program to the descriptive exploitation of connections between informational and entropic (and/or probabilistic) concepts. Interestingly, the information-entropy(-probability) connections as the core of the thermoinformational program are not only the elements on which the predictive and explanatory capacity of each proposal rests, but they are also elements that aim to provide consistency to the whole program[101]. For the Brillouinian tradition (Section 3.2), that the negentropy of the system corresponds to the negative information of the agent is simply a fundamental or primitive principle of its proposal (NPI), whose content cannot be tested in the laboratory and must be defended from possible objections of conceptual inconsistencies and interpretative incoherencies[102].

My next task is precisely to evaluate the way in which this exploitation of informational concepts within informational thermophysics contributes epistemically to the generation of successful explanations and predictions. But before proceeding with such a task, let me illustrate how I will proceed with this epistemic analysis with one of the last historical links of the thermoinformational research program. This is precisely the case of the intellectual work of the Israeli physicist Arieh Ben-Naim (2008a, 2008b, 2010, 2015), undoubtedly one of the greatest exponents of informational thermo-physics in recent years in terms of the radicality of his interpretative proposal and his constant effort to popularize this current of thought within and outside the physics community. A quite significant sample of his work can be found in his *A Farewell to Entropy. Statistical Thermodynamics Based on Information* (Ben-Naim 2008), whose title makes explicit his central intellectual goal: namely, to redefine the foundations of thermophysics on informational conceptual architectures and to eliminate all entropic notions from this physical domain.

[100] As an illustration, Lakatos mentioned Newton's law as the core of the Newtonian research program in classical mechanics and gravitation "In Newton's programme the negative heuristic bids us to divert the *modus tollens* from Newton's three laws of dynamics and his law of gravitation. This 'core' is 'irrefutable' by the methodological decision of its proponents: anomalies must lead to changes only in the 'protective' belt of auxiliary, 'observational' hypotheses and initial conditions." (Lakatos 1978, p.48).

[101] It should be noted I are not addressing (at all) here the question of what it takes for a thermophysical theoretical proposal to be verified or refuted.

[102] On the entropy-information connections as the hard core of the thermoinformational program, complex conceptual architectures and interpretative scaffolds (the 'protective belt' in Lakatosian terms [1978, p.49-50]) are erected to (i) provide consistency to this 'nuclear connection' in the face of certain conceptual problems and (ii) allow its applicability in concrete models. Illustratively, the conceptual identity between $S_G$ and $H_{IT}$ as the hard core of the Jaynesian program historically faced the problem of not being directly applicable to individual physical systems. A possible solution to this problem was historically provided by the proposal of Caves (1990), who anchored the $S_G$-$H_{IT}$ identity to the conceptual ability of (epistemically interpreted) algorithmic information K to refer to individual objects, thus reconstructing a new architecture around the core of the Jaynesian program (Section 4.2.3). The evolution of such 'protective-belt-like' conceptual architectures and interpretive scaffolds would not only serve the function of reinforcing the core of the thermoinformational program but also allow it to be conceptually exploited in explanations and predictions within particular application scenarios.

Undoubtedly, this proposal will be paradigmatic of the most recent stage in which the thermo-informational program finds itself.

In line with earlier thermo-informational authors, Ben-Naim sought to exploit descriptively informational concepts to explain no less than the Second Law of TD "Using information as a fundamental concept makes the understanding of the Second Law much easier. It also removes the mystery that has befogged entropy and the Second Law for a long time (...) [Missing information] is the best way of describing what is the quantity that changes in a spontaneous process" (Ben-Naim 2008, p.251). Another primary intellectual task Ben-Naim intended to accomplish in *A Farewell to Entropy* is the use of informational notions to interpret the content of entropy concepts, which (should such a strategy prove epistemically fruitful) would enable him to satisfy his intellectual goal of showing the reducibility of entropy within thermophysics "the concept of missing information not only contributes to our understanding of what is the thing that changes and which is entropy, but it also brings us closer to the last and final step in understanding entropy's behavior" (ibid, p.15). Apart from this theoretical goal, Ben-Naim also defended an informationalist ontological framework in the traditional line of Wheeler (Section 4.3): "This will not only remove much of the mystery associated with the unfamiliar word entropy, but will also ease the acceptance of John Wheeler's view of regarding 'the physical world as made of information, with energy and matter as incidentals'" (ibid, p.31).

These strategies for explaining the Second Law of TD and interpreting entropy concepts ultimately depend on a descriptive exploitation of informational concepts by Ben-Naim. This strategy is characteristic of certain thermo-informational syncretism along the lines of Tribus and McIrvine (1971) or Bekenstein (1973) since the 1970s (Section 4.1), integrating conceptual resources and interpretative strategies from both the Brillouinian and the Jaynesian traditions. These conceptual resources and interpretations (e.g., connecting $H_{IT}$ and $S_G/S_B$ and OI) is precisely what historically connects his proposal with the 'hard-core' of the broadly-conceived thermoinformational program. Illustratively, Ben-Naim consistently employs (see citations above) a Brillouinian concept of 'missing information' (Section 3.2) to interpret entropy epistemically and informationally, complementarily relating information (OI) and entropy (presumably $S_B$). Also, Ben-Naim identifies jaynesianly the concepts of information $H_{IT}$ and entropy (presumably $S_G$): "without entering into the controversy about the question of the subjectivity or objectivity of information, whatever it is, I believe that the entropy is identical, both conceptually and formally, with Shannon's measure of information." (Ben-Naim 2008, p.30).

From this (so to say) 'hard core' of conceptual relations, Ben-Naim went on to connect other thermophysically significant notions such as $S_{TD}$ "If one accepts the probabilistic interpretation of the entropy, and agrees on the meaning of Shannon's information, then the interpretation of the thermodynamic entropy as thermodynamic information becomes inevitable." (Ben-Naim 2008, p.xxi). Consequently, from the methodological framework that I will unfold during this section I argue that the effective achievement of Ben-Naim's intellectual goals (i.e., to informationalize and eliminate entropy concepts within thermal physics) will be legitimized on the basis of the epistemic success of his strategy of explaining the laws of TD and interpreting thermophysical concepts. Moreover, this epistemic success will depend substantially on the consistent and coherent use of thermophysical concepts such as missing information and conceptual connections such as the one postulated between $S_{TD}$ and $H_{IT.}$ Having set out this illustrative application of the forthcoming evaluative proposal, in the following sections I will develop in more detail how the use of informational concepts can effectively contribute to the attainment of knowledge or theoretically legitimize certain intellectual claims within the disciplinary field of thermophysics.

## 5.2. Framing Conceptual Misuses in Informational Descriptions

The informational research programme/tradition in classical thermophysics is constitutively defined (as opposed to other programmes in classical thermal physics, such as BSM or GSM) by its exploitation of informational concepts to obtain knowledge about thermophysical reality. As I argue, the epistemic capabilities of thermo-informational approaches depend on exploiting descriptively certain conceptual connections (see Section 2.5). In this section I analyse the two core conceptual connections on which virtually all informational explanations, predictions and interpretations in the history of thermal informationalism (from von Neumann in the late 1940s to Ben-Naim in the 2010s) are articulated: namely, the connection (either identificationist or complementarist) between $H_{IT}$ and $S_G/S_B$ and the identification between $H_{IT}$ and OI.

### 5.2.1. Conceptual identification between Shannon's Information $H_{IT}$ and Gibbs Entropy $S_G$

One of the central conceptual relations on which the main epistemic strategies and theoretical goals of information thermophysics are deployed (both historically and synchronically) is the conceptual identification between $H_{IT}$ information and $S_G$ entropy (and in the case of uniform distributions, also $S_B$). As I noted in Section 2.5, there is a general consensus in the literature that the terminological and syntactic-formal relationship between $H_{IT}$ and $S_G$ is self-evident. Firstly, the fact that $H_{IT}$ and $S_G$ have the same technical name is prima facie a historically contingent fact (Section 3.1). Secondly, there is no consensus on whether this 'superficial' (formal) connection between the two concepts is sufficiently robust to legitimize a 'deep' (semantic) identification between the meanings of $H_{IT}$ and $S_G$. There is a critical current[103] that defends that this formal

[103] Among these authors, we find Kenneth Denbigh in the 1980s: "There are, of course, good mathematical reasons why information theory and statistical mechanics both require functions having the same formal structure. They have a common origin in probability theory, and they also need to satisfy certain common requirements such as additivity. Yet, *this formal similarity does not imply that the functions necessarily signify or represent the same concepts*." (Denbigh 1981, p.113, Italics are mine). Of course, this argument also appears within the philosophical realm, where it was questioned: "Is the syntactic identity of von Neumann[-Gibbs]-Shannon-Weaver entropy really evidence of

similarity between both notions responds only to the merely technical need to define a measure with certain mathematical properties, and not to the existence of a non-explicit semantic link between $H_{IT}$ and $S_G$. Since the terminological-formal relationship between $H_{IT}$ and $S_G$ is not sufficient to identify both concepts, let us delve into their possible semantic relationships. Recall from Section 2.5.5, the main descriptive tactic to attempt to bridge such a semantic disconnect between $H_{IT}$ and $S_G$ is to redefine $H_{IT}$ systematically over the representational resources of SG, i.e., over the phase space $\Gamma$ of a physical substance, along the lines proposed by Frigg (2004) or Shenker (2020). Without elaborating on these proposals, I am going to evaluate now what are the main semantic constraints to identify (without further justification) both concepts.

### *Semantic Disconnection between $H_{IT}$ and $S_G$*

The semantic disconnection between $H_{IT}$ and $S_G$ (or $S_B$) can be found in the notion of 'state' (and its corresponding 'space of possible states') on which the technically-defined meaning of both depends, as I posited in the analysis of Section 2.5.4 and 2.5.5. Recall that the meaning of $S_G$ depends on densities $\rho$ defined over a particular region of the phase space $\Gamma$ of a molecular system, i.e., the statistical space made up of all its possible microstates. Otherwise, the meaning of $H_{IT}$ depends on the particular region occupied by a particular message *m* (i.e., proportional to the probability of occurrence of each individual symbol $s_i$ composing it) within the space of possible messages generable from the source S. While the phase space providing the meaning of $S_G$ is (synchronically) constructed from the possible values of molecular positions and velocities, the space of possible messages is progressively (diachronically) constructed as each individual symbol $s_i$ composing the message is selected among the multiple alternatives that the alphabet possesses. This representational distinction between the 'possible state spaces' underlying $S_G$ and $H_{IT}$ was pointed out by Wicken: "The formal identity of the Shannon and Boltzmann equations results from general demands on the properties of a "state" (...) Whereas [$S_B$] is based on the variety of alternative microstates among which the system moves, [$H_{IT}$] is based on states as events deriving from choices. Since a symbol set, such as a die, expresses alternatives, it becomes tempting to talk about the sequences so generated as "possessing" entropies"" (Wicken 1987, p.184).

The semantic gap between $H_{IT}$ and $S_G$ ultimately depends on the conceptual impossibility of defining the 'microstates[104]' and 'macrostates' of a physical system in a thermophysically

anything physical? (...) How surprising is it that this would be addressed in both cases by logarithmic function?" (Ladyman and Ross 2007, p.216).

[104] While the phase space $\Gamma$ on which $S_G$ is defined presupposes a concept of 'microstate' (i.e., determination of all molecular position and velocity values), the $H_{IT}$ space of possible messages $\Lambda$ does not possess a homologous representational resource because it does not make sense to speak of the three-dimensional 'position' or 'velocity' of each symbol $s_i$ within a particular sequence. As might be expected, this conceptual difference between the state spaces underlying $S_G$ and $H_{IT}$ implies that their attached probability notions of SP and IP will also be different. While $S_G$ is defined on the probability $\rho$ of a set of microstates in $\Gamma$, it would not make sense to claim that $H_{IT}$ is defined on the probability $\rho$ of (i) a set of positional-velocity values in of (ii) a particular symbol sequence *m* in $\Lambda$, but properly of the probability distribution of the source S with each individual symbol $s_i$ of the message[#]. In fact, while the probability

meaningful way by means of the representational resources provided exclusively by the space of possible messages[105] $\Lambda$. Interestingly, if $H_{IT}$ happens to be semantically defined over a phase space $\Gamma$, then $H_{IT}$ happens to become a mere trivial reformulation of $S_G$ whose only additional contribution is the logarithm in binary basis and the lack of Boltzmann constant $k_B$. At this point there seems to be two possible paths. On the one hand, some informational authors such as Ben-Naim (2008, p.30) defended that the temperature-energetic dimensions of entropic notions are unnecessary "the units of entropy (J/K) are not only unnecessary for entropy, but they should not be used to express entropy at all. The involvement of energy and temperature in the original definition of the entropy is a historical accident, a relic of the pre-atomistic era of thermodynamics (...) [reject Boltzmann's constant $k_B$] will also facilitate the identification of the thermodynamic entropy with [$H_{IT}$]". However, this elimination of the multiplicative constant of $S_G$ (identified with $S_{TD}$) to conceptually connect this notion with $H_{IT}$ can only be understood in Ben-Naim's case more as a desideratum than as a justification of the robustness of such a connection. On the other hand, the other possibility of defending a connection between $S_G$ and $H_{IT}$ is through some interpretative fixation of the meaning of $H_{IT}$ in relation to $S_G$.

density $\rho$ (i.e., representing the ensemble in GSM) on which $S_G$ is defined can be modified with the dynamical evolution of the system, for example with the approach of a gas to its equilibrium state, the probability density of the source S on which $H_{IT}$ is defined remains invariant from the beginning of the message generation. Similarly, there are also no conceptual connections between the dynamical transition between 'microstates' of the system in phase space and the transition of symbols in the space of possible messages: while the former is represented by a continuous 'path-like' function (i.e., Hamiltonian or Liouvillean) connecting microstates, the latter would be represented by the transition between a first 'selection-like' collapse of the distribution over the source S in a particular symbol-region to a second collapse. "This [$H_{IT}$] meaning of "state" bears no relation to that in thermodynamics. Neither do transitions between states. In thermodynamics, the latter involve new distributions of matter and energy among available (...) states; in communications, they involve the production of new letters in generating messages. They involve choices, and the generation of unique informational sequences." (Wicken 1987, p.183).

[105] Interestingly, the remarkable semantic disconnection between $S_G$ and $H_{IT}$ would not only be justified on the conceptual impossibility of articulating a physically relevant notion of microstate via $\Lambda$, but also on the inability to articulate a thermophysically meaningful concept of macrostate by means of the representational resources of $H_{IT}$. That is, it is not possible to encode macroscopic values $A_i$ (e.g., temperature, heat, etc.) simply by means of probability distributions over a communicative source S in a thermophysically meaningful way without somehow appealing to GSM-like mechanisms such as canonical distributions: "insertion of a probability distribution p into formula [$H_{IT}$] does not yield [$S_{TD}$] unless the probability distribution is a canonical distribution (...) we cannot associate a quantity of heat with the change of [$H_{IT}$]. For we have no rule to associate its change with the exchange of heat with the surroundings. The rule [Clausius inequality] that associates heat transfer with entropy holds only for [$S_{TD}$] and, indeed, defines it. No other entropy can satisfy it without at once also being [$S_{TD}$]" (Norton 2005, p.393). As Norton points out, even encoding heat values Q over individual symbol occurrence probabilities, there is no theoretically-defined procedure in IT for associating changes in Q with thermophysically meaningful changes in $S_{TD}$ (as with the canonical ensemble in the context of GSM). One could conventionally state that the generation of infrequent messages such as $m_1$ = 'XZXVXX' versus frequent ones such as $m_2$ = 'DAD' implies an increase in system heat Q, but this would only be thermophysically meaningful if this message can be expressed by a 'canonical ensemble' that connects Q values to certain sets of microstates. This conceptual disconnection at the 'macrostate' level not only implies that $H_{IT}$ (defined over $\Lambda$) is semantically disconnected with the non-probabilistic and indirectly observable concept $S_{TD}$ (defined over the space of observable variables of a system, see Section 2.2), but that any possible relation to this notion is conceptually mediated by $S_G$.

*Interpretative Disconnection between $H_{IT}$ and $S_G$*

Some authors in the history of thermal informationalism have claimed to defend the existence of a deep (i.e., non-merely-formal) conceptual connection between $S_G$ and $H_{IT}$ not on the basis of their technically defined meaning, as I have argued above, but on the basis of some particular way of interpretatively delimiting the meaning of these concepts[106]. Additionally, one might now ask whether the theoretical architecture[107] underlying $H_{IT}$ can contribute in a satisfactory way to specifying its identification with its formulaically related entropic measure $S_G$. Philosophers of physics such as Wüthrich (2019) have argued that the possibility of interpretively exploiting the content of $H_{IT}$ constitutes not only an opportunity to solidify its conceptual relation to $S_G$ (as posited from the Jaynesian and/or syncretic tradition à la Bekenstein [1973]), but also a platform from which to also show the enormous interpretive gap between $S_G$ and $H_{IT}$:

> "information [$H_{IT}$], arguably, is an inadmissible concept in fundamental physics. For there to be information in the first place, there must be a communication system in place, a physical set-up such that the concept of information is applicable. In Shannon's mathematical theory of communication (Shannon 1948), for there to be communication, there must be an information source of a message, a transmitter sending a signal, via a potentially noisy channel, to a receiver, which receives the signal and decodes it for the destination. (…) Even subtracting the intentionality, and abstracting from the personhood of the destination, we are still left with an ineliminable minimum level of complexity required for the signal to be interpreted as the transmission of information" (Wüthrich 2019, p.14)

In this paragraph Wüthrich argues that the possibility of successfully interpreting $H_{IT}$ as a thermophysically relevant concept in the SM domain (thus intertwining it with $S_G$) depends conceptually on specifying the thermophysical content of surrounding communicative-theoretic elements such as the communication channel CH, the information source S or the destination D. As might be expected in this direction, the enormous interpretative openness (for example) of the CH channel has been explored throughout the history of physical informationalism, from a process of measurement by Rothstein (1951) (Section 3.2) to its explicit omission by Ben-Naim (2008) (Section 5.1). Note that this complex interpretative task is not theoretically trivial, since the canonical definition of $H_{IT}$ by Shannon technically depends on the capacity (Lombardi et al. 2016a) of the communicative channel. As underlined in Chapters III and IV, the virtual totality of informational authors who exploited a plausible conceptual connection between $H_{IT}$ and $S_G$ (e.g., Jaynes, 1957a; Bekenstein, 1973) strategically omitted the monumental task of specifying in a technically precise and conceptually consistent way the thermophysical content of the

[106] This is paradigmatically the case of the classical Jaynesian identification between $S_G$ and $H_{IT}$, which ultimately depends on epistemically reconceptualizing their probabilities as representations of the degree of agents' beliefs (see Section 2.4). Because of the central role played by the ordinary notion of OI information in such Jaynesian descriptions, I will leave its detailed analysis for later.

[107] This theoretical architecture of $H_{IT}$ is properly constituted by the conceptual resources (technically defined or not) that make up the IT-based communicative model of message transmission by which the meaning of $H_{IT}$ is usually specified (Section 2.3).

communication channel, the noise, the information source S, the destination D, and so on. The main reason for arguing that this task is not only intellectually exhaustive but conceptually impossible is that the elements of Shannon's IT underlying $H_{IT}$ are 'functionally' defined (and interpreted) concepts, e.g., a source S is any entity that performs the function of selecting symbols from an alphabet. In this sense, the constitutively structural notion of 'phase space' underlying $S_G$ would not be conceptually rich enough to allow defining in an interpretatively coherent way the inherently functional notions (e.g., 'source' or 'destination`) underlying $H_{IT}$. Although following Wüthrich (2019) in defending that one might have good reasons to argue for the impossibility of coherently interpreting the elements of $H_{IT}$ within the theoretical scope of $S_G$, my main argument lies precisely in the historical fact (Chapter III and IV) that no informational proposal has shown this to be de facto feasible.

Our results in the previous sections lead us to a key dilemma: (i) either we tried to omit the interpretable content of $H_{IT}$ by reconceptualizing it as a strictly formal concept (so that its connection with $S_G$ becomes almost trivial), (ii) or we have tried to reconceptualize the resources of $H_{IT}$ in an incomplete or interpretatively incoherent way (so that its connection with $S_G$ becomes unsatisfactory). The superficial formulaic similarity between $S_G$ and $H_{IT}$ has historically served as a basis for assuming (or even delegating to future proposals) a semantic and/or interpretative connection between the two concepts. However, I argue that (a) the representational incompatibility between the phase space $\Gamma$ of $S_G$ and the message space $\Lambda$ of $H_{IT}$ suggests that the semantic inconsistency of their relationship; as well as (b) the inherent 'functional' character and the enormous conceptual complexity underlying the theoretical notions under $H_{IT}$ (e.g., source, destination, etc.) suggests the impossibility of interpretively connecting $S_G$ and $H_{IT}$ in a satisfactory and coherent way. Thus, contrary to informationalist expectations, the formal similarity between $S_G$ and $H_{IT}$ hides many unbridgeable semantic and interpretative disconnections between the two concepts. Finally, I argue that the conceptual misuse does not lie in claiming to connect $H_{IT}$ and $S_G$/$S_B$ semantically or interpretatively. This is as legitimate a conceptual claim as any other in the context of thermophysics. I argue that what does constitute conceptual misuse is to descriptively exploit such a semantic-interpretative connection without further justification, since (as I have argued above) the task of specifying the SM content of $H_{IT}$ is by no means trivial.

### 5.2.2. Conceptual identification between Shannon’s Entropy $H_{IT}$ and Ordinary Information OI

Besides the conceptual identification between $H_{IT}$ and $S_G$, the other pivotal conceptual misuse on the architectural core of the thermo-informational program is identifying the concepts of $H_{IT}$ information and OI information. Such identification is usually performed by means of a sleight-of-hand use of $H_{IT}$ and OI, wherein there is an explicit (verbal or symbolic) use of a particular notion $C_1$ while simultaneously implicitly employing the characteristic virtuous functions of a different notion $C_2$, exploiting an existing terminological link[108] between $C_1$ and $C_2$. As can be extracted from the previous historical assessment (Chapters III and IV), the history of informational thermal physics is replete[109] with instances of conceptual identification (via sleight-of-hand) between the technical-statistical notion $H_{IT}$ and the theoretically-underdetermined notion OI. An illustrative example is: "*our information* [OI] about an isolated system can never decrease (only by measurement can new information be obtained) (...) the *entropy of information theory* [$H_{IT}$] is (...) a straightforward generalization of the entropy concept of statistical mechanics" (Rothstein 1951, p.90. Italics are mine). Note in this quote that while $H_{IT}$ information is explicitly mentioned, at the same time the first instance of the term "information" denotes implicitly an agential ("our") and semantic ("about") informational concept à la OI.

Initially, the connection between the concepts $H_{IT}$ and OI resides in the fact that both are considered in the literature as notions belonging to the conceptual landscape of information (Section 2.3) and terminologically linked by the label 'information'. As I argued at length in Section 2.5.3, the possibility of conceptually relating $H_{IT}$ and OI non-superficially (i.e., on a semantic and interpretative level) resides in a limited domain of application of what I called the $H_{IT}$-OI vertex of the 'Golden Information Triangle', outlined by Timpson (2013, p.30). However, this common conceptual space between $H_{IT}$ and OI (extremely limited to irrelevant applications) is overshadowed by the countless cases in the literature that show their profound disconnection. This a toy example of Aczél and Daróczy (1975): suppose that the probability that a crow is black vs. white is 0.99 vs. 0.01 except in the case of having a white mother, where the probability that the crow is white is 0.5.  In this scenario, by obtaining the information (OI) that an unobserved crow has a white mother, the $H_{IT}$ entropy (interpreted as 'uncertainty') about the color of the crow paradoxically increases[110]. Another interesting case is that of Uffink (1995, p.234), where the $H_{IT}$ of the probability for the location of the house keys increases when I obtain the information (OI) that they are not in my pocket (namely, the place where I assumed I was most probable to find them). As the latter author notes, although Jaynes was well-aware of the conceptual disconnection between $H_{IT}$ and OI, this did not prevent him from encouraging the sleight-of-hand between the two I "one can easily invent situations where acquisition of a new piece of knowledge (that an event previously considered improbable had in fact occurred) can cause an increase in the entropy" (Jaynes, 1957b, p.186).

[108] An illustrative example of conceptual sleight of hand not between $H_{IT}$ and OI but between $S_G$ and $S_{TD}$ can be found in the quote "there is not only an analogy between entropy [$S_G$] and information [$H_{IT}$], but an identity between the thermodynamic entropy [$S_{TD}$] and Shannon's measure of information [$H_{IT}$]" (Ben-Naim 2008, p.22). Here, Ben-Naim explicitly employs the concept of $S_{TD}$ to identify it with $H_{IT}$, while in the first stay of the term 'entropy' the statistical concept $S_G$ is implicitly used to exploit its formal analogy with $H_{IT}$, implicitly using both notions.

[109] I argue that the pervasiveness of this tactic is due to the aim of proponents of this intellectual current to capitalize simultaneously (i) on the interdisciplinary popularity and technical-theoretical consistency of $H_{IT}$ and (ii) the capacity of OI to capture agency and semantic-epistemic relevance.

[110] This example of Aczél and Daróczy (1975) lies outside the 'Golden Information Triangle' described in Section 2.5. It also constituted an example of the interpretative disconnection between OI and $H_{IT}$ assessed in Section 5.2.3.

Fortunately, from the celebrated assessments of Carnap and Bar-Hillel (Section 3.4) to the present day there is within the philosophical literature a sort of consensus on the conceptual disconnection between $H_{IT}$ and OI: "S(P) [$H_{IT}$] is a technical conception of information, which should not be taken as an analysis of the various senses of 'information' in ordinary discourse. In ordinary discourse, information is often equated with knowledge, propositional content, or meaning." (Frigg 2008). While acknowledging this disconnection, the Dretskean (1981) and post-Dretskean[111] (e.g., Skyrms, 2010; Martínez, 2019; or Isaac, 2019) programs held out the intellectual promise of undermining the 'asemantic dogma' of IT and developing a technical notion from $H_{IT}$ that mimics the attractive conceptual properties of OI, namely, its agential and semantic-epistemic character (see Section 2.4.4). Isaac[112] (2019) builds on Skyrm (2010) to argue that the semantic information that a signal $e_1$ transmits about an event $e_2$ can be encoded by vectors composed of $H_{IT}$-derived log probability ratios $P(e_2|e_1) / P(e_2)$. In the case of Martínez (2019), semantic information can be extracted from $H_{IT}$ by rate-distortion sweet spot extraction, where the channel rate measures the faithfulness with which the information is transmitted from the source.

However, as Dretske claimed, even if there is a theoretical possibility of developing a technical-theoretical semantic informational notion from $H_{IT}$, there is no guarantee that this semantic latent in Shannon's IT will be epistemically relevant in a given disciplinary context (as in this case is classical statistical physics): "the ordinary, semantically relevant, sense of information (...) is something to be distinguished from the concept of meaning. (...) Whether [communication] theory is capable of furnishing an adequate account of information, in the semantically relevant sense, is a question about whether it can provide an illuminating account of that commodity capable of yielding knowledge" (Dretske 1981, p.46). Philosophers of physics such as Timpson and Lombardi (Section 4.4) have been highly skeptical about the possibility of effectively developing a semantically relevant technical concept of information "Shannon's theory, taken in itself, is purely quantitative: it ignores any issue related to informational content. Shannon information [$H_{IT}$] is not a semantic item: semantic items, such as meaning, reference or representation, are not amenable of quantification. Therefore, the issue about possible links between semantic information and

[111] The results (or properly, its lack of results) of the original Dretskean program have been criticized from the main recent trends of informational naturalization of semantics à la Isaac (2019) or Martínez (2019), who argues that the resources linked to $H_{IT}$ and IT (including his model of signal transmission) suffice to generate a notion of relevant semantic information "informational content in the Dretskean tradition is not by a long shot all there is to information theory. (…) it does mean that no Dretske-style "semanticized information" needs to be recognized, over and above the quantities studied in information theory proper" (Martínez 2019, p.1215)

[112] Interestingly, Isaac argues for the semantic latent of Shannon's IT by appealing to the proposal developed by Turing in the 1940s: "Yet Shannon's was only one of two parallel endeavours to mathematically analyse information. A formal apparatus analogous to Shannon's had already been developed independently at Bletchley Park by Turing and colleagues in their daily attempts to crack the Enigma code. While much of the maths was the same (in particular, the appeal to log probabilities as the measure of information, Good [1979]), the goal of Turing's project was radically different, namely to infer from an opaque string of symbols its intended meaning and, more generally, the Enigma machine settings encoding all German messages that day. Thus, whereas Shannon's project was unconcerned with meaning per se, Turing's was focused on meaning above all else" (Isaac 2019, p.104).

Shannon information [$H_{IT}$] is a question to be faced once the concept of Shannon information is endowed with a sufficiently clear interpretation" (Lombardi et al. 2016b).

### 5.2.3. Conceptual Misuses under Informational Descriptions

I have just delimited the two main conceptual misuses that underlie (directly or indirectly) most descriptive strategies employed by informational thermophysicists, from Shannon's bandwagon to the most recent proposals such as Ben-Naim's. In fact, this descriptive reliance on particular non-justified conceptual connections is usually linked to other relations between notions that have not been technically specified either. Illustratively, a canonical way of extending the misidentification between $H_{IT}$ and $S_G$ is through its extension with a conceptual identification[113] between $H_{IT}$ and $S_{TD}$: "It can be shown that Shannon's function and Clausius' function are the same. (...) In retrospect a logician can show that with these criteria [consistency and additivity] Shannon was bound to produce a measure that would be consistent with thermodynamic entropy. Once it is recognized that the two subjects derive from common considerations it is straightforward to derive one from the other" (Tribus and McIrvine 1971, p.180, 182).

I must conclude this section by arguing that the mere fact of aiming to connect semantically and interpretatively two distinct concepts does not constitute a conceptual misuse per se. The history of modern physics constantly shows us that the extension of the applicative domain of a scientific concept (and its semantic connection with other notions) can eventually be epistemically fruitful (e.g., Wilson 2006). What properly constitutes a 'conceptual misuse' in this case is (i) employing an identification between $H_{IT}$ and $S_G/S_B$ without technically justifying (à la Frigg [2004] or Shenker [2020], see Section 2.5.4) how this connection can be consistently established, and (ii) interchangeably employing $H_{IT}$ and OI without rigorously specifying (à la Isaac [2019] or Martínez [2019]) how this connection can be draw. As I will argue below, these conceptual connections constitutive of the thermo-informational programme are not only exploited descriptively in a theoretically unjustified way but (and here the key to my analysis) they also do not epistemically contribute to the explanation, prediction and interpretation of thermal phenomena.

[113] For example, let us attend to the following statement "The concept of an inherent connection between the entropy of Clausius [$S_{TD}$] and the intuitive notion of information [OI] preceded Shannon's work by many years" (Tribus and McIrvine 1971, p.182). Note that the descriptive tactic expressed by Tribus and McIrvine in this quote hinges on a significant number of conceptual connections that require specification, namely: that there exists (at a minimum) (i) a robust conceptual connection between $S_G$ (or $S_B$) and $S_{TD}$ (Section 2.5.3), and (ii) the ability of OI to fix interpretatively the entire meaningful content of $S_G$ (or $S_G$). In this sense, the degree of conceptual misuse derivable from Tribus and McIrvine's (1971) descriptive tactic would be proportional to the degree of disconnect we find after evaluating conceptual architecture (i) and (ii). Illustratively, arguing that $S_{TD}$ and OI possess an 'inherent connection' would entail a series of misuses from the most explicit (i.e., OI requires an agent, but $S_{TD}$ is agent-dependent) to the less obvious, such as a conceptual sleight-of-hand between $S_{TD}$ and $S_G$ that would extend the OI-based epistemic interpretation of $S_G$ over the notion of $S_{TD}$.

## 5.3. Informational Explanations in Statistical Thermal Physics

One of the main ways of obtaining knowledge in the physical sciences is through the explanation of certain phenomena in their respective domains of research. In this direction, it is evident that one of the main epistemic virtues of thermophysics is its ability to explain the macroscopic behavior of certain substances from the evolution of observable quantities (in the case of TD) or from the microscopic dynamics of their molecular components (in the case of SM). Due to the nature of this informational research program, in the following I focus mainly on the domain of SM explanations. Among the main illustrative cases of SM-explanations one could find: how a substance undergoes a phase transition from a solid to a fluid state or under what circumstances certain observable properties of a macroscopic substance remain stable. To restrict my argument, I further focus on those explanations of the paradigmatic thermal behavior of the approach of a substance to equilibrium, illustratively the free expansion of a gas in the volume of its container.

### 5.3.1. Framing Explanations and Explanatory Power in Statistical Mechanics

Although the first systematic evaluation of thermophysical explanations dates back at least as far as Sklar (1993), the consolidation of the philosophical debate on explanations of equilibrium processes is relatively recent (Allori 2020). Due to the enormous heterogeneity of factors involved in such explanations (i.e., statistical modelling, mechanical properties or dynamical hypotheses), it is not evident which model of scientific explanations could best give a philosophical account of their epistemic functioning, namely: deductive-nomological, inductive-statistical, causal, unificationist or pluralistic model[114]. In this direction, authors such as Myrvold (2016) argue that certain 'ontological' interpretations of probabilities play a decisive role in the explanation-SM of systems' approach to equilibrium, thus combining a model of statistical (i.e., appeals to probabilities) and causal (i.e., probabilities represent physical properties) explanation. Regardless of the particular model of explanations one might hold, any thermophysical description that aims to explain the approach to equilibrium should at least specify the following elements:

> "A system behaves as it does because of its dynamics, together with initial conditions. Explanations of relaxation to equilibrium will have to involve an argument that the dynamics, together with initial conditions of the right type, yields that behaviour, plus an explanation of why the sorts of physical processes that give rise to the sorts of systems considered don't produce initial conditions of the wrong type" (Myrvold 2016, p.33)

[114] Illustratively, authors like Wilhem (forthcoming) propose to use the deductive-nomological model to formalize the 'typicalist' explanations of the equilibrium approach defended from Neo-Boltzmannian quarters in the context of BSM (Section 2.4), differentiating them from those other SM-explanations of equilibration that appeal directly to probabilistic concepts.

This idea in the philosophy of SM-explanations (Lazarovici and Reichert 2015, p.704; Myrvold, 2016, p.33) derives from the fact that any explanation of this thermal behavior (i.e., the thermal equilibrium approach) must at least follows from (i) the dynamics of the molecular components of the systems and (ii) the initial conditions in the evolution of the system. Of course, beyond these common elements each particular proposal on the market will diverge depending on various parameters, such as the fine-grain of the dynamics. While Frigg and Werndl (2011a) argue that what would explain the equilibrium approach is precisely that the dynamics of molecular systems is 'epsilon-ergodic', other authors such as Albert (2000) emphasize that the explainability of this phenomenon depends on representing the initial conditions by a uniform distribution over a tiny macrostate (i.e., the so-called 'Past Hypothesis', Section 2.4). This 'minimalist' proposal relies on the fact that what we might call the 'power' (indicative of the quality of explanatory knowledge it generates) of an SM explanation depends ultimately on the possibility of descriptively specifying (i) the initial conditions, and (ii) the dynamical behavior of microstates, that theoretically enable us to abduce the reasons underlying that such macroscopic behavior. To elaborate more on this idea, I propose the following criterion.

> **Explanatory Power Criterion:**
>
> A description $C_A$ of the thermal behavior of a molecular system allows to develop a powerful explanation of this phenomenon if its descriptions $C_i$ refer explicitly to (i) the initial conditions and (ii) the particular dynamic evolution of microstates, and this descriptive specification would allow to validly abduce how the evolution of the actual microstate of the system might give rise to such thermal behavior.

By 'powerful' SM explanation of a macroscopic thermal phenomenon such as the free expansion of an ideal gas I mean that explanation which explicitly specifies how the evolution of the actual microstate[115] $x$ (independently of considering $x$ being contained in a macrostate-region $\Gamma_M$, in a density-ensemble $\rho$ or in a constrained-region $\Gamma_C$) gives rise to such thermal behavior. Initially, an SM explanation of this phenomenon from the mere statistical behavior of certain average values will not count as 'powerful'. Although those values are ultimately defined on sets of microstates, their reference to individual microstates of the system can only be specified indirectly by means of, for instance, an averaging principle (See Section 2.5). Moreover, for an explanation to be properly statistical mechanical it must rely on the descriptive specification of molecular positions and velocities, and not only on observable- properties $A_i$ (as in the case of purely TD-phenomenological explanations) or merely statistical values. It should be remarked that an SM explanation is 'powerful' does not imply that it is 'the' explanation of that phenomenon, it simply

[115] I should stress that in attempting a descriptive specification of the actual microstate of the system we are not necessarily committing ourselves (even if this is the case) to a molecularist ontology, as some authors in the literature suggest (see McCoy 2020). In seeking to specify descriptively the actual microstate of the system, we simply aim to highlight how the current system state is represented from the theoretical framework (SM) in which we find ourselves, which is in no other way than by molecular position and velocity values. In this sense, our Explanatory Power Criterion does not necessarily give preference to ontic models (à la Salmon [1984]) of SM explanations over other models of scientific explanations, e.g., nomological, statistical-inductive, unificationist, etc.

means that it possesses the characteristics to satisfactorily account for that behaviour within the SM theoretical framework.

The above criterion establishes two necessary conditions (or 'minimalist desiderata') (i) and (ii) for a description-explanans $D_y$ resulting from a descriptive tactic to abductively generate powerful explanations of the equilibrium approximation as explanandum. Although (i) (ii) are sufficient to develop my argument, these necessary conditions are not prima facie incompatible with additional necessary or sufficient conditions that we might find in the literature (see Allori 2020). To illustrate this criterion, let us approach a case of a non-informational (in this case Neo-Boltzmannian) explanation of the equilibrium approximation. Let us initially consider a description $D_x$ in which we specify (i) the macrostate $\Gamma_M$ representing the initial condition of the system, and (ii) the 'typicality' (Section 2.4) of the microstates of $\Gamma_M$ as a dynamic property that allows us to abduce that a gas will be in equilibrium at $t_1$. Thus, following Wilhelm's (2019) model, the argumentative structure underlying this 'typicalist' (non-informational) explanation is:

| | |
|---|---|
| Description-Explanandum $D_x$: | (i) & (ii): Microstates in $\Gamma_M$ typically evolve to $\Gamma_{Meq}$ |
| Initial Condition (i): | The actual microstate $x$ is in $\Gamma_M$ at $t_0$ |
| Dynamic Property (ii): | Most microstates in $\Gamma_M$ at $t_0$ will be in $\Gamma_{Meq}$ at $t_1$ |
| Abducted Explanans $D_y$: | It is *plausible* that the microstate $x$ is in $\Gamma_{Meq}$ at $t_1$ |

As seen in the argument above, I propose (along the lines of Psillos 1999, Ch. 4]) that underlying any powerful explanation generally lies an abductive reasoning that 'plausibly' connects the content of a description-explanans with the behavior-explanandum in an 'inferentially valid' way. An abduction[116] (broadly understood as 'ampliative plausible inference' see Douven [2021]) would be 'valid' if the content of the explanandum $D_y$ follows plausibly from the abducted explanandum $D_x$ in a semantically consistent way, i.e., in virtue of the technically defined meaning of the concepts, and syntactically correct, i.e., according to theoretically defined procedures. This pragmatic notion of 'inferential validity[117]' allows us to ensure that the inferred content can not only be true (albeit fortuitously) but is dependent on the meaning of the descriptions upon which they are articulated. In this sense, the typicalist explanation of the approach of a gas to equilibrium set forth above unfolds upon a valid abduction, and because in description $D_x$ we also semantically specify (i) its initial condition and (ii) a microstate-based dynamic property (i.e., typicality), this explanation

[116] On the inadequacy of deductive reasoning in this context, see Lazarovici and Reichert (2015, p.704).

[117] Let us clarify that an inference will be 'valid' (not in a formal-deductive sense [Prawitz 2012], but broadly pragmatic) whenever the content of what is inferred depends semantically on the content of the descriptions used to deploy the inference. For example, suppose we infer that the molecules of a freely expanding gas will eventually occupy the entire container 2V containing the gas by assuming an increase in $S_B$ from the description $D_z$ = 'this substance tend to 'mimic' the volume of their containers'. In this sense, even if the inference is true, this inference made by the agent does not follow semantically from the content of any concept of entropy, but from a notion inconsistent with the theoretical principles of thermophysics (i.e., 'mimic the container'). Therefore, such an inference will be invalid.

will be 'powerful', providing us with meaningful knowledge in the case of being successful. Again, the debate on whether a thermophysical explanation is successful or not is distinct from the debate on whether an explanation 'can' (in terms of its 'power') be successful or not.

Finally, this 'minimalist' criterion constitutes a neutral evaluative tool that allows us prima facie to assess the plausible explanatory power of any theoretical framework in SM, be it BSM, GSM or the multiple informational proposals, without presupposing any privilege among them. Illustratively, Wallace's (2011) Gibbsian explanation based on his 'Simple Dynamical Conjecture' is as likely to constitute a powerful explanation as Albert's (2000) Neo-Boltzmannian proposal based on the 'Past Hypothesis', except that the initial conditions are represented differently by ensemble-based probabilities densities and Boltzmannian macrostates, respectively. However, the key to why this is a legitimate criterion for my task lies precisely in its ability to be satisfied even from informational proposals, and that it would be accepted by its main proponents (Brillouin, Jaynes, Bennett, etc.). Let us now consider an illustrative example of an informational SM explanation of a free expansion (such as the one mentioned above) that can be qualified as 'powerful'. Suppose we have the description-explanandum $D_x$ = 'Agent *a* possess well-defined semantic information $I(t_0)$ about the actual microstate of the system being in $\Gamma_M$ at $t_0$' and $D_y$ = 'Agent *a* possess well-defined semantic information $I(t_0\text{-}t_1)$ about how the actual microstate of the system will typically or with high probability transition to $\Gamma_{Meq}$ at $t_1$', wherein it is informationally specified both (i) the initial condition of the system and (ii) the microstate-based dynamical property, allowing to statistical mechanically abduce such macroscopic behavior. Therefore, an SM explanation based on the informational descriptions $D_x$ and $D_y$ would (if this well-defined semantic information[118] concept consistently refers to molecular properties in $\Gamma$) possess the capability to be 'powerful'. Insofar as such powerful SM explanations are prima facie compatible with the descriptive tactics of any thermoinformational tradition[119] (Brillouinian, Jaynesian, etc.), this 'Explanatory Power Criterion' could also be applicable within the informational research program in thermal physics. Having clarified these points, let us now proceed to use this evaluation of the explanations deployed from thermophysical informationalism.

### 5.3.2. Explanations in the Brillouinian Tradition

---

[118] The technically-defined information concept I presuppose here is both agent-based and semantically relevant within SM, since it must quantitatively determine how scientific agents refer to the individual microstates and to the dynamics that give rise to such macroscopic behavior. Otherwise, such factors could not be validly abducted. In principle, the candidate semantic information concept could be the one engineered by Floridi (2011), or the one naturalistically developed by Isaac (2019) or Martínez (2019) (see Section 5.2.2). However, whether these semantic information concepts can be theoretically relevant within an SM framework remains an open question.

[119] The tentative example of 'robust' informational SM explanation used would be paradigmatically compatible with the Brillouinian program, since it is based on consistently defining quantities of semantic information (in our example, this concept would not be its 'bound information' $I_{Bb}$ but another with technically well-defined semantic properties) over Boltzmannian macrostate-regions $\Gamma_{Mi}$ (see Section 3.2) and over the dynamical evolution of their microstates. For the example to be compatible with the Jaynesian framework, these technically-defined quantities of semantic information (incorporated by us) must be defined over regions of phase space bounded by the macroscopic constraints C (see Section 3.3) and likewise over the dynamical evolution of their microstates.

One of the main intellectual claims of the informationalism that emerged in the early 1950s with Shannon's bandwagon (Section 3.1) was the exploitation of informational conceptual resources to explain thermal phenomena. As certain descriptive tactics and interpretative strategies were progressively consolidated within the scientific community, the development of informational explanations of the fundamental theoretical elements of TD and SM also began to become standardized within certain intellectual trends (or even traditions) within this field. This was precisely the case of Rothstein (1951) (as argued in Section 3.2.2), who sought to exploit by means of certain descriptive tactics the possibility of giving an explanatory account of TD principles by means of the then omnipresent Shannon $H_{IT}$ concept: "we can formulate the statistical expression of the second law of thermodynamics rather simply in terms of information: *our information about an isolated system can never decrease* (only by measurement can new information be obtained)" (Rothstein 1951, p.90. Italics are mine).

If we evaluate the descriptive tactics underlying this explanation, its first distinctive feature of the is a conceptual sleight-of-hand (see Section 5.2.2) between the Shannon $H_{IT}$ notion of information (e.g., "the entropy of information theory..." [ibid]) and the ordinary OI notion, whose agent-centric and semantic-epistemic character is explicitly manifested in the expression "our information about" in the above quotation. Immediately linked to this element we also find the assumption of a conceptual generalization of $S_G$ in $H_{IT}$: "the entropy of information theory is, except for a constant depending on the choice of units, a straightforward generalization of the entropy concept of statistical mechanics" (ibid). Based on this descriptive tactic, Rothstein developed an informational redescription $D_x$ of the Second Law of TD (i.e., our information about a system cannot decrease) as explanans of a non-informational statistical description $D_y$ of the Second Law as explanandum. However, this explanation cannot be powerful precisely because of its inability to refer (i) to the initial conditions or (ii) to the microstate-based dynamics of the system satisfying the Second Law. That is, from the idea $D_x$ = 'our information (OI) about the system cannot decrease' it is not possible to validly abduce the microstate-based dynamic reasons underlying the thermo-statistical fact that $D_y$ = 'the entropy of a closed system cannot decrease'. This is because the descriptive tactic underlying Rothstein's explanation is semantically inconsistent (i.e., $H_{IT}$ does not technically refer to the amount of OI that an agent about a system) and/or interpretatively incoherent (i.e., the OI an agent possesses over the macroscopic initial conditions of a system far from equilibrium decreases as the system equilibrates). As it was noted in Section 3.2, this informational explanation in thermophysics proposed by Rothstein (ibid.) was conceptually refined and interpretatively reinforced (although this does not imply an epistemic advance, as we shall see) with the fine-grained theoretical proposal of Brillouin (1956, 1962):

> “Let us suppose that we have additional *information on the state of the gas*: for instance, we may happen to know that the gas, at a certain earlier instant of time, occupied a smaller number V’. This is the case if the gas is in a container V’ and we suddenly open the connection with another volume V’’ […] After we open the volume V’’, the gas flows in, density oscillations takes place between the two volumes, and the steady state is progressively established with a uniform density through

> the volume V. Increase of entropy and loss of information proceed together. We may say that the *gas progressively "forgets" the information*" (Brillouin 1962, p.157. Italics are mine)

First, let us evaluate the descriptions underlying this Brillouinian explanation. As it was analysed in Section 3.2.3, Brillouin's NPI entails a conceptual architecture dependent on his notion of bound information $I_{Bb}$, from which he intends to reinterpret Boltzmann entropy $S_B$ by means of the ordinary notion of information OI, i.e., by quantifying "information on the state of the gas" (ibid.). Thus, the tactic Brillouin adopts above incurs (in a Rothstein [1951]-like fashion) a 'conceptual sleight-of-hand' identification between $I_{Bb}$ and OI, ambiguously exploiting the technically-defined character of the former and the non-technical ability of the latter to intuitively specify the epistemic dynamics of agents in thermo-statistical contexts (Section 5.2). Because of this configuration of his descriptive tactics, the epistemic key to the explanation deployed by Brillouin (1962, p.157) rests precisely on the ordinary notion of OI information. This author strategically deploys his informational description $D_x$ of a gas approaching equilibrium (i.e., "gas progressively "forgets" the information" [ibid, p.157]) as explanans of a non-informational Boltzmannian description $D_y$ of the equilibration process as explanandum. According to the above criterion this Brillouinian explanation would not be powerful because, although it allows to determine abductively (i) the initial conditions (representing the initial volume V' by a Boltzmannian macrostate $\Gamma_{M0}$), it would be unable to do the same with (ii) the dynamics of the actual system's microstate approaching equilibrium. That is, in the Brillouinian informational explanation of the equilibration process, the specification of the microstate-based dynamic reason underlying the fact that a gas will expand freely throughout the volume of its container is conceptually not possible.

| | |
|---|---|
| Brillouinian Description $D_x$: | 'The gas progressively forgets information during $t_0$ - $t_1$' |
| Initial Condition (i): | The actual microstate *x* is at $I_{Bb}(\Gamma_V)$ at $t_0$ |
| Dynamic Property (ii): | (non-conceptually specified) |
| Abducted explanans $D_y$: | It is *plausible* that the actual microstate *x* is in $\Gamma_{2V}$ at $t_1$ |

From the informational description $D_x$ = 'gas progressively "forgets" the information' (ibid.) it would be impossible to validly and non-trivially (i.e., in a semantically consistent way) abduce the dynamic reasons underlying the thermo-statistical fact that $D_y$ = 'the steady state is progressively established with a uniform density through the volume V' (ibid.). This lack of explanatory power of the Brillouinian program has essentially two sources in the above descriptive strategy. Firstly, that its explanatory key ultimately rests on OI implies that the agent-dependent, semantic and epistemic character of OI should contribute to validly abduce the dynamic reasons underlying the free expansion of a gas. Without further theoretical specification of how the OI information possessed by agents about the microstates of molecular substances would allow abduction of dynamical properties, this Brillouinian concept of OI information (or explicitly its $I_{Bb}$ notion) is irrelevant for generating thermophysical explanations. A similar thesis has been defended by

Earman and Norton in the particular case of Brillouinian explanations of the Maxwellian demon failure: "Brillouin's labeling of the quantity [$I_{Bb}$] as 'information' is intended, of course, to suggest our everyday notion of information as knowledge of a system. But those anthropomorphic connotations play no role in the explanation of the Demon's failure (...) The anthropomorphic connotations of human knowledge play no further role" (Earman and Norton 1999b, p.8). In addition to the explanatory irrelevance of OI, Brillouin deploys in the paragraph above an inconsistent use of the term 'information', so that sometimes it expresses an agent-dependent concept "we have additional information" and sometimes it refers to a property of the observed system “gas progressively "forgets" the information".

Secondly, the lack of power of Brillouinian explanations also depends constitutively on the semantic robustness and epistemic contribution of their NPI-mediated connection between entropy $S_B$ and information $I_{Bb}$. I argue that apart from allowing conceptual sleight-of-hand between OI and $H_{IT}$, employing information $I_{Bb}$ actually employs the thermophysically content of the notion of $S_B$ entropy upon which it is defined. Although $I_{Bb}$ and $S_B$ are terminologically distinct, both constitute de facto the same concept (i.e., the semantic content of both is reduced to $S_B$), where the term 'bound information' serves only the function of implicitly invoking the notion OI. This critical line starts from Denbigh "bound information is nothing more than a name given by Brillouin to an "entropy change" encoded in Boltzmann-type complexions" (Denbigh 1981) and was later taken up by Earman and Norton "*All that matters for the explanation* is that the quantity I is an oddly labelled quantity of entropy and such quantities of entropy are governed by the Second Law of thermodynamics." (Earman and Norton 1999, p.8. Italics are mine). In this sense, Brillouinian explanations do not add any additional epistemic value to any other non-informational or conventional explanations that could be deployed from the BSM conceptual apparatus. In short, explanations derived from the Brillouinian informational approach have two defining characteristics: namely, they are epistemically parasitic on the representational resources of BSM (i.e., $S_B$, or $\Gamma_M$, see Section 2.2); and the descriptive exploitation of information concepts in Brillouinian explanations do not contribute to abduce the dynamic behavior of the actual microstate underlying the (macroscopic) approach to thermal equilibrium.

### 5.3.3. Explanations in the Jaynesian Tradition

Unlike Brillouinism, which claims to be an informational program with explanatory pretensions in the thermophysical domain, the Jaynesian program (see Section 3.3) was not originally developed to generate explanations in the field of thermophysics (or any other domain of physics). This manifests itself in that for Jaynes, the entire theoretical framework of MEP-based SM should be interpreted not as a platform for directly describing thermal phenomena, but otherwise descriptions of the inferential capabilities of agents: "in the problem of prediction, the maximization of entropy is not an application of a law of physics, but merely a method of reasoning" (Jaynes 1957a, p.630). In spite of this non-explanatory pretension, Jaynes eventually generated explanations exploiting the conceptual architecture of his MEP, as in the following case: "We finally see the *real reason for the second law*: since phase volume is conserved in the dynamical evolution, it is a fundamental requirement on any reproducible process that the phase volume W compatible with the final state cannot be less than the phase volume $W_0$ which describes our ability to reproduce the initial state" (Jaynes 1965, p.396. Italics are mine). Again, let us analyse the descriptive tactic underlying this Jaynesian explanation of the second law. First, according to the epistemic Jaynesian interpretation of the phase volume $W_0$ under $S_G$ entropy, it (implicitly connected to OI) describes the ability of agent-observers to identify the actual microstate within a region, from where we could interpretively derive the idea that Liouville's theorem would represent the ability of agents to retrodict and predict the past and future microstate of the system[120], respectively (see Parker 2011, footnote 11).

According to Neo-Jaynesian authors such as Parker (2011), the MEP conceptual architecture is powerful enough to be able to informationally explain thermophysical principles such as the Second Law or the Maxwellian demon failure: "the Jaynesian offers an explanation of the content of the Second Law [and the approach to equilibrium], in that it explains the impossibility of exploiting the statistical nature of thermal systems to generate work at no entropic cost." (Parker 2011, p.855). Otherwise, I argue that it is not possible to generate powerful explanations in the thermophysical domain from descriptions generated from the Jaynesian tradition. My central argument to support this thesis is that explanations dependent on Jaynesian descriptions such as $D_x$ = 'the ability of the agent to reproduce the initial state cannot decrease with evolution' do not allow us to validly abduce (i) the initial conditions and (ii) the dynamic factors underlying the statistical redescription of the Second Law (and the equilibrium approximation) that assumes as explanandum $D_y$ = 'the impossibility to statistically exploit the thermal systems to generate work without entropic cost'.

| | |
|---|---|
| Jaynesian Description $D_x$: | ‘our ability to reproduce the initial state cannot decrease’ |
| Initial Condition (i): | (the actual microstate *x* is in $\Gamma_C(V)$ at $t_0$) |
| Dynamic Property (ii): | (non-conceptually specified) |
| Explanans Abducted $D_y$: | It is *plausible* that the gas *x* is in $\Gamma_C(2V)$ at $t_1$ |

The impossibility of validly abducting the microstate-driven dynamic rationale of the thermo-statistical description $D_y$ from the thermo-informational description $D_x$ should be explored in the explanatory contribution of the notions of information and in the epistemic dependence on conceptually inconsistent and interpretatively incoherent descriptive strategies. Firstly, the

[120] Note that these types of informational redescriptions are underpinned by descriptions based on fine-grained $S_G$, although the Jaynesian demonstration of the de facto equilibrium approximation depends constitutively on coarse-graining $S_G$ procedures (see Section 3.3).

specification of system initial conditions in Jaynesian explanations of equilibration depends on assigning densities by maximizing $H_{IT}$, interpreted as a measure of the amount of uncertainty (or OI information) of the agent about microstates, under macroscopic constraints[121] like $\Gamma_C(V)$. In this sense, the explanatory contribution of these informational concepts is provided by the interpretation of the numerical maximization of $H_{IT}$ as an effective representation of the 'decrease' in the amount of uncertainty or lack of information (OI) of the agent about the microstate. As remarked in Section 3.3, due to the Jaynesian interpretation this plausible explanatory contribution of $H_{IT}$ and OI is ultimately reduced to the explanatory contribution of epistemic probabilities, raising the "explanatory question: Why do our probabilistic assumptions work so well in giving us equilibrium values" (Sklar 1993, p.193). But this is a debate unrelated to the one that concerns us here: namely, whether epistemic SM-probabilities are explanatorily irrelevant (as Albert [2000] or Frigg [2011a] argue) or can contribute to explaining the epistemic dynamics of agents, as Uffink (2011) proposes (Section 2.4). In any case, the agent-dependent and semantic-epistemic properties that Jaynes interpretatively attributes to $H_{IT}$ (implicitly connecting with OI) are conceptually irrelevant and insufficient to explanatorily abduce the molecular-dynamic reasons underlying the equilibrium approach. Thus, $H_{IT}$ and OI seems to become epistemically inert in the Jaynesian explanations of thermal behaviors.

Of course, Jaynesian explanations (based on coarse-grained descriptions) of the approach to equilibrium will be conceptually independent of the molecular dynamics underlying such macroscopic behavior[122] 'the dynamics of the system does not play any role in Jaynes' derivation of the microcanonical distribution' (Frigg 2008, p.94). This dynamical disregard underlying Jaynesian non-equilibrium can be understood as a phase-averaging introduction at t of a coarse-grained description $\bar{\rho}$ mutually inconsistent with the fine-grained description $\rho$, because the system is statistically simultaneously characterized at t as possessing distinct molecular properties modeled by $\rho$ and $\bar{\rho}$. Because eventually introducing a coarse-grained description $\bar{\rho}$ removes the agent's epistemic access to (i) the initial condition of the system, any explanation dependent on coarse-grained descriptions that does not explicitly specify (ii) the dynamic property would (according to the Explanatory Power Criterion) constitute a non-powerful explanation of the equilibrium approximation. In Uffink's words: "the increase of entropy and the approach to equilibrium would thus apparently be a consequence of the fact that we shake up the probability density repeatedly in order to *wash away all information about the past*, while refusing a dynamical explanation for this procedure" (Uffink 2010, p.196. Italics are mine).

### 5.3.4. Recent Explanations in Syncretic Informationalism

[121] Technically, $\Gamma_C(V)$ encodes the phase region constrained by the macroscopic value V.

[122] As we saw in Section 3.3, this dynamic disregard is not a non-indigenous feature of non-equilibrium MEP, but appears originally in the context of non-equilibrium GSM (Section 2.2.3) from which the Jaynesian program is conceptually nourished.

Since the early 1970s informational proposals (e.g., Tribus and McIrvine, 1971; Bekenstein, 1973) systematically combined descriptive tactics and interpretative strategies of Brillouinism and Jaynesianism (see Section 4.1). One of the main claims of this informational current is the explanation of phenomena and notions within thermal physics "the information-theory treatment of thermodynamics clarifies the concept of equilibrium" (Tribus and McIrvine 1971, p.186). This will to vindicate the explanatory capacity of informationalism within classical thermophysics was historically prolonged for several decades until our time, where curiously enough the epistemic success of these informational explanations was assumed within certain scenarios of scientific popularization "The laws of thermodynamics (...) are, underneath it all, laws about information" (Seife 2007, p.2) and was replicated by certain proposals of recent years.

Ben-Naim is one of the current informational physicists who most strongly defend the explanatory value of this line of descriptive tactics (Section 5.1). He intended to develop an explanation so ambitious as to systematically account for various macroscopic thermal processes (e.g., mixing of two substances, expansion of a gas or isothermal compression of a liquid) by means of certain informational concepts: "Using information as a fundamental concept makes the understanding of the Second Law much easier (...) We shall see that the MI is the best way of describing what is the quantity that changes in a spontaneous process" (Ben-Naim 2008, p.251). Let us now focus on the explanation that Ben-Naim deploys on the paradigmatic case of a gas expanding from a volume V to the total volume of a container 2V. This explanation is based on specifying that property that causally accounts for the entropy increase $S_B$ that occurs during an equilibration process such as the expansion of a gas "in an expansion process I, we lose real information. We know more about the locations of the particles before the process than after the process. It is *this loss of information that is responsible for the change in the MI in the expansion process*" (ibid. p.281. Italics are mine).

What is the descriptive tactic that allows Ben-Naim to make such a claim? As it was noted in Section 5.1, this author describes the initial state of an expanding gas by means of a certain amount of 'missing information' $H_i$ = In $m^N/N!$ (i.e., formally identical to $S_B$ except for the constant $k_B$ and its logarithmic basis) encoding the amount of information (OI) that the agent possesses about the positions of the N molecules of the system and dividing the volume of the vessel V into m cells, resulting in $m^N/N!$ possible configurations. In the same way Ben-Naim describes the amount of missing information about the molecular positions once the system reaches its equilibrium state by $H_f$ = In $(2m)^N/N!$ From these conceptual elements, the author describes informationally this process of expansion of a gas in a container as a difference of missing information positional $H_i$ - $H_f$ = N In 2. One could argue that it is not possible to powerfully explain the expansion of an ideal gas by means of Ben-Naim's informational description $D_x$ = 'the amount of missing information that an agent possesses about the position of an expanding gas doubles in N In 2 once the gas reaches the equilibrium state' precisely because, although it allows to specify (i) the initial condition of the system (via $H_i$ = In $m^N/N!$), it would not be possible to determine (ii) the dynamic properties that allow to validly abduce such behavior. As with the explanations of Brillouinism or

Jaynesianism, the lack of success and power of this syncretic explanatory line of Ben-Naim (2008, p.281) depends on certain inconsistencies and incoherencies that underlie his descriptions.

| | |
|---|---|
| Information Description $D_x$: | MI about micro-positions of a gas is doubled during $t_0$ - $t_1$' |
| Initial Condition (i): | The missing information at $t_0$ is MI($\Gamma_V$) |
| Dynamic Property (ii): | (non-conceptually specified) |
| Abducted Explanans $D_y$: | It is *plausible* that the gas expands to 2V at $t_1$ |

Initially, the lack of power of Ben-Naim's informational explanation rely on the fact that the concept 'missing information' is merely a terminological reformulation of the $S_B$ concept, whose only conceptual difference is the omission of the constant $k_B$. However, this deliberate omission of the units of energy and temperature would not help to specify the molecular-dynamic reasons underlying the free expansion of a gas. One could even argue that the suppression of $k_B$ in the missing information eliminates $S_B$'s own conceptual connection between observable quantities $A_i$ and microscopic properties (see Section 2.2), thus eroding the possibility of generating a powerful explanation on this descriptive tactic. As in the case of Brillouinian explanations descriptively based on $I_{Bb}$, Ben-Naim's descriptive tactic based on the notion of 'missing information' does not constitute a significant strategic advantage (although it does constitute a disadvantage) over other plausible non-informational explanations deployed on BSM, e.g., typicalist, ergodic, etc.). Because the conceptual relationship between Ben-Naim's missing information and $S_B$ is redundant and quasi-trivial, the only possibility for his descriptive tactic to contribute to his explanation of thermal phenomena depends on the interpretative contribution of the OI concept. Recall that OI underlie his statistical notion of 'missing information' identified with $H_{IT}$: "the measure of "information" as defined by Shannon also retains some of the flavor of the meaning of information as we use in everyday life" (ibid, p.17). But as with Brillouinian and Jaynesian explanations, the lack of technical specification of meaning and the interpretative ambiguity of OI imply that it is impossible to validly abduce any technically-defined dynamic property from this informational concept. Therefore, from recent informationalist proposals such as that of Ben-Naim (2008) it would not be possible to either successfully or powerfully explain paradigmatic thermal phenomena such as the expansion of an ideal gas.

One might ask whether the introduction of the concept of information K since the late 1980s by Zurek's proposal (Section 4.2) constitutes a significant advance in terms of explanatory power. As one can extract from the analysis in Section 2.3 and Section 5.2, the main tactical conceptual contribution made by K with respect to $H_{IT}$ is the semantic possibility of specifying individual objects referentially, which in the case of Zurek's (1990) interpretative proposal implies referring to individual microstates of molecular systems. At first sight, this contribution of K-information might suggest a better ability to informationally describe the dynamic evolution of individual microstates (in this case, the Hamiltonian of individual microstates). However, one can defend that this apparent conceptual virtue of information K with respect to information $H_{IT}$ does not result in a real explanatory advantage of K with respect to $H_{IT}$ in the domain of molecular thermophysics. The information concept K is conceptually unnecessary and insufficient for the generation of powerful explanations for the same reasons as $H_{IT}$. It is conceptually unnecessary because K does not contribute any further to that which $S_G$ (Jaynesianism) or $S_B$ (Brillouinism) contribute to the description of molecular-dynamic factors underlying thermal behaviors. Conceptually insufficient because since K is constitutively lacking notions like $k_B$, it is unable (with respect to $S_G$ and $S_B$) to consistently connect observational quantities $A_i$ and the behavior of sets of microstates.

### 5.3.6. Explanatory Powerless of Information Physics Proposals

Up to this point I have evaluated whether the main historical programs of thermophysical informationalism (i.e., Brillouinism, Jaynesianism, and their syncretic branches) can powerfully explain the thermal behavior of certain substances from the dynamics of their molecular components, such as the free expansion of a gas. Initially, the only distinctive character of informational explanations in thermophysics lies in their descriptive exploitation of informational concepts, all of them derived from OI, $H_{IT}$ and K (Section 2.3). Through the evaluation above, I have concluded that all these informational concepts cannot provide any powerful explanatory value in the domain of classical thermostatistical physics.

First, the ordinary non-technical concept OI claims central explanatory weight in informational thermophysics, either directly through its introduction into interpretive strategies linked to theoretical notions such as $H_{IT}$ (e.g., Ben-Naim 2008) or $S_B$ (Brillouinianism) or indirectly through epistemic probabilities (Jaynesianism). However, that an agent possesses OI about a system is neither sufficiently well-defined nor theoretically necessary to validly abduce the dynamic behavior underlying a given process. Briefly returning to the debate on the relevance of epistemic SM-probabilities (Section 2.3.3), I argue that modeling the informational (OI) dynamics of agents by means of credences, Humean laws or other forms of epistemic SM-probabilities may be explanatorily relevant to account (i) for what we should rationally believe or expect regarding the evolution of the system (as Uffink [2011] defends), but not directly (ii) for the dynamics of the system itself . Of course, one could indirectly infer certain properties of (ii) from (i). Even considering this restricted sense of 'explanation', the theoretically-underdefined concept of OI alone does not contribute to systematically connect epistemic SM-probabilities with the factors (initial state and dynamics) that would allow us to powerfully explain a thermal behavior.

Second, the concept $H_{IT}$ brings no additional explanatory value to its ability to conceptually emulate $S_G$ (or $S_B$) just as K brings no additional explanatory value to its ability to referentially specify individual objects based on $H_{IT}$. However, besides being unnecessary (i.e., due to the availability of BSM and GSM) and insufficient (i.e., due to their conceptual inability to consistently connect observational $A_i$ values and molecular properties) in generating powerful explanations, the notions $H_{IT}$ and K require, in order to explain, a surplus of structure extremely difficult to interpret or justify from the realm of thermophysics.

Additionally, all the thermo-informational explanations evaluated depend on exploiting (or 'epistemically parasitizing') to some degree the conceptual and technical-procedural resources underlying the explanations deployed from conventional or non-informational thermophysics (Anta 2021a). Illustratively, Brillouinian explanations of the approach to equilibrium depend constitutively on the meaning of conceptual elements (e.g., macrostates, phase volume, etc.) that allow to powerfully explain these same thermal phenomena from a Neo-Boltzmannian perspective. Similarly, possible Jaynesian explanations in non-equilibrium domains depend conceptually on the semantic consistency (or in this case, their presumed lack thereof) of Gibbsian-like procedures such as coarse graining. Therefore, due to the inability of informational concepts to contribute explanatorily, the possibility of information thermophysics (in its various traditions) to generate SM-powerful explanations ultimately depends on its parasitic exploitation of explanatory resources of conventional or non-informational thermophysical programs, like BSM or GSM.

In short, I argue that the main historical attempts to exploit information concepts to generate powerful explanations of thermophysical phenomena from the various informational programs have proved to be little (if at all) successful. But this historically contingent fact should not be taken as a destructive criticism, but rather allows us to propose two routes for the informational program to become a truly explanatory physics framework. On the one hand, from an informational framework it could be conceptually possible to develop technically precise concepts of semantic information capable of modelling the capacity of agents to epistemically interact with thermally behaving molecular systems. On the other hand, one could assume (e.g., from a Jaynesian point of view) that thermo-informationalism does not aim to be explanatorily powerful with respect to physical phenomena but with respect to the psychological phenomena that scientists experience when interacting with these phenomena. I will come back to this idea later on.

## 5.4. Informational Predictions in Statistical Thermal Physics

In addition to explanations, other central ways of obtaining knowledge in physics are through the prediction of the evolution of certain properties. Thus, another of the main epistemic virtues of thermophysics is its potential to predict thermal patterns from the evolution of their observable quantities (within TD) or the dynamics of their particles (within SM). Illustratively, from a Gibbsian framework one could predict the future velocity distribution of the molecules of an expanding ideal gas inside a thermally insulated container. To restrict the scope of my argument I will focus on those predictions that we may encounter in non-equilibrium contexts, illustratively the free expansion of a gas in the volume of its container. My task in this section consists precisely in carefully evaluating whether we can obtain meaningful (or 'advantageous') knowledge about certain thermo-macroscopic behaviors from the predictions of the main informational traditions.

### 5.4.1. Framing Predictions and Predictive Advantages in Statistical Mechanics

The question about predictions in the SM theoretical domain has been exhaustively studied from the philosophy of thermophysics, playing a decisive role in a vast plurality of central debates such as that of time direction, the asymmetry between prediction and retrodiction or the Past Hypothesis (Sklar, 1999; Albert, 2000; Frigg, 2008). Although historically the general philosophy of science has consistently focused its attention on the ability of predictions to empirically support (if advantageous) the content of a theory, here I will focus on the ability of different predictions generated from a research program (i.e., informational thermophysics) to provide us with meaningful knowledge of a particular domain. A decisive characteristic feature is that, unlike SM-explanations (which do not necessarily rely on probability notions, as in the case of typicality-based explanations, see Section 5.3), SM-predictions are constitutively statistical in that their underlying descriptions depend on technically exploiting probabilistic resources[123] (Section 2.4).

However, that a prediction can be computationally feasible (and subsequently observationally testable) does not imply that these predictions will provide us with meaningful knowledge about the phenomenon described from a particular descriptive tactic. In order to assess below which descriptions allow us to generate epistemically relevant predictions in the field of statistical thermophysics I propose (in parallel with explanations) an again 'minimalist' criterion of predictive advantage. Unlike the assessment of the power of particular informational explanations, the analysis of informational predictions will be carried out comparatively (i.e., with respect to their virtues vis-à-vis non-informational predictions). This analysis of predictive advantages has as a prerequisite that the future values calculated from a certain statistical description based on molecular properties or phase space can be empirically verified, but it must also ensure that the observable values depend on or follow validly from the content of the description. This criterion could be formulated as follows.

[123] In order to be able to predict the values observationally that the system will take at future times of dynamical evolution in a 'non-probabilistic' way it would be necessary to determine the number of microstates contained in the macrostate $\Gamma_M$ macroscopically accessible for us at the beginning of its dynamical evolution. This would only be a computationally exhaustive task (over $10^{23}$ position-velocity values per individual microstate $x$) and analytically intractable, since the number of microstates in $\Gamma$ to be determined is continuously infinite. Although with an ideal number of resources we could solve this problem (e.g., by numerical discretization techniques or partitions), we would still have to solve approximately $6 \times 10^{23}$ equations, six equations of motion associated with each individual microstate trajectory. However, as Shenker reminds us "This is an idealization, and such calculation is doubly impossible: the system is too complex, and the number of microstates, and hence of trajectory segments, is a continuous infinity" (Shenker 2020, p.11). Thus, the introduction of probabilistic notions (either static SP or dynamic DP) into the conceptual strategies of thermophysics becomes an indispensable technical requirement for the deployment of effective predictions of thermal patterns from the statistical modelling of their molecular dynamics.

**Criterion of Predictive Advantage:**

> An informational description $D_A$ that allows describing the future evolution of a thermal behavior enables an advantageous prediction (as opposed to non-informational descriptions $D_B$) if (a) the empirically measurable values $A_i$ that macroscopically describe such behavior follow validly from information concepts in $D_A$, and (b) if $D_A$ is more pragmatically and epistemically virtuous than $D_B$ in obtaining predictive knowledge.

Like the minimalist criterion of explanatory power, this new minimalist evaluative framework of predictive advantage (applicable from any strategic, BSM, GSM, or informational framework) is also deployed on a descriptive tactic that allows describing thermal behaviors, and depends on its 'inferential validity', so that the content of future observational values follows correctly and consistently from the exploited notions. To evaluate its performance, let us apply this criterion to the already paradigmatic case of an expanding gas. Let us suppose that we have a description of the initial observational state of the gas (i.e., located in the right-hand-side subvolume $V_R$ of the container) by a Lebesgue-uniform distribution over the macrostate-region $\Gamma_{M0}$. From this descriptive setting one could predict the macroscopic values of the expanding gas at time $t_1$ by exploiting Lioville's theorem and introducing DP dynamic probabilities (Section 2.4), so that we predict that the probability of the gas being in the volume associated with A1 is P(A1).

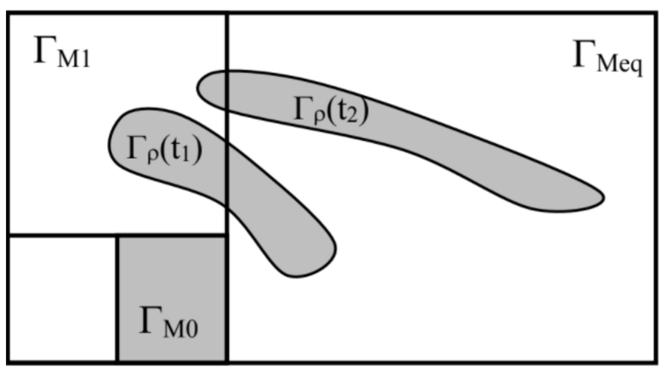


**Figure 6**. Phase portrait of a gas expanding in a vessel 2V. Macrostate-region $\Gamma_{M0}$ represents the initial condition of the gas, e.g., trapped in the right half of 2V. Once the partition is removed, the microstates in $\Gamma_{M0}$ dynamically evolves as a dynamical blob $\Gamma_\rho(t_1)$ in the phase space $\Gamma$ of the gas-box system. A similar diagram can be found in Hemmo and Shenker (2012, Fig 6.2).

Does this constitute an advantageous prediction (according to the above criterion) as opposed to, for example, a non-statistical prediction? Firstly, the fact that the probabilistic value P = 0.52 of $A_i$ (i.e., defined as $P[\Gamma_\rho(t_1)]$) is validly inferred from the description $D_A$ employed, including the dynamic blob, the macrostate $\Gamma_{M1}$ and the dynamic probabilities DP. The validity of this predictive-statistical inference is determined by the fact that the predictive potential of $P[\Gamma_\rho(t_1)]$ follows (i) in a semantically consistent way from the physical significance of the macrostate-region

$\Gamma_{M1}$ and the density ρ, and (ii) in a syntactically correct way according to the procedure set by the macrostate-blob overlap via dynamic probabilities (i.e., $\Gamma_{\rho}(t_1)$ )$\Gamma_{M1}$). Secondly, this description $D_A$ brings significant epistemic and pragmatic virtues with respect to another fine-grained description $D_B$ derived from partitioning $\Gamma_{Mi}$ into non-observationally distinguishable sub-cells $\Gamma_{Ni}$. $D_A$ is more epistemically virtuous than $D_B$ because $\Gamma_{Mi}$ represents experimentally relevant values of observable quantities $A_i$ and $\Gamma_{Ni}$ represents arbitrary values of those quantities $A_{i.}$ $C_A$ is also more pragmatically virtuous than $D_B$ because predictive exploitation of cells $\Gamma_{Ni}$ is much computationally expensive than the exploitation of microstate-regions $\Gamma_{Mi}$. Thus, a prediction that validly relies on description $D_A$ (illustrated in Figure 6) will constitute an advantageous prediction over strategies that can be deployed from description $D_B$. Once these evaluative resources have been exposed, my next objective is to apply them to the main predictions found in the history of informationalism in statistical thermophysics, evaluating the quality of the predictive knowledge they provide us with.

### 5.4.2. Predictions in the Jaynesian Tradition

The aim to exploit informational concepts to technically improve (or conceptually justify) the predictive capabilities of thermophysics arose in the context of the early 1950s with the emergence of Shannon's bandwagon (Section 3.1). However, this aim was historically consolidated with the Jaynesian informational proposal articulated on MEP (Section 3.3) to reconceptualize SM as a discipline whose central epistemic virtue is its ability to predict, understanding its SM-framework as a 'predictive statistical mechanics'. In order to evaluate the success of Jaynesian predictions, I should first methodically analyse the MEP-dependent descriptive tactic that allows us to describe the future observational values of the modelled system[124].

Let us assume as an illustrative applicative scenario wherein someone intends to predict Jaynesianly the macroscopic volume at $t_1$ of a freely expanding gas from a subvolume V. First, we would tactically generate a uniformly distributed density ρ on Γ with respect to V (as a macroscopic constriction CV) that describes the initial state of the gas at $t_0$ during this expansion process. Next, ρ can evolve via Liouvillean dynamics through the phase space Γ of the gas up to $t_1$. Subsequently a Jaynesian coarse-graining of the description ρ is performed, deriving a new coarse-grained description $\bar{\rho}$ that maximizes the $H_{IT}$ values at $t_1$ and from which the probability of the gas being characterized by the volume 2V is predicted to be P(2V) = 0.82. One might wonder now whether this type of Jaynesian prediction is or can be advantageous over non-informational predictions, paradigmatically Gibbsian predictions. However, besides its indisputable potential to

[124] Recall that the conceptual architecture of MEP (Jaynes 1957a) is articulated on the derivation of a microstatistical description (i.e., density over Γ) that maximizes $H_{IT}$ information, being interpretatively justified as follows: a MEP-derived density ρ imply a minimal commitment to the information (OI) that an agent lacks about the actual microstructure and with respect to its constraint-represented observational knowledge C of the system (Section 3.6). It is precisely this descriptive tactic from which Jaynesian predictions can be generated.

generate empirically correct predictions, my main goal here is to assess whether (and how) the Jaynesian exploitation of information concepts contribute to make these predictions successful.

First, one could argue that the success of Jaynesian predictions depends conceptually on the fact that in assigning prior or initial probabilities $\rho$, maximizing $H_{IT}$ consistently corresponds to a commitment to the agent's lack of information (OI) about the molecular properties of the system. However, (as I have suggested in the previous evaluation) the fact that the probability of the microstate of a gas being at region $\Gamma_{2V}$ (i.e., representing the gas occupying a volume of 2V) is $P(\Gamma_{2V}) = 0.82$ cannot validly follow from the agent lacking information about the actual microstate, due to the referential ambiguity and technical indefiniteness of OI (via epistemic probabilities). That is, it is completely underdetermined whether the amount of information represented by $P(\Gamma_{2V}) = 0.82$ or $H_{IT}(\Gamma_{2V})$ is (i) sensitive or not to measurement instruments (i.e., how much the agent's epistemic access to the actual microstate is mediated), (ii) sensitive to the knowledge of the dynamics (e.g., if the agent knows that the system is Hamiltonian or not), (iii) possessed by an individual agent or by an epistemic community, and so on. Then, the everyday concept of information (OI) would not contribute to increase the predictive power that the epistemic Jaynesian interpretation of $H_{IT}$ and SM-probabilities might have. Recall that my task is not to evaluate whether MEP is a good predictive tool or not (in fact, it is an excellent one) but whether the Jaynesian exploitation of information concepts truly contribute to this predictive success.

Thus, one can think that the contribution of information concepts to the success of these Jaynesian predictions comes from their technical refinement and semantic consistency with $H_{IT}$. Otherwise, the $H_{IT}$ notion only contributes (with respect to its formally analogous concept $S_G$) the intentional omission of any conceptual connection between macroscopic properties $A_i$ and microstatistical ones via $k_B$. As Uffink (1995, p.233) argued, $H_{IT}$ is not necessary (or the unique method[125]) for MEP to successfully predict macroscopic $A_i$ values, since the consistency requirement of MEP (see Section 3.3) could also be satisfied by maximizing formally analogous quantities, such as $S_G$ or Renyi entropy[126]. In this sense, the Jaynesian-MEP algorithm of prior probability assignment would prima facie allow predicting the same empirically-correct numerical results as the Gibbsian (1902) algorithm of prior probability assignment for microcanonical ensembles (technically dependent on $S_G$ maximization instead of $H_{IT}$), although in a much theoretically simpler fashion. That is, the part of the conceptual architecture of Jaynesianism that would allow explaining the predictive empirical success (according to this evaluation) of this theoretical framework does not depend on any informational concept at all, either interpretively from OI or formally from $H_{IT}$. Thus, even when acknowledging that the Jaynesian program allows us to generate successful

[125] Interestingly, the uniqueness theorem for $H_{IT}$ was disregarded even for Shannon himself: "This theorem and the assumptions needed for its proof *are in no way necessary* for the present theory. It is given chiefly to lend a certain plausibility to some of our later definitions. The real justification of these definitions, however, will reside in their implications (Shannon, 1948, p.393. Italics are mine).

[126] Briefly, Renyi entropy $S_R(\rho) := 1/\ 1-q \log \sum_{k=1} \rho^q$ is a q-order mathematical generalization (developed out of the theoretical context of communication theory) of Shannon entropy that does not satisfy the branching axiom (Section 2.3.3), see Frigg and Werndl 2011a p.121-122.

predictions in a more pragmatically simpler way than GSM, it should be argued that this predictive success seems to be completely independent of the informational character of their descriptions[127].

### 5.4.3. Predictions based on Algorithmic Information

I have just concluded from the previous assessment that the non-technical concept OI and the theoretical $H_{IT}$ do not epistemically contribute to the generation of successful predictions from the Jaynesianism framework. As argued in Section 4.2, a certain sector of traditional Jaynesianism was intellectually transformed since the 1990s by the theoretical proposal of Caves (1990), incorporating the notion of algorithmic information K for the description of individual microstates *x* of the target system. Following the thermophysical-algorithmic framework put together by Zurek (1990), Caves suggested that (i) the boundary condition and (ii) the Hamiltonian dynamics of each system (i.e., basic conceptual elements in any SM-description) are algorithmically simple, due to the logarithmic number of bits required to specify them. For the case of a microcanonical ensemble defined uniformly over the phase region $\Gamma_{Ri}$, Caves states "To specify a typical microstate *j*, one gives the length of the specifying string [$r_i$], which requires [log log $\Gamma_{Ri}$] bits, and then gives the entire string [$r_i$], which requires [log $\Gamma_{Ri}$]" (ibid, p.105). Thus, the number of bits required to specify a microstate within a microcanonical ensemble is as follows.:

$$I(r_i) \approx \log \Gamma_{Ri} + \log \log \Gamma_{Ri} = S_G(\Gamma_{Ri}) + \log S_G(\Gamma_{Ri}) \quad (1)$$

This allows to conceptually reinforce the prior probability assignment algorithm provided by the Jaynesian-SM, noting certain practical limits "any probability assignment for a macroscopic system must have algorithmic prior information much smaller that the algorithmic information of a typical microstate; otherwise, no practical memory could store the required algorithmic prior information" (ibid., p.106). However, one might now ask whether this concept of algorithmic information K contributes to the generation of advantageous predictions as opposed to predictions derivable from BSM or GSM. In a first pass, this contribution could not be understood in terms of the immense number of computational resources required by algorithmic descriptions of the microstate molecular systems (which are super-logarithmic and only well-defined for microcanonical assemblies) as opposed to Jaynesian informational descriptions. The exploitation of these algorithmic descriptions à la Zurek-Caves turns the prediction-SM of macroscopic $A_i$ values into a task either computationally infeasible (in the case of macrocanonical ensemble) or exhaustive in terms of algorithmic resources (e.g., time, memory, etc.); which constitutes a significant practical disadvantage compared to the predictive efficiency of GSM or BSM.

[127] In fact, a plausible non-informational reformulation of Jaynes MEP could be given. For instance: the most appropriate distribution ρ to represent our inability to determine the actual microstate of the system is the distribution ρ maximizing the Gibbs entropy $S_G$ or Renyi entropy $S_R$ of the set of microstates $\Gamma_R$ compatible with our observational knowledge of the system.

On the other hand, I would argue here that algorithmic information K is not only epistemically unnecessary and insufficient to explain thermal behaviors (Section 5.3), but also epistemically unnecessary (i.e., because it serves the same function as $S_G$ or $S_B$) and insufficient (i.e., because of its inability to systematically connect observables $A_i$ and sets of microstates) to statistically predict the evolution of macroscopic quantities for precisely the same reasons. Thus, even if the algorithmically-enhanced Neo-Jaynesian descriptions of Caves allowed to calculate the number of bits required to specify individual microstates *x*, this would not necessarily imply that the eventual prediction of thermo-empirically meaningful values is validly derived from the 'contingent compressibility' (description-dependent) of phase regions and not directly from the molecular patterns underlying certain thermal behaviors. In this sense, the concept of algorithmic information K would not epistemically contribute to the generation of advantageous predictions over non-informational ones, at least as K has been historically exploited from the thermo-informational tradition around Zurek (1990).

### 5.4.4. Recent Predictions in Syncretic Informationalism

Next, I argue that recent informational predictions (again I take Ben-Naim [2008] as referent) are not only not epistemically advantageous over non-informational ones, but also generate conceptually incorrect and interpretatively incoherent predictions. Let us focus on how Ben-Naim (ibid., p.265-268) predicts the evolution of macroscopic quantities $A_i$ in a coupled assimilation process with expansion, where an ideal gas composed of identical molecules $N_A$ is initially uniformly distributed over a container whose volume 2V is separated by a partition into two contiguous equal volumes $V_A$ and $V_B$. After removing the partition, the process of assimilation and expansion of the gas begins, where finally the $N_A$ gas molecules will entirely fulfil 2V.

Ben-Naim's theoretical treatment of this assimilation-expansion process constitutes an informational redescription of the famous Gibbs paradox (1902, p.167), where the mixing of two gases increases their entropy even when their components are identical and whose conceptual foundations have been exhaustively treated in the literature (e.g., Denbigh & Redhead, 1989; Sklar, 1993; Ainsworth, 2012). This author argues that, in the case where molecules identical at the onset were differently labeled as $N_A$ and $N_B$, one could predict which assimilation-expansion process would result in an increase of N In 2 MI (computed via Sackur-Tetrode formula, see Ben-Naim, [2008, p.266]). On the other hand, from the initial assumption that the molecules are identical $N_A = N_B$ would predict that the MI does not change during this same process. From here, Ben-Naim infers from MI-dependent tactics that it is properly the assimilation of $N_A$ molecules into the gas composed of $N_B$ molecules (and vice versa) that cause a decrease in the missing information (OI) of the agent on the microstate of the system, so that the expansion should increase its MI. However, from here it can be predicted that the missing information that an agent possesses would increase after the disassimilation of a gas (i.e., separation of $N_A$ and $N_B$ into two separate volumes $V_A$ and $V_B$, respectively), which is in direct contradiction with the fact that the information (OI) that the agent possesses about the molecular positions on $N_A$ and $N_B$ evidently increases due to their macroscopic distinction into $V_A$ and $V_B$. Therefore, from Ben-Naim's descriptive tactic it would be possible to derive invalid predictions of certain processes (e.g., assimilation plus expansion) regarding the meaning of the concepts (mainly MI) used to SM-describe such processes.

This type of informational predictions is not only deployed on semantically inconsistent descriptions, but also generate remarkable inconsistencies within their interpretative scaffolding, as Bar-Hillel himself pointed out "Finding that assimilation decreases the MI (or the entropy) is even more shocking than finding that the mixing does not change the MI (...) it is hard to accept and harder to visualize why the deassimilation increases MI and assimilation decreases MI" (Ibid, p.276). As seen in this quote, the author confesses that the prediction that the processes of deassimilation or separation would increase the lack of information (OI) (when intuitively one would expect the opposite) is presented as inconsistent with respect to his interpretative strategy, whereby the meaningful content of $S_B$ would ultimately be reduced to how 'In W' represents the agent's lack of information (OI) about the microstate of the molecular system. This constitutes a mere informational reformulation of the Gibbs paradox, so that in the macroscopic separation of two sets of $N_A$ and $N_B$ molecules the agent's information would increase (i.e., such separation would grant more knowledge to the agent) and decrease (i.e., how it is predicted from MI) simultaneously. With the pretense of solving this 'informational' interpretive paradox, Ben-Naim (ibid., p.271-284) introduces an ad hoc distinction between 'real' and 'non-real information (OI)' such that "we also gained real information by the assimilation process" (ibid., p.281).

I now argue that this ad hoc interpretative adjustment is only a distraction from one of the main conceptual shortcomings of Ben-Naim's 'missing information': namely, that MI values are (numerically and arithmetically) insensitive to thermophysically significant changes of macroscopic properties $A_i$, mainly volume V and amount of substance (Section 5.3). In this sense, Ben-Naim's informational architecture is not only unnecessary and insufficient to generate advantageous predictions (e.g., against GSM) of macroscopic behaviors. Also, the predictive knowledge we would obtain along these lines (e.g., our positional information (OI) about this system will decrease in $t_1$) is irrelevant from a strictly TD-phenomenological point of view. Thus, from Ben-Naim's informational descriptive tactic, not only would we generate predictions that are largely disadvantageous with respect to non-informational tactics (due to the insensitivity of the concept MI to macroscopic properties $A_i$), but it would also allow us to make predictions based on semantically and interpretatively incompatible descriptions.

### 5.4.5. Predictive Disadvantages of Information Physics Proposals

So far, I have just analysed the epistemic capacity of certain key informational descriptions (i.e., Jaynesianism, Caves or recent informationalism) to obtain meaningful predictive knowledge in the field of thermophysics. The first point that should be stressed in this direction is the inability of the various informational thermophysics approaches to generate what are known as 'novel predictions' in the domain of classical statistical-molecular physics. Illustratively, the Ben-Naim theoretical program developed during the first decades of the 21st-century claims to predict exactly the same range of molecular phenomena that Boltzmann or Gibbs claimed to predict statistically at the end of the 19th-century: the adiabatic expansion of a gas, the mixing of two fluids or the changes of state of substances. In this sense, informational thermophysics would not extend the domain of predictable phenomena with respect to BSM or GSM, which is prima facie disappointing with respect to other fields of contemporary physics[128] (as general relativity extends the domain of phenomena with respect to Newtonian mechanics by allowing the prediction of the orbit of mercury). However, one might think that predictions deployed on informational descriptions are advantageous over those deployed on non-informational descriptions.

From this systematic evaluation of predictions, I conclude that the main informational concepts (OI, $H_{IT}$, and K, Section 2.3) have not provide any significant epistemic advantage in their exploitation to generate predictions. Predominantly, Jaynesian-MEP proposes that assigning a density that numerically maximizes $H_{IT}$ information (interpreted as the agent's lack of information [OI] about the actual microstate) can be predictively advantageous over other algorithms in SM. Firstly, the vague epistemic connotations that $H_{IT}$ carries as a measure of uncertainty (OI) of the agent do not play any role in the computation of the dynamic DP probabilities that allow to make SM-predictions, but only serve the purpose of justifying[129] the use of this statistical-inference algorithm. As for instance the white-crow example (Aczél & Daróczy, 1975) showed (see Section 5.2.2), by obtaining the information (OI) that an unobserved crow has a white mother, the $H_{IT}$ entropy (interpreted as 'uncertainty') about how probable is the color of the crow will increase. This simple illustrative example shows not only that $H_{IT}$ as a measure of uncertainty (OI) does not contribute at all to the predictive efficiency of MEP, but that it generates important interpretative incoherencies, as Jaynes himself once suggested: "This paradox shows that 'information' is an unfortunate choice of word to describe entropy expressions" (Jaynes, 1957b, p.186).

Secondly, the statistical properties of $H_{IT}$ do not provide any significant predictive advantage that was not already provided de facto by its structurally analogous $S_G$ or Renyi entropy, since both measures can equally satisfy the 'consistency requirements' underlying MEP (i.e., [1] $H_{IT}$ is a continuous over $\rho$, [2] for all $n$, $H_{IT}(p_1,\ldots,p_n \in \rho)$ is symmetrical in $p_n$, and [3] for all $n$, $H_{IT}$ has

[128] As detailed in Chapter VII, Lakatos (1978, p.49, 52-54) argued that the inability of a scientific research program to predict new phenomena within a particular domain is a clear sign that the program in question is not 'theoretically progressive', and that it should consequently be regarded as a degenerate program.

[129] As Uffink (1995, p.232) pointed out: "The justification of the MEP is then that the maximum entropy distribution is the distribution that correctly corresponds to a maximal amount of uncertainty [OI]. It represents the only probability assignment that is 'maximally noncommittal with regard to missing information', i.e. that, while obeying the constraint, does not assume any information which we actually do not have." Also in the words of Frigg: "Why is MEP compelling? The intuitive idea is that we should always choose the distribution that corresponds to a maximal amount of uncertainty, i.e. is *maximally non-committal with respect to the missing information* [OI]. But why is this a sound rule? In fact MEP is fraught with controversy; and, to date, no consensus on its significance, or even cogency, has been reached." (Frigg 2008, p.91. Italics are mine).

the form $-k \Sigma p_i \log p_i$ for some positive constant $k$, see Section 3.3.3). On the contrary, there is no straightforward way of extending $H_{IT}$ into the Γ-continuum since "more mathematical ingredients are needed in order to obtain a meaningful measure of information for continuous probability distributions than are mentioned in Shannon's formulation" (Uffink 1995, p.234). This means that $H_{IT}$ is not the only measure that allows MEP to be an effective SM-probabilities assignment algorithm[130]. Then, discarding the epistemic contribution of OI and $H_{IT}$, the predictive success of MEP would ultimately depend on its procedural differences with respect to other predictive devices in SM like the Gibbsian algorithm.

Additionally, the specification of individual microstates by information K (Caves 1990) would also not allow generating more efficient or successful predictive mechanisms than the conventional ones, but extremely more computationally costly. In fact, as in the case of Ben-Naim (2008), the descriptive exploitation of informational concepts can easily lead to semantically inconsistent and interpretatively incoherent predictions, such as that in the separation of two sets of molecules our information about molecular positions increases and decreases simultaneously. Finally, although Jaynesianism (as opposed to other informational programs) has historically proven to be a more instrumentally advantageous predictive framework than GSM, here I conclude that the informational concepts originally employed by Jaynes (1957a) (Section 3.3) can only be considered as heuristic devices to introduce his proposal into the scientific community. The enormous predictive success of MEP is due to MEP being a powerful statistical inference algorithm in SM and not to the informational concepts he used to launch his theoretical proposal into Shannon's bandwagon.

## 5.5. Informational Interpretations of Statistical Thermal Entropy

Along with explanation and prediction, another of the main sources of knowledge in thermophysics is provided by the understanding of phenomena and concepts within this field. Such understanding could be obtained from several intertwined lines, essentially (i) from certain explanations that make intelligible (see, de Regt 2017) the dynamic factors causing a process or (ii) from different interpretative strategies. In this section I focus on assessing the capacity of a given interpretation (informational, in this case) of a concept to cognitively enable us to understand the meaning of this concept in a more fruitful way than from its non-interpreted content. In particular, I focus my analysis on the capacity of different informational interpretations of entropy concepts to provide us with a meaningful understanding of these concepts. In order to analyse how

[130] In this regard, Uffink argues that "the strategy of justifying the MEP by means of Shannon's uniqueness theorem seems to fail. First of all, the assumptions on which the theorem rests are not all compelling, but contain also conventional elements. Secondly there is still the problem of extension to the continuum. Thirdly there are examples showing that the entropy expression has properties which do not correspond completely to what one would intuitively expect of an information measure." (Uffink 1995, p.234).

these interpretative strategies might provide us with a robust understanding and make intelligible the various aspects of the meaning of the entropy concept, in this section I will assess how an interpretation can achieve this epistemic function.

### 5.5.1. Framing Interpretative Strategies in Thermophysics

As I have explored in the historical analysis (Chapters III and IV), the diachronic unfolding of informational thermophysics is replete with a vast plurality of interpretative strategies of information concepts based on exploiting information notions.[131]. In this section I will focus on those interpretive strategies in the history of this intellectual program (i.e., based on informational descriptions) that aim to elucidate the meaning of thermophysical entropy. The question of whether informational interpretations provide intelligibility in the conceptual landscape of entropy has been addressed only during the last decades within the literature, e.g., "the notion of information may be helpful in promoting good intuitions about the notions of entropy" (Shenker 2020, p.3) or illustratively "The missing information metaphor is well defined and can be quite useful, especially in understanding that descriptions of macroscopic matter necessarily discard enormous amounts of information about system details, working ultimately with a small number of macroscopic variables such as pressure, volume and temperature" (Leff 2007).

Note at this point that the question is not properly whether entropy concepts have any explicit syntactic or semantic connection to some notion of information, as has been widely discussed in the literature (e.g., Frigg and Werndl, 2012; Shenker, 2020). Otherwise, the focus should be on whether the multiple resources linked to a given interpretation of a concept (e.g., metaphors, analogies, etc.) can satisfy the epistemic function of providing understanding of the content of that concept by increasing its intelligibility. But for this understanding to have robust epistemic value it must be more than a mere subjective sensation from the agent (see Trout 2002); that is, such an understanding must enable the agent to employ the meaning of the concept in an inferentially valid way (see footnote 122). In this direction, I propose here to extend and to some extent reformulate the model of "intelligibility" (i.e., as pragmatic factors that make their understanding possible) of scientific theories developed by de Regt (2017, Sect. 4.4) to the case of scientific concepts, to develop what I would call ‘Criterion of Interpretive intelligibility’.

[131] Here, one should properly differentiate between 'interpretive scaffolding' (i.e., representational resource that allows us to specify the meaning or referential mechanism of a concept) and 'interpretive strategy', as the descriptive exploitation of the meaning of a concept to enhance its understandability. Interestingly, this distinction between 'interpretive scaffolding' and 'interpretive strategy' corresponds to the two main functions of Fregean 'sense': namely, the capacity to specify the reference of a concept (e.g., as $H_{IT}$ allows us to specify the reference of $S_G$ within Jaynesianism) and the cognitive value that a concept brings (e.g., as OI helps us to understand the meaning of $S_B$ in Brillouinism), respectively. In this sense, the ability of interpretive strategies (as opposed to explanatory or predictive ones) to endow the content of a concept with intelligibility depends contextually on the cognitive capacities of the agents and not only on the content itself (de Regt 2017).

**Criterion of Interpretative Intelligibility**:

> An interpretation of a thermophysical concept $C_{TD}$ based on an informational description $D_A$ will be an intelligible interpretation for a competent agent about the content of $C_{TD}$ if only such a $D_A$-driven intelligible interpretation provides the competent agent with the cognitive capacity to qualitatively exploit $C_{TD}$ in an inferentially valid way.

Again, note that this minimalist criterion does not impose a sufficient condition for a given interpretation to contribute to the understanding of a concept, but merely a necessary condition that ensures that the understanding obtained can contribute to gaining knowledge about the phenomenon to which the concept refers. For example, the fact that certain thermodynamic quantities can be reconceptualized from the movement of particles makes it possible for agents to understand certain thermal phenomena such as the free expansion of a gas by visualizing (as a cognitive resource that makes inferential exploitation possible) the dynamics of its molecular components, which allow them to predict that the molecules will eventually disperse throughout the volume of the container: "The general picture of moving gas particles allows us to make qualitative predictions of macroscopic properties of gases in particular situations. The kinetic theory has qualities that provide tools which can be employed by physicists who possess the relevant skills in order to gain understanding of the behavior of phenomena in its domain" (from Regt 2017, p.106)[132]. Finally, that an agent can make valid inferences (see Section 5.3) about the interpretatively specified content of a concept C properly means that the content of the inferred follows semantically from the content of C.

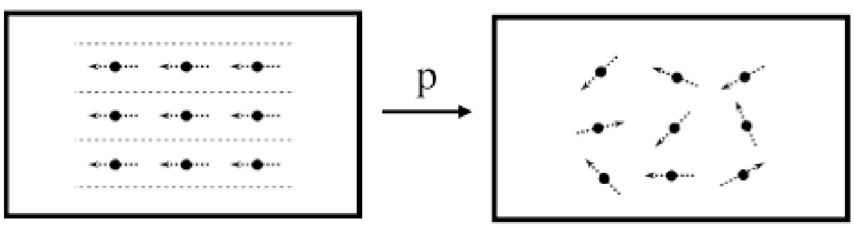


**Figure 7**. Adiabatic Demagnetization of a Gas.

Let us illustrate it with its application in the case of the interpretation of entropy as disorder. Suppose that a (minimally competent) agent adopting such an interpretation of the meaning of entropy is confronted with the concrete phenomenon of adiabatic demagnetization of a gas the agent will be able to infer (without making quantitative calculations), just by visual recognition of regular patterns in the spatial configuration of the gas molecules that the entropy of the gas will increase by eliminating the magnetic field passing through it. Such a prediction relying on disorder-based interpretation will be incorrect, because the $S_{TD}$ of the gas does not increase, but

[132] Following this line defended by de Regt, whether the intelligibility of $S_B$ or $S_G$ can provide comprehensibility to $S_{TD}$ ultimately depends on the non-interpretive conceptual connection between both notions (more on this point later).

remains constant throughout the demagnetization process. However, the key here derived from this criterion is that any conceptual possibility of inferring (e.g., statistical, explanatory-abductive, predictive, etc.) from this interpretation of the meaning of $S_{TD}$ will not result in an inferentially valid process, regardless of whether certain occasions may result in correct inferences. Thus, the disorder-based interpretive strategy constitutes an interpretation of the concept of entropy that is not only epistemically unfruitful, but also not robust in the sense already specified. To this point, I must supplement these analytical tools with a minimal desideratum for understanding the meaning of scientific concepts: namely, that the concepts $C_i$ upon which an interpretation of concepts $C_f$ is deployed be more (or at least equally) intelligible than the content of the interpreted concept for as many cognitive profiles of agents as possible[133]. Now that I have clarified what a robust and intelligible interpretation of a concept consists of, let us apply it to the case at hand.

### 5.5.2. Informational Interpretation A: Entropy as Lack of Microstatistical Information

> "Entropy and information are very closely related. In fact, entropy can be regarded as a measure of ignorance. When it is known only that a system is in a given macrostate, the entropy of the macrostate measures the degree of ignorance the microstate is in by counting the number of bits of additional information needed to specify it, with all the microstates treated as equally probable" (Gell-Mann 1994, quoted on Ben-Naim 2008)

One of the main ways of understanding entropic concepts from informational notions is through a 'complementarist' strategy of assimilating entropy as a measure of the uncertainty or properly 'lack of information' (OI) possessed by the observer about the system's microstate[134]. But does the understanding of $S_B$ or $S_G$ entropy as a measure of the agent's uncertainty about the actual microstate constitute a robust interpretation of this concept? Let us take as an illustrative case the free expansion of a gas from a subvolume V to the full volume 2V of its container. Suppose now that someone adopting interpretation A aims to inferentially exploit his uncertainty regarding the actual microstate of the gas at $t_0$ given a set of macroscopic initial conditions (e.g., initial volume V), proportional to the volume of microstates $\mu(\Gamma_M)$ compatible with V. This thermophysically competent agent will predict (Section 5.4) qualitatively according to the evolution of the gas that its uncertainty will increase and that the microstates on which this ignorance is based will change, because the actual microstate *x* will transit through macrostates-regions $\Gamma_{Mi}$ (i) with a larger number of microstates and (ii) different from the initial $\Gamma_{M0}$, correspondingly.

---

[133] Therefore, an optimal interpretation or the 'best interpretive system' in Lewisian terms of a thermophysical concept such as entropy should combine in a balanced way: (a) sufficient technical specificity to make valid inferences (satisfying our criterion), with (b) the cognitive advantages provided by the conceptual resources of an interpretation highly intelligible even to less theoretically competent agents.

[134] As explored in Section 3.1, this line of interpretation historically predates Shannon's bandwagon in the 1950s, the quote by Lewis (1930) or Wiener's proposal (1949) during the 1940s being some of the most illustrative examples.

This A-based predictions are essentially correct according to BSM, and its phase-geometric intelligibility (i.e., uncertainty proportional to the macrostate measure $\mu[\Gamma_M]$) may be fruitful for understanding the operation of statistical descriptions of an approach to equilibrium. However, through the informational interpretation A of the meaning of $S_B$ or $S_G$ it would be impossible for the agent to semantically derive (and thus also not understand) that this increase-displacement of the agent's uncertainty linked to the gas expansion necessarily entails an increase of $S_{TD}$ in the environment of the gas-box system. Therefore, even if A is an intelligible interpretation of the character of $S_B$ or $S_G$ as a system microstates counter (i.e., expressed as '$\log_e$ W'), this would specify the content of this concept in a way that makes it insensitive to a $S_{TD}$ production (i.e., encoded in Boltzmann constant '$k_B$'): "[Jaynesian A interpretation of $S_G$] does not claim to make direct contact with the thermodynamic entropy as it is usually conceived in thermodynamics, this is an inevitable feature of such interpretations" (Parker 2011, p. 855 ). If one assumes that the validity of any inference derived from $S_{TD}$ depends on interpretatively specifying the thermophysically meaningful relationship between volume of microstates $\Gamma_M$ and $S_{TD}$ production, the interpretation A of entropy as ‘missing microstatistical information’ will lack intelligibility.

### 5.5.3. Informational Interpretation B: Entropy as Microscopic Indistinguishability

While the interpretative strategy A can be applied to any thermo-statistical concept of entropy, there are other types of interpretative lines applicable exclusively to $S_B$ or $S_G$. This is the case of the epistemic and 'complementarist' interpretation of the $S_B$ concept as a measure of the agent's inability to distinguish between distinct molecular microstructures, originally developed by Boltzmann (1896) and reconceptualized within the Brillouinian tradition by relying informational concepts (Section 3.2). This interpretation B tries to specify the meaning of $S_B$ through the microscopic resolution capacity of the observer to distinguish the microscopic configurations of the observed system. "This lack of information introduces the possibility of a great variety of microscopically distinct structures, which we are, unable to distinguish from one another" (Brillouin 1962, p.160). From interpretation B, the agent would understand the entropic behavior of the system in such a way that the observer's ability to distinguish the actual microscopic configuration (among a group of macroscopically indistinguishable configurations) decreases during this expansion process. Suppose that the agent exploits this description to predict qualitatively (infer without the need for quantitative calculations) that the gas will generate entropy $S_B$ during its expansion in the vessel[135]. Such a prediction is not only correct, but will constitute a valid inferential process if the agent has derived such a conclusion through its semantic

[135] The qualitative reasoning underlying such a prediction would be something like this. (P1) The ability of agent a to distinguish the microstate *x* of the gas is inversely proportional to its entropy $S_B$. (P2) The ability of agent a to distinguish the microstate *x* of the gas will decrease during the free expansion of a gas. (C) Therefore, the entropy of the gas will increase during its free expansion.

assimilation of the generation of entropy $S_B$ as a thermophysical consequence of its number of indistinguishable microscopic configurations increasing during its expansion.

Nevertheless, the possibility that the agent can make such valid inferences depends initially on the cognitive resources linked to the interpretation B of $S_B$, making it possible for its content to become sufficiently intelligible to cognitively enable its inferential exploitation. At this point I would argue that the fundamental problem lies in the fact that the main proponents of this interpretative strategy are forced to 'epistemically parasitize' the cognitive resources linked to other interpretations in order to make intelligible the phenomenon to which the concept refers: "Since any one of these [macroscopically indistinguishable] different microstructures can actually be realized at any given time, the lack of information corresponds to actual disorder in the hidden degrees of freedom" (Brillouin 1962, p. 160)[136]. In this case, whether the agent can understand the $S_B$ of the freely expanding gas increases with increasing epistemic inability to obtain information (OI) about the molecular configuration would depend conceptually on cognitively exploiting the degree of molecular disorder (OO) in the different microstructures contained in the observable macrostate. This occurs, for example, by imagining how the molecular configuration of the gas at the end of the expansion are much less visually regular than the molecular configuration of gas at the beginning of the process (of course, assuming that the gas was initially in a crystal-like ‘regular’ Past State). Therefore, the robustness of an interpretive strategy B of $S_B$ would ultimately depend on developing a version of this proposal that is non-parasitic on the intelligibility resources of the disorder interpretation, a task that has not yet occurred.

### 5.5.4. Informational Interpretation C: Entropy as Shannon’s (or Algorithmic) Information

Next, I will evaluate an interpretative strategy that aims at specifying the meaning of entropy, particularly the notion $S_G$, through its identification with a notion of information (namely the technical concept $H_{IT}$): namely, that initially suggested by von Neumann (1949) and subsequently developed from the Jaynesian tradition (Section 3.3) or the later syncretic current. First, and unlike the informational interpretation A, the possibility for the agent to understand the meaning of entropy through this interpretation C depends on his understanding of the concept of information $H_{IT}$. Otherwise, the user might assume (because of the intelligibility of the everyday notion of information [OI]) that $H_{IT}$ information refers to the agent's epistemic semantic capacity over the microscopic configuration of the system, and not strictly a function of the binary-based logarithm of the probability over a macrostate $\Gamma_M$. Thus, the number *n* of bits of the system in the macrostate-region $\Gamma_M$ corresponds to the number of binary partitions of the region $\Gamma_M$ needed to specify the actual microstate *x* of the system. However, since the phase space of a classical system is

[136] That is, that the lack of information (via interpretation B) provides intelligibility would depend on the fact that the number of microstructures indistinguishable from the point of view of observation corresponds to the degree of molecular disorder (OO) that we find in the system.

continuous, the number of bits n of the system at each moment of its evolution will be uncountably infinite $n$ = r for each moment $t_i$ of its dynamical evolution[137].

However, that the conventional amount of $n_0$ bits at the beginning of the gas expansion and $n_f$ at the end of the process not only does not provide cognitive resources to increase the intelligibility of the entropic behavior of the gas, but it would also in no way allow to conceptually understand why an increase of $n_0$ - $n_f$ bits would necessarily imply a generation of a certain amount of $S_{TD}$ into the system's environment. That is, according to the previous evaluation, one cannot statistically predict or correctly infer the macroscopic values of the system from the interpretation C in an invalid way, e.g., predicting (correctly) that the amount of entropy $S_{TD}$ produced is proportional to the volume expanded by the gas does not semantically depend on bit differences. As in the A-interpretation, but even rejected by some adherents of this position: "The involvement of energy and temperature in the original definition of the entropy is a historical accident, a relic of the pre-atomistic era of thermodynamics (...) The equation for the entropy would be simply S = In W and entropy would be rendered dimensionless!" (Ben-Naim 2008, p.30). In conclusion, the interpretative exploitation of $H_{IT}$ not only conflates the meanings of $S_B$-$S_G$ and $H_{IT}$ as mere microstate counters in Γ, but also renders the concept of $S_{TD}$ entropy incomprehensible because it does not illuminate how (or why) certain macroscopic values like T or Q should be associated with sets of microstates.

One might ask whether the technical concept of algorithmic information K (substituting or complementing $H_{IT}$ information) would contribute epistemically within the C-interpretation to improve our understanding of thermophysical entropic notions. As I remarked in Section 4.1, this new line of interpretation was originally proposed by Bennett, who stated that "the macrostate's statistical entropy in binary units, measures the extent to which the distribution is spread out over many microstates; [K($x$)], the microstate's algorithmic entropy, measures the extent to which a particular microstate $x$ is not concisely describable, and [$H_{IT}$($x$)], the algorithmic entropy of $x$, is the number of bits required to describe the distribution" (Bennett 1982, p.938). In the sense that Bennett notes above, the only conceptual contribution of K in the specification of $S_B$ or $S_G$ content is its ability to determine the number of bits required to describe the distribution of microstates upon which entropy notions $S_B$ or $S_G$ are defined. Because the algorithmic information K does not constitutively allow one to bound the factor (recall Section 5.3) relating number of microstates to (empirically significant) entropy $S_{TD}$ production, then it could not be interpretively employed to understand the thermophysically meaningful content of $S_B$ or $S_G$.

### 5.5.5. Informational Interpretation D: Entropy as Epistemic Probability

[137] While in quantum physics only the Planck length imposes an objective limit on the discretization-partitioning of the phase space (see Frigg 2008), in the classical case there is no unconventional way to obtain the bits "the amount of [Shannon information] (...) depends on the accuracy we choose in locating the particle; the finer the size of the cell (...) the larger the [Shannon information]" (Ben-Naim 2008, p.254).

Finally, another of the most historically relevant informational interpretative strategies of entropic concepts (particularly $S_G$ entropy) are those dependent on subsidiarily interpreting SM-probabilities as representations of the agent's information (OI). This interpretative line, intertwined with strategy C, was historically developed from the MEP-Jaynesian tradition capitalizing on epistemicist positions à la Tolman (1938) in the context of GSM (Section 3.3). Currently, this informational interpretation D is often employed in the literature to shed light on (and even justify) Gibbsian technical procedures such as the introduction of coarse-grained descriptions in a way that "washes away all information about the past" (Uffink 2010, p.196). According to this interpretative line D, the agent would understand the meaning of the $S_G$ entropy of the freely expanding gas through the idea that the Liouvillean dynamics of the distribution during the evolution of the gas would represent (more or less consistently) how the information (OI) that the agent possesses during his study of this process is transformed.

In this interpretive scenario, the increase in $S_G$ entropy during coarse graining would imply that the agent loses information (OI) (or epistemic access) about the initial macroscopic conditions of the system. This is because to achieve $S_G$ entropy increase it is necessary to obtain a new coarse-grained statistical description $\bar{\rho}$ by phase-averaging the values of the initial distribution $\rho$ (Section 2.2.3). Note that while the Gibbs entropy $S[\rho]$ will remain (contrary to TD-predictions) constant during evolution according to Liouville's theorem, the value of the coarse-grained Gibbs entropy $S[\bar{\rho}]$ will increase (except for thermal equilibrium, which remains the same) with respect to $S[\rho]$. Since the initial probability distribution $\rho$ represents the agent's macroscopic knowledge about the initial conditions of the system (i.e., when the agent knows that the gas is initially in volume V), this observational knowledge would be immediately removed or, at least, would cease to be epistemically accessible when a new distribution $\bar{\rho}$ is introduced[138]. Of course, real scientific agents do not forget so easily. However, interpretation D omits this fact: the only epistemic states considered are those represented in $\rho$ and $\bar{\rho}$, so the coarse-graining translation from $\rho$ to $\bar{\rho}$, would abstractly model a forgetting of the initial epistemic states represented in $\rho$.

According to interpretation D, understanding the significance of the gas entropy increase through the agent's loss of information (OI) about the initial conditions of the gas should be inferentially exploitable in a valid way. However, the fact that this agent infers (correctly) that the gas will dissipate a minimum amount of energy based on the interpretation D of the entropy does not depend on or does not follow semantically from the meaning of this concept, but from the theoretical validity of the coarse-graining procedure of the probability distributions to fit the macroscopic values of the system. Regardless of whether the 'loss of information' underlying GSM coarse-graining can be justified in an objective and non-anthropomorphic way (see Robertson 2020), the fact remains that it does not contribute, together with the associated cognitive resources, to endow the meaning of $S_G$ with intelligibility in cases of systems approaching thermal

[138] According to this interpretation-D subsidiary to an epistemic interpretation of probabilities, the agent would 'forget' or lose information (OI) about the probability that a given microstate of the gas was the current one when it macroscopically occupied volume V due to the entropy increase.

equilibrium. In any case, interpreting an increase in $S_G$ as a loss of observational information found in the distributions could eventually constitute a robust interpretation of the meaning of the concept of coarse-grained $S_G$ entropy, but in no case of its fine-grained version, mainly because it would immediately cease to be valid in scenarios in which coarse-grained is not necessary, i.e., any process that is not far from thermal equilibrium.

### 5.5.6. Can Informational Interpretations provide Robust Understanding of Entropy Concept?

At this point, it would be interesting to compile the conclusions derived from the analysis on each of the main interpretations of entropy (mainly $S_B$ and $S_G$, but in no case $S_{TD}$) deployed from informational descriptions. Firstly, I have assessed how the technical concepts underlying interpretations C and D are (at least in a context of thermophysical practices) more unintelligible than entropy concepts, making it substantially more difficult to understand their meaning contrary to our desideratum. The satisfactory understanding of theoretical concepts such as 'number of bits' $H_{IT}$ or K (or the introduction of coarse-grained statistical descriptions, in the case of D) necessarily requires that the agent possesses a highly technical knowledge in advanced statistics and theoretically specialized telecommunications or computer theory. Of course, this prima facie technical knowledge will not provide the agent with any strategic advantage to satisfactorily understand the thermophysically meaningful content of the entropic notions.

Additionally, one may wonder whether the exploitation of non-technical and non-theoretically defined informational concepts (all of them linked to OI) such as 'lack of microstatistical information' (A), 'indistinguishability of microscopic information' (B), or 'loss of observable information' (D) in the interpretation of thermo-statistical concepts $S_B$ or $S_G$ entropy may contribute to a certain understanding of their meaning. Authors such as Leff (2007) or recently Shenker (2020) defend this idea. I argue that certain properties of OI information (i.e., its agency, semantic-epistemic character, Section 2.4) may become relevant to understand not the thermo-empirical content of $S_B$ or $S_G$ entropy, but its statistical character as a microstates counter or probabilistic measure of a phase region Γ. Since "there is no real information relevant to thermodynamics beyond that provided by the macroscopic state specification" (Wicken 1987, p.192), the only informational concepts that would contribute to illuminate the deeply empirical meaning of the thermo-statistical $S_B$ or $S_G$ entropies are those that technically refer to the energy and temperature values encoded in such entropy quantities via Boltzmann constant $k_B$. However, none of the main informationalist interpretative strategies one can find throughout the history of modern thermophysics (i.e., A-D) employ a notion of information significant enough from a thermoempirical point of view to be able to validly infer why the entropy produced[139] by heating 1kg of water from 303º K to 353º K will be approximately 0.64 kJ/K[140].

Of course, it should be recognised that the everyday notion of information (OI) can cognitively assist (both practitioners and students) in providing intelligibility to the statistical character of the entropic concepts $S_B$ and $S_G$, although due to its conceptual vagueness this plausible assistance is highly context-dependent. In this direction, the robustness of the partial intelligibility that informational interpretations can provide us necessarily depend on their conceptual capability to theoretically specify how scientific agent-observers epistemically relate to thermophysical phenomena by means of these $S_B$ or $S_G$ statistical entropy concepts. Illustratively, a plausible robust informational interpretation should consistently (maybe quantitatively) specify the semantic-epistemic dynamics underlying expressions such as 'lack of information (OI) about the actual microstate', namely: how should we understand such agentiality, what is the vehicle of such quantity, what is the representational mechanism that allows such information to refer to a microstate, etc. This 'plausible' informational interpretation would involve undertaking the colossal theoretical task of informationally naturalizing (maybe from information-driven naturalistic programs like Skyrm's [2010] or Martínez' [2019], see footnote 117) the agent-based semantic and epistemic dynamics underlying the theoretical-experimental processes of obtaining knowledge about thermophysical phenomena.

The key to my argument is that all these informational interpretations (A-D) seem incapable of endowing $S_B$ or $S_G$ entropy with intelligibility as thermo-empirically meaningful measures of the amount of entropy $S_{TD}$ produced during a TD-irreversible process, being particularly useless for understanding $S_{TD}$. It is precisely in this sense that informational interpretations of entropy concepts appear to us as interpretive strategies that are not very technically precise according to the previous assessment, although they can be heuristically useful in certain pedagogical scenarios. Moreover, they can be prone to confusion (e.g., employing $H_{IT}$ and OI interchangeably) by less competent conceptual users. For these reasons, the main historical proponents of these strategies found it necessary to parasitize epistemically on other non-informational interpretative strategies such as 'molecular disorder' (e.g., Brillouinian tradition [Brillouin 1962, p. 160]) or 'energy dispersion' (e.g., Jaynesian tradition [Jaynes 1957a]) to reinforce the wide-scope intelligibility of their proposal. In short, although the suggestive notion of information (OI) can heuristically assist in our understanding of the statistical meaning of entropy, the main informational interpretations

[139] In this direction, Tolman (pre-informational epistemic interpretations of entropy) famously suggested that: "the statement sometimes made that the entropy of a system is a measure of the degree of our ignorance as to its condition. From a precise point of view, however, it seems more clarifying to emphasise that entropy can be regarded as a quantity which is thermodynamically defined with the help of its relations to heat and temperature (...) and statistically interpreted with the help of the analogous relation between mechanical quantities" (Tolman 1938, p. 561).

[140] Of course, someone (perhaps Ben-Naim) might argue that the lack of $k_B$ in $H_{IT}$ is not necessary to infer observational values, since one could prima facie calculate those values once $H_{IT}$ has been computed. While this is true, the lack of $k_B$ in computing entropy values makes that process a mere algebraic procedure with no physical meaning that can be understood by an everyday scientific agent, so that this thermophysical meaning should be interpretively specified only a posteriori, at the end of the algebraic computation.

of entropy (A-D) are not well-suited to epistemically contribute to generate a thermo-physically meaningful and technically precise understanding of these concepts.

## 5.6. Deflating Informational Ontologies in Thermal Physics

Finally, I evaluate how certain alleged achievements within the informational program in thermophysics were exploited to support certain worldviews. In this sense, the defense of informational ontologies (Section 4.4) since the 1990s constitutes one of the most significant theoretical objectives of the informational research program in physics in the last decades. Following Chakravartty's (2017) distinction between 'epistemic achievements' and 'epistemic goals' within the debate on scientific realism, I argue that the theoretical goal of sustaining a realist ontology or attitude could only be coherently justified in a scientific context by epistemic achievements obtained from this domain: "scientific realism is a position concerning the actual epistemic status of theories" (Chakravartty 2017). But before assessing these informational realist positions, let us return to the philosophical debate on this issue outlined in Section 4.4.

### 5.6.1. Returning to the Philosophical Debate on Information Ontologies in Physics

As a starting point for the assessment of physics-informational realism, let us recall the evaluative proposals of Timpson (2013) and Lombardi et al. (2016). Initially, Timpson (2013) analyses how certain interpretative uses of informational concepts in physics (illustratively, Penrose's [1998] reconceptualization of quantum teleportation as a flow of 'informational substances') incur a categorical error between abstract-types and concrete-tokens, which makes possible a hypostatization of the notion of information (Section 4.4.3). This author develops a deflationary framework with respect to inflationary attitudes based on this type of conceptual misuse (which I include among those analysed in Section 5.2), employing an analytical explanation of how the historical-linguistic transition between the verb 'to inform' and the noun 'information' made possible the categorial confusion that one would observe in the metaphysical proposals of Wheeler (1990), Stonier (1990) or Landauer (1991).

At first sight, it seems prima facie untenable that the whole sophisticated conceptual architecture underlying the ontologies linked to Wheeler's (1990) 'it from the bit' or Landauer's [1995]) 'information is physics') derives from a mere confusion of grammatical categories among professional scientists. Lombardi et al. (2016a) argue that the analytic-linguistic nature of Timpson's (2013) argument invalidates it as a criterion for assessing the theoretical legitimacy supported by ontological-informational frameworks in physics. Although I fundamentally agree with Lombardi et al. (2016a) on this point, I argue that Timpson's (2013) deflationary proposal is not deficient because it is analytic but precisely because its explanatory factors are unnecessary and insufficient to account for the emergence of ontological claims in informational physics. It is unnecessary, since the ability of any physicist to sustain an informational ontological proposal is independent of his competence with the philosophical-analytic categories of type-token or its abstract and concrete nature (respectively), and with the grammatical categories ‘verb-name’. In other words, it should be be conceptually possible for a physicist to be competent (i) with the analytic (type-token) and grammatical (noun-verb) categories, and at the same time (ii) have reason to believe that a bit-increment in the description of a molecular system has some physical meaning in reality[141]. Therefore, Timpson's (2013) deflationary proposal based on (i) is not necessary to explain (ii) this type of ontological attitude in physics. Although I fundamentally agree with Timpson's proposal that these ontological attitudes derive from conceptual misuses, I argue that these inconsistently and incoherently employed concepts are not philosophical-analytic or grammatical-linguistic, but properly thermal and informational (i.e., entropic, probabilistic, and informational, see Chapter II). "These discussions [on Wiener's problem] are philosophically interesting when one admits that it is physics and not grammar that is the best clue for discovering the content of the physical world." (Lombardi et al. 2016b, p.227).

In response to this deflationary proposal by Timpson, Lombardi et al. (2016a) sought to develop another deflationary framework for evaluating this type of ontological attitudes in physics (Section 5.4.4). Like the first author, they focused their analysis on how the interpretive specification of the meaning of Shannon information $H_{IT}$ conditions the belief that it is part of the physical reality "The relationship between the word 'information' -in Shannon's formal context- and the different views about the nature of information is the logical relationship between a mathematical term and its interpretation" (ibid, p.228). That is, according to Lombardi et al. the possibility of believing that $H_{IT}$ information has a physical nature is conditioned by the ability to connect this notion interpretatively with physically meaningful concepts. In this direction, one might ask what is the interpretation of $H_{IT}$ that allows us to realistically engage with its content. According to the authors, any interpretative strategy is equally legitimate for holding ontological attitudes about $H_{IT}$ as long as such specification of its content is 'useful' in a given scientific or engineering application scenario (see Lombardi et al. 2016a, p.2009). Thus, the deflationary perspective of Lombardi et al. would be significantly conditioned by an interpretive pluralism about informational concepts.

Having defined their interpretatively-pluralistic position, Lombardi et al. (2016b) apply their proposal in the domain of informational physics "What only matters is that all those items [energy, fields, and information] inhabit the world of physics, that is, according to physics they are part of the furniture of the world. And this implies that *contemporary physics offers no grounds to deny the possibility of a non-trivial and meaningful physical interpretation of the concept of*

[141] Additionally, that the verb 'inform' has historically derived into the noun 'information' is by all accounts insufficient to satisfactorily explain why in the 1990s a group of physicists began to advocate that information is part of our physical reality. Even assuming a hard Whorfian hypothesis whereby our linguistic resources determine our cognitive capacities (Scholz et al. 2015, Section 4.4), the lexical genesis of the noun 'information' from the verb 'to inform' occurs during a time interval (i.e., the late Middle Ages, Capurro & Hjørland 2003) several centuries separate from the intellectual genesis of the ontological attitudes about information analysed here (Chapter IV).

*information*" (ibid, p.227. Italics are mine). But is it true that it is possible to generate a physically meaningful interpretation of the $H_{IT}$ concept from modern physics? Focusing exclusively on the domain of thermophysics, Lombardi et al. (ibid.) argue that it is possible to generate a thermophysically relevant interpretation of $H_{IT}$ simply through its conceptual connection with $S_B$ or $S_G$. Charitably, this interpretation can be 'useful' in that it would allow us to quantify the bits that are necessary to describe the macrostate of a molecular system. But despite this instrumental utility generated by the formal connection between $H_{IT}$ and $S_B$-$S_G$, from the exhaustive analysis of the semantic and interpretative disconnection in Section 5.2.2 one can extract that the only way to provide robust thermophysical content to $H_{IT}$ would be to transform it into $S_B$ or $S_G$. That is, that the number of bits needed to describe an expanding gas increases exponentially during this process does not provide us with relevant knowledge about, for example, the distribution of molecular properties underlying a macroscopic thermal process[142]. Thus, the interpretation of $H_{IT}$ by means of $S_G$ or $S_B$ could only be useful in a pragmatic sense (e.g., changing from a natural logarithm quantity to a binary logarithm quantity), and not useful in a properly epistemic sense, as defended in Section 5.2.

Therefore, according to the proposal of Lombardi et al. (2016b), for an interpretation of $H_{IT}$ content to be thermophysically meaningful and thus amenable to ontological attitudes, it would suffice for it to be instrumentally useful in the thermophysical domain. Although I agree with Lombardi et al. that the sustainability of info-realist positions should be assessed based on the use of concepts underlying such a position, in the following I will argue (against these authors) that it is not enough for an interpretation of an informational concept to be 'instrumentally useful' but that it should be 'epistemically useful' (i.e., required to obtain knowledge) within thermophysics[143]. That is, the 'unrestricted' interpretative pluralism on which these authors base their deflationary framework has the opposite effect of allowing almost any physical interpretation (under the requirement of being instrumentally useful) of an informational concept to be legitimized to believe that its content is part of physical reality.

### 5.6.2. Deflating Information Ontologies in Physics from a Khinchean Perspective

Having evaluated the advantages and disadvantages of the main philosophical positions on how to analyze informational ontologies in physics, I will now develop my proposal. I start from the premise (obtained above) that, in order to sustain a realistic attitude towards an informational concept, it must not only be instrumentally useful within thermophysics (as proposed by Lombardi et al. 2016b), but also epistemically useful in knowledge generation. However, it would also be reasonable to assume that the claim to incorporate the content of a concept C within our ontology

[142] Notice that the evaluation of the information-based interpretive strategies of Section 5.5 supports this idea as well.
[143] Illustratively, making binary partitions of phase space can be 'instrumentally useful' for specifying the microstate of a molecular system (since it is pragmatically relevant to obtaining knowledge), while defining observable values over phase space 'epistemically contributes' to obtaining knowledge about this system (since it is necessary for obtaining this kind of knowledge).

should not depend on this C providing any kind of knowledge, but on a form of knowledge that guarantees the epistemic contribution of the concept in the domain in question. For example, it would not be sufficient for an informational concept to provide understanding to a particular agent with respect to another thermophysical concept, since it is difficult to individuate the concept's contribution to this understanding from any other psycho-cognitive or pragmatic factor. In this direction I have heuristically adopted what is known in the recent literature on mathematical Platonism as the 'enhanced indispensability argument' (Baker 2009), which states that:

> Enhanced Indispensability Argument: We ought rationally to believe in the existence of any entity which plays an indispensable explanatory role in our best scientific theories.

This argument reinforces the indispensability argument, originally from Quine (1976) and Putnam (1979a), broadly stating that we ought to commit ourselves to the existence of all entities that are 'indispensable' (not just explanatorily, but in a broad sense) simpliciter to our best theories, such as statistical thermophysics[144]. This proposal is aimed to unify the results obtained in both debates in order to carefully assess the feasibility of informational ontological proposals. Interestingly, this analysis (and its conclusions) bifurcates in two significantly different directions depending on whether we characterize informational notions as mathematical concepts or not. Let us initially assume that the technical notions of information (except for OI, for obvious reasons) are constitutively mathematical and content neutral.

As it is explored in Section 3.4.1, the idea that Shannon's IT constitutes a mere statistical formalism, and that $H_{IT}$ is exclusively a mathematical modelling tool was popularized in the 1950s by Alexander Khinchin (1957) and later by Katz (1967) and other authors. Lombardi et al. (2016a), who advocated this line of thought in their philosophical proposal, called this position a 'formalist view' of informational concepts. According to this 'Khinchinian' view[145], the $H_{IT}$ concept is no more than a statistical tool with no defined content and no other 'proper' theoretical domain (i.e., communication theory or signal transmission theory) than the one on which it is consistently applied, such as discrete signal sequences (Shannon 1948), ideal gases in isolated containers (Jaynes 1957a) or the event horizon surface of a black hole (Bekenstein 1973). If we assume this general position on Shannon's theory[146], it should be easy to reach the conclusion that "Information theory approach is not a miracle device to arrive at a solution of any statistical problem. It is only a general way to formulate the itself consistently" (Katz 1967, p.i).

[144] Interestingly, the question about the physical reality of information has been treated from the philosophy of science independently of the key problem of scientific realism (except by Ladyman and Ross [2007]), in general, and of the problem of mathematical Platonism in which Baker's (2009) enhanced indispensability argument arises, in particular.
[145] Note that this Khinchinian view fits perfectly with the interpretative pluralism of Lombardi et al. (2016a), where the content of any informational concept is provided only by the interpretation we adopt about it.
[146] Although this perspective is addressed mainly to Shannon's concept, it could equally well be extended to the concept of algorithmic information K, where Kolmogorov (1965) holds a formalistic position with respect to AIT analogous to that which Khinchin (1957) holds with respect to IT.

At this point of the argument, we should ask ourselves whether informational notions qua mathematical concepts could contribute to explain scientific phenomena, such as in our case thermophysical behaviors. Within the recent debate on whether mathematical structures can explain scientific phenomena (Baker 2005, 2009), Juha Saatsi (2011) argues that there are two ways in which a concept can contribute to an explanation in the natural sciences: namely, (i) by playing a merely 'representational' role, offering formal instruments to describe the explanans, or (ii) by playing a properly 'explanatory' role, conceptually determining the factors that account for the phenomenon to be explained. Note that this distinction is orthogonal to the previously remarked distinction between the 'practical utility' and 'epistemic utility' of a concept in an explanation. Having distinguished between these two functions, I follow Saatsi (ibid.) in defending that informational notions qua mathematical elements can contribute in the explanation of a thermophysical phenomenon without this being a properly explanatory contribution: "Mathematics can help us learn about the world by virtue of playing a representational (and, derivatively, inferential and justificatory) role in science (...) our endeavour to separate the representational and the explanatory functions of mathematics is helpful because it indicates how mathematics can play a knowledge-conferring role without being explanatory per se" (Saatsi 2011, p.145, 153). But according to the 'enhanced indispensability argument', we should rationally believe in the physical existence of the content of an informational-mathematical concept if it is explanatorily relevant (i.e., it plays a properly explanatory role) in physics, in general, and in statistical thermophysics, in particular.

From the analysis developed in Section 5.3 it follows that the use of mathematical concepts of information such as $H_{IT}$ or $I_{Bb}$ within the main explanations of thermal phenomena does not fulfil a properly explanatory function. The reason, as I have detailed extensively, is that the use of $H_{IT}$ as a mathematical tool within Jaynesian (Section 5.3.3) or recent informational (Section 5.3.4) explanations only fulfils the 'representational' function of allowing the encoding of the microstatistical data involved in the explanans in a highly useful way (thanks to the logarithmic function of $H_{IT}$) and by means of bits. Likewise, the complementation of these explanans with the K-notion would only contribute 'representationally' to the delimitation of the number of bits required to specify an individual microstate (Section 5.3). Furthermore, the capitalization of other informational concepts such as $I_{Bb}$ or $I_L$ in Brillouinian (Section 5.3.2) explanations exclusively possess the 'representational' function of reformulating in negative or difference terms non-informational quantities such as $S_B$ (Denbigh, 1981; Earman and Norton, 1999) or the NIB degrees of freedom, respectively. In short, because informational notions qua mathematical structures are unable to contribute explanatorily in obtaining thermophysical knowledge, then we should not rationally believe in their physical existence.

### 5.6.3. Deflating Informational Realism in Physics

Once I have analysed the lack of rationale for a realist attitude towards informational concepts in physics from a Khinchinian or formalist stance, the reader may now rethink my argument assuming that these informational concepts are not properly mathematical (although they require mathematical elements for their formulation) and that they also have a certain theoretically defined and/or interpretatively specifiable content. With this change of premise, we move from the debate on the plausibility of informational Platonism to the debate on scientific realism. From this stance on information notions, $H_{IT}$ can be assimilated as a concept about certain statistical properties in signal transmission, about the degree of uncertainty of an epistemic agent, or (if reinterpreted coherently) about the statistical distribution of molecular properties of a physical system. And the same goes for the rest of the technical concepts of information. Departing from the enhanced indispensability argument, the main idea could be stated as follows:

> **Criterion for Realistic Attitudes**: it is reasonable to adopt a realistic attitude only toward the existence of the content of those information-theoretic concepts (i.e., qua non-mathematical) that are (a) explanatorily relevant or (b) capable of generating novel valid predictions in physics, and not only representationally or pragmatically useful in knowledge elicitation.

The latter requirement (b) is incorporated to satisfy those authors (e.g., Kitcher, 1993, p.140-149; Votsis, 2011) who consider that it is those theoretical concepts that allow the 'generation of novel predictions' (and not the 'explanatory role', as occurs in mathematical Platonism [Baker 2009]) that are most amenable to realist attitudes[147]: "those aspects of the theory that are essential to the derivation of such novel predictions are the parts of the theory most worthy of realist commitment" (Chakravarty 2017). The key here is that the 'inferential validity' status on the criteria of explanatory power (Section 5.3.) and predictive advantage (Section 5.4) allows us to connect the knowledge obtained abductively and predictive-statistically (respectively) with the content of the concepts exploited to obtain such results.

Returning to the analysis developed in Section 5.3 and Section 5.4, it follows that the exploitation of the (interpretively specified) content of information-theoretic concepts such as $H_{IT}$, $I_{Bb}$ or K within the main explanations or predictions of thermal phenomena does not fulfil a properly epistemic function[148]. Due to the inability of $H_{IT}$ to consistently encode macroscopic properties $A_i$ of the molecular system, explanations relying on $H_{IT}$ content will lack explanatory power and their statistical predictions will be disadvantageous (if not incoherent) against alternative non-informational predictions. As argued in Section 5.3.5, the complementation of these explanations

[147] As the reader might have noticed, this 'rationality criterion' is nothing more than a heuristic framework that allows us to define the reasonableness of a given realist attitude about the content of an information concept not simply on the basis of the mere usefulness of interpreting it physically (Lombardi et al. 2016a), but on the explanatory and predictive role of the content of this concept in providing us with thermophysical knowledge.

[148] Recall that the use of $H_{IT}$ as a theoretical notion in Jaynesian (Section 5.3.3 & 5.4.2) or information-syncretic strategies à la Ben-Naim (2008) (Section 5.3.4 & 5.4.4) would only allow to redescribe microstatistical properties exploited explanatorily and predictively by bits.

or predictions with the content of the algorithmic-information K notion only allow to specify the number of bits required to specify the actual microstate *x*, but this does not increase their explanatory power or predictive advantages (if anything the computational cost of both strategies increases in a practically unsustainable way). Also in this direction, the exploitation of other theoretical concepts from informational programs such as $I_{Bb}$ in the Brillouinian tradition (Section 5.3.2), only contributes to redescribe by negative quantities or differences of non-informational quantities such as $S_B$ (Denbigh, 1981; Earman and Norton, 1999). Compiling this evaluative line I conclude that, informational notions (qua notions with theoretically or interpretatively specific content) are unable to contribute in an explanatorily powerful or predictively advantageous way in obtaining thermophysical knowledge. Therefore, according to the heuristic criterion one would have no reason to rationally believe in the physical existence of informational concept content.

This evaluative diagnosis will be indistinct with respect to the type of realist attitude we maintain towards the content of the informational concept in question, namely, reifying information as a substance or considering it as a structural notion. Initially, I argue that 'ontologically naive' reificationist attitudes (Timpson, 2013; Ladyman & Ross, 2007) such as "Information is a quantity which may be transferred from one system to another (...) information, like energy, also possesses a physical reality" (Stonier 1990, p.26-27) are nourished by the same epistemic incapacity from which other more theoretically refined attitudes, such as the information-theoretical structural realism of Ladyman and Ross (2007), are nourished. The main difference (at least at a conceptual level) between the two positions lies in the fact that the former is based on an accumulation of conceptual misuses notably superior to the latter. But, contrary to what Timpson (2013) proposes, these realist positions would not be explained by misuses of analytical 'type-token' or grammatical 'verb-name' concepts, but of physical and informational concepts. For example, specifying 'the transfer of information from one system to another' as a process in which a substance-property 'information' (or 'infostuff' according to Ladyman and Ross, 2007, p.189) transits spatiotemporally from one system to another incurs a misinterpretation of the communicative channel as a sort of spatiotemporally extended pipe and of information as a fluid that contains it.

Ladyman and Ross (ibid.) sought to develop a viable physical-informational ontology, avoiding the conceptual misuses underlying the reification-prone attitudes (typical of the metaphysical-informational pretensions seen in Section 4.3) and conceiving of that information that deserves to be the object of realistic attitudes as the modal-statistical structure of our best physical theories. In particular, these authors propose that we should adopt a realist perspective with respect to the less redundant (maximally compressible, along the lines of Davies [1990], Section 4.3.3) statistical data provided by fundamental physics as the most epistemically successful theory. However, the Ladyman-Ross criterion underlying the adoption of selective realism with respect to this type of information is none other than an algorithmically-refined notion of 'simplicity' or lack of redundancy. That is, even if our most epistemically successful theories generate certain incompressible data about physical reality, there is no guarantee that the physically meaningful knowledge provided by such statistical information is due to its lack of redundancy. To conclude, whether physical reality can be algorithmically compressible or not can only be a 'desideratum', and as such would not constitute a sufficiently robust criterion to justify from a scientific domain that we should rationally believe in the existence of something.

In short, regardless of how philosophically sophisticated our realist attitude to the physical nature of information may be, the only thing that would provide a sufficiently rational justification for believing in the existence of the content of a theoretical concept is that the attainment of meaningful knowledge about a physical domain is dependent on that content. The assessment of the explanatory power (Section 5.3) and predictive capability (Section 5.4) of informational thermophysics aimed to show how the historical exploitation of informational notions (qua theoretical concepts) has not epistemically contributed to the obtention of thermophysical knowledge. Therefore, according to this analysis, there are no grounds solid enough to rationally justify the belief that informational properties and structures are part of our physical reality.

### 5.6.5. Can Information be Physical?

During this section I have analysed how the informational ontological proposals that emerged from this intellectual current (e.g., "It from the Bit" [Wheeler 1990] or "Information is Physical" [Landauer 1991]) were unjustifiably deployed due to the lack of epistemic achievements dependent on the use of informational concepts that would rationally justify our belief in their existence. My argument to sustain this thesis bifurcates in two directions depending on whether we consider such concepts as mathematical-formal or as theoretical, generating a dilemma. On the one side (and following the enhanced indispensability argument of Baker [2009], see Saatsi [2011]), if we consider à la Khinchin (1957) that information notions are mathematical concepts, their mere representational function and pragmatic utility of their formal tools in informational explanations of thermophysical phenomena would not be sufficient to justify a realist attitude about these properties. On the other side, if we consider informational notions as theoretical concepts, then the lack of epistemic contribution of their (interpretatively specified) content in informational explanations and predictions of thermophysical phenomena would also not be sufficient to justify our belief in their physical existence. Thus, my final thesis in this direction is not so much that information does not possess (or is conceptually impossible to possess) a physical nature, but properly that the history of informational thermophysics has not provided us with compelling reasons to justify our belief that this might be the case.

# Chapter VI

## Conclusion

The main aim of this work has been to analyse the historical and conceptual foundations underlying information physics (and in particular, informational thermophysics) as a program of scientific research or as a current of scientific thought. This research program was consolidated in the fifties with the aim of employing informational notions for the study of the thermal behavior of macroscopic substances, and after more than seven decades it still enjoys a certain popularity within the physics community. As I have discussed in detail (Chapters III and IV), informational thermophysics did not unfold historically in a progressive-homogeneous manner, but rather in a diachronically convoluted fashion via the confluence of different intellectual pretensions, research programs and theoretical motivations. However, in the previous Chapter V I concluded that the descriptive exploitation of informational concepts from 'informational' thermophysics has proven throughout its history (i.e., 1940s-present) not to provide us with a robust knowledge of thermophysical phenomena, but also epistemically inferior to other non-informational programs in classical thermostatistical physics. This is due to (i) the powerlessness of its explanations of thermal behavior such as the free expansion of a gas, (ii) the lack of contribution of information concepts to generate advantageous (regarding GSM and BSM) statistical prediction of observable values, and (iii) its uselessness in providing intelligibility to the thermophysical character of the concepts of entropy. This idea does not imply that the informational tradition does not have other virtues. For example, the Jaynesian program provides an extremely useful framework for making statistical inferences in classical thermophysics, while the Brillouinian use of informational concepts allows one to illuminate the statistical character of entropy on a pedagogical level. Otherwise, my core thesis is that the descriptive exploitation of informational concepts by these thermo-informational traditions since the late 1940s has not historically led to a significant advance in our knowledge of thermophysical reality. Let us now draw some conclusions about this idea.

### 6.1. The Intellectual Bubble of Thermophysical Informationalism

The first of the conclusions I intend to defend is that informational thermophysics has not provide us with new (and significative) knowledge about classical statistical physics and how this was not an impediment to its historical development. It should be recalled that informational thermophysics originated historically in the middle of the Shannon bandwagon (Section 3.1). In this context of gestation, some of the main intellectual promoters of this movement such as von Neumann (1949) explicitly stated that a future tactical exploitation of informational concepts from thermophysics could be epistemically beneficial in this scientific domain: "I have been trying to justify the suspicion that a theory of information is needed (...) Thermodynamical concepts will probably enter into this new theory of information. There are strong indications that information is similar to entropy and that degenerative processes of entropy are paralleled by degenerative processes of information." (Von Neumann, 1958, p.62f.). Here von Neumann suggests that, after the molecular-statistical reformulation of TD concepts in the SM domain, Shannon's IT would historically constitute the most evolutionarily sophisticated conceptual-theoretical framework from which to develop thermophysics.

This kind of theoretical promise served as motivation for a whole generation of physicists (Chapter III) to develop systematic theoretical proposals to employ informational notions in this physical domain, notably that of Brillouin (1956) and Jaynes (1957a). But as I argued, the Brillouinian exploitation of bound information $I_{Bb}$ does not allow us to generate powerful explanations of thermal behavior (Section 5.3) or the use of $H_{IT}$ and OI information from the Jaynesian MEP does not provide any predictive advantage over its non-use (Section 5.4). Furthermore, I have also pointed out how the informational interpretive strategies of Brillouin and Jaynes are unable to render the thermophysical content of entropy concepts comprehensible (Section 5.5). These important shortcomings within the field of thermo-statistical physics seem to indicate that the exploitation of informational concepts would be (against the promises of the most TD/SM-informationalization enthusiasts) epistemically irrelevant within this consolidated scientific domain. As suggested in Section 3.4 with the Carnapian critique, this real lack of epistemic success of Brillouinian and Jaynesian informational tactics was overlooked and historically overshadowed by the enormous fervor aroused by the future theoretical possibilities of informational thermophysics. This began to inflate the epistemic bubble of this intellectual current.

Nevertheless, the (then unperceived) lack of epistemic successes linked to the exploitation of informational concepts from the Brillouinian tradition (Section 3.2) and the Jaynesian program (Section 3.3) might simply imply that these two early thermo-informational proposals were deficient, but not that the thermo-informational program could not eventually become explanatorily and predictively superior to the non-informational proposals. Interestingly, thermo-informationalism as a broadly conceived theoretical program (since it encompassed several heterogeneous traditions) increased exponentially from the 1960s to the 1990s within the scientific community, mortgaging its capacity to generate putative knowledge within classical thermo-statistical physics. During this period, the intellectual bubble (i.e., confidence of the scientific community regarding its epistemic potential) of this current did nothing but inflate, e.g., "it is certain that the conceptual connection between information and the Second Law of thermodynamics is now firmly established" (Tribus and McIrvine 1971, p.188). As I evaluated in

Chapter V, the emergence of new syncretic and information-algorithmic approaches in the 1980s did not constitute any real advance in terms of explanatory power, predictive advantages or comprehensibility of our thermophysical reality with respect to historically previous traditions like BSM or GSM. The algorithmic complexity of a density does not allow one to powerfully explain (or render intelligible) why a gas eventually expands through a container (Section 5.3) or to advantageously predict the observable values that characterize this process (Section 5.4).

This inability of the informational program to obtain knowledge about the thermophysical reality can not only be extracted from a philosophical evaluation (as has been the case in Chapter V) but also by attending to the main theoretical advances and epistemic successes obtained in classical thermal physics during this period. Among them one can find: (i) the derivation of the TD-limit by Khinchin (1949) and later in the sixties-seventies (e.g., allowing to powerfully explain certain thermal behaviors at equilibrium), (ii) a computable approximation to the Boltzmann equation by Lanford (1975) (e.g., allowing to predict microstatistical values with overwhelming accuracy) or (iii) the so-called BBGKY hierarchy, which allowed to effectively compute the evolution of densities by means of Boltzmann proto-equations as well as to predict complex molecular behaviors (e.g., transport phenomena) in gases and liquids (see Uffink 2007, p.1006-1063). Interestingly, the epistemic achievement of all these significant theoretical advances in the field of classical thermo-statistics was historically independent of the intellectual pretensions and strategies of informationalism (Chapter III and IV) unfolded during the second half of the twentieth century. Surprisingly, after more than seven decades of epistemic infertility, the informational program in classical thermal physics continues to enjoy today in certain sectors (mainly those linked to scientific popularization) a huge and undeserved reputation:

> “The laws of information had already solved the paradoxes of thermodynamics; in fact, information theory consumed thermodynamics. The problem in thermodynamics is, in truth, a special case of information theory. Now that we see that information is physical, by studying the laws of information we can figure out the laws of the universe” (Seife 2006, p.87)

Interestingly, the very possibility of making such claims is a direct consequence of the enormous intellectual inflation suffered by this intellectual current in the last two decades "in the last several decades there has been significant growth in a field called "The Physics of Information"" (Maroney and Timpson 2018, p.103). Note that these radical informational stances within current classical thermophysics not only emerge in outreach settings but also within professional scientific practices, as is the case of Ben-Naim (2008) (Section 5.1). The inability of Ben-Naim's informational approach to powerfully explain, advantageously predict, or intelligibly interpret certain thermal behaviors of macroscopic substances (Section 5.3, 5.4, and 5.5) implies that the claim to exploit informational concepts in an epistemically fruitful way in thermophysics remains a mere promise after more than six decades of historical development. In fact, Ben-Naim's 2008 informational explanations of the equilibrium approximation of an ideal gas are equally powerless (i.e., due to their common factors, see Section 5.3) as the Brillouinian explanations developed in the 1950s. In this sense, we find a strong lack of correlation between the amount of intellectual output generated by informational thermophysics (books, papers, conferences, etc.) during 1950-2010 and the continued lack of theoretical results or epistemic successes generated by this program during this same period. This profound mismatch between the high credentials generally associated with this program and its lack of results is the result of this intellectual inflation.

This dynamic shed some light of suspicion on the status of these theoretical proposals as a successful program. Illustratively, the epistemic capacity of Einstein's general relativity in 1915 to (i) powerfully explain the advances in the perihelion of mercury and (ii) predict advantageously against Newtonian theory that this would generate a curvature of light (i.e., proven in the famous solar eclipse of May 29, 1919) was decisive for the acceptance of this conceptual framework within gravitational physics. In contrast, the exploitation of information concepts in Brillouin's 1956 or Jaynes' 1957a proposals did not contribute at all to explain powerfully (competing against BSM) or predict advantageously (competing against GSM) why and how the molecular properties of a gas approaching equilibrium are distributed. Unlike explanatory-predictively successful proposals such as general relativity, the prolonged acceptance of thermo-informational theories for more than seven decades[149] (1950s-2010s) could only be explained by factors independent of their ability to provide us with knowledge (e.g., intellectual fashions[150], suggestive concepts, multidisciplinary interests, etc.). Ultimately, the possibility of generating a significant informational proposal within physics has remained a mere promise since its very origin: "A new territory was conquered for science when the theory of information was recently developed" (Brillouin 1956, p.ix).

## 6.2. Informational Thermophysics as Degenerate Science

The second of the final conclusions I intend to defend concerns the status of informational thermophysics as a theoretical proposal (or properly, a collective of theoretical proposals) in this physical domain. In this sense, the disability of the thermo-informational traditions to provide us with meaningful knowledge allows us to characterize in Lakatosian terms this historically unfolded sequence of theories ($T_\alpha = T_1, T_2, ... T_i$) as a 'degenerate research program'. Famously Imre Lakatos (1978, p.33-35) introduced in *The Methodology of Scientific Research Programmes* (where he seeks to enhance Popperian falsificationism from a historicist position à la Kuhn) the distinction between 'progressive research programs' and 'degenerating[151] research programs', differentiating

[149] As it was defended in Section 4.1, a significant part of the sustained informational enthusiasm in the 1990s stem from the black hole application of information concepts (e.g., Bekenstein 1973).

[150] Of course, the existence of intellectual trends or 'fashions' within contemporary physics is a historical fact not incompatible with considering modern physics as one of our (if not 'the') most epistemically successful scientific disciplines. This is similar to what Penrose (2016) recently argues in his *Fashion, Faith, and Fantasy in the New Physics of the Universe*, where he defends (i) that string theory research programs should be considered intellectual fashions, but also (ii) that there is still theoretical and epistemic progress within quantum gravity.

[151] Being charitable to Lakatos (1978), the use of the term "degenerating" was not intended to be merely patronizing. Otherwise, the characterization of a program as "degenerating" (and its scientific product as "degenerated") is intended to indicate the inability of a research program to generate knowledge. This is precisely the sense in which I use it.

between those sequences of theories that produce an advance with respect to our knowledge of a physical domain and those sequences that do not generate such an advance. In his own words:

> "Let us say that such a series of theories is theoretically progressive (…) if each new theory has some excess empirical content over its predecessor, that is, if it predicts some novel, hitherto unexpected fact. Let us say that a theoretically progressive series of theories is also empirically progressive (…) if some of this excess empirical content is also corroborated, that is, if each new theory leads us to the actual discovery of some new fact. Finally, let us call a [research program] *progressive if it is both theoretically and empirically progressive, and degenerating if it is not*" (Lakatos 1978, p.34-35. Italics are mine)

Due to the fine grain of this evaluation (focusing not on theories, in general, but on particular theoretical proposals, see Chapters III and IV) the Lakatosian criterion of 'theoretical progression' based on novel predictions does not serve to analyse whether there is indeed theoretical progress in obtaining significant knowledge about a particular domain, as I already pointed out in Chapter V. Moreover, this criterion would imply that all classical thermo-statistical research programs (e.g., BSM, GSM, etc.) are equally progressive, due to the implausibility of novel thermal behavior. For adjusting these evaluative tools to the purpose of this dissertation, I extend Lakatos' framework as follows: a 'cluster' $T_\alpha = T_i,...T_f$ of historically unfolded thermo-informational theories[152] is 'theoretically progressive' if the explanatory power and predictive advantage of the latter theoretical proposals $T_f$ is significantly higher than those of the initial theoretical proposals $T_i$. Otherwise, the cluster of theoretical proposals defined by $T_i,...T_f$ will be 'theoretically degenerating'. Illustratively, the Boltzmannian BSM program can be characterized as 'theoretically progressive' because the Neo-Boltzmannian proposals of the 1990s are epistemically superior (e.g., it allows to generate typicalist explanations of the equilibrium approximation, it enables feasible statistical predictions, etc.) to Boltzmann's proposal in 1872.

Let us now apply this evaluative framework on informational thermophysics. Initially, we order the cluster of all the theoretical-systematic proposals $T_\alpha$ of this program (exclusively in their use of informational tactics) according to their mutual interrelation and historical momentum $T_\alpha = T_i,...T_f$ with $T_i$ being Brillouin's proposal in 1956 (Section 3.2.) and $T_f$ Ben-Naim's proposal in 2008 (Section 5.1). Without listing exhaustively all the elements of $T_\alpha$, it should be mentioned that these include syncretic proposals developed during 1960-1980 as that of Tribus (Section 4.1) or the proto-metaphysical frameworks of the 1990s (Section 4.3). On the basis of this Lakatosian approach, I argue that the cluster of theoretical proposals $T_i,...T_f$ that connects historically-intellectually (after approximately 52 years) Brillouin's proposal with Ben-Naim's can be characterized as theoretically non-progressive and therefore degenerating. The reason, as previously discussed, is that (i) the lack of explanatory power of $T_i$ and $T_f$ has the same conceptual

[152] Unlike Lakatos (1978), this 'cluster of proposals' does not imply to defend (from a Hegelian-like perspective) that theoretical proposals unfold historically in a sequential-linear fashion. This 'cluster of proposals' aims to be a merely a heuristic resource to evaluate sets of proposals historically connected by relationships of intellectual influence in which a common use of concepts, conceptual relationships and interpretative strategies is inherited.

roots (Section 5.3), and (ii) both $T_i$, and $T_f$ are predictively disadvantageous against the capabilities of non-informational frameworks such as GSM or BSM (Section 5.4). Thus, thermophysics dependent on information concepts from 1956 to 2008 (Chapter III, IV & V) would constitute a paradigmatic case of what Lakatos (1978) called a degenerate research program or simply degenerate science. It should be recalled at this point that the idea of the theoretical progression of informational thermophysics as a degenerate or non-progressive science has already been suggested in the philosophical literature (except in non-Lakatosian terms) by Norton: "We found a literature [on thermophysical computationalism] based on unsound principles and methods. More recent work in that literature has not improved matters. It takes that same unsound foundation and adds layers of more elaborate theorizing. (…) the totality remains incoherent, with its unsound foundations now obscured by the sheer mass of the new theorizing" (Norton 2011a, p.196).

The characterization of 'degenerate science' (Lakatos, 1978; Mbugua, 2015; Musgrave & Pigden, 2021) is particularly useful in the case of informational thermophysics as a stream of scientific thought and even in its claim to form a physical discipline. This is because any historical proliferation of theoretical proposals under a common conceptual framework should report certain progress in the capacity of the new proposals to provide us with meaningful knowledge within an empirical scope with respect to the capacity of the old proposals; otherwise, this theoretical production will be a mere inert intellectual exhibition. This is precisely the case of thermo-informationalism. My philosophical assessment of the conceptual foundations of informational thermo-physics leads us to characterize it as a degenerate scientific tradition. Likewise, the historical analysis of this intellectual current has shown that the diachronic succession of theoretical proposals does not lead to a real advance in the knowledge they provide about physical reality. According to Lakatos (1978, p.58; Musgrave & Pigden, 2021, Section 2.2), the historical persistence and its acceptance (within an active scientific community) of a degenerating program in the face of the existence of progressive programs within the same domain (e.g., BSM) constitutes a scientific but 'irrational' or 'uncritical' act on the part of its promoters.

Of course, that informational thermophysics has historically generated degenerative science is an historically contingent fact that would not justify defending (via a pessimistic meta-induction) that it will be impossible to develop in the future new informational proposals that can be epistemically fruitful in the empirical domain of thermophysics. My criticism is not so much that the classical thermo-informationalist program is a 'degenerate science', simpliciter, but properly that it is a degenerate science insofar as it claims to provide us with novel significant knowledge about the thermophysical reality. In this sense, I defend the plausibility of maintaining a Jaynesian attitude (Section 3.3.3) towards the whole informational research program, so that its intellectual aim was never to be a research program about a particular area of physical reality (competing against BSM or GSM) but properly about the cognitive dimension of scientific agents in their theoretical-experimental interaction with thermophysical phenomena[153]. That is, I argue that thermo-

[153] This is an interpretative line similar to the one that Maroney and Timpson (2018, p.103) seem to have adopted with respect to informational physics (not only classical, but also classical) focused on computational information

informationalism could have viability as a scientific program (i.e., generating theoretical and epistemic progress) devoted to developing a naturalized epistemology of classical thermostatistical physics, along the lines of what originally von Neumann envisioned in the 1940s (recall Section 3.1.2). Although if one conceives informational thermophysics as a naturalized epistemology (and therefore as a cognitive science and not a properly physical one), its historical development would only have generated a set of theoretically underdeveloped proposals so far. After seven decades, the consolidation of information physics as a naturalized epistemology remains an incomplete task that would require not only an enormous intellectual effort on the part of the physics community, but also close interdisciplinary collaboration with cognitive scientists and even philosophers.

However, my philosophical task here has not been to predict the future evolution of these of thermo-informational research programs, but to evaluate the conceptual foundations of the main intellectual products already historically generated from this current. In this sense, the radical novelty of this critical thesis compared to other critical proposals in the philosophical literature such as Carnap (1977), Earman and Norton (1999), Albert (2000), Norton (2005, 2011), Hemmo and Shenker (2012), Timpson (2013) or Wüthrich (2017) is that I have not only specified why certain particular informational theses (e.g., informational exorcisms, $S_{TD}$ and $H_{IT}$ identification, physical applicability of OI) are incorrect or incoherent. Standing on the shoulders of those giants, I have moreover historical-philosophically assessed how the systematic exploitation of information concepts in the last seven decades by different research programs has generated epistemically flawed proposals within the domain of classical thermostatistical physics. This is precisely my modest novel contribution to the rich literature on this thermophysics research program. Having finished these comments on the object of my analysis, I should end this chapter by reflecting on the methodological framework that has allowed us to carry out this evolution.

## 6.3. Critical Epistemic Role of Philosophy in Contemporary Physics

The last of my conclusions in this chapter is a defense (from the methodological framework and from the results obtained) of the critical contribution of HPS to contemporary physics. But before detailing this central feature of this procedure, let me point out what my pretensions have not been. The role of my historical analysis of this intellectual current (Chapters III and IV) has not been merely to describe a series of events or the evolution of certain terminological usages, but to offer certain interpretative keys to explore the intertwined web of intellectual goals, conceptual architectures and disciplinary motivations that underlies the history of informational thermophysics. Complementarily, the function of my conceptual assessment (Chapters II and V) of this line of scientific thought cannot be prescriptive, as I believe that philosophical analysis does

processing: "So information processing can continue to be studied in the abstract, independent of the means of physical instantiation. However, the interplay with physics means that sometimes physics will suggest new ways of analysing information processing tasks, enriching both fields, and logics which best match our current best physics will be the ones best suited for understanding the physical resource costs of real information processing devices."

not possess the legitimacy to tell a certain sector of the scientific community how to go about its work. By qualifying informational thermophysics as 'degenerative science' we do not prescribe how this program should have been properly carried out, I simply state a historical contingent fact about the epistemic achievements of this physical tradition.

Adopting Chang's (2004, p.237) and also Torretti's (2006) defense of HPS as 'complementary science' (in relation to 'specialist science'), the role of my historical-conceptual analysis can be neither descriptive nor prescriptive, but critical or properly critical-constructive: "Generating doubt may seem like the precise opposite of generating knowledge, but I would argue that constructive skepticism can enhance the quality of knowledge, if not its quantity" (Chang 2004, p.237). Authors such as Torretti (2006) argue that philosophical criticism can indeed contribute to progress in certain areas of physics. Particularly, the critical role of historical-conceptual analyses effectively contributes to generate knowledge not only historical (e.g., how certain conceptual uses evolve) or philosophical (e.g., the illegitimacy of certain metaphysical positions), but also scientific in a certain disciplinary domain, due to its capacity to access questions to which specialist science does not have access because of its characteristics. In this direction, Chang (2004, p.237) remarked that "HPS can recover useful ideas and facts lost in the record of past sciences, address foundational questions concerning present science and explore alternative conceptual systems (...) HPS can enlarge and deepen the pool of our knowledge about nature; in other words, HPS can generate scientific knowledge".

Assuming HPS as a complementary science, with this fine-grained historical-conceptual analysis of informational thermophysics I have modestly aimed to produce knowledge about certain conceptual architectures, interpretative scaffolds and epistemic strategies underlying one of the domains of modern physics. The generation of this type of 'complementary' knowledge is inaccessible to specialist physicists due to its institutional requirements, but it would be accessible to the history-philosophy of statistical thermophysics it has (as I have tried to show) the analytical and evaluative resources necessary to shed a minimum of light on the conceptual foundations of this discipline. Thanks to the critical and constructive function of this evaluation, I have obtained one of the central results of this work: namely, that the main uses of informational concepts in classical thermo-statistics during 1940s-2010s have been unable to provide us with significant knowledge about this physical domain. Modestly and being optimistic, these philosophical results could eventually contribute within the classical thermal physics community (i) to offer evaluative tools to detect the intellectual inflation (i.e., particularly when these are historically extended in several decades) of certain conceptual and interpretative uses, and, more importantly, (ii) to foster critical attitudes among the specialist classical statistical physics community.

In short, throughout this thesis I have tried to vindicate the importance of philosophy in bringing a critical character to the advancement of contemporary physics. It is a commonplace to state that there is a widespread perception of a recent proliferation of conspiratorial and pseudo-rational attitudes (e.g., the rise of flat-Earthers in the 2010s, COVID-19 denialism, etc.) in any sphere of public life today, with criticism being exercised as a mere banal act in the face of the epistemic status of science. It is another commonplace to assert that modern fundamental physics undoubtedly constitutes the intellectual sphere that provides the highest quality knowledge about our empirical reality[154]. What is not commonplace (and this is where this analysis comes into play) is to point out that there have been certain intellectual trends within modern physics, as is the case of informational thermophysics, where uncritical attitudes towards its constant lack of epistemic success allowed its historical persistence up to the present day. The fundamentally critical-constructive role that philosophical reflection has played regarding this broad intellectual movement, from Rudolf Carnap in the 1950s to Orly Shenker, John D. Norton or Christopher Timpson in recent years, has made it possible to detect the enormous intellectual inflation that this research program has undergone up to the present day. Therefore, this work should be understood as a modest sample as well as a defense of the critical capacity that the philosophy of physics possesses to generate a higher quality knowledge about the physical reality we inhabit.

[154] This idea is defended, for instance, by physicalist authors such as Ladyman and Ross (2007, Chapter 2).

# Bibliography

Aczél, J., & Daróczy, Z. (1975). *On Measures of Information and their Characterizations*. NY Academic Press.

Adriaans, P., (2020). Information, *The Stanford Encyclopedia of Philosophy* , Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2020/entries/information/>.

Albert, D. Z. (2000). *Time and Chance*. Harvard University Press.

Allori, V. (ed.) (2020). *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*. World Scientific.

Anta, J. (2021). Can informational thermal physics explain the approach to equilibrium? *Synthese* 1, 1-24.

Arabatzis, T. (1995). *The Electron: A Biographical Sketch of a Theoretical Entity*. Dissertation, Princeton University.

- (2006). *Representing Electrons. A Biographical Approach to Theoretical Entities*. University of Chicago Press.

Aspray, W. (1985). The Scientific Conceptualization of Information: A Survey. *Annals of the History of Computing*, 7 (2), 117-140.

Avery, J.S. (2012). *Information Theory and Evolution*. World Scientific.

Bais, F.A., & Farmer, J.D. (2007). *The Physics of Information*. arXiv:0708.2837v2, 1–65.

Balescu, R. (1997). *Statistical Dynamics: Matter out of equilibrium*. Imperial College Press.

Batterman, R. (2002). *The Devil in the Details*. Oxford University Press.

Baras, D., & Shenker, O. (2020). Calling for explanation: the case of the thermodynamic past state. *European Journal for Philosophy of Science* 10 (3), 1-20.

Bartels, A. (1995). Chains of meaning: A model for concept formation in contemporary physics theories. *Synthese* 105 (3), 347-379.

-(2010). Explaining Referential Stability of Physics Concepts: The Semantic Embedding Approach. *Journal for General Philosophy of Science / Zeitschrift für Allgemeine Wissenschaftstheorie* 41 (2), 267 - 281.

Bar-Hillel, Y. (1955). An examination of information theory. *Philosophy of Science* 22 (2), 86-105.

- (1964). *Language and information: Selected essays on their theory and application*. Reading, Addison-Wesley.

Bar-Hillel, Y., & Carnap, R. (1953). Semantic information. *The British Journal for the Philosophy of Science*, 4, 147–157.

Bekenstein, J. D. (1972). Black holes and the second law. *Lettere al Nuovo Cimento*, 4, 737-740.

- (1973). Black Holes and Entropy. *Physical Review D*, 7, 2333-2346.

- (2006). *Of gravity, black holes and information*. Di Renzo Editore.

Belfer, I. (2014). Informing Physics: Jacob Bekenstein and the Informational Turn in Theoretical Physics. *Physics in Perspective* 16, 69-97.

Bennett, C.H. (1982). The thermodynamics of computation - a review. *International Journal of Theoretical Physics* 21, 905–940. http://www.research.ibm.com/people/b/bennetc /bennettc1982666c3d53.pdf

- (1987). Demons, Engines and the Second Law, *Scientific American* 257, 108—116.

- (1988). Notes on the History of Reversible Computation, *IBM Journal of Research and Development* 32, 16—23. Reprinted in Leff and Rex (1990), pp. 281—288.

- (1990). Dissipation, anisotropy, and the stabilization of computationally complex states of homogeneous media. *Physica A: Statistical and Theoretical Physics 163*(*1*), 393-397.

- (2003). Notes on Landauer's principle, reversible computation, and Maxwell's Demon. *Studies in History and Philosophy of Modern* Physics 34(3), 501–510.

Bérut, A; Arakelyan, A; Petrosyan, A.; Ciliberto, S; Dillenschneider, R., & Lutz, E. (2012). Experimental verification of Landauer's principle linking information and thermodynamics. *Nature* 48, pp. 187-190.

Ben-Naim, A. (2008). *A Farewell to Entropy. Statistical Thermodynamics Based on Information*. World Scientific.

Berkovitz, J; Frigg, R & Kronz, F. (2006). The ergodic hierarchy, randomness and Hamiltonian chaos. *Studies in History and Philosophy of Modern Physics* 37 (4), 661-691.

Bohm, D. (1989). Meaning and Information. In Pylkkänen ed. *The Search for Meaning*. Wellingborough: Crucible

Boltzmann, L. (1872). Weitere Studien uber das W ¨ armegleichgewicht unter Gasmolekulen. *Wiener Berichte,* 66, 275–370. In (Boltzmann 1909) Vol. I, paper 23.

- (1877). Bermerkungen uber einige Probleme der mechanische Warmetheorie *Wiener Berichte*, 75, 62–100. In (Boltzmann 1909) Vol. II, paper 39.

- (1896). *Lectures on gas theory*. University of California Press.

Bridgman, P. W. (1927). *The Logic of Modern Physics*. Macmillan.

Brillouin, L. (1949). Life, Thermodynamics, and Cybernetics. *American Scientist* 37 (4), 554-568

- (1951). Maxwell's Demon Cannot Operate: Information and Entropy. I, *Journal of Applied Physics* 22, 334—337. Reprinted in Leff and Rex (1990), pp. 134—137. 'Maxwell's Demon Cannot Operate: Information and Entropy. II', *Journal of Applied Physics* 22, 338—343.

- (1953). The Negentropy Principle of Information, *Journal of Applied Physics* 24, 1152—1163.

- (1956). *Science and Information Theory*. Academic Press.

- (1962). *Science and Information Theory: Second Edition*. Dover.

- (1964). *Scientific Uncertainty and Information*. Academic Press.

Brudno, A. A. (1978): The Complexity of the Trajectory of a Dynamical System. *Russian Mathematical Surveys* 33, 197-98.

Burgers, J. (1954). Entropy and disorder, *British Journal of Philosophy of Science* (5) 70-71.

Butterfield, J.. & Earman, J. (eds.) (2007). *Philosophy of Physics*. Elsevier.

Bynum, T.W. (2008). Norbert Wiener and the rise of information ethics. In M. J. van den Joven & J. Weckert (eds.), *Information Technology and Moral Philosophy*. Cambridge University Press.

Caves, C. M. (1990) Quantitative Limits on the Ability of Maxwell Demon to Extract Work from Heat', *Physical Review Letters* 64, 2111—2114.

- (1993) Information and entropy, *Physical Review* E47, 4010—4017.

- (1994) Information, Entropy and Chaos, in J. J. Halliwell, J. Perez Mercader and W. H. Zurek (eds), Physical Origins of Time Asymmetry. *Cambridge: Cambridge University Press*, pp. 47—87.

Callender, C. (1999). Reducing thermodynamics to statistical mechanics: The case of entropy. *Journal of Philosophy* 96 (7): 348-373.

- (2011). The past histories of molecules. In Claus Beisbart & Stephan Hartmann (eds.), *Probabilities in Physics*. Oxford University Press. pp. 83-113.

Carnap, R.. & Bar-Hillel, Y. (1952). An Outline of a Theory of Semantic Information, *Technical Report* No. 247, Research Laboratory of Electronics, MIT.

Carnap, R. (1945). Two Concepts of Probability, *Philosophy and Phenomenological Research*, 5, 513-532.

- (1963). The Philosophy of Rudolf Carnap. In Paul Arthur (ed.) *The Library of Living Philosophers 11*. Open Court.

- (1977). *Two Essays on Entropy*. Shymony (eds.), University of California Press.

Carnot. S. (1824). *Reflexions sur la puissance motrice du feu*. Bachelier. Re-edited and translated in E. Mendoza (Ed.) (1960*). Reflections on the motive power of fire*, Dover.

Chaitin, G. J. (1966). On the Length of Programs for Computing Finite Binary Sequences, *J. Association for Computing Machinery*, v. 13, No. 4, 547–569.

- (1975a). Randomness and Mathematical Proof. *Scientific America*.

- (1975b). A Theory of Program Size Formally Identical to Information Theory. *Journal of the ACM*, 22 (3), 329–340https://doi.org/10.1145/321892.321894

- (1977). Algorithmic information theory, *IBM Journal of Research and Development*, v.21, No. 4, 350–359.

- (1987). *Information, Randomness & Incompleteness*. World Scientific.

Chang, H. (2004). *Inventing Temperature: Measurement and Scientific Progress*. Oup Usa.

-(2019) "Operationalism", *The Stanford Encyclopedia of Philosophy* (2019 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2019/ entries/operationalism/>.

Cherry, C. (1957). *On Human Communication: A Review, a Survey, and a Criticism.* MIT Press.

Clausius, R. (1850) *On the Motive Power of Heat, and on the Laws which can be deduced from it for the Theory of Heat*. Poggendorff's *Annalen der Physick*, LXXIX. Dover Reprint.

- (1854). Ueber eine veränderte Form des zweiten Hauptsatzes der mechanischen Wärmetheoriein". *Annalen der Physik und Chemie*. 93 (12): 481–506.

- (1856). *On the Application of the Mechanical theory of Heat to the Steam-Engine*. In: Clausius, R. (1865). The Mechanical Theory of Heat – with its Applications to the Steam Engine and to Physical Properties of Bodies. London: John van Voorst, 1 Paternoster Row. MDCCCLXVII.

- (1862). Ueber die Anwendung des Satzes von der Aequivalenz der Verwandlungen auf die innere Arbeit, Viertelsjahrschrift der Zuricher naturforschenden Gesellschaft, 7, 48. Also in Annalen der Physik und Chemie, 116: 73–112 (1862), English translation in Philosophical Magazine, 24, 81–97, 201–213.

- (1865). *The Mechanical Theory of Heat – with its Applications to the Steam Engine and to Physical Properties of Bodies*. London: John van Voorst, 1 Paternoster Row. MDCCCLXVII.

Conway, F., & Siegelman, J. (2005). *Dark hero of the information age: in search of Norbert Wiener, the father of cybernetics*. Basic Books.

Costa de Beauregard, O., & Tribus, M. (1974). Information theory and thermodynamics, *Helv. Phys Acta* 47, 238 247.

Cover, T., & Thomas, J. (1991). *Elements of Information Theory*. Wiley.

Cox, R. T. (1946). Probability, Frequency and Reasonable Expectation. *American Journal of Physics*. 14, 1–10.

Cuffaro, M. E., & Fletcher, S.C (eds.) (2018). *Physical Perspectives on Computation, Computational Perspectives on Physics*. Cambridge University Press.

Cyranski, J.F. (1978). Analysis of the maximum entropy principal “debate”. *Foundations of Physics* 8 (5-6): 493-506.

Davies, P. (1990). Why is the Physical World so Comprehensible? In Zurek, W. (ed.) (1990a). *Complexity, Entropy, and the Physics of Information*. Addison-Wesley.

Davies, P., & Gregersen, N. H. (2010). *Information and the Nature of Reality. From Physics to Metaphysics*. Cambridge University Press.

Denbigh, K. G. (1981). How Subjective is Entropy? *Chemistry in Britain* 17, 168—185. Reprinted in Leff and Rex (1990), pp. 109—115.

- (1989) Note on Entropy, Disorder and Disorganization. *The British Journal for the Philosophy of Science*, 40, 323-332.

Denbigh, K. G., & Denbigh, J. S. (1985). *Entropy in Relation to Incomplete Knowledge*. Cambridge University Press.

Dennett, D. (1991). Real Patterns. *The Journal of Philosophy*, Vol. 88 (1)., 27-51

Deutsch, D. (1985). Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer, *Proceedings of the Royal Society of London A*, 400: 97–117.

Dias, P., & Shimony, A. (1981). A Critique of Jaynes’ Maximum Entropy Principle, *Advances in Applied Mathematics* 2, 172–211.

Dretske, F. (1981). *Knowledge and the Flow of Information*. MIT Press.

Duncan, T. L., & Semura, J. S. (2007). Information Loss as a Foundational Principle for the Second Law of Thermodynamics. Foundations of Physics 37, 1767-1773.

Earman, J., & Norton, J. (1998). Exorcist XIV: the wrath of Maxwell's demon. Part I. From Maxwell to Szilard. *Studies in History and Philosophy of Modern Physics* 29 (4), 435-471.

- (1999). Exorcist XIV: the wrath of Maxwell’s Demon. Part II. From Szilard to Landauer and beyond. *Studies in History and Philosophy of Modern Physics* 30(1), 1–40.

Ehrenfest, P. & Ehrenfest-Afanassjewa, T. (1911). *Begriffliche Grundlagen der Statistischen Auffassung in der Mechanik. Enzyclopadie der Mathematischen Wissenschaften* Klein. F. and Muller, C. (eds.). Teubner, 3–90. English translation (1959). *The conceptual foundations of the statistical approach in mechanics*. Cornell University Press.

Ellis, R. (2006). *Entropy, Large Deviations, and Statistical Mechanics*. Springer.

Feynman, R. (1963). Ratchet and pawl. In *The Feynman lectures on physics*, Addison Wesley.

- (1986). Quantum Mechanical Computers, *Found. Phys*. 16, 507-531.

- (1996). *Feynman Lectures on Computation* (eds) A. J. G. Hey and R. W. Allen. Addison—Wesley.

Fisher, R. A. (1934). Probability, Likelihood and Quantity of Information in the Logic of Uncertain Inference. *Proceedings of the Royal Society A*. 146 (856), 1–8.

Floridi, L. (2011). *The Philosophy of Information*. Oxford University Press.

- (2019). *The Logic of Information*. Oxford University Press.

Fredkin, E., & Toffoli, T. (1982). Conservative logic. International Journal of Theoretical Physics, 21, 219.

Frieden, R. (1999). *Physics from Fisher Information*. Cambridge University Press.

- (2004). *Science from Fisher Information: A Unification*. Cambridge University Press.

Friedman, K. (1973). Replies to Tribus and Motroni and to Gage and Hestenes, *Journal of Statistical Physics* 9, 265–69.

Friedman, K., & Shimony, A. (1971). Jaynes' Maximum Entropy Prescription and Probability Theory, *Journal of Statistical Physics* 3, p.381–384.

Frigg, R. (2004). In What Sense is the Kolmogorov-Sinai Entropy a Measure for Chaotic Behaviour? Bridging the Gap Between Dynamical Systems Theory and Communication Theory. *British Journal for the Philosophy of Science* 55 (3), 411-434.

- (2008). A field guide to recent work on the foundations of statistical mechanics. In Dean Rickles (ed.), *The Ashgate Companion to Contemporary Philosophy of Physics*. (pp. 991-96). Ashgate.

Frigg, R., & Hartmann, S. (2006). Models in Science. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Stanford.

Frigg, R., & Hoefer, C (2015). The Best Humean System for Statistical Mechanics. *Erkenntnis* 80 (3), 551-574.

Frigg, R., & Werndl, C. (2011a). Entropy - A Guide for the Perplexed. In Claus Beisbart & Stephan Hartmann (eds.), *Probabilities in Physics*. Oxford University Press. pp. 115-142.

Frigg, R., & Werndl, C. (2011b). Explaining Thermodynamic-Like Behavior in Terms of Epsilon-Ergodicity. *Philosophy of Science* 78 (4), 628-652.

Frigg, R., & Werndl, C. (2012). A new approach to the approach to equilibrium. In Yemima Ben-Menahem & Meir Hemmo (eds.), *Probability in Physics. The Frontiers Collection*. Springer. pp. 99-114.

-(2019). Statistical Mechanics: A Tale of Two Theories. *The Monist* 102 (4), 424-438.

Gabor, D. (1946). *Theory of communication*. London: Institution of Electrical Engineering

- (1954). Communication Theory and Cybernetics, *Transactions of the IRE. Professional Group on Circuit Theory* 1, no. 1: 19–31.

Gibbs, J.W. (1902). *Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundation of Thermodynamics*. Yale University Press.

Gleick, J. (2011). *The Information: A History, a Theory, a Flood*. Pantheon Books.

Goldstein, S. (2019). Individualist and Ensemblist Approaches to the Foundations of Statistical Mechanics. *The Monist*, 102, 439–457.

Goldstein, S. Lebowitz, J, Tumulka, R., & Zanghi, N. (2020). Gibbs and Boltzmann Entropy in Classical and Quantum Mechanics. In Allori, V., (Eds.), *Statistical Mechanics and Scientific Explanation Determinism, Indeterminism and Laws of Nature-World Scientific*. (pp.519-581), World Scientific.

Grunbaum, A. (1973). Is the Coarse-Grained Entropy of Classical Statistical Mechanics an Anthropomorphism?, In his *Philosophical Problems of Space and Time*, Springer, 646–65.

Haglund, J. (2017). Good Use of a 'Bad' Metaphor. Entropy as Disorder, *Sci & Educ* 26, 205-214.

Harshman, N. L. (2016). Physics and information. In Floridi. L. (eds.) *The Routledge Handbook of Philosophy of Information*, 7-14. Routledge.

Hartley, R. (1928). Transmission of information. *Bell System Technical Journal*, 7, 535–563.

Hayles, N.K. (1990). *Chaos bound: orderly disorder in contemporary literature and science*, Cornell University Press.

Hawking, S. (1974). Black hole explosions?. *Nature*. 248 (5443), 30–31.

Helmholtz, H. (1983). *Wissenschaftliche Abhandlungen* (Barth, Leipzig, 1883), Vol. 2, p. 972. The paper, entitled "The thermodynamics of chemical processes," was reprinted from the Sitzungsberichten der Akademie der Wissenschaften. J.A. Barth.

Hemmo, M., & Shenker, O. (2006). Von Neumann’s Entropy Does Not Correspond to Thermodynamic Entropy. *Philosophy of Science* 73 (2), 153-174.

- (2012). *The Road to Maxwell’s Demon*. Cambridge University Press.

- (2019). The physics of implementing logic: Landauer's principle and the multiple-computations theorem”, *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 68, 90-105.

Hoefer, C. (2019). *Chance in the World: A Humean Guide to Objective Chance*. Oup Usa.

Hoefer, C., & Martí, G., (2020). Realism, reference and perspective. *European Journal for Philosophy of Science*. 10 (3), 1-22.

Hosoya, A., & Shunsuke, F. (2018). Informational theory of relativity, *Progress of Theoretical and Experimental Physics.* 2018(11), p.1-33.

Isaac, M. A. (2019). *The Semantics Latent in Shannon Information*. British Journal for the Philosophy of Science 70 (1), 103-125.

Jauch, J. M. & Baron, J. G. (1972). Entropy, information and Szilard's paradox. *Helv. Phys. Acta* 45, 220 – 232.

Jaynes, E.T. (1957a). Information theory and statistical mechanics, *Phys. Rev*. 106, 620-630.

- (1957b). Information theory and statistical mechanics. II, *Phys. Rev*. 108, 171-190.

- (1965). Gibbs vs Boltzmann Entropies, *American Journal of Physics*, 33, 391–8.

- (1978). Where do we stand on Maximum Entropy? In *Proceedings of the Maximum Entropy Formalism Conference*, Massachusetts Institute of Technology.

- (1983). *Papers on Probability, Statistics and Statistical Physics*; Ed. by R. D. Rosenkrantz, D. Reidel publishing Co., Dordrecth, Holland.

- (1990). Probability in Quantum Theory; in *Complexity, Entropy, and the Physics of Information*, W. H. Zurek (ed.), Addison-Wesley.

Jaynes, E.T. (2003). *Probability Theory: The Logic of Science*. Cambridge University Press.

Katz, A. (1967). *Principles of Statistical Mechanics: The Information Theory Approach*. Front Cover.

Khinchin, A. I. (1949). *Mathematical Foundations of Statistical Mechanics*. Dover Publications.

- (1957). *Mathematical Foundations of Information Theory*. Dover Publications.

Klein, M. (1972). Mechanical Explanation at the End of the Nineteenth Century. *Centaurus* 17(1), 58-82

Kline, R. (2015). *The Cybernetics Moment: Or Why We Call Our Age the Age of Information*, Johns Hopkins University Press.

Kindi, V., & Arabatzis, T. (2008). The Problem of Conceptual Change in the Philosophy and History of Science. In Stella Vosniadou (ed.), *Handbook of Research on Conceptual Change*. Routledge.

Köhler, E. (2001). Why von Neumann Rejected Carnap’s Dualism of Information Concepts. In Rédei M and Stölzner (eds.) *John con Neumann and the Foundations of Quantum Physics*, 97-134. Kluwer Academic Publisher.

Kolmogorov, A. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. English translation by Morrison, N. in (1956) *Foundations of the theory of probability,* Chelsea Publishing Company.

- (1965). Three approaches to the quantitative definition of information. *Problems of Information Transm.*, 1, p.4–7.

- (1983). Combinatorial foundations of information theory and the calculus of probabilities. *Russian Mathematical Survey*, 38 (4), 29-40.

Ladyman, J. (2018). Intension in the Physics of Computation: Lessons from the Debate about Landauer's Principle. In *Physical Perspectives on Computation, Computational Perspectives on Physics*, Cuffaro, E. and Fletcher, S. (eds.) Cambridge University Press, 219-239.

Ladyman, J, Presnell, S, Short, A. J., & Groisman, B (2007). The connection between logical and thermodynamic irreversibility, *Studies in the History and Philosophy of Modern Physics*, 38, 58–79.

Ladyman, J, and Ross, D. (2007). *Every Thing Must Go. Metaphysics Naturalized*. Oxford University Press.

Ladyman, J. Ross, D., & Kincaid, H. (2013). *Scientific Metaphysics*. Oxford University Press.

Ladyman, J. and Robertson, K. (2013). Landauer defended: Reply to Norton. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 44 (3), 263-271.

Lakatos, I. (1978). *The Methodology of Scientific Research Programmes* (*Philosophical Papers*: Volume 1), J. Worrall and G. Currie (eds.), Cambridge University Press.

Landauer, R. (1961). Irreversibility and Heat Generation in the Computing Process, *IBM Journal of Research and Development* 5, 183—91. Reprinted in Leff and Rex (1990), pp. 188—196.

- (1991). Information is Physical. *Physics Today* 44, 5, 23.

- (1999). Information is Inevitably Physical. Submitted to *Feynman Lectures on Computation*, vol. 2, edited by A.J.G. Hey. Addison Wesley.

Larsen, K. (2005). *Stephen Hawking: A Biography*. Greenwood Publishing Group.

Laudan, L. (1977). *Progress and Its Problems*. University of California Press.

Lavis, D. A. (2005). Boltzmann and Gibbs: An attempted reconciliation. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 36 (2), 245-273.

- (2011). An objectivist account of probabilities in statistical physics. In Claus Beisbart & Stephan Hartmann (eds.), *Probabilities in Physics*. Oxford University Press. p.51-72.

Lazarovici, D. (2019). On Boltzmann versus Gibbs and the Equilibrium in Statistical Mechanics. *Philosophy of Science* 86 (4), 785-793.

Lazarovici, D. & Reichert, P. (2015). Typicality, Irreversibility and the Status of Macroscopic Laws. *Erkenntnis* 80 (4), 689-716.

Lebowitz, J. (1993). Macroscopic laws, microscopic dynamics, time's arrow and Boltzmann's entropy. *Physica A: Statistical Mechanics and its Application*s 194 (1–4), 1-27.

Leff, H. S. & Rex, A. F. (1990). *Maxwell's Demon: Entropy, Information, Computing*. Princeton University Press.

- (1994). Entropy of Measurement and Erasure: Szilard's Membrane Model Revisited, *American Journal of Physics* 62, 994—1000.

- (2003). *Maxwell's Demon 2: Entropy, Classical and Quantum Information, Computing*, Instit. of Physics Publishing.

Leff, H. (2007). Entropy, Its Language, and Interpretation, *Found Phys* (37), 1744-1766.

Lewis, G. N. (1930). The symmetry of time in physics, *Science* 71, 568-577.

Lewis. D. (1994). Humean Supervenience Debugged, *Mind*, 103: 473–490.

Lloyd, S. (1988). *Black Holes, Demons and the Loss of Coherence: How complex systems get information, and what they do with it* (Ph.D. thesis). The Rockefeller University

- (2000). Ultimate physical limits to computation. *Nature*. 406 (6799). 1047–1054.

- (2006). *Programming the Universe: A Quantum Computer Scientist Takes On the Cosmos*, Knopf.

Loewer, B. (2004). David Lewis's Humean Theory of Objective Chance. *Philosophy of Science* 71 (5), 1115-1125.

Lombardi, O; Fortin, S., & López, C. (2016a). Deflating the deflationary view of information. *European Journal for Philosophy of Science* 6 (2), 209-230.

Lombardi, O., Holik, F., & Vanni, L. (2016b). What is Shannon information? *Synthese* 193 (7), 1983-2012.

MacKay, D. (1969). *Information, Mechanism and Meaning*. MIT Press.

Maher, P. (2007). Explication Defended, *Studia Logica*, 86(2): 331–341. doi:10.1007/s11225-007-9063-8

Malament, D., & Zabell, S. (1980). Why Gibbs phase averages work. The role of ergodic theory. *Philosophy of Science* 47 (3), 339-349.

Mari, L. (1999). Notes towards a qualitative analysis of information in measurement results, *Measurement*, 25(3), 183–192.

Maroney, O. (2009). Information Processing and Thermodynamic Entropy, *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2009/entries/information-entropy/>.

Martínez, M. (2019). Representations are Rate-Distortion Sweet Spots. *Philosophy of Science* 89, p.1214-226.

Maxwell, J. C. (1878). 'Diffusion', Encyclopedia Britannica (9th ed) 7, 214. Reprint in Maxwell (1953), pp.625-646.

Mbugua, K. (2015). Explaining Same-Sex Sexual Behavior: The Stagnation of the Genetic and Evolutionary Research Programs. *Journal for General Philosophy of Science* 46(1), 23-43.

McCoy, C. D. (2020). An Alternative Interpretation of Statistical Mechanics. *Erkenntnis* 85 (1), 1-21.

Mindell, D.A. (2004). *Between Human and Machine Feedback, Control, and Computing before Cybernetics*. Johns Hopkins University Press.

Misner, C; Thorne, K., & Zurek, W. (2009). John Wheeler, relativity, and quantum information. *Physics Today*. 62 (4), 40–46.

Myrvold, W. (2011). Statistical mechanics and thermodynamics: A Maxwellian view. *Studies in History and Philosophy of Science Part A* 42 (4):237-243.

- (2016). Probabilities in Statistical Mechanics. In Christopher Hitchcock & Alan H'Ajek (eds.), *The Oxford Handbook of Probability and Philosophy*. Oxford University Press. p. 573-600.

Musgrave, A., & Pigden, C. "Imre Lakatos", *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2021/entries/lakatos/>.

Nersessian, N. (2008), *Creating Scientific Concepts*, MIT Press.

Nickles, T. (2017). Historicist Theories of Scientific Rationality, *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2021/entries/rationality-historicist/>.

Nolte, D. D. (2018). *Galileo Unbound. A Path Across Life, the Universe and Everything*. Oxford University Press.

Norton, J. D. (2005). Eaters of the lotus: Landauer's principle and the return of Maxwell's demon. *Studies in the History and Philosophy of Modern Physics*, 36, 375–411.

- (2011a). Waiting for Landauer. Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 42 (3), 184-198.

- (2011b). Challenges to Bayesian Confirmation Theory. In Prasanta S. Bandyopadhyay & Malcolm R. Forster (eds.), *Handbook of the Philosophy of Science, Vol. 7: Philosophy of Statistics*. Elsevier B.V. pp. 391-440.

- (2012). Approximation and Idealization: Why the Difference Matters. *Philosophy of Science* 79 (2), 207-232.

- (2013). The End of the Thermodynamics of Computation: A No Go Result. *Philosophy of Science* 80 (5),1182-1192.

- (2016). The impossible process: Thermodynamic reversibility. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 55, 43-61.

- (2018). Maxwell's Demon Does not Compute. In Michael E. Cuffaro and Samuel C. Fletcher, (eds.), *Physical Perspectives on Computation, Computational Perspectives on Physics*. Cambridge University Press.

Nyquist, H. (1924). Certain Factors Affecting Telegraph Speed. *Bell System Technical Journal*, 3, 324-346.

Parker, D. (2011). Information-Theoretic Statistical Mechanics without Landauer's Principle. *British Journal for the Philosophy of Science* 62 (4), 831-856.

Pauli, W. (1999). *Wissenchaflicher Briefwechsel Wolfgang Pauli smit Bohr, Einstein, Heisenberg und anderen*, vol. 4.2, ed. By Karl von Meyenn, Springer-Verlag.

Penrose, R. (1970). *Foundations of Statistical Mechanics*. Oxford University Press.

- (1998). Quantum computation, entanglement and state reduction, *Phil. Trans. R.Soc. Lond.A*, 356(1743), 1927–1939.

- (2006). *The Road to Reality: A Complete Guide to the Laws of the Universe*. Vintage.

- (2016). *Fashion, Faith, and Fantasy in the New Physics of the Universe*. Princeton University Press.

Piccinini, G. & Maley, C. (2021). Computation in Physical Systems, *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2021/entries/computation-physicalsystems/>.

Piechocinska, B. (2000) Information Erasure, *Physical Review* A, 61, 62314, 1-9; rep in Leff and Rex (2003, Ch. 4.4).

Pincock, C. (2012). *Mathematics and Scientific Representation*. Oxford University Press USA.

Popper, K.R. (1959). The propensity interpretation of probability. *British Journal for the Philosophy of Science* 10 (37), 25-42.

Prawitz, D. (2012). The epistemic significance of valid inference. *Synthese* 187 (3):887-898.

Prigogine, I. (1980). *From Being to Becoming: Time and Complexity in the Physical Sciences*. W. H. Freeman.

Psillos, S. (2012b). Causal Descriptivism and the Reference of Theoretical Terms. *Perception, Realism, and the Problem of Reference*. Raftopoulos, A., & Machamer, P. (Eds.). Cambridge University Press.

de Regt, H. (2017). *Understanding Scientific Understanding*. Oup Usa.

Raymond, R. (1950). Communication, Entropy, and Life, *American Scientist* 38, 273—278.

- (1951). The Well-Informed Heat Engine, *American Journal of Physics* 19, 109—112. Reprinted in Leff and Rex (1990), pp. 138—141.

Reichenbach, H. (1949). Philosophical Foundations of Probability, *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press: 1–20. Reprinted in S. Sarkar (ed.) *Logic, Probability and Induction,* Garland.

Rodd, P. (1964). Some Comments on Entropy and Information, *American Journal of Physics* 32, 333—335. Reprinted in Leff and Rex (1990), pp. 145—147.

Ridderbos, K. (2002). The coarse-graining approach to statistical mechanics: How blissful is our ignorance? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 33 (1), 65-77.

Robertson, K. (2020). Asymmetry, Abstraction, and Autonomy: Justifying Coarse-Graining in Statistical Mechanics. *British Journal for the Philosophy of Science* 71 (2), 547-579.

Rothstein, J. (1951). Information, Measurement, and Quantum Mechanics. *Science*, 114, 171-175.

- (1952). Information and Thermodynamics, *Physics Review*, 85: 135.

Rovelli, C. (1996). Black Hole Entropy from Loop Quantum Gravity. *Physical Review Letters*. 77 (16).

Ryle, G. (1949). *The Concept of Mind*. Hutchinson & Co.

Schrödinger, E. (1994). *What is Life?* (11th reprinting ed. 2004). Canto.

Schumacher, B. (1995). Quantum coding. *Physical Review A: Atomic, Molecular and Optical Physics*, 51, 2738–2747.

Seife, C. (2007). *Decoding the Universe: How the New Science of Information is Explaining Everything in the Cosmos from our Brains to Black Holes*. Ed Penguin Group.

Segal, J. (2003). *Le zéro et le un: Histoire de la notion scientifique d'information au 20e siècle*, Éditions Syllepse.

Shannon, C. (1948). *A Mathematical Theory of Communication*. In *Collected Papers*. Sloane NJA and Wyner AD (eds), IEEE Press.

-(1956). The Bandwagon. In *Collected Papers*. Sloane, N.J.A and Wyner, A.D (eds), IEEE Press, 1993.

Shannon, C. & Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press.

Shapere, D. (1985). Reason and the Search for Knowledge. *Philosophy of Science* 52 (2), 310-312.

Shenker, O. (2000). Logic and Entropy. http://philsci-archive.pitt.edu/id/eprint/115.

- (2020). Information vs. entropy vs. probability. *European Journal for Philosophy of Science* 10 (1), 1-25.

Shenker, O., & Hemmo, M. (2012). *The Road to Maxwell's Demon*. Cambridge University Press.

Shimony, A. (1973). Comment on the Interpretation of Inductive Probabilities, *Journal of Statist. Physics* 9, 187–91.

- (1985). The status of the principle of maximum entropy. *Synthese* 63, 35–53.

Styer, D.F. (2008). Entropy and evolution, *American Journal of Physics*, 76, 1031–1033

Sklar, L. (1993). *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics,* Cambridge University Press.

Skyrms, B. (1981). Maximum Entropy Inference as a Special Case of Conditionalization. (unpublished).

- (2010). *Signals: Evolution, Learning, and Information*. Oxford University Press.

Smoluchowski, M. (1912). 'Experimentell nachweisbare, der üblichen Thermodynamik widerprechende Molekularphänomene', Physikalische Zeitschrift 13, 1069-1080.

Sokal, A. (1996). Revelation: A Physicist Experiments With Cultural Studies. In *Editors of Lingua Franca*, 49–54.

Solomonoff, R. (1964). A formal theory of inductive inference. *Information and Control*, 7(1–22), 224–254.

Stonier, T. (1990). *Information and the Internal Structure of the Universe: An Exploration Into Information Physics.* Springer-Verlag.

- (1992). *Beyond Information: The Natural History of Intelligence*. Springer-Verlag.

- (1997). *Information and Meaning: An Evolutionary Perspective*. Springer-Verlag.

Suárez, M. (2004). An Inferential Conception of Scientific Representation. *Philosophy of Science,* 71, pp. 767-79.

Suárez, M., & Pero, F. (2019). The Representational Semantic Conception. *Philosophy of Science* 86 (2), 344-365.

Swanson, J. (1960). Physical versus Logical Coupling in Memory System. *IBM Journal* 4, 3: 305-310.

Szilard, L. (1929) On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings', in *The Collected Works of Leo Szilard: Scientific Papers* (MIT Press, 1972), pp. 120—129. Reprinted in Leff and Rex (1990), pp. 124—133.

Ter Haar, D. (1954). *Foundations of Statistical Mechanics*. Constable.

Thims, L. (2012). Thermodynamics ≠ Information Theory: Science's Greatest Sokal Affair, *Journal of Human Thermodynamics*, 8(1):1-120

Timpson, C. G. (2013). *Quantum Information Theory and the Foundations of Quantum Mechanics*. Oxford University Press.

Tisza, L. & Quay, P. (1963). The statistical thermodynamics of equilibrium. *Ann. Phys*. 25:48–90 (1963)

Torretti, R. (2006). *Can science advance effectively through philosophical criticism and reflection?* Preprint. http://philsci-archive.pitt.edu/2875/

Tribus, M. (1963). *Information theory and thermodynamics, Heat Transfer, Thermodynamics and Education*: Reedited (1979) in (eds. Johnson, H. and Boelter, L) Anniversary Volume, 348-68.

Tribus, M., &. McIrvine, E. (1971). Energy and information, *Scientific America*. 225, (1971), 179-188.

Trout, J. D. (2002). Scientific explanation and the sense of understanding. *Philosophy of Science* 69 (2):212-233.

Tolman, R. C. (1938). *The Principles of Statistical Mechanics*. Oxford University Press. Reissued (1979). Dover.

Uffink, J. (1995). Can the maximum entropy principle be explained as a consistency requirement? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 26 (3), 223-261.

- (2007). Compendium of the foundations of classical statistical physics. In J. Butterfield, & J. Earman, (Eds.), *Handbook for philosophy of physics*. Elsevier.

- (2011). Subjective probability and statistical physics. In Claus Beisbart & Stephan Hartmann (eds.), *Probabilities in Physics*. (pp. 25-49) Oxford University Press.

- (2014). Boltzmann's Work in Statistical Physics, *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2017/entries/statphys-Boltzmann/>.

van Lith, J. (2001). Ergodic theory, interpretations of probability and the foundations of statistical mechanics. *Studies in History and Philosophy of Modern Physics* 32 (4), 581-94.

Vedral, V. (2010) *Decoding Reality: The Universe as Quantum Information*. Oxford University Press.

Von Helmholtz, H. (1882) *Wissenschaftliche Abhandlungen*, Three volumes: second volume 1883, third volume 1895, Johann Ambrosius Barth.

Von Mises, R. (1957). *Probability, Statistics and Truth*. Revised English edition, Macmillan.

Von Neumann, J. (1933) *Mathematische Grundlagen der Quantenmechanischer Gesamtheiten*, Springer-Verlag; translated (1955) by Beyer R as *Mathematical Foundations of Quantum Mechanics* Princeton University Press.

- (1949). The Role of High and of Extremely High Complication, Fourth University of Illinois lecture, in *Theory of Self-Reproducing Automata*, Burks AW (ed.), Reprinted in (1966) University of Illinois Press, pp. 64–73.

- (1958). *The Computer and the Brain*. Yale University Press.

von Plato, J. (1989). De finetti's earliest works on the foundations of probability. *Erkenntnis* 31 (2-3), 263-282.

- (1994). *Creating Modern Probability*. Cambridge University Press.

Wallace, D. (2012). The necessity of Gibbsian statistical mechanics. *PhilSci archive*. http://philsci-archive.pitt.edu/15290/.

- (2017). Inferential vs. Dynamical Conceptions of Physics. in O. Lombardi, S. Fortin, F. Holik, C. López (eds.), *What Is Quantum Information?* Cambridge University Press, 179-204.

- (2020). The necessity of Gibbsian statistical mechanics. In Allori, V., (Eds.), *Statistical Mechanics and Scientific Explanation Determinism, Indeterminism and Laws of Nature-World Scientific*. World Scientific: 583-616.

Wheeler, J. (1990). Information, Physics, Quantum: The Search for Links. In Zurek, W. (Eds.), *Complexity, Entropy, and the Physics of Information*. Addison-Wesley.

- (1998). *Geons, Black Holes, and Quantum Foam: A Life in Physics*. W.W. Norton & Co.

Wheeler, J. A., & Zurek, W. H. (eds) (1983). *Quantum Theory and Measurement*. Princeton University Press.

Wicken, J. (1987). Entropy and information: Suggestions for common language. *Philosophy of Science* 54 (2): pp.176-193

Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. Cambridge University Press.

- (1950). *The Human Use of Human Beings*. DaCapo Press.

Wilhelm, I. (2019). Typical: A Theory of Typicality and Typicality Explanation. *British Journal for the Philosophy of Science*:

Wilson, M. (2006). *Wandering Significance: An Essay on Conceptual Behavior*. Clarendon Press.

- (2017). *Physics Avoidance: And Other Essays in Conceptual Strategy*. Oxford University Press.

Wright, A. (2016). The Physics of Forgetting: Thermodynamics of Information at IBM 1959–1982, *Perspectives on Science* 24, no. 1 (2016), 112–141.

Wüthrich, C. (2019). "Are Black Holes about Information?" In. Richard Dawid, Karim Thébault, and Radin Dardashti (eds.), *Why Trust a Theory? Epistemology of Fundamental Physics*, Cambridge University Press, 202-223.

Yockey, H. (1992). *Information Theory and Molecular Biology*. Cambridge University Press.

Zurek, W. H. (1984). Maxwell's Demon, Szilard's Engine and Quantum Measurements, in G. T. Moore and M. O. Seully (eds), *Frontiers of Nonequilibrium Statistical Mechanics* (New York: Plenum). Reprinted in Leff and Rex (1990), pp. 249—259.

Zurek, W. H. (1989). Thermodynamic Cost of Computation, Algorithmic Complexity and the Information Metric, *Nature* 341, 119—124.

Zurek, W. (ed.) (1990). *Complexity, Entropy, and the Physics of Information*. Addison-Wesley.