The Concept of Entropy in Statistical Mechanics and Stochastic Music Theory: On the Possibility of a Formal and Philosophical Isomorphism

Ivano Zanzarella
ivano.zanzarella@rub.de

Term Paper

March, 2021

Abstract

Originally appeared in the field of thermodynamics, the concept of entropy, especially in its statistical acceptation, has found applications in many different disciplines, both inside and outside science. In this work we focus on the possibility of drawing an isomorphism between the entropy of Boltzmann's statistical mechanics and that of Xenakis's stochastic music theory. We expose the major technical aspects of the two entropies and then consider affinities and differences between them, both at syntactic and at semantic level, hereto particularly referring to the philosophical problem of the asymmetry of time.

Contents

1 Introduction

The concept of entropy firstly appeared at the beginning of the 19th century in the field of thermodynamics (see [Müller 2007], [Uffink 2007]). The notion was introduced in 1865 by the German physicist Rudolf Clausius and Later developed, in a probabilistic and statistical acceptation, by James Clerk Maxwell, Ludwig Boltzmann and Josiah Willard Gibbs, which made of it one of the core notions of the newborn discipline of statistical mechanics. Since then, methods, concepts and mathematical formalism of this new discipline have proven to be powerful theoretical means also for analyzing and explaining different phenomena than for example thermodynamic ones. Among them, especially the statistical and probabilistic interpretation of the concept of entropy as a measure of uncertainty, randomness, disorder, disorganization, unpredictability – with its relative mathematical formalism – has been recognized as having great theoretical efficacy within a large and heterogeneous number of disciplines, from cosmology, logic, biology, informatics and economics to hermeneutic, philosophy and the arts. So, for instance, Ralph Hartley and Claude Shannon used it as a probabilistic measure of the quantity of information contained in a given number of messages flowing from a source – laying, in so doing, the groundwork of modern information theory. And still, outside science, the Greek-French architect, engineer, music theorist and composer Iannis Xenakis, father of stochastic music, applied it in his theory of sound and of music composition for describing (or prescribing) the development of sounds and musical compositions.

From the philosophical point of view, it appears very interesting that a single concept like that of entropy has been suitable for being applied in so many heterogeneous fields and for so many different phenomena. Still, even more astonishing is the fact that this concept – traditionally a scientific one – has been suitable for being applied in music, a discipline for centuries now no longer regarded as scientific – and this not just in the form a blurry poetical metaphor, but in that of a consistent and rigorous mathematical notion. However, although lots of words have been spent about the concepts of thermodynamic entropy, informational entropy, etc. and about the relationship between them, as a clear lack of literature thereon confirms, no thorough epistemological analysis has been ever made about musical entropy and about the relationship it has with other entropies. Indeed, such an analysis is needed in order to deepen our general understanding of the concept of entropy and of why it works for so many different phenomena, even outside science. Therefore we shall undertake this epistemological analysis in the following, in particular, focusing on the relationship between the entropy of stochastic music theory and that of statistical mechanics.

More specifically, this analysis will be aimed at ascertaining the possibility of drawing an *isomorphism* between the two entropies. Thus, after a summarized exposition, in [Section 2], of the most important technical aspects of the concept of entropy as it is applied in Boltzmann's statistical mechanics and Xenakis' stochastic music theory, we will compare in [Section 3] the two applications both at formal and syntactic and at semantic and philosophical level. In particular, in [Section 3.2] we will consider whether one of the philosophical problems traditionally connected with the concept of entropy in statistical mechanics, namely that of the asymmetry of time, also applies in a similar way for musical entropy and, if so, to what extent. In [Section 4] we will finally derive an explicit conclusion from this comparison, namely that an isomorphism between the two entropies is possible – both at formal and at philosophical level, even if not a total one.

2 Entropy in Science and Music

2.1 A Brief Historical and Contextual Introduction to the Concept of Entropy

As mentioned above, the concept of entropy firstly appeared in the field of thermodynamics. Curiously, it did not emerge in the context of the pure scientific research, but in that of an engineering problem, namely how, and how far it is possible to improve the efficiency of a heat engine [Müller 2007, Ch. 3].

Since the first empirical observations – allegedly during the 17th century – about the relationship between heat and mechanical work, the realization of a perfect heat engine – an engine capable of wholly transforming heat in useful work – had been challenging many technologists and inventors. The first who gave an answer to this problem and found a "limit" for the improvement of a heat engine was the French engineer and physicist Sadi Carnot [Carnot 1824]. He stated – in what is today known as "Carnot theorem" – that a heat engine which acts by transferring heat from a warmer body to a cooler one and converting some of the absorbed heat into mechanical work cannot ever have efficiency (ratio between absorbed heat and produced work) greater than that of a particular ideal engine – the "Carnot engine" – operating on an ideal reversible thermodynamic cycle – the "Carnot cycle". Being able to transfer heat from the cooler body to the warmer one, this ideal machine has in fact maximum efficiency: the whole amount of heat exchanged in the system is equal to that of useful work done by it.

After Carnot's research, the German physicist Rudolf Clausius could demonstrate [Clausius 1854] that in the case of such an ideal heat engine with maximum efficiency, the sum of the ratios of the heat exchanged between the bodies and their respective temperatures is always equal to zero, whereas in the case of any other actual – i.e. less efficient – heat engine, it is (generally) less than zero, for the Carnot engine represents the upper limit of the efficiency of all heat engines. From this statement – which today bears the name of "Clausius Inequality" – important implications can be derived. Firstly, an actual heat engine cannot produce a quantity of work greater than the total amount of heat exchanged (First Law of Thermodynamics). Furthermore, heat transfers between two bodies do not leave the total amount of heat in the system unchanged: part of the heat exchanged in the system get irremediably lost and cannot be converted into useful work (Second Law of Thermodynamics, after the formulation of Kelvin-Plank). This statement has been demonstrated to be (also mathematically) equivalent to Clausius' own formulation of the Second Law of Thermodynamics: if only in an ideal reversible engine the work produced in the system by transferring heat from the warmer body to the cooler one can exactly correspond to the amount of work needed by the system for cyclically allowing a thermal transfer from the cooler body to the warmer one, in a real engine, in which not the whole amount of heat is converted into work, this kind of reverse path would be impossible without additional work from outside the system. The result of this is namely that, unless external work is done on the system, it is impossible to have an actual engine whose *sole result*[1] is the transfer of heat from a body of lower temperature to a body of higher temperature [Clausius 1854, 486].

Thus, only in a Carnot engine, independently from the "direction" of the thermal exchange, the aforementioned ratio between exchanged heat and temperature would remain constant. Now, exactly at this point Clausius introduced the word "entropy" (symbolized with "S") for indicating this quantity which in reversible (and isothermal) thermal processes does not undergoes changes [Clausius 1865, 390]. On the contrary, since, as said, in the heat transfer taking place in irreversible engines part of the heat gets lost, entropy is never conserved in actual thermal processes and instead always bound to increase – that being one of the most important consequences of the Second Law of Thermodynamics.

Clausius' thermodynamic entropy is a state function. This means, it just measures the ratio between exchanged heat and temperature in relation to individual states of a thermal system, states in which the system is in thermodynamic equilibrium. In other words, it does not regard its non-equilibrium states, the states in which it instead undergoes changes. Furthermore, this concept of entropy refers to macroscopic aspects of a thermodynamic system (like for example the temperature), without mentioning its microscopic ones (for example the behavior of the molecules). It is a non-probabilistic concept and offers «no intuitive interpretation as a measure of disorder, disorganization, or randomness (as is often claimed)» [Frigg and Werndl 2011, 117].

Fundamentally different was the approach to thermodynamics which had been developing not from an engineering problem, but from a truly scientific hypothesis: that the macroscopic properties of a thermodynamic system depend on, and are explainable by, the mechanical behavior of its microscopic components[2]. So, for example, changes in the pressure of a gas in a container are made dependent on changes of the momentum of the molecules composing it – a change occurring due the continuous collisions of them with the walls of the container; or still, changes

[1] *Sole result* means here *spontaneously*, i.e. without an external work.

[2] The first scientist to propose this hypothesis was Daniel Bernoulli in [Bernoulli 1738].

in the temperature on the (mean) kinetic energy of the molecules, etc. Already Clausius – though anticipated by August Karl Krönig [Krönig 1856] – tried to interpret thermodynamic processes in such kinetic terms [Clausius 1857], actually laying the groundwork of modern kinetic theory of gases. However, a genuine turning point within this new approach to thermodynamics came with the research of James Clerk Maxwell and Ludwig Boltzmann.

With the development of kinetic (or mechanical) theory of heat, concepts and tools of Newtonian mechanics began to be employed with regard to the microscopic constituents of a thermodynamic system in order to explain its macroscopic behavior. Yet, even if Newtonian mechanics was basically deterministic, the new "laws" of thermodynamics appeared to have rather a probabilistic and statistical character. So was Maxwell's final formulation [Maxwell 1867] of his *distribution function* – later improved and extended by Boltzmann (see [Boltzmann 1868], [Müller 2007, 92-94] and [Uffink 2007, Sec. 3-4]), which gives the *probability* that a certain number of particles (atoms or molecules) of a gas in a container have a certain velocity (where the system is supposed to be in a state of thermodynamic equilibrium at a specific time). This number, in fact, cannot be deterministically set, for the particles, due to incessant collisions, undergo a continuous change of their velocities. The only way in which it is possible to interpret that number is then just as a *mean* or *expectation value*, so that the distribution function rather amounts to a *probability distribution*[3]. Moreover, starting from the presuppositions of the new kinetic approach, a first probabilistic interpretation was given by Boltzmann also to the concept of thermodynamic entropy [Boltzmann 1872], which now began to be related not only, as in Clausius, to the macroscopic properties of a thermodynamic system, but also (in a mathematically equivalent way[4]) to the microscopic objects composing it (see [Section 2.2]).

So, with Maxwell and Boltzmann, who firstly introduced probability in physics, a new discipline was born: that of statistical mechanics – label coined by Josiah Willard Gibbs, who systematized and generalized their results [Gibbs 1902]. As pointed out

[3] On the possible ways of interpreting the Maxwell-Boltzmann distribution see [Frigg and Werndl 2011, 123-124].

[4] See [Frigg and Werndl 2011, 127] and [Section 2.2].

above, many concepts of this new discipline turned out to be good theoretical means for explaining and describing different phenomena in other disciplines. And thus was the concept of entropy after the statistical and probabilistic interpretation by Boltzmann and Gibbs, which has then become a fundamental notion in Shannon's information theory and Xenakis' theory of music.

In the following two sections we shall analyze in detail how in particular Boltzmann and Xenakis intended the concept of entropy in their respective fields and theories. This will later allow us to follow the epistemological aim of this work of ascertaining if an isomorphism can possibly be drawn between these different applications of it.

2.2 Entropy in Boltzmann's Statistical Mechanics

Boltzmann gave essentially two formulations of the concept of entropy: in [Boltzmann 1872] and in [Boltzmann 1877]. In both papers, his aim was to justify the Second Law of Thermodynamics in the terms of the new *paradigm* represented by the kinetic theory of gases. This also meant to attempt to justify Clausius's thermodynamic entropy in kinetic and probabilistic terms. In the following we shall analyze these two formulations in all their technical details[5].

With regard to the 1872-formulation, let us examine an isolated thermodynamic system like a gas of n particles of mass m in a container of volume V. Given that each particle of the system has three degrees of freedom with respect to position and momentum, the system and its evolution over time can be mathematically described by considering a $6n$-dimensional phase space[6], that we will call the *γ-space* (X_γ), in which each specific *microstate* of the system (i.e. all its particles' having a certain positions and momenta at a certain instant of time) will be denoted by a vector

[5] Given the intricate and old-fashioned form in which Boltzmann presented his results in the original papers (acknowledged among others even by [Müller 2007, 95]) we will follow in this analysis – rather systematically than historically aimed – [Frigg and Werndl 2011] and [Uffink 2007], which in turn refer back mostly to [Ehrenfest and Ehrenfest-Afanassjewa 1911] and [Tolman 1938].

[6] This "geometrical" language for describing thermodynamic systems was introduced by [Gibbs 1902] on the idea that changes in the system are "changes in phase" (see [Tolman 1938, 43 and ff.]).

$x_\gamma \in X_\gamma$. X_γ is the Cartesian product of n copies of the 6-dimensions phase space X_μ (μ-*space*), relative to one single particle. The microstate of one single particle will be thus denoted by the vector $x_\mu \in X_\mu$, where $x_\mu = (x, y, z, p_x, p_y, p_z)$, with $\vec{r} = (x, y, z)$ and $\vec{p} = (p_x, p_y, p_z)$, respectively, for position and momentum coordinates. From the kinetic theory of gases, we know that a given macrostate of the system at a certain time t (like the gas' having specific pressure and temperature) is describable by considering its microstate at that same time, that is, by individuating positions and momenta at time t of the particles. To this aim, Boltzmann refers to his improved and extended[7] version of the theoretical and mathematical means for the analysis of the microstates of a thermodynamic system which was already developed in the decade before by Maxwell, namely the previously mentioned *Maxwell distribution*, in modern form:

$$f(x_\mu, t) = \frac{\chi_V(\vec{r})\,(2\pi mkT)^{-\frac{3}{2}}}{||V||} \exp\left(-\frac{\vec{p}^{\,2}}{2mkT}\right), \tag{1}$$

where $\chi_V(\vec{r})$ is the characteristic function of the set V, k the Boltzmann's constant, T the temperature of the gas, $||V||$ the volume of the container and $\vec{p}^{\,2} := p_x^2 + p_y^2 + p_z^2$.

Today known as *Maxwell-Boltzmann distribution*, $f(x_\mu, t)$ gives the (mean) number of particles in the system with a position and momentum value lying within the infinitesimal interval $(x_\mu,\ x_\mu + dx_\mu)$ at time t. The study of the overall dynamic of the system can be carried out considering the evolution of $f(x_\mu, t)$ over time, namely

$$\mathrm{H_B}(f) := \int_{X_\mu} f(x_\mu, t) \log f(x_\mu, t)\, dx_\mu, \tag{2}$$

[7] As mentioned before, the *Maxwell distribution* [Maxwell 1867] describes only the speeds of the particles of a thermodynamic system, whereby the system is supposed to be at time t in thermodynamic equilibrium and the particles not interacting with each other except for very brief collisions without consequences on the overall microstate of the system. Later, Boltzmann extended the *Maxwell distribution* also to the description of the energies of the particles and to states of non-equilibrium of the system. The result was a nonlinear integro-differential equation, the so-called *Boltzmann equation*, which expresses – always in terms of probability distribution – how the number of particles having a certain energy (as well as a certain momentum and position) at time t *varies* within an infinitesimal time interval as consequence of collisions (*Stöße*) between particles (see [Boltzmann 1868], [Boltzmann 1872, 345 (Abh.)] and, for an overview, [Müller 2007, Ch. 4] and [Uffink 2007, Sec 4.2]).

Relying on mechanical arguments, Boltzmann then moves on in his 1872 paper with the discussion of one of the most important results proposed in that work, namely the meaning of his famous H-theorem: due to incessant collisions of the particles with each other, the quantity $\mathrm{H_B}(f)$[8] can only decrease towards a minimum value, after whose reaching it remains constant. This value is reached when $f(x_\mu, t)$ becomes the Maxwell-Boltzmann distribution [Equation 1], i.e. exactly when the system reaches the state of equilibrium, where collisions have no longer effect on the positions and momenta distribution of the particles. He then passes to highlight the close connection existing between $\mathrm{H_B}(f)$ and the thermodynamic entropy in equilibrium states, which takes the modern mathematical form of:

$$S_{\mathrm{B,f}}(f) := -kn\mathrm{H_B}(f), \tag{3}$$

where $S_{\mathrm{B,f}}$ is the *fine-grained* (or *continuous*) *Boltzmann entropy*[9] (therefore "B" and "f" in subscript). $S_{\mathrm{B,f}}$ can only increase during the dynamic evolution over time of the system towards equilibrium, and reaches its maximum – remaining then constant – once the system has finally reached it and its microscopic distribution has thus become the Maxwell-Boltzmann distribution [Equation 1].

In this way, Boltzmann provides a first kinetic proof of why entropy can only *increase* (or at least remain constant), delivering, in so doing, the coveted «mikroskopische Deutung des zweiten Hauptsatzes der Thermodynamik» [10]. Moreover, as we have seen, the kinetic interpretation of the Second Law of Thermodynamics and of

[8] Boltzmann calls this quantity E.

[9] More in particular, the connection between Boltzmann entropy $S_{\mathrm{B,f}}$ and the thermodynamic entropy can be can be better recognized considering the Sackur–Tetrode Formula which describes the thermodynamic entropy of a monatomic ideal gas:

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

where S_{TD} is the thermodynamic entropy, T_0 the temperature of the gas and V_0 its volume. It can be shown that $S_{\mathrm{B,f}} = S_{\mathrm{TD}}$ up to an additive constant. It remains open to question if this result remains also valid for systems different from ideal gases (see [Frigg and Werndl 2011], [Emch and Liu 2013], [Reiss 1965], [Uffink 2007]).

[10] [Boltzmann 1872, 116 (Brush 1976)]. See also [Boltzmann 1872, 313-314 (Abh.)] and [Uffink 2007, 46] on the comparison between Boltzmann's and Clausius' entropy.

the concept of entropy expressed in [Equation 2] and [Equation 3] crucially rests in itself upon a probabilistic and statistical notion, namely the Maxwell-Boltzmann distribution. In his paper, in fact, Boltzmann appears very aware of the fact that a kinetic explanation of the behavior of a thermodynamic system cannot but be a statistical and probabilistic one, for

> « [...] die Moleküle der Körper sind ja so zahlreich und ihre Bewegungen so rasch, daß uns nie etwas anderes als jene *Durchschnittswerte* wahrnehmbar wird. [...] Denn die Moleküle sind gleichsam ebenso viele Individuen, welche die verschiedensten Bewegungszustände haben, und nur dadurch, daß die Anzahl derjenigen, welche *durchschnittlich* einen gewissen Bewegungszustand haben, konstant ist, bleiben die Eigenschaften des Gases unverändert. *Die Bestimmung von Durchschnittswerten ist Aufgabe der Wahrscheinlichkeitsrechnung. Die Probleme der mechanischen Wärmetheorie sind daher Probleme der Wahrscheinlichkeitsrechnung*»[11].

The validity of Boltzmann's 1872 conclusion was however criticized[12] already by his contemporaries, which questioned many aspects and theoretical presuppositions of that research: from the specific role of probability in the argumentation and the exact mechanical characterization of the particles collisions to the generality of his H-theorem – pivotal for making his notion of entropy work – which is actually still today regarded as problematic[13]. For this reason, Boltzmann reformulated his kinetic theory of gases and, with it, the fundamental concept of entropy. He did this in an important paper of 1877 which will become, as we will see in the next sections, a true source of inspiration for several generations of scientists and mathematicians, from Max Plank and Albert Einstein to Claude Shannon and indirectly, on the boarder with the "mathematized" arts, even to Iannis Xenakis.

In his 1877 paper, Boltzmann intended to highlight even more the role of probability theory in thermodynamics, as well as the link between probability and entropy/Second Law[14]. Let us see briefly how in the following.

[11] [Boltzmann 1872, 316 (Abh.)], my italics.
[12] See [Uffink 2007, 974-983] and [Ehrenfest and Ehrenfest-Afanassjewa 1911, 35-36] for an overview.
[13] See [Emch and Liu 2013, 92-105] and [Uffink 2007, 962-974].
[14] This is also the reason why many commentators (see [Uffink 2007, 55], [Klein 1973, 83] and [ter Haar 1955]) consider this paper as clearly marking the transition from kinetic theory to statistical mechanics.

We take again the $6n$-dimensional phase space X_γ for describing a thermodynamic system consisting of a gas in a vessel, as well as the phase space X_μ relative to one single particle of it. Since the system is completely isolated with constant energy and volume V, the state of each particle (its position and momentum at time t) can be represented only by points from a limited region of the phase space X_μ, which we will call the *accessible region of* X_μ, $X_{\mu,a}$.

We consider now a finite number of partitions[15] $\omega = \{\ \omega_i \mid i = 1, \ldots, l\ \}$ of $X_{\mu,a}$ defined in the following way: $\forall i, j \in \{1, \ldots, l\}$ with $i \neq j$, $\omega_i \neq \varnothing$, $\omega_i \cap \omega_j = \varnothing$, and $\bigcup_{i=1}^{l} \omega_i = X_{\mu,a}$. These partitions define cells on the 6-dimensional accessible region $X_{\mu,a}$, which are taken to be rectangular with respect to the position and momentum coordinates and to have fixed volume $\delta\omega$, defined through the Lebesgue measure μ on X_μ, $\mu(\omega_i) = \delta\omega$ for all $i = \{1, \ldots, l\}$. Now, for each $x_{\mu_i} \in X_{\mu,a}$, i.e. for each *microstate* of the system, we define a *distribution of state* as follow: $Z := \{n_1, \ldots, n_l\}$. This distribution indicates the number n_i of particles that, at a certain time t_i, are in a state (i.e. have certain position and momentum) which is represented by points of $X_{\mu,a}$ contained in the cell ω_i. It indicates, simply said, *how many* particles are in which cell ω of the accessible region $X_{\mu,a}$ at a certain instant of time. Obviously is $\sum_{i=1}^{l} n_i = n$ the total number of the particles of the system, and, for different time instants t_i, there are different Z_i distributions, where the ratio Z_i/n can be interpreted as being equivalent to the Maxwell distribution – $f(x_\mu, t_i)$ [Equation 1]. Moreover, different distributions Z_i individuate univocally different *macrostates* of the system. Now, the association between the microstates $x_{\mu_i} \in X_{\mu,a}$ and a macrostate of the system depending on the particles distribution Z_i is defined in these terms: $Z_i := \{\ x_{\mu_i} \in X_{\mu,a} \mid Z(x_{\mu_i}) = Z_i\ \}$. What it is important to notice here is that $Z(x_{\mu_i})$ is however not a bijective function. This means that possibly many microstates with the *same distribution* can correspond at time t_i to a same macrostate. In fact, the microstates can have different *internal arrangements*: even if the distribution (i.e. the number n_i of particles being in the cell ω_i) remains unchanged, the state of the

[15] Boltzmann considered this kind of discretization as a mere mathematical tool and not as a physical hypothesis, as Max Plank will soon do (see [Müller 2007, 99]). This idea will be very important also for Xenakis, as we shall see in [Section 2.3].

particles within the cells can undergo rearrangements, for example by permutation. The number W of the possible arrangements compatible with a given distribution of state Z is determined in general by using ordinary combinatorics:

$$W_Z = \frac{n!}{\prod_{i=1}^{l} n_i!}. \tag{4}$$

Now, we can finally define a new concept of entropy, namely as:

$$S_{\mathrm{B,c}} := k \log W_Z, \tag{5}$$

which is called *combinatorial* (or *coarse-grained*) *Boltzmann entropy*[16] (therefore "c" in subscript), which is thus directly proportional to the number W_Z of possible arrangements of a particles state distribution corresponding to a given macrostate of the system. This means that the more arrangements are compatible with a given particle distribution, the greater the entropy value will be.

Now, in order to understand the inherent probabilistic character of combinatorial entropy is moreover decisive to say that *all* arrangements of a given distribution are regarded as *equiprobable*. Thus, entropy is, more precisely, the measure of the number of possible arrangements coherent, *with equal probability*, to a given distribution. Now, Boltzmann states [Boltzmann 1877, 164-167 (Abh.)] that the system spontaneously tends to evolve *always* toward states with *more* arrangements coherent with a same distribution, until it reaches the state with *most* arrangements, namely the equilibrium, where entropy is therefore at maximum. If, within the evolution of the system toward equilibrium, the number of equally probable arrangements coherent with a distribution increase, this also means that the randomness, the unpredictability, the disorganization and the disorder of the system grow, reaching the maximum at equilibrium. Entropy is in fact a measure of such disorder and randomness in the sense that the greater it becomes, the less it is possible to know about the actual arrangement of the particles of the system on the basis of their distribution (where, we repeat, all arrangements

[16] It can be shown that $S_{\mathrm{B,c}}$ is formally (nearly) equivalent to $S_{\mathrm{B,f}}$ (see [Uffink 2007, 56-57], [Frigg and Werndl 2011, 126-127] and [Müller 2007, 99-101]).

compatible with a distribution are equiprobable). At the same time, however, entropy can be identified, Boltzmann says, with the probability of the total state of the system: in fact, equilibrium is also the *most probable* state – since the *largest* number of arrangements correspond to its distribution – to the effect that the evolution of open thermodynamic systems always and spontaneously takes place from less probable to more probable states and never in the opposite way.

Recapitulating, a thermodynamic system generally evolves towards an equilibrium state, i.e. the most disordered and, at the same time, the the most probable one, whereby entropy is the measure of these disorder and probability. All this eventually amounted for Boltzmann to providing a new kinetic and probabilistic justification of the Second Law of Thermodynamics and of why entropy has constantly to increase[17].

Boltzmann's entropy formula [Equation 5] can also take another form inasmuch as it is applied not to arrangements with respect to distributions, but, in a formally equivalent way[18], to distributions (i.e. microstates) with respect to macrostates, where by distribution is meant how the particles-points of the phase space of the system are distributed in the cells. In this new form, it can represent the probability to find, given a macrostate of the system at a certain time, the corresponding distribution of the particles-points within the cells of the system phase space. Thus, assuming that in most systems $n_i \gg 1$ and using Stirling's approximation formula for the factorials,

[17] Yet, if [Müller 2007, in pt. 101] presents this conclusion as a more or less direct and explicit consequence of Boltzmann's research, [Frigg and Werndl 2011, 125], drawing on [Ehrenfest and Ehrenfest-Afanassjewa 1911] and other commentators, regards it rather as a postulation in need of further assumptions and justifications (such as ergodicity or tipicality).

[18] See [Frigg and Werndl 2011, 125-127], [Uffink 2007, 56-57], [Müller 2007, 99-103], [Ehrenfest and Ehrenfest-Afanassjewa 1911, Ch. 2], [Tolman 1938, Ch. 4].

$\log n! \approx n \log n - n$, we can derive from [Equation 5]:

$$
\begin{aligned}
S_{\mathrm{B,c}} &= k \log W_Z \\
&= k \log \left(\frac{n!}{\prod_{i=1}^{l} n_i!} \right) \\
&= k \log(n!) - k \log(n_1!) - \ldots - k \log(n_\ell!) \\
&\approx (nk \log n - n) - (n_1 k \log n_1 - n_1) - \ldots - \\
&\qquad - (n_\ell k \log n_\ell - n_\ell) \\
&= -k \sum_{i=1}^{\ell} n_i \log n_i,
\end{aligned}
$$

whereby, letting $p_i = n_i/n$ be the probability of finding a randomly chosen particle in the cell ω_i, we therefore have:

$$S_{\mathrm{B,c}} = -nk \sum_{i=1}^{\ell} p_i \log p_i, \tag{6}$$

which links the macrostates of the system with the possible microstates of it (i.e. possible distributions of the particles in the cells), expressing, for a given macrostate, the probability to find at the microlevel its correspondent distribution[19]. At equilibrium, the probability to find a randomly chosen particle in a given cell is equal for all cells, since the particles-points are distributed uniformly within the phase space. As in the case of arrangements and distributions, however, to a same macrostate several different microstates can correspond. In fact, due to the uniform distribution of the particles in the phase space and hence the equiprobability of all possible particles microstates (i.e. the equiprobability to find a particle in each of the cells of the phase space), at equilibrium a very large number of microstates are compatible with this

[19] The two formulations of entropy are equal up to the additive constant $nk \log(\delta\omega)$ relative to the volume of the cells of the phase space, which we omitted for simplicity and irrelevance to the next argumentations. For the same reasons we avoided to derive this entropy formula in terms of proportionality between the number of microstates and the volume of the phase space. For a detailed derivation of the formula in current terms, see [Frigg and Werndl 2011, 125-127] or [Uffink 2007, 56-57].

macrostate of system, indeed the largest, and this is also the reason why the entropy of the system at equilibrium is at maximum. For non-equilibrium states, where the particles do not occupy quite homogeneously all regions of the phase space, but rather have all a specific and uniform value for position and momentum (i.e. they occupy a specific region or cell of the phase space), the probability to find the microstate of the system corresponding to a given macrostate of it is ascertainable again through [Equation 6], i.e. through finding out how many particles occupy which cells. And, the less *all* cells have the *same probability* to contain particles with position and momentum values compatible with that given macrostate, the lower the entropy and the more ordered the system will be.

[Equation 6] never appears in the original scientific production by Boltzmann in the form in which we presented it. However, as it can be observed in some of his later writings [Boltzmann 1898], the idea behind it should have been already clear to him. The modern formula was proposed later by Gibbs [Gibbs 1902], who generalized it to different kinds of systems and system states, even more highlighting its statistical and probabilistic character. As a measure of the disorder and unpredictability of a system, the probabilistic notion of entropy expressed in [Equation 6] became a powerful tool also in order to account for other kinds of phenomena.

This was for example the case of Shannon and his information theory (see [Shannon 1948] and [Shannon and Weaver 1949]). Shannon's aim was to provide a mathematical law for quantifying (and predicting) the quantity of information contained in a certain number of messages flowing from a source. Relaying on some previous ideas by Hartley, [Hartley 1928] he proposed to make this quantity (mathematically) dependent on the probability of each message. The more probable a message is, the less information we gain from its reception. From messages flowing from a source with exactly the same probability, we gain no information, since all messages have equal probability to contain information. As in the case of thermal equilibrium, where all microstates are equiprobable, when all messages have the same probability to contain the same amount of information the entropy (*Shannon information entropy*) is at maximum. On the contrary, the less probable a message is, the more uncertainty there is about its information content, the more information we gain from its reception removing the uncertainty. Information and uncertainty are for Shannon, in fact, two sides of the same coin. Shannon noticed that the only mathematical equation able to represent this relation between probability and information, as well as to satisfy other important formal requirements[20], was

$$H = -K \sum_{i=1}^{n} p_i \log_2 p_i, \tag{7}$$

where H is the entropy, named by Shannon after Boltzmann's H-theorem[21], K a positive constant, p_i the probability of the messages, and the binary logarithm referring to the fact that the entropy is expressed in bits, unit of measurement of information [Sedgewick and Wayne 2011, 185]. [Equation 7], as we can see, is formally very similar to Boltzmann's entropy formula in the form of [Equation 6]. And also the probabilistic concept of entropy it expresses, giving probabilities not to particles microstates but to messages, is by and large comparable to that expressed by the latter.

A lot of words have been spent until now about the relationship between Boltzmann's and Shannon's entropy. In particular, in epistemology, questions have been investigated such as if they both can be considered isomorphic concepts, if they interpret and apply probability in the same way, and still, if physical system as the thermodynamic ones are accountable in terms of information[22]. It would be surely very interesting to go over these epistemological and philosophical analyses and then also consider if a relationship evem exists between Shannon's information theory and the stochastic theory of music by Xenakis' we will present in the next section. Furthermore, such considerations would allow us to spot possible differences in the

[20] Continuity, additivity, monotonicity, branching and bit normalization (see [Shannon 1948, 10-11], [Frigg and Werndl 2011, 118]).

[21] In Boltzmann thermodynamics, it can be in fact shown that if [Equation 1] does not vary very much with respect to each cell of the phase space, [Equation 6] and [Equation 2] - i.e. the H-theorem - can be considered approximately equal up to the additive constant relative to the volume of the cells [Footnote 19] (see [Frigg and Werndl 2011, 127]).

[22] Remarkable in this respect are for example the pioneering papers by Edwin Thompson Jaynes (in particular and [Jaynes 1957a] and [Jaynes 1957b]), who argued that Boltzmannian statistical mechanics can be seen just a particular application of Shannon's information theory

ways in which the concept of entropy is extended both *inside* science (e.g. from thermodynamics to information theory) and *outside* science (e.g. from thermodynamics to music theory). Some remarks will be done about these subjects in [Section 3]. Yet, since such analysis would require more space than that we presently have, they will not pretend to be exhaustive. Rather, we shall concentrate in this work more on the relationship between Boltzmann's statistical mechanics and Xenakis' stochastic music theory.

2.3 Entropy in Xenakis' Stochastic Music Theory

Iannis Xenakis has been one of the most important avant-guard composer of the 20th century. Known for having integrated mathematics into music theory and music composition – especially stochastics and probability theory, statistics and set theory – he was also one of the firsts to employ computers and algorithms in music composition. Trained architect and engineer, he often "translated" his architectural into musical works and vice versa on the basis of the mathematical models and structures underlying them[23]. The synthesis of his lifelong research through music, mathematics and philosophy is expressed in the book *Formalized music* [Xenakis 1992], partly appeared in French language during the Sixties and containing other papers published in different specialized journals of music and mathematics. In the following we will have an overview of the main principles underlying his music theory and observe what role which kind of concept of entropy plays in it.

The history of the relationship between mathematics and music traces back at least to classical antiquity. Still in early modern times, scientists used to deal with music-theoretical problems, and music theoreticians and composers borrowed from mathematics principles and concepts for their works. Some of these principles even belonged to areas of mathematics such as the combinatorics or the developing probability theory and were applied by modern composers in the attempt to evade the strict determinism of tonal music and to randomize the process of music composition

[23] See for example Xenakis' *Philips Pavilion* at Brussels World's Fair of 1958, whose design was «pin-pointed» [Xenakis 1992, 10] from *Metastasis*, a musical work premiered in 1955.

[Gardner 1970].

Considering the role of chance in sound phenomena and bringing chance in itself into music composition has been also Xenakis' own major aim. Indeed, one of his first observation in [Xenakis 1992, Ch. 1] is that most of the theories of sound and of music – tonal, serial or dodecaphonic music – rely on deterministic assumptions and models. This has as consequence that, on the one hand, the evolution in time of complex and disordered sound events (indeed the most common in our experience, like for example the sound made by a car or the song of the cicadas) cannot be comprehensively described in any way, and that, on the other, it is impossible to produce musical compositions whose structure and evolution in time are truly ruled by randomness and chance, i.e. truly outside the deterministic constraints those theories impose. Even serial and dodecaphonic music, that have purported to have broken the tradition of the determinism of tonal music by introduction of atonality, falls again, after Xenakis, into a new form of determinism, namely that of the series. Thus, a new theory of sound, of music and music composition is needed, which must be able to account, in terms of description or "prescription", for the randomized and stochastic evolution of complex sonic events and musical compositions. This theory should contain pure sound or tonal music just as particular cases. And the way for providing this new theory is referring to mathematics, and in particular to probability theory and logic. In fact, Xenakis considers the application of probabilistic and statistical methods to sound phenomena and music – i.e. nothing but sound phenomena which can be created by the composer – a natural consequence of their introduction into the sciences (in particular, statistical mechanics and quantum theory) as powerful tools for explaining the world, which apparently behaves in some respects not deterministically. As sound is part of this world, it is reasonable to think that a probabilistic, *stochastic theory of music* can offer some useful insights into its nature and the way it behaves, as well as some effective hints to composers for making sound evolve in their compositions in a truly stochastic and indeterministic way [Xenakis 1992, 4].

Furthermore, as every deterministic or indeterministic theory, the stochastic theory of music should be subjected, for Xenakis, to the laws of logic, in particular general algebra, which operate on individual objects or set of objects with the aid of some primitive operations (union, intersection, negation), properties and relations (equivalence, implication, and quantifications) [Xenakis 1992, 4]. Indeed, for Xenakis «music [...] may be defined as an organization of these elementary operations and relations between sonic entities or between functions of sonic entities» [Xenakis 1992, 4], where these sonic entities in fact constitute the "empirical objects" of music.

It is thus worth starting our introduction to stochastic theory of music from Xenakis' characterization of the primitive material of music, namely the sonic entity, the sound[24]. Xenakis basically shares the modern physical account of sound as complex sound wave as described for example by [Helmholtz 1862]. Thus a complex sound, say a C_3 played on a piano, is composed by a large number of simpler sinusoidal waves (principal tone, overtones, summational tones, differential tones, etc.) which behave according to Fourier Analysis in producing the final sound. In particular, the quality of the sound is determined by the quantity and the amplitude of these simpler sound waves. Like Boltzmann and Gibbs in thermodynamics, Xenakis chooses to construct a geometrical representation of "macroscopic" sonic events. To this aim – almost recalling the particle-wave dualism of quantum physics – he considers the simple sound waves constituting these sonic events as *sound grains*, elementary sonic particles or sonic quanta, represented by points in a three-dimensional space with coordinates frequency, intensity and duration. Complex sounds are thus assemblages, *clouds* or *ensemble* of simple sound grains evolving over time. The macroscopic condition of every possible sound event becomes explainable, as in statistical mechanics, by referring to the microscopic condition of its granular constituents and its evolution over time. In other words, it suffices to consider the position of the sonic grains in the grains space relative to the complex sound in terms of their having specific frequency and intensity coordinates as a function of time.

Like the phase space describing thermodynamic systems, however, the grains space is not infinite, for sounds are perceivable by human ears up to certain limits

[24] By "sound" or "sonic event" Xenakis means not only the "musical sound" distinguished from the noise, but every possible auditory phenomenon perceivable by the human ear. In the following we use the word "sound" in this acceptation.

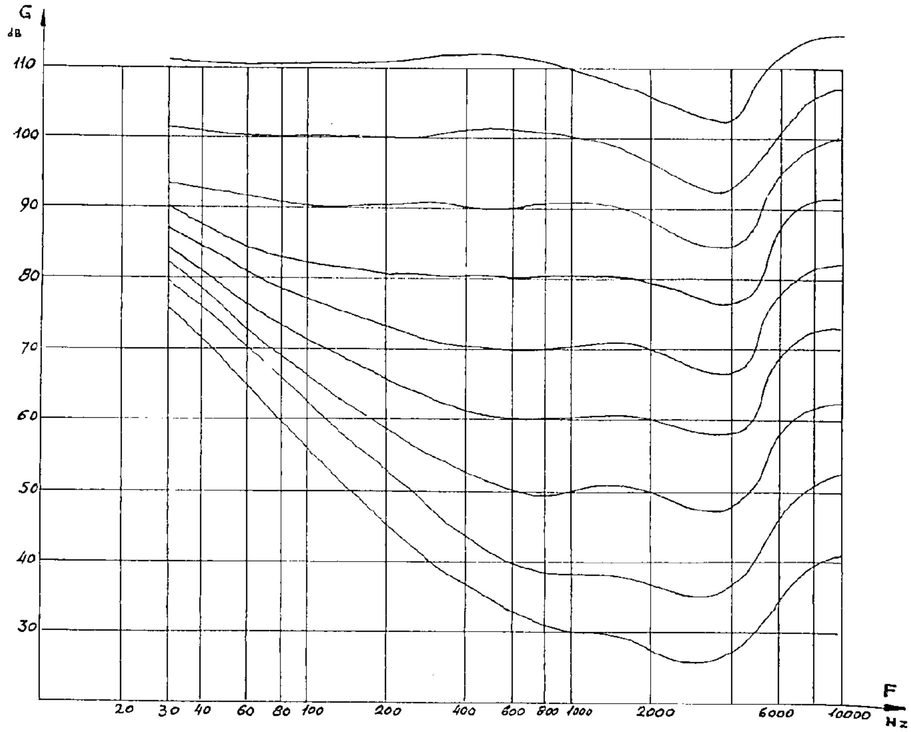

Figure 1

of duration, frequency and intensity. These limits have been firstly discovered in 1933 by Harvey Fletcher and Wilden A. Munson [Fletcher and Munson 1933], who proposed the famous equal-loudness contour diagram representing namely the *audible area* [Figure 1], which gives for example the thresholds for the minimum perceptible duration of a sound as a function of its frequency F and its intensity G, or of its perceivable intensity in decibels compatible with its minimum frequency and duration. Thus we have to consider only this audible subregion of the grains space, i.e. only that containing the possible grains configurations microscopically corresponding to macroscopically audible sounds. For simplicity, we can operate a transformation on the curved space defined by the Fletcher-Munson diagram in order to gain, without alterations, a more regular space of rectangular form for representing the audible subregion of the grains space [Figure 2]. If we now assume that the (limited) duration of a given macroscopic sonic event can be divided into a large, but finite, number

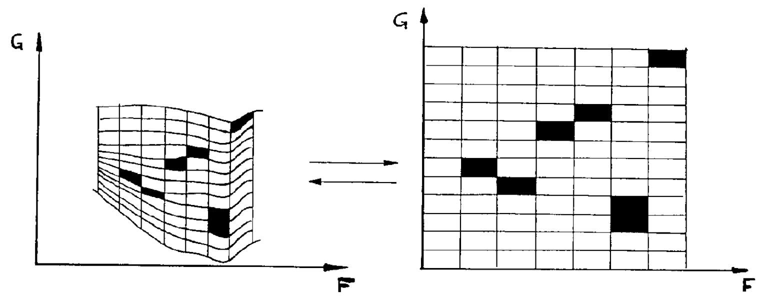

Figure 2

of slices of time ΔT_i, with $i \in [0, n]$, small at will and invariably equal in size, we can individuate for each T_i, T_{i+1}, where $T_{i+1} - T_i = \Delta T_i$, a certain microsound configuration, i.e. a certain distribution of the grains along the two dimensions of frequency F and intensity G coherent with the macrostate of the sound event in that interval of time. By flattening for simplicity the grains in the duration slice ΔT on the two-dimensional frequency-intensity space, we find that the entire macroscopic evolution of the sound event over time can be described by a finite series of n *screens* S_i [Figure 3], with $i = \{0, \ldots, n\}$ and $\bigcup_{i=0}^{n} S_i = FGT$, that succeed one another in time with different distributions and configurations of grains *in*[25] them. As in statistical mechanics [Tolman 1938, 43-44], it is not important – if not completely impossible – considering directly the *single* microconstituents (single grains) of the sound and their respective behavior. In fact, we always refer to *clouds* of grains and to their evolution over time. Then, what is important for microscopically explaining the macroscopic condition of a sound at time T_i, is knowing the total number of the grains at time T_i – what Xenakis calls *density* – and their distribution over the i-th screen (i.e. their having specific frequency and intensity values at that instant of time). In other words, we need to know how many sound grains of the cloud occupy which region of the screen at time T_i, which means, how many sound grains

[25] We use the preposition "in" instead of "on" for speaking about grains with regard to screens even though, being screens two-dimensional objects, this may be grammatically incorrect. However, this will help us hereafter to keep well in mind that, as to the screens, we only artificially flattened their three-dimensionality, and that the grains *in* them exist in the their "time thickness" ΔT.

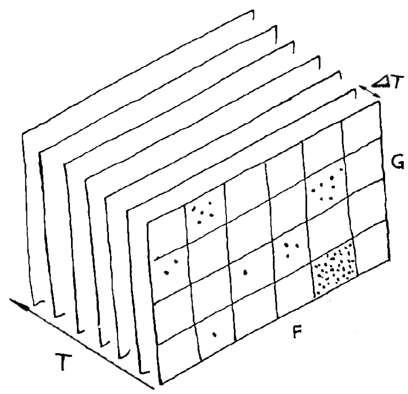

Figure 3

have at this time instant specific frequency and intensity coordinates compatible with a given macrostate of the sound at that same time. The careful reader will already notice a certain resemblance between this problem and that answered by the Maxwell-Boltzmann formula (see [Equation 1]). Xenakis actually considers the possibility to apply an adjusted version of this equation for calculating the distribution of the sonic grains within the screens. Yet, he notices that this would require firstly a reformulation of the geometrical environment of scalar type so far adopted – the grains space FGT should be indeed construct as a vector space – and secondly the introduction of the concept of speed [Xenakis 1992, 55-56]. Albeit mathematically possible, maybe due to practical reasons Xenakis does not further undertake this task and continues with his arguments based on the scalar grains space. Nevertheless, he shows that the mathematical[26] and probabilistic ideas behind the Maxwell-Boltzmann distribution are very clear to him. In fact, the distribution of the grains in the screens at certain instants of time is anyway deterministically detectable [Xenakis 1992, 52], as the grains of natural sounds[27] "fluctuate" around an equilibrium position of *mean*

[26] It can be noticed from [Equation 8], [Equation 9] and [Equation 10], in fact, that the mathematical idea of the Boltzmann-Maxwell distribution (specifying the *mean number* of microscopic objects lying in a certain region of a geometrical space individuated by an infinitesimal interval) is also applied by Xenakis for calculating the distribution of the grains within the screens with respect to coordinates D, T, F and G one at a time.

[27] An exception to this would be represented by electronic sound, whose grains are characterized by fixity.

frequency and intensity. Also the density – i.e. the number of the grains in each screen – varies from one screen to another around a mean value. This eventually amount to say that the number of grains occupying at T_i a given region of S_i can be only *probabilistically* interpreted, just as a *mean* or *expectation* value – actually like in the case of particles distributions in thermodynamic systems.

How, then, does Xenakis account, from the mathematical point of view, for the distribution of the sound grains in the screens? First of all, he evidently embraces the theoretical presuppositions of the combinatorial approach of the "second" Boltzmann and of Gibbs: the distribution of the grains within the screens $FG\Delta T$ is only *statistically* determinable. Moreover, since we are not interested in the single grains, but in knowing which region of the screen a *cloud* of grains occupies, in order to find out that, we can take screens as divided into cells $\zeta = \{ \zeta_j \mid j = 1, \ldots, l \}$ with $\bigcup_{j=1}^{l} \zeta_j = S_i$, defined as in [Section 2.2][28] and of equal volume $\Delta F \Delta G \Delta T$ (remember: screens have a thickness ΔT) [Figure 4], and consider the density ΔD of the grains for each cell, i.e. the (mean[29]) number of grains in it.

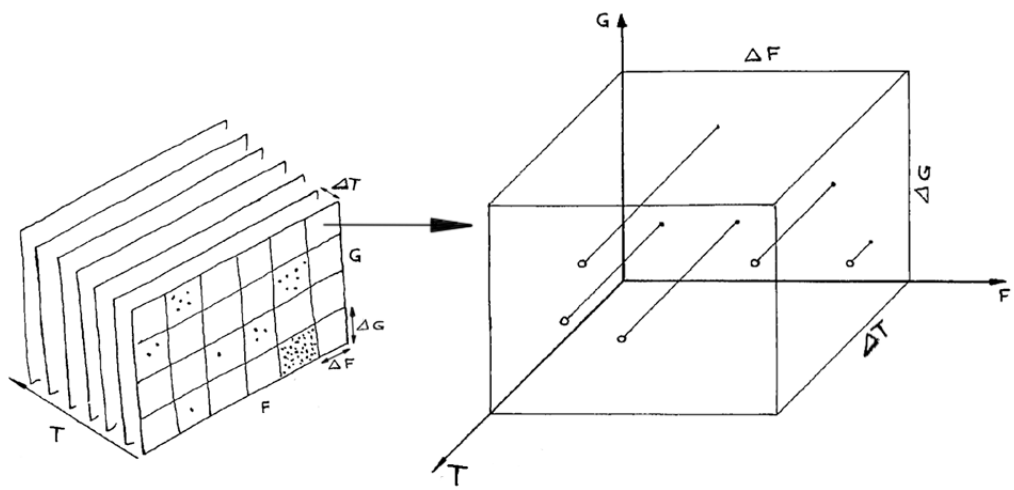

Figure 4

[28] The only difference being the possibility that $\zeta_j = \varnothing$ for some $j \in \{1, \ldots, l\}$.
[29] "Mean", because, as we said, in natural sound the distribution of the grains in the screens and therefore also in their relative cells is not deterministically, but only probabilistically definable.

It is possible calculate the density ΔD of the cells by the *mean density* D of the screen (ratio of the number of grains in it and its total volume) using Poisson's formula:

$$P_\mu = \frac{\mu_0^\mu}{\mu!} e^{-\mu_0} \tag{8}$$

which expresses, given the fixed screen mean density μ_0, the probability P_μ for a specific density μ (i.e. a given number of grains) to occur in a certain cell of the screen.

At this point, known the (mean) density of the grains per cell, we can calculate – always in a statistical way – the distribution of the grains along the other dimensions of the screen: the time T, the frequency F and the intensity G. Instead of giving a method for calculating this distribution directly with reference to *all* the three dimensions of the grains space *at one*, which would be mathematically very complex, Xenakis simplifies the calculation proposing *stochastic laws* – derived from the theory of continuous probability – for the distribution with reference to *one dimension at a time* [Xenakis 1992, 12-16; 52-54; 323-327].

As to the dimension of the time, the law that statistically gives the distribution of the grains along ΔT, i.e. the portion of the time axis T related to a screen, is

$$P_x = ce^{-cx} dx, \tag{9}$$

where c is the linear density of the points on the axis, x the length of every possible segment taken on it and P the probability that the i-th segment will have a length x_i between x and $x + dx$. Hereby we consider the *points* on ΔT to be the *projections* t_j over this portion of the time axis T of the grains in the "thick" screen $FG\Delta T$. These points thus individuate segments *within* ΔT of length $t_{j+1} - t_j = x_j$, representing the duration of each grain of the screen. Then, [Equation 9] expresses the *probability* to find on ΔT, taken on it *any* (even infinitesimal) duration δt, segments $x_j = t_j, t_{j+1}$ with $x_j = \delta t$[30].

[30] This same argument could be generalized from the duration of a single screen to that of the whole sound event: the statistical distribution of the grains along T can be also described through [Equation 9].

With respect to frequency F and intensity G (as well as to each further variable one wishes to consider as characteristic of sound), the formula for calculating the distribution of the grains along ΔF and ΔG, the other two dimensions of the cells [Figure 4], is

$$\theta(\gamma)d\gamma = \frac{2}{a}\left(1 - \frac{\gamma}{a}\right)d\gamma, \tag{10}$$

where a is the length of the segments ΔF or ΔG on their respective axes. This equation thus gives the probability that a certain number of segments within ΔF and ΔG, defined, as in the case of the time, by points deriving from the projection of the grains in $FG\Delta T$ on the frequency F or intensity G axes, and representing intervals respectively of pitch or intensity, will be found to have a length between γ and $\gamma + d\gamma$, for *any* γ with $0 \leq \gamma \leq a$.

So, these are the different mathematical tools for statistically calculating the distribution of the sonic grains in a cell ζ_j of the screen S_i at time T_i with regard to its four dimensions[31] $\Delta F \Delta G \Delta T \Delta D$, provided an overall screen mean density D_i in advance. They allow to individuate, for every chosen δt and γ, sets[32] of durations, frequencies and intensities, $\delta t = \{\delta t_1, \ldots, \delta t_n\}$ and $\gamma = \{\gamma_1, \ldots, \gamma_n\}$, containing any possible value of these variables within the range of $\Delta F \Delta G \Delta T \Delta D$ compatible, with a certain probability, with δt or γ. In this sense, by them a probability distribution is assigned to the elements of these sets, formally in the following way[33]: $P = (p_1, \ldots, p_n) := (p(\delta t_1), \ldots, p(\delta t_n))$, with $p_i \geq 0$ and $p_1 + \cdots + p_n = 1$. This means, every duration segment δt_j on ΔT has a certain probability p_i to correspond to a fixed duration segment δt. So, for instance, for $\delta t_1 = \delta t$, $p_1 = p(\delta t_1) = 1$, assuming that this probability is 0 for all the other elements.

There is obviously a correspondence between these sets, so that, for example,

[31] We regard density as fourth dimension of the cell.

[32] These sets are finite for the following reasons: the number of possible perceivable duration intervals is limited (see equal-loudness contour diagram); the duration ΔT of the screens is limited; the overall duration T of complex sonic events is limited; the audible area FG is limited. They are furthermore totally ordered as subsets of totally ordered sets (respectively T, F and G).

[33] We consider in the following only the case of δt, which is however completely similar to that of γ.

grains with a given frequency and intensity have a certain duration, etc. If in a natural sound these correspondences are clearly independent of us, in music they can be set, even randomly, by the composer, who derives those values and correspondences by specifying the initial parameters, i.e. giving, for example, concrete values to D (mean screen density), j (number of cells the screens are divided into), δt, γ etc. and applying the above formulas to them [Xenakis 1992, 54; Ch. 1; Ch. 3]. In any case, by calculating with the same statistical procedures the grains distribution for *all cells* along their respective dimensions, and then considering the arithmetic means of the obtained values, we should be able to generalize our knowledge about the distribution of the grains at a certain "thick" instant of time to the whole screen. In this way, we would have a statistical, microscopic, *description* of the macroscopic state of a complex sonic event at that same instant of time, may this sound be natural or musical[34]. But, as said, the composer could use these mathematical procedures also for actively building screens. In so doing, he would be able to *prescribe*, through a manipulation of the microscopic components of the sound, the macroscopically perceivable properties of a (musical) sound event at specific instants of time of it. Examining in depth how stochastic music composition exactly works, how the values of the variables are concretely calculated and how the correspondences between them are set would however lead us too far away from the aim of the present work. Nevertheless, what has been generally said about that is already sufficient in order to introduce Xenakis' concept of entropy.

Entropy is defined, with respect to each of the four dimensions $\Delta F \Delta G \Delta T \Delta D$ of a screen cell, as

$$H = -K \sum_{i=1}^{n} p_i \log p_i. \tag{11}$$

As for the time, the minimum entropy is given in the case of simultaneity or isochrony (if they are emitted at regular intervals of time) of the grains. In such a case, in fact, we would have that $\delta t_1 = \ldots = \delta t_n = \delta t$, that is, only one element in the set δt,

[34] These mathematical methods are useful, in fact, not only for describing sonic events in general or composing new music, but, as Xenakis points out, also for analyzing and describing in a more accurate way *musical* works of the past [Xenakis 1992, 4].

namely δt. Its *variety* (the number of *distinct* elements in a set or group) would be 1. Correspondingly, we would have $p(\delta t_1) = \ldots = p(\delta t_n)$, thus $p_1 = \ldots = p_n = 1$ (since the probability of the duration is $p_i = 1/n$ and the set has $n = 1$ elements), hence

$$H = -K \sum_{i=1}^{n} p_i \log p_i = -K \log 1 = 0.$$

On the contrary, if the variety of the set is different from 1 and, *theoretically*[35], $n \to \infty$, then also $H \to \infty$, inasmuch as $p(\delta t_1) = \ldots = p(\delta t_n) = 1/n$, i.e. inasmuch as *all* durations of the grains, emitted according to [Equation 9] become, in a fairy long succession of screens (see [Footnote 31]), *equiprobable*.

As to the other dimensions, the argument unfolds quite similarly. The overall entropy of the cell is given by the arithmetic mean entropies calculated for each dimension, and that of the whole screen by the arithmetic mean of the entropies of the single cells composing it. For mean entropy $H = 0$, the cell or the screen will contain just one single sonic grain corresponding to a *pure sound*, with specific and constant frequency and intensity and emitted at regular intervals of time (the variety of the sets δt and γ will be 1) [Figure 5A]. In the case of maximum mean entropy (theoretically ∞, see [Footnote 35]), cells or screens will contain grains homogeneously distributed over all their regions, i.e. will have each one a different value for frequency, intensity and duration. The state of maximum entropy will thus correspond to a *white sound* [Figure 5C] – that of malfunctioning (analog) radios and televisions.

[35] This is however impossible in practice: in fact, firstly, there is a limit, as already pointed out, for the perceptibility of duration intervals, and secondly, a sonic event or a musical composition always has a *finite* number of grains and screens (see [Footnote 32]). So, all in all, n will always depend, in a real sonic event, on the overall duration of this latter, and thus will also the probability distribution over the set δt of the duration intervals.

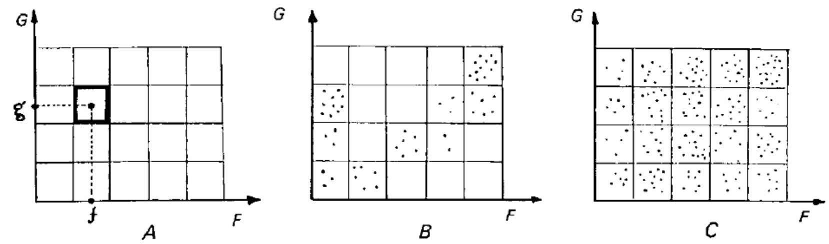

Figure 5

And now something very fascinating: «Between these two limits the grains may be distributed in an infinite number of ways with mean entropies between 0 and the maximum and able to produce both the Marseillaise and a raw, dodecaphonic series» [Xenakis 1992, 64]. Between these two limits (e.g. [Figure 5B]), with the means of statistics, every possible perceivable natural sonic event is describable or prescribable, every musical work of the past is explainable and every new musical work can be composed.

So far we have discussed just about single screens, as descriptions at the microlevel of the state of a complex sonic event at a certain instant of time. Nothing has been said about the evolution in time of complex sonic events, that is, about the succession of the screens and how this happens and works. Xenakis distinguishes between *causes* and *modes* of this succession. The cause of a transition from a screen to another is to be traced back to the *physical* structure of the sound – which has been analyzed for example by [Helmholtz 1862] – or the *logical* structure of a musical composition – which Xenakis thoroughly analyzes in [Xenakis 1992, Ch. 6-8] using linear algebra, in particular interpreting these transitions as operations on vectors from a vector space with dimensions F, G and T. Again, an examination of these structures is something that can and need not be undertaken here. Rather than in the causes, what we are instead interested in modes in which the transitions between screens *generally* take place.

First of all, speaking about transitions between screens means speaking about transitions between different configurations of sonic grains along frequency, intensity and duration. Thus, we can describe or prescribe a transition from these three different points of view. A *transition* is formally a succession of two *terms*, whereby the second one is called the *transform*. The terms can be denoted for instance by the letters of the alphabet $(a, b, c, \ldots)$. A *transformation* is finally a collection of transitions and can take the following form:

$$\downarrow \begin{array}{llll} a & b & c & \cdots \\ a & c & c & \cdots \end{array}$$

where $(a, b, c, \ldots)$ are for example pitches, intensities, durations, etc. A transformation can be also represented by a matrix. Taking the previous example:

$$\begin{array}{c|ccc} \downarrow & a & b & c \\ \hline a & 1 & 0 & 0 \\ b & 0 & 0 & 0 \\ c & 0 & 1 & 1 \end{array}$$

So, transformations of screens are generally the result of transformations of pitches, intensity, durations, etc, correspondingly represented by matrices. In music composition, in order to prescribe a certain overall evolution of the sound, the composer will construct a protocol for such transformations and engender the relative matrices, which may be independent from each other or possibly coupled.

Screen transformations represented by matrices wherein only values 1 and 0 appear are actually just a particular case of transformations, namely determined ones, with univocal and closed transitions. The general case is in fact that of *stochastic* transformations, whose transitions are many-valued and depend on probabilities. In the corresponding matrices, the values 1 and 0 are namely replaced by relative frequencies which specify the probability of a certain transition. Are these probabilities constant over a long period of time and dependent only on the first term of each transition, then we have a particular case of stochastic transformation, called *Markov chain*:

$$a\,b\,a\,b\,b\,b\,a\,b\,a\,a\,b\,a\,b\,a\,b\,a\,b\,b\,b\,b\,a\,b\,a\,a\,b\,a\,b\,b\,a\,a\,b\,a\,b\,b\,a\,b\,a\,a\,a\,b\,a\,b\,b\,a\,a\,b\,b\,a\,b\,b\,a$$

The real frequencies of each transition occurring in the transformation are:

$$\begin{array}{lrl} a \to b & 17 & \text{times} \\ a \to a & 6 & \text{times} \\ \hline & 23 & \text{times} \end{array} \qquad \begin{array}{lrl} b \to a & 17 & \text{times} \\ b \to b & 10 & \text{times} \\ \hline & 27 & \text{times} \end{array}$$

from which we can construct a matrix of transition (i) as in the example above, and also a matrix of transition probabilities (ii) by converting the real into relative frequencies:

$$\text{(i)} \quad \begin{array}{c|cc} \downarrow & a & b \\ \hline a & 6 & 17 \\ b & 17 & 10 \end{array} \qquad \text{(ii)} \quad \begin{array}{c|cc} \downarrow & a & b \\ \hline a & 0.26 & 0.63 \\ b & \underline{0.74} & \underline{0.37} \\ & 1.00 & 1.00 \end{array}$$

Now, the mode in which a screen transformation – and thus the evolution over time of a sonic event – take place, always implies certain changes in the overall value of the entropy of the screens of this event. If a transformation takes place in a determined mode (with regard to *all* the dimensions of the screens), that is, if all transitions in it occur according to a matrix of transition probabilities of the form

$$\begin{array}{c|cc} \downarrow & a & b \\ \hline a & 0 & 1 \\ b & 1 & 0 \end{array}$$

the overall entropy of the screens within the transformation remains unchanged. Since all transitions occur deterministically, the transformation brings about neither an increase nor a decrease of disorder along the temporal evolution of the sonic event[36]. In the opposite case, if all transitions occur with the same probability, thus according

[36] A case in which all transitions (must) always occur deterministically is that, for example, of tonal music.

to a matrix of transition probabilities of the form

$\downarrow$	a	b
a	0.5	0.5
b	0.5	0.5

the transformation would be completely indeterminate and totally ruled by chance. The transitions would be all equally unpredictable and thus the disorder and the entropy of the transformation at maximum.

As said, transformations of screens are the result of transformations of pitches, intensities, durations, etc. Then, the overall entropy of the former (the extent to which they bring about changes in the entropy of the screens over time) corresponds to nothing but a mean entropy of the latter. Thus, a musical composition might develop, for example, deterministically as to time, towards disorder as to frequency, and so on, whereby the overall entropy of the composition, and of the general screen transformation involved in it, would be a mean value of the entropies of the transformations occurring along each of its three dimensions.

Hence, between the two limits set by determined and indeterminate transformations, we have stochastic transformation, in which transitions take place each with a different probability (see [Matrix (ii)] above). From the point of view of the entropy, these transformations are classifiable under three groups:

1. "Neutral" transformations, i.e. those in which the overall value of entropy does not change;
2. Transformation towards order, i.e. those bringing about decreases in the overall value of entropy;
3. Transformation towards disorder, i.e. those bringing about increases in the overall value of entropy.

Group (1) is that of determined and indeterminate transformations, which keep the value of entropy unchanged, respectively, at its minimum (zero) and its maximum (theoretically ∞). Conversely, to group (2) or (3) belong stochastic transformations inasmuch as the probabilities of their transitions are set as, or develop towards, those of the transitions in determined or indeterminate transformations. Differently – in most cases – from natural sound, in music compositions many kinds of transformations can be applied to screens in the course of the same composition. Thus, in this respect, the variation of entropy over time becomes for the composer an additional aesthetic criterion for shaping music [Xenakis 1992, 75-78].

3 Epistemological Analysis of a Possible Isomorphism

After having examined from the technical point of view the applications that Boltzmann's statistical mechanics and Xenakis' music theory provide of the concept of entropy, we are ready to consider more nearly the problem posed in the introduction, namely, whether and to what extent it is possible to consider them *isomorphic*. An explicit conclusion on this will be reached in [Section 4], after that the affinities and differences between the two applications will be highlighted both on the syntactic and formal [Section 3.1], and on the semantic and philosophical [Section 3.2] level.

3.1 Formal Affinities and Differences

Already in the previous section some remarks have been done about the formal affinities between statistical mechanics and stochastic music, not only as to the concept of entropy in itself, but also with regard to the formal mechanisms used in those theories for describing different kinds of phenomena, on the basis of which this concept is established. Now we shall more closely discuss whether and to what extent, from the formal point of view, an "isomorphism" between entropy in statistical mechanics and entropy in music can be drawn. Obviously, this discussion cannot be exhaustive and extensive here. In fact, the analysis of the formal affinities between different theories always requires great amounts of effort and time. However, since in the case of statistical mechanics and music this analysis has not yet been undertaken – as a lack of literature thereon also confirms – we firmly believe that the few hints we will give in the following on this subject may be a good theoretical starting point for further research.

Our discussion will especially focus on these four aspects: the geometric description of the systems, the concept of distribution/density, the mathematical formalism and the use of probability.

As we saw in [Section 2.2], one of the great intuitions of Boltzmann, later further and more rigorously articulated by Gibbs, was that of using a geometrical language for the microdynamical representation of the states and the evolution over time of macroscopic systems such as the gases. Thermodynamic systems, for example, can be represented by a so-called *phase space* whose points correspond to their microstates with their particles having a specific coordinate for position and momentum at a specific time. Is a system isolated, with constant number of particles, volume and energy, then only a subspace of this phase space can represent it, namely its *accessible region*, containing all possible microstates compatible with that system in terms of position and momentum of its particles. For studying the distribution of the particles of the system as points within this bounded phase space – or of the microstates, insofar as we refer to a 6-dimensions or a $6n$-dimensions phase space – it is finally useful to discretize it in a finite number of equal cells.

From this point of view, the influence of statistical mechanics on Xenakis' music theory is very clear. Firstly, a macroscopic sonic event is explained by referring to its microscopic constituents, the sonic grains. Then, this sonic event is also geometrically representable, in particular by a three-dimensional space whose points correspond to sonic grains with specific duration, frequency and intensity coordinates. This space is also bounded, since the range of the audible frequencies, intensities and durations is limited, as the Fletcher-Munson diagram shows. The microscopical explanation of a sound event can only take place in a probabilistic and statistical way. To this aim, it is useful to discretize this space in equal cells. Moreover, in order to account for the evolution of the sound over time, it is also useful to consider screens, i.e. sets of cells in succession along one axis of the space, that of the time.

Although some of the original ideas by Boltzmann are clearly applied in Xenakis' argumentation – for example the discretization of the geometrical space – no direct reference in [Xenakis 1992] is made to the papers of him we have examined before. Indeed, his name appears just among those of the first scientists who generally spoke of entropy. From our point of view, this historical unawareness by Xenakis is certainly not worrisome in itself. [Xenakis 1992] pretends in fact not to be an history of statistical mechanics. Neither, very probably, Xenakis had access, in the initial phase of his research (mid-fifties), to a vast literature about the history and the foundations of statistical mechanics, that has indeed grown only in recent times. Yet, this is important for us in order to appreciate the extent to which Xenakis' stochastic theory of music differs from the statistical mechanics which has developed from Boltzmann's approach.

As to the geometrical representation, an important difference is that regarding the typology of spaces employed. The interpretation by Xenakis of the sound in its evolution over time as a discrete succession of screens (which could be regarded as its microstates) is mostly due to the fact that the geometric space used for representing the sound is a scalar space, containing points and not vectors as the thermodynamic phase space. This is in fact a vector space with double the degrees of freedom of the system it represent, in which therefore continue variations for example of particles' position and momentum can be described as can be described functions of time. As mentioned, a formulation of stochastic theory of music based on vectors would be generally possible, and this with no changes in content. Nevertheless, it would be not very convenient, in Xenakis' opinion, especially for what concerns the practical applications of the theory to music composition. Furthermore, a difference also lies between the space partitions of the phase space in statistical mechanics and those of the sonic grains space. The former are conceived, at least in the Boltzmannian framework, as mere theoretical means for mathematically explaining the microstate of the system. The particles, that is, points in the phase space can occupy equally all cells, and indeed actually occupy them homogeneously when the system is in equilibrium. Conversely, it is not a formal requirement for the cells of the grains space to be occupied by grains. This means, there may be cells of the space which are never occupied along the entire evolution of the sound. And this actually happens since not all regions of the audible area, as the Munson-Fletcher diagram shows, are equally perceivable by human ears (we are more receptive for frequencies and intensities at the center of the diagram and less for those on the borders).

From the respective characterization of the geometric space representative of the systems, other affinities and differences between statistical mechanics and stochastic music theory derive which especially regard the concept of distribution/density, and consequently the mathematical formalism and the statistical "laws" used for explaining macroscopic systems in microscopic terms. In both theories, the concept of distribution or density refers to the way in which the particles-points occupy the regions of the geometric space considered. This distribution is only statistically calculable, since the particles are generally unstable, for different reasons, in their position in that space. These being the general points in common between the two theories as to the concept of distribution, important dissimilarities also arise in how this concept is really intended in both of them, especially from the mathematical point of view. The most important one is that, whereas in statistical mechanics the distribution is a function of time and as such is calculated for example through [Equation 1] or [Equation 2], it is not in stochastic music theory where it is conversely calculated referring to a discrete and instantaneous state of the macroscopic sonic event, namely a screen or a cell of a screen. As a result, time does not appear as a variable in [Equations 8-10] either. As mentioned, this partly depends on a focused theoretical choice that Xenakis takes due to practical constraints, namely that of disentangling the dimensions of the grains space within the calculation of the distribution of the sonic grains. So, the calculation of the distribution of the grains with respect to time, for example, is reduced to calculating durations, i.e. the (in itself time-independent) linear density of the punctual projections of the grains on a limited segment (mono-dimensional by definition) of the axis of time. Radically different is the approach of statistical mechanics, where no such disentanglement of the phase space dimensions is present. Again, the reason for this is the kind of geometric space by which the systems are respectively represented. The vector phase space of statistical mechanics allows it to calculate the distribution of the particles as a function of the time, where in stochastic music theory this is not possible unless the grains space is constructed too as a vector space, to which thus adequately adapted mathematical tools of the form for example of [Equation 1] would be applicable, as Xenakis says.

Continuing on the mathematical formalism, a glaring affinity is that of the formulas used in both theories for calculating entropy, namely that between [Equation 6] and [Equation 11]. Beyond that, there are however differences in content between the two. [Equation 6] expresses the probability to find, given a macrostate of a system at a certain instant of time, the corresponding distribution of the point-particles in the cells of the phase space, where it is assumed that there are many possible distributions, i.e. microstates, compatible with that macrostate. [Equation 11], first of all, does not rest on the same kind of combinatorial arguments, for it is assumed, as far as emerges form Xenakis' reasoning, that there is a bijective correspondence between the microstates, say, of frequencies, intensities, etc. and the macroscopically perceivable frequencies, intensities, etc[37]. The equation expresses how much we can infer about the distribution of the grains along each of the different dimensions of the grains space, knowing the variety (i.e. the number) of possible duration, frequency, etc. intervals and how likely they are. Thus, again, entropy regards single dimensions and is related to time insofar as one calculates it for each of the screens that succeed each other in the evolution of the sonic event over time. Furthermore, the series in [Equation 11] is not taken to be an always convergent series. For infinite varieties, in fact, the sonic event can theoretically evolve with entropy values approaching infinite. On the contrary, the series of [Equation 6] must always converge towards a maximum finite value, namely that of the entropy at the system's equilibrium state. This eventually amounts to say that not only, depending on what kind of transformations the evolution of a sonic event takes place after (see [Section 2.3]), we can have a limitlessly and monotonically increasing entropy, but also that the overall entropy of the sonic event can (or can be set to) decrease or remain constant. In other words, a natural evolution of sonic events towards a state of equilibrium does not exists,

[37] This is clear considering for example the case of duration: it is impossible to have on the axis of time two or more segments representing the duration of a single sonic grain, because in this case the segments would be indistinguishable and the durations they represent necessarily simultaneous. In other words, we have the possibility to consider that grain as to time always by one single duration, i.e. one single segment on the axis of time.

and this evolution has indeed no privileged direction in terms of entropy and time. This is actually an evident difference between the concepts of entropy in statistical mechanics and stochastic music theory, as well as a problem worth of philosophic considerations that will be done in the next [Section 3.2].

[Equation 6] also express the extent to which one can infer about the distribution of the particles-points in the phase space knowing about the macrostate of the system, whereby, if the system is in equilibrium and the entropy at maximum, it becomes impossible to make any such inference, since the largest possible number of distributions are with the same probability compatible with the macrostate. Indirectly, by interpreting the systems of statistical mechanics through a Shannonian informational approach, entropy becomes in this sense also the measure of the disorder and uncertainty connected with a system and of the information we gain removing this uncertainty. Actually, [Equation 11] seems to conform more to this informational interpretation of the concept of entropy in statistical mechanics – which is even clearer if one looks at the direct references Xenakis does in [Xenakis 1992] to the writings of Shannon and to the ideas by Boltzmann as mediated by the American information theorist. Indeed, from the formal point of view, the role played by equiprobability and its connection with the notions of uncertainty and entropy are the same both in statistical mechanics and in stochastic music theory, if we interpret them both informationally. Yet, a difference may arise if one considers, in stochastic music theory, entropy and uncertainty not with respect to single cells or screens (whereby the notions are related to the distribution of the grains), but, generalizing, with respect to entire transformations of screens. As we know, transformations are collections of transitions occurring between screens with a certain probability. The more equally likely all transition probabilities are, the more uncertainty there is and the higher the value of the entropy of the entire transformation is. Uncertainty is thus related here *not* to possible distributions of microscopic objects compatible, *at a specific time*, with the macrostate of a certain system (as clearly in statistical mechanics and partly in stochastic music theory), but directly to possible evolutions *over time* of macroscopic systems, in this case, of sonic events. The uncertainty, and so the entropy are at maximum when, considering a screen transformation, all possible screens[38] could succeed in time, with equal probability, a given screen – when, in other words, every possible evolution of the sonic event is equally probable. On the contrary, there is no uncertainty when all transitions between screens in a transformation is completely determined. This means that it is obviously easier to predict the development of a piece of music, given initial chords or notes, if this is a classical one (with only determined transitions possible) than if it were a stochastic one (with transitions ruled by probabilities)[39]. In the same sense, it would be easier to predict the information content of messages emitted by a source on the basis of deterministic principles (hypothetically, for instance: every third time you toss the coin, it must be tail), than if the source emitted informative messages with certain probabilities. The possibility of interpreting in an informational way unpredictability and entropy in stochastic music theory, put it actually nearer to information theory than to statistical mechanics[40]. As mentioned above, it would be surely interesting, from the epistemological point of view, considering the extent to which also these two theories are syntactically and semantically comparable.

Here we shall however continue with the last point of our inquiry into the formal affinities and differences between statistical mechanics and stochastic music theory, namely the interpretation of the notion of probability involved in them. There is actually plenty of literature in history and philosophy of physics about the notion of probability in Boltzmann and, in general, in statistical mechanics. Although it is impossible to review the whole debate about this topic in the present work[41], some general remarks about Boltzmann's use of probability will allow us to understand whether probability is namely used in the same way also by Xenakis.

Let us start from the long quotation by [Boltzmann 1872, 116 (Brush 1976)] in [Section 2.2]. In particular Boltzmann says (my italics):

[38] That is, all possible configurations of the grains in the grains space compatible with the macroscopic sound at a certain instant of time.

[39] Interesting experimental research has been done in this respect for confirming this claim (see e.g. [Manzara et al. 1992]).

[40] Obviously, unless an interpretation of statistical mechanics in terms of information theory is taken into account, like for example that of Jaynes (see [Footnote 23]).

[41] See for an overview [Uffink 2007], [Frigg 2008], [Frigg 2009], [Frigg 2010a], [Frigg and Werndl 2011].

> «[...] die Moleküle sind gleichsam ebenso viele Individuen, welche die verschiedensten Bewegungszustände haben, und nur dadurch, daß die Anzahl derjenigen, welche *durchschnittlich* einen gewissen Bewegungszustand haben, konstant ist, bleiben die Eigenschaften des Gases unverändert».

As it can easily be noticed from the second clause of the quotation, for Boltzmann the necessary statistical and probabilistic character of the dynamical explanation of the macroscopic behavior, in this case, of thermodynamic systems is not only due to our epistemic and empirical limitations in accounting for the exact number, positions and speeds of the microscopic constituents of the system. Indeed, the macroscopic properties of the system *in themselves* are considered to be the result of observer-independent probabilistic factors, namely of the fact that the number of the microscopic elements of the system with the same *mean* speed remain constant. As we know, this is the meaning of the Maxwell-Boltzmann distribution [Equation 1], which Boltzmann extends and develops in his H-theorem [Equation 2] and in its definitions of entropy. Thus, following [Uffink 2007] and [Frigg and Werndl 2011], we also agree on the fact that the notion of probability in Boltzmann should be interpreted to a large extent ontically, with the "laws of probability" being actually objective empirical laws, and more in particular in a frequentist way: drawing repeatedly microscopic elements from a system, for example molecules from a gas, Boltzmann's statistical laws would give us the relative frequencies with which we can find particles randomly taken from the system with specific position or momentum values. As we have seen, this is especially clear in Boltzmann's combinatorial arguments.

The question is now: is probability in Xenakis also interpretable in frequentist terms? As already pointed out at the top of [Section 2.3], Xenakis' musical research is aimed at introducing in music, against the strict deterministic schemes of serialism and dodecaphony, an extended idea of causality, as it has been finding a place in science (especially quantum and statistical mechanics) already since the first decades of the twentieth century, namely the causality involving probabilistic laws and containing determinism just as a particular case [Xenakis 1992, 8-10]. In this sense, Xenakis clearly accepts the idea that natural phenomena could be described by empirical laws of probabilistic form, that is, the idea that nature behaves in some respects indeterministically in itself. And this obviously applies also to sonic events: natural sonic events such as «the collision of hail or rain with hard surfaces or the song of cicadas in a summer field»[42], or still human-dependent sonic event such as those produced by «a political crowd of dozens or hundreds of thousands of people» or by «a Geiger counter in the proximity of a radioactive source» share *all* a stochastic character and can *all* be described using the same statistical laws, those that finally the composer can just take and apply in music in order to run away from the determinism of the tradition. In this sense, Xenakis generally interprets probability ontically. For him, probabilities are surely «part of the "furniture of the world"» [Frigg and Werndl 2011, 119]. Furthermore, from this ontic perspective it becomes also possible to attribute to Xenakis a frequentist interpretation of probability: he for example defends the idea that stochastic phenomena (as those just mentioned) evolve, for the the law of large numbers, asymptotically «towards a stable state», namely a στόχος, from which the adjective "stochastic". This law, [Xenakis 1992, 16] points out, is also that to which densities, durations, frequencies, etc. can be subjected. For example, as we have seen, taking a sound event with theoretically infinite duration i.e. infinite number of grains, [Equation 9], the statistical law for calculating the distribution of the sonic grains along the axis of time, will show an equalization of the probabilities of the durations. This means that, independently of how one chooses the reference duration segment δt, the relative frequency with which one will find duration segments equal to it will be, in this case, always the same.

Yet, if we take the possible informational interpretation of stochastic music theory, and thus of the probabilistic concepts involved in it in terms of uncertainty or unpredictability, we may be persuaded to consider Xenakis' use of probability understandable in some respects rather within the context of the epistemic theories of probability (objectivist or subjectivist). Especially in music, where the sound, as succession in time of screens, is artificially constructed by the composer, it seems that the probabilities with which the screens succeed each other and the sound evolves – those appearing in the matrices of [Section 2.3] – are not ontically relative to the object (sound in itself), but epistemically to the subject, as (rational or subjective)

[42] All quotations until the end of the section are from [Xenakis 1992, 8-10] unless otherwise indicated.

credences conditioned by the musical system considered. The uncertainty about the next incoming term of a succession of screens can be epistemically taken as corresponding to the degree of belief of a subject in actually proposing a specific screen as possible next term. This degree would be at maximum if for example the succession takes place within tonal harmony, which is a totally determined music system. In fact, it is a certainty for the subject[43] that a screen representing for instance a leading tone will be always followed by one representing the tonic. This certainty, however, cannot clearly be regarded as ontic, because it is actually inexistent outside the specific music system of tonal harmony (and obviously in the sound in itself either)[44].

Thus, the use of the notion of probability in Boltzmann and Xenakis is quite identical, unless one gives an informational interpretation to stochastic music theory, which leads to the necessity of acknowledging probability there under epistemic categories. Indeed, it would be a goal of the aforementioned comparative analysis of information theory and stochastic music theory that of ascertaining if the informational interpretation of this latter and a Jaynesian, informational interpretation of Boltzmann's statistical mechanics would be also comparable with each other as to the possibility to understand probability in both theories (only) in epistemic terms.

3.2 Philosophical Affinities and Differences: The Problem of Asymmetry of Time

In the last section we discussed the formal affinities and differences between the applications that statistical mechanics and stochastic music theory provide of the concept of entropy. In this discussion, a philosophically very interesting point has emerged: sonic phenomena do not tend towards an "equilibrium state" and, consequently, "musical entropy" does not tend to a finite maximum value, nor has it necessarily to increase, since there exist simplifying (or neutral) transformations

[43] Assuming that he or she has some acquaintance with the system.

[44] This indeed suggests that hereto an *objectivist* epistemic approach would be preferable to a subjectivist one, since, in the most cases, a music system influences the "musical believes" of all (rational) subjects in exactly the same manner.

(see [Section 2.3]) implying a decrease (or no changes) of the entropy value associated with the sonic screens as the sound develops over time. In other words, stochastic music theory does not establish a specific and constrained direction for the evolution of a sonic event or – even more – of a musical composition, say, a direction from disequilibrium to equilibrium, from order to disorder, or from lower to higher entropy.

This is eventually what most radically characterizes musical entropy as a concept *in meaning* very different from the entropy of statistical mechanics. In fact, as we have seen in [Section 2.2], statistical mechanics and thermodynamics generally regard as *empirically possible* for their target systems only *one* kind of evolution, namely that directed at states of always greater equilibrium and higher entropy, whereby evolutions taking place in other directions are considered as impossible, or rather *not typical* or (very) *improbable*.

That the macroscopic systems of statistical mechanics and thermodynamics always evolve in a specific direction is something we also experience in our everyday life. Looking at two photos, the one showing a red cup of milk and a black cup of coffee, and the other the same red cup full of cappuccino and the black one empty, we are able without any doubt to put both photos in the correct order, namely, the first before the second. No other ordering would be possible: in fact, separating coffee from milk, once they have been mixed, is something we know to be utterly impossible. The same would apply for a "collection of photos" ideally representing two gases with different temperatures, thus with their respective molecules having two different speeds. The photo showing the gases and their molecules still separated from each other would be ordered by anyone *before* that taken after the vessels containing each of the gases have been connected. Here all molecules of both gases would appear mixed after a certain period of time and it would become impossible, through natural, non-demoniac[45] processes, to separate again the warmer from the cooler gas, i.e. the

[45] The reference here is to the so-called *Maxwell Demon*, a thought experiment conceived by Maxwell (firstly appeared in 1867 within a private correspondence and then presented in [Maxwell 1871]). It involves a supernatural being, namely a "demon", able to violate the Second Law of Thermodynamics by dividing again two previously mixed gases at different temperatures, i.e. by separating individually the faster from the slower molecules of a system in thermodynamic equilibrium. Many attempts have been done in order to construct an actual mechanism serving as such a demon, yet it has been proven that the complex formed by the thermodynamic system and

faster from the slower molecules. Hence, again no other ordering of the photos would be possible.

Now, if we were to assign to all these photos, as we ordered them, points from an hypothetical time axis, we would then immediately see that the constrained direction in which thermodynamic systems evolve towards states of greater equilibrium and higher entropy also has a temporal meaning: it takes places namely *asymmetrically* with respect to the time. In other words, this constrained evolution is also bounded to a constrained direction of the time, in the sense that, as entropy cannot decrease, time cannot be reversed. For example, given time instants $t_1 < t_2 < t_3$, we have that a thermodynamic system in a state of non-equilibrium and low entropy at time t_2 has necessarily to evolve to one closer to equilibrium with higher entropy at t_3, where it is *highly improbable* that it evolved from an equilibrium state with high entropy at time t_1. In this way, the constrained direction in which both entropy and time increase eventually individuates a concept which is today know after Sir Arthur S. Eddington as the "*arrow of time*" [Eddington 1928], which, for the laws of thermodynamics and statistical mechanics, can only point in one direction, generally to the future.

As we have seen in [Section 2.3], Boltzmann explained the unidirectionality of the evolution of thermodynamic systems by referring to microdynamical and probabilistic arguments, and, in particular, by treating the states of greater equilibrium and higher entropy as the *most probable* ones among all possible states. Some believe that he did it consistently and comprehensively[46], but other, already among contemporaries[47], have seen instead a real difficulty (if not, as Loschmidt did, a real paradox) in basing coherently time-asymmetric descriptions, like those of macroscopic thermodynamical systems, on dynamic and mechanical laws that are conversely *symmetric* with respect to the time, i.e. reversible and quasi-periodical [Callender 1997, 224-225]. In this sense, deriving the former from the latter would mean to accept firstly that, for every entropy-increasing evolution of the system, there should also exist an equal and

---

the mechanism does not violate as a whole the Second Law of Thermodynamics (see [Maroney 2009]).

[46] See e.g. [Lebowitz et al. 1993], [Lebowitz 1994].

[47] See e.g. [Loschmidt 1876], [Poincaré 1889], [Poincaré 1893], [Zermelo 1896]; see also [Uffink 2007, 64-73].

opposite entropy-decreasing one, and secondly that each system evolving towards states with higher entropy will eventually return, given an inconceivably long interval of time, to some past state with lower entropy.

Such a behavior is, as we know, never observed and, in order to come up with the theoretical difficulties of Boltzmann's approach, also other alternative explanations of the time-asymmetric evolution of thermodynamic systems have been proposed from within statistical mechanics. Some of these appeal for example to the notion of *typicality* (see e.g. [Goldstein 2001]) and argue that this evolution and the direction in which it takes place are somehow *typical* of the systems [Frigg and Werndl 2011, 126], [Frigg 2009]. These approaches are not unproblematic either [Frigg 2010b], and in general that of the asymmetry of time is a problem, in thermodynamics [Callender 2016] and physics in general, still far from having a definitive answer.

In fact, what still remains mostly troubling about the "problem of the arrow of time" in statistical mechanics is exactly the question of the compatibility between the symmetrical laws of dynamics and the asymmetrical laws of thermodynamics. Whereas the former theoretically allow entropy increases in both time directions, the latter – as everyday experience also confirms – prohibit it. In order to solve this consistency problem, and have a complete and coherent physical explanation of why systems evolve only towards equilibrium and higher entropy, we essentially have two options: change the "rules of the games", i.e. explain thermodynamic systems no longer by the reversible dynamical laws we currently use[48], or simply *assume* that entropy always increases [Callender 1997, 227], as theoretically satisfying as this latter option might be.

Conversely, some claim that we do not need to solve the problem of the arrow of time because this is not really a problem, but only the result of an erroneous understanding of the foundations of statistical mechanics. [Callender 1997, 228-230], for example, argues that we can regard statistical mechanics and its predictions about

[48] This means, in this case we should use non-dynamical laws or alternative dynamical laws, which are however irreversible. To this respect, lots of theories have been proposed – within both classical and quantum mechanics – which try to justify the impossibility of entropy-decreasing processes always in *dynamical* terms (see e.g. [Sklar 1993], [Albert, D. 1996], [Callender 1996] and [Callender 1997]).

the *theoretical possibility* of past states of higher entropy as holding as generalized in all models of our fundamental theory *only* under an *ad hoc* condition that, in agreement with observation, indeed excludes this possibility, i.e. only by the *assumption* that past states are always states of lower entropy. Since however, from the point of view of classical mechanics, we are actually allowed not to maintain this condition and instead to believe that someday entropy may really start to decrease (as for example cosmologist T. Gold predicted), statistical mechanics should be regarded not as true in *all* physically allowed models, and therefore just as a *special science*, whose laws are *ceteris paribus* laws, and whose models are just elements of a *proper subset* of the set of models of classical mechanics.

That the problem of time is the result of a mistaken view in the foundations of statistical mechanics is also the opinion of [Ben-Naim 2017a] and [Ben-Naim 2017b], who argues that entropy is (even formally) not associated with the "arrow of time". In fact, if we return, within a Jaynesian theoretical framework, to an interpretation of the entropy of statistical mechanics based on *Shannon's measure of information* (see [Section 2.1] and [Section 3.1]), so that it turns to indicate the *uncertainty* about the distribution of the microstates (or particles) of a systems in its corresponding phase space, we would have a formulation of entropy independent of the time. Shannon's measure of information (also known as *Shannon entropy*) is in fact defined on *any* probability distribution and as such is a very general concept. By applying it to the probability distribution of positions and momenta of the particles of a given system, Ben-Naim shows that the entropy of statistical mechanics (in all formulations of [Equation 2], [Equation 5] and [Equation 6]) can be considered as proportional to the *maximum* of Shannon's measure of information, which is in itself not a function of time. Thus, the constrained direction in which the system of statistical mechanics evolve would be no longer associated with a constrained direction of the time. In other words, as it is possible in information theory to remove the uncertainty by gaining information, and thus to *reduce* information entropy, by interpreting the entropy of statistical mechanics in informational terms, we could consistently accept, at least from the theoretical point of view, that also an entropy decrease or a *negentropy*[49] in

[49] Introduced originally by [Schrödinger 1944], the term *negentropy* – contraction of *negative entropy*

thermodynamic systems is possible.

Now, Ben-Naim's thesis is especially important for us in order to gain isight into the problem of the arrow of time connected with the concept of entropy in stochastic music theory. As we said above, the evolution of sonic phenomena does not follow a constrained direction. There is no evolution towards states of higher entropy and disorder. On the contrary, sound can evolve towards states of equal or also lower entropy.

Consider, we are given with a book of screens corresponding to the development over time of a certain sonic event (including a musical composition of any stile), where each screen has its mean entropy value. However, the pages of this book – i.e. the screens – display no numbering and therefore we have to order them. The question now is: does a "correct" ordering exist, similarly to the case of the photos of the cups or the gases? Given that the entropy or the degree of disorder of the screens are not an ordering criterion, the answer can be affirmative just if the book is a *descriptive* one – just if, in other words, we know what sonic event the book as a whole refers to and what kind of transition matrices are related to the screen transformations commonly occurring in this particular sonic event. If we know, for example, that the book refers to the noise made by a squeaking door, we would be able – up to a certain approximation – to arrange the screens in a succession corresponding to the evolution over time of this kind of sonic event. The same would hold in the case the book refers to a musical composition and one knows this latter: the screens could be arranged as to reproduce the development of the composition. Even if less accurately, also knowing just the style which the composition belongs to would lead to the same goal: knowing that, for instance, classical music allows deterministically only certain screen transitions and that stochastic music allows transitions according to certain probabilities, already amounts to have working criteria for providing a *correct* ordering of the screens of the book. Moreover, this ordering would be also constrained and univocal: arranging the screens differently would yield a description of different sonic

– refers to processes in which a decrease of the entropy is possible. Mathematically, it is nothing but entropy with negative sign and as such indicates the degree of order, predictability or information (see also [Wu et al. 2020]).

events or different musical compositions than the ones actually considered. And if we were to associate points from the axis of time to the screens which we "correctly" ordered, we would notice that the constrained evolution of the sonic event, i.e. the constrained succession of the screens, also imply a time asymmetry, i.e. a constrained direction of the time. In fact, reversing the time, and thus the succession of the screens would yield again different sonic events or musical composition than the ones indeed taken into account.

Yet, as we know, Xenakis' sonic screens have not only a *descriptive*, but also a *prescriptive* purpose (see [Section 2.3]). Thus, it is possible to *create* new sonic events or musical composition by arranging screens in a certain way. For example, one can construct a new musical composition by reversing the descriptive succession of screens relative to an existing composition[50]. In this sense, the reversibility of the time would be made possible in music[51], and consequently also entropy decreases (negentropy). Suppose in fact that the entropy value associated with each of the screens of the book relative to the original composition have be set as such to increase monotonically along the succession: prescribing a new musical composition on the basis of the reversal of the original composition would mean now to compose a book of screens in which these succeed each other with a monotonically decreasing entropy value, equal to that of the original succession, but of opposite sign. Entropy and time would be reversed: «much like a god, a composer may create the reversibility of the phenomena [...] and, apparently, invert Eddington's "arrow of time"» [Xenakis 1992, 255]. Still, one could construct a new sonic event by rearranging and permutating freely the screens of a descriptive succession relative to a certain initial sonic event[52], producing, for example, a new succession in which screens succeed each other with

[50] Classical and dodecaphonic counterpoint, for example, take abundantly advantage of this method of musical composition: certain melodic lines or harmonic structures are composed, whose *retrograde* (or *cancrizans*) is then taken and employed inside or outside the same composition.

[51] Whether the reversal of a musical composition is definable as aesthetically evaluable music or as still belonging to the musical style of the original composition is a question of music aesthetics that as such is *epistemologically irrelevant*. Aesthetics cannot exclude, in fact, the mere theoretic possibility of such reversibility in music.

[52] If the sonic event is a musical composition, the same of [Footnote 51] holds as to the result of this free rearrangement.

non-monotonic, irregular changes in the associated entropy values.

Both in the case of the reversal and in that of the rearrangement, no affirmative answer could be provided to the question of whether something like a "correct" or empirically constrained screen ordering exists: in fact, every possible permutation, rearrangement or reversal of the screens of a descriptive succession yield a new sonic event, and, in general, every possible succession of screens – also not derived from a descriptive succession and instead freely chosen within the range of audibility – yields a physically possible sonic event. There are not sonic events, i.e. screen successions, which are theoretically prohibited, or (within the audible area) empirically impossible or improbable, and this because there is not a general constrained direction of the evolution of a sonic event, like that necessarily towards states of higher entropy of the systems of statistical mechanics. Since screen transformations are not necessarily entropy-increasing, entropy cannot serve as a criterion for finding the "right" screen succession. And if there is not constrained direction of the evolution of sonic events, then there is not in music even a constrained direction of time, an *arrow of time*.

That musical entropy is not associated with the arrow of time appears ever more reasonable if one consider the possible informational interpretation of stochastic music theory we already referred to in [Section 3.1]. As we have seen, musical entropy is in some respect nearer to the informational notion of entropy than to the entropy of statistical mechanics. In informational sense, musical entropy can express firstly the *uncertainty* about the distribution of the grains in a screen. Hereto, it is just related to time inasmuch as one calculates it for each of the ΔT's (i.e. screens) that succeed each other in the evolution of the sonic event. Secondly, it can also express the uncertainty about the possible transformations of a screen, i.e. the uncertainty about the the next incoming term of a given screen transition. We know that, as highlighted by Ben-Naim, informational entropy is independent of the time. It is thus straightforward to see that musical entropy too, if interpreted in such informational terms, is definitely not a function of time, and – similarly to the case of the entropy of statistical mechanics if interpreted in the same informational terms – clearly not associated with the arrow of time.

To sup up: differently from statistical mechanics, entropy-related constraints for a specific and necessary direction of the evolution of sonic events are absent in music. The succession of the screens of a noise could be reversed or repeatedly rearranged and we would have again (empirically possible) noises. The succession of the screens corresponding to a musical composition could be similarly reversed or repeatedly rearranged and we would have new musical compositions. The problem of the arrow of time does not exists in music as in statistical mechanics, and time can be reversed in music as far as new sonic events can be constructed, *in a theoretically consistent way*, as reversals of descriptive screen successions related to actual sonic events. Inasmuch as these successions contain entropy-increasing screen transitions, the sonic events constructed as their reversals will contain correspondent entropy-decreasing transitions, and vice versa. Expressing it with [Callender 1997], in music and acoustics in general are thus to be found some of the domains for which models, predictions and theoretical constraints of statistical mechanics about entropy and time are not true.

4 Conclusion

We discussed in [Section 2] all the most important technical aspects of the concept of entropy in both statistical mechanics and stochastic music theory. Then, in [Section 3.1] we passed to analyze thoroughly the formal and syntactic affinities and differences between this concept as applied by both disciplines, and in [Section 3.1] the semantic and philosophical ones. In particular, to this respect we found out that there exists a substantial difference in meaning between the two applications of entropy. If for statistical mechanics there is generally an empirically univocal and constrained direction of the evolution of the systems and consequently a univocal and constrained direction of the time, in music we have nothing like that, and the problem of the arrow of time itself does not arise.

Having all this in mind, we can finally come to the conclusion of our epistemological inquiry and thus attempt to give an answer to our initial question, namely if there exists an *isomorphism* between the entropy of statistical mechanics and that of stochastic music theory, and, if so, to what extent. The answer may appear quite straightforward to the reader: an isomorphism exist, but it is not a total isomorphism. Many formal and philosophical aspects of the statistico-mechanical and musical application of entropy coincide, but many others – perhaps the majority or/and the most important ones – do not, unless one accept to interpret both statistical mechanics and stochastic music theory informationally. In this case, as we have in fact noticed, characteristics of both theories that otherwise differ, would instead be more comparable with each other.

Furthermore, even if both theories share some formal aspects and a similar mathematical framework, they definitely characterize, as to their respective *meanings*, their respective target phenomena in different ways. For example, stochastic music theory applies, surely in a successful way, some formal devices, mathematical tools and theoretical presuppositions of statistical mechanics: among others, the idea, that macroscopic systems can be described by their microstates, the idea that this description can only be a statistical and probabilistic one, etc. However, it does not characterize its objets in the same sense in which statistical mechanics does for its own: if, for instance, this latter attributes mechanical properties to the particles, the grains of stochastic music theory are apparently not characterized in the the same mechanical sense – and indeed fall short, in general, of a clear (metaphysical) characterization. This may be perhaps also the reason why several philosophical problems of statistical mechanics – like for example those connected with the use of reversible dynamical and mechanical laws at the microlevel for explaining time-asymmetric macroscopic processes – do not arise in the case of stochastic music theory, being attributed here to microscopic objects no mechanical nature.

Therefore, in the end, we can say that the entropies of statistical mechanics and stochastic music theory are not quite the same concept, but rather different applications, each with its own peculiarities, of a same general idea which partly refers to the heuristic notion of the measure of disorder and disorganization – or uncertainty, in informational terms – within the general theoretical context of a microscopic explanation of macroscopic phenomena.

It however remains very fascinating the fact that this idea can be applied in domains so different from each other, among which even one generally considered extra-scientific, namely music. Indeed, the question arises: does some epistemic gain or enrichment derive from applying the same concept or idea in different domains and in different ways? The answer to this question is not easy. Yet, it is very probably a positive one: extending a concept throughout different domains gives the opportunity not only to understand more of the concept itself, to discover new aspects and theoretical possibilities of it, but also to understand more about the nature of the phenomena this concept refers to and of the characteristics these phenomena share with each other. Last but not least, the (consistent) application of a concept in different domains, each with its different language and forms of expression, also opens the possibility for new, and in some cases "non-conventional", ways of communicating and transmitting that concept: is it possible, for example, to *understand* something about the general concept of entropy as applied, say, in statistical mechanics or information theory, maybe just by listening to a stochastic musical composition?

To these and other interesting philosophical questions we shall attempt to give an answer in a forthcoming work.

5 References

Albert, D. (1996). *On the Character of Statistical-Mechanical Probabilities.* Oral presentation, PSA96, Cleveland, OH, November 3, 1996.

Ben-Naim, A. (2017a). Entropy, Shannon's Measure of Information and Boltzmann's H-Theorem. *Entropy*, 19(2):48.

Ben-Naim, A. (2017b). Is Entropy Associated with Time's Arrow? *arXiv preprint, arXiv:1705.01467.*

Bernoulli, D. (1738). *Hydrodynamica, sive, De viribus et motibus fluidorum commentarii: opus academicum ab auctore, dum Petropoli ageret, congestum.* Sumptibus J. R. Dulseckeri, Basel.

Boltzmann, L. (1868). Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten. *Wiener Berichte*, 58:517–560. In Boltzmann, L. (1909). *Wissenschaftliche Abhandlungen*, Vol. I, pages 49-96.

Boltzmann, L. (1872). Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen. *Wiener Berichte*, 66:275–370. In Boltzmann, L. (1909). *Wissenschaftliche Abhandlungen*, Vol. I, pages 316-402; and in Brush S.G. (1976), *Kinetische Theorie II*, pages 115-225.

Boltzmann, L. (1877). Über die Beziehung zwischen dem zweiten Hauptsätze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung resp. dem Sätzen über das Wärmegleichgewicht. *Wiener Berichte*, 76:373–435. In Boltzmann, L. (1909). *Wissenschaftliche Abhandlungen*, Vol. II, pages 164-223.

Boltzmann, L. (1898). *Vorlesungen über Gastheorie.* J. A. Barth, Leipzig. 2 Vol.

Boltzmann, L. (1909). *Wissenschaftliche Abhandlungen.* Hasenöhrl F. (ed.). Johann Ambrosius Barth, Leipzig. Volumes I-III.

Brush, S. (1976). *Kinetische Theorie II: Irreversible Prozesse Einführung und Originaltexte.* Vieweg+Teubner Verlag, Braunschweig.

Callender, C. (1996). *Explaining Time's Arrow.* Thesis submitted for Ph.D., Rutgers University. Friedman, M. (1974), "Explanation and Scientific Explanation", *Journal of Philosophy*, 71:5-19.

Callender, C. (1997). What is 'the problem of the direction of time'? *Philosophy of*

Science, 64:S223–S234.

Callender, C. (2016). Thermodynamic Asymmetry in Time. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition.

Carnot, S. (1824). *Réflexions sur la puissance motrice de feu et sur les machines propres à développer cette puissance*. Bachelier, Paris.

Clausius, R. (1854). Über eine veränderte Form des zweiten Hauptsatzes der mechanischen Wärmetheorie. *Annalen der Physik*, 169(12):481–506.

Clausius, R. (1857). Über die Art der Bewegung, welche wir Wärme nennen. *Annalen der Physik*, 100:353–380.

Clausius, R. (1865). Über verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie. *Annalen der Physik*, 201(7):353–400.

Eddington, A. S. (1928). *The Nature of the Physical World*. Macmillan Co. and Cambridge University Press, New York-Cambridge (En.).

Ehrenfest, P. and Ehrenfest-Afanassjewa, T. (1911). Begriffliche Grundlagen der Statistischen Auffassung in der Mechanik. In *Encyklopädie der Mathematischen Wissenschaften mit Einschluß ihrer Anwendungen*, volume IV-4.II. Klein, F. and Müller, C.

Emch, G. G. and Liu, C. (2013). *The logic of thermostatistical physics*. Springer Science & Business Media, Heidelberg & Berlin.

Ernst, G. and Hüttemann, A. (2010). *Time, Chance, and Reduction: Philosophical Aspects of Statistical Mechanics*. Cambridge Universtiy Press, Cambridge.

Fletcher, H. and Munson, W. A. (1933). Loudness, its definition, measurement and calculation. *Bell System Technical Journal*, 12(4):377–430.

Frigg, R. (2008). A Field Guide to Recent Work on the Foundations of Statistical Mechanics. In Rickles, D., editor, *The Ashgate Companion to Contemporary Philosophy of Physics*, pages 99–196. Aldershot & Burlington, Vt.: Ashgate.

Frigg, R. (2009). Typicality and the Approach to Equilibrium in Boltzmannian Statistical Mechanics. *Philosophy of Science*, 76:997–1008.

Frigg, R. (2010a). Probability in Boltzmannian Statistical Mechanics. In Ernst, G.

and Hüttemann, A (eds.), pages 92-118.

Frigg, R. (2010b). Why typicality does not explain the approach to equilibrium. In Suárez, M., editor, *Probabilities, Causes and Propensities in Physics*, pages 77–93. Springer, Berlin.

Frigg, R. and Werndl, C. (2011). Entropy - a Guide for the Perplexed. In Beisbart, C. and Hartmann, S., editors, *Probabilities in physics*, pages 115–142. Oxford University Press, Oxford.

Gardner, M. (1970). Mathematical games. *Scientific american*, 222(6):132–140.

Gibbs, J. W. (1902). *Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundation of Thermodynamics*. C. Scribner, New York.

Goldstein, S. (2001). Boltzmann's approach to statistical mechanics. In *Chance in physics*, pages 39–54. Springer.

Hartley, R. V. L. (1928). Transmission of Information. *Bell Syst. Tech. Journal*, 7.

Helmholtz, H. (1862). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. F. Vieweg & So., Braunschweig. En. tr. by A.J. Ellis (1885), Dover Publications, Mineola (NY) 1954.

Jaynes, E. T. (1957a). Information theory and statistical mechanics. *Physical review*, 106(4):620.

Jaynes, E. T. (1957b). Information theory and statistical mechanics. ii. *Physical review*, 108(2):171.

Klein, M. J. (1973). Mechanical explanation at the end of the nineteenth century. *Centaurus*, 17(1):58–82.

Krönig, A. K. (1856). Grundzüge einer Theorie der Gase. *Annalen der Physik*, 99:315.

Lebowitz, J. (1994). Time's arrow and Boltzmann's entropy. In J.J. Halliwell, J. P.-M. and Zurek, W., editors, *Physical Origins of Time Asymmetry*, pages 131–146. Cambridge Universtiy Press, Cambridge.

Lebowitz, J. L. et al. (1993). Boltzmann's entropy and time's arrow. *Physics today*, 46:32–32.

Loschmidt, J. (1876). Über die Zustand des Warmegleichgewichtes eines Systems von Körpern mit Rucksicht auf die Schwerkraft. *Wiener Berichte*, 73:128–142, 366–372.

Manzara, L. C., Witten, I. H., and James, M. (1992). On the entropy of music: An experiment with bach chorale melodies. *Leonardo Music Journal*, 2(1):81–88.

Maroney, O. (2009). Information Processing and Thermodynamic Entropy. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2009 edition.

Maxwell, J. (1871). *Theory of Heat*. Longman and Co., London.

Maxwell, J. C. (1867). On the Dynamical Theory of Gases. *Philosophical Transactions of the Royal Society of London*, 157:49–88.

Müller, I. (2007). *A History of Thermodynamics. The Doctrine of Energy and Entropy*. Springer, Berlin.

Poincaré, H. (1889). Sur les tentatives d'explication mécanique des principes de la thermodynamique. *Comptes Rendus de l'Académie des Sciences (Paris)*, 108:550–553.

Poincaré, H. (1893). Le mécanisme et l'expérience. *Revue de Métaphysique et de Morale*, 1:534–537.

Reiss, H. (1965). *Methods of thermodynamics*. Dover Publications, Mineola, N.Y.

Schrödinger, E. (1944). *What is Life?* Cambridge Universtiy Press, Cambridge.

Sedgewick, R. and Wayne, K. (2011). *Algorithms*. Addison-Wesley, Boston.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell Syst. Tech. J.*, 27(3):379–423.

Shannon, C. E. and Weaver, W. (1949). University of Illinois Press, Urbana and Chicago.

Sklar, L. (1993). *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics*. Cambridge Universtiy Press, Cambridge.

ter Haar, D. (1955). Foundations of Statistical Mechanics. *Reviews of Modern Physics*, 27:289–338.

Tolman, R. C. (1938). *The Principles of Statistical Mechanics*. Clarendon Press, Oxford.

Uffink, J. B. (2007). Compendium of the Foundations of Classical Statistical Physic. In Butterfield, J. E. a. J., editor, *Handbook for the Philosophy of Physics*, pages

923–1074. Elsevier.

Wu, K., Nan, Q., and Wu, T. (2020). Philosophical Analysis of the Meaning and Nature of Entropy and Negative Entropy Theories. *Complexity*, 2020.

Xenakis, I. (1992). *Formalized Music: Thought and Mathematics in Composition*. Pendragon Press, Hillsdale, New York.

Zermelo, E. (1896). Über mechanische Erklarungen irreversibler Vorgange. *Annalen der Physik*, 59:793–801.

Notes on the Illustrations

All the images used in this work have been taken from [Xenakis 1992]. In [Figure 3] and [Figure 4] some minor and non-altering elements (arrows, symbols for the name of the axes, etc.) have been added for facilitating the understanding of Xenakis' arguments as we exposed them.