Abstract

This paper will reconsider the logical path of reducing the second law of thermodynamics to mechanics up to the point where it seems to drop the ball. The thesis of an off-centered interpretation of the mechanical reduction is presented in the light of a promising (although unusual and contested) conceptualization of the second law. This will be accomplished in a conceptual way and strictly between the borders of classical physics.

1 Introduction

The endeavor to reduce thermodynamics to first principles of classical mechanics implies the description of irreversible macroscopic phenomena in terms of what is considered a more fundamental microscopic reality. I will call this the reductionist project. This theme is usually conflated with discussions about the origin of an alleged arrow of time. This can produce misunderstandings, especially in treatments where different arrows of time are considered and mathematical manipulations, physical explanation and philosophical assumptions are intermingled.

1 For contacts: l.serafino@ack.edu.kw
**By definition,** the reductionist project has to face a well known difficulty usually called *paradox*: macro-systems described by thermodynamics have a time directed, irreversible phenomenology, despite the fact that underlying laws of mechanics are time symmetric. How to overcome this paradox has been a challenge in physics for centuries and I will focus on it here because this forms the main conceptual problem for the reductionist project. The issue is if classical mechanics, i.e. *particles* and *trajectories* satisfying Newton’s dynamical principles, is enough or not to resolve the paradox. Historical accounts of the problem (Boltzamnn work, *Stosszahlansatz*, Loschmidt’s and Zermelo’s objection to Boltzmann, etc.) are well described in some recent reviews [1,2]. There is no need to repeat them here.

If and up to what extent the reductionist project succeeded or not is still a matter of debate and it depends also on one’s own specific scientific community background. For kinetic theory scholars the question has been closed, at least conceptually, in the 1975 with the *Lanford’s mathematical results* [3, p. 36]. From a different point of view others question it [4]. A look at the recent scientific production in philosophy of physics shows that it remains a hot topic [5]. The main aim of this paper is to present a clear picture of the reductionist conceptual path to the mechanical explanation of the second law up to where it seems to give up. I will then propose a provocative conceptual framework to overcome this limitation. I will do this without trespassing the borders of classical physics and in a pure qualitative way. In what follows I will assume only a basic knowledge of mechanics and thermodynamics namely the second law of thermodynamics in its classical formulation (Kelvin–Planck and Clausius postulates) and the related concept of entropy.

First remark. One thing is to reconcile conceptually macro-irreversibility with micro-reversibility; another story is to derive mathematically hydrodynamic equation able to predict macroscopic behavior in accord with experimental evidences. For me, one thing is physical reality and another is mathematics. I will keep these two tasks clearly separated and I will focus on the conceptual side here. I will also consider separately the reductionist project and the discussion about the emergence of an arrow of time because I think the latter is a separated and more general issue.
As commonly happens in foundation of science, confusion arises when the language we use is not clear and terms are not properly defined. **Reductionism** is a position in scientific metaphysics that insists that higher-level domain phenomena can be explained by reference to the properties of the lower-level entities that make them up ([6] p. 696). Canonical examples are chemical properties that are supposed to be reducible to quantum mechanics and genetics that is supposed to be reducible to molecular biology. Reduction can be understood as a kind of **explanation**. In trying to reduce thermodynamics to mechanics we assume the atomic nature of matter obeying Newton’s law to be fundamental; macroscopic irreversibility should be logically derived from them. It is very common in this field to find papers with expressions like “the origin of irreversibility” or “the source of irreversibility” and so on. We can agree that “looking for the origins” means “look for an explanation of”, explanation that can be reductionist or not. In any case I think that using this term can be confusing and misleading especially in technical (i.e. mathematical) accounts of the field.

The idea of deducing truths about the macroscopic world starting from Newton’s law can appear at the outset very delicate because the ambition is to explain something that has been conceived a principle of physics in a more fundamental way, something that cannot be questioned, always and necessarily true. If we embark in this mission we refuse to consider the second law as an axiom but it becomes something that deserves to be explained. If it succeeds, it will no longer be correct to affirm:

*The explanation of irreversibility is in the second law of thermodynamics*...

Instead

*The explanation of that macroscopic phenomenology that we define (in a way to specify) irreversible and that is described by the second law, is [some mechanical explanation here]...*

The point is now if it is possible to provide coherent and logical argument to motivate the reductionist project. The following discussion is a simplified version of the way the reductionist project has been conceived in its main elements since the original ideas of Ludwig Boltzmann at the end of XIX century [7, p. 83].
2 The coin of the reductionist project

To this aim, let’s consider the archetypal example of process of interest in thermodynamics: the free expansion of an isolated gas in a box illustrated in the figure 1. As flipping a coin can tell us everything about probability theory, the gas-in-the-box model contains everything we need here. If we assume the atomic ontology of matter, we have small spheres ceaseless zigzagging randomly all around with a given kinetic energy. In this model, the direction of a particle can change only by elastic collisions with other particles or with the container’s walls. Other classical assumptions are that the size of the particles is negligible in comparison to the distance traveled and that particles exert no intermolecular forces on one another.

Why if we remove the internal constrain after some time we will observe a configuration like in B? Simple, inertia! Particles now are free to fill the space available. The system will reach a configuration like in figure 1-B where the number of particles is equally distributed in both sides of the box. In the thermodynamics jargon this condition specifies that the system has reached a new equilibrium: macroscopic parameters will remain the same with time and entropy is at its maximum.

Let’s consider it from the mechanical point of view. We have particles moving around. Why we do not observe the opposite process where all the spheres go back spontaneously to a configuration like in figure 1-A? Well, this is not impossible but it requires a series of very low probability events. For the set of particles in the box, there is no a priory reason to expect that some velocities directions can be privileged. All have equal right to take place inside the box for each particle.

Figure 1
Out of the all possible combination of positions and velocities in every direction, only very specific arrangements can generate the reverse movie for the gas: particles on the left side all moving without disturbing each other towards the breach and at the same time particles of the right side provided with velocities that prevent them to go on the left as represented in figure 2.

Again, it is not theoretically impossible that a similar configuration will be generated out of the random particles impacts, it is just highly unlikely. If we imagine that in a mole of gas the number of elements is of the order of $10^{23}$, the probability of an event like this is extremely low. This is just an example for the gas in a box in figure 1 but for other configurations similar arguments hold.

So the explanation of this irreversible phenomenology:

- the gas is initially confined in the state figure 1-A;
- the gas goes spontaneously from state figure 1-A to state figure 1-B;
- the gas remains in a figure 1-B-like state: particles are dispersed throughout the available space;

it is understandable starting from classical mechanical BUT we need to add a probability ingredient to complete the picture. Probability reasoning is what usually enters in an explanation when we deal with large numbers like in this case.

This is why commonly it is said that thermodynamics is reduced to “statistical” mechanics: because of this unavoidable probabilistic ingredient. Usually these arguments
are elaborated inside the frame of **Gibbs ensemble** more formal approach but the moral of the story is exactly the same [8].

At the end of the general explanation of the reductionist project sketched above, some important remarks are in order. In our mechanical derivation **apparently**:

- **There is nothing mysterious in the probability arguments used.** What we call probability is just a *relative frequency* out of the total possible velocity configuration that the system can have. The number of particles is huge but finite so no need to worry about asymptotic behaviors or measure theory hurdles.

- **We do not need to take into account chaos theory or extra randomness properties.** Since what we observe is just the result of basic classical mechanics knowledge and some probability reasoning.

- **We do not need to take into account interactions with the environment.** The gas is constrained in a portion of space whose borders represent an important constituent of the system. At this point there is no reason to affirm that external environmental perturbations play a role in the described irreversible phenomenology.

- **We do not need to take into account ignorance of the observer.** There are many B-like states and maybe observer cannot distinguish between them but it make no sense to affirm that this is a kind of physical explanation of what he is observing.

- **We don’t need to take into account quantum mechanics or even some more exotic stuff.** It is enough evident I suppose.
From the conceptual point of view the reductionist project succeeds in proposing a clear and parsimonious framework able to explain observed thermodynamic behavior. Once the conceptual basis are set, statistical mechanics continues the project with the definition of macroscopic variables in terms of a mechanical ontology. So temperature will be related to translational kinetic energy, pressure is connected to momentum transfer and so on. A further step is to derive kinetic equations able to describe approach to equilibrium time-evolution of systems in agreement with experimental results. So far so good, is the tale already over? Not even in your dreams, the best is yet to come.

3 From the past hypothesis to the Landauer’s principle

If we accept the reductionist project in the representation sketched above we have also to accept that what we usually refer to as irreversibility is an epiphenomenon and that the second law - as a macroscopic description of it - is open to violations as a result fluctuations of the system. We cannot predict when a configuration shown in figure 2 will occur, maybe in billions and billions of years. At this point one is naturally led to wonder if experimentally reproducible violations are possible. Again, for the reductionist there is no reason to believe that in principle this is not possible but this clash horns with the following matter of fact:

**Crucial question:** if it is only a matter of classical mechanics, why is it experimentally so hard to reproduce initial conditions that lead to second law violation?

A common answer to this question uses an argument that it is generally known as the past hypothesis: it is hard to reproduce anti-thermodynamic initial condition because the Universe is in a low entropy state, and so it was yesterday, and the day before and so on. Continuing with this line of reasoning we end assuming that the Universe initial state (the so called Big Bang) was in a condition of low entropy [9].

This answer, if we accept the reductionist justification, is not valid on logical basis. It implies a circular reasoning way of thinking in which the second law seems to
re-enter surreptitiously as a form of explanation to put an end to the story in an ambiguous way. Are we trying to reduce thermodynamics to mechanics or not? The reductionist project cannot give up so early.

Assuming for a moment that it makes sense to talk about the entropy of a baby Universe, if we define it as a quantity proportional to the number of microstate corresponding to a given macrostate at equilibrium, I can accept the proposition

*In the past the entropy of the Universe was less than now*

But this cannot answer the crucial question because entropy in itself does not have any causal power to prevent us to arrange an anti-thermodynamic behavior. Entropy is not a driving force; it is just a description of a state like volume.

The point at issue is located elsewhere and the argument goes as follows. Going back to the *crucial question*, let’s consider again the gas in the box because the behavior of this is what we want to elucidate. The main point here is the distinction between macroscopic state described by thermodynamics and microscopic configuration. When we operate on this system macroscopically for example with a compression we treat the gas as a collective and we cannot select specifically single microstates.

To obtain a state like in *figure 2 as initial condition* we need to arrange a very particular configuration of positions and velocities out of the broth of randomly moving particles. So we need the help of a *demon* able to manipulate velocities and position at the particle level. To this day, the possibility to put such a hell’s creature into practice is out of human technological possibilities *but* even if one day it will be realized there is a strong argument that this will not invalidate the second law. Here is where a new piece of the puzzle enters into the story: the *Landauer’s principle*.

To put it simply, this information-theoretic principle states that even the most basic form of information elaboration (in this case about particles velocities arrangements) is inevitably accompanied by generation of heat [10]. According to the common interpretation, even assuming the most performanting one, such a demon cannot avoid generating more entropy in the Universe than the amount reduced acting on the
system of interest. This theoretically prevents the possibility of an experimentally reproducible second law violation. Of course a lot has been said about Landauer’s argument, its validity and its connection to second law violations and studies in this direction go on day in and day out [11]. A recent article published on *Nature* shows experimental evidences concerning the validity of the principle [12], thus I will take it seriously in what follows.

It seems that following the logical path of the reductionist project led us eventually to a suspicious situation. Putting together all the pieces of the reductionist rationale developed so far and the Landauer’s result we get trapped in this fascinating:

**Assertion 1.** *The mechanical explanation of the second law leaves open the possibility of a spontaneous violation (through random fluctuation) but not to an experimentally reproducible violation.*

This is really intriguing from a philosophical point of view. Does this means that the second law is really something more fundamental than what the mechanical explanation is showing? Does this nullify the entire reductionist castle? Is there something that has been omitted or neglected in the logic of derivation? To arrange a given initial condition to the gas we need to interact with it with some devices that obey to physical laws too. Here is where external world enters into the picture. The system plus the experimental device now must be considered jointly and it is to this new enlarged system that we need to apply a thermodynamic rationale. In essence, Landauer’s principle says that there is an unavoidable energy cost in every interaction – however feeble it could be - between the particle and the apparatus below which is not possible to go. This in turn account for the increase in final entropy for the enlarged system.

---

2 The working logic of the well-known usual *Maxwell demon* is different from the one presented here but theoretical limits are the same.
The difference between a spontaneous and a reproducible violation lies in the fact that in the former case the system is assumed to be perfectly isolated. The point now is that this precondition is of course physically not tenable in reality since an exchange of energy with surrounds cannot be prevented in the strict sense [13]. In a more realistic scenario we can imagine that even in the case where we can observe an improbable event like in figure 2 during our lifespan this does not imply a violation of the second law if we take into account the total dissipation of energy in the environment through walls or in some other ways.

This seems to lead us to the need of an interventionist explanatory stance i.e. the inescapable role played by interactions of the system of interest with the external environment [14]. In any case at this point the status of the reductionist project seems more shaky than initially imagined. If assertion 1 is not totally satisfactory, we need to deduce that the reductionist project is hiding something and it can offer only a (however good) partial picture of observed irreversible phenomenology. In the next section I will put forward an elucidation about the limits of the reductionist project appealing to a different perspective of the second law.

4 Energy spreading and the off-centered interpretation of the reductionist project

Now that we have reached this stage, we can accept that assertion 1 complete the story and don’t care too much, keeping a statistical interpretation of the second law for good. Otherwise, if this asymmetry between random violation and reproducible violation leaves a bitter taste in mouth what remains to do is to try to motivate why even random violation have no concrete possibility to appear. Landauer’s principle, experimentally verified, seems to tell something deeper that pushes in this direction. This must be done in a way that keeps what the reductionist project so far discussed was able to offer and gives further insights to resolve what seems to be lacking. In what follows for the sake of moving out from this impasse I will elucidate one possible way out that leads us to take
into account the interventionist stance i.e. to see how the fact that the system cannot be strictly isolated from the surroundings can be a significant ingredient for the issue at stake.

The central example in this paper has been the free expansion of a classical ideal gas but this does not tell us the whole story. In the gas-in-the-box model we have particles moving around in trajectories enclosed in an ideal vessel but this is hiding the fact that the second law is telling us something about energy and not about particles. In this sense the mechanical explanation attempt is off-centered and in the rest of this section I will develop and clarify this interpretation.

In what follows I will endorse a conceptualization of the second law of thermodynamics as energy spreading ([15] p. 77) and I will try to elucidate how this can shed light on the issue at stake. In my opinion this vision becomes also coherent with the information-theoretic approach envisaged during Landauer’s principle discussion above. Further, even if the reductionist project seems to capitulate and the second law gaining the upper hand, it will appear clear how thermodynamics can be reconciled peacefully with mechanics and even more. The main pint is to make sense how the sentence “exchange of energy with surrounds cannot be totally prevented” can lead to the desired direction.

Canonical treatments of the second law associate it with increasing disorder, missing information and the like. A different, albeit contested, conceptualization is to look at the second law as connected to energy dispersal tendencies. The second law refers to this property of energy. Entropy represents the measure of this dispersal. This strengths and limits of this interpretation has been critically considered in recent works [16].

Basically, we need a shift in our focus form a dynamical account to an energy account with its tendency to spread temporally and spatially. This becomes evident when we consider more complex thermodynamic systems than the ideal gas model. The energy in a molecular system can be shared by potential, translational, rotational, vibration, electronic, and intermolecular modes. In this more general understanding, the second law postulates that energy spreads and seeks out all available storage modes and this is maximal when thermodynamic equilibrium exists [17].
This overturning of perspective leads us to recast the reductionist project from a different standpoint. Through the lens of this spreading conceptualization we can reduce thermodynamics to the following ontological basic elements: space, time, energy, its (with an intentional abuse of language) “inertial” properties and forces able to curb it locally. The second law appears to refer to a fundamental property of energy when it is released from constrains. In this interpretation every discussion about incompatibility between thermodynamics and mechanics must be reconsidered.

In the mechanical description of the ideal gas-in-the-box model particles and their dynamics are the leading actors used to describe the system. Energy enters as a property of particles in the form of translational kinetic energy that is conserved during dynamical evolution. The ideal nature of the interface between particles and the surroundings is assumed to be truly insulating where only perfect elastic collisions are admitted, this represents the cut between what is inside and the external environment. The same idealization appears in considering collision between particles and the lack of other internal energy storage modes.

Because of this idealization, this model can tell a lot, but not the complete story. This does not mean that it is totally flawed and indeed up to a given extent it succeeds in describing the behavior of the gas as we have seen. This is an approximation that takes into account only partial dissipation modes of energy. In this sense the isolated-gas-in-the-box model is a misleading example – a biased coin we can say – even if it represents the canonical way to introduce the topic. Here the whole discussion has been grounded on it exactly for this reason: to stress its limits. Basically what is missing in this derivation is the unavoidable energy exchange due to interaction with the surroundings, that, however small, it will never vanish in real cases.

We can use this argument to see how even random, unpredictable second law violations only apparently can have place. When an ideal gas is allowed to expand in a larger volume, this results in a greater dispersion of the total particles kinetic energy. According to the molecular view, we can expect that the system can – although with a tiny probability - also reproduce spontaneously a dynamical configuration apparently anti-thermodynamic. Nothing in the dynamics prevents this possibility for velocity
directions. In this scenario the second law is apparently violate unless one considers the realistic **total energy spreading balance** including the surroundings and all possible storage modes. The total energy dissipation gives rise to an overall greater entropy state against the reduction inside the container and there is no conflict with the second law of thermodynamics. In this scenario it must be accepted that the only apparently anti-thermodynamic behavior allowed for the gas inside the box must fall in this typology.

A strict mechanical portrayal with its primitive ontology in terms of particles and their trajectories will offer only a partial account of what is actually happening. This is the very place where apparent contrast between mechanics and thermodynamics arises. Because of its ontological stance, when mechanics focuses on ideal models it sacrifices energy expression. Thus for examples it keeps perfect elastic properties of physical bodies and discards energy dissipation modes. In a mechanical description it is assumed that we can push this idealization as far as it is desired. From this it follows that any mechanical representation of underlying energy properties will be partial and even “conflicting” by definition.

The main aim of this paper was to review the logical path of the reductionist project i.e. to see if higher-level domain phenomena can be explained by reference to the properties of the lower-level entities. Thus we have reached a stage where it appears clear where the limits of this project are located and this justify the following

**Assertion 2. In the reductionist project, the ontological priority to mass particles and their trajectories over energy produces an off-centered account of irreversibility.**

In thermodynamics, energy and its manifestations acquires an ontological primacy in its scientific endeavor. Further, thermodynamic macroscopic point of view enables it to grab energy properties that reveal themselves manifestly only at a more general and collective scales of analysis. This standpoint allows thermodynamics to emphasize
energy nature that is not possible to fully accomplish with a mechanical description in terms of perfect particles and their trajectories.

Of course thermodynamics has its own idealization too but it is in the position of disclosing something deeper about energy that mechanics, from its ontological stance, is no able to do. Mechanics can take into account energy conservation properties but has limitations in describing spreading properties that can be fully accomplished only through a comprehensive vision that prioritize energy in all its expressions.

Collective behavior of particles is a way through which energy spreads globally. This must be considered as a fundamental fact of nature in the same way as speed of light is and so it cannot deduced but – for the moment - accepted. Locally, energy can be temporally and spatially constrained but globally it will result in a more dispersed configuration in any case. In this light mechanics of many particles become intelligible through thermodynamics and not the opposite. Particle kinetic energy is one storage mode through which energy can be dispersed spatially and temporally.

Thermodynamics speculation about energy reflects something deeper than mechanics can do. This can appear paradoxical but it is perfectly in line with a vision that gives ontological priority to energy and its spreading properties. Mechanics and thermodynamics reconcile each other under this conceptual umbrella.

As said, this does not mean that kinetic theory is fundamentally flawed and its results are there to prove it. What it is stressed is that it can provide only a partial story because only internal translational kinetic energy spreading modes are taken into account. What has to be stressed is that the attempt of deriving irreversibility from an ontology that focuses on material bodies dynamics to the detriment of energy will not be completely successful. In the idealized gas expansion model this does not happen by definition. In such a situation irreversibility can be derived only through approximations of some sorts. This is the reason why mathematical derivation irreversibility purely from

3 Unless there is evidence that an explanation can be located at a deeper level of physical reality (quantum, elementary particles).
mechanics always requires an *ad-hoc* time asymmetric ingredient [18]. Boltzmann, Fokker-Planck, master equations and so on are examples of this [19].

5 **Interlude: the generating power of the Universe**

Speculating about energy at a more philosophical level, two concepts come to mind in the form of a binary opposition: *global freedom* vs. *local constraints*. Freedom enters through *inertial properties* and constrains through *forces*. Thermodynamics appears to be a close interplay between these at an aggregate level. A never-ending ballet between energy’s freedom desire of spreading around and opposite tendencies to constrain it. In a sense, the Universe seems in a perennial state of strain and this is unveiled in fluxes and gradients that keep it alive. It is thanks to this cosmic rule that complex structures, including life matter, emerges out of turmoil. This *global freedom tendencies / local constraints duality of energy*, at a metaphysical level, embodies the creative power of the Universe whereas the second law focuses more on the demolishing part of the story.

6 **The ascent of mechanics to thermodynamics**

The scope of this article was directed primarily at reconsidering the logical reductionist path and to see up to where it can lead us. This led us to assertion 1. At that point we had two possibilities, assuming that we want to take seriously the implications of Landauer’s principle i.e. even for the most fundamental form of interaction, elaboration of another system information, energy dissipation is unavoidable. One possibility was to accept the verdict about the difference between reproducible and unpredictable second law violation. The other was trying to elucidate if and why even this latter possibility is actually negated. This has been reached basically trough an interventionist *explanans* in the frame of a specific interpretation of the second law. This route implicitly asserts that the reductionist project must be reconsidered; it is not

---

4 Global in the sense of “valid for the universe as a whole”
possible to fully account for irreversibility from mechanics but the “mechanical” part of the story plays an important role in the overall framework.

The spreading metaphor of the second low has been used to try to shed light on this from a different point of view. This led the discussion to the limits of a mechanical interpretation of thermodynamics summarized in the concept of off-centered interpretation. Assertion 2 states that an ontology that prioritize physical bodies over energy will sooner or later fail to take into account all the consequences of the second law in purely mechanical terms. Of course the spreading view has its own limits and it’s open to criticism. It has been used here as an interpretive tool able to reconsider an interventionist stance with the limits of the reductionist project.

This interpretation seems to be able to reconcile thermodynamics and mechanics but the status of the second law as fundamental principle remains untouched. Speculating even further, if we adopt an ontology that assigns priority to energy over other physical entities, it eventually makes sense to talk about an ascent of mechanics to thermodynamics. In this apparently bizarre suggestion, we can see mass properties like inertia as manifestation of the more general spreading tendencies of energy. This open up a new interesting arena of speculation also about the possibility to explain energy properties at a deeper level and at the and we can say that there’s a long way to go yet to consider the reflection about this matter a close chapter.

Bibliography
