

On the Deduction of the Carathéodory's Axiom of the Second Law of Thermodynamics from the Clausius and Kelvin Principles

Pedro F. da Silva Júnior*

Universidade Federal de Pernambuco, Caruaru, PE, Brazil

Abstract

Carathéodory's formalism for classical thermodynamics is a rich alternative approach to this theory, although unpopular with students and physics professors. This approach dispenses with the content of thermal machines for the presentation of the second law of thermodynamics. In this paper, we discuss Carathéodory's formalism historically, and show how Carathéodory's axiom of the second law of thermodynamics is derived, didactically, from the Clausius principle and the Kelvin principle. In addition, also providing an objective character for this paper, in the sense of seeking to popularize the teaching of Carathéodory's formalism in disciplines of classical thermodynamics at undergraduate level, we guide the reader to obtain the entropy and mathematical content of the second law of thermodynamics through this formalism. Finally, considering the wide reviewed literature, the proof we gave for deducing the Carathéodory's axiom of the second law of thermodynamics from the Clausius principle is new.

Keywords: Carathéodory's formalism, Carathéodory's axiom, second law of thermodynamics, Clausius principle, Kelvin principle.

1. Introduction

In physics, we are often presented with distinct but equivalent conceptual and mathematical approaches to the same theory. By equivalent we mean that these different descriptions for the same theory obtain, without any loss of physical content, the same final results. Moreover, generally, the paths and methods used by each description differ enormously from each other. These distinct descriptions for the same theory we call *formalisms*. A famous case of formalisms in physics occurs in classical mechanics, where we have the formalisms due to Newton, Lagrange, and Hamilton.

Another discipline of physics that allows the use of formalisms is classical thermodynamics. Classical thermodynamics is the perspective of thermodynamics that studies physical systems from laws that generalize the observations made about the *macroscopic* behavior of these systems. To do this, classical thermodynamics ignores the *microscopic* nature of matter. This differs from another famous perspective of thermodynamics that considers for its description the *microscopic* nature of matter and advances in *statistical mechanics*. This other perspective on thermodynamics is *statistical thermodynamics*. It is not statistical thermodynamics that we will be dealing with here.

Classical thermodynamics is a highly solidified and well-established area of physics throughout the scientific community. Moreover, it is a theory whose content is considerably popular, from the most basic levels of science education, to higher level courses in physics and related areas, such as chemistry, engineering, etc. Thus, the fundamental theoretical material of classical thermodynamics should cause little strangeness to the reader. Specifically about the existing formalisms for classical thermodynamics, and also about its own teaching at the undergraduate level, it is notable the traditional presentation of this subject from the perspective of the efficiency of thermal engines – or thermal machines.

Majority in current textbooks, this traditional formalism makes use of the two equivalent experimental principles of the second law of thermodynamics – the Clausius and Kelvin principles – in conjunction with Carnot's theorem of thermal machines, to obtain the *entropy* and the mathematical content of the *second law of thermodynamics* – the principle of entropy increase. This traditional formalism was built by Clausius on historical and technical developments by names like Carnot, Clapeyron, and Kelvin [1]. Thus, we name this traditional formalism here as *Clausius's formalism*, inspired also by the teaching literature of classical thermodynamics [2].

*Correspondence email: pedro.fsilva2@ufpe.br

However, as we have previously announced, the Clausius's formalism is not the only possible one to describe classical thermodynamics. Another famous formalism originated in the work of Gibbs [3] who, in a series of papers between 1873 and 1878, advocated analytical methods for describing classical thermodynamics. In doing so, Gibbs influenced several popular works of classical thermodynamics that appeared later [4, 5]. These works constructed a classical thermodynamics of postulates, introducing fundamental thermodynamic notions such as *entropy* in the form of elementary concepts from which the other concepts of the theory are derived [6]. In general, we call here *Gibbs's formalism* the formalisms due to the pioneering work of Gibbs.

Besides these, there is also the *Carathéodory's formalism* for classical thermodynamics, which is the formalism that most interests us in this paper. The Carathéodory's formalism differs from the other two formalisms mentioned above in that it does not need the thermal machines used in the Clausius's formalism and, although it also makes use of postulates, or axioms, it does not do so using the same methodology involved in the Gibbs's formalism. The Carathéodory's formalism plays an important role in the construction of the theory of classical thermodynamics itself, as the following quote tells:

Entropy was discovered by a somewhat circuitous path through the efficiency of heat engines, a finding that in hindsight could appear serendipitous. Were we just *lucky* to have discovered something so fundamental in this way? Can it be seen *directly* that entropy as a state variable is contained in the structure of thermodynamics, without the baggage of heat engines? It can, as shown by Constantin Carathéodory in 1909.

([7] — p. 149)

These are the first words of James H. Luscombe in the introduction to the tenth chapter of his recent¹ *Thermodynamics* [7]. Despite its recognized importance² Carathéodory's formalism is almost completely unknown today by most physics professors and students. Examples of didactic productions that teach the Carathéodory's formalism are also currently scarce. That said, we present in the 2 section a short historical discussion of the background and methods of

Carathéodory's formalism. We also show, in sections 3, 4, and 5, how Carathéodory's formalism connects with one of the most important results of classical thermodynamics: the second law of thermodynamics, formulated from the principles of Clausius and Kelvin.

2. Carathéodory's formalism

In a 1909 paper published in the *Mathematische Annalen*, the mathematician Constantin Carathéodory [10], proposed a formalism for classical thermodynamics that obtained the results of the theory³ starting with two axioms, one for the first law and the other for the second law of thermodynamics, in such a way that the development of thermodynamic concepts took place in terms of considerations arising from mechanical concepts.

The axiom used by Carathéodory for the second law of thermodynamics was the greatest innovation in his work. This axiom was not based on experiments, nevertheless, from it emerged the *entropy* and its *classical mathematical content*. Although this axiom is eventually known as the *second Carathéodory's axiom* [12], given the existence of a Carathéodory's axiom also for the first law of thermodynamics, our major focus here is the study in particular of the *second law of thermodynamics*. For this reason, we will henceforth refer to Carathéodory's axiom for the second law of thermodynamics only as *Carathéodory's axiom*.

Several authors have dedicated themselves to the mission of disseminating Carathéodory's ideas over the years, namely: Sears [12], Chandrasekhar [13], Buchdahl [14], Landsberg [15], Dunning-Davies [16], among others. In particular, Max Born, one of the forerunners of quantum mechanics, defended Carathéodory's view for a large part of his life. With publications [17, 18] and also with harsh criticism of Clausius's formalism, Born sought to popularize Carathéodory's formalism by saying, for example, in a 1921 paper:

- a) *There is no other area of physics where considerations are applied that bear any resemblance to the Carnot cycle and its corollaries.*
- b) *One has to admit that thermodynamics, in its traditional model, has not yet realized the logical ideal of separation between the physical content and the mathematical description.*
- c) *It is necessary to do a removal of rubble, which a tradition*

¹Luscombe presents in his book a modern approach to the Carathéodory formalism, working it with the concept of vector fields.

²Carathéodory's work is a milestone in terms of the descriptive foundations of an analytic classical thermodynamics [8, 9].

³The full expected results of thermodynamics do not arise from Carathéodory's formalism. The same is also true of the Clausius's formalism. The greatest example of this is the third law of thermodynamics, whose precise and complete content can only be obtained by the advent of quantum mechanics. Without this inclusion, both formalisms produce the content of the third law of thermodynamics in a *textit{ad hoc}* fashion. In this sense, we can say that these formalisms provide what would be expected as the classical mathematical content of the theory. For more details on this subject, we suggest that the reader consult Tisza's book [11].

full of too much piousness hitherto, has not dared to remove.

([17])

In correspondence to Einstein [19], exactly about this publication [17], that he submitted to the current *Physikalische Zeitschrift*, Born wrote:

Frankfurt a.M.
12 February, 1921

Dear Einstein

... I have done little theoretical work. I have recently written an account of Carathéodory's thermodynamics, which will appear shortly in the *Physikalische Zeitung*. I am very curious to know what you will say about it.

([19] — p. 53)

Having already published the papers that would elevate his name to the rank of one of the greatest in the history of science - on relativity and the photoelectric effect - positive feedback from Einstein on the Carathéodory's formalism could certainly have given this matter a different reception in the physics community. But, despite Born's request, there is no mention in the subsequent correspondences between Born and Einstein [19] of the latter's concern over Carathéodory's work. Afterwards, and still on his attempt to popularize the work of Carathéodory, the frustration of Born is confirmed when he states, referring by *classical method* to what we here call Clausius's formalism:

My interpretation of Carathéodory's thermodynamics did not have the effect I had hoped for of displacing the classical method which, in my opinion, is both clumsy and mathematically opaque.

([19] — p. 55)

After Born, possible obstacles to the use and dissemination of Carathéodory's formalism were investigated by various other authors. In this context, besides the apparent historical bad luck that we have cited, pedagogical and mathematical obstacles that may have contributed to the unpopularity of the Carathéodory's formalism were pointed out by Zemansky [20] in 1966. Zemansky wrote that for the presentation of the second law of thermodynamics:

Due to the fact that Carathéodory's axiom was not based directly on experience and

that the proof of his theorem was long-winded and difficult, most physicists and textbook writers ignored the Carathéodory treatment [...]

([20] — p. 915)

Curiously, prior to Zemansky's account there were already papers that solved the issues he pointed out. For example, in 1964, Landsberg [21] proved that the *Carathéodory's axiom* can be deduced from Kelvin's principle. This result was further investigated by Tittelauer and Van Kampen [22] a year later, in 1965. In that same year, Dunning-Davies [23] proved the reciprocal of Landsberg's conclusion, establishing the *equivalence between Carathéodory's axiom and Kelvin's principle*. So, even though it is not based directly from experimental facts, Carathéodory's axiom was shown to be equivalent to Kelvin's experimental principle, so that Carathéodory's axiom can be seen as just another statement of the second law of thermodynamics [14].

On the other hand, the *Carathéodory's theorem* was made demonstrable with short arguments, related to the geometry of the *thermodynamic space* [14, 17, 18], as well as, related to the use of the concept of vector fields [7]. These demonstrations, however, follow a level of simplicity that does not contemplate all the mathematical content of Carathéodory's theorem in its general version, as Boyling [24] showed. Nevertheless, for the teaching of classical thermodynamics, the justifications given by the aforementioned authors in the demonstration of this theorem are, as they propose to be, sufficient, and do not impair the following of the physical results of the theory [14].

As is already clear, overcoming these two difficulties pointed out by Zemansky to the Carathéodory's formalism was not enough to make this formalism spread later in physics. But then, were there any other major difficulties, besides those originally indicated by Zemansky, with the use of the Carathéodory's formalism? And in particular in the context of the presentation of the second law of thermodynamics? We argue with this paper that no. And so, we aim to introduce some of the methods and meanings of the Carathéodory's formalism to the reader, specifically in the context of the presentation of the second law of thermodynamics.

Therefore, in the following sections, we seek to didactically introduce the reader to the efforts we have cited related to connecting the Carathéodory's axiom with the Clausius principle and the Kelvin principle. As a novel result from the extensive literature reviewed in this paper, we prove the *direct deduction of the Carathéodory's axiom from the Clausius principle*. The section 3 deals with the deduction of the Carathéodory's

axiom from Kelvin principle, according to Titulaer and Van Kampen [22]. The section 4, on the other hand, deals with our deduction of the Carathéodory's axiom from the Clausius principle. We consider these steps to be the most important to the reader with regard to the connection of Carathéodory's formalism with the second law of thermodynamics.

Next, seeking to contribute in some measure to the popularization of Carathéodory's formalism in classical thermodynamics courses at the undergraduate level, we show to the reader in the section 5 a glimpse of how the *entropy* and the mathematical content of the *second law of thermodynamics* arise as a direct consequence of the application of Carathéodory's theorem. To this end, Carathéodory's theorem was demonstrated in the subsection 5.1 in a simple way, from Born's argument [17]. The *entropy* itself, on its turn, was covered soon after, in subsection 5.2.

Also, for a good understanding of what follows, a basic knowledge of the fundamental concepts of classical thermodynamics is sufficient for the reader: *system and neighborhood; thermal reservoir; equilibrium and thermodynamic space; coordinates, state and thermodynamic processes; first and second law of thermodynamics, etc.* However, whenever necessary, for the sake of emphasis, we will briefly discuss some of these concepts.

Finally, this paper is not intended to defend the superiority of the Carathéodory's formalism with respect to the other formalisms of classical thermodynamics. Hence, we only try here to indicate the possibility of using the Carathéodory's formalism in the teaching of classical thermodynamics.

3. Carathéodory's axiom from Kelvin's principle

There are some differences in the literature regarding the writing of the experimental principles of the second law of thermodynamics. Thus, even if merely related to a slightly different choice of words by each author, these differences can cause confusion in the interpretation of the statements of these principles [25]. In an effort to avoid such situations, this paper will state both the Carathéodory's axiom and the Clausius and Kelvin principles as found in the classic book *An Introduction to the Study of Stellar Structure* by the 1983 Nobel Prize in Physics winner Subrahmanyan Chandrasekhar.

It then follows Kelvin's principle, according to Chandrasekhar [13]:

In a cycle of processes it is impossible to transfer heat from a heat reservoir and con-

vert it all into work, without at the same time transferring a certain amount of heat from a hotter to a colder body.

([13] — p. 24)

What Kelvin's principle – which will be referred to hereafter for short as (K) – says is that, during any thermodynamic cycle, it is not possible for a system to fully convert heat Q absorbed from a thermal reservoir into work W without, during the same cycle, there also being heat given off from the system to another system at a lower temperature. In other words, let Q be the heat absorbed by a system from a thermal reservoir during any thermodynamic cycle, and let W be the work related to the interaction of the system with its neighborhood during that cycle. Then, by (K), the following equality at the end of the cycle is impossible, with $Q > 0$

$$Q = W. \quad (1)$$

Observe that (K) prominently works with the concepts of heat, work, and temperature, in an *a priori* fashion. That is, in this statement of the second law of thermodynamics, heat, work, and temperature are elementary thermodynamic concepts.

Understood (K), the Carathéodory's axiom follows, also according to Chandrasekhar [13]:

Arbitrarily near to any given state there exist states which cannot be reached from an initial state by means of adiabatic processes.

([13] — p. 24)

There is much to comment on the Carathéodory's axiom – which will be referred to hereafter, for short, as (AC) – but initially it is necessary to understand the word *state* in this context. Classical thermodynamics deals with macroscopic systems in *equilibrium situations*, that is, in situations where the coordinates, or variables, of the system are well defined. Indeed, when *equilibrium* is established, analogously to the case of classical mechanics, we have the complete characterization of the thermodynamic system by the value of its independent thermodynamic coordinates. And, when this is done, the state of the system is defined and is expressed by the set of these independent thermodynamic coordinates that characterize the equilibrium. So, for classical thermodynamics, *situations of equilibrium are equivalent to defining the state of the thermodynamic system.*

That said, let's get to the content of (AC). What (AC) states is that, given any particular state of a thermodynamic system, there will be other states that the

system cannot reach through adiabatic processes. Adiabatic processes are processes that occur without energy exchange in the form of heat between the system and its surroundings. Note that (AC) makes no distinction between reversible or irreversible processes. Notice here that (AC) assumes as a major elementary thermodynamic concept the notion of an adiabatic process. This construction is present in Carathéodory's formalism in order to get away from the direct concept of heat flow, exchanging it for the idea of *processes not purely mechanical*⁴.

But, to really understand the statement of (AC) we need to analyze the meaning of the term "reach" in it. To get from one state to another, the thermodynamic system needs to perform a process and then reach a new equilibrium situation, thus configuring a new thermodynamic state. To say that there are states that cannot be "achieved" by adiabatic processes means to say that there are equilibrium situations that cannot be achieved if we use an adiabatic process as a way to do so. Put another way, *given any state of a thermodynamic system, we cannot subject the system to arbitrary adiabatic processes.*

We now seek to show that (K) \Rightarrow (AC). To do this we need only show that if (AC) is false, then (K) is also⁵. Then, according to Titulaer and Van Kampen [22], we first prove that (K) \Rightarrow (AC) for *reversible processes*. Reversible processes are, as the name suggests, processes that are amenable to being executed in both "time directions", since all situations of the system during these processes are equilibrium situations. For example, if we can conceive of a system consisting of a gas confined in a container bounded by a frictionless moving piston, we can lower the piston by depositing grains of sand one by one onto it, thereby reversibly compressing the gas in the container, so that the gas will always remain in equilibrium in this process. On the other hand, if we remove the grains of sand from the piston one by one, we may, at some point, return to the exact same equilibrium situation that we initiated this reasoning, returning the gas to its original thermodynamic state. This reversible temporal behavior defines reversible processes. Therefore, reversible processes are represented by continuous curves in thermodynamic space.

Thus, let be a thermodynamic space defining the thermodynamic states for a model thermodynamic system with a usual set of three independent thermodynamic coordinates: θ , the *empirical temperature* of the system, measured by some measuring instrument, on some temperature scale; x_1 and x_2 , two coordinates related to the mechanical behavior of the thermodynamic

system. For example, $x_1 = V$ and $x_2 = M$ being respectively the *volume* and *magnitude of magnetization* of the system. The choice of three independent coordinates for the characterization of the model thermodynamic system under discussion is simply because of the initial facility of working in three dimensions. However, naturally, what will be argued below also holds for a larger number of thermodynamic coordinates.

This thermodynamic space, represented in Fig. 1, characterizes any thermodynamic state by the measured values of the coordinates (θ, x_1, x_2) , for example: (θ^0, x_1^0, x_2^0) . Consequently, such a state is uniquely identified as a point A in said space such that, $A = (\theta^0, x_1^0, x_2^0)$.

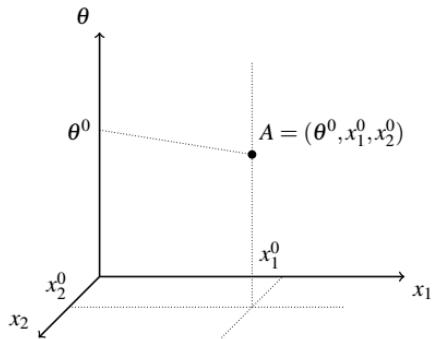


Figure 1: Thermodynamic space of coordinates θ , x_1 and x_2 .

Moreover, because we are dealing with *thermodynamic systems*, it is always possible to choose the *empirical temperature* as one of the independent thermodynamic coordinates. Next, let \mathcal{P} be the following *reversible cycle* given in the thermodynamic space of Fig. 1 and represented in Fig. 2.

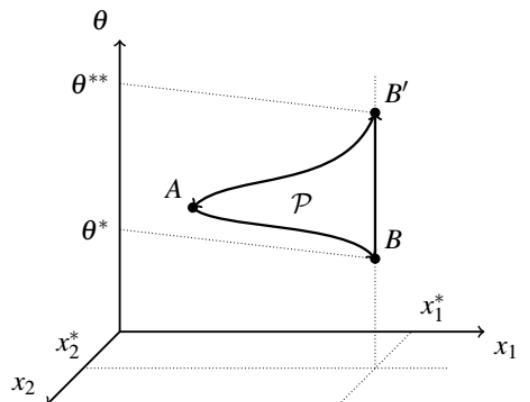


Figure 2: Reversible cycle \mathcal{P} constructed by intermediate processes between points A , B and B' in thermodynamic space. Since \mathcal{P} is reversible, the direction of travel in the cycle can be chosen arbitrarily. Here the direction $A \rightarrow B \rightarrow B' \rightarrow A$ is chosen.

The *reversible cycle* \mathcal{P} of Fig. 2 is set up such that: the process $A \rightarrow B$ is assumed to be adiabatic, so

⁴For an extensive treatment of this question, we suggest the reader consult section 8 of Buchdahl's book [14].

⁵In other words, if it is true that the negation of (AC) provides the negation of (K), then since (K) is true, (AC) will also be so.

$\mathcal{Q}_{A \rightarrow B} = 0$; the process $B \rightarrow B'$ does not escape the line that preserves the values of $x_1 = x_1^*$ and $x_2 = x_2^*$, so there is no work involved from B to B' , however there is heat absorbed $\mathcal{Q}_{B \rightarrow B'} > 0$ from B to B' due to the temperature difference between B and B' , $\theta^{**} - \theta^* > 0$; finally, the process $B' \rightarrow A$ is also supposed to be adiabatic, so $\mathcal{Q}_{B' \rightarrow A} = 0$. This closes the \mathcal{P} cycle. For this cycle we have constructed it is important to realize that since \mathcal{P} is reversible, this whole idealized construction for \mathcal{P} could be reversed by inverting the cycle and taking $B' \rightarrow B$ in reverse such that, in this sense, there would be heat given up $\mathcal{Q}_{B' \rightarrow B} < 0$ from B' to B .

However, a more careful look at the \mathcal{P} cycle we have constructed reveals that if it is possible, then (AC) is false. Indeed, one can approximate B' to B arbitrarily. Moreover, since \mathcal{P} is reversible, both the adiabatic process $A \rightarrow B$ and the adiabatic process $A \rightarrow B'$ can be realized. But if B' is approximated arbitrarily from B , and from A both B and B' can be reached by reversible adiabatic processes, then from A all states on the same line $x_1^* x_2^*$ of B and B' can be reached by reversible adiabatic processes. Hence, also approximating the line $x_1^* x_2^*$ arbitrarily close to A , we would have that: *arbitrarily close to A there are states that can be reached from A by reversible adiabatic processes, thus falsifying (AC) for reversible processes.*

But since we have so far no real – physical – justifications for the validity of (AC), let us suppose that \mathcal{P} is indeed possible and then (AC) is indeed false. Let us now apply the first law of thermodynamics to \mathcal{P} . This gives us, thanks to the additivity of energy

$$\Delta E_{\mathcal{P}} = \Delta E_{A \rightarrow B} + \Delta E_{B \rightarrow B'} + \Delta E_{B' \rightarrow A}. \quad (2)$$

Since \mathcal{P} is a thermodynamic cycle, $\Delta E_{\mathcal{P}} = 0$. Applying the characteristics of \mathcal{P} to (2), we have

$$-W_{A \rightarrow B} + \mathcal{Q}_{B \rightarrow B'} - W_{B' \rightarrow A} = 0. \quad (3)$$

Note that the effective algebraic contributions of the quantities of work that appear in \mathcal{P} are related to the realization of work of the system in the neighborhood, or of the neighborhood on the system. Then, the expression (3) can be rearranged. Naming W the *effective work* involved in the path of \mathcal{P} , we get

$$\mathcal{Q} = W. \quad (4)$$

But this is analogous to what equation (1) says, and so equation (4), which comes from the assumption that (AC) is false, falsifies (K). Hence, for *reversible* processes $(K) \Rightarrow (AC)$. However, we know that we also deal with *irreversible* processes in classical thermodynamics. These processes are, as the name suggests, processes that can only be realized in a single “time direction”,

since the intermediate situations of the system in an irreversible process are not equilibrium situations. Even though classical thermodynamics is a theory that studies only equilibrium situations, in it a qualitative analysis of irreversible processes is also possible. This is due to the fact that in classical thermodynamics the initial and final situations of irreversible processes *are always equilibrium situations*. Furthermore, it is from irreversible processes that the true meaning of *entropy* and *second law of thermodynamics* emerges.

In terms of thermodynamic space, irreversible processes cannot be represented as the usual continuous curves in this space, unlike reversible processes. Thus, since only the initial and final states of an irreversible process are defined, it is usual to represent it as a dashed line in thermodynamic space, connecting its initial and final states. It is natural that we try to evaluate the relation $(K) \Rightarrow (AC)$ for irreversible processes as well. So let be the *irreversible cycles* \mathcal{P}_1 and \mathcal{P}_2 in the same previous thermodynamic space as Fig. 1, represented, respectively, in Fig. 3 and Fig. 4.

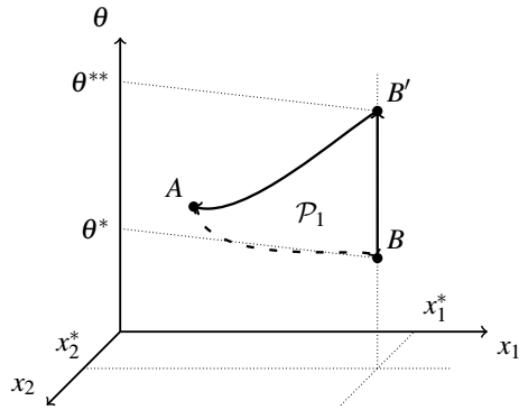


Figure 3: Irreversible cycle \mathcal{P}_1 constructed through intermediate processes between points A , B and B' in thermodynamic space. The intermediate process $A \rightarrow B$ is irreversible. Therefore, \mathcal{P}_1 can only be realized in the direction $A \rightarrow B \rightarrow B' \rightarrow A$.

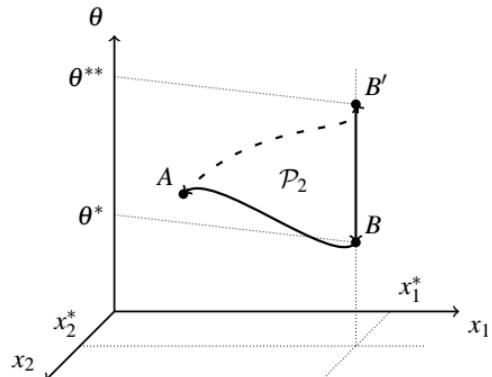


Figure 4: Irreversible cycle \mathcal{P}_2 constructed through intermediate processes between points A , B and B' in thermodynamic space. The intermediate process $A \rightarrow B'$ is irreversible. Therefore, \mathcal{P}_2 can only be realized in the direction $A \rightarrow B' \rightarrow B \rightarrow A$.

The \mathcal{P}_1 and \mathcal{P}_2 cycles are constructed analogously to that posed for the \mathcal{P} cycle in the reversible case. In the cycle \mathcal{P}_1 : the irreversible process $A \rightarrow B$ is assumed to be adiabatic, so $Q_{A \rightarrow B} = 0$; the reversible process $B \rightarrow B'$ holds on the line where $x_1 = x_1^*$ and $x_2 = x_2^*$, so $W_{B \rightarrow B'} = 0$, however $Q_{B \rightarrow B'} > 0$ thanks to the temperature difference between B and B' , $\theta^{**} - \theta^* > 0$; finally, the reversible process $B' \rightarrow A$ is also supposed to be adiabatic, so $Q_{B' \rightarrow A} = 0$. This closes the cycle \mathcal{P}_1 . It should be pointed out that thanks to the *irreversibility* of \mathcal{P}_1 it can only be traversed in the direction $A \rightarrow B \rightarrow B' \rightarrow A$.

Similarly, in the cycle \mathcal{P}_2 : the irreversible process $A \rightarrow B'$ is assumed to be adiabatic, so $Q_{A \rightarrow B'} = 0$; the reversible process $B' \rightarrow B$ holds on the line where $x_1 = x_1^*$ and $x_2 = x_2^*$, so $W_{B' \rightarrow B} = 0$, however $Q_{B' \rightarrow B} < 0$ thanks to the temperature difference between B' and B , $\theta^* - \theta^{**} < 0$; lastly, the reversible process $B \rightarrow A$ is also supposed to be adiabatic, so $Q_{B \rightarrow A} = 0$. This closes the cycle \mathcal{P}_2 . Similarly as for the cycle \mathcal{P}_1 , thanks to the *irreversibility* of \mathcal{P}_2 it can only be traversed in the direction $A \rightarrow B' \rightarrow B \rightarrow A$.

Now suppose that both irreversible cycles \mathcal{P}_1 and \mathcal{P}_2 are simultaneously possible, i.e., both cycles can be performed. It turns out that if this is the case, then (AC) is false. In effect, again, we approximate B' to B in an arbitrary way. Moreover, if we suppose \mathcal{P}_1 and \mathcal{P}_2 to be simultaneously possible, then as a consequence both irreversible processes $A \rightarrow B$ and $A \rightarrow B'$ are also possible. Then, if B' is approximated arbitrarily from B , and from A one can reach both B and B' by irreversible adiabatic processes, then from A all states on the same line $x_1^* x_2^*$ as B and B' can be reached by irreversible adiabatic processes. Hence, also approximating the line $x_1^* x_2^*$ arbitrarily close to A , we would have that: *arbitrarily close to A there are states that can be reached from A by irreversible adiabatic processes, thus falsifying (AC) for irreversible processes.*

At this point, compared to the argument for the reversible case, the attentive reader should have already figured out what the next step is to be. Again, in principle we have no physical argument that prevents (AC) from being false in the irreversible case. So suppose that (AC) is really false and then \mathcal{P}_1 and \mathcal{P}_2 are simultaneously possible. Then, repeating the argument for the reversible case and applying the first law of thermodynamics to both \mathcal{P}_1 and \mathcal{P}_2 , we find that, at the end of each of these cycles

$$Q = W. \quad (5)$$

In the expression (5) Q is the heat absorbed, or ceded, by the system in interaction with an appropriate thermal reservoir for each of the cycles, and W is

the *effective work* related to the interactions of the system with its neighborhood also for each of the cycles. Assuming that \mathcal{P}_1 and \mathcal{P}_2 are simultaneously possible, it is clear that an equality analogous to the expression (5) could be written for each of these irreversible cycles: one in which $W = Q > 0$ at the end of the cycle, related to cycle \mathcal{P}_1 , and one in which $W = Q < 0$ at the end of the cycle, related to cycle \mathcal{P}_2 . The first equality falsifies (K), since it expresses exactly the same content as the expression (1), which is forbidden by (K).

Therefore, for *irreversible* processes $(K) \Rightarrow (AC)$. A direct argument for this conclusion can be found in chapter 5 of Landsberg's book [15]. A few considerations should be made about this result. Note that the true phenomenology behind (K) in connection with (AC) is only revealed from the study of *irreversible* processes. Since, if (K) dealt with the impossibility that at the end of a cycle we have $W = Q < 0$, instead of $W = Q > 0$, nothing would be changed in our analysis of the reversible case to show that $(K) \Rightarrow (AC)$. Removing this apparent mathematical ambiguity from the reversible study of (K) only occurs with the analysis of irreversible processes, revealing the true physical character of (K). To avoid overextending ourselves in this discussion, the argument concerning the reciprocal of this relationship between (AC) and (K) will not be presented here; however, it is short, and can be consulted in Dunning-Davies's paper [23].

4. Carathéodory's axiom from Clausius's principle

We now present our proof of *deduction of Carathéodory's axiom from Clausius's principle*. Equally as was done in the previous section, we establish the Clausius principle as Chandrasekhar [13]:

It is impossible that, at the end of a cycle of changes, heat has been transferred from a colder to a hotter body without at the same time converting a certain amount of work into heat.

([13] — p. 24)

We shall attempt to perform a similar analysis for the Clausius principle — which will be referred to hereafter, in abbreviated form, as (C) — as we did for (K) in the previous section. What (C) says is that, during any thermodynamic cycle, it is not possible for a system to absorb a certain amount of heat Q from a body at a lower temperature than that of the system and then fully transfer that same amount of heat Q to a body at a higher temperature than that of the system, without, during this cycle, there

being the conversion of some amount of work W into additional heat. Here, the mentioned bodies whose system comes into contact in the described cycle are bodies that preserve their respective temperatures when interacting with the system. This makes it implicit that these bodies in contact with the system performing the cycle are thermal reservoirs.

So, in the scheme of the Clausius principle, we have a thermal reservoir with a temperature lower than the temperature of the system, and a thermal reservoir with a temperature higher than the temperature of the system. These studied thermal reservoirs are usually given the suggestive name *thermal sources*; *cold source* for the thermal reservoir under lower temperature than that of the system, and *hot source* for the thermal reservoir under higher temperature than that of the system. That is, suppose a system describes a thermodynamic cycle that absorbs a certain amount of heat \mathcal{Q}_c from a cold source, and then rejects another certain amount of heat \mathcal{Q}_h to a hot source, without there being in that cycle any realization of *effective work* to be converted into additional heat. By (C), the following equality at the end of the cycle is impossible

$$|\mathcal{Q}_c| = |\mathcal{Q}_h|. \quad (6)$$

That is, at the end of the cycle we cannot have equality between the magnitudes of the quantities of heat that were absorbed and rejected, respectively, from the cold source, and to the hot source. In (6) we must write the modulus of the quantities of heat in the cycle, because, as a function of the interaction with the system, heat rejected to a source is, of course, algebraically negative.

Thus, already familiar with the content of (AC), we wish to show that $(C) \Rightarrow (AC)$. As before, for this it is sufficient for us to show that if (AC) is false, so is (C). We will show a proof for this relation first for *reversible* processes. So, let \mathcal{P}' be first the *reversible cycle* depicted in Fig. 5 and given in the same thermodynamic space that we are already used to working in, for the same model thermodynamic system used earlier. Notice that \mathcal{P}' runs through the cycle of points $B \rightarrow C \rightarrow D \rightarrow A \rightarrow B$ in thermodynamic space.

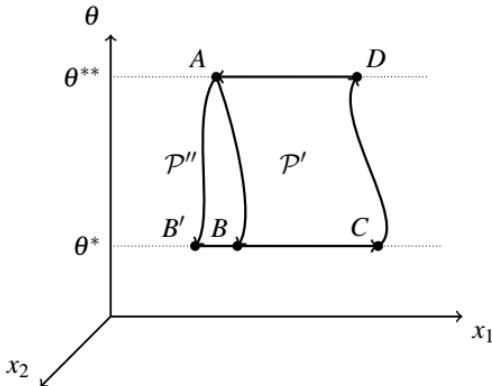


Figure 5: Reversible cycle \mathcal{P}' constructed by intermediate processes between points A, B, C and D in thermodynamic space. Since \mathcal{P}' is reversible, the direction of travel in the cycle can be chosen arbitrarily. Here the direction $B \rightarrow C \rightarrow D \rightarrow A \rightarrow B$ is chosen. The stretch $A \rightarrow B' \rightarrow B$ and the cycle \mathcal{P}'' will be discussed later.

We construct the reversible cycle \mathcal{P}' from Fig. 5, by going through the cycle of points $B \rightarrow C \rightarrow D \rightarrow A \rightarrow B$, so that the system describes, during \mathcal{P}' : the process $B \rightarrow C$, whose temperature θ^* remains constant during the contact of the system with the neighborhood, but there is the absorption of a certain amount of heat $\mathcal{Q}_{B \rightarrow C} > 0$ of the system from the neighborhood; the process $C \rightarrow D$, which is assumed to be adiabatic, hence $\mathcal{Q}_{C \rightarrow D} = 0$; the process $D \rightarrow A$, whose temperature θ^{**} remains constant during the contact of the system with the neighborhood, but there is the rejection of a certain amount of heat $\mathcal{Q}_{D \rightarrow A} < 0$ from the system to the neighborhood; finally, the process $A \rightarrow B$, which is also supposed to be adiabatic, so $\mathcal{Q}_{A \rightarrow B} = 0$. This closes the cycle \mathcal{P}' .

Here, some important considerations should be noted: (i) in principle nothing prevents a cycle like \mathcal{P}' from being constructed, (ii) in general during \mathcal{P}' we have $|\mathcal{Q}_{B \rightarrow C}| \neq |\mathcal{Q}_{D \rightarrow A}|$, with *effective work* being converted into additional heat, iii) the intermediate processes $B \rightarrow C$ and $D \rightarrow A$ of \mathcal{P}' , by preserving the temperature of the system in contact with its neighborhood, indicate that during these processes the system is in contact with *thermal sources*, where naturally the *cold source* is that at temperature θ^* and the *hot source* is that at temperature θ^{**} , and iv) since \mathcal{P}' is reversible, this whole idealized construction for \mathcal{P}' could be reversed by reversing the cycle and taking in reverse an absorption of heat from the hot source, and a rejection of heat to the cold source.

Next, we take note of the cycle \mathcal{P}'' , which can also be seen in Fig. 5 and runs through the cycle of points $B' \rightarrow B \rightarrow C \rightarrow D \rightarrow A \rightarrow B'$ in thermodynamic space. In \mathcal{P}'' the point B' has been chosen such that we have $|\mathcal{Q}_{B' \rightarrow C}| = |\mathcal{Q}_{D \rightarrow A}|$ during the realization of \mathcal{P}'' , with the intermediate process $A \rightarrow B'$ also assumed to be adiabatic.

Given the unrestricted possibility of the occurrence of \mathcal{P}' , if it is also true that \mathcal{P}'' is possible along the lines of what has been constructed, then it means that the adiabatic processes $A \rightarrow B'$ and $A \rightarrow B$ are simultaneously possible. However, if $A \rightarrow B'$ and $A \rightarrow B$ are simultaneously possible, then (AC) is false. In fact, we can approximate B' from B arbitrarily. And, if we assume \mathcal{P}' possible, the reversible adiabatic processes $A \rightarrow B'$ and $A \rightarrow B$ become simultaneously possible. If we also arbitrarily approximate the line of points in thermodynamic space whose temperature is θ^* , to the line of points whose temperature is θ^{**} , we would have that: ar-

bitrarily close to A there are states that can be reached from A by reversible adiabatic processes, thus falsifying (AC) for reversible processes.

However, we can see that if we assume the falsity of (AC), the execution of \mathcal{P}'' immediately gives us the falsity of (C). Indeed, if (AC) is false then $A \rightarrow B'$ and $A \rightarrow B$ are simultaneously possible, in particular \mathcal{P}'' is possible, and as a consequence during \mathcal{P}''

$$|\mathcal{Q}_{B' \rightarrow C}| = |\mathcal{Q}_{D \rightarrow A}|. \quad (7)$$

Which provides, by the expression (7), the same as the expression (6), which is forbidden by (C). Hence, for *reversible* processes $(C) \Rightarrow (AC)$. One would naturally expect $(C) \Rightarrow (AC)$ also for the *irreversible* case. In fact, an argumentation as to the validity of the relation $(C) \Rightarrow (AC)$ for *irreversible* processes follows naturally from what has already been shown with the analogous analysis that $(K) \Rightarrow (AC)$ for the *irreversible* case. For this argument would require the construction of two *irreversible* cycles similar to \mathcal{P}' and \mathcal{P}'' , such that in that analogous to \mathcal{P}' the process $A \rightarrow B$ would be assumed to be adiabatic and irreversible, and in that analogous to \mathcal{P}'' the process $A \rightarrow B'$ would be assumed to be adiabatic and irreversible. The other processes in these cycles would be reversible. We next would take the assumption that these irreversible constructed cycles are simultaneously possible.

So, repeating the same analysis and verifying conclusions similar to those obtained in the *irreversible* case of the $(K) \Rightarrow (AC)$ relation, we would show that, in order not to violate (C), (AC) is also true for *irreversible* processes. To avoid repeating these same steps and the same arguments made before, which would make the discussion here unnecessarily dull, this step will not be developed in the present paper.

5. Road to entropy

Now, armed with the validity of the Carathéodory's axiom, deduced from the principles of Clausius and Kelvin in the previous sections, and the content of the *Carathéodory's theorem*, which we shall see next, we shall show how to obtain the *entropy* and the mathematical content of the *second law of thermodynamics* from the Carathéodory's formalism. With this goal in mind, we will first need to talk a bit about some formal aspects of Carathéodory's formalism. Seeking to combine detail, fluidity, and didactic character in the present section, we divide it into two subsections: one, 5.1, to deal with the mathematics itself and the theorem used in Carathéodory's formalism, and another, 5.2, to deal with

a possible path⁶ possible for obtaining the *entropy* and the mathematical content of the *second law of thermodynamics* by this formalism.

5.1. Mathematical Requirements and Carathéodory's Theorem

The first of the formal aspects of Carathéodory's formalism that we must study is the mathematical interpretation that this formalism gives to the *thermodynamic coordinates*. Quantities like the *empirical temperature* θ , the *volume* V , among others, which we usually call *thermodynamic coordinates* when they characterize the state of a thermodynamic system, appear in Carathéodory's formalism in a formal perspective that draws a great parallel with the concept of the generalized coordinates of classical mechanics [14], which characterize a mechanical system.

Other quantities related to a thermodynamic system, such as *heat* \mathcal{Q} , *work* W , and the *energy* E , are identified in Carathéodory's formalism as a kind of *generalized function* of the thermodynamic coordinates. Mathematically, the thermodynamic coordinates are expressed as quantities x_i , where the index i varies according to the number of thermodynamic coordinates under analysis. For the generalized functions of the thermodynamic coordinates we have announced, on the other hand, we write that they are functions $\chi_i = \chi_i(x_j)$, where the index j tells us that the χ_i are not necessarily functions that depend on all the thermodynamic coordinates considered.

And it is precisely because of this fact that the χ_i are not always *state functions* in the sense of characterizing the state of a thermodynamic system. Since they do not necessarily contain a dependence with all the thermodynamic coordinates that define that state. An example of a function χ_i that is actually a state function is the energy E of a thermodynamic system, since it has dependence with all the thermodynamic coordinates that define the state of that system. An example of a χ_i function that is not a state function is the heat \mathcal{Q} , which is related to the interactions of the thermodynamic system with its neighborhood and thus has no dependence with all the thermodynamic coordinates that define the state of the system.

Already notice here the physical distinction that this formalism provides by telling us mathematically that energy is a function of state of the system and therefore is directly linked to the characterization of its state, while heat is not. In other words, the introduction of these ideas already makes it clear that a thermodynamic system can possess energy, but it cannot possess heat, since the latter depends on the thermodynamic process carried

⁶There are other possible paths for this [7, 14, 15, 16].

out. Moving on in this discussion, what we often see in classical thermodynamics is

$$\delta\chi^* = \sum_{i=1}^n \chi_i(x_j) dx_i. \quad (8)$$

Careful analysis of the expression (8) is crucial to what follows⁷. What (8) says is that the sum of the product between the generalized functions χ_i and the infinitesimals of the thermodynamic coordinates gives us an infinitesimal of some particular generalized function χ^* . That is, in classical thermodynamics the infinitesimals of χ_i are given by expressions analogous to the one in (8). Turning to the symbols used in (8), both δ and d refer to infinitesimal quantities. But, as usual in physics, for an infinitesimal quantity that characterizes an *exact differential*⁸ we reserve the symbol d . Otherwise, we give the symbol δ for an infinitesimal quantity that is not necessarily an exact differential, and we call it an *inexact differential*.

If the quantity $\delta\chi^*$ is an exact differential, we replace the symbol δ with the usual d and have that the generalized function χ^* is actually a state function, having dependence with all thermodynamic coordinates x_i . This case makes the expression (8) more familiar, when we have

$$d\chi^* = \sum_{i=1}^n \frac{\partial \chi^*}{\partial x_i} dx_i. \quad (9)$$

That is, in this case the χ^i are the partial derivatives of χ^* with respect to the thermodynamic coordinates x_i and the state function χ^* can be obtained via a ordinary integration⁹ from the expression (8). Again, exemplifying, and from what has been previously discussed, we have as an immediate example of a thermodynamic quantity that defines an exact differential, energy, and as one that doesn't, heat. A test that we can always apply to check whether the quantity $\delta\chi^*$ is an exact differential or not is to evaluate, in the generalized functions χ_i of $\delta\chi^*$ in (8), whether

$$\frac{\partial \chi_k}{\partial x_l} = \frac{\partial \chi_l}{\partial x_k} \quad \text{with } (k, l = 1, 2, \dots, n). \quad (10)$$

The test of equation (10) evaluates whether the crossed partial derivatives of any pairs (χ_k, χ_l) of the

χ_i , to the corresponding pair of the thermodynamic coordinates (x_l, x_k) with naturally $(k, l = 1, 2, \dots, n)$, are identically equal to each other. The reader may recognize that this test follows from the *Clairaut-Schwarz theorem*, from the *differential calculus of several variables*. If equation (10) is satisfied for any pair of indices, $(k, l = 1, 2, \dots, n)$, the quantity $\delta\chi^*$ will be an exact differential.

Next, we also want to pay special attention to the important situation in which the expression (8) nullifies, that is, when

$$\delta\chi^* = \sum_{i=1}^n \chi_i(x_j) dx_i = 0. \quad (11)$$

For the equation defined in (11) we give here the name¹⁰ of *differential equation associated with $\delta\chi^*$* . It is important for us to note that the solutions of these differential equations associated with $\delta\chi^*$ will always be a set of thermodynamic coordinate values, which can be visualized, in the geometric perspective of thermodynamic space, as a set of points given by these thermodynamic coordinate values in that space. In particular, when $\delta\chi^*$ is an exact differential, that set of points in thermodynamic space forming the solutions of the differential equation associated with $\delta\chi^* = 0$ is, from equation (11)

$$\chi^* = \chi^*(x_1, x_2, \dots, x_n) = c. \quad (12)$$

Where in equation (12) c is a constant. Equation (12) is a *hypersurface of n dimensions* in thermodynamic space with n thermodynamic coordinates. Then, fixed a hypersurface $\chi^*(x_1, x_2, \dots, x_n) = c$, for a given value of c , one naturally establishes the solutions of (12), which are *hypercurves* in this space. At first abstract, this conclusion materializes much more intuitively when we work with three thermodynamic coordinates and then the equation (12) translates into a familiar *surface in 3 dimensions*, $\chi^*(x_1, x_2, x_3) = c$, in the related three-dimensional thermodynamic space. Furthermore, the solutions of $\chi^*(x_1, x_2, x_3) = c$ become *curves* in this space. It is worth noting that it is from three thermodynamic coordinates that we usually model and study most thermodynamic systems from classical thermodynamics at the undergraduate level.

Finally, before we talk about Carathéodory's theorem, we need to deal with when the quantity $\delta\chi^*$ in (8)

⁷An elegant name for the quantities expressed by (8) is that of *pfaffian differential forms* [26]. However, because this is merely a technical term, and would add little to the discussion of the material in this paper, we do not use this nomenclature. For the reader who wants to delve deeper into the rich mathematical theory of pfaffian differential forms, we suggest consulting Sneddon's book [26].

⁸i.e., an infinitesimal quantity whose variation only depends on its initial and final values.

⁹Com $\delta\chi^*$ by defining an exact differential $d\chi^*$, taking its integral need not specify any particular set of values of the thermodynamic coordinates, which would define a path for the integration to be done. In that case, the integration of $d\chi^*$ evaluates only the initial and final values of χ^* .

¹⁰These equations are customarily called *Pfaff differential equations* [26].

does not constitute an exact differential, but becomes one when it is multiplied by a certain function, which we call the *integrating factor*. In fact, analogous to χ_i , let η be any generalized function of thermodynamic coordinates which at first is not an exact differential, so the infinitesimal of η is $d\eta$. It turns out that in some cases, when we multiply $d\eta$ by some other generalized thermodynamic coordinate function μ , we get the validity of the following relation

$$d\sigma = \mu d\eta. \quad (13)$$

That is, in some cases, we can multiply an inexact differential $d\eta$ by an appropriate generalized function of the thermodynamic coordinates μ , so that we get an exact differential $d\sigma$. When this happens, we say that $d\eta$ is an *integrable inexact differential* by μ , hence the suggestive name *integrating factor* for μ . This action, given in (13), is also called *integrate the equation* differential associated with $d\eta$. And $d\eta$ being a *integrable inexact differential*, the solutions of its respective associated differential equation are also of the form of the expression (12). But in which cases can we do this *integration*? Except for the cases restricted to two or three thermodynamic coordinates in study¹¹, the result that generalizes the answer to this question, and gives physical meaning to all this previous mathematical construction, is the *Carathéodory's theorem*.

It then follows Carathéodory's theorem, adapted to the notation of the present paper, from what appears in H. A. Buchdahl's [14] book of classical thermodynamics *The Concepts of Classical Thermodynamics*:

If every neighbourhood of any arbitrary point A contains points B inaccessible from A along solution curves of the equation $\sum_{i=1}^n \chi_i(x_j)dx_i = 0$, then the equation is integrable.

([14] — p. 62)

Before we prove Carathéodory's theorem, let's make evident to the reader the familiar physical content it has from Carathéodory's axiom, which we have already discussed extensively in the 3 and 4 sections. First, let's recapitulate that the heat \mathcal{Q} related to a thermodynamic system is here mathematically identified as a generalized function of thermodynamic coordinates according to (8), so for n any x_i thermodynamic coordinates, we have

$$\delta\mathcal{Q} = \sum_{i=1}^n \mathcal{Q}_i(x_j)dx_i. \quad (14)$$

Naturally, with the corresponding *differential equation associated with \mathcal{Q}* , given by

$$\delta\mathcal{Q} = \sum_{i=1}^n \mathcal{Q}_i(x_j)dx_i = 0. \quad (15)$$

Note that equation (15) already gives us, precisely, the mathematical description of a *infinitesimal adiabatic process*. Now, if we translate and apply Carathéodory's theorem to \mathcal{Q} , and to points in the thermodynamic space of any n thermodynamic coordinates we want to study, we can write that: *if any neighborhood of any point A contains points B that are inaccessible by A from the solution curves of the equation $\delta\mathcal{Q} = 0$, then this equation is integrable*.

But, we note here a congruence between the premise of Carathéodory's theorem applied to \mathcal{Q} , and Carathéodory's axiom, which we remind the reader again, already translating it to the context of thermodynamic space¹²:

Carathéodory's axiom. *arbitrarily close to any given point there are points that are inaccessible from an initial point by means of adiabatic processes.*

And, as we also recall, adiabatic processes are processes in which the equation (15) occurs throughout its execution. Also, realize that an adiabatic process is exactly the physical concept that is mathematically represented in this formalism by means of the solution curves of the equation (15). In other words, *adiabatic processes are the solution curves of the equation (15)*. Note the elegant distinction, and at the same time connection, that Carathéodory's formalism makes between the physical substance and the mathematical substance of classical thermodynamics.

Thus, we rewrite Carathéodory's theorem as follows: *if arbitrarily close to any given point there are points inaccessible from an initial point by adiabatic processes, then the equation $\delta\mathcal{Q} = 0$ is integrable*. Or, to summarize:

Carathéodory's theorem. *if Carathéodory's axiom holds, then $\delta\mathcal{Q} = 0$ is integrable.*

Carathéodory's theorem speaks in mathematical terms when an equation of type (11) is integrable, and Carathéodory's axiom points out in physical terms that heat always is. For mere initial simplicity, we prove Carathéodory's theorem for any three thermodynamic coordinates (x_1, x_2, x_3) . As we had announced earlier, we follow the arguments of Born [17].

¹¹Suggest reading Buchdahl's book [14] for the reader who wishes to study these simpler cases in depth.

¹²Remember, we have already justified the univocal relation between thermodynamic states and points in thermodynamic space.

Proof. Consider that, as our initial hypothesis, the Carathéodory's axiom holds. That is, consider that an arbitrary point B in 3-dimensional thermodynamic space is inaccessible from an also arbitrary point A, however close B and A are, by the solutions of $\delta\mathcal{Q} = 0$ in that space. Let us now say that there is a second arbitrary point C accessible to A by $\delta\mathcal{Q} = 0$. Then A and C are accessible to each other by the solutions of $\delta\mathcal{Q} = 0$. But C must also be inaccessible to B by these solutions of $\delta\mathcal{Q} = 0$, because otherwise, through passing through C, B would be accessible to A, which contradicts our initial hypothesis. Therefore, the points accessible to A define a surface in the thermodynamic space containing A such that this surface also contains all the points accessible to A by the solutions of $\delta\mathcal{Q} = 0$. Now, since this property of inaccessibility of points in space was exemplified by A, with arbitrary A, it must also hold for any other point in that thermodynamic space. Thus defining, for each point chosen, and by the fact that there are always points arbitrarily close to the point one chooses, a surface in three-dimensional space, $\sigma(x_1, x_2, x_3) = c$, which contains all the respective points accessible from the arbitrary point chosen. There are thus, as a consequence of our inaccessibility hypothesis, and of there being inaccessible points arbitrarily close to the point that is chosen, several neighboring surfaces that do not intersect¹³, $\Sigma(x_1, x_2, x_3) = c$, containing the points that are accessible to each other. On these surfaces we must have $d\sigma = 0$ and $\delta\mathcal{Q} = 0$, from which we conclude that $\delta\mathcal{Q}$ and $d\sigma$ must be proportional quantities on these surfaces. That is, there exists $\mu = \mu(x_1, x_2, x_3)$ such that

$$d\sigma = \mu \delta\mathcal{Q}. \quad (16)$$

■

For the general case of n thermodynamic coordinates the above argument is the same, exchanging only the term and the construction of *surfaces* for the generalization of *hypersurfaces*. It follows from the above proof that writing μ as a quantity that relates to the differentials $d\sigma$ and $\delta\mathcal{Q}$ exactly as in the form expressed in equation (16) is not mandatory. In other words, according to Carathéodory's theorem, mathematically we only need μ to express the proportionality that exists between $d\sigma$ and $\delta\mathcal{Q}$. Thus, for reasons that will become clear later, we will write for $\delta\mathcal{Q}$ and $d\sigma$, instead of the equation (16), the following expression

$$\delta\mathcal{Q} = \mu d\sigma. \quad (17)$$

Note at this point one of the striking features of

Carathéodory's formalism: ambiguities and general mathematical conclusions will necessarily be subject to physical evaluations of their meaning. Let us now, in order to better understand the equality (17) and obtain the *entropy* and the mathematical content of the *second law of thermodynamics* by the formalism developed here, seek to better study the integral factor μ and the state function σ that we have just discovered.

5.2. Entropy and absolute temperature

In order to obtain the *entropy* and the mathematical content of the *second law of thermodynamics* by the Carathéodory formalism, we must analyze the situation of two thermodynamic systems in purely thermal contact with each other, as well as adiabatically isolated from the surrounding neighborhood. And at this point, we shall again deal with any n number of thermodynamic coordinates for the thermodynamic systems under study, assuming the validity of Carathéodory's theorem for this general case as well. Consider then two thermodynamic systems, which we will refer to as K_A and K_B , in purely thermal contact with each other, i.e., K_A and K_B can interact only via heat exchange. Furthermore, as we said, K_A and K_B are adiabatically isolated from their surrounding neighborhood.

If we look at the set K_A and K_B globally, we can say that both systems form a compound thermodynamic system K_C . Looking at this composite system K_C , in mathematical terms, we have

$$\delta\mathcal{Q}_C = \delta\mathcal{Q}_A + \delta\mathcal{Q}_B. \quad (18)$$

When thermal equilibrium is established between K_A and K_B , both individual systems will have the same empirical temperature θ at the end of a given time¹⁴, which will also naturally be the same as the empirical temperature of the composite system K_C at equilibrium. Now, being $(x_1, x_2, \dots, x_{n-1}, \theta_A)$, $(y_1, y_2, \dots, y_{n-1}, \theta_B)$, e $(x_1, x_2, \dots, x_{n-1}, y_1, y_2, \dots, y_{n-1}, \theta_A, \theta_B)$, the respective n thermodynamic coordinates of K_A , K_B and K_C , we will have, at thermal equilibrium, $\theta_A = \theta_B = \theta$. The coordinates x_i and y_i are the thermodynamic coordinates that provide the mechanical behavior of the respective systems K_A and K_B . These findings will be important later on. By applying equation (17) to equation (18), we immediately have

$$\mu_C d\sigma_C = \mu_A d\sigma_A + \mu_B d\sigma_B. \quad (19)$$

Rearranging (19) in terms of $d\sigma_C$, we obtain

¹³This verification is simple, as Landsberg discusses [15].

¹⁴In general, the time for thermodynamic equilibrium to occur during any thermodynamic interaction – mechanical, chemical, etc. – between thermodynamic systems is called the relaxation time. A beautiful discussion of this concept in its macroscopic aspect can be found in Callen's book [4].

$$d\sigma_C = \frac{\mu_A}{\mu_C} d\sigma_A + \frac{\mu_B}{\mu_C} d\sigma_B. \quad (20)$$

As justified by Carathéodory's theorem, the σ quantities that arise from the integration of $\delta\mathcal{Q}$ quantities are state functions, and thus are functions of all thermodynamic coordinates that characterize the state of a thermodynamic system. From the expression (20), we also see that $\sigma_C = \sigma_C(\sigma_A, \sigma_B)$. That is, we can write

$$d\sigma_C = \frac{\partial \sigma_C}{\partial \sigma_A} d\sigma_A + \frac{\partial \sigma_C}{\partial \sigma_B} d\sigma_B. \quad (21)$$

So, comparing the expressions (21) and (20), we have

$$\frac{\mu_A}{\mu_C} = \frac{\partial \sigma_C}{\partial \sigma_A}; \quad (22a)$$

$$\frac{\mu_B}{\mu_C} = \frac{\partial \sigma_C}{\partial \sigma_B}. \quad (22b)$$

The equations (22) tell us that the quotients $\frac{\mu_A}{\mu_C}$ and $\frac{\mu_B}{\mu_C}$ are such that

$$\frac{\mu_A}{\mu_C} = \frac{\mu_A}{\mu_C}(\sigma_A, \sigma_B); \quad (23a)$$

$$\frac{\mu_B}{\mu_C} = \frac{\mu_B}{\mu_C}(\sigma_A, \sigma_B). \quad (23b)$$

But, we know that in principle, by Carathéodory's theorem, the quantities μ are functions of the thermodynamic coordinates defining the states of their respective parent thermodynamic systems, i.e., and already for the thermal equilibrium between K_A and K_B

$$\mu_A = \mu_A(x_1, x_2, \dots, x_{n-1}, \sigma_A, \theta); \quad (24a)$$

$$\mu_B = \mu_B(y_1, y_2, \dots, y_{n-1}, \sigma_B, \theta); \quad (24b)$$

$$\mu_C = \mu_C(x_1, \dots, x_{n-1}, y_1, \dots, y_{n-1}, \sigma_A, \sigma_B, \theta). \quad (24c)$$

Thus, pay attention that to reconcile equations (24) and (23) the quantities μ should not depend on their respective thermodynamic coordinates that provide the mechanical behavior of the respective system related to μ . Otherwise, the quotients given in the equations (23) could not depend on the σ quantities alone. Hence, by this analysis, the most general form of the μ quantities must be¹⁵

$$\mu_A = \mu_A(\sigma_A, \theta) = t(\theta)f_A(\sigma_A); \quad (25a)$$

$$\mu_B = \mu_B(\sigma_B, \theta) = t(\theta)f_B(\sigma_B); \quad (25b)$$

$$\mu_C = \mu_C(\sigma_A, \sigma_B, \theta) = t(\theta)f_C(\sigma_A, \sigma_B). \quad (25c)$$

Note that, only thanks to the thermal equilibrium between the systems K_A and K_B are we able to write the equations (25), which show that the integral factors μ can be written with respect to a function of universal character¹⁶ $t = t(\theta)$, independent of the thermodynamic system, which depends solely on the empirical temperature θ of the equilibrium between K_A and K_B and which is therefore common to both systems K_A and K_B .

The universal content of this function $t = t(\theta)$ motivates us to define the so-called *absolute temperature* T of [28] thermodynamic systems, minus an arbitrary constant k , so that

$$T = kt(\theta). \quad (26)$$

The algebraic sign of the arbitrary constant k defining the absolute temperature in equation (26) is a convention¹⁷ and Carathéodory's formalism makes this clear. The historical choice was $k > 0$. Moreover, it was to obtain this direct proportionality between the temperature scales T and $t(\theta)$ given in (26) that we chose to work with equation (17), instead of equation (16). This suggests for the absolute temperature the same experimental methods as are used for the empirical temperature determination, as discussed by Buchdahl [14].

If we substitute the result of equations (25), together with equation (26), into equation (17), we will have, for any thermodynamic system

$$\delta\mathcal{Q} = \frac{T}{k} f(\sigma) d\sigma. \quad (27)$$

Isolating the terms of (27) with the dependency on σ

$$\frac{\delta\mathcal{Q}}{T} = \frac{1}{k} f(\sigma) d\sigma. \quad (28)$$

We already know that the quantity σ is a state function, as is, for example, the energy E of a thermodynamic system. We then call the σ function *empirical entropy*, in allusion to its relation to the empirical temperature of a thermodynamic system. Next, we define

$$d\mathcal{S} \equiv \frac{1}{k} f(\sigma) d\sigma. \quad (29)$$

The quantity \mathcal{S} is clearly also a state function, and is called the *absolute entropy*, or just the *entropy* of the thermodynamic system, also alluding to its relationship with the absolute temperature. The measurement of entropy is also independent of the particular properties of

¹⁵A more detailed argument for this step can be found in Zemansky's paper [20].

¹⁶We should point out that the Clausius formalism also indicates that temperature is related to the integral factor of heat [27]. In the Gibbs formalism, although also present, this fact is not very relevant.

¹⁷We strongly recommend the reader to consult section 11 and appendices [A-6] and [A-8] of Pauli's book [28], for an enlightening discussion of this subject.

each thermodynamic system. Finally, substituting equation (29) into equation (28), we obtain

$$dS = \frac{\delta Q}{T}. \quad (30)$$

The expression (30) is the mathematical content of the *second law of thermodynamics* for *reversible processes*. Why (30) is only valid for reversible processes becomes clear as we try to evaluate the entropy S from (30) for irreversible processes. If we try to do this, we will immediately see that a simple integration of dS is not possible to obtain dS for irreversible processes. For such an integration, the absolute temperature T in (30) would not even be defined in the intermediate steps of any irreversible process we would consider evaluating. This means that we need another strategy to study S in irreversible processes.

However, fortunately, Carathéodory's axiom provides us with that other strategy quickly. To do this, consider, again for pure simplicity, a thermodynamic space with three thermodynamic coordinates, x_1, x_2 and S . Then consider the schematic of this thermodynamic space in Fig. 6 with two supposedly adiabatic irreversible processes, \mathcal{P}_+ and \mathcal{P}_- , both starting from the same arbitrary point A .

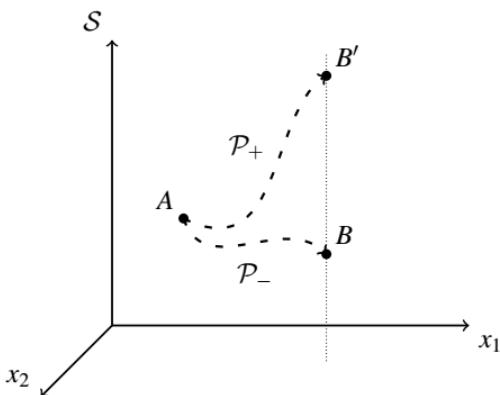


Figure 6: Thermodynamic space of the coordinates x_1, x_2 and S , with the assumed irreversible adiabatic processes \mathcal{P}_+ of $A \rightarrow B'$, and \mathcal{P}_- of $A \rightarrow B$.

That done, suppose both of these processes, \mathcal{P}_+ and \mathcal{P}_- , simultaneously possible. Then, the entropy variation in the course of these two processes could be, from Fig. 6, either positive or negative. Positive during \mathcal{P}_+ , negative during \mathcal{P}_- . But, we can arbitrarily approximate B' from B , and since we assume \mathcal{P}_+ and \mathcal{P}_- simultaneously possible, it follows that all points on the line containing B and B' can be reached from A by irreversible adiabatic processes. Also approximating the line containing B and B' from A , we will have that: *arbitrarily close to A there are states that can be reached from A by irreversible adiabatic processes, thus falsifying Carathéodory's axiom for irreversible processes.*

The following is derived from this argument: *to not violate Carathéodory's axiom, the entropy of any thermodynamic system in the course of an irreversible adiabatic process must always either increase or decrease, never both* [28]. Again here, and perhaps even more importantly, Carathéodory's formalism impels us to note the distinction between the physical substance of classical thermodynamics and the mathematical construct developed to describe it. That is, the choice for the algebraic sign of the entropy variation in irreversible adiabatic processes becomes *arbitrary*. As the reader may already know, we choose the positive sign for this variation and write

$$\Delta S_{Q=0} > 0. \quad (31)$$

The expression in (30) is the mathematical content of the *second law of thermodynamics* for *irreversible processes*. Putting the expressions (30) and (31) together, when we take an adiabatic process in (30), we get

$$\Delta S_{Q=0} \geq 0. \quad (32)$$

The expression (32) is the general mathematical content of the *second law of thermodynamics*.

Finally, we would like the reader to take the time to observe the essence of Carathéodory's formalism. In it, despite the mathematical investment that is required, the physics of classical thermodynamics is largely reflected and requested, while the mathematical construction accurately distinguishes the physical content of the theory from the purely formal and abstract. Carathéodory's formalism still provides a vast scope for pedagogical connection between classical thermodynamics and classical mechanics, starting with the concept of generalized coordinates, as Buchdahl [14] points out. Also, by being able to reproduce the results of classical thermodynamics simply by assuming the validity of the Carathéodory axiom at first, we draw a parallel between this axiom and the one found in Hamilton's formalism – *Hamilton's principle* –, again in the context of classical mechanics, as Pippard says [29]. In other words, despite the misfortune of the unpopularity of the Carathéodory formalism over the years, this approach can definitely add much to the teaching of classical thermodynamics, and can be used as a viable alternative to the Clausius and Gibbs formalisms.

6. Conclusion

In this paper, we cover some of the construction of the Carathéodory formalism for classical thermodynamics in relation to the other best known formalisms of that

theory. In the section 2 we discuss some of the origin and tenor of the Carathéodory formalism in the context of its historical unpopularity over the years. We advocate for Carathéodory's formalism by fostering it as a viable alternative to teaching classical thermodynamics.

In furtherance of this cause, we seek to didactically show the reader how Carathéodory's axiom of the second law of thermodynamics can be deduced from the principles of Clausius and Kelvin: we show $(K) \Rightarrow (AC)$ in the 3 section, and $(C) \Rightarrow (AC)$ in the 4 section. From the extensive literature reviewed, we have given a new proof for $(C) \Rightarrow (AC)$.

In addition, we guide the reader in the 5 section through one of the possible paths that lead to obtaining the *entropy* and the mathematical content of the *second law of thermodynamics* from this formalism. Thus, we hope that this work will serve in some measure to popularize the Carathéodory formalism in disciplines of classical thermodynamics at the undergraduate level, also contributing to the teaching of classical thermodynamics itself.

Acknowledgments

Thanks are due to Paulo Henrique Ribeiro Peixoto and Gustavo Camelo Neto, professors at UFPE.

References

- [1] I. Müller, *A History of Thermodynamics: The Doctrine of Energy and Entropy* (Springer, Berlin, 2007).
- [2] M. J. Oliveira, *Equilibrium Thermodynamics* (2 ed, Springer, Berlin, 2017).
- [3] J. W. Gibbs, *The Collected Works of J. Willard Gibbs. Volume I: Thermodynamics* (Longmans, London, 1928).
- [4] H. Callen, *Thermodynamics and an Introduction to Thermostatistics* (2 ed, Wiley, New York, 1985).
- [5] E. A. Guggenheim, *Thermodynamics: An Advanced Treatment for Chemists and Physicists* (8 ed, North-holland, Amsterdam, 1986).
- [6] M. J. Oliveira, Rev. Bras. Ensino Fís. **41**, 1, (2019).
- [7] J. H. Luscombe, *Thermodynamics* (CRC Press, Boca Raton, 2018).
- [8] A. Thess, *The Entropy Principle: Thermodynamics for the Unsatisfied* (Springer, New York, 2011).
- [9] E. H. Lieb, J. Yngvason, Physics Reports **310**, 1, p. 1-96 (1999)
- [10] C. Carathéodory, Math. Ann. **67**, p. 355-386 (1909).
- [11] L. Tisza, *Generalized Thermodynamics* (MIT Press, Massachusetts, 1966).
- [12] F .W . Sears, American Journal of Physics **34**, p. 665 (1966).
- [13] S. Chandrasekhar, *An Introduction to the Study of Stellar Structure* (Dover Publications, New York, 1939).
- [14] H. A. Buchdahl, *The Concepts of Classical Thermodynamics* (Cambridge University Press, London, 1966).
- [15] P. T. Landsberg, *Thermodynamics and Statistical Mechanics* (Dover Publications, New York, 2014).
- [16] J. Dunning-Davies, *Concise Thermodynamics: Principles and Applications in Physical Science and Engineering* (2 ed, Woodhead Publishing Limited, Sawston, 2007).
- [17] M. Born, Physik. Z. **22**, 218, 249, 282, (1921).
- [18] M. Born, *Natural Philosophy of Cause and Chance* (Clarendon Press, Oxford, 1949).
- [19] M. Born, A. Einstein e H. Born, *The Born-Einstein Letters* (The MacMillan Press, London, 1971).
- [20] M. W. Zemansky, American Journal of Physics **34**, 10, p. 914-920 (1966).
- [21] P. T. Landsberg, Nature **201**, p. 485-486 (1964).
- [22] U. M. Titulaer, N. G. Van Kampen, Physica **31**, 7, p. 1029-1032 (1965).
- [23] J. Dunning-Davies, Nature **208**, p. 576-577 (1965).
- [24] J. B. Boyling, Comm. Math. Phys. **10**, 1, p. 52-68 (1968).
- [25] J. Dunning-Davies, arXiv:1103.4360 (2011). <https://arxiv.org/abs/1103.4360>. Último acesso em agosto de 2020
- [26] I. Sneddon, *Elements of Partial Differential Equations* (McGraw-Hill, New York, 2006).
- [27] M. W. Zemansky, R. Dittman, *Heat and Thermodynamics* (7 ed, McGraw-Hill, New York, 1997).
- [28] W. Pauli, *Thermodynamics and the Kinetic Theory of Gases: Volume 3 of Pauli Lectures on Physics* (Dover Publications, New York, 1973).
- [29] A. B. Pippard, *The Elements of Classical Thermodynamics* (Cambridge University Press, London, 1966).