ON A CLASS OF THEORIES OF MECHANICS - PART I

Jayme Vaz Jr.

RELATÓRIO TÉCNICO Nº 50/91

Abstract. In this paper we discuss the foundations of Mechanics based on the concept of convolutive sets. We show that there exist an infinity of expressions for the action which satisfy the laws of Mechanics. It follows that it is possible to formulate a general class of theories of Mechanics, of which the Classical Mechanics is only one particular example. We discuss how to formulate these theories, leaving their explicit construction for another paper. Among such theories there is one that is *formally* identical to Quantum Mechanics but does not involve the concept of probability.

Universidade Estadual de Campinas
Instituto de Matemática, Estatística e Ciência da Computação
IMECC - UNICAMP
Caixa Postal 6065
13.081 - Campinas - SP
BRASIL

Setembro - 1991

ON A CLASS OF THEORIES OF MECHANICS – PART I

Jayme Vaz Jr.

Departamento de Matemática Aplicada – IMECC-UNICAMP
Universidade Estadual de Campinas
CP 6065 – 13081 – Campinas, S.P., Brazil

Abstract

In this paper we discuss the foundations of Mechanics based on the concept of convolutive sets. We show that there exist an infinity of expressions for the action which satisfy the laws of Mechanics. It follows that it is possible to formulate a general class of theories of Mechanics, of which the Classical Mechanics is only one particular example. We discuss how to formulate these theories, leaving their explicit construction for another paper. Among such theories there is one that is *formally* identical to Quantum Mechanics but does not involve the concept of probability.

1 Introduction

Mechanics is the part of Physics that studies motion through the formulation of a physical theory. We shall see that this is not just a trivial sentence: indeed, the essence of the present paper is to understand what that definition of Mechanics does mean. The problems we meet are that, in order it to make sense, the concepts of motion and physical theory should already be defined; but they entail *delicate* questions: especially the concept of motion. A satisfactory definition of a physical theory can be given (sec.3), but that of motion deserves a careful discussion. Thus, we can properly say that the *objective* of this paper is to *understand* the *meaning* of:

Ph.D. student in Applied Mathematics

Definition 1 Mechanics is the part of Physics that studies *motion* through the formulation of a physical theory.

The reality of motion has always been an object of intense philosophical inquiry, especially at the time of the Greek philosophers. To quote two well-known examples, there are the views of (a) Heraclitus, according to whom the "kinēsis" fills the entire reality, and (b) Parménides, defended by Zeno through his four arguments, according to whom there is no reality in the motion. Nowadays the interest on this question remains, mainly, for physicists. One reason is the advent of Quantum Mechanics and the debate about its interpretation. Even for those physicists to whom the philosophical questions arising from the reality of motion present no interest, the question "what is motion" has a fundamental importance. In fact, an objective answer to that question has to be given since the study of a not well-defined object is questionable.

Several related topics are discussed in sec.2. Some ideas discussed in a previous work [1] are widely used, mainly the ones of *conjunctive* and *convolutive sets*. Definitions are given in order to disentangle some concepts involved in the notion of motion (particularly those concerning the dynamical variables and the independent variables) and of its spacetime representation. Next, we ask about the nature and the properties of motion (which are the basic questions of Mechanics according to def.1) and search for answers in the laws of Mechanics. The conclusion that there is not a unique answer to these questions (sec.3) is one consequence of our approach. Rather than disappointing, this is indeed a welcome fact and we exploit it accordingly.

In fact, what we do is to search for an expression for the action such that it satisfies our formulation of the law of Mechanics as a kind of variational principle. This is a crucial step and accordingly we do not start with a definition of the action as the time integral of the lagrangian (the action functional) for a given path (more on this issue is discussed in sec.4). Since a price has to be paid in order to get advantages, it is necessary to use some quantities which do not have a standard denomination in the literature. We try to avoid confusions whenever they can arise, by calling attention to the different mathematical nature of the objects we consider.

From our approach it naturally emerges the possibility of formulating an infinity of other theories of Mechanics besides the Classical one: i.e., we have a *class* of theories of Mechanics. These are interpreted according to [1] and once we define conjunctive dynamics and convolutive dynamics we see that Classical Mechanics describes the former while the other theories, called Generalized Mechanics, describe particular cases of the latter (sec.5).

We discuss also in this paper how to formulate this general class of theories of Mechanics, leaving the explicit construction for another paper [2]. Then the real importance of our formulation of the Generalized Mechanics does appear: there exists one theory among these that is formally identical to Quantum Mechanics but does not involve the concept of probability. Then, we discuss the relation between Generalized Mechanics and Quantum Mechanics. Ref.[1] provides the basis for this paper and [2]. These works, and the inspiring one of Bernardes [3], put in evidence the generality of the idea of convolutive sets [1], according to which they are interpreted.

2 What is Motion ? Some Definitions and a Question

The first thing about motion that we naturally have in mind and ask is: motion of what? Roughly speaking, motion is a translation of bodies in space. Therefore, if there is a motion then there exists some kind of system, for example a body, which is in motion. This naive conception of motion could work for some purposes, but physicists cannot be satisfied with it. The problem is not whether it is or not general enough to include other situations we would like to see included (such as propagation of waves, flow of energy, etc.), but that of (dangerously) drawing a too simple picture of motion from this conception.

The relation between the concepts of motion and system can be said to be of the "chicken and egg" type. In fact, the dichotomy between *the motion* and *the system* is not justifiable, except at the linguistic level. We can easily see that one supposes the other, and vice-versa, looking at the following two simple situations: (a) when we say that there exists a motion we believe that there exists a system which is in motion; (b) when we say that there exists a system, for example, a body, we believe that, once we are looking at it, we can close our eyes and at a later time open them and still find this system, and in doing this we suppose the motion in the sense that this body is either at rest, which is a particular state of the motion, or in another state of motion. A reasoning like the latter is used explicitly in particle physics where we infer the existence of particles from tracks in bubble chambers. Bohm [4] has already recognized this "chicken and egg" problem and discuss it through what he calls the holomovement. A general discussion involving the presence of "chicken and egg" problems in scientific topics is provided by Bernardes [5].

On the other hand, even being this dichotomy unjustifiable, it must be respected. One of the consequences of the "scientific analysis" method is the use of a language based on the concepts of *noun* and *conjunction*, i.e., the ordinary language. *Motion* and *system* are nouns, which means that they are entities of autonomous conceptual existence. It must be clear that (a) being different nouns does not mean that they are unrelated, and (b) being different nouns implies that to them there must be ascribed different adjectives (the adjectives qualify the nouns). The use of this terminology (due to Bernardes and already used in [1], and which is, in our opinion, very intuitive and directly related to the linguistic nature of the problem [5]) can be made clear by another simple example: wine and grape are different nouns and have autonomous conceptual existence; to them there is room for different adjectives, for example, a dry wine, a rotten grape, etc., and they are related, since wine is made from the fermented juice of grapes. In other words [1], we are considering the terms *motion* and *system* and not the values that each term has [6]. In any case, what matters is to understand that once (a) there exists a kind of "cyclic" relation between motion and system, and (b) both concepts have autonomous existence, then one of them must be taken as the primitive one. Which one of these concepts is to be taken as the primitive one is somewhat arbitrary, i.e., we think that it just is a matter of choice, since what could be classified as "real" is the *relation* between them[1]. Bohm [4] supposes the motion, or better speaking, his more general notion of holomovement, as the primitive notion (" ... the holomovement is undefinable ... ") and the system as an aspect of it (" ... the word "electron" should be regarded as no more than a name by which we call attention to a certain aspect of the holomovement ... "). We adopt a different view. In any case, what matters is to understand the "wholeness" of the situation and, once we take one concept as primitive, we define the

[1]It is appropriate to say that, in a general case, the process of category formation is not so simple, and that categorization involves both an arbitrary and a phenomenological aspect (see ref.[7]). However, the nature of this process does not matter to our purposes and is beyond our interests.

other and close this kind of vicious circle.

We consider the concept of *system* as a *primitive notion*. By system we mean *particles* and *fields*, that is, *physical systems*, but since the consideration of fields can be naturally done once we understand that of particles, we consider throughout this paper only particles and for simplicity only *one* particle. The consideration of the concept of system as a primitive notion is in accordance with our intuitive conception of motion as translation of bodies in space. But for a first account of what is motion we need another element. This is a spacetime theory [8].

In order to put this explicitly, let us consider in more detail the idea of the motion. Motion comes to mind through an act of perception [4] from an observer in a reference frame and is realized when, excluding obvious exceptions easily considered, in the same reference frame a given body is in different positions at different times. Thus, the definition of motion first requires the understanding of what is an observer, a reference frame and space and time, i.e., we need a *spacetime* theory. Once we accept a model for a spacetime theory, all these concepts are well-defined. Each of these models contains a structure called *spacetime*, which consists of a Hausdorff, connected, paracompact 4-dimensional real manifold and geometrical objects (for example, the metric field) defined in this manifold characterizing its geometry. In all spacetime theories, including the newtonian theory of space and time, an *observer* is a timelike curve pointing to the future and a *reference frame* is a timelike vector field such that each one of its integral lines is an observer. More details can be found in [8]. Since we will not consider in this paper the relativistic case, we adopt as a model for spacetime the *galilean spacetime* [9] which is an affine manifold. It can be viewed as a fiber bundle [10] with $\mathbb{R}$ as the base space, $\mathbb{R}^3$ as the typical fiber and Galilei group as the structural group. These considerations provide the precise definitions we need for the above concepts.

Now, to make our ideas work, we must take care of and note the fact that *two different nouns* are involved in that act of perception, that is, the *motion* and the *system*. These are *different nouns* and, even being related, we must respect this difference when we discuss the attributes of a noun. This is the *requirement* of *linguistic* nature already discussed [1,5]. In order to put this down explicitly, let us denote motion by $\mathcal{M}$ and system by $\mathcal{S}$. In [1] we saw that, if π is a property of a noun, then π(noun) is an adjective that defines a *state* of this noun. Since we have here two different nouns, we state:

Definition 2 Let π_M be a property of the motion ($\mathcal{M}$) and μ a measure [1]. We call $\pi_M(\mathcal{M})$ a *state of the motion* and the state variable [1] $\xi_M = \xi_M(\mathcal{M}) = \mu \circ \pi_M(\mathcal{M})$ a *dynamical variable* (DV).

Definition 3 Let π_S be a property of the system ($\mathcal{S}$) and μ a measure. We call $\pi_S(\mathcal{S})$ a *state of the system* and the state variable $\xi_S = \xi_S(\mathcal{S}) = \mu \circ \pi_S(\mathcal{S})$ an *independent variable* (IV).

Definition 4 Let $\pi_M^1, \ldots, \pi_M^n$ be the set (a conjunctive [1] set)[2] of all the properties of motion (possibly an infinity) and $f^{(i)} : \mathcal{D} \to I\!R$ certain mappings ($i = 1, \ldots, n$) ($\mathcal{D}$ is a set that is to be defined). We say that the *dynamics of motion* is given when it is given an element $\Upsilon \in \mathcal{D}$ (that defines *the* state of motion) such that $\xi_M^i = \mu \circ \pi_M^i(\mathcal{M}) = f^{(i)}(\Upsilon)$, i.e., Υ qualifies completely the noun motion.

Definitions 2 and 3 satisfy our linguistic requirement of giving different names to different things. The relation between these definitions and our usual notions of DV and IV is easily seen (and will be clarified in what follows). The use of the adjectives *dynamical* and *independent* are justified in each case. In the former we are considering a concept related to *change* and in the latter a concept which we choose as primitive and basic. This agrees with our intuitive ideas. Whatever relation may exist between them, it must be found later and not be imposed a priori. The role of def.4 will become clear later (and specially in [2]). Note that Υ instead of being defined as an element of $\mathcal{D}$ could be defined as a minimal set, if it exists, of DV $\xi_M^j, (j = 1, \ldots, k)(k < n)$, and that it remains to define the set $\mathcal{D}$ in def.4. However, the nature of this set is a *particularity* of a given theory of Mechanics (def.10), as we will see.

Remark 1 Before we continue, a brief digression is useful. In [1] we have defined a measurable property and, according to this, DV and IV are measurable properties. But (since there are very interesting points [4] involving the etymological root of the word "measure") to avoid confusions we must once and for all say what it means here. So, to define a measure we need to define a measurement [1]. According to that definition of a measurement,

[2]Whenever we write set we mean a conjunctive set.

it could be not only a "true" measurement – in the physical sense of an interaction between what is to be measured and a measuring apparatus – but in the case of a "true" measurement any two ones must give the same result [obviously, in these two ones what is to be measured is supposed to be prepared in the same way and the measurement is supposed to be an ideal one (no experimental errors) and made under the same conditions][3]. We did not need this distinction in [1], but here it is necessary in order to formulate a *physical* theory. In the sense of [1] a measurable property can properly be said to be an enumerable property and measuring is to "enumerate" (that is, to ascribe numbers). According to that sense, it is not necessary for a measurable property to satisfy the condition that any two measures give the same result: but this is necessary for a *physical* property[4]. To set up this distinction we state

Definition 5 Let π be a measurable property and μ_1 and μ_2 any two measures. If $\mu_1 \circ \pi = \mu_2 \circ \pi$ then π is called an observable property (non-observable property otherwise) and the state variable $\mu \circ \pi = \mu_1 \circ \pi = \mu_2 \circ \pi$ an observable.

From this definition it follows the obvious proposition: If π_1 and π_2 are observables properties and for given measures μ_i and μ_j we have $\mu_i \circ \pi_1 = \mu_j \circ \pi_2$, then π_1 and π_2 are the same observable property. The proof follows from def.5 and that of a measure as a bijection [1]. Note that μ_i and μ_j are measures that are made using the *same* measuring apparatus and on the *same* object.

---

[3]This statement needs a brief explanation. A scientific fact requires objectivity and reproducibility to be recognized as such. This means that, once we are studying a scientific fact, any two measures of a related quantity must a priori be equal. The differences that eventually appear in different results are in practice due to experimental errors (because of several reasons like noise and finite size of the measuring apparatus [1]). However, what matters is the belief that the sources of these errors are not in what is being measured; because, if this is the case, then it is not a scientific fact, according to the above criteria. Thus, that statement follows.

[4]Of course, it must be defined what is meant by a physical property. In order to avoid unnecessary (for our purposes) discussions about the nature of a physical property (and thus of physical reality, which is the question at the basis of the Bohr-Einstein debate) we mean by a physical property simply a property which can be studied by means of Physics and is thus constrained to satisfy the requirements of objectivity and reproducibility of a scientific fact.

Now, we look for the relation between π_M and π_S or between ξ_M and ξ_S. This relation exists, because there is a relation between $\mathcal{M}$ and $\mathcal{S}$ as expressed by the above act of perception. We need, however, something more polished for our purposes. A full arrangement of this idea will naturally yield our demanded relation. For what follows, in a coordinate representation [10], a point of the galilean spacetime is written as $(t, x) \in \mathbb{R} \times \mathbb{R}^3$.

Definition 6 Let $\pi_S = \mathsf{x}$ be the property of $\mathcal{S}$ of having a location in space, the IV $x = \mu \circ \mathsf{x}(\mathcal{S})$ be its position, t a parameter and x_t be the position of $\mathcal{S}$ at instant t. Let the proposition $x_{t_0}(\mathcal{S}) = x_0^{\#}$ be true, where $x_0^{\#} \in \mathbb{R}^3$ is the numerical value of x. If for an instant $t > t_0$ the proposition $x_t(\mathcal{S}) = x_0^{\#}$ is false, then we say that a motion $(\mathcal{M})$ occurred. If for this t, say t_n, $x_{t_n}(\mathcal{S}) = x_n^{\#}$, we write[5] $\mathcal{M} = \{n, 0\}(\mathcal{S})$ for a motion between t_0 and t_n.

Def.6 is valid for any $\mathcal{S}$. Then we may consider $\mathcal{M}$ independently of $\mathcal{S}$ and def. 6 as a "representation" of $\mathcal{M}$. This is a *very important point* since with it a *general* theory of $\mathcal{M}$ (which is the objective of Mechanics - def.1) can be formulated with *no reference* to $\mathcal{S}$ except for the *spacetime representation* of $\mathcal{M}$ given by def.6 – that is, we can *represent* motion in spacetime – and also because this fits the points above discussed concerning $\mathcal{M}$ and $\mathcal{S}$. Def.6 also exhibits a representation of $\mathcal{M}$ in the so called (extended) configuration space, which proves to be very natural and convenient, specially when we have other degrees of freedom besides the three translational ones considered in this paper. The *configuration space* U is a differentiable manifold whose dimension is equal to the number d of degrees of freedom of the system. When time is included, we have the *extended configuration space* $V = U \times \mathbb{R}$. In this paper we have $d = 3$ degrees of freedom, $U = \mathbb{R}^3$ and $V = \mathbb{R}^3 \times \mathbb{R} = \mathbb{R}^{3+1}$.

Now, remember that t is a varying continuous parameter. In order to simplify the notation and the ideas that follow, we will consider t as a discrete parameter t_i $(i \in \mathbb{N})$, and take the continuum limit at the end of the results. Also, instead of def.6, it is better to consider

Definition 7 Let M be the set $\{\{i, j\}(\mathcal{S}) \mid i, j \in \mathbb{N},\ i \neq j,\ \text{all } \mathcal{S}\}$, called the motion. By a spacetime representation of M we mean a bijection $\mathcal{R}(\mathcal{S})$:

[5] From now on, we will omit the superscript # on x and the reference to $\mathcal{S}$ on $\{n, 0\}$ whenever no confusion arises.

$M \to V \times V$ (here $\mathcal{R}(\mathcal{S}) : M \to I\!R^{3+1} \times I\!R^{3+1}$), $\{i, j\} \mapsto \mathcal{R}\{i, j\} = (x_i, t_i; x_j, t_j)$.

Now, given $\{n, 0\} \in M$ and $\{1, 0\}, \{2, 1\}, \ldots, \{n, n-1\} \in M$ and according to the principle of fragmentation (PF) (see [1] and references therein) the *set* $\{n, 0\}$ can be described by its elements $\{1, 0\}, \ldots, \{n, n-1\}$, i.e.,

$$\{n, 0\} = [\![\{1, 0\}, \{2, 1\}, \ldots, \{n, n-1\}]\!].$$

The question that must be put forth, according to [1], is then: Is the set $\{n, 0\}$ a conjunctive set, $\{n, 0\} = \bigcup_{k=0}^{n-1} \{k+1, k\}$, or a convolutive set, $\{n, 0\} = *_{k=0}^{n-1} \{k+1, k\}$? According to [1], the answer depends on the property which we are considering. If we denote this property by π, then it must be clear that we want to describe $\pi\{n, 0\}$ by *defining* it *from* $\pi\{1, 0\}, \pi\{2, 1\}, \ldots, \pi\{n, n-1\}$ (it must be stressed that we do not know beforehand the expression for $\pi\{n, 0\}$) and that it is $\pi\{1, 0\}, \pi\{2, 1\}, \ldots, \pi\{n, n-1\}$ that *are* already defined (because they are related to the properties of the system whose notion is taken as the primitive one)[6]. In the case of a conjunctive set relative to π, we have $(\mu \circ \pi)\{n, 0\} = \sum_{k=0}^{n-1} (\mu \circ \pi)\{k+1, k\}$, while for a convolutive set relative to π this equality does not hold: $(\mu \circ \pi)\{n, 0\} \neq \sum_{k=0}^{n-1} (\mu \circ \pi)\{k+1, k\}$. This is discussed in more details in the next section.

3 The Laws of Mechanics and "the Answer"

There are several formulations of the laws of Mechanics, of which the most beautiful are expressed in terms of a variational principle. However, it is not this "anthropomorphic" character that makes the variational principle more useful or important. It is because of its relation to the representation of a motion that we consider a variational principle as the expression of the laws of Mechanics. Furthermore, and this is a crucial step, we only suppose that there exists a variational principle, searching for it by means of general considerations. The reasons why we do so will become clear in what follows.

In order to fix notations, let us consider functions $\mathsf{f} : \mathcal{M} \to I\!R$ such that $f = \mathsf{f} \circ \mathcal{R}^{-1} : V \times V \to I\!R, (x_i, t_i; x_j, t_j) \mapsto f(x_i, t_i; x_j, t_j)$. If instead of x_i

[6] An illustrative example of what we are meaning by this is given by the way we show that fractal sets are an example of convolutive sets in Ref.[1].

and x_j we have different values x'_i and x'_j, we can define $f_h(x_i, t_i; x_j, t_j) = f(x_i + h_i, t_i; x_j + h_j, t_j), h = (h_i, h_j) = (x'_i - x_i, x'_j - x_j)$, and

$$\Delta f_h(x_i, t_i; x_j, t_j) = (f_h - f)(x_i, t_i; x_j, t_j) = f_h(x_i, t_i; x_j, t_j) - f(x_i, t_i; x_j, t_j).$$

The differential of f, denoted by df, is the linear part of Δf_h, i.e., $(\alpha \in I\!R)$

$$\Delta f_h(x_i, t_i; x_j, t_j) = df_h(x_i, t_i; x_j, t_j) + R(h), \quad \lim_{h\to 0} \frac{R(h)}{\|h\|} = 0$$

$$df_{h_1+\alpha h_2}(x_i, t_i; x_j, t_j) = df_{h_1}(x_i, t_i; x_j, t_j) + \alpha df_{h_2}(x_i, t_i; x_j, t_j).$$

This definition is easily generalized to functions from Banach space into another Banach space [10]: for example, $f : I\!R^m \to I\!R^n$; but for our purposes we need to consider only functions with values in $I\!R$ or $\mathbb{C}$.

Let $h = (x'_n - x_n, x'_0 - x_0)$ and such that $h = 0$ means that the values of x_n and x_0 are given and fixed. Therefore, we state

Postulate 1 (The Laws of Mechanics) Let there be a motion between x_0 at t_0 and x_n at t_n, i.e., $\{n, 0\} \in M$, which according to the PF can be described by its elements $\{k+1, k\}(k = 0, 1, \ldots, n-1)$. Then there exists a *non-observable* property of motion $\mathcal{A} : M \to \mathbb{C}$, called *the* $\mathbb{C}$*-action* (or *the* $\mathbb{C}$*-principal function*), such that:

(i) for the set $\{n, 0\}$ and $\{n, 0\} \mapsto \mathcal{A}\{n, 0\}$ it holds:

$$d\mathcal{A}\{n, 0\} \mid_{h=0} = 0;$$

(ii) for the elements $\{k+1, k\}$ and $\{k+1, k\} \mapsto \mathcal{A}\{k+1, k\}$ it holds:

$$\mathcal{A}\{k+1, k\} = i\mathcal{I}\{k+1, k\},$$

where $\mathcal{I}$ is *the* $I\!R$*-action* such that:

$$\mathcal{I}\{k+1, k\} = \ell_k \mu(\Delta t_k),$$

where ℓ_k is the value assumed in a point x'_k between x_k and x_{k+1} by a *continuous function which characterizes the system*, and $\mu(\Delta t_k)$ is the measure of the interval $\Delta t_k = t_{k+1} - t_k$.

It is important to note that, because the action is *not* an observable property, it does *not* need to be $\mathbb{R}$-valued (whenever no confusion arises, we call the $\mathbb{C}$-action simply the action – more on this issue is discussed in remark 4). Note that it is the $\mathbb{R}$-valued action of the elements $\{k+1,k\}$ that we define by relating it to the properties of the system. The use of a $\mathbb{C}$-valued action instead of a $\mathbb{R}$-valued one in condition (ii) is a matter of convenience. On the other hand, we have *not* defined an expression for the $\mathbb{C}$-valued action of $\{n,0\}$. Our *objective* is to *find* such an *expression* for $\mathcal{A}\{n,0\}$ *from* the ones of $\mathcal{A}\{k+1,k\}$ and such that $\mathcal{A}\{n,0\}$ *satisfies* the condition (i) above.

The postulate 1 follows the spirit of the scientific analysis method as discussed in [1]. We suppose the *fragmentation* when we "break" $\{n,0\}$ into its parts $\{1,0\},\ldots,\{n,n-1\}$ and *analysis* when we study these parts and their relationship. Now, the question of last section can be reformulated as:

Question 1 Relative to the action $\mathcal{A}$, is the set

$$\{n,0\} = [\![\{1,0\},\{2,1\},\ldots,\{n,n-1\}]\!]$$

a conjunctive set or a convolutive set ?[7]

In other words, we want to know how to write $\mathcal{A}\{n,0\}$ from $\mathcal{A}\{1,0\}$, $\mathcal{A}\{2,1\},\ldots,\mathcal{A}\{n,n-1\}$.

First, let us consider $A=\mathcal{A}\circ\mathcal{R}^{-1}$, $\mathcal{A}\{n,0\}=A(x_n,t_n;x_0,t_0)\equiv A(n,0)$, $\mathcal{A}\{k+1,k\} = A(x_{k+1},t_{k+1};x_k,t_k)\equiv A(k+1,k),(k=0,1,\ldots,n-1)$. Note that only the values x_n and x_0 are given and fixed, the other values $x_1,x_2,\ldots,x_{n-1}$ are not yet defined. So, in general we write the *ansatz*

$$A(n,0)=\mathcal{F}_1\left(\mathbf{O}[1,\ldots,n-1]\left(\mathcal{F}_2\left(\omega[A(1,0),A(2,1),\ldots,A(n,n-1)]\right)\right)\right),$$

where $\mathbf{O}[1,\ldots,n-1]$ is an operation – to be defined – acting over the values of $x_1,\ldots,x_{n-1}$ that tells us which values are to be chosen (for example, take $x_1=10,x_2=21,\ldots,x_{n-1}=2$ or take $x_k=+1$ or -1 if k is odd or even, etc.); and where $\mathcal{F}_1$, $\mathcal{F}_2$ and ω are functions, which are also to be defined. Note that they must be regarded as parts of a whole operation $\mathbf{W}[1,...,n-1]$. Note also that

$$\omega[A(1,0),A(2,1),\ldots,A(n,n-1)]=$$

[7]In ref.[1] we have defined conjunctive and convolutive sets only in relation to $\mathbb{R}$-valued properties. However, the generalization to $\mathbb{C}$-valued properties is trivial.

$$= \omega[A(x_1,t_1;x_0,t_0), A(x_2,t_2;x_1,t_1),\ldots,A(x_n,t_n;x_{n-1},t_{n-1})] =$$
$$= \omega[A'(x_0,x_1,\ldots,x_n)] \equiv \omega[A(x_0,x_1,\ldots,x_n)],$$

that is, in the limit $\epsilon \to 0$ we have a functional[8] $A[\gamma]$, where $\gamma : I\!R \ni [t_0,t_n] \to U, t \mapsto x(t) \mid x(t_0) = x_0 \,\text{and}\, x(t_n) = x_n$. It is important to bear in mind that the above expression for $A(n,0)$ is only a convenient way of writing the dependence of $A(n,0)$ on $A(1,0),\ldots,A(n,n-1)$. This convenience will become clear later (see Remark 3, which explains why we explicitly consider the functions $\mathcal{F}_2$ and ω, instead of only one function that plays the role of the composite function $\mathcal{F}_2 \circ \omega$).

We expect however the operation $\mathrm{O}[1,\ldots,n-1]$ and the functions $\mathcal{F}_1$, $\mathcal{F}_2$ and ω to be not completely arbitrary. One immediate consequence of the hypothesis of homogeneity of space is that the operations over each $x_1,\ldots,x_{n-1}$ must be the same, i.e., $\mathrm{O}[1,...,n-1]$ can be written as the same operation $\bigwedge_k$ over x_k, $\mathrm{O}[1,...,n-1] = \bigwedge_1 \ldots \bigwedge_{n-1}$. This operation must also satisfy other physical hypotheses and be restricted to values of x_k that satisfy some possible constraints of the system. We must also have, because $\mathcal{A}$ is $\mathbb{C}$-valued, that $\mathcal{F}_2 : \mathbb{C} \to X$ and $\mathcal{F}_1 : X \to \mathbb{C}$ for some set X, and $\omega : \mathbb{C} \to \mathbb{C}$. Therefore, because we take t as a discrete parameter, we shall write

$$A(n,0) = \lim_{\epsilon\to 0} \mathcal{F}_1 \bigwedge_1 \ldots \bigwedge_{n-1} \mathcal{F}_2\omega[A(1,0),A(2,1),\ldots,A(n,n-1)]$$
$$\equiv \mathcal{F}_1 \bigwedge_{[\gamma]} \mathcal{F}_2 \omega(A[\gamma]) \qquad (1)$$

where we have considered the limit $\epsilon = \Delta t_k = t_{k+1} - t_k \to 0$, $(k = 0,1,\ldots,\text{n-1})$ and used this expression in order to *define* the right hand side. Depending on the definition of the operator $\mathrm{W}[1,\ldots,n-1] \equiv \mathcal{F}_1 \bigwedge_1 \ldots \bigwedge_{n-1} \mathcal{F}_2$ it may, or not, be necessary to introduce some normalization factor in order for this limit to exist, but we do not need to consider this point as yet.

Another immediate requirement that $A(n,0)$ has to satisfy is postulate 1. This is what we intend to do: find an expression for the action such that it satisfies the laws of Mechanics as expressed by postulate 1. Before doing this, we must recognize another exigency. If for the action $A(n,0)$

[8] Although $A(k+1,k)$ and $A[\gamma]$ are different mathematical objects, the use of the same letter to denote them is customary and should not cause confusion.

we must have $dA(n,0)\mid_{h=0}=0$, then what can be said about $dA(k+1,k)$ (remember that $dA(k+1,k)=idI(k+1,k)$, where $I(k+1,k)\in\mathbb{R}$ according to condition (ii) of postulate 1)? If we pay attention to the fact that in postulate 1 we have x_n and x_0 given and fixed, we immediatly see that, because $x_1, x_2, \ldots, x_{n-1}$ are not given and not fixed (as yet), the same condition must *not* hold for the individuals $dA(k+1,k)$. This is due to the fact that $dA(1,0), \ldots, dA(n,n-1)$ are *not* independent at all. The value of each $dA(k+1,k)$ depends on $h_k \equiv (x'_{k+1}-x_{k+1}, x'_k - x_k)$ and then only $(n-1)$ terms $dA(k+1,k)$ from a total of n can assume a value independently from the others. Then, we impose

$$\sum_{k=0}^{n-1} dA_{h_k}(k+1,k) = i\sum_{k=0}^{n-1} dI_{h_k}(k+1,k) = 0. \qquad (2)$$

Remark 2 It must be clear that we cannot impose $\sum_{k=0}^{n-1} a_k dA_{h_k}(k+1,k) = 0$ $(a_k \in \mathbb{R})$ because of: (a) the linearity of dA (by definition) $a_k dA_{h_k}(k+1,k) = dA_{a_k h_k}(k+1,k)$, and (b) the non-independence of h_k $(k=0,1,\ldots,n-1)$. Remember that with x_n and x_0 given and fixed, we have $h_0=(0,x'_1-x_1)$ and $h_{n-1}=(x'_{n-1}-x_{n-1},0)$. When we choose a value for x'_1-x_1, say r_1, we have $h_0=(0,r_1)$ determined and $h_1=(r_1,x'_2-x_2)$; for $x'_2-x_2=r_2$ we have $h_1=(r_1,r_2)$ determined and $h_2=(r_2,x'_3-x_3)$, and so on. Now, when we choose $x'_{n-1}-x_{n-1}=r_{n-1}$ we have not only $h_{n-2}=(r_{n-2},r_{n-1})$ determined but also $h_{n-1}=(r_{n-1},0)$. With $a_k dA_{h_k}=dA_{a_k h_k}$, then $h'_0 = (0,r'_1)=a_0h_0=a_0(0,r_1)$, $h'_1=(r'_1,r'_2)=a_1h_1=a_1(r_1,r_2)$, $\ldots$, $h'_{n-1} = (r'_{n-1},0)=a_{n-1}h_{n-1}=a_{n-1}(r_{n-1},0)$, from which we must have $a_0=a_1=\ldots=a_{n-1}\equiv a$. As at least one of the a_k's must be non-zero (definition of linear dependence), then $a_k=a\neq 0$ for $k=0,1,\ldots,n-1$, which yields (2).

We are nearly in the condition to use expression (1) for postulate 1. To do this we need some hypotheses. One reason is clear: since we have not made any hypothesis concerning linearity, $\mathbf{W}[1,...,n-1]=\mathcal{F}_1\bigwedge_1\ldots\bigwedge_{n-1}\mathcal{F}_2$ could be in general a non-linear operation and this non-linearity could come from $\bigwedge$ or from $\mathcal{F}_1$ and $\mathcal{F}_2$ or from both. Other reasons will appear after their introduction. These hypothesis are, for f and g suitable functions,

Hypothesis (a): $\bigwedge_k$ is a linear operation over x_k,

$$\bigwedge_k [g_1(x_k) + cg_2(x_k)] = \bigwedge_k g_1(x_k) + c\bigwedge_k g_2(x_k);$$

Hypothesis (b): $g_1 \bigwedge_1 \ldots \bigwedge_{n-1} g_2(f_{(h_0,h_1,\ldots,h_{n-1})}(x_0, x_1, \ldots, x_{n-1}, x_n)) = 0,$

$$\forall h_0, h_1, \ldots, h_{n-1} \Longrightarrow (f_{(h_0,h_1,\ldots,h_{n-1})}(x_0, x_1, \ldots, x_{n-1}, x_n)) = 0;$$

Hypothesis (c): $g(x_k) \geq 0 \Longrightarrow \bigwedge_k g(x_k) \geq 0;$

Hypothesis (d): $\mathcal{F}_1$ and $\mathcal{F}_2$ are diffeomorphisms, $\mathcal{F}_1 : \mathbb{C} \to \mathbb{C}$, $\mathcal{F}_2 : \mathbb{C} \to \mathbb{C}$, and $\mathcal{F}_1 = \mathcal{F}_2^{-1}, \mathcal{F}_2 = \mathcal{F}_1^{-1}$.

Note that, even with these hypotheses, we are still situated at a very general level. They are very general (or obvious, as for hypothesis(b)), and we still work with a non-linear operation. The hypothesis (a) says where the non-linearity (if it exists) comes from. The hypothesis (c) will be clear when used (it will be necessary in [2]), the same being the case for hypothesis (d).

Now we can use eq.(1) for postulate 1. We have, using hypothesis (a),

$$dA_h(n,0)\mid_{h=0} = \mathcal{F}_1' \bigwedge_1 \ldots \bigwedge_{n-1} \mathcal{F}_2'(d\omega[A(1,0), A(2,1), \ldots, A(n,n-1)]) =$$

$$\mathcal{F}_1' \bigwedge_1 \ldots \bigwedge_{n-1} \mathcal{F}_2' \left[\frac{\partial \omega}{\partial A(1,0)} dA_{h_0}(1,0) + \ldots + \frac{\partial \omega}{\partial A(n,n-1)} dA_{h_{n-1}}(n,n-1) \right] = 0.$$

With hypothesis (b) and using (2) for instance in $dA_{h_0}(1,0)$, we have

$$\left[\frac{\partial \omega}{\partial A(2,1)} - \frac{\partial \omega}{\partial A(1,0)} \right] dA_{h_1}(2,1) + \ldots$$

$$\ldots + \left[\frac{\partial \omega}{\partial A(n,n-1)} - \frac{\partial \omega}{\partial A(1,0)} \right] dA_{h_{n-1}}(n,n-1) = 0.$$

Remembering that $dA_{h_k}(k+1,k) \neq 0$, it follows that ω must satisfy the system of differential equations

$$\frac{\partial \omega}{\partial A(1,0)} = \frac{\partial \omega}{\partial A(2,1)}, \ldots, \frac{\partial \omega}{\partial A(1,0)} = \frac{\partial \omega}{\partial A(n,n-1)}.$$

The general solution of this well-known system can be easily found by a change of variables, from $A(1,0),\ldots,A(n,n-1)$ to $A'_0 = A(1,0)+A(2,1)+\ldots+A(n,n-1)$, $A'_1 = A(1,0)-A(2,1),\ldots,A'_{n-1} = A(1,0)-A(n,n-1)$. Denoting by W the function ω written in the variables A', we obtain a system whose determinant of the matrix of the coefficients is always non-vanishing, and the solution is the trivial one $\partial W/\partial A'_1 = \ldots = \partial W/\partial A'_{n-1} = 0$. Then our general solution is a function of variable A'_0 alone, i.e.,

$$\omega[A(1,0),A(2,1),\ldots,A(n,n-1)] =$$
$$\omega[A(1,0)+A(2,1)+\ldots+A(n,n-1)], \quad (3)$$

where we have seen that

$$\sum_{k=0}^{n-1} A(k+1,k) = \sum_{k=0}^{n-1} A(x_{k+1},t_{k+1};x_k,t_k) \equiv A(x_0,x_1,\ldots,x_n)$$

in the limit $\Delta t_k = \epsilon \to 0$ is a functional of the paths γ joining x_0 and x_n, $\gamma : I\!R \ni [t_0,t_n] \to U, t \mapsto x(t) \mid x(t_0) = x_0, x(t_n) = x_n$. Moreover, and this is the principal conclusion we draw from (3), it is a *local* functional [11].

Remark 3 This result is a justification for what we said in the beginning of this section, i.e., that the expression we wrote for $A(n,0)$ is only a convenient way of writing the dependence of $A(n,0)$ on$A(1,0),\ldots,A(n,n-1)$. The only role played by ω is to permit us to arrive at (3), that is, to conclude that $A[\gamma] = \lim_{\epsilon\to 0} A(x_0,x_1,\ldots,x_n)$ is a local functional. The "localization property" of this functional is a very important one. Now, since the consideration of the function ω had achieved its purposes, we can join it in $\mathcal{F}_2$ and rewrite the hypothesis (d) or simply take it equal to the identity. So, from now on we will take $\omega \equiv 1$.

In order to write an expression for this functional in the limit $\epsilon \to 0$ it is necessary to define a measure $\mu(\epsilon)$ (condition (ii) of postulate 1). We adopt a Lebesgue-Stieltjes measure [12] $\mu_F(t)$ defined from a given generating function $F(t)$ and define

$$I(k+1,k) \equiv \ell^0_k \mu_F(\epsilon),$$

where now we write ℓ^0_k for the value assumed in a point x'_k between x_k and x_{k+1} by a function $\mathcal{L}^0$ which we suppose to be continuous. It is this function

$\mathcal{L}^0$ **that** ***characterizes the mechanical system.*** **Furthermore, if we remember def.6, we conclude that $\mathcal{L}^0$ must be a function of x and $\dot{x} = dx/dt$ (and maybe t) since it is necessary to give the value of x at two distinct times to define a motion.** With the additional hypothesis that $F(t)$ is absolutely continuous and defining $\mathcal{L} \equiv \mathcal{L}^0 F'(t)$, where $d\mu_F(t) = dF(t) = F'(t)dt$, we write

$$\lim_{\epsilon \to 0} \sum_{k=0}^{n-1} \ell_k^0 \mu_F(\epsilon) \equiv \int_{t_0}^{t_n} d\mu_F(t) \mathcal{L}^0(x, \dot{x}, t) = \int_{t_0}^{t_n} dt \mathcal{L}(x, \dot{x}, t).$$

In the next section we show that $\mathcal{L}(x, \dot{x}, t)$ is the usual lagrangian of the system; therefore it is $\mathbb{R}$-valued, as assumed in postulate 1.

Definition 8 Let Γ be the set of all paths $\gamma : \mathbb{R} \ni [t_0, t_n] \to U, t \mapsto x(t) \mid x(t_0) = x_0, x(t_n) = x_n$. The functional $I : \Gamma \to \mathbb{R}, \gamma \to I[\gamma]$,

$$I[\gamma] = \lim_{\epsilon \to 0} \sum_{k=0}^{n-1} \ell_k \epsilon \equiv \int_{t_0}^{t_n} dt \mathcal{L}(x, \dot{x}, t), \tag{4}$$

is called *the $\mathbb{R}$-action functional* ($\mathcal{L} : TU \times \mathbb{R} \to \mathbb{R}$ is called the *lagrangian* (TU is the tangent bundle) and $A[\gamma] = iI[\gamma]$ is *the $\mathbb{C}$-action functional*).

Remark 4 Note that the quantity that is related to the system through the lagrangian in def.8 is called the *action functional* and the one related to the motion in postulate 1 is called the *action*. This is a very important distinction and must not cause confusion. What might cause some confusion is the fact that, in Physics, what we have called the *action functional* is sometimes simply called the *action*, a denomination which does not take into account the mathematical nature of this object[9]. We will see in the next section that what we call the *action* is sometimes called the *action function* or the *principal function*[10]. In any case, our denomination must not cause confusion because we are considering in each case different mathematical

---

[9]For example: Landau and Lifshitz [13] call it simply action, Sudarshan and Mukunda [14] properly call it the action functional and Arnold [9] also makes this correct distinction but not in an insistent way.

[10]For example: in Classical Mechanics (remember that we do not have considered Classical Mechanics as yet) it is called action function by Arnold [9] and Hamilton's principal function by Saletan and Cromer [15].

objects. The same holds for the $\mathbb{R}$-action functional and the $\mathbb{C}$-action functional. What we must have in mind, and this is a very important point, is that we are considering the *same* property, in one case for the set and in the other for the elements of this set. When we say "the action" we refer to a property called action of the set $\{n,0\}$ and the same for each element $\{k+1,k\}$ of this set. When we say "the action functional" we refer to a quantity *defined from* this property of each one of those elements $\{k+1,k\}$ and which has the particular property of being a local functional.

Finally, introducing (4) in (3) and then in (1) and using hypothesis (d) (the justification of it is given in the next section), we obtain a final expression for $A(n,0)$. We can summarize our results for this expression for $A(n,0)$ as a proof for the

Proposition 1 There exists an infinity of expressions for the $\mathbb{C}$-action $A(n,0)$ such that it satisfies the laws of Mechanics as expressed by postulate 1 and which can be written in the general form

$$A(n,0) = \lim_{\epsilon\to 0} \mathcal{F}^{-1} \bigwedge_{1} \dots \bigwedge_{n-1} \mathcal{F}\left(\sum_{i=0}^{n-1} A(i+1,i)\right) \equiv \mathcal{F}^{-1} \bigwedge_{[\gamma]} \mathcal{F}(A[\gamma]), \qquad (5)$$

where $\bigwedge_i$ and $\mathcal{F}$ satisfies the hypothesis (a)-(d) and $A[\gamma]$ is the $\mathbb{C}$-action functional given by $A[\gamma] = iI[\gamma]$ where $I[\gamma]$ is the $\mathbb{R}$-action functional (def.8).

All the efforts of this section where made towards an answer to question 1. This answer is provided once we know how to write $A(n,0)$ from $A(1,0),\dots,A(n,n-1)$ and we have hoped that the requirement of the action to satisfy the laws of Mechanics tell us how to do this. But this has given us only the proposition 1. This, however, must be viewed as a welcome fact, that we will discuss in the next sections.

In any case, given an expression for the action we can formulate a physical theory in order to describe the motion. As we said, this will be discussed in the next sections, but before doing this we must explain what is meant here by a physical theory, i.e.,

Definition 9 (Bernardes') A physical theory is a general scheme, from which we can formulate propositions about observable properties and such that it is based on axiomatic propositions which involve at least one non-observable property[11].

Definition 10 A physical theory that satisfies the laws of Mechanics as expressed by postulate 1 is called a theory of Mechanics.

Remark 5 From definitions 1, 9 and 10 we see that a theory of Mechanics must formulate propositions about the observable properties of the *motion* (def.2 and def.5) from a general scheme provided by the laws of Mechanics. According to def.4, this scheme must provide a way to know $\Upsilon \in \mathcal{D}$. On the other hand, since the concept of system is taken as a primitive notion, the description of the system is provided by the description of motion (remember sec.2). In other words, in a general case the system is to be described by an element Υ that belongs to a set $\mathcal{D}$ whose definition depends on a given theory of Mechanics (more on this issue is discussed in [2]).

4 Conjunction and Classical Mechanics

We saw in the preceding section that, given an expression for the action, we can formulate a physical theory in order to describe the motion. Proposition 1 does not, however, give a unique answer to question 1. We must have another tool in order to decide for one expression for the action among the infinite many provided by (5).

At this point, any choice among the ones in (5), that is, any hypothesis concerning the nature of the set $\{n, 0\}$, is not a necessary one. The necessary condition provided by the laws of Mechanics is equally satisfied by any expression in (5).

Let us see one example. One choice among those is provided by the acceptance of

Definition 11 (The Principle of Conjunction (PC)–in mechanics) [12] Relative to the action, the set $\{n, 0\} = [\![\{1, 0\}, \{2, 1\}, \ldots, \{n, n-1\}]\!]$ is

[11]Note that Bernardes' definition distinguishes theory and phenomenology, with the latter being a scheme which involves only observables properties.

[12]The PC plays a fundamental role in what could be called Classical Physics. In what follows we discuss its role in Mechanics; in Thermodynamics this is done in the work of Bernardes [3] which contains a full account of it.

a conjunctive set, i.e.:

$$\{n,0\} = \bigcup_{k=0}^{n-1}\{k+1,k\} \Longleftrightarrow A(n,0) = \sum_{k=0}^{n-1} A(k+1,k).$$

Now we have

Proposition 2 The theory of Mechanics that satisfies the hypothesis of the validity of the PC is Classical Mechanics (CM).

To prove this proposition, remember first that $\bigwedge_k$ in (5) is a linear operation over the allowed values of x_k that are compatible with the constraints of the system. In order to have $A(n,0) = \sum_{k=0}^{n-1} A(k+1,k)$ this operation must select only one value $x_k^{\#}$ among all the allowed ones, since then $A(n,0) = \mathcal{F}^{-1}\mathcal{F}(\sum_{k=0}^{n-1} A(k+1,k)) = \sum_{k=0}^{n-1} A(k+1,k)$. We do not yet know what this value is, but in the limit $\epsilon \to 0$ we will have $A(n,0) = A[\gamma] \mid_{\gamma=\bar{\gamma}}$ where $\bar{\gamma}$ is a given path which constitutes the set of the values $x_k^{\#}$. In other words, the operation $\bigwedge_{[\gamma]}$ says that we must take among all the paths $\gamma \in \Gamma$ the path $\bar{\gamma}$. In order to know which path is it, we use postulate 1. Then we have $dA(n,0) \mid_{h=0}= d(A[\gamma]) \mid_{\gamma=\bar{\gamma}}= 0$ and because $\bigwedge$ is a linear operation, $d(A[\gamma]) \mid_{\gamma=\bar{\gamma}}= (\delta A[\gamma]) \mid_{\gamma=\bar{\gamma}}$ (where we use δ to denote the differential of a functional as it is customary [11]); and for the path $\bar{\gamma}$, $(\delta A[\gamma]) \mid_{\gamma=\bar{\gamma}}= 0$. As we see from the definition of $A[\gamma] = iI[\gamma]$, eq.(4), the condition $(\delta A[\gamma]) \mid_{\gamma=\bar{\gamma}}= 0$ implies the condition $(\delta I[\gamma]) \mid_{\gamma=\bar{\gamma}}= 0$, that is the Hamilton principle of CM (of course, in CM it does not matter if we start working with $A[\gamma]$ or $I[\gamma]$; thus, we shall not be worried with this point).

The condition $(\delta A[\gamma]) \mid_{\gamma=\bar{\gamma}}= 0$ is the necessary condition [11] for the differentiable functional $A[\gamma]$ to have an *extremum* for $\gamma = \bar{\gamma}$. Then $0 = (\delta A[\gamma]) \mid_{\gamma=\bar{\gamma}}= \underset{\gamma}{\text{ext}}(\delta A[\gamma]) = d(\underset{\gamma}{\text{ext}} A[\gamma])$ and, for the $\bigwedge_{[\gamma]}$ operation, $\bigwedge_{[\gamma]} = \underset{\gamma}{\text{ext}}$.

Definition 12 The operation $\bigwedge_{[\gamma]} = \underset{\gamma}{\text{ext}}$ is called Hamilton operation.

The Hamilton operation is an operation over *all* the paths $\gamma \in \Gamma$ which selects among them the one that gives the extremum[13] of $A[\gamma]$. Thus $\underset{k}{\text{ext}}$

[13]Indeed, we have in general a stationary value of $A[\gamma]$ or a minimum for sufficiently small time intervals, as well-known; but throughout this paper and [2] we omit such details and consider it simply as an extremum.

consists in selecting among all the allowed values of x_k the value $x_k^{\#} = \bar{x}(t_k)$, where $\Gamma \ni \bar{\gamma} : t \mapsto \bar{x}(t)$. We can say that CM is the theory of Mechanics that is based on the choice of $\bigwedge$ to be the Hamilton operation, and that such a choice has as its foundation the acceptance of the validity of PC. Note that this choice explains hypothesis (c) of sec.3, since it could be $-\underset{\gamma}{\text{ext}}$ (the negative of $\underset{\gamma}{\text{ext}}$), some redefinitions being enough for this. Note also that $A(n,0) = \underset{\gamma}{\text{ext}} A[\gamma] = i\,\underset{\gamma}{\text{ext}} I[\gamma] = iI(n,0)$, where $I(n,0)$ is what is usually call Hamilton's principal function or the action function. Our $\mathbb{C}$-action could properly be called the $\mathbb{C}$-Hamilton's principal function.

This particular choice has particularly remarkable consequences. They are well-known facts in the realm of CM, but we must now look at them as mere consequences of the choice we made of the $\bigwedge$ operation to be the Hamilton operation. Let us discuss now some points that are quite important for us:

PC Acceptance Consequence (a): We saw that def.6 forwards a natural and convenient representation of the motion in the configuration space and in def.7 we have defined its representation, $\mathcal{R}\{n,0\} = (x_n, t_n; x_0, t_0)$. It does not say anything about what happens between t_0 and t_n. It only introduces a language that permits us to speak about the motion. This language is based on the concepts of paths. Now, it is the law of Mechanics that says something about the motion and *how* to use this language to speak about it. Once accepted the PC we arrive at CM, and then the action is given by the value of the action functional for a path $\bar{\gamma}$ which gives its extremum among all the possible paths $\gamma \in \Gamma$ between (x_0, t_0) and (x_n, t_n). Thus, in CM we can represent the motion in the configuration space as a given path (the so called actual path) among all paths joining those points.

PC Acceptance Consequence (b): Being a theory of Mechanics, CM can formulate propositions about the observable properties of motion which are DVs. Because of consequence (a) we can describe the motion at any instant t as a point (x,t) in the configuration space, where x is an IV (see def.6). This leads us to "identify" in CM the concepts of DV and IV. Also, as it is well-known, it is enough to know the action (for example by a solution of Hamilton-Jacobi equation) to solve any mechanical problem. According to def.4 the action defines the dynamics of motion. On the other hand, it is also enough to know what is the path $\bar{\gamma}$ (for example by a solution of the

Lagrange equations) which gives the extremum of the action functional. So, a function $t \mapsto x(t)$ also defines the dynamics of motion in CM. Then, all the DV are determined once we know $x(t)$, which is an IV. Therefore, although the concepts of DV and IV are different, there is an "identification" of them in CM because their values are equal.

PC Acceptance Consequence (c): The above consequences lead to another result. Instead of asking for a value of a DV we can in CM ask for a value of an IV. Because the IVs describe the system, CM can be viewed in this sense as a genuine theory of the system; that is, we can define (apart from mass) a particle in CM as the mapping $\gamma : \mathbb{R} \to U, t \mapsto x(t)$. The use of the adjective "genuine" is justified by the discussion we made in sec.2 and by remark 5. In CM all the propositions that we formulate about the motion are propositions about the system because of consequence (b). Any other theory of Mechanics also describes the system (remark 5) but not "genuinely", in the sense that its propositions are to be refered to DVs and not to IVs (remember that the IVs are parts of a language that we introduce in order to speak about the motion and which refer to properties of the system).

Now, since the action defines the dynamics of motion, we state

Definition 13 If the motion is such that the action is a conjunctive property, i.e., $\{n, 0\} = [\![\{1, 0\}, \{2, 1\}, \ldots, \{n, n-1\}]\!]$ is a conjunctive set relative to the action, then we will say that the motion has a *conjunctive dynamics*.

5 Convolution and Generalized Mechanics

The preceeding discussions put forth some questions. For example: (a) Is the PC hypothesis valid ? (b) Is there any other choice of $\bigwedge$ instead of the Hamilton operation ?

The answer to question (a) is immediate. First, experiences has shown that CM has a limited validity. If there is something wrong with CM, then it seems reasonable that the problem is in the acceptance of PC. In fact, all our other hypotheses seem very natural and we have shown in [1] that there exist sets (fractal sets) which can be considered as not being conjunctive sets.

If the answer to (a) is negative, then we expect the answer to (b) to be affirmative. We show that this is indeed the case in [2]. There exist an infinity of possibilities. A remarkable fact is provided by one of these: it will lead us to a theory that is formally identical to Quantum Mechanics. This will be discussed in [2].

However, none of those points are necessary for what we intend to do here. It is enough for us answering the question: Is it necessary, at the point we arrived at the end of sec.3, to make any choice of $\bigwedge$? Surely the answer is *no*! Proposition 1 offers to us not one possibility but an infinity. We can formulate not one theory of Mechanics, but a large class of them.

Proposition 1 defines a class of theories of Mechanics. Given an operation $\bigwedge$, by (5) we know how to write an expression for the action. Moreover, this might be such that we do not have $A(n,0) = \sum_{k=0}^{n-1} A(k+1,k)$. In this case the set $\{n,0\} = [\![\{1,0\},\{2,1\},\ldots,\{n,n-1\}]\!]$ is a convolutive set relative to the action, $\{n,0\} = *_{k=0}^{n-1}\{k+1,k\}$, where $*$ denotes convolution [1].

Definition 14 Any theory of Mechanics such that the set $\{n,0\} = [\![\{1,0\},\{2,1\},\ldots,\{n,n-1\}]\!]$ is a convolutive set relative to the action, is called a *Generalized Mechanics*(GM).

Definition 15 If the motion is such that the action is a convolutive property, i.e., $\{n,0\} = [\![\{1,0\},\{2,1\},\ldots,\{n,n-1\}]\!]$ is a convolutive set relative to the action, then we will say that the motion has a *convolutive dynamics*.

Once we recognize the possibility of formulating a class of theories of Mechanics we must ask: How to do it ? In other words, we must answer

Question 2 In a GM, what are the variables that play the role of DV, that is, what are the observables of motion if it has a convolutive dynamics ?

Question 3 How do these DVs evolve with time, that is: what are the equations of motion in GM ?

The answer to these questions is given in [2]. To such an end, we need the help of another postulate that we will discuss now.

Postulate 2 Every GM has a Classical Limit.

Proposition 3 A necessary condition for a GM to have a Classical Limit is to be congruous with CM.

First, let us explain what is meant by a congruity between two GM. In fact, congruity is used to express an obvious linguistic requirement. Suppose we have two operations $\bigwedge_1$ and $\bigwedge_2$ and the correspondent expressions for the actions A_1 and A_2; A_1 (A_2) is the action (the principal function) in GM1 (GM2). Let π be a property of motion such that in GM1 the DV ξ_1 is given from A_1 by $\xi_1 = f(A_1)$. We have here a DV ξ_1 that have a certain name (call it "name"). Thus there is the proposition: "name" is obtained from the action by f. Now in GM2 we have the DV ξ_2 that corresponds to the property π and whose name is "name" because π is the same. Thus in GM2 we have $\xi_2 = f(A_2)$. Because f is the same we say that GM1 and GM2 are congruent.

Now consider a GM and CM. In the Classical Limit $\bigwedge_{[\gamma]} \to \underset{\gamma}{\mathrm{ext}}$ and $\xi_{GM} \to \xi_{CM}$ in a certain limit that depends on $\bigwedge_{[\gamma]}$. If $\xi_{CM} = f(\underset{\gamma}{\mathrm{ext}} A[\gamma])$ and $\xi_{GM} = g(\bigwedge_{[\gamma]} A[\gamma])$ then in the Classical Limit $\xi_{GM} \to g(\underset{\gamma}{\mathrm{ext}} A[\gamma]) = f(\underset{\gamma}{\mathrm{ext}} A[\gamma])$. But $\underset{\gamma}{\mathrm{ext}}$: $A[\gamma] \mapsto \underset{\gamma}{\mathrm{ext}} A[\gamma] = A_{CM}$ is clearly *onto*, since every A_{CM} is the image under $\underset{\gamma}{\mathrm{ext}}$ of some functional $A[\gamma]$ (that is, if there is a motion, then there exists some system which is in motion), and then $g = f$, i.e., CM and GM are congruent. This condition is clearly non-sufficient, since we cannot conclude from it that $\bigwedge_{[\gamma]} \to \underset{\gamma}{\mathrm{ext}}$ in a certain limit[14]

Remark 6 Note that we did not give much importance to $\mathcal{F}$ in (5); but now hypothesis (d) concerning it in sec.3 is clear. It is a necessary one for CM as we saw in proposition 2 and now for postulate 2. A convenient way to skip taking into account is to define

$$\Phi \equiv \mathcal{F} \circ A \tag{6}$$

[14]From this it follows that the exigency of congruity between a GM and CM is more general than the one of having a classical limit. Thus, with postulate 2 we intend to consider only systems with a classical analog. This is not, however, a restrictive assumption since it is the exigency of congruity that is used in practice in [2].

and work with Φ instead of A. Thus, from (5) we have

$$\Phi(n,0) = \lim_{\epsilon\to 0} \bigwedge_1 \ldots \bigwedge_{n-1} \mathcal{F}\left(\sum_{k=0}^{n-1} A(k+1,k)\right) \equiv \bigwedge_{[\gamma]} \mathcal{F}(A[\gamma]). \qquad (7)$$

On the other hand, although the use of $\Phi(n,0)$ instead of $A(n,0)$ seems natural (mainly because $\bigwedge$ is a linear operation), the need to know $\mathcal{F}$ cannot be avoided, because it is for $A[\gamma]$ (not $\Phi[\gamma]$) that we have defined an expression via lagrangian. In spite of this, it should not be regarded as an additional difficulty since $\mathcal{F}$ can be easily found. Suppose that we have a motion between x_0 at t_0 and x_n at t_n, that is, $\{n,0\} \in M$, characterized by $\Phi(n,0)$, and then a motion between x_n at t_n and x_s at t_s, $\{s,n\} \in M$, characterized by $\Phi(s,n)$. The whole situation can be viewed as a motion between x_0 at t_0 and x_s at t_s, $\{s,0\} \in M$, characterized by $\Phi(s,0)$. Thus, it is legitimate to suppose that $\Phi(s,0)$ can be obtained from $\Phi(n,0)$ and $\Phi(s,n)$, that is, the motion between x_0 at t_0 and x_s at t_s can be described from the ones between x_0 at t_0 and x_n at t_n and between x_n at t_n and x_s at t_s. In other words, we suppose that the set of all $\Phi(i,j)$ constitutes a semigroup[15], i.e., that we have an associative composition law m such that $m : (\Phi(s,n), \Phi(n,0)) \mapsto m(\Phi(s,n), \Phi(n,0)) = \Phi(s,0)$. Now from the expression for these quantities we must get

$$m\left\{\lim_{\epsilon\to 0} \bigwedge_1 \ldots \bigwedge_{n-1} \mathcal{F}\left(\sum_{k=0}^{n-1} A(k+1,k)\right), \lim_{\epsilon\to 0} \bigwedge_{n+1} \ldots \bigwedge_{s-1} \mathcal{F}\left(\sum_{k=n}^{s-1} A(k+1,k)\right)\right\} =$$

$$= \lim_{\epsilon\to 0} \bigwedge_1 \ldots \bigwedge_{n-1} \bigwedge_n \bigwedge_{n+1} \ldots \bigwedge_{s-1} \mathcal{F}\left(\sum_{k=0}^{s-1} A(k+1,k)\right).$$

Since on the right hand side we have an operation $\bigwedge_n$ which does not appear in the arguments of m, it is legitimate to suppose that m involves $\bigwedge_n$. Thus, we suppose m to be such that it holds

$$\bigwedge_n \left[\lim_{\epsilon\to 0} \bigwedge_1 \ldots \bigwedge_{n-1} \mathcal{F}\left(\sum_{k=0}^{n-1} A(k+1,k)\right)\right] \left[\lim_{\epsilon\to 0} \bigwedge_{n+1} \ldots \bigwedge_{s-1} \mathcal{F}\left(\sum_{k=n}^{s-1} A(k+1,k)\right)\right] =$$

[15] In fact, it constitutes a group, as we will see in [2], but we do not need to consider that point here.

$$= \lim_{\epsilon\to 0} \bigwedge_1 \cdots \bigwedge_{n-1} \bigwedge_n \bigwedge_{n+1} \cdots \bigwedge_{s-1} \mathcal{F}\left(\sum_{k=0}^{n-1} A(k+1,k) + \sum_{k=n}^{s-1} A(k+1,k)\right).$$

Once we suppose that the individual limits exist and that the order of the $\bigwedge$'s may be changed, we arrive at the following condition, that $\mathcal{F}$ must satisfy:

$$\mathcal{F}\left(\sum_{k=0}^{n-1} A(k+1,k)\right)\mathcal{F}\left(\sum_{k=n}^{s-1} A(k+1,k)\right) = = \mathcal{F}\left(\sum_{k=0}^{n-1} A(k+1,k) + \sum_{k=n}^{s-1} A(k+1,k)\right). \quad (8)$$

From this condition we conclude that $\mathcal{F}$ must be an exponential function, and that

$$\Phi(n,0) = \lim_{\epsilon\to 0} \bigwedge_1 \cdots \bigwedge_{n-1} \exp\left(\sum_{k=0}^{n-1} A(k+1,k)\right) \equiv \bigwedge_{[\gamma]} \exp\left(A[\gamma]\right). \quad (9)$$

It is this expression that we shall study in [2].

Acknowledgments: The author thanks Professor N.Bernardes for having introduced him to such kind of problems; to Professor W.A.Rodrigues Jr. and Professor E.Recami for reading the manuscript, discussions and suggestions; to Dr. J.E.Maiorino and Dr.A.L.T.Carvalho for helpful comments, and CNPq and FINEP for financial support.

References

[1] Vaz Jr., J., Hadronic Journal, **14**, 27, (1991).

[2] Vaz Jr., J., "On a Class of Theories of Mechanics - part II", subm. for publication.

[3] Bernardes, N., "Holotropic Theories of Thermodynamics", a series of papers to be subm. for publication.

[4] Bohm, D., *Wholeness and the Implicate Order*, Routledge & Kegan Paul, Inc., London (1980).

[5] Bernardes, N., article published in the Science supplement of Folha de S.Paulo, June 16th., (1989).

[6] See, for instance, Zinov'ev, A.A., *Logical Physics*, D.Reidel Publishing Co., Dordrecht (1983).

[7] See, for instance, Gardner, H., *Cognitive Science*, Basic Books, New York (1985), and references there in, particularly Rosch, E., "Natural Categories", Cognitive Psychology, **4**, 328, (1973).

[8] Rodrigues Jr., W.A., Scanavini, M.E.F., and Alcantara, L.P., Found. Phys. Lett., **3**, 59, (1990).

[9] Arnold, V.I., *Mathematical Methods of Classical Mechanics*, Springer-Verlag, New York (1978).

[10] von Westenholz, C., *Differential Forms in Mathematical Physics*, North-Holland Publ. Co., Amsterdam (1981).

[11] Gelfand, I.M. and Fomin, S., *Calculus of Variations*, Prentice-Hall, Inc., Englewood Cliffs, NJ (1963).

[12] Riesz, F. and Sz-Nagy, B., *Functional Analysis*, Frederick Ungar Publ. Co., New York (1965).

[13] Landau, L. and Lifshitz, E., *Mechanics*, Addison-Wesley Publ.Co., Inc., Reading, Mass. (1960).

[14] Sudarshan, E.C.G. and Mukunda, N., *Classical Dynamics: A Modern Perspective*, John Wiley & Sons, Inc., New York (1974).

[15] Saletan, E.J. and Cromer, A.H., *Theoretical Mechanics*, John Wiley & Sons, Inc., New York (1971).

RELATÓRIOS TÉCNICOS — 1991

01/91 **Um Método Numérico para Resolver Equações de Silvester e de Ricatti** — *Vera Lucia da Rocha Lopes* and *José Vitório Zago.*

02/91 **"Regge-Like" Relations for (Non-Evaporating) Black Holes and Cosmological Models** — *Vilson Tonin-Zanchin* and *Erasmo Recami.*

03/91 **The Exponential of the Generators of the Lorentz Group and the Solution of the Lorentz Force Equation** — *J. R. Zeni* and *Waldyr A. Rodrigues Jr.*

04/91 **Tensornorm Techniques for the (DF)-Space Problem** — *Andreas Defant* and *Klaus Floret.*

05/91 **Nonreversibility of Subsemigroups of Semi-Simple Lie Groups** — *Luiz San Martin.*

06/91 **Towards a General Theory of Convolutive Sets (With Applications to Fractals)** — *Jayme Vaz Jr.*

07/91 **Linearization of Holomorphic Mappings of Bounded Type** — *Jorge Mujica.*

08/91 **Topological Equivalence of Diffeomorphisms and Curves** — *M. A. Teixeira.*

09/91 **Applications of Finite Automata Representing Large Vocabularies** — *Cláudio L. Lucchesi* and *Tomasz Kowaltowski.*

10/91 **Torsion, Superconductivity and Massive Electrodinamics**
Cartan's Torsion Vector and Spin–0 Fields — *L. C. Garcia de Andrade.*

11/91 **On The Continuity of Fuzzy Integrals** — *G. H. Greco* and *R. C. Bassanezi.*

12/91 **Optimal Chemical Control of Populations Developing Drug Resistance** — *M. I. S. Costa, J. L. Boldrini* and *R. C. Bassanezi.*

13/91 **Strict Monotonicity of Eigenvalues and Unique Continuation** — *Djairo G. de Figueiredo* and *Jean-Pierre Gossez.*

14/91 **Continuity of Tensor Product Operators Between Spaces of Bochner Integrable Functions** — *Andreas Defant* and *Klaus Floret.*

15/91 **Some Remarks on the Join of Spheres and their Particular Triangulations** — *Davide C. Demaria* and *J. Carlos S. Kiihl.*

16/91 **Sobre a Equação do Telégrafo e o Método de Riemann** — *L. Prado Jr.* and *E. Capelas de Oliveira.*

17/91 **Positive Solutions of Semilinear Elliptic Systems** — *Ph. Clément, D. G. de Figueiredo* and *E. Mitidiere.*

18/91 **The Strong Coupling Constant: Its Theoretical Derivation from a Geometric Approach to Hadron Structure** — *Erasmo Recami* and *Vilson Tonin-Zanchin.*

19/91 Time Analysis of Tunnelling Processes, and Possible Applications in Nuclear Physics — *Vladislavi S. Olkhovsiky* and *Erasmo Recami.*

20/91 Procedimento, Função, Objeto ou Lógica? — *M. Cecília Calani Baranauskas.*

21/91 The Relation Between 2-Spinors and Rotations — *W. A. Rodrighes Jr.* and *J. R. Zeni.*

22/91 Boundaries for Algebras of Analytic Functions on Infinite Dimensional Banach Spaces — *R. M. Aron, Y. S. Choi, M. L. Lourenço* and *O. W. Paques.*

23/91 Factorization of Uniformly Holomorphic Functions — *Luiza A. Moraes, Otilia W. Paques* and *M. Carmelina F. Zaine.*

24/91 Métrica de Prohorov e Robustez — *Mario Antonio Gneri.*

25/91 Cálculo de Funções de Green para a Equação de Schrődinger pelo Método de Expansão Tipo Sturm-Liouville — *L. Prado Jr.* and *E. Capelas de Oliveira.*

26/91 On the Weierstrass-Stone Theorem — *João B. Prolla.*

27/91 Sull'Equazione di Laplace nell'Universo di De Sitter — *E. Capelas de Oliveira* and *G. Arcidiacono.*

28/91 The Generalized Laplace Equation in Special Projective Relativity — *E. Capelas de Oliveira* and *G. Arcidiacono.*

29/91 The Projective D'Alembert Equation — *E. Capelas de Oliveira* and *G. Arcidiacono.*

30/91 The Generalized D'Alembert Equation in Special Projective Relativity — *E. Capelas de Oliveira* and *G. Arcidiacono.*

31/91 A General Algorithm for Finding the Minimal Angle between Subspaces — *Alvaro R. De Pierro* and *Alfredo N. Iusem.*

32/91 Scalar Curvature on Fibre Bundles — *Maria Alice B. Grou.*

33/91 Sur la Dimension des Algèbres Symétriques — *Rachid Chibloun, Artibano Micali* et *Jean Pierre Olivier.*

34/91 An Inverse Column-Updating Method for Solving Large-Scale Nonlinear Systems of Equations — *José M. Martínez* em *Mário C. Zambaldi.*

35/91 Parallel Implementations of Broyden's Method — *Francisco A. M. Gomes* and *José M. Martínez.*

36/91 Equivalência Elementar entre Feixes — *A. M. Sette* and *X. Caicedo.*

37/91 Unique Ergodicity for Degenerate Diffusions and the Accessibility Property of Control Systems — *Luiz San Martin.*

38/91 Unobservability of the Sign Change of Spinors Under a 2π Rotation in Neutron Interferometric Experiments — *J. E. Maiorino, J. R. R. Zeni* and *W. A. Rodrigues Jr.*

39/91 Disappearance of the Numerically irrelevant Solutions (NIS) in Non-Linear Elliptic Eigenvalue problems — *Pedro C. Espinoza.*

40/91 Positive Ordered Solutions of a Analogue of Non-Linear Elliptic Eigenvalue Problems — *Pedro C. Espinoza.*

41/91 On von Neumann's Variation of the Weierstrass–Stone Theorem — *João B. Prolla.*

42/91 **Representable Operators and the Dunford-Pettis Theorem** — *Klaus Floret.*

43/91 **Simultaneous Approximation and Interpolation for Vector-Valued Continuous Functions** — *João B. Prolla.*

44/91 **On Applied General Equilibrium Analysis** — *José A. Scaramucci.*

45/91 **Global Solutions to the Equations for the Motion of Stratified Incompressible Fluids** — *José Luiz Boldrini* and *Marko Antonio Rojas-Medar.*

46/91 **A characterization of the set of fixed points of some smoothed operators** — *Alfredo N. Iusem* and *Alvaro R. De Pierro.*

47/91 **Lyapunov Graphs and Flows on Surfaces** — *K. A. de Rezende* and *R. D. Franzosa.*

48/91 **On the Multiplicative Generators of Semi-Free Circle Actions** — *J. Carlos S. Kiihl* and *Claudina Izepe Rodrigues.*

49/91 **A Priori Estimate and Existence of Positive Solutions of Nonlinear Cooperative Elliptic Equations Systems** — *Marco Aurelio S. Souto.*