FROM THE HISTORY OF PHYSICS TO THE DISCOVERY OF THE FOUNDATIONS OF PHYSICS

ANTONINO DRAGO

Translated from the Italian by David and Michael Braithwaite

CONTENTS

INTRODUCTION

1. To whom is this book addressed?

First of all, who are the potential readers to whom this book is addressed?

Several paragraphs involve university level Mathematics and an understanding of the themes of this book presupposes that we know the fundamental elements of university physics and mathematical analysis.

It is therefore addressed to science graduates: firstly, to professional scientists, engaged in pure or applied research, for whom this book will provide an overview of the subjects studied at the University in a specialized and without a unitary connection; also to teachers of Physics, Chemistry, Computer Science, Mathematics, who can find in this text new cultural and educational tools for their profession which will enable them to convey a cultural vision of science and integrate the teaching of science with other school subjects. It will also be of interest to science undergraduates who are eager to expand the notions and technicalities of their studies with a vision that allows them to succeed in the cultural adventure they are undertaking; Foundations of science through this text but is of the last part concerning teaching methodology provides teachers with new insights that will enable them to improve on traditional teaching.

A reader who is eager to learn the average culture and has received his or her notions of mathematics from scientific high school or little more, would have to make the minimum effort required to acquire a greater mathematical and physical culture in order to follow, with knowledge of the facts, the text's review of the facts of the main physical theories, involving their theoretical and mathematical foundations. With such an effort, all those who are educated in philosophy should be able to enter into the exposition of the topic and grasp the new proposal of this book: to offer an understanding of a theory of Physics that includes its foundations.

2. The novelty of this text of History of Physics: the search for the Foundations

Here we pursue the strategy of achieving a deeper understanding in the study of Physics. But how? Not so much with the addition of more and more physical discoveries (almost as if the new ones explain the Foundations of the previous ones, as the usual refresher courses for teachers hope to do), or with the study of an increasingly advanced mathematics (the dream of set theory to interpret formally any scientific enterprise), or with the study of classical mathematical logic (which would provide the keystone of scientific construction), but considering Physics in the context of the History of Physics by which we mean the experience of scientific theories accumulated over the centuries according to an evolution that is to be interpreted, as the meritorious works of Koyré and Kuhn began to do.

Let us then consider the History of Physics. In recent times teachers have shown greater interest in the History of Physics and attributed greater importance to it in the scientific-mathematical curriculum. High school physics ministerial programs also hint at a teaching approach with a historical character. It is clear that learning Physics is much more interesting if each concept and physical law is placed in the context of the historical process that determined it, but to gain a full understanding of the history of Physics as a whole, the problem of its Foundations needs to solved: that is, we propose a History of Physics that goes in search of the Foundations of Physics, in order to achieve an understanding of the meaning of five centuries of that scientific adventure that has become the new and solid rationality, apparently superior to any other previous one. This means that it is necessary to study not only and not so much the History of Physics, but also the History of the Foundations of Physics.

We note that university-level teaching neglects the latter issue; it does not enable students to achieve a comprehensive interpretation of the theories studied or, in particular, to recognize the foundations of each theory that they study in different courses (perhaps in different ways). E.g. the degree course in Mathematics does not lead to an understanding of the relationship between Analysis and Algebra or Physics degree course do not develop an understanding of the relationship between Mechanics and Thermodynamics. Formulae, techniques and also theories are taught, but without philosophical implications, which, even though they always informed the beginnings of a theory, are considered mere scaffolding to be dismantled, once the formal results have been obtained, so that all that remains is what is objective and unquestionable. This attitude reduces the subject matter to its skeletal structure, which in fact is univocal and excludes any doubt about whether there might be alternatives to the framework presented. That is, it suggests that the foundations of Physics are simply mathematical techniques and the corresponding experimental practices, so that these supposed foundations are conceived as unquestionable and monolithic. This is due to the fact that for centuries no attempt (philosophical or cultural) to conceive the Foundations of Physics has survived the advances of science. In particular, the crisis of the early 1900s led to new theories which broke their moorings with respect to any previous conception, sailing out into an open sea in which any compass other than scientific inventiveness is dispensed with by the physicist.

How then will the search for the Foundations be carried out? To begin with we can take up again the questions posed above, those concerning the relationship between Algebra and Analysis in Mathematics or that between Mechanics and Thermodynamics. We immediately realize that it is difficult to answer them because both concern a level of reflection that is superior to the single theories; they concern the relationships between theories; that is, they require a ‘supertheory’, which of course will not necessarily to be scientific; therefore they concern a level of reflection to which university teaching, exclusively technical, has not prepared us at all; therefore, although we have studied the subjects at University well, we do not have the tools to deal with the above questions.

But we can find a key to the problem by asking a more restricted, and therefore more approachable, question: why both Physics and Mathematics degree courses have two examinations in Mechanics.

First Mechanics is studied in Physics I and then again Mechanics as rational Mechanics.

If out of around twenty exams needed to graduate in Physics two must cover the same topic, it means that the adjective “rational”, which is added in the title of the second examination, is highly significant within the curriculum, that is, it has a fundamental value for Physics as a whole.

If we look for an explanation for this in the contents of rational mechanics, we find that this subject expresses a “superior” point of view; because while the mechanics of Physics I is linked to experiments and does not want to dispense with the inductive capacities of the scientist and the student, the latter uses mathematics intensively, both to define the initial concepts (point, time, force, mass, etc.) and to apply the most advanced mathematics possible (for example, in the past calculus of variations, currently differential geometry) to solve problems; moreover, models of reality are constructed in accordance with a primarily mathematical consistency, so that only theoretical schemes of deductive nature are obtained.

Thus rational mechanics includes its particular conception of the foundations of physics: they are constituted by advanced mathematics. Galilei was already saying this with the famous phrase: “The great book of the Universe [...] is written in squares circles and triangles”; Descartes mantained a similar opinion; then Newton constructed through an advanced mathematics the laws governing the motions even in the heavens; Lagrange made his mathematics an instrument of near omnipotence in the solution of the problems of Physics; Laplace made calculus an instrument of almost theoretical omniscience; the mathematical physics of the late nineteenth century conceived Physics and more in general the entire science of nature by means of differential equations, the solutions of which represented theories, already constructed or potential.

But then, if this was the case, why is there no teaching of rational thermodynamics, or of rational electromagnetism? And then, if the mathematics (of analytic geometry and of infinitesimal analysis) are the foundations and the premise for physics, why do we have important theories, such as thermodynamics, that do not fit into this mental scheme?

Above all, it is not possible to answer the following question: why this conception of the Foundations of Physics was so resoundingly disavowed with the discovery of light quanta (after the emergence of matter quanta: electrons, atoms, molecules)? These imposed a discrete vision of reality, that is, outside the advanced mathematics of continuity which had not only served as the basis for all theoretical physics for two and a half centuries but had also been the basis of Euclidean geometry for two millennia. Why had mathematical physics, which seemed to have grasped the foundations of physics, not predicted the quanta?

Does accepting quanta mean returning to the pre-Galilean mathematical backwardness? This was one of the most perplexing questions facing those physicists who had to build Quantum mechanics in the early 1900s. They seemed to have found an answer by finding a way to rework the schema of the continuum (Schroedinger’s equation), but they could not erase the discreteness of the atomic spectra or of the eigenvalues of a discontinuous quantity; and that subsequently the solution found was partial is demonstrated by the fact that after about 1960, in Physics the most advanced mathematics has become the algebraic mathematics of symmetries, not the analytical mathematics of differential equations.

As can be seen, by asking that question about the doubling up of the teaching of

Mechanics in the two degree courses (Mathematics or Physics) we realize that, implicitly, the teaching of Physics at university has a particular conception of the Foundations of Physics. It arose with Lagrange and Laplace in the late 1700s and then was developed above all by the French scientists of the early 1800s to become the conception of the foundations of science at the end of the 1800s.

When in the early years of 20th century it was radically called into question by the discovery of quanta, a solution was achieved by the highly abstract mathematics of Quantum mechanics (Hilbert space) and the differential mathematics of general relativity; therefore, university teaching as a whole suggests that the problems of the new physics have been solved using a mathematics that was even more abstract and sophisticated than that used in the theoretical physics of the nineteenth century. And obviously it continues to believe that the teaching of Physics at high school is essentially incapable of presenting advanced mathematics; and therefore it is conceived as an inevitably repetitive teaching of the simplest things alone, viz. those that are at the level of the little mathematics that the students know. This is a purely mathematical and highly abstract way to conceive the Foundations of Physics.

3. *The novelty of the method: the comparison of the various formulations of the same scientific theory*

But in this way we have obtained a negative conclusion, that is, how the Foundations of Physics should not be conceived. How can a positive result be found? It is the teaching of rational mechanics itself that suggests a strategy for a new understanding of the foundations of physics. Indeed, it is also the only teaching that presents several formulations of the same theory: in fact it makes known the principle of Maupertuis, the principle of virtual works, the Lagrangian, the Hamiltonian. This part of university teaching provides a glimpse of the theoretical world that stand above the single physical theory: it is the world of the different formulations of the same theory; that is, they concern the same field of phenomena of nature and therefore experimentally verified laws are always the same, but each of the aforementioned formulations sees as fundamental different concepts (action, work, T-V, energy H) and uses a mathematics that is not the same mathematics as that of the other formulation (calculus of variations, minimization of work, partial differential equations). This then can be that level we were looking for to understand the fundations of scientific theories; it calls for a comparative study of several formulations of the physical theory considered.

It is here that Ernst Mach's programme played a crucial role. He was the first to present the history of a physical theory, Mechanics, starting from a concept (work) which differs completely from the Newtonian concepts (force, space and time, but not work or energy in the original formulation). As a good energetist he wanted to show that there was actually an alternative to the Newtonian paradigm actually based on a principle that current teaching methodology do not even name (in mechanics), the principle of the impossibility of perpetual motion. The result is a profoundly new vision of mechanics, to the point of being accepted for a century as the possible alternative to the dominant vision (inspiring Planck to go beyond classical physics, and Einstein to introduce a critique and an alternative to the Newtonian theoretical system).

But the Mach that devoted so much space to illustrating all the formulations of mechanics, when considering Lagrange's formulation, which introduces advanced mathematics on a huge scale, expresses the opinion (not argued) that they are all equivalent, except for differences of a technical nature. Thus his program of emphasizing an alternative to the Newtonian paradigm was blocked by the myth of the extraordinary power of higher mathematical tools for theoretical physics (this was the same myth of Comte who moreover saw in the Lagrangian the apex of mechanics and all physics).

Mach's program must in my opinion be taken up again and reformulated in current terms. First, from a study of the History of Physics, we see that for two and a half centuries every physical theory has been formulated in several ways, even if university teaching, to simplify the subject to students, presents only the most important formulation (and moreover not with complete honesty, keeps silent about the fact that it is not the only one possible). So there is a vast subject of study in addition to that of studying Newtonian mechanics only. Thus here we propose an examination of the history of Physics to discover, first of all, the various formulations of each physical theory taught at university.

With this research program, History is no longer an investigation into past curiosities, or scientists' biographies, or the unstoppable growth of theories, but rather the discovery of various formulations, ignored by university teaching. In particular it is an investigation into radically different formulations from the dominant one, formulations that have been either neglected or undervalued due to preconceptions typical of the past. Now such alternatives can be re-evaluated and compared with the usual formulations, so as to obtain a broad and articulated cultural picture of Physics and from the variety of their different characteristic features derive the foundational characteristics of all of them.

Given the plurality of formulations of the same theory, it will be necessary to explain how they may differ. With respect to the same experimental laws, they present different concepts, different techniques and different methods; these differences between the formulations certainly constitute profound theoretical aspects that are more general than the theoretical contents of each individual formulation and therefore certainly represent fundamental aspects of the particular theory so variously formulated.

This comparison invites the reader to rise to a new level, which is intermediate between that of a theory of Physics (with its experiments, concepts – such as space, time, force, field, entropy, etc. –, its mathematical techniques – algebraic or differential equations –, its individual laws – eg $f = ma$ – its principles), and the very general level of philosophy of knowledge (including philosophical ideas that may be seen to form the basis of physics – determinism, experimentalism, rationalism, etc.. –); rather it is the level of a comparison of various formulations of a given physical theory.

We are thus taken to a level that still concerns theoretical physics, but which is much more general than that of a single formulation, because it opens our mind to an investigation of Physics from a cultural perspective; now Physics is no longer conceived as a closed system of experimentally confirmed ideas, but is also a theoretical system with a variety of possibilities. On the other hand, this type of investigation, which is more general than that within a specific formulation, does not lead us to trespass on the terrain of traditional philosophy, - where the crisis of the philosophy of knowledge is still unresolved -, mainly due to its inability to account for tumultuous growth of science in past centuries.

In particular, a comparison between the different formulations will reveal how fundamental the physical relationship between physics and mathematics and between physics and logic are. It may seem an almost banal discovery but let us remark that these two relationships are not explored by the teaching of Physics at university. Although it claims to promote the study of everything that is scientific, in reality, these objective and formally defined relationship are completely disregarded. By filling this void, the teaching of physics, weighed down today by the encyclopedic accumulation of disparate experiments and laws, will regain a cultural sense, where culture means the ability to see the objective aspects of reality from a higher and broader perspective.

Proceeding in this way we will find that the Foundations of Physics are, almost trivially: Mathematics and Logic. But in reality this proposition will not be trivial because it will be seen that even in mathematics and logic each theory has multiple formulations; and therefore the Foundations of a physical theory consist of a structure of two relations with Mathematics and Logic, where each of the latter branches into various formulations. It will be by no means trivial to discover the differences among the various formulations of each of the three branches of science.

Finally, based on the analysis of many physical theories, it will be concluded that the Foundations of every scientific theory are constituted by two fundamental dichotomies: the dichotomy regarding the type of infinity, actual (AI) or potential (PI) and the dichotomy of the type of organization of the theory, deductive from few axioms (AO) or problem-based (PO) (or equivalently on the type of logic, either classical or intuitionist). This conception of the fundations of Physics also derives from the lesson we received from the crisis of the early 1900s: since then, the AI of the mathematics (eg the infinitesimals) of classical physics contrasted with the PI mathematics of the discrete included by Quantum theory; the AO of Newtonian theory (with its axioms of absolute space, absolute time and cause-force) with the PO of both special Relativity, based on a problem (how to reconcile Electromagnetism with Mechanics) and the first historical formulation – by Heisenberg – of the theory on quanta.

Once the Foundations of Physics are clarified, then the themes of the History of Physics can be expounded succintly, leaving aside the myriad biographies, anecdotal facts, the failed experiments, the controversies, etc. which may also attract the interest of students, but which are highly distracting, because they would require ten times more time than is currently available. The history of physics presented here, based on its foundations, is not only shorter, but is also the best and most appealing History of Physics, because it is an account of the crucial historical developments according to a logic which is intrinsic to the subject of study.

THANKS

I thank Fabio Bevilacqua, Danilo Capecchi and Pietro Cerreta for their precious advice. The text was first revised by Raffaele Pisano. I am very grateful to Nicola Amoruso who further revised the text and improved it.

BIBLIOGRAPHY

It is conceived as an introduction not to research on the topics covered, but as extensions of them from a cultural point of view. References to specialist writings on the various subjects of the following text can be easily retrieved with an internet search engine.

The texts of Koyré and Kuhn of History of Physics are mentioned below. Then, to widen the study of the evolution of Physics, C.C. Gillispie, *On the Edge of Objectivity. An Essay of the History of Scientific Ideas*, Princeton: Princeton U.P., 2016. A classic in the history of physics that is also a critical work that played a decisive role in the subsequent history of theoretical physics is E. Mach's book *The Science of Mechanics. A Critical and Historical Account*, Open Court, Chicago, 1960. More widely, the biography and scientific works of each scientist can be consulted on the *Complete Dictionary of Scientific Biography*. New York: Scribner & Sons, 2007, 16 vol.s.

For the History of Mathematics, a very useful and lively book is that of C.B. Boyer, *History of Mathematics*, Wiley, New York, 1968; a much broader and more complete one is that of M. Kline, *Mathematical Thought from Ancient to Modern Times*, Oxford U.P. Oxford, 1990; or its synthesis, more conceptual and philosophical: *Mathematics. The loss of certainty,* Oxford U.P., Oxford, 1972. On the Foundations of Mathematics see K. Meschkowski, *Evolution of Mathematical Thought*, Holden-Day, New York, 1965. On the revolution of non-Euclidean geometries a classical introduction is R. Bonola, *Non-Euclidean Geometry: A Critical and Historical Study of Its Development*, Dover, New York, 2018; a more technical but very informative book is that by B. A. Rosenfeld: *History of Non-Euclidean Geometry*: *Evolution of the Concept of a Geometric Space*, Springer, Berlin, 1988.

On classical logic there are innumerable works. On the non-classical logic a quick introduction is that of K. Meschkowsky's book, already mentioned. At a specialist's level a complete history of logic is that of D. Gabbey and F. Guethner (eds.), *Handbook of History of Logic*, Elsevier, New York, vol.s 14.

CHAPTER 1

HISTORY AND FOUNDATIONS OF THERMODYNAMICS: THE DISCOVERY OF THE TWO DICHOTOMIES

In this chapter we begin an exploration of the Foundations of Thermodynamics, a physical theory that more than any other lends itself to reflecting on possible alternatives in the current Mathematics of theoretical physics. The mathematics of this theory is in fact relatively simple so that it is easy to understand the origin of the problems on this point.

1.1 The importance of the relationship between Physics and Mathematics

> "The unreasonable effectiveness of mathematics in the natural sciences" E. Wigner (1970)

Since the nineteenth century, science has been conceived, as if the experiment completely exhausted its contents (positivist conception of science). However, this was not Galilei's conviction. He claimed that it was necessary to make "reasoned experiments, that is, to link the questioning of experience to a hypothesis, which concerned not only intuitive concepts, but also e.g. Mathematics. For example, he worked hard to understand whether the law of falling bodies was a proportion between, on one hand, the space covered and, on the other, either the time or the square of the time. Yet, even present philosophers of science conceive science reductively; they discuss intuitive concepts, or the relationship between a law and experiments, but they do not take into account the mathematics used, in part because according to the positivist conception that has long dominated, mathematics is merely a functional, philosophically neutral tool of physics. We will see that compared to previous philosophical discussions on science, the highlighting in what follows of the role of mathematics in theoretical physics will give additional understanding that is of crucial importance.

Modern physics was born on the scientific basis provided by Euclidean geometry. With Descartes geometry became analytical; space was mathematized in all its individual points and provided the ideal place where physicist's mind could operate. So that since the time of Galilei and Descartes physics took the measurement of lengths in space as its cultural background; the passing of time was also formalized with a continuous variable, that could be represented on a circle or on a straight line. The same applies to the results of measurements of any other variable, as e.g. the values of temperature measurements on the segment of a thermometer. Space and time became the main founding concepts of theoretical physics and, on this basis, it was thought that every physical quantity varied in the mathematical continuum. This was the origin of the preconception that physical theory concerns only continuous quantitative aspects of phenomena contained in continuous space and time.

Therefore until 1900 all physicists, with rare exceptions, thought that the continuum of real numbers was the only mathematical basis for describing reality, and indeed they competed with mathematicians to derive the characteristic properties of the continuum, because these were of common interest for the more productive advancement of both sciences. For a long time mathematicians believed that, among the subjects they dealt with, those useful to Physics constituted the most reliable and relevant part of the whole Mathematics. This close link between Physics and Mathematics was long considered without possible alternatives.

The first doubts about it arose with the birth of non-Euclidean geometries, since these made it clear that space could be conceived in radically different ways. Moreover almost all philosophers who supported the atomist conception cultivated the doubt that the continuum was not indispensable for physical theory. But it was only in the early 1900s that it was proved experimentally (optically, by means of the ultra-microscope) that matter is composed of molecules (as Chemistry had already hypothetically suggested); in addition, theoretical physics introduced light quanta. Thus all nature appeared corpuscular and the mathematics used until then seemed completely abstracted from this reality. Without going into the problem now, suffice it to say that this crisis was then superseded by theoretical physics whose leaders were successful in reiterating the previously dominant model of continuity.

However the alternative of the discrete did not die. The great physicist M. Born in fact affirmed in 1949:

> The concept of the point of the mathematical continuum has no direct meaning: physically it is not significant to say that the result of a measurement is an irrational number ($\sqrt{2}$ o π)... So that today the propositions of the type: the result of a certain measurement is π should be eliminated from physics. This does not mean eliminating the concept of the real number, which remains indispensable for applying infinitesimal analysis. Rather it is necessary to describe the physical quantities hazily with the help of real numbers.[1]

(He proposed to introduce the probability, a suggestion that was not followed later). Later, in 1970, another great physicist, Wigner, wrote an article whose title is very significant: "The unreasonable effectiveness of mathematics in natural sciences". Wigner marvels at the fact that mathematics is such a suitable tool for theoretical physics, both in the invention phase of theories and in the convergence of the results found by the theory with reality.

> We are in a position similar to that of a man who was provided with a bunch of keys and who, having to open several doors in succession, always hit on the right key on the first or second trial. He became skeptical concerning the uniqueness of the coordination between keys and doors.

Wigner concludes that:

1) the enormous usefulness of mathematics in natural sciences borders on the mysterious

---

[1] M. Born, "Continuity, Determinism, and Reality", *Danske Videnskabernes Selskab Mathe- matisk Fysiske Meddelelser*, **30** (1949), 1- 26.

("a miracle"), because there is no rational explanation for it all;

2) there exists the problem of the non-uniqueness of the mathematics of our physical theories; but this, says Wigner, is "the nightmare of theoretical physicists "![2]

The synthesis between discrete and continuous, carried out by Quantum mechanics, has not therefore solved all the problems; physicists still question the meaning of mathematics in their theories and in particular on the use of the continuum. We will return to the synthesis of Quantum mechanics at the end of this book, when we have accumulated more experience concerning the relationship between physical theories and mathematics.

Meanwhile, we note that all the aforementioned concerns the translation into Mathematics of the position of a body in space, of the many values of a physical quantity, or the translation of experimental data collected into a mathematical formula that expresses them. But in the history of Physics the relationship between Physics and Mathematics has become symbiotic for another reason, that of having introduced a formidable calculation tool into theoretical physics: infinitesimal analysis. Although the real first physical theory (mathematized) was the Descartes' geometric optics, usually we consider to be the first theory that which in reality was born fifty years later, i.e. Mechanics, because Newton founded this theory basing it on that important mathematical tool.

It was precisely by means of infinitesimal analysis that Newton's mechanics constituted itself into a theory of great theoretical and predictive capacity; it is with this Mathematics that 1) the set of infinite possible solutions is summarized by a single differential equation, 2) every solution gives the evolution of the system in all the past and the future, and 3) the theory structurally always provides new laws of nature through the solutions of suitably modified differential equations; since then, physical theory could also be developed for calculations only. With this instrument the physicist started to progress towards the conquest of new and more powerful theories, with which he acquired a hopefully all-inclusive theoretical capacity. Therefore mathematical analysis has become (as Born says) an "indispensable" component of theoretical Physics, thus creating a physical-mathematical link, which from Newton onwards seemed unquestionable. At the end of the 19th century in Physics the mathematical physics that studied differential equations applicable to reality represented the culmination of this perspective.

Let us try to understand this mathematical theory better by studying its history

1.2 For a more correct history of classical mathematics

> There is no religious denomination in which the misuse of metaphysical expressions has been responsible for so much sin as it has in mathematics. (L. Wittgenstein, *Culture and Value*, 1929)

[2] E. Wigner, "The Unreasonable Effectiveness of Mathematics in the Natural Sciences," C*ommunications in Pure and Applied Mathematics*, 13 (1960) I , 1-14.

Usually one fact is overlooked: that the analysis we are familiar with today is not at all the analysis as it was conceived and used for a century and a half after its birth. In reality, analysis was born according to three foundations, of which that of infinitesimals was only the most important one besides two minor foundations, of which one was that of Newton's "first and ultimate reasons". This expressed, though vaguely, the mathematical process of a limit as a ratio (or "reason") of two increments (e.g of space and time to obtain velocity). The second minor foundation was intended to be concrete: Newton considered mathematical variables as mechanical quantities: the derivative at a point of a curve-function was considered the speed of a material mass-point that travels through it at that point. This foundation did not refer to actual infinity (i.e. the idea of attributing concrete reality to the infinite; for example, to consider the point at infinity of a straight line on a par with all the others), but only to potential infinity (which is that of counting with natural numbers, for which the infinite is the unlimited: one never reaches the number higher than all the others). (Even Leibniz, in the last period of his life, wanted to base analysis on potential infinity alone; but he remained uncertain, so much so that he raised the dilemma between this infinity and actual infinity to "a labyrinth of human reason").

What had the lion's share was the third foundation of analysis, certainly Leibniz's and perhaps also Newton's. It was explicitly based on the concept of infinitesimal dx, understood as the inverse of actual infinity; hence, we were referring to an idealistic abstraction – actual infinity - to which, moreover, an operation (inverse) was incorrectly applied; it was therefore a semi-scientific, metaphysical entity. It caused confusion in the foundations of all mathematics because, accepting it as a mathematical entity, it was not known what number it was: in fact it was not zero (otherwise the calculations on it would have been useless), nor was it any number different from zero (otherwise it would have been a quantity like all the others with which we do the usual calculations); therefore it constituted an exception to the usual definitions of the Mathematics of the time, all well defined with precise operations of calculation; whereas, introducing that strange way of operating in Mathematics one no longer knew, for example, whether equality was still valid when a member of it included an infinitesimal. The problem was roughly solved by talking about a "quantity that tends to zero", although the word "tends", for a single number has no mathematical meaning.

In the first decades of the 1800s Cauchy began a long reform of the foundations of analysis to make it "rigorous", that is without infinitesimals. He relied on the concept of limit, defined by the technique of ε-δ that we learned at high school. This reform was completed in the 1870s, when Dedekind and Weierstrass defined real numbers precisely in the way it was taught in high schools using the same technique as ε-δ. On the basis of this technique infinitesimals were eliminated and as a result many aspects of higher analysis were clarified and significant advances accomplished.

But, in fact, this new analysis, called "rigorous", still made essential use of actual infinity. This was emphasised very early, even if textbooks on Mathematics (and also on History) do not usually remember it. In 1882 the German mathematician Paul Du Bois Reymond pointed out a criticism that may be summarized as follows:

> Large or small, an interval always remains an interval between its two extreme points. If abruptly, without a logical justification, instead of an interval [as is determined by the ε-δ

technique], we make a point [the final value of the limit operation] appear, this action arbitrarily produces a new concept without deducing it from what preceded it. On the other hand, a slow and gradual fusion of two [above-mentioned] points (the ends of the range that we are going to restrict (by increasing approximations) into one, has absolutely no mathematical sense. Either the two points are separated by an interval, or there is only one point; an intermediate situation cannot exist.

That is, the usual reasoning contains an act of thought with which (according to the optical illusion of two parallel lines that narrow more and more as their points are seen at an increasing distance from the observer) within an interval the desired limit point is selected and is thus exactly identified, excluding all the infinite other possible points of the same interval. But this act is merely idealistic. In fact, since the time of Cantor we know that there are as many points in an interval, however small it is, as in a long segment, in a line and in a space, even in *n* dimensions space; choosing a point with the technique of the $\varepsilon - \delta$ then is a worse problem than finding a needle in a haystack. In other words, we can describe the usual definition of passing to the limit as follows: one appeals to the experience of an actually performed calculation of approximation for some very small value of δ; this, however, is then incorrectly extended to *all* its possible values, including the value $\delta = 0$, that is, that, and only that, which corresponds to having restricted the interval at the exact limit point.

Only in some special cases (such as that in which it is known that "the limit is an integer") does our algorithm approximate gradually to the limit and exactly determines this limit value. Therefore in the process of transition to the limit the "rigorous" analysis, which claims always to obtain the limit point through technical means alone, in reality uses actual infinity to "leap" to the only final point. This fact implies that his process is not effective; in other words, actual infinity has remained hidden within the definition of Cauchy's limit which although he purported to suggest a non-metaphysical mathematical process, superseding the infinitesimal. Only apparently, therefore, did the reform of analysis, carried out in the sec. XIX, exclude actual infinity from Mathematics.

Not only Du Bois Reymond but also others opposed "rigorous" Mathematics and more in general the use of the actual infinity. In this regard, many books on the History of Mathematics quote the Kronecker's famous sentence: "God made the integers, all else is the work of man" but because they do not agree with him, but because they associate Kronecker's rational opposition with the ancient hostility of religion to modern science. But opposition to actual infinity did not remain restricted to Kronecker; it was supported by mathematicians of the caliber of Poincaré, Borel, Lebesgue. However, this opposition did not detract from the absolute confidence with which at the end of the nineteenth century "rigorous" mathematics imposed itself as the final solution to the problems that had until then been unresolved, and, since the beginnings of the last century, was considered the only possibility of rationally founding infinitesimal analysis.

Meanwhile mathematical theories had proliferated. Set theory had claimed to enclose them all in its formalism. However, mathematics went through a deep crisis in the early 1900s. The crisis was due to the fact that contradictions were discovered in Russell's set theory (called "paradoxes" if it is hoped to solve them easily; or "antinomies" if they are believed to be intrinsic to the theory, such that the theory cannot avoid them). A contradiction within a theory implies that it must be abandoned, because according to the laws of logic, any proposition follows from contradiction; as a consequence, it is then impossible to distinguish between true propositions and false propositions, thus rendering any logical discourse pointless.

With the birth of the antinomies, the great project of the original set theory was blocked. Moreover, owing to this crisis, Russell stated: “Mathematicians no longer agree on how to define the number 1”. In fact today we know: 1) the intuitive definition, learned in elementary schools, or in any case in similar intuitive ways; 2) Peano’s definition, which sees 1 as the successor of 0; 3) the definition of set theory, which defines 1 as the common element of the class of sets with a single element (this definition is impredicative, ie it refers to a totality of which the defined object is a part; nevertheless this it is the only definition that set theory can give: modern mathematicians maintain that impredicative definitions have the right to exist in mathematics, as long as they do not give rise to contradictions) and still more definitions ... So that Russell insisted: “Mathematics may be defined as the subject in which we never know what we are talking about, nor whether what we are saying is true.”

Most scholars of set theory reacted by redefining set theory; this gave rise to what was called ZFC[3] (Zermelo and Fraenkel); C stands for Zermelo’s *choice axiom*, which in classical mathematics has been used from the beginning of the 20th century: “Given an infinite family of infinite sets, an element can be chosen from each set and with these representative elements form a new infinite set.” Who has actually done this in his life? Nobody. One can only “think” it, whatever this “thinking” means. It is natural to ask: is it correct to introduce pure ideas into mathematics? Various mathematicians reject this axiom while classical mathematicians maintain it, using it, for example, when they demonstrate the well-known initial theorem of Analysis I, that of Heine-Borel (“A pointwise continuous function on a closed and limited interval is uniformly continuous there”).

However the ZFC theory, besides being highly criticized, is too difficult. It is therefore not taught at university, so that the set theory that is taught, called “naive”, is the original one, subject to Russell’s contradictions, but these are hidden by the teaching methodology.

In 1905 Brouwer proposed an “intuitionist” foundation of Mathematics. He was the most courageous, brilliant and tenacious supporter of a general alternative to the mainstream. He based Mathematics on the (temporal) intuition of numbers (the number two as the two parts of the flow of time as divided by an interior act of a mathematician establishing the “now” and so on for the other natural numbers),

> (Mathematics arises when the subject of **two-ness**, which results from the passage of time, is abstracted from all special occurrences. The remaining empty form [the relation of \(n\) to \(n+1\)] of the common content of all these **two**-nesses becomes the original intuition of mathematics and repeated unlimitedly creates new mathematical subjects; dalla *Stanford Enc. of Phil.*).

according to which Mathematics is only that experienced by the intuition of the

[3] On this subject see M. Kline: *Mathematical Though from the Ancient to Modern Times*, Oxford: Oxford U:P:, 1978.

mathematician, not that conceived in the abstract, formally. [4]

Others also sought alternative solutions to classical mathematics. In 1918 H. Weyl proposed an "elementary mathematics" that corresponded to the essential finiteness of physical operations, but in addition the process of extrapolation from a series of numbers, in the same way as the theoretical physicist extrapolates a function-law from a series of experimental data. In 1920 Duhem predicted an entire new Mathematics, closer to Physics but also more sophisticated: the "Mathematics of approximation", which coincides in fact with today's constructive mathematics, which we will examine shortly. In 1936, the physicist Bridgman tried to define a set theory in a strictly operational way.[5]

Despite opposition, Brouwer's ideas eventually gave rise to an entire mathematics that diverges from "rigorous" mathematics even in the definition of a real number. In the following we will not consider Brouwer's intuitionist mathematics, which in some ways is difficult, but briefly, its reduction to "constructive mathematics" by Bishop (1967).

To illustrate it, let us recall the definition of Cauchy of a real number: we must give two series of numbers a_n and b_n such that, for any small ε at will a n_o can be found such that for all $n>n_o$ the inequality $|a_n - b_n| < \varepsilon$ holds.

To define the limit of a function $f(x)$, let us say that for any small number ε at will a number δ can be found such that for $|x- x'| < \delta$ then

$|f(x) - f(x')| < \varepsilon$.

But these definitions do not specify:

1° whether the series or function was actually calculated or not:

2° if the words "can be found" (replaced in some texts by the word "exists") indicate a merely potential idea, or an actual process of calculation; in the second case we should say "the way to calculate must be found"; which means that we need to find an algorithm that actually approximates to the result; for example, long ago algorithms were found to calculate $\sqrt{2}$, π, e and and many other irrational numbers; but such an algorithm cannot be found for any number, series or function, because the set of algorithms (being finite sets of finite operations) can be put in one-to-one correspondence with the set of integers, whereas the real numbers cannot, i.e. they have a higher cardinality and higher cardinalities have series and functions;

3° as Du Bois Reymond observed, there is no guarantee that it is possible to specify an exact single number only from its approximations.

It is therefore clear that constructivists require that from the definition of a real number the word "exists" is always replaced by the words "can be calculated". Hence it is necessary to know the algorithm to calculate the sequence or the functions as well as the algorithm for calculating the increasing approximations; and we must be satisfied with only approximations to the final result of the exact real number (which, however, can be exactly determined under specific additional conditions; eg, that it is an integer).

Consider other differences between classical mathematicians and constructivist

---

4 See for an elementary introduction K. Meschkowski, *Evolution of Mathematical Thought*, Holden-Day, New York, 1965 and for an advanced reflection on the subject M. Kline, *Mathematics. The loss of certainty,* Oxford U.P., Oxford, 1972.

5 P.W. Bridgman, "A physicist's second reaction to Mengelehre", *Scripta Math.* **2** (1934) 101-107, 224-234.

mathematicians. Let the number $a=0, a_1 a_2 a_3 a_4 a_5$... be given where

a_n is equal to 0 if $2n = p_1+p_2$ with p_1 e p_2 two prime numbers, otherwise it is 9.

The number a is well defined constructively, because there is a precise algorithm that calculates it, digit by digit. In fact the property $2n=p_1+p_2$ holds for all the numbers so far calculated (the test can be done for some n), but we do not know if it holds for any natural number (Goldbach conjectured it in the XVIII century) because today we do not have a theorem that proves it or falsifies it, nor do we have a counterexample to the conjecture). As a result we have a real number a of which we do not know all the digits a_n, but only a large number of them; in particular we do not even know where it is located exactly on the line of real numbers (although it is certainly very close to zero: it can be just zero, but it can also be different from zero). This type of number has been called "fugitive" because of the lack of a precise location. Instead for classical mathematics (both "rigorous" and that of the infinitesimals) this number is well defined and well placed, it does not matter if we humans do not know where it is located; in the world of mathematical ideas the number is totally decided and therefore it has its precise place on the real number line. In summary, we can say that for classical mathematics a mathematical entity exists if in the (Platonic) world of numbers it does not lead to contradiction; while for constructive mathematics, a mathematical entity exists only if there is a computational algorithm that constructs it.

If then in the previous example of a real number the definition is slightly changed then there is an even more radical divergence. We define $a_i = 0$ if for all i Goldbach's conjecture holds, 9 otherwise. Here we do not know how to calculate the number a, from its first digit after 0. Therefore constructivists reject this number; while classical mathematicians maintain that it is well defined and therefore exists. As for the single numbers, the divergences over functions and the functionals (functions of functions) are non-denumerable (even if they do not affect the numbers and the functions that are most used in the practice of Mathematics, which has a solid correspondence with reality).

Moreover it can be seen that some basic theorems of classical analysis do not hold in constructive mathematics. For example, Weierstrass's theorem: "A closed and bounded infinite set of points has at least one point of accumulation". The simplest proof is the one that proceeds in the following way. We represent this set of points on a segment AB of a line. We cut this segment in half; then infinite points will be either to the right or to the left. If, for example, they are on the right, then we split this segment in half again and proceed in the same way for every further half part which is found to have infinite points. Thus we come to delimit the point of accumulation more and more closely; in fact it is defined precisely by this succession of intervals. But this demonstration is not constructive for two coinciding reasons: when the AB segment is divided in half, how do we know in which half there are infinite points and in which not? Suppose up to 10 million points were counted both on the right and on the left; what is decided? At what stage of this process will we be authorized to decide? In fact, it is not possible to decide on the infinite sets as we do on the finite sets: it is possible only if one has a precise algorithm, which in this case is missing, because the set is completely generic. The other reason is that when we split the segment in half, we cannot use the logical law of the excluded third, saying: either infinity is to the right or to the left; precisely because we know that this law holds only if applied to a finite number of elements or an infinite number provided that we have a finite algorithm calculating it; which however exists in only a denumerable number of cases, not in all non-denumerable cases of a generic set of points within a continuum.

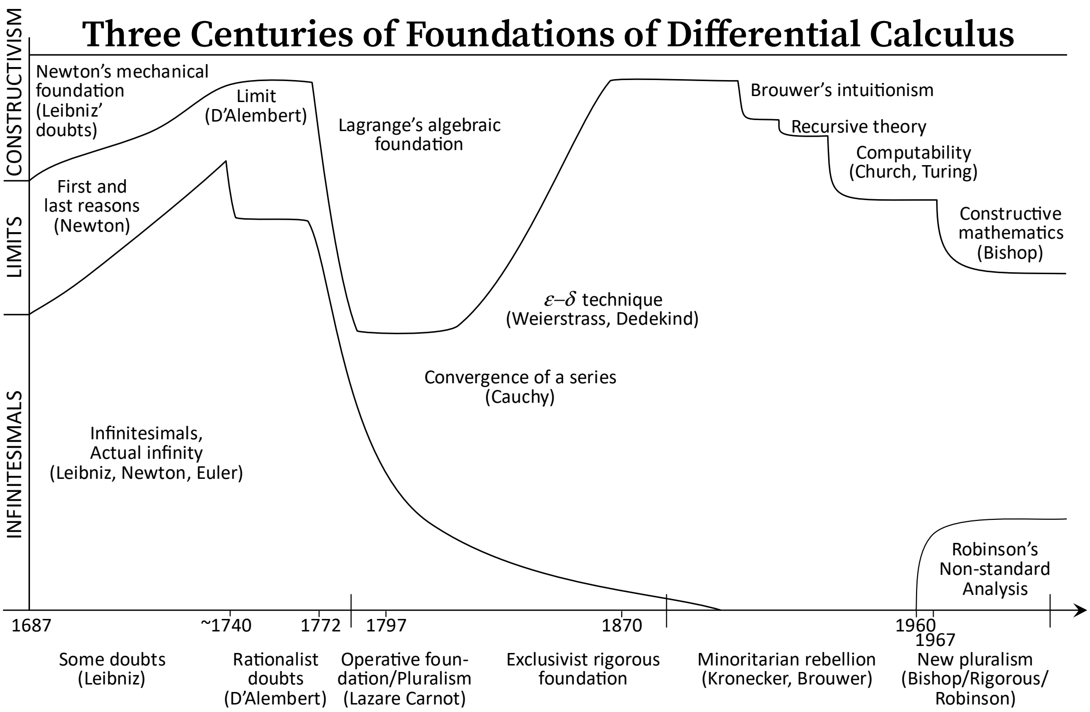
Moreover it can be seen that in constructive mathematics the derivative operation on a point-wise function is repeatable (provided that the function is constructive and we can carry out the process of approximation to the limit constructively, that is, with an actual calculation, but the integral operation requires more than pointwise continuity, e.g. uniform continuity.

Through these examples we realize that there are many differences between the two types of Mathematics. They can be summarized as follows: constructive mathematics, always requiring an approximation algorithm, is not able to define in general the single point with infinite precision and is therefore not able to decide the trichotomy (whether either x=0, or x>0, or <0) or to identify the precise crossing point of two curves or the maximum point of a function. It can only approximate to the latter two.

In general, the characteristic property of Mathematics that includes non-constructive axioms is the acceptance of the so-called "limited principle of omniscience": "Given a subset of $E(n)$ of natural numbers and given a n_o, either an $n'=n_o$ exists in E or $\forall n$: $n \neq n_o$". For example, for the subset of prime numbers, $E(p)$, it is possible to decide this property with a precise algorithm (we calculate all the prime numbers below n_o; they are a finite set and therefore exhaustible); but it is tantamount to omniscience to hypothesize having such a decision algorithm for any subset of natural numbers.

Surprisingly, since 1960, the foundation of infinitesimals has been recognized as a first, even if rough, approximation of a logically well-founded theory, called "non-standard analysis". This is achieved by extending the numbers by removing the Archimedean property ("Given a quantity, there are multiples and submultiples of the given quantity"); in fact infinitesimals dx do not change if they are multiplied by a real quantity. They then turn out to be instances of numbers called hyper-real, which form a larger field than classical real numbers. Non-standard analysis is now used routinely, since, besides being equivalent to the traditional mathematics of infinitesimals, it simplifies all concepts and calculations of analysis. Moreover it can be said that physicists, while accepting the reform of rigor, actually reserved the right to use infinitesimals whenever it facilitated their calculations; see the case of calculating the moment of inertia of a cylinder in rational Mechanics (we will consider another example of the use of infinitesimals in Thermodynamics).

TABLE 1.1: HISTORY OF THE THREE FOUNDATIONS OF DIFFERENTIAL CALCULUS



Legenda: The intersect of the horizontal axis with a vertical line detaches a segment whose three parts represents more or less the percentages of the mathematicians following one of the three foundations of differential calculus.

So today the situation of the foundations of this calculus resembles that of its origin in that three foundations coexist. This makes it even clearer that the opinions of mathematicians on any kind of number, including infinitesimals, differ variously.

To this variety of foundations the vast majority of mathematicians react, according to the very influential Bourbaki group, by ignoring the problem of the Foundations, preferring to seek maximum effectiveness in calculations and in the production of new theorems; in fact, the Bourbaki philosophy states that if other contradictions are discovered in the usual mathematics, then they will be studied and resolved at the moment in which they might occur; for now there is no point bandaging one's head before it is broken. Hence current university teaching, rejecting the study of the problem of the Foundations (which however may implicitly appear in the optional final year courses that use formal methods, such as that of Mathematical Logic). The novelty of the present book is to introduce the problems of the Foundations through the History of mathematical and physical theories, understood in the breadth of all their main formulations. The reference to these different formulations will replace the use of specific formal methods to investigate the Foundations and therefore will simplify the study of those readers who are supposed to know the main physical and mathematical theories.

1.3 The exception of Thermodynamics

Let us now consider whether the above illustrated pluralism in the foundation of the differential calculus does not also concern theoretical physics, where, after the discovery of quanta, it cannot be doubted that the discrete also plays an essential role.

It is interesting to quote from a paper by M. Born that first expresses the enthusiasm for the happily progressive application of infinitesimal analysis in theoretical physics and the sense of wonder at the strangeness of Thermodynamics with respect to Mathematics:

> The mathematical tools that the physicist needs in order to represent the classical domain of his science seem to be narrowly restricted in scope. Systems of partial differential equations dominated that period of physics which lies completely behind us today in this age of quantum physics. Furthermore, there is only a surprisingly small number of differential equations that present themselves repeatedly. Indeed, there is, e.g., no domain of continuum physics in which Poisson's equation does not play a role. That fact, which is already conspicuous for the beginner, does not rest upon mere happenstance, ..
>
> but upon a principle of research that arises from economy of thought. The formulas that mathematics provides are there already, and indeed a relatively small number of them have been worked out to the extent that the physicist can then begin anything. Therefore, he will convert and manipulate his empirical material and the laws that he obtains from it until they take on a prearranged form. In classical physics, the logical treatment of a domain is first considered to be complete when it has been reduced to a chapter in "normal" mathematics.[6]

But Born noticed the existence of a surprising exception: classical Thermodynamics. In this theory the methods used to found the principles are far from the traditional "normal" practice.

> There is only one conspicuous exception: viz., classical thermodynamics. The methods that are usually applied in order to derive the basic laws in that discipline depart completely from the otherwise-customary methods. One already sees that in the fact that there is no other domain of physics in which arguments and conclusions are applied that have any similarity to the process of Carnot cycles. If one further asks which formulas and theorems of mathematics will actually be used for the thermodynamic inferences then one will hardly be able to characterize them as such. The physical theories whose presentation one should appeal to are so singular in their details that nothing seems to remain after one deducts that physical content. But that cannot be the case. Thermodynamics then culminates in a typical mathematical assertion, namely, the existence of a certain function of the state parameters – viz., entropy – and gives prescriptions for its calculation. One would have to admit that thermodynamics, in its traditional form, has still not realized the logical ideal of separating the physical content from the mathematical representation.

---

6 M. Born, "Kritische Betrachtungen zur Traditionelle Darstellung der Thermodynamik", *Phys. Zeit.*, **22** (1921) 218-224, 249-254, 282-286.

This “theoretical ideal”, says Born, was finally reached by the contribution of Carathéodory (we will examine it later and we will see, however, that it is insufficient).

In fact, still today, among all the physical theories that are encountered by a student, Thermodynamics has many mathematical features that are out of the ordinary (which, however, contrary to what Born says, do not seem to depend on the lack of “a correct division” between physical contents and the mathematical representation”):

1) Time t is not an essential variable of the theory. When the theory speaks of it it is to underline, through the concept of irreversibility, that t cannot go backwards, whereas in many other physical theories the equations are invariant under temporal inversion $t \rightarrow -t$.

2) Space is not a founding concept; the only spatial quantity is the volume V which has no particular properties with respect to temperature T and pressure P. Consequently we lose the fundamental variables, time and space, with which in theoretical physics we usually derive and integrate, namely, the usual application of calculus.

3) Each variable is defined as positive. The newly constructed variable, entropy, is also defined as positive (by virtue of the III principle).

4) There is a duality between intensive variables (degree 0 homogeneous, such as temperature T and pression P) and extensive variables (degree 1 homogeneous, such as volume V, total energy U, entropy S; in physical terms, joining two equal systems, the variables of the second type double their values, while the values of the first type remain the same. This duality does not exist in Mechanics, nor in Electromagnetism.

5) Thermodynamics as a whole speaks of transformations, but (until recently) the algebra of the group theory has not been applied to it.

6) Irreversible processes have not been studied for a century; perhaps because they are represented with an inequality $\Delta S \geq 0$ that was unusual in the Mathematics of theoretical physics

7) The principles have long been expressed by means of “impossibilities”, which (Truesdell) made their translation into traditional Mathematics impossible.

8) The observation the student makes immediately when he studies the Thermodynamics of Carnot - Clausius - Kelvin is that it does not make essential use of infinitesimal analysis. Yet its conceptual aspect is extraordinarily abstract, because it concerns the universe of thermal transformations. It seems useful to note Theobald’s opinion on the lack of differential equations in Thermodynamics: here P, V and T are measurable independently and their numerical relations exhaust the physical content of their relations. Instead in Mechanics force F, mass m and acceleration a cannot be measured independently of each other; therefore, besides the algebraic relation $F = ma$ which connects them, there exists the corresponding differential equation. Moreover Theobald is of the opinion that while Mechanics and Electromagnetism start from descriptions of the neighborhood of a point and reach a general theory on all the space, therefore they must assume that the continuum is valid in both local and global reality, Thermodynamics is interested in global systems only and states nothing about their

local constitution; therefore it does not assume the continuum a priori.[7] We might add that in fact Thermodynamics built its theory only from the moment when (S. Carnot) it ceased to concern itself with the internal physical structure of heat and bodies, that is, when it ceased to concern itself with infinitesimal analysis (the only mathematics that Fourier, Laplace, Poisson had previously considered suited to describing thermal phenomena in bodies.

9) In the formula of the first principle the mathematical sign for "equal" between heat Q and work W does not mean their convertibility or interchangeability. In fact if, to simplify the reasoning, we consider a cycle, such that the variation of total energy U is null $\Delta U = 0$, it follows that (disregarding the correct sign, which is in fact conventional) $Q=W$; but this last equality cannot hold; because it is negated by the second principle (which rather requires us to write $Q\text{-}Q'=W$ (as well as the Kelvin principle, which requires two distinct thermal sources and in any case requires that the transformation of Q into W always has an efficiency lower than 100%). That formula $Q = W$ in reality, that is according to the second principle, means an "equivalence" between Q and W; but the word "equivalence" in Physics does not have a specific mathematical symbol. In conclusion, this first principle uses mathematics in a way that is inappropriate. If we then ask why the mathematical calculations of Thermodynamics, based on this equality-equivalence, never lead to errors, we must remember that these calculations only use state functions (P,V,T,S), no longer W and Q; that is, we work in a space of state variables where every mathematical aspect is accurately defined. However, it is at the same time abstract, because it is only partially connected with reality, which is given by Q and W; these are calculated at the end of calculations with the state variables, as variables concerning applications only. (Already Poincarè, reflecting on physical theory, had strongly criticized the first principle and had concluded that it is only an interpretative schema imposed by our mind on nature).

10) Clausius' formulation translates the relation $\sum_n Q_n / T_n = 0$, which holds in a system of n Carnot sub-cycles of the cycle being considered, to Clausius' integral $\int dQ/T$ on a closed path. This step is unjustified according to the definition of integral given by Riemann, taught in Analysis, since only particular sub-cycles (those of Carnot) are taken into consideration and not all the possible sub-domains of the original cycle.

11) It is not clear why, to establish the Thermodynamics, in which δQ and δW are not exact differentials, the mathematical problem of exact differential was not used, immediately after it was clarified mathematically and solved in general (1813). It took a good 90 years to arrive at Carathéodory's formulation.

12) The third principle of Thermodynamics is disconcerting due to its extreme mathematical simplicity; it merely states that one variable, entropy, is always defined as positive (as are the other thermodynamic variables). Moreover its usual formulation $\lim_{T\to 0} S(T) = 0$, is incompatible with classical mathematics; in fact in this mathematics every limit process reaches the limit point $S = 0$, which however in Thermodynamics does not exist because the third principle says it is inaccessible by experimental means; however, in classical Mathematics there is no way to express a limit process that cannot reach the

[7] D. W. Theobald, *The concept of Energy*, Spon, London, 1966.

final point. So that formula, which many Thermodynamic texts repeat, is rather a symbolic formula, not a real formula of classical Mathematics.

13) The "oddities" intrinsic to Thermodynamics may be connected with the strange way, characteristic of physicists of Thermodynamics, of indicating the non exact differential with δ, or a $đ$; and of writing $P=P(V,T)$ e $P=P(S,T)$, although in both cases the two functions $P(\ , \ )$ have a mathematical form that is certainly different; the partial derivative between brackets, or also $(\partial P/\partial V)_T$. In the last two cases the thermodynamicists use an abbreviated notation to emphasize the physical meaning of the formula; moreover, when in the second case they indicate the partial derivative of F with respect to the variables y and z, keeping x fixed, their notation takes into account the equation of state between the three variables and indicates only one of them as fixed: $(\partial F(x,y)/\partial x)_z$; this notation seems strange to mathematicians because it does not hold for more than three variables.

14) Among all the theorems of usual physics, Carnot's theorem is the only one to use reasoning by absurdity; that is, it derives an experimental statement from what is impossible, unreal! And note that even the equivalence theorem between the Kelvin statement and that of Clausius is shown by absurdity. So the only formalism in these theorems is not mathematics, but logic.

It can therefore be seen that all these problems do not depend on experiments; you could make thousands of experiments without being able to decide any of the problems dealt with, given that, as they involve the relationship of the theory with Mathematics, they are typically theoretical. They indicate an essential component of physical theory that no laboratory result can call into question.

1.4 Carathéodory's formulation of Thermodynamics

In the early 1900s, the period of maximum crisis in theoretical physics, Born (of whom we previously reported the disbelief over the great simplicity of Mathematics in Thermodynamics) urged the great Greek mathematician Carathéodory to propose a new formulation. In 1909 he found it, introducing, for the first time in an essential manner, Analysis into Thermodynamics. This mathematization seemed, even more than Boltzmann's statistical mechanics, to bring Thermodynamics back to Mechanics.[8]

Carathéodory treated the whole subject of the theory on the basis of the problem of the non-exact differential of heat, finding for the first time appropriate theorems to solve the problem in $n>2$ variables, as the case of Physics requires. (The reader may skip the following two pages of mathematical proofs)

His mathematical problem is that of the heat differential. Let us begin with the simplest case of $n = 2$ variables.

$$dQ = dU + pdV. \qquad (1.1)$$

This differential is not exact, in the sense that, equivalently:

---

[8] C. Carathéodory, *"Untersuchungen über die Grundlagen der Thermodynamik", Math. Ann.*, **67** (1909) 355-383. English translation in J. Kestin (ed.), *The second law of thermodynamics*, Stroudsburg: Dowden, 229-256.

1) its integral on a closed path is not zero. Physically: if a transformation returns to the starting state, the amount of heat is not zero; take, for example, the energy lost by a cyclist to cycle uphill; it is not compensated by the energy acquired in a subsequent descent of the same height, because the energy lost to overcome frictions is irrevocably lost;

2) its integral exists, but depends on the path of transformation. Physically: two states *A* and *B* can be linked by an adiabatic process, thus $\int dQ = 0$; or by an isobaric and an isochoric process and thus $\int dQ \neq 0$; (intuitively: when you go into the woods to collect mushrooms, the paths in the woods are not equivalent).

3) *dQ* is not equal to the differential of a function: that is, there is no primitive function. Physically: there is no thermodynamic variable such that its integral is equal to $\int dQ$.

In Mathematics, 1.1 corresponds to a differential in two variables, which, with absolute generality, is given by: $dP = M(x,y)dx + N(x,y)dy$.[9]

Theorem 1. A necessary and sufficient condition for *dP* to be exact is that (provided that *M* and *N* are equipped with continuous derivatives) it is

$$\partial M(x,y)/\partial y = \partial N(x,y)/\partial x \qquad (1.2)$$

i.e. Schwartz' equality of the mixed partial derivatives.

Proof. The condition is necessary. In fact, if *dP* is the exact differential of a function *F*, then (1.2) must hold, because it represents the equality of the mixed partial derivatives (which is true whenever *F* is continuous with its first derivatives).

The condition is sufficient. It is a matter of constructing the primitive function of *dP*. First, we can hold a variable constant, that is, $x = x_0$ and thus conceive *N* as a function of *y* alone; now let us try to derive the primitive function of $N(x,y)$. Given that only one variable has remained, the integration is always possible, leaving an additive constant, of course a constant with respect to *y* and therefore function of x_0:

$$F(x_0,y) = \int_{\alpha(x)}^{y} N(x_0,v)\,dv + \phi(x_o)$$

where the integral is from *x* to the continuous function $\alpha(x_0)$, whose curve defines the field of definition.

If now we proceed inversely and derive with respect to *y* (or *v*), by construction of the function *N* we obtain $\partial F/\partial y = N(x_0,y)$. But we do not know whether, deriving *F* with respect to every *x*, $\partial F/\partial x = M(x,y)$ also holds. Let us try to impose it. We will first need to calculate this partial derivative:

$$\frac{\partial F}{\partial x} = \int \frac{\partial N(x,v)}{\partial x}dv - N(x,\alpha(x))\frac{d\alpha(x)}{dx} + \frac{dj(x)}{dx}$$

However for the hypothesis (1.2), the integrating function can be replaced by $\frac{\partial M(x,v)}{\partial v}$, which gives the integral $\int_{\alpha(x)}^{y} \frac{\partial M(x,v)}{\partial x}dv$ of a function *M* derived in the same integration variable; then we obtain:

[9] On the differential form and its solvability, see I. Sneddon, *Elements of Partial Differential Equations*, Dover, New York, 1957.

$$\frac{\partial F}{\partial x} = M(x,y) - M(x,\alpha(x)) - N(x,\alpha(x))\frac{d\alpha(x)}{dx} + \frac{d\phi}{dx} .$$

This expression can, as we wish, be equal only to *M (x, y)* if the rest of the second member is null; if that is

$$\frac{d\varphi(x)}{dx} = M(x,\alpha(x)) + N(x,\alpha(x))\frac{d\alpha(x)}{dx};$$

It is so if we choose the constant $\varphi(x)$, so far arbitrary, such that it is the solution of

$$\phi(x) = \int \left(M(u,\alpha(u)) - N(u,\alpha(u))\frac{d\alpha(u)}{du}\right)du + K ,$$

which is certainly possible under broad assumptions. In this case *dP* is equal to the differential of the function *F* (*x, y*)

$$F(x,y) = \int M(x,v)\,dv + \int \left(M(u,\alpha(u)) + N(u,\alpha(u))\frac{d\alpha(u)}{du}\right)du + K \quad \text{. (C.D.D.)}$$

Now let's verify with (1.2) that the differential *dQ* is not exact. We use for variables those that are suggested by the usual notation (1.1), ie *U* and *V*:

$$\frac{\partial^2 Q}{\partial U \partial V} = \frac{\partial P(U,V)}{\partial U} \neq \frac{\partial^2 Q}{\partial V \partial U} = \frac{\partial 1}{\partial V} = 0 \qquad (1.3)$$

In fact, the first two derivatives are different from 0; for example, given a gas in a blocked cylinder (*V* = constant), if we increase the internal energy *U* (for example by heating), *P* changes necessarily. Instead the second pair of derivatives gives 0.

Now, when a differential in two variables is not exact, it can become exact by multiplying it by an appropriate integrating factor. In fact, by multiplying *M* and *N* by a function, the derivatives in (1.2) give additional terms that can make the two members equal.

Let us verify it in the case of Thermodynamics. To facilitate the calculations, however, let us restrict the system to being a perfect gas with variables *T* and *V*. Then, for a mole we have

$$dQ = c_v dT + \frac{RT}{V} dV .$$

(1.2) gives

$$\frac{\partial c_v}{\partial V} = 0 \neq \frac{\partial}{\partial T}\frac{RT}{V} = \frac{R}{V}.$$

But if we multiply *dQ* by what, in another way, we know has the role of integrating factor to obtain entropy, *1 / T*, we have

$$\frac{dQ}{T} = \frac{c_v dT}{T} + \frac{R}{V} dV = c_v d(\ln T) + R d(\ln V).$$

Treating ln *T* and ln *V* as new variables, then the differential *dQ / T* is clearly exact, since given that c_v and *R* are constants, their derivatives are both null and therefore equal to each other.

Now it is easy to show that it is always possible to find the integrating factor of a differential form on two variables; but this does not hold on more than two. In fact we note that the differential *dP* defines, through *M* and *N*, a variable vector field, whose generic vector has components *M (x, y)* and *N* (*x, y*). Equating *dP* with zero expresses geometrically the condition of perpendicularity at every point between the line element, with components *dx* and *dy*, and the vector components *M* and *N*. It can be seen that this equality gives rise to a differential equation that can be solved under broad hypotheses; therefore we can always find a family of solutions $F(x, y) = C$. It is clear then that there will always be an integrating factor function $\rho(x,y)$ such that $\rho(x,y)dP=F(x,y)$. Geometrically, we can imagine that the multiplication by the integrating factor links together all the line elements that are perpendicular to the component vector M and N in the respective application points so that the line elements together form a family of well-ordered curves which are arranged so as to cover the entire plane, without the curves ever crossing (because otherwise at that point two distinct infinitesimal line elements would be perpendicular to the same vector).

But Thermodynamics studies an isolated system not only the exchange of mechanical work, as it expresses the differential $dU + PdV$; but also works of any type (hydrostatic, electric, magnetic, chemical, nuclear, etc.) and above all, this theory studies the problem of interacting systems (therefore with P_1, V_1 and P_2, V_2 and P_3, V_3, etc.); that is, in these cases the differential of Q is in $n > 2$ variables: $dU + P_1dV_1 + P_2dV_2 + P_3dV_3 +$ etc. Obviously also in these cases the differential of Q is not exact. Therefore the exact differential theory is useful for Thermodynamics only if it solves the problem in more than two variables, when the $dP = 0$ condition alone is not sufficient to determine the integrating factor in all cases.

When building his formulation of Thermodinamics Carathéodory's problem was essentially to find a physically significant mathematical condition that would ensure the integrability of dQ in more than two variables. We see, again geometrically and in a particular case, the idea with which Carathéodory solved the problem of finding a condition of integrability of the non-exact differential. Let us return to the case of two variables, taking a very simple differential form: $dx\text{-}dy = 0$. Its solution curves are the straight lines $x\text{-}y = C$, that is, the straight lines that form 45° with the x-axis.

Take the point A $(0,0)$ and draw the circle of radius r. The line $x = y$ meets it at two precise points, B and B'. Then we note that, if we want to keep the condition $dP = 0$, from A we can only go along the line $x = y$ and so we can only reach the points B and B' of the circle. Also as r shrinks, we see that there is always an infinity of points close to (0,0) that cannot be reached from it with curves for which $dP = 0$. Since the point A can be chosen at will, we can conclude that: if there is an integral factor $1/\rho$ for each point A there are points, as close as one wishes to A, not reachable along the curve $dP = 0$

So let's go back to the non-exact differential in $n>2$ variables. Carathéodory showed that in this case in some neighbourhoods of the point A all neighboring points are reachable along curves $dP = 0$. Therefore Carathéodory showed the following theorem excluding this possibility:

<u>Theorem</u>: A necessary and sufficient condition for a differential in more than two variables to be integrable is that the following condition holds: in every neighborhood of each point there are points not-reachable along a curve which is a solution of $dP = 0$.

Thus Carathéodory was able to achieve a very interesting mathematical reformulation of Thermodynamics and in doing so produced *the first axiomatic theory of the history of Physics*. At the beginning of its formulation Carathéodory needs to define an adiabatic wall, that is, insulating with respect to heat exchanges, without defining what is heat, because this concept will be defined only later by the theory; he overcomes this difficulty by referring to the common knowledge of the "diathermic" walls, i.e. that they do not allow the space enclosed by them to change temperature.

<u>Principle I</u>. "To make a system (homogeneous or heterogeneous) pass from an initial state S_1 to a final state S_2 by means of a process that uses diathermal walls the same amount

of work is always necessary. The latter does not depend on the reversibility or irreversibility of the process. The work done on the system is expressed by the integral of dW, while

$$U_2 - U_1 = \int dW$$

and it is called the increase of the internal energy of the system U. According to this definition the internal energy is a function of state and is fully defined when the constant of integration is determined ".

More generally, if we remove the diathermal walls the same previous transformation leads to a difference between the internal energy U and the work done on the system W; this

$$\int_{S_1}^{S_2} dQ = U_2 - U_1 - \int_{S_1}^{S_2} dL$$

difference is called "absorbed heat":

.

This is Carathéodory's version of the usual I principle. Note that he does not treat heat as an intuitive physical entity, but he derives it from work and internal energy.

He derives from the previous formula the differential expression

$$dQ = \sum_i^n X_i dx_i ;$$

which can easily be recognized as not being an exact differential. To ensure the existence of the integrating factor ρ on $n<2$ variables, then we need to resort to the Carathéodory theorem, or rather to the only sufficient condition; his proposition then assumes the role of the II principle of Thermodynamics:

Principle II: In every neighbourhood of the thermodynamic state S_0 there are infinite states S that are not reachable from S_0 by means of an adiabatic transformation, that is, for which $dQ = 0$.

Note that the adiabatics are transformations in which there is no heat exchange, that is, they only involve mechanical phenomena. Carathéodory's statement underlines therefore the fact that thermodynamics is essentially irreducible to mechanics, because close to any thermodynamic state there are states that are not mechanically reachable by the first.

If dQ has an integrating factor then $dQ/\rho = dF$. It can be shown that ρ is a function only of t, the empirical temperature of the system, and it is shown that this function $\rho(t)$ is proportional to that of the perfect gas thermometer and therefore it is called absolute temperature T. In this case dQ/T, which we now know to be equal to the differential of a state function, is therefore equal to an exact differential of a new variable dS, where S is precisely the entropy of the system.

1.5 Undecidibility of Carathéodory's formulation

The entire formulation above, however, depends on an axiom which is the most abstract idea one can suggest to a physicist; Carathéodory himself, in his original article, admitted that it is very difficult for Physics to verify his second principle, because the sensitivity of all measuring instruments is finite; therefore we cannot speak of all the neighbourhoods of a point, unless we go beyond the sensitivity of the instruments; moreover, the axiom requires carrying out such a verification for any state of the system and such a verification on an infinite number of states can never be completed.

On the other hand, in constructive Mathematics the single point cannot be found in general with absolute precision; much less can we speak of the infinite number of neighbourhoods of each point. Therefore for constructive Mathematics this principle is to be rejected.

At this point it can be doubted that constructive Mathematics forces one to discard the whole of Thermodynamics. In reality, Here the undecidability of $PV = nRT$ or other experimental laws have not been questioned, but only that of a principle.

To clarify the point, consider a physical theory in general.[10] It includes five types of mathematical entities:

1) *Constants*, for example R, or c, or ε, or μ, etc.. These constants are simple numbers, which are determined experimentally with a certain precision (say ten decimal digits). They are therefore rational numbers and do not present any problems in constructive mathematics.

2) *Discrete variables* (such as atomic numbers) *and continuous variables* (such as pressure P, temperature t, spatial location x, etc.). The former are not a problem for constructive Mathematics. The latter can be thought to vary on the real constructive numbers, without this presenting problems (also because their values correspond to the results of physical measurements, which always give rational numbers).

3) *Formulas* that link together various variables and constants. They might appear to present problems because they always include an equality, which in constructive mathematics is undecidable in general. But every formula expresses an experimental law, whose equality is always verified with a certain approximation, that given by the finite precision of the measurements of all its variables. Its equality is therefore only approximate and holds within a certain error. It is therefore a constructive equality.

4) *Mathematical techniques*, such as differential equations. These are not clearly experimental, so they are not approximate. Therefore in constructive mathematics the corresponding finite difference equations must be studied to see if their solutions are the same as the difference equations of classical mathematics or not. In general, but not always, the first hypothesis applies. It is therefore necessary to see when the solutions of differential equations required by theoretical physics are constructive.

5) *Principles*. These may have simple mathematical formulas (such as the I and III principles of Thermodynamics), for which what was said in the above point 3 is true; or they may have verbal expressions, such as the Carathéodory axiom or the principle of inertia. In this case it is necessary to analyze the verbal expression carefully to verify whether the *words* used and their overall *physical sense* do not include undecidabilities.

Previously we have examined, according to step 5, the principles of Carathéodory thermodynamics and we have declared that one of them is undecidable. All this does not change anything about the law $PV = nRT$, or other aspects of the theory that are related to physical measures. Constructive mathematics does not therefore reject the experimental laws of Thermodynamics, but the principle of Carathéodory's formulation, which as we have

---

[10] A. Drago, "Relevance of Constructive Mathematics to Theoretical Physics", in E. Agazzi et alii (eds.): *Logica e Filosofia della Scienza, oggi,* Bologna: CLUEB, 1986, vol. II, 267-272.

seen is very marginally linked to experimental reality; and consequently rejects Carathéodory's formulation. Ultimately, constructive mathematics does not reject a physical theory, but particular formulations of that physical theory.

Let us now examine the traditional formulation of Thermodynamics according to constructive Mathematics. We have already noted that its Mathematics has a very elementary level. But there is an idealizing concept: the perfect gas; however this concept can always be approximated to by suitable physical systems and does not therefore present problems. There are no differential formulas with the exception of the defining formula of entropy. But it can be recalled that its original formula is an algebraic formula $\Sigma Q_i/T_i$. The critical point is the step from the summations $\Sigma Q_i/T_i$ to the integral $\int dQ/T$, but this step is not essential in order to account for the experimental reality; moreover it is not a Riemann integral, which would require the calculation for all the decompositions of the domain of definition. Here, instead, only Carnot's decomposition into sub-cycles is used, with transformations that are assumed to be reversible and hence uniformly continuous; it is therefore the operation of limit to the integral as was carried out by Archimedes; it is certainly constructive.

We therefore conclude by stating that *there is at least one constructive formulation of Thermodynamics, the usual Carnot-Kelvin-Clausius formulation*. And this was historically the first formulation.

Rethinking the Mathematics of Thermodynamics in the light of constructive Mathematics, we note a further "oddity" with respect to those already recognized in the usual illustration of this physical theory. Carnot's theorem is demonstrated without calculus, by *ad absurdum* reasoning. This fact is inexplicable in the context of the traditional relationship between mathematics and physics. But according to constructive mathematics it is very clearly justified: Sadi Carnot wanted to obtain the maximum among the efficiency functions of thermal machines; this result is undecidable. In fact, due to the fugitive numbers a maximum of a function is undecidable; a maximum function among a set of functions is even more undecidable. Therefore, his theorem was obliged to use not (constructive) mathematics, but logical arguments alone, based on empirical observations concerning machine technology.

Furthermore this traditional formulation is by no means a primitive theory. It it can be noted that in this formulation there is an operational and constructive affirmation which implies that the Carathéodory axiom is a special case of it. In fact, the existence of the Carnot cycle in every state S of the system implies that in such a state there exist both adiabatic and isothermic transformations distinct from each other; this fact implies that in the neighbourhoods of each state *S* exist states infinitely close to it that are inaccessible with adiabatic transformations: this is true at least for the states of the isotherm passing through S, because these states, by hypothesis, do not lie on the adiabatic. Therefore the aforementioned operative statement on the existence of the Carnot cycle is also a suitable hypothesis for solving Carathéodory's mathematical problem of the exact differential.

Ultimately, constructive mathematics has led us to recognize differences not only regarding numbers and functions, but also the principles of physical theories and formulations of the same theory, some of which are accepted, others not.

The following can be concluded from this:

1) The first formulation of Thermodynamics is traditionally accused of being primitive for not having used the advanced techniques of infinitesimal Analysis. In reality we can think of that formulation as an implicit attempt to question the traditional relationship between classical Mathematics and Physics.

2) The common opinion regarding constructive Mathematics is true, that is, that, being more restricted than classical, it does not give the same physical results as classical analysis; but this is true not in the sense that it is constitutionally incapable of reproducing what is experimental in physics; but in the sense that, as we have seen in the case of Thermodynamics, it does not accept some well-known formulations of physical theories; that is, it separates the different formulations into two classes: those that can be expressed and those that cannot be expressed constructively.

3) The two classes of formulations thus obtained correspond to two different approaches in Physics, clearly recognizable in the history of Thermodynamics. The first is specifically thermodynamic, for example, S. Carnot's, and its characteristic feature is to be global. The second emphasizes the mechanical content of Thermodynamics, both in concepts (for example, Carathéodory's definition of heat through what is not mechanical energy) and in Mathematics (the classical Analysis of differential equations); its characteristic feature is to be local.

When schools of constructive Mathematics first appeared at the beginning of the 20th century, there was a general scepticism concerning their capability of recovering the practice of mathematicians (especially differential equations). This scepticism over the possibility that constructive Mathematics could bring about innovations in Physics was even greater and persisted even longer. It was only after the birth of Bishop's constructive mathematics that there was a first attempt to introduce constructive Mathematics in Physics, and precisely in Thermodynamics (considered a particularly interesting physical theory because the latest formulations are based on very simple mathematics, even elementary algebra alone) and in Quantum mechanics (which constitutes a great challenge to demonstrate the suitability of constructive Mathematics in Physics). But the authors[11] encountered various problems of undecidability and concluded that it was difficult or perhaps impossible to express the axioms of these theories constructively. That is, they suspected that the defect lies in constructive Mathematics, which is too weak for theoretical physics. Instead here, using the various formulations of Thermodynamics, *it was shown that the defect is possibly in the particular formulation chosen for studying a physical theory.*

1.6 Thermodynamics and constructive mathematics: reversibility

But does constructive mathematics have only a selective, obstructive function? Or can it also indicate how to best reconstruct the corresponding formulations, such as those invented by

---

[11] W.K. Burton, "Constructive Thermodynamics", in H.A. Schmidt, K. Schütte, E.J. T hiele (eds.), *Contributions to Mathematical Logic,* Amsterdam: North-Holland, 1968, pp. 75-89; D. Bridges, "Towards a Constructive Foundation of Quantum Mechanics", in F. Richman (ed.): *Constructive Mathematics,* Heidelberg: Springer, 1981

scientists who were not aware of it? Or even, now that we know more about the foundations of Mathematics, what is the most appropriate Mathematics for thermodynamic theory?

To answer the question we will refer to the first and best known formulation, that of Carnot-Kelvin-Clausius.

Let us observe once again the characteristic and basic concept of Thermodynamics, reversibility. It has not received a clear definition either in the past or in recent times (it is easy to verify this by comparing several textbooks). The various definitions can be classified into two main types:

1) those that refer to the *effects* produced by a reversible process, regardless of how it occurs;

2) those that specify the characteristics of the process, that is to say, in what way it is carried out.

Type 1 definitions refer to a set of entities that are difficult to delimit with precision: the environment and all agents in nature would have to be reviewed to establish whether the effects on the system and the external environment are the same when the process is reversed. "A colossal verification, which no one will ever be able to perform," says Sommerfeld's text on Thermodynamics. Type 2 definitions are more practicable. They almost always use the concept of "quasi-static process" (to which some authors even reduce the whole concept of reversibility; as if the viscous processes, obviously quasi-static, could be reversible).

But the concept of an "quasi-static" process requires that a state be at the same time in equilibrium and yet belong to the dynamics of a process, so that each state would be represented not by a real number, which may indicate only a single fixed state, but by a "dynamic" number that "tends" towards the given number, and therefore by that number which was defined as "tending" to zero: the infinitesimal!

For physicists this is a conceptual trap because the infinitesimals have no operational meaning. To get out of it, what was suggested by E. Mach should be recalled. Thermodynamics is the first physical theory that introduces "unknown" quantities, temperature and heat, for which new instruments of measurement are necessary. It was natural to wonder whether these variables were continuous or not; this doubt led to paying greater attention to the physical conditions necessary to ensure continuity, in other words: reversibility is a prerequisite for the use of the mathematical continuum.

However, we would add that, if we want to introduce the mathematical continuum with precision through reversibility, we must choose a particular conception of the mathematical continuum. We therefore abandon the attempts pursued by many physicists to find only physical conditions to define the concept of reversibility, but rather see its definition as a conceptual problem of the relationship between mathematics and physics.

If, after excluding infinitesimals, we turn to "rigorous" Mathematics, then the state of the system can only be represented by an isolated point, but in this case we cannot, by examining only an isolated point, derive the properties of a dynamic process involving, of course, an infinity of points and, *vice versa*, we cannot derive from the dynamics of the process the exactly representative point of our state, since the limit of the idealistic $\varepsilon-\delta$

technique leads to a leap from approximating intervals to a single point among the infinite points of the last interval considered by us. Thus we note that it was this inadequacy of "rigorous" Mathematics that forced physicists to return to using the previous mathematics of infinitesimals and thus speak of the quasi-static process. Instead, in constructive Mathematics we know that a state can be represented in general with a small continuous segment (albeit increasingly reducible in length), within which each parameter can still vary. And in fact this is the right way to think of it if our thinking accepts the fact that all our measurements are necessarily inaccurate (even if perfectible). Of all the types of continuum, the constructive one appears therefore well suited to the experimental nature of physical theory.

Is it also in defining the physical concept of reversibility? In fact, a fundamental clarification of this concept was given when a distinction was made between reversibility and invertibility. It is also important to clarify the history of the concept of reversibility. This concept was introduced by S. Carnot in 1824, but was certainly derived from the concept of geometric motion of his father, Lazare Carnot; which defined such motion as invertible motion.

A moment of reflection indicates that we can define as a locally invertible process a process which at each point-interval can proceed indifferently in both directions, always using the same tools and agents. (It is clear that which invertibility is a local condition, it does not ensure reversibility, because this is a global condition on the whole process; therefore, on finite sections of the process and also on the entire route, invertibility requires additional conditions, such as the total lack of friction that acts on each finite segment).

Let us see what this definition of invertibility involves:

1) it refers to an interval that can become so small that it reaches the limit of sensitivity of the measuring instruments. This allows us to disregard as secondary, since they are very small, a whole series of effects concomitant with the considered variation, and therefore we can limit ourselves to considering a small number of physical quantities, the main ones;

2) the transformation, although dynamic, is constituted by a series of states that are equilibrium states; because, if we were in a state that is not in equilibrium, to be able to reach equilibrium again we would be forced to proceed in a specific direction, and this action would thus lose the possibility of proceeding indifferently back and forth at every moment;

3) the system cannot have heat exchanges with the outside due to finite temperature differences, since the transfer of heat in the opposite direction would have to recover from the environment the previously dispersed heat and would have to make it converge entirely into the system; this action cannot use the same means as the direct process and moreover it is practically impossible to do so in such a way as to equalize the balance of the direct transport of heat and its inverse transport;

4) there cannot be hysteresis cycles, otherwise on the pivot points the transformation could proceed in two directions.

Let us translate the above definition into mathematical terms. In fact it is required that on the small segment which represents the point of the constructive continuum, the following operations can be performed: 1) the limit from the right, 2) that from the left, 3) and to verify that they coincide. Therefore, point by point, the condition of local invertibility ensures (from the Physics side) that it represents a state of equilibrium and (from the side of constructive Mathematics) the "pointwise continuity" of each variable.

In constructive Mathematics there are many undecidabilities. We note then that the undecidability of some problems seems implicit in the traditional Thermodynamics theory, since all of Carnot's reasoning (as also that of the father in Mechanics) is based on an analogous principle, the impossibility of perpetual motion; which is a methodological principle, not an axiom, because this principle is a statement about the complex of possible physical operations: there is no finite set of physical operations such as to generate a movement that persists for all time. Similarly, every undecidability is a statement about the complex of actual mathematical operations: there is no finite set of operations through which we can solve a problem for all the values of a given variable.

Continuing the revisiting of the relationship between Mathematics and Thermodynamics, we move from simple variables to the functions of variables. At least in Sadi Carnot, the central problem of Thermodynamics concerns the search for a function: the function of maximum efficiency of thermal machines. In classical mathematics the problem of a maximum function on a continuous set of functions is generally solved by differential operations. But nineteenth-century physicists, who followed classical mathematics, failed to find the solution to this thermodynamic problem. Constructive mathematics, on the other hand, warns that the problem of finding a particular maximum function is in general undecidable: in fact there is no algorithm (which is based on a number of steps that can be numbered at most) that detects exactly a precise function (which is defined on a continuous infinity of points). Carnot seems to have understood this impossibility, because he solved the problem with *ad absurdum* reasoning, and not with the Mathematics of that time.

1.7 Thermodynamics and constructive Mathematics: uniform continuity and integration operation

Let us then observe how the traditional formulation, seen as a whole, uses mathematics. At first glance it seems that current Thermodynamics (with the many other mathematical oddities) as well as its original formulation of S. Carnot's Thermodynamics (with the proofs by absurdity, with its impossibilities and with the choice of an elementary mathematics) express in effect a tenacious refusal of infinitesimal calculus. In reality it would first be necessary to specify the infinitesimal operations that are being referred to. Usually we think of the derivative operation, forgetting however that it was only since the time of Leibniz and Newton that derivation was seen as a primary operation, whereas the whole of antiquity had considered integration to be the primary operation: the calculation of the areas instead of the calculation of the tangent, i.e. the calculation on a finite area, instead of calculating a property in the neighbourhood of a point. In fact, Thermodynamics has a global, and not local, approach to the system being studied. The above suggests following the older tradition.

In fact the fundamental mathematical problem of Thermodynamics is not to differentiate a variable (as it is in mechanics, where there is the problem of calculating space and time differentials, whose ratio gives the speed of a material point and where there is the calculation of the differential of the momentum, a differential which is equal to a force for the time differential). It is rather that of the calculation of a quantity (work, or other) considered between two different states, that is, of an integral. Here S. Carnot was clear:

> The production of motive power is therefore due ... to the passage [of heat] from a more or less warm body to a colder body "(p. 10); so that "... it is not enough to produce heat; we must also obtain cold; without which heat would be useless". (p. 11).

That is, S. Carnot's fundamental problem is that of a calculation that considers a transformation between two different states: that is, an integration.

In Physics it is clear that a locally invertible process does not ensure uniform continuity, which is a global property. In classical Mathematics, Heine-Borel's theorem says that a pointwise continuous function is uniformly continuous in a closed and bounded interval. In this Mathematics there is therefore no need for additional conditions to be added to a pointwise function in order to have the uniform continuity ensuring reversibility. But the proof of that theorem considers an infinity of intervals of continuity and from each of these extracts a point to form an infinite set of points; that is, it makes essential use of the Zermelo axiom which includes actual infinity (AI). This is exactly what constructive mathematics does not accept and with it the proof of this theorem. In constructive Mathematics a function is integrable when it is uniformly continuous; and therefore a pointwise continuous function is not integrable, unless an additional condition occurs that makes it uniformly continuous. Also in Physics, the integration of an invertible thermodynamic function requires the addition of other physical conditions; physical experience tells us that reversible processes must not only be invertible, but also free of friction, viscosity, hysteresis and turbulence. We can intuitively say that an invertible process is reversible when it identifies a succession of states (in the sense of the physical continuum composed of lin elements) that all have the same extension; this is possible only if there is no friction, or some similar phenomenon, essentially disruptive of the uniformity of the sequence of the line elements. The absence of disruptive processes thus ensures the uniform continuity of the thermodynamic functions.

With this, constructive mathematics clarifies an obscure point in thermodynamics, that is, it re-evaluates the concept of reversibility as something more than the invertibility of quasi-static processes; and assigns a precise mathematical role (moving from pointwise continuity to uniform continuity) to the physical conditions added to invertibility in order to obtain reversibility.

Once again physical concepts are parallel to those of constructive mathematics.

1.8 Thermodynamics and constructive Mathematics: the primitive entropy function and the derivative operation

Let us start by noting that, of all thermodynamic variables, one, heat (which can only be integrated if the process is reversible) has an integral that depends on the path of integration; it is clearly seen physically; a thermal machine that receives a certain amount of heat and that works between two temperatures releases heat at a lower temperature; the more work it is able to produce, the less the quantity of heat it releases; and it is clear that there are better machines than others (which corresponds to the different methods of performing an integration). To apply Mathematics here a way must be found to make the integral independent of the path of integration; for example, in physical terms, to seek maximum efficiency. On the other hand, steam machines themselves suggested to Carnot a privileged path of integration, the cycle. Thus we find that S. Carnot's reasoning about the cycle, which seemed to be an engineering concept, is actually mathematical.

In his cycle he considers one isotherm that transmits heat to the system, while the other isotherm releases it; these two transformations are connected by two transformations that do not give interaction between heat and work: either isochore or adiabatic transformations; they can be understood as auxiliary variables, i.e. infinitesimal as dx in analysis, and therefore with a negligible contribution. Then it is enough to carry out calculations of the yield of the cycle on only the two isotherms, which is exactly what S. Carnot did in a long footnote;[12] here he uses advanced mathematics to calculate on only two isotherms, at t_1 and at t_2. work W, which therefore depends only on Q/t_1 and Q/t_2.

Thus, by inventing that particular cycle (Carnot's cycle: two isotherms and two adiabatics), he may have realized that it is necessary to integrate not Q, but dQ / t, that is the heat evaluated with respect to the temperature level of the source, the said level being different for each isothermal transformation. Note that surprisingly Sadi Carnot almost always distinguished the two concepts Q and Q / t with two distinct words, "chaleur" and "calorique"!

When the transformations are reversible, the integrand function of the integral of dQ / t- is uniformly continuous; thus it is certainly integrable in constructive Mathematics. Therefore the calculation of the integral can be performed on any partition of the cycle, even taking a single example of partition of the definition domain, that of a network of Carnot sub-cycles (which is exactly how it is generally taught, though it cannot be justified through "rigorous" analysis).

Returning to the constructive way of obtaining the primitive function of the integral of dQ / t, let us note that with it we obtain a new thermodynamic quantity, entropy S, which, given that way it is constructed, is a state function. With its addition to the other thermodynamic variables, the state is actually defined with a set of variables that are well established theoretically (P, V, T, U, S) with respect to a set that a priori could be much wider. At this point, given any three variables, an equation of state holds for them (as is true for a perfect gas for which $PV = nRT$ holds for the three state functions P, V and T); therefore each of the three can be considered (as is the time t in mechanics) as a useful parameter for integrating and also deriving; and in fact it is at this point in the theoretical development of thermodynamic theory that the derivatives appear: the Maxwell relations,

---

[12] S. Carnot, *Riflections on the Motive Power of Fire* (1824), Dover, New York, 2005, fn. 18.

$$(\frac{\partial S}{\partial P})_t = -(\frac{\partial P}{\partial T})_P; \quad (\frac{\partial S}{\partial V})_P = (\frac{\partial V}{\partial T})_S; \quad (\frac{\partial P}{\partial T})_V = (\frac{\partial S}{\partial V})_T; \quad -(\frac{\partial V}{\partial T})_P = (\frac{\partial S}{\partial P})_T.$$

which define the thermodynamic potentials, ie the most important state functions of theoretical thermodynamics. Thus we see that in the development of thermodynamic theory the derivation operation comes only after that of integration unlike what occurs in mechanics).

1.9 The third principle as a choice of the type of mathematics

Finally, constructive Mathematics provides a very clear solution to the problem of the mathematical formula of the third principle. In this Mathematics in general the limit points of a limit process are not reached, but only approximated to; exactly as the physical content of that principle requires: "$S = 0$ is approximable, but there is no set of operations such as to achieve the null value". This is precisely the version of some authors, the most physically correct, while others state that *lim* for $T \to 0$ of S is 0; although $S = 0$ does not exist in Physics; but these authors rely on classical Mathematics, which provides no other formula.

Thus a principle of this kind acquires a new foundational meaning; it becomes a choice of the kind of Mathematics and therefore belongs simultaneously to the theory and to the method for constructing the theory. It emphasizes that when developing a theory it is necessary to choose a specific relationship between Physics and Mathematics. In the past physicists did not perceive such a necessity; in the 17th Century infinitesimal Analysis and glorious Mechanics were developed at the same time and this connection appeared extremely productive; this fact led to the belief that this kind of mathematics-physics relationship was the only possible one. If, on the other hand, we reflect on the fact that since this relationship pertains to the foundations of the theory, then every theory should discuss it and establish it as its first principle. As a consequence, the third principle whose content concerns the relationship between Physics and Mathematics – because it declares a choice on the kind of Mathematics (constructive Mathematics) –, in reality should be established first, with an explicit declaration of choice for constructive Mathematics. To state it briefly, it suffices to say that every Thermodynamic variable is essentially positive or that all thermodynamic functions are homogeneous; a statement that has already been proposed as the fourth principle of the traditional formulation.[13]

With this the traditional thermodynamic theory appears in a new light: certainly not as a theory that is already harmonized with some traditional mathematical theories, as is Newton's Mechanics, nor as a theory that (as Truesdell thinks) did not use calculus out of ignorance of its previous intensive use in Mechanics. Rather, thermodynamic theory develops by defining and specifying its mathematics step by step inasmuch as it establishes its fundamental concepts. Its mathematics is constructive, which, although recently formalized, coincides with the mathematics corresponding to physical operativism. In the history of theoretical physics this

---

[13] P.T. Landsberg, "Is Thermodynamics an Axiomatic Discipline?", *Bull. Inst. Phys. and Phys. Soc.*, (1964) 150-156. P.T. Landsberg, "Born's Centenary: Remarks about Classical Thermodynamics", *Am. J. Phys.*, 31 (1983) 842-845.

kind of mathematics had remained in the shadows because theoretical mechanics had made extensive use of the powerful infinitesimal analysis which, using actual infinity (AI), promised the resolution of every problem or at least an initial substantial theoretical model of it. But the resolution of foundational problems is not a question of technical power.

Ultimately, in this chapter the exploration of the history of Thermodynamics in the light of constructive Mathematics introduced a fundamental dichotomy, that between the kind of Mathematics using AI (the classic, both that of infinitesimals, and the "rigorous" of ε-δ) and the constructive Mathematics which uses only PI.

CHAPTER 2

HISTORY AND FOUNDATIONS OF CLASSICAL MECHANICS: THE CONFIRMATION OF THE TWO DICHOTOMIES

Having clarified that there is a choice of the type of Mathematics in the foundations of Thermodynamics, let us establish whether the same choice applies to the foundations of the prestigious science of Mechanics; at the same time we will recognize another fundamental dichotomy, that of the type of organization.

2.1 The choice of the type of mathematics in Newton's mechanics

> Absolute, true, and mathematical time, of itself, and from its own nature flows equably without regard to anything external … Relative, apparent and common time is a sensible time, the accurate measure (whether accurate or inequable) of duration is obtained by means of motion….
>
> Absolute space, in its own nature, without regard to anything external, remains always similar and immovable. Relative space is some movable dimension or measure of the absolute spaces; which our senses determine by its position to bodies; and which is vulgarly taken for immovable space…
>
> Newton 1687

> Scientists, as far as they could, have purged their construction of metaphysics;
> but for what they could not, they used it as a key to understanding the whole universe.
>
> E. Burtt, 1924

We have seen in the previous Chapter that Thermodynamics is generally based on a mathematics that is essentially different from that of Carathéodory's formulation, which is normal in theoretical Physics. We have also found that the pluralism of mathematics also corresponds to a pluralism of essentially different formulations of the same physical theory.

But is this a specific result of the Thermodynamic theory, a theory that many consider "strange" and too different from the others? Or is it an example of a radical theoretical dichotomy which concerns every physical theory?

To find an answer to this question let us look at Mechanics, the theory that most of all represents theoretical physics, and let us ask ourselves if there is the possibility of formulating it on the basis of constructive, rather than classical mathematics. Having found among the various formulations of thermodynamic theory one that uses constructive mathematics, our principal aim will be to study the history of Mechanics in order to find possible alternatives to the dominant formulation.

Before Newton there had been an intense debate concerning the foundations of physics. Suffice it to mention two of the protagonists of this debate: Galilei and Descartes. The extraordinary thing about Galilei's work is that he was fully aware of two dichotomies, the one relating to the type of Mathematics and the one concerning the type of organization; he highlighted them above all in his last two works, the most important. The *Dialogue Concerning the Two Chief World systems* (1632) is not only a discussion about Astronomy, aimed at understanding if it is the Earth that revolves around the Sun or viceversa, but also a discussion about the whole of reality, including the very organization of our way of seeing reality itself. In fact, he was very familiar with the model of the deductive theory, suggested by Aristotle for the organization of a scientific theory; as a young man he had followed lectures on the subject in Rome and had written notes (which are kept in the national edition of all his works). But he never organized his theory deductively; rather his last two works are written in the dialogic manner followed by Plato; or more precisely, in the last work, *The Discourses on Two New Sciences* (1638), deductive passages demonstrating mathematical results are intertwined with investigative passages in which problems are discussed dialogically. This choice is a sign that Galilei wanted to organize his scientific theory in a new way that differed from the dominant mode at that time, that of Euclid. Moreover in this last work he dedicates a whole day (the third) to discussing the decomposition of a segment into infinite rather than finite parts and concludes without having opted for either of the two possibilities. He therefore did not decide between the two alternatives, but showed that he was aware of them and wanted them to form the basis his science.

Descartes was also aware of these dualisms, but followed resolutely the model of the deductive organization of the theory and deplored the "lack of certain principles" in Galileo's work, that is, a deductive organization. Moreover he (like Huygens) opposed the use of AI, both in Mathematics and in Physics. It is therefore certain that Newton, in constructing his theories with the contributions of previous scientists, was also aware he was making choices that were more or less different from theirs.

Let us then examine the choices Newton made in constructing his physical theory, although he did not declare his choices or discuss them in his writings. Newton clearly chose AI, since he introduced into physics infinitesimal analysis, in which its fundamental element, an infinitesimal, was commonly conceived as the inverse of AI.

Note that recently (1960) it was shown that Newton's mathematics, infinitesimal analysis, was almost modern non-standard Analysis and that the number field adopted by Newton (that of infinitesimals, in modern terminology, hyperreal numbers), was broader than that of the ε-δ real numbers of classical mathematics.

Let us see precisely what implications Newton's choice of mathematics had for theoretical physics. We note that the three possible types of mathematics, non-standard analysis, "rigorous" calculation and constructive analysis, present three different interpretations of the basic mathematical concepts. Let us start with the concept of point.

The point in non-standard analysis is an infinitesimal, that is, "a point that tends to approach point 0"; instead in "rigorous" analysis it is an exact and isolated point (because in this Mathematics it is possible to identify exactly every single point belonging to the continuum); while in constructive analysis it is in general an interval, which however can diminish without limit. As a result, every physical concept, the definition of which derives from the concept of point, can have a different meaning in each of the three mathematics. For example, the first principle of Newton's dynamics implies that one is able to decide when a force in space is exactly zero (because even a force approximately equal to zero could lead to unpredictable results, given that the first principle does not define the effects of a force, which will be known only with the second principle). In constructive mathematics the problem considered above is undecidable (essentially because of the fugitive numbers), while in the other two types of Mathematics the concept of $F = 0$ is legitimate and is represented by a point within the line representing the range of values of force F.

Other examples of concepts with different meanings are those defined by *limit properties*. We have already seen that a limit can be a value which can be approximated to any desired level of accuracy, or be an ideal limit, going beyond any possible experimental data and therefore detached from them. For example, Newton attributed to matter the concept, obtained by idealization, of *a perfectly hard body* (so that its form is fixed and therefore does not rebound in a collision). This concept is well known to historians of physics, since the entire development of mechanics in the 18th century was hampered by this ideal concept. Another example is the case of the ether, a concept whose characteristics were obtained as ideal limit properties of the properties of bodies traversed by electromagnetic vibrations. Both of these concepts can be represented in both non-standard and rigorous analysis, but are unacceptable in constructive mathematics, because they are limits located beyond approximations.

Let us now analyze the experimental laws expressed through mathematical functions. Of course, every experimental law is always an approximate equality between approximate values of physical magnitudes. To make the calculations easier each Mathematics represents it through an equality. However, non-standard and rigorous mathematics idealize approximate equality in an exact equality (notice that this idealization does not undermine experimental measurements, which are incapable of appreciating the difference between the idealized value and the experimental result). Constructive mathematics, on the other hand, cannot in general establish whether $f(x) = 0$ and therefore always considers equality to be approximate, yet with an increasingly perfectible approximation. Ultimately, this Mathematics retains all the heritage of experimental laws accumulated so far by physicists. It has, however, the problem, as we have seen in Thermodynamics, of whether the principles are constructive or not, which, being sentences that synthesize myriads of experimental facts (e.g. a body in motion continues its motion until a force changes its motion), are always idealizing propositions regarding them.

The various Mathematics may differ from one another even in relation to some mathematical functions used in theoretical physics. In fact, some functions may exist in rigorous analysis and not exist in constructive analysis. Consider for example the function that Dirichelet defined as equal to 0 on rational numbers and to 1 on irrational numbers (it is approximated by means of the limit of the expression $[\cos(2\pi n!x)]^m$ as n and m approach ∞).

In constructive mathematics it cannot be decided whether a given number x is a rational or an irrational number because the resolution of this problem requires an examination of an actual number of digits and the algorithms to decide this problem (e.g. in the cases of π, $\sqrt{2}$, etc.), given that each one is composed of a finite number of rules of construction, they in total at most a countable infinity, that is, an infinity less than that of the numbers of the continuum: therefore a function of this type cannot be defined. In the other two Mathematics it is perfectly legitimate.

Are there examples of similar functions in theoretical physics? For example, are there Newtonian functions based essentially on actual infinity? Clearly, already at the level of variables, absolute space and absolute time have precisely this characteristic, both because they are ideal limits of experimental concepts and because they want to consider every single point, both spatial and temporal, with infinite (actual) precision. At the beginning of the paragraph the magniloquent expressions with which Newton defined these magnitudes as absolute are in contrast with "common" measurable and experimental magnitudes. More than two centuries later, the dramatic crisis of the concept of the ether led physicists to abandon these undecidable (and even non-operational) functions.

Let us now analyze the differential equations of a physical theory. Consider Newton's differential equation for classical mechanics. Here both non-standard analysis and "rigorous" mathematics constantly search for new techniques to solve the most difficult problems and in the hope that they will all be finally solved. But constructive mathematics shows simply that this differential equation cannot always be solved with a general algorithm. Let us look at the simplest case.

First note that the differential equation typical of Newtonian dynamics

$$m\frac{d^2x}{dt^2}=f\left(\frac{dx}{dt},x,t\right)$$

can be translated into $y=dx/dt$ and $mdy/dt=f(x,y,t)$
so that it is then possible to proceed with the study of first order differential equations alone.

In constructive mathematics it is shown that this type of differential equation is solvable in general when f is Lipschitzian and is uniformly continuous. In classical mathematics, on the other hand, it is enough that that function is pointwise continuous on the interval of definition. Then with the Heine-Borel theorem it is shown that the function is also uniformly continuous. However that theorem makes use of the Zermelo axiom and therefore it is unacceptable for constructive mathematics.

The case of pointwise continuous but not uniformly continuous functions is therefore incompatible with the use of constructive mathematics. In Mechanics we have such functions when we consider force functions that are not limited: for example, central forces (such as gravity), which for r = 0 give infinite force. Then there are discontinuous forces, those that occur in instantaneous impacts. Thus, when constructive mathematics is adopted, the differential equation of the second principle of Newton's dynamics is in some cases undecidable, i.e. it cannot be solved with the same general algorithm holding for all other cases. In the cases just mentioned of unlimited or discontinuous forces, the equation can still be used, but only as a heuristic principle. In conclusion, it is capable of suggesting the resolution of many mechanical problems, but not all problems.

It follows that this differential equation can no longer represent a principle-axiom, which by its nature should be universal, that is, cover all real physical cases. Because of this fact, constructive mathematics does not accept Newton's formulation, which furthermore is based on an incompatible inertia principle since it requires f = 0 with absolute precision. Actually, the metaphysical character of the Newtonian formulation had already been pointed out as soon as it appeared by famous scholars: Leibniz, Berkeley, and then, subsequently, by D'Alembert, L. Carnot, Lagrange, Mach, Poincaré.

If we then examine some of the best known of the other formulations of Mechanics, we find a seemingly discouraging result. The variational formulations (Maupertuis' minimal action approach, the various minimum principles, Gauss's least squares method) require solving an equation of the type

$$Adt=0.$$

But we know that $x = 0$ exactly is undecidable, equality of a function with 0 more so and that of a functional even more so. All these formulations are therefore unacceptable for constructive mathematics. The Lagrange formulation based on a differential equation that is even more complex than $F = ma$, certainly repeats the problems of the preceding differential equation.

In conclusion, the differences presented by the three different types of Mathematics lead to different ways of formulating the basic principles and concepts a physical theory, so that it may happen that a particular formulation of a physical theory agrees with only teo types of Mathematics: for example, Newtonian Mechanics is only compatible with classical mathematics and non-classical analysis. And it can be the case, as in classical mechanics, that there is apparently no acceptable formulation in constructive mathematics. In the following we will see however that a more accurate inspection of history of mechanics will suggest a formulation of classical mechanics whose mathematics is constructive.

2.2 Leibniz's search for an alternative to Newton's mechanics

The principles of Newtonian Mechanics were almost exclusively formulated by Newton. They were formulated almost three centuries ago and for two centuries remained the foundation of a science, Mechanics, which in the 1700s included more and more fields of phenomena in its theory, so that for a long time it was thought that physical theory was simply the theory of mechanics. One consequence was that Newton was seen as the Aristotle of the modern era; no one more than he had such a lasting influence on the mentality of modern man.

But constructive mathematics cannot accept Newton's formulation, based on infinitesimals and differential equations; and usually in Mechanics (unlike Thermodynamics) no one is aware of any other formulation that is so simple mathematically as to be directly recognized as being based on constructive mathematics. For this reason it is worth carrying out a supplementary investigation into the history of physics to see if there have been attempts, and with what scope, to formulate a mechanical theory different from Newton's.

For a long time it was extremely difficult to evaluate the scope prefisco extent of the thought of Newton's contemporary, Leibniz (1646-1716) on the subject of dynamics. His scientific contribution was certainly wide and profound. This contribution includes the mathematical theory of infinitesimal analysis (of which he was perhaps the sole inventor); moreover, he laid the foundations of modern mathematical logic. But his theory of mechanics seemed either too naïve (because based on very elementary examples of mechanical motions) o metaphysical (because including principles and considerations of this nature) or very sophisticated (in the applications of infinitesimal analysis to some problems).

The problem of an accurate appraisal of his contribution to mechanics stems from various factors. Leibniz was enormously productive and the material to be taken into consideration by an historian is very broad, with manuscripts still to be published. Furthermore, Leibniz's thought evolved: he was first a Cartesian, then a follower of Huygens, and finally his thought underwent an independent development. In addition, Leibniz's dynamics, compared with the contemporary Newtonian theory of 1687, is clearly different, but incomplete, so that it requires a specific analysis. Moreover, Leibniz's whole scientific programme remained incomplete, since he believed that infinitesimal analysis constituted merely one part of that very general mode of reasoning by symbols that he had envisioned. Finally, let us add the fact that Leibniz was considered by most commentators to be the great metaphysician (he is usually remembered for saying that everything is conceivable with the metaphysical "monads" and finally that this world is "the best of all possible worlds", etc.). The consequence was that a proper study of his scientific texts in order to find a coherence in his thought was not carried out. A sign of this lack of interest is the fact that several of his mechanical works long remained untranslated into other languages (e.g. Italian).

It should be noted, however, that several times after his death, Leibniz was partially re-evaluated, acknowledging his anticipation of many subsequent scientific ideas. For example, in the nineteenth century his anticipation of the energy conservation law was recognized; and in the XX century his anticipation of modern mathematical logic, of Relativity and then of the foundation of theoretical physics on symmetries became manifest.

Leibniz was always clearly opposed to Newtonian Mechanics and even to the very name mechanics (it is he who by way of opposition invented the word "dynamics"), as he was to the concept of attraction at a distance (accused of being a kind of Aristotelian "occult quality", given that Newton left it unexplained) and to the whole logical structure of Newton's theory. Already in response to Descartes regarding the correct mathematical notion of *vis viva* he had launched the programme of a "dynamic reform", which he pursued with even more determination after the publication of Newton's *Principia*, which he never accepted. Thus Leibniz wrote to Honoratus Fabri:

> We should have inquired whether complex natural phenomena could not be derived from other known and studied phenomena. It is in fact useless to assume possible causes in the place of true causes, when the true and certain causes are before our eyes. So I believe that with my example more intelligent things could be proposed for research; to treat natural philosophy [read: theoretical physics] in the future without imaginary hypotheses and presuppose such causes whose actual [read: experimental] reality is well established in nature. No one, as far as I know, has so far tried to explain phenomena starting from phenomena, to explain *the congery of particularities by moving from a few general phenomena*, which is the real demonstrative procedure of physics.[1]

The following will be a reconstruction of his mechanics, obtained by extracting the most significant parts from his last works, those written between 1690 and 1698.

2.3 Leibniz's principles

Leibniz stated strongly that there are two general principles of human knowledge, the first is the *principle of non-contradiction* (it is absurd that both *A* and *not A* at the same time);

[1] G.W. Leibniz 1677, "Letter to H. Fabri" (my emphasis)

the second is *the principle of sufficient reason*, which he stated as follows:

> nothing is without reason, that is, every truth may be demonstrated [by means of an] a priori [method, i.e. it may be], deduced from the concepts of its terms, although it is not always in our power to arrive at this analysis....[2]

Note that Leibniz's last sentence explains clearly that we are forced to state the first sentence, which has two negations, because it is not always possible to establish with certainty the derivations we are looking for (this is the meaning of the second sentence).

A doubly negated sentence is used in this quotation. This occurs in many areas of science, but has so far gone unnoticed, one reason being that in natural languages there are ambiguities in the use of a double negation. A careful examination is therefore needed to understand when a doubly negated sentence coincides with a single, but reinforced, negation ("I do not want go no further "), or it is a merely emphatic affirmative expression ("I don't have more than one euro"), or, it is a hidden double negation because one negation is understood (for example: who would ever object ?[= nobody objects] It is only apparently false that ... [= it is not true that it is false])

Let us begin to clarify the meaning of the statements for which ¬¬A 🎭 A does not hold by asking ourselves what their domain of definition is. It is a fuzzy set. In such a case ~~then~~ its double complementation does not bring us back to the initial set, because we do not know if the points of the blurred contour of ¬¬A are the same as those of the blurred contour of A. We see here another difficulty that scientists have encountered in recognizing the above sentences as scientific: imprecision. Such imprecision seemed foreign to mathematics and rigorous science, which in general idealize reality with clear concepts, even at the cost ~~perhaps~~ of detaching themselves from reality. Until a few decades ago mathematicians stated with Hilbert: "In Mathematics there is no such thing as the sentence <We will not know even in the future>"; sooner or later everything will definitely be decided and therefore any indecision and imprecision will come to an end. However, 1900 mathematicians discovered infinite undecidable problems in the foundations of all mathematics).

At the level of logic, the statement sentence proposition $\neg\neg A$ which does not imply A, concerns a set or a physical reality that we do not know how to decide completely: we cannot state A; but, on the other hand, it is absurd that it is not A. In science there are many situations of this kind. Traditionally, the theorist has tried to avoid them, looking for the certainty of a precise statement, as if every statement of the theory could be decided experimentally either positively or negatively, in all possible cases. In fact in theory very often some concepts and some situations are idealized (we have already seen some principles used in the various formulations of Thermodynamics; in mechanics Newton idealized space as absolute and time as absolute, force as a cause, etc.).

Nothing strange then if a scientific theory, using idealized concepts, reduces the propositions $\neg\neg A$ to A, or to something similar to A. For example in the collision of the bodies the time of impact is very short compared to the duration of the trajectories. Therefore the statement "It is not true that t is not zero" holds. The search for certainty here suggests reducing the sentence to $t = 0$, which in itself is an absurdity; but if we then assume that in practice we must take

2 G.W. Leibniz 1686, "Letter to Arnaud", July.

into account that $t = 0$ is always approximable with very small t values, all is correct. (This 'all is correct', even if through a restriction that remains implied, is actually taken by theoretical physicists as proof that idealization is unrestrictedly valid!).

Now, let us look at the question from the point of view of the type of logic. In classical logic the law $\neg\neg A = A$ holds. In what logic are we if this law is no longer valid, as in our case? Extensive studies by mathematical logicians, aimed at reducing as far as possible the distance between classical and non-classical logic, have, however, had to conclude that this distance is irreducible, essentially by virtue of the aforementioned logical law, which either is or is not valid. It has been shown that the non-validity of this law is typical of (almost) all non-classical logics (from intuitionist to minimal).[3]

Here we see another difficulty in recognizing that a proposition ¬¬A does not imply A. In this case it belongs to non-classical logic, that is to a type of logic that has always been accused of being abstruse, or unrealistic, or deviant with respect to the ideal rationality of classical logic. Yet the opposite is true. To the conscious activity of the I, classical logic seems a source of certainty because all questions are decided through it with a clear distinction between "true" or "false". But it should be noted that in doing so, such a logic remains abstracted from reality and from life, where we almost always encounter situations that are complex, ambiguous and elusive, and therefore far removed from absolute precision; e.g. will it rain or not? Does he love or not love me? (this also applies to computers, based essentially on the approximation of real numbers and therefore unable to decide with a calculation whether π has infinite digits or not). But science, from Newton onwards, has always been presented as certain and without doubt about what has already derived from hard experimenta data. As a result, non-classical logics have always been considered not strictly scientific.

Returning to the principle of sufficient reason, its double negation introduces us to a logic that has no absolute certainties (or A or $\neg A$), but rather to an inductive, heuristic logic, which has not cut itself off from creativity: $\neg\neg A$ represents e.g. a future possibility. With words typical of the language of Leibniz: experimental science is not expressed by "necessary" propositions (that is, necessary to our mind, which is a direct creation of the Divine), as are mathematical truths (which, according to the tradition of that time, are the expression of Truth itself). Experimental science, on the other hand, is instead expressed by "contingent" truths, that is, they depend on imperfect external reality, as Leibniz says with precision: each of them is such that its contrary proposition does <u>not</u> imply <u>contradiction</u> (once again note the double negation, which cannot be substituted by the corresponding positive statement; because the latter would promote contingent truths to priori, i.e. necessary, truths). Today we express this concept by saying: every physical law is valid within certain limits (those of experimental errors); its contrary, entering the domain of values not yet established experimentally, merges the false with the possible and is therefore not absurd. In this sense it is contingent.

---

3 See the texts indicated for non-classical logic. The distinction was defined precisely in the work of D. Prawitz e P.-E. Melmnaas, "A survey of some connections between classical intuitionistic and minimal logic", in *Contributions to Mathematical Logic,* Amsterdam: North-Holland, in H. A. Schmidt, K. Schuette e H.-J. Thiele, eds., 1968, pp. 215-229; J.B. Grize,"Logique", in J. Piaget (ed..): *Logique et la connaissance scientifique*, in *Encyclopédie de la Pleyade*, Paris: Gallimard, 1970, 135-288, 206-210.

In particular, physics cannot claim to be based on perfect equalities and absolute identities, but rather on the search for the "indiscernible" (another double negation, when it is noted that "discerning" means distinguishing, declaring "unequal"), i.e. an approximate equality. That which is stated above eliminates at the root the preconception that traditional theoretical physics has maintained for centuries i.e. that there is only the principle of non-contradiction and therefore only classical logic. It is possible to reason through non-classical logic and this can also be done in Physics.

These aspects of mechanical theory, which belong to theoretical physics, may seem to involve philosophy. If we look for a principle of Leibniz's that is only physical in nature and which is linked to these ideas, we find that general principle, which Stevino and Huygens had already enunciated and applied fruitfully to establish laws of Physics: the impossibility of perpetual motion. We note that the statement of this principle is still a double negation (since "perpetual" means "without end", as Stevino rightly pointed out). Therefore the statement of the impossibility of perpetual motion cannot be an a priori proposition, nor is it evident to reason, but is the product of an enormous number of common experiences, and is proved by a *reductio ad absurdum* ("If it were possible, then the whole world would be different from the way it is; for example, work would be done without expenditure of energy"). Therefore this principle is ~~perfectly~~ connected with the philosophical principle of sufficient reason, because both follow the same non-classical logic.

2.4 Fundamentals of a dynamic of interaction

The impossibility of perpetual motion unifies the principle of sufficient reason with the essentially heuristic method of subsequent Leibnizian physical theory. Precisely because perpetual motion is impossible, the causes of an observed motion have to be sought. Thus, in his physics the concept of cause is not metaphysical, but is derived from a real impossibility. This theoretical orientation, therefore, does not accept the concept of force-cause presented as metaphysical by Newton, who considered gravitational force as the only force in the world because it was the expression of God's intervention in all places. Moreover in $f = ma$ the force is impossible to define without implying an arbitrary assumption: the static force is equal to the dynamic force; for example, if the wire of a dynamometer that supports a body breaks, the force that is exerted on the free body is the same as that acting on the body when stationary. This is an assumption because we have no means of comparing the static force, defined through other static forces, with the corresponding dynamic force, which is defined with the dynamic magnitude, acceleration Instead, according to the development of Newton's theory it is precisely the concept of force that should lead us to explain precisely the transition from the static to the dynamic of in the theory of Mechanics.

The concept of force is present also in Leibniz's writings, but his is an abuse of language (which other authors will continue to repeat until the end of 1870), because he applied this word to the concept of kinetic energy, or other types of energy. Moreover in his theory the specific concept of force is useful only if it has an important specification: theoretical reversibility in the transition from force to effects: "*Causa aequat effectum*" (not as in

Newtonian force-cause, which, wanting to represent God's intervention in the world, makes the transition from a metaphysical cause to a physical effect); from the effects it must be possible to completely reconstruct the cause.

Much less did Leibniz want to link his concept of force to the mathematics of differential analysis, knowing well the artificiality and idealized nature of the infinitesimal (he called them "beings of reason"), in contrast to the contingent-experimental character of physics. Instead he made extensive use of combinatorics, algebra, and the idea, still not formalized, of vector quantity; it was on this last idea that he based his polemic regarding the theory of impact of Descartes, who believed that the magnitude of the total momentum, considered as a scalar *mv*, was conserved, whereas Leibniz correctly pointed out that the actual conserved quantity is the total . considered as a vectorial quantity.

Moreover his contingentist conception, based on imprecision, does not ascribe primary importance to Newtonian forces which bring together in an idealizing fashion all the various influences of the external world on the body considered as uniform but rather to the interactions of bodies with one another ("everything is connected"); and in these relationships, the principle of action and reaction. Thus the scientific method of Leibniz focusses on bodies together with their interactions that are objectively given by their reciprocal collisions; in other words, its method is not analytical (analysis of a complete system by breaking it down into its elementary bodies and then reconstituting the system with those elements), but globalistic.

In this perspective (considering the system as a whole) it is natural that the principle of inertia also has a different version from that of Newton; it can indeed be enunciated, precisely according to the double negation of the principle of sufficient reason, through the doubly negated phrase: "The <u>in</u>-<u>difference</u> of bodies to rest or to motion". Which, at least according to Koyré, is the true theoretical content of this principle (and not Newton’s animistic "persevering").

This perspective, not anchored to absolute truths and essentially heuristic, is also devoid of all the other Newtonian certainties, or rather it is in contrast with them: for example. *space and time are not absolute*. Leibniz's controversy with Clarke (representing Newton’s position) regarding these concepts is famous. Space for Leibniz is not the "sensory [ear] of God", it does not exist a priori, it is a relative concept, the order of relations between things themselves and it exists only in so far as things exist. Likewise, time, which is simply an ordering between events, is also relative.

If there is no longer absolute space, the concept of space is relative to the observer and it then becomes natural to consider the composition of different movements as a basic theoretical fact. The composition of movements is clear to Leibniz when one observer is in motion with respect to another. We can recall the classic example of the boat in uniform motion, on whose deck the movements take place as on land, while from land their values are obtained by adding the speed of the boat to the speeds as measured on the boat. Furthermore, the vectorial nature of velocity is clear and therefore the principle of composition of velocities, the so-called parallelogram, is also clear.

2.5 Theory of impact of bodies

In particular, the field of phenomena that for Leibniz is most attractive and stimulating is that of the *phenomena of impact*. Impact is a terrestrial phenomenon that is extraneous to the abstractions of astronomy (science without experiments repeatable by the scientist). Moreover, the phenomenon of impact is not comprehensible within the Newtonian schema, both because the equation $f=ma$ is not applicable (since in this case the forces are instantaneous, i.e. they represent singularities of differential equations), and because Newton had suggested a schematization that Leibniz did not share: according to Newton, God created the world through perfectly hard bodies, such that they are invariable in form even in the most violent collision. Note that hard bodies always maintain their shape, do not deform and therefore cannot rebound. This makes it impossible for them to conserve the kinetic energy they had before the impact. Leibniz rejects this type of ideal body; emphasizing rather the role of *elastic bodies*, which, by changing their shape on impact, conserve energy and therefore can rebound on impact.

Leibniz also stresses that "our mind seeks conservations" (Descartes also followed this idea, justifying it with the conservation of the movement impressed by God at the time of the creation of the world). In other words, the many theoretical possibilities of understanding reality induce our mind to look for something stable, invariant. The controversy with Descartes leads him to consider as the main invariant not the *vis* (understood by Descartes as mv) but rather the *vis viva*, mv^2, that is, without the constant factor, today's kinetic energy, ½ mv^2. This technical contribution is universally recognized: it is he who ascribed far greater importance to mechanical energy than Huygens, his physics teacher, had done.

Leibniz was therefore able to write the fundamental laws of the theory of collisions of elastic bodies as conservations of certain quantities, the invariants of motion. These laws had been anticipated by Huygens, but Leibniz was the first to give them prominence and generalized them as follows (leaving out the vector signs):

$$1)\ \textit{principle of relativity of motion}: v - V = -(v' - V') \quad \textit{i.e.}\ v + v' = V' + V$$

$$2)\ \textit{conservation of momentum}: mv + MV = mv' + MV' \quad \textit{i.e.}\ m(v - v') = M(V' - V)$$

$$3)\ \textit{conservation of kinetic energy}: mv^2 + MV^2 = mv'^2 + MV'^2 \quad \textit{i.e.}\ m(v^2 - v'^2) = M(V'^2 - V^2)$$

Although I put together all three equations for the sake of beauty and harmony, two of them would suffice, since, taking any two of these equations, we can derive the other.

Thus the first and second give the third as follows. From the first we have $v+v'=V'+V$; from the second we have $m(v-v')=M(V'-V)$; multiplying LHS's and RHS's of these and equating, $m(v+v')(v-v')=M(V+V')(V-V')$; which gives $m(vv-v'v')=M(VV-V'V')$, i.e. the third equation.

Similarly the first and third give the second, since dividing LHS (RHS) of $m(vv-v'v') = M(VV-V'V')$, which is the third, by LHS (RHS) of the first, $v+v' = V'+V$, and equating, gives $m(vv-v'v')/(v+v') = M(VV-V'V')/(V'+V)$ from which it follows that $m(v-v') = M(V'-V)$ i.e. the second equation.

Finally, the second and third equation give the first, since the third, $m(vv-v'v') = M(VV-V'V')$, divided by the second, i.e. by $m(v-v') = M(V'-V)$ gives $m(vv-v'v')/m(v+v') = M(VV-V'V')/M(V'+V)$ which gives $v+v'=V'+V$, which is precisely the first equation.

I will add only one observation, namely that many distinguish between [perfectly] hard and soft bodies, and again distinguish the hard bodies between elastic and non-elastic hard bodies; and for each type of body they construct different laws. But it can be conceived that in nature bodies are hard-elastic [or rather more or less elastic] .... Well, Nature needs this Elasticity of bodies to obtain the execution of the great and beautiful laws that the its infinitely wise Author designed; among which those two laws of Nature, which I was the first to make known, are not of lesser importance; of which *the first is the law of conservation of energy...*; and *the second is the law of continuiy,* by virtue of which, among other effects, every change must take place in imperceptible steps and never by leaps. This means that nature does not admit non-elastic hard bodies [that is, ideal perfectly hard bodies]....

However, it must be admitted that although bodies must be elastic in nature, in the sense I have just explained, nevertheless elasticity often does not appear much in the masses or bodies we use, even when these masses are formed by elastic parts: they resemble a sack full of hard balls, which also give way to a modest impact, without restoring the shape of the sack, as do soft bodies or those bodies that yield without sufficiently recovering [their initial shape]. This is because their parts are not sufficiently united to transfer their changes to the whole. From this it follows that during the collision of these bodies a part of the energy is absorbed by the particles that make up the mass, without this energy being restored to the whole; and this must always happen when the compressed mass does not recover [its initial shape] perfectly. [This is true] even if it often happens that a mass proves to be more or less elastic depending on the different types of collision; it is also proved by water, which yields to modest pressure yet causes a cannonball to rebound.

When the parts of the bodies totally absorb the energy of impact (as when two pieces of loam or clay collide), or in part (as in the collision of two wooden balls, which are much less elastic than two balls of marble or tempered steel); when a certain amount of energy is absorbed by the parts, there is an equal loss of absolute energy and relative speed; that is, [this is the case] in the first and third equations; which are no longer true, since what remains after impact has become less than what existed before the collision due to that part of the energy that was diverted elsewhere. But this does not at all concern the amount of progress [read: vector momentum], or the second Equation. Moreover the movement of this total progress is maintained, it alone, even when the two bodies [being perfectly soft bodies] after impact accompany each other with the speed of their common centre of gravity, as do two balls of loam or clay. But for the semi-elastic bodies (such as two wooden balls) it also occurs that after impact the bodies move away from each other, even if with a decrease of the first Equation, owing to that energy that was absorbed by them on impact causing them to be deformed. And, based on experiments on the degree of elasticity of this wood, what should happen to two balls of this material in any type of collision or impact could be predicted. However, this total energy loss, or this defect of the third Equation, does not derogate at all from the inviolable truth of the law of conservation of the same energy in the world, since what is absorbed by the particles is by no means lost in the universe, although it is true for the total energy of the competing bodies.[4]

Note the description of impact with practically modern concepts, and above all how the last sentence prefigures conservation of energy other than merely mechanical energy.

One novelty of this approach to mechanical theory is that the phenomenon of impact does not concern the trajectories of points in space, but only the invariants expressed as a function of the velocities of bodies; consequently the impact equations must be solved only for velocities. The previous conservations lead therefore simply to algebraic equations in unknown velocities. This agrees with Leibniz's rejection of infinitesimal analysis in the Foundations of Physics and substantiates that relationship with elementary mathematics that

4 G.W. Leibniz, *Essay de dynamique sur les lois du mouvement*, 1698, in A. Drago, *La riforma della dinamica di G.W. Leibniz*, Benevento: Hevelius, 2004, pp. 122-131. Refer to the "Introduction" to this book for a more detailed presentation and illustration of Leibniz's anticipation of an alternative formulation of mechanics. For a general view on this mechanics see: "The birth of an alternative mechanics: Leibniz' principle of sufficient reason". in H. Poser et al. (eds.): *Leibniz-Kongress. Nihil Sine Ratione*, 2001, Berlin. voI. 1, pp. 322-330.

Leibniz had envisioned for theoretical physics.

It should also be noted that for Leibniz the case of continuous forces can be obtained by passing to the limit of an infinite series of impulsive forces, that is to say, collisions. Therefore, all the physics treated by Newton can be subsumed under the formulation of mechanical theory based oncollisions.

However, it should be noted that the mathematical laws of Leibniz’s dynamics lack the mathematical treatment of the collisions of plastic bodies (which Wallis had given separately). Leibniz did not therefore obtain all the laws necessary for a complete theory of collisions of bodies.

Leibniz also studied terrestrial mechanical phenomena, of which those of heavy bodies are of particular importance. For ease of reasoning, Leibniz limited himself to the exemplary case (used by him many times) of a single falling body, which falls or ascends a certain distance. It is in this context that Leibniz perfected what he had already specified in his controversy with Descartes, a first definition of work: force times distance travelled, $f{\cdot}ds$ (in his language: power times velocity), but he is not able to take into account the angle formed between the two quantities.

Leibniz also studied the interaction between bodies according to the principle of action and reaction and Torricelli's principle: "The centre of gravity of a system of bodies subject to the action of gravity cannot ascend." Today we are well aware that this principle is a restricted (although more precise than several others used since ancient times) form of the principle of virtual work, the mathematical statement of which was formulated for the first time by a follower of Leibniz (J. Bernoulli), just one year after his death (1617) With it the theory of Leibniz, which was always in continuous evolution, would have had a more specific and formal physical principle than that of the impossibility of perpetual motion. With his formula he would have been able to achieve a complete reformulation (as we will see below).

2.6 A second method of founding mechanical theory: Lazare Carnot’s formulation

> *The sciences are like a beautiful river, whose course is easy to follow, when it has assumed a certain regularity; but if we want to go back to the source, we never find it, because it is everywhere; in a way it is spread over the entire surface of the earth: similarly, when we want to go back to the origins of the sciences we find nothing but obscurity, vague ideas, vicious circles; and we lose ourselves in primitive ideas.* (L. Carnot 1783)

In the following we will describe a formulation of Mechanics that was wrongly neglected until a few decades ago and would prove to be be very important: that of Lazare Nicolas Marguerite Carnot (father of Sadi Carnot, the founder of Thermodynamics), famous leader of the French revolution, head of the armed forces during the Victory (against the European monarchist armies, allied in 1793 to crush the newborn revolution), at the time well known in Europe as a geometer, analyst and theoretical physicist of Mechanics. Strangely it was then completely neglected, until in 1971, one of the major historians of physics, C.C. Gillispie, published an authoritative study which re-evaluated his formulation of mechanics and in general the whole of his scientific work. This revaluation, however, did not emphasize what follows below, namely that Carnot's Mechanics constitutes a real alternative to that of Newton, being a concretization of Leibniz's project (L. Carnot declared himself a follower of Leibniz in his works of Geometry and Analysis).

Carnot published two books on mechanics.[5] The first book considers machines in the most general terms possible, that is, according to the definition: "machine" is *everything that transmits movement*; and since according to the Carnotian concept space is full, then no mechanical phenomenon is outside such a theory of machines, which is therefore defined as "the science of the communication of movements".

Furthermore, Carnot clearly distinguished two different organizations of mechanical theory. At the end of the first edition and in the preface of the second he wrote some pages on the two ways of organizing the theory and on the "two methods of considering the principles of mechanics". The first method consists in considering the principles as metaphysical; in particular that of force, from which the whole theory is derived and is only verified at the end; the second method aims to organize a theory on the basis of simple and real ideas (remember Leibniz's letter to H. Fabri), and then considers all the principles of the theory in the same way as other physical and mathematical entities, as ideas derived from experience. We can add that the objective is, according to Leibniz's teachings, to obtain the invariants of motion that our mind is seeking. Carnot specified that the theory can be developed as an alternative to that of Newton's, which starts from the static and then proceeds to the dynamic. Instead one can, as he himself does in the *Essai*, proceed from the dynamic and then consider the static of system equilibria as a special case.

Probably he had to submit the second edition of the book to the judgment of those who, like Laplace, considered deductive organization and Newton's principles to be the pinnacle of theoretical physics. Carnot decided therefore, in his last work, to accept the challenge of organizing Mechanics also deductively from principles and starting from the static, but he did so in his own way.

2.7 *The seven hypotheses*

The first part of the *Essai sur les machines en général* is dedicated to *principles*, called "hypotheses", thus emphasizing the need for verification through experience; they are declared by the author to be "inductively derived from the best observed phenomena ".[6] Here, Lazare Carnot has the merit of being the first to state the principles of mechanics that are exclusively experimental (that is, in a way that is today considered necessary, especially following the profound crisis in the foundations caused by the failure of the notion of an

---

5 L. Carnot 1783, *Essai sur les machines en général*, Dijon: Defay (Italian transl. *Saggio sulle Macchine*, Napoli : CUEN, 1994; English transl. Berlin : Springer, 2020); L. Carnot, *Principes fondamentaux de l'équilibre et du mouvement*, Paris : Deterville, 1803.

6 R. Dugas, *Histoire de la Mécanique*, Neuchâtel:Griffon, 1950, p. 312.

abstract ether and Einstein's critical analysis of the concept of simultaneity with respect to space and time measurements). Furthermore, it should be noted that he declared that he did not want to go back to explaining many fundamental concepts since it would mean engaging in metaphysical discussions. Matter, time, space, motion, movement therefore must be taken as primitive concepts.

Let us then consider Lazare Carnot's seven hypotheses (p. 49).

> First hypothesis: A body, once put in a state of rest, would not be able by itself to leave that state, and, once set in motion, would not be able by itself to change either its speed or its direction.
>
> Second hypothesis: If different parts of any system of bodies in equilibrium are acted upon by new forces, which alone would be in mutual equilibrium, then the equilibrium of the system will not be disturbed.
>
> Third hypothesis: When several forces, whether active and passive, are in equilibrium with each other, each of them is always equal and directly opposite to the resultant of all the others.
>
> Fourth hypothesis: In a system of bodies the motions or motive forces, which destroy each other in each instant, can always be decomposed into equal and directly opposite pairs, relative to the straight line that joins the moving bodies to which they belong. In each of these bodies these forces can be seen as destroyed by the action of the other.
>
> Fifth hypothesis: The action that two contiguous bodies exert on each other by impact, pressure or traction, does not depend at all on their absolute velocity, but only on their relative velocity. That [action] which two bodies communicate to each other by means of bodies interposed between them, is gradually transmitted from one to the other by means of these intermediate bodies: in this way the action always resolves itself into a series of immediate actions between two contiguous bodies.
>
> Sixth hypothesis: The quantities of motion, or motive forces, which bodies impart to each other by means of wires or rods, are directed along such wires or rods; and those that bodies impart to each other by impact or pressure, are directed along the perpendicular to their common surface, coming out of the point of contact.
>
> Seventh hypothesis: When two equal bodies that collide centrally are perfectly hard [read: plastic], they always move in company after the collision; that is, along the line of their reciprocal action which, according to the previous hypothesis, is always perpendicular to their common surface and through the point of contact. When bodies are perfectly elastic, they separate after the impact with a relative velocity equal, but in the opposite direction, to that at which they were travelling immediately before the impact. If the bodies are neither perfectly plastic nor perfectly elastic, they separate with a greater or lesser relative velocity, according to the degree of elasticity.

(In the statement of the last hypothesis we have substituted "plastic" for "hard", because, according to Gillispie, that is the meaning of this last word in L. Carnot's theory).

2.8 The principle of inertia

We now turn to a detailed analysis of the hypotheses of Carnot's mechanics. He himself, at the end of his commentary on the first hypothesis, wrote about the "principle known as the law of inertia". However, Carnot's version differs markedly from Newton's.

One first difference between the two versions consists in the fact that Newton refers to all bodies at all times and in all places, while Carnot circumscribes the proposition to a restricted set of situations, in which it can be affirmed that a body is at rest or is in motion. These situations are indicated with a deliberately imprecise premise: "A body, once put at rest ...". It is thanks to this premise that Carnot's version avoids the problem implicit in Newton's statement, that is, being able to select a perfectly uniform and rectilinear motion, which would take place "as long as ..."; that is, it avoids the problem of having to decide when, on the path (possibly infinite) of the body, there exists a force which has an exact non-zero value. Therefore, in the principle of inertia, Carnot, correctly, does not require the verification of the absence of forces ($F = 0$) on the whole path, nor does he assert their presence.

Carnot is well aware of this. He argued that it is not possible to judge with certainty "if a motion is absolute [as does Newton since he is referring to absolute space and absolute time], or if there is a movement or a drag force" and that it took "a lot of effort to correct this error". Therefore the statement of the first hypothesis does not consider the general verification of the absence of forces nor does it claim to provide rules for the verification of the state of rest or motion. In general such verifications are impossible, and would be circular anyway with the definition of inertial reference system.

Therefore Carnot deliberately used the expression " A body, once put at rest ...", referring to the particular circumstances (experimental and / or theoretical) in which we know how to decide whether a body is at rest or in uniform rectilinear motion. This evaluation remains our judgment, of an empirical and occasional, rather than of a general nature. In fact, in the history of Mechanics, the principle of inertia has been applied only when, taking the earth, considered as stationary, as the initial reference system it was believed that it was possible to indicate approximately a state of uniform rectilinear motion or rest, with gradually more and more sophisticated corrections for more sophisticated situations. At this point we understand that the situation described by Newton's statement, the absence of forces, is an idealization of the real one, an idealization which he obtained both by passing to the limit F=0 exactly and by positing a reference frame as absolute.

A further problem, equivalent to the previous one (to establish if $F = 0$ exactly), consists in Newton's claim to establish exactly when a body is in a state of rest, and to distinguish it from a state of motion; that is, decide whether $v = 0$ exactly (not whether $v<$ 🎁!). Carnot's wording "Once ..." also avoids this problem of deciding whether $v = 0$ exactly, which is impossible in reality.

Moreover, since the Carnot statement requires, in the premise, that we have already somehow succeeded in defining operationally the state of rest or uniform rectilinear motion, this motion, once identified, *can* be used, in that particular situation, in such a way as to operationally define the inertial reference system and the instrument for measuring time. Then indeed, starting from Carnot's first hypothesis we can progressively construct the entire conceptual scheme of dynamics without any vicious circle. In Newton's Mechanics, on the other hand, the three concepts, posited at the idealized limit of experience, 'statement of the principle', 'reference system', 'clock' must be defined by referring one to the other, that is, forming a circular definition of such entities.

Carnot therefore suggested a version of the principle of inertia that fits very well with what has actually been and what is being done, while that of Newton (as also that taught today), which claims to establish universally valid properties, turns out to be a tautological proposition.

2.9 *Inertia and animate beings*

Carnot's statement goes on to state that in the above conditions the body cannot “by itself” change its state of motion. The expression goes back to Aristotle and it is legitimate to wonder why Carnot repeats it. In reality Carnot wants to exclude definitively the theory of *impetus*, according to which a body, if it is in motion, has previously had impressed on it a propulsive force. If one pays close attention, this theory is still implied in Newton's proposition, when he uses the (moralistic) words "persevere", or "continue" (it is precisely their meaning of an effort performed by the body that makes the imprecise words "as long as", discussed above, seem natural).

Let us ask ourselves if Aristotle's words could have generated confusion. He attributed capacity for autonomous movement only to beings endowed with a soul and denied it to inanimate beings. Carnot repeats the aforementioned words, but substantially disagreeing with Aristotle, because for him the principle of inertia can be applied equally well to animate beings, as he writes expressly.

> An animal is subject, like inanimate bodies, to the law of inertia; that is, the general system of the parts that compose it cannot on its own give itself any progressive movement in any sense. For example, if you place a horse on a horizontal plane [without friction, slippery], it will be impossible for the horse to impress the slightest movement on its center of gravity in any horizontal direction; nevertheless the horse has the capacity to advance each of its limbs in the direction it wants and this distinguishes it from inanimate bodies; but at the same time as it moves a paw to one side, another part of his body will recoil to the same degree, since in the internal system of this animal the principle of equality between action and reaction is as valid as in the inert matter; in such a way that it is only thanks to the friction of its legs against the ground on which it finds itself, that it manages to move forward, impressing on the earth itself a quantity of motion, imperceptible to us, equal and opposite to that which it acquires.[7]

It is the first time in the history of physics that such considerations are lucidly developed. In Carnot they derive from his refusal to attribute the variations of motion to forces of a metaphysical kind, in this case animistic. In fact when he says "by itself" he is stating that a body has no autonomous force, so that the explanation for changes of motion, whatever it may be, must be sought outside that body. Carnot himself maintains it:

> [...] Any body that changes its state of rest or uniform and rectilinear movement always does so as a result of the influence or action of some other body, to which it at the same time impresses

[7] L. Carnot 1803, p. 246.

> a momentum equal to and directly opposite to that which it receives [...]. Every body therefore resists its change of state; and this resistance, which is called the force of inertia, is always the same and directly opposite to the quantity of motion it receives.

This formulation therefore avoids the need in theoretical physics to discuss causes which are not acting on material bodies (metaphysical force-causes) and opens the way to the statement of the principle of action and reaction, which concerns precisely the experimentally verifiable interaction between material bodies.

2.10 Comparison of the versions of the principle of inertia in L. Carmot and Newton

Carnot therefore restricts his reasoning/formulation to what is experimentally observable, considering motion only in a finite space, that which can be controlled with measurements which take into account experimental error. Understood in this way, the principle of inertia indicates a method of investigation: if in the observed body there is a change of motion, other bodies should be looked at, in order to find those acting on it.

We also note that, like all methodological principles, it is precisely a doubly negated sentence: "... it <u>cannot</u> <u>change</u> ...." where the word "change" is to be considered negative because it requires an explanation on the part of the physicist regarding the state of rest.

This double negation is not merely a figure of speech, without importance for physics, because, unlike figures of speech, it does not possess an affirmative equivalent. Indeed if the two negations are removed we would have an abstract word, with no physical meaning. In fact we obtain precisely the affirmative sentence of Newton's version that corresponds to L. Carnot's doubly negated sentence: when from this the two negations are removed one obtains precisely the "perseveres" or the "continues" of Newton's version! Here we see the logical distance between the two types of organization that the two physicists used for their theories of mechanics.

Now let us consider for a moment the theoretical importance of a doubly negated proposition, as it is used here by L. Carnot. It is the consequence of a way of conceiving the whole theory to which it belongs. In fact, if a proposition of non-classical logic is translated (removing the two negatives) into the corresponding affirmative proposition of classical logic, we obtain unverifiable content; therefore in a physical theory it can be valid as long as the theory allows idealization. In an Aristotelian organization (of a theory) some principles may be very abstract compared to experimental facts so that some consequent deductions may also be non-operative with respect to reality, provided they are compatible with experimental data (for example, in Newtonian mechanics, absolute motion as derived from the fundamental concepts of space and time. This gives a large degree of freedom to the theoretical physicist.

Note, however, that there are chemical, physical and mathematical theories that have a different organization; they been formulated without abstract principles, but are each based on a problem (PO). First, classical chemistry (which wants to solve the problem of how many elements of matter there are) and, originally, thermodynamics (which wants to solve the problem of the efficiency of conversions of heat into work); in addition some mathematical theories, such as Lobachevsky's non-Euclidean geometry (which wants to solve the problem of the number of parallels).

The fundamental problem of such theoretical organization is expressed with a proposition for which the law $\neg\neg A \rightarrow A$ does not apply. In this type of organization $\neg\neg A$ cannot be idealized in A, because it would mean eliminating the problem. E.g. if it is stated that heat equals work, then there is no longer a problem; whereas the problem is expressed by the corresponding doubly negated proposition: "It is not true that heat is not work"; and therefore it is necessary to find how it is converted.

Thus the formal characteristic of the PO appears to be the presence of doubly negated propositions, beginning with the one that expresses the problem of the theory; they cannot be replaced by equivalent positive statements because there are no operational means to prove them. Then it is clear that, while in a deductive theory we deduce from the vertex of the first principles a pyramid of theorems, which succeed one another to infinity without ever closing their series, in a PO theory the reasoning is essentially cyclic, in the sense that, given the problem posed $\neg\neg A$ and the direction indicated to solve it, the theory reveals as many contents as possible of that proposition A which at the beginning could not be made true with the simple intellectualistic operation of suppressing the two negations.

Returning to the principle of inertia, note that in general in Lazare Carnot's version there is no:

a) force-cause;
b) infinite motion in time (t varies between $+\infty$ and $-\infty$);
c) perfectly rectilinear and uniform motion;
d) motion that takes place in a perfectly empty space;
e) space and time as absolutes.

These concepts, on the other hand, are present in Newton's version, which seems to postulate an infinite inertial rectilinear motion, which in fact would be a perpetual motion. Lazare Carnot's thinking is far from stating this type of infinite motion because in the commentary to the seven hypotheses he notes that there are always limitations due to other bodies, as well as inevitable energy losses. And therefore for him the fundamental principle holds (referred to in the "Preface" of the book *Principles*): "Perpetual motion is impossible".

We then conclude that Carnot omitted undue idealizations and, in particular, the concepts obtained by passing to the ideal limit concept of the experimental properties (for example the perfectly hard body, there exist bodies that approximate to it, but is not realized by any real body). Already in paragraph 1.6 it was said that this principle of the impossibility of perpetual motion can be translated operationally by stating that there is no finite set of physical operations (i.e. a machine) that allows us to achieve an endless movement. This principle corresponds therefore to the physical translation of the basic principle of constructivism, that is, the restriction of our ability to conceive something to the case where we can calculate it with precise operations.

Having established this connection, it is natural that Carnot's version of the principle of inertia can be expressed through constructive mathematics. The language of Carnot ("A body, once put at rest ...") seems to be taken precisely from constructive mathematics which would express itself in this way in the case of an undecidable problem in general, but solved in particular cases. For example, in constructive mathematics it is undecidable for which point a function $f(x)$ has an endpoint, but when the function is the quadratic expression of two variables, then we can establish the point for which *f(x)* has a null derivative by examining whether the discriminant is 0 (this problem is decidable in constructive mathematics because there is an algorithm to calculate the square root of the discriminant). In the same way, Lazare Carnot's statement means that we should restrict ourselves to those particular cases in which we are able, in some opportune manner, to actually determine whether $v=0$ o $v=cost$.

2.11 *From a static to a dynamic force; interactions*

The second hypothesis concerns the forces applied to the systems of bodies in equilibrium with each other. Here we are dealing with statics, where the forces are well defined and therefore are also accepted by Carnot. The idea of the principle in question translates the fundamental idea of Leibniz and D'Alembert (superimposing on the given system a system of forces with null resultant). In modern terms, we should add that the bodies are rigid, since otherwise, in addition to the indicated effects of translation and rotation, the effects of deformation or even breaking would also have to be taken into account . But this omission by Carnot has no influence on his theory, since in fact he does not then apply his hypothesis to deformable bodies.

The third hypothesis concerns the gradual transition from a static (equilibrium) to a dynamic situation (subject to forces). Carnot was able to establish this transition to the extent in which he also considers passive forces, that is forces opposed to the velocity of the body. From the Newtonian point of view of the metaphysical force-cause they are a contradiction in terms. Lazare Carmot was the first to define them correctly.

In such a system of forces in equilibrium each of them is equal, but in the opposite direction, to the resultant of all the others; the latter is defined by Carnot as that force which, evaluated in any given direction, is equal to the sum of all the others evaluated in the same direction. Note that the (still undefined) vectorial character of the forces is well highlighted here, although in the simplified form of specifying the mathematical representation of the vectors by means of their projections along a specified direction. Newton, on the other hand, is unclear on this point; e.g. he wants to deduce the well-known rule of the parallelogram of forces as a Corollary of his three principles, that is, improperly, from physical theory.

With the fourth hypothesis we pass to dynamics in the strict sense, since momenta are considered as well as the "motive powers", which however in the usual definition are each the product of mass by acceleration and therefore differs from the momentum by virtue of an operation of differentiation of the velocity. Here Carnot wants to consider on the same par the impact of bodies and continuous motion in a synthetic way.

The contribution of this hypothesis constitutes a broadening of the principle of inertia in that it expresses the conservation of the overall momentum for an (isolated) system of interacting bodies, rather than of one body.

From this principle derives the conception that any change in the quantity of motion derives from an impact (or a continuous series of infinitesimal impacts) with other bodies and that, provided that one extends the system to include the acting bodies, it is possible to abandon the Newtonian schematizaion of a system as a single body subject to unspecified forces coming from its environment. In reality this hypothesis is the traditional principle of action and reaction, as also stated in Carnot's last sentence:

> It therefore seems certain that in general, whenever one body impresses movement on another, it in turn receives the same amount in the opposite direction; [this occurs] at least as long as the collision is direct and is exercised between two bodies only. But the analogy suggest that the same thing must happen regardless of the number of bodies and whatever the directions of their movements. All the phenomena of nature confirm this important law, which is usually expressed by saying that the reaction is always equal and opposite to the action. (L. Carnot 1803, p. 60)

Note that Carnot speaks of "analogy" because in the statement of the fourth hypothesis he refers, for the first time in the history of theoretical physics, to a partition into pairs of quantities of motion (called, generically, "forces") which the bodies of a system exchange with each other. This partition is not an experimental operation; it is superimposed on reality, but it is compatible with experiments and everything appears as if it were actually being done. It is very important to express it as a specific principle (Some call it the fourth principle of Newton's dynamics).

Note that here also Carnot idealizes reality. This is not, however, suggested by an operation whereby an unattainable situation is approached as a limit, as is the case with Newton with the absence of friction or with the absolute void, but rather by mathematical operations which at present pertain to vector algebra.

2.12 Analysis of the principles of mechanics through mathematical logic

The previous critical comments on the principles of mechanics can be summarized and illuminated in modern terms by means of an simple formalism of mathematical logic. It will show that the principles of dynamics cannot have many versions and that the physicists mentioned have already explored them all, even if they were reasoning only intuitively.

It has been pointed out (Nagel) that all the laws of physics can be expressed by a statement preceded by two quantifiers: "Every" ($\rightarrow$), and "There exist one" (). A possible form of the current statement of the principle of inertia might be the following: "For every body *x* there exists a *y* (i.e. a system, composed of an inertial system, a closed system and a clock) such that if the body *x* is in *y*, then it is either at rest or is in rectilinear and uniform motion" (that *y* must include the three elements indicated is confirmed by every analysis of the principle of inertia).

The statement can be formalized as follows

$$A = \rightarrow x\ y : P(x,y) \qquad (2.1)$$

A the statement of the principle of inertia;

x a body;

y a compound system, formed by a closed system, which is inertial and in which a clock has been defined

P(x,y) a predicate regarding x and y: "if x is in y, then it is either at rest or in rectilinear and uniform motion ".

All this, however, belongs to classical logic, which does not concern itself with establishing whether the functions and operations indicated by the statement can actually be calculated (let alone declared operational). In this context the principle of inertia announces that *y exists* but does not explain how to find it.

At the beginning of the century a technique was suggested (by the logician T. Skolem) to formally translate the existential quantifier on the variable y of the predicate into a constructive function. It is a question of finding in the set C of functions of constructive mathematics a function $\alpha(x)$ such that it actually calculates for each x the corresponding y, and then replacing the existential quantifier $\exists$ with it.

$$A_{\neg} = \forall x P(x, \alpha(x)) \qquad (2.2)$$

Now (2.2) says that for every x the predicate P holds for x and for $\alpha(x)$: "Every body x is at rest or in uniform rectilinear motion in a system y given by $\alpha(x)$", where $\alpha(x)$ is a function that belongs to constructive mathematics and actually associates a compound system y to every body x

If we consider constructivism as physical operationalism, then to translate (2.1) into (2.2) we must find, for each body x, an effective procedure $\alpha(x)$ that, from the knowledge of body x alone, provides an inertial system, an isolated system and a clock. Two centuries of fruitless research convincingly demonstrate that it is not possible to derive so many entities from x operationally.

However the aim of relativizing (2.1) operationally can be achieved by forcing the predicate, that is through one of these choices:

1) Replace the quantifier $\exists$ in (2.1) by a constant value y_0. This proposition follows:

$$\forall x A(x, y_0) \qquad (2.3)$$

that is: "Every body x is at rest or in uniform rectilinear motion if it is placed in y_0"; that is, if its motion is measured with respect to a precise inertial system, in a precise closed system and equipped with a precise clock.

(2.3) corresponds to setting the clock and the reference system in two very different ways:

a) in the manner followed by physicists since the time of Galilei, that is, with an empirical clock, with a closed system experimentally verified and with the terrestrial reference frame, but changing them from time to time according to subsequent improvements for specific cases;

b) in the idealistic way suggested by Newton, that is idealizing the previous empirical method with the transition to the limit, a limit that transcends experience itself: introducing the ideal concepts of space and absolute time, which stipulate the clock once and for all (as an actualization of absolute time) and the inertial system (as an actualization of absolute space), also implicitly suggesting that we are always capable, in principle, of verifying whether our system is isolated or not, or whether $F = 0$ or not.

2) Accept the fact that we generally do not know the function ✔ *(x)*, but sidestep the problem of the existential quantifier by saying: "In particular circumstances that experimental physics can specify: →*x A(x,y)*", without giving further explanation as to what experimental physicists should do. This is what Mach does (in his famous criticism of Newton's principles that paved the way for Einstein).

3) Deny the physical importance of these problems, classifying them as metaphysical and state only what we are able to derive from the experimental observations of body *x*: the impossibility that the single body under observation changes its state of motion on its own when it is at rest or in rectilinear and uniform motion; ultimately, no quantifiers and →. This is what L. Carnot does, expressing everything through finite quantities only and experimentally.

We can also apply this technique of mathematical logic to the third principle of dynamics: "For every action there is an equal and opposite reaction". To express this principle we will not take into account the problem of determining in which inertial system it holds, since this problem will have been solved already with the first principle.

To translate it into a formula, we give the imprecise words "action" and "reaction" the only possible meaning, that of "forces". Then we can specify it this way: "For every force measured on a body *x*, there exists a force resultant of the forces acting on a system of bodies x_i such that:

$\vec{R}_x = \sum_{i=1}^{n} \vec{r}_{x_i}$, with $\vec{R}_x$ equal and opposite to $\vec{A}_x$. Formulated in mathematical logic:

$$\vec{r}_{x_i}\left[\left(\sum_{i=1}^{n} \vec{r}_{x_i} = \vec{R}_x\right) \wedge \left(\vec{R}_x = -\vec{A}_x\right)\right] \tag{2.4}$$

The principle says that if a physicist measures a force on a certain body x, he is able to derive the "reactions" on the other bodies x_i, which acted on the first. We know from the previous case of the principle of inertia that the crucial point of the formula (2.4) is the existential quantifier. In fact, the quantifier index is already a problem: how many bodies? Their number could be higher than any preset limit. For example, before the discovery of Neptune (1846) suggested by Leverrier's calculations, the actions of the bodies of the solar system seemed complete, that is, they seemed to come from a precise number of planets; later, however, due to unforeseen deviations of the known planets from the calculated trajectories, we had to add other planets. Then, since we cannot generally know the exact number of bodies involved in (2.4), we must conclude that the statement is metaphysical.

If then somehow we come to know that the bodies are *n*, with *n*> 2, what is the procedure required to find the reaction? To answer, it is necessary to know the vector calculation of forces. But, according to Newton, the parallelogram rule, instead of preceding principle III, is a consequence of it (Corollaries I and II). The Newtonian statement of principle III is therefore once again metaphysical, unable to tell us how to compose forces in order to discover the reaction for the case *n*> 2.

And if in any case the parallelogram rule is valid between the given *n* vectors, on the

Cartesian axes it gives three mathematical relations, which are added to the three mathematical relations ; in total we have on each axis two relations to solve a problem which, however, depends on n degrees of freedom (the bodies). There remain therefore ∞^{n-2} possible solutions. Clearly we will never have an operative procedure to determine in general a single solution.

Let us therefore try to replace the existential quantifier with a constructive function:

$$\forall \vec{A}_x \, \alpha\left(\vec{r}_{x_i}\right) \left[\left(\sum_i \vec{r}_{x_i} = \vec{R}_x \right) \wedge \left(\vec{A}_x = -\vec{R}_x \right) \right] or \; \forall \vec{A}_x \, \alpha\left(\vec{r}_{x_i}\right) \left[\vec{A}_x = -\sum_i \alpha\left(\vec{r}_{x_i}\right) \right] \qquad (2.5)$$

where the second and last term, $\alpha_i\left(\vec{r}_{x_i}\right)$, represent actual procedures for finding the various reactions (in the second case $\vec{r}_{x_i}$ and thus be able to calculate the equality of the second proposition of (2.5)). Given everything that has been said above this function $\alpha_i(\vec{r}_{x_i})$ clearly cannot be found.

Let us therefore try to solve the problem, as in the case of inertia, by establishing the values on which it is existentially quantified. This operation is equivalent to establishing the number of bodies and the procedure for finding them (in particular, assuming a priori the dependence of the force on distance *f(r)*). This is what is done experimentally.

If, on the other hand, we want to transcend experience and move to an ideally universal situation, then we can fix the number of bodies, certainly not an infinite number, but the simplest situation, two alone, hoping that then the procedure for the interaction of n bodies can be repeated in pairs. So in the case of two bodies, this addition gives us, assuming we know the vector calculus, an indication: the reaction sought is along the line of action, in the opposite direction to the action. All this constitutes the only, manifestly very small, indication that Newton's statement offers regarding the complexity of the problem.

But even in this way we do not know at what distance the body we are looking for is, unless we already know the force function $F(r)=1/r^2$ rather than, for example, $1/r^5$. Once again Newton's version of the principle solves the problem by assuming a particular force, gravity, as universal, that is, imagining that this is the only possible one in nature.

Now it is clear that Newton wanted to infer from a very particular case of interaction (n = 2 with a known force), a statement about a situation generalized to any force and to an unlimited number of bodies. That is, Newton has removed almost all the propositions that should constitute and limit (2.4), leaving only the last one. In it, , left alone and detached from any body, can only be understood metaphysically: it is considered a cause. Therefore also the first force is detached from the body it is acting on and becomes a force in a metaphysical sense; as a result the proposition is completely abstracted from bodies. Moreover, it is further generalized: the concept of force is replaced by the concept of "action ".

For his part, Mach says nothing of interest in this regard, although he uses this principle as the defining basis for the initial concept of mass.

Lazare Carnot, on the other hand, not knowing how to provide generally a constructive function ✔ that finds the reactions, correctly begins his third hypothesis by saying: "Once ..." we are able to establish experimentally that the equality between a set of forces ...; that is, Carnot restricts himself to those cases in which we can establish experimentally and conventionally that equality that Newton assumes established a priori. Moreover Carnot does not abstract from bodies, since he specifies that "force" is essentially the variation of the momentum, which is linked to the body by mass. Furthermore he limits himself to a finite region of space, and implicitly limits the number of bodies to what can be found there given the bodies have finite dimensions; that is, he limits it to a finite number (even if unspecified). Finally the vector calculation necessary to compose forces is correctly expressed by Carnot. (For this calculation he adds a further hypothesis to clarify his procedure: his fourth hypothesis suggests that between the forces in equilibrium everything occurs as if each pair of bodies is independent of all the others). Given all this he can state his third hypothesis as a correspondence between Physics and Mathematics, that is, between physical equilibrium of forces and vectorial calculation of opposing forces. Having done this he has exhausted both its mathematical and physical content, since the function ✔ summarizing a general method to find operationally the reactions of the single bodies, cannot be specified further.

It must be concluded that here once again Carnot has clarified the actual content of the third Newtonian principle and that, from the point of view of mathematical logic, no other type of solution to the problem of how to formalize the principles of dynamics can be expected.

2.13 Hypotheses on Impact

The content of the fifth hypothesis is that in a collision, or pressure or traction between two bodies, everything depends on the relative velocities of the bodies, not on the absolute velocities. This type of hypothesis concerns only interactions. So we have before us the first formulation of an entire physical theory based on interactions and not on the schema of causes and effects.

The second part of this hypothesis takes into consideration the interaction between two bodies (mediated by intermediate bodies) and states that it is always transmitted through bodies that are in direct contact with each another. Note that, according to the Cartesian conception of space, Lazare Carnot identifies it with matter. For him, therefore, there is no void, rather there are only more or less light bodies, all in contact with each other, direct or indirect. From this follows the complete generality of Carnot's position with respect to interactions between bodies, valid for any interaction (known at the time) in space. The Newtonian gravitational force might also be explained in terms of exchange of momentum between contiguous bodies.

In modern language, the content of the sixth hypothesis specifies the geometry of interactions. According to Carnot this can be derived either inductively from experience, or through reasoning alone. For example, in the case of impact, he argues that, with respect to the aforementioned perpendicular, there is no reason why the direction of the exchanged momentum is inclined in one direction rather than the other (this is an example of the use in Science of the principle of sufficient reason).

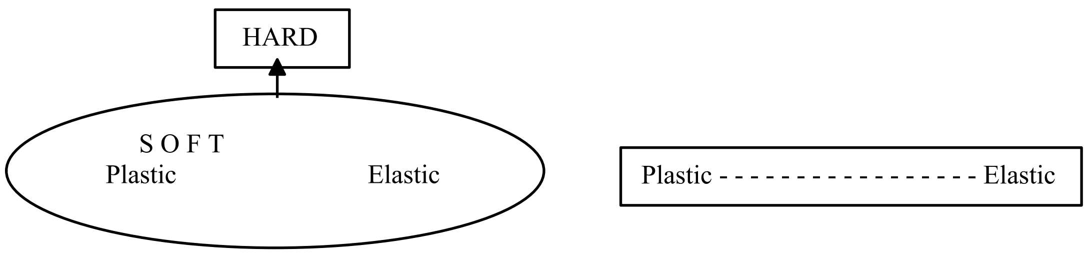
The seventh hypothesis expresses the behaviour of bodies in collisions. In this principle we note, in effect, a repetition concerning the direction in which the bodies interact through contact. Indeed the part of the seventh principle that deals with this problem could easily be omitted, since it was already contained in the sixth principle. But for Carnot repetitions merely detract from elegance.

This hypothesis is crucial, since the schematization of perfectly hard bodies hindered, for a century and a half after Newton, an understanding of conservation laws, in particular the conservation of energy. The following table 2.1 illustrates the ancient schematizations (of Wallis and Newton: on the left) and modern ones (on the right) of the bodies with respect to impact (see summary table below).

Carnot's hypothesis also includes elastic bodies, for which (if the masses are equal) the relative velocity after the impact is the same, but with the opposite sign of the one before impact, and partially elastic bodies, for which the relative velocity after impact has a lower absolute velocity value than it had before impact. This consideration is only qualitative, but it is specified by Carnot subsequently through an index of elasticity of bodies (as is done in our times). Consequently Carnot's schematization for colliding bodies turns out to coincide with the modern one: it considers perfectly plastic bodies and ~~the~~ perfectly elastic bodies, where bodies between those extremes can be specified by an index (the modern one varies between 0 and 1). This shows that he had gone completely beyond the Newtonian conception of perfectly hard bodies.

TABLE 2.1

THE TWO CLASSIFICATIONS OF BODIES WITH RESPECT TO IMPACT: NEWTONIAN AND MODERN



Wallis and Newton   Huygens, Leibniz and L. Carnot

	fixed form	*variable form*
Without rebound	**[perfectly hard]**	**plastic**
With rebound	------------	**elastic**

It should also be noted that by imagining a series of impacts or impulses any continuous force can (in principle) be obtained, so that the Newtonian case of continuous force is brought back to the particular case of impact. The fact that a process requiring a limit is needed for this step shows clearly that by doing so we make the transition from an operational and experimental theory like that of impact to a theory involving a limit, which can therefore lead to non-operational idealizations, precisely as occurs in Newton.

2.14 Critical considerations on the seven hypotheses

We observe that Carnot's seven hypotheses contain both the first principle (first hypothesis) and the third principle (fourth hypothesis) of Newtonian mechanics. But the second principle, $F = ma$, is missing. In fact, this formula is never used by Carnot (in any his works on Mechanics), since for him, as for D'Alembert, it represents a metaphysical cause; for both of them the force of the dynamics (that according to them precedes the static), F, is not equal to the static force and can therefore be defined by only identifying it with ma. In fact, there is no way to define the dynamic force of a body that had been suspended from a gauge for measuring its static force and then separated from it, since it is not known whether the dynamic force exercised on the body after the separation is equal to the previous static force.

Let us now see how Carnot's mechanics develops mathematically. The concept of mass has no operational definition in Carnot's works. In the *Essai* mass is not defined, while in the final work we read that the mass of a body is (repeating the definition of Descartes) "the actual space" that it occupies, or (repeating Newton's definition) its "real quantity of matter". Jammer criticizes Carnot for failing to define it,[8] but at that time the defect was common to all, particularly to the Cartesian school. Note that it was only a hundred years later that an operational definition of mass was formulated, when Mach put forward a definition, which nevertheless received heavy criticism, as Jammer himself notes; it starts from the equality $m_1a_1 = m_2a_2$ of the interaction of two bodies, yet in order actually to give the definition of a single, preliminary concept it assumes the third principle. This is not surprising, as so far it actually has not been possible to define the mass without the second or third principle. We therefore conclude that Carnot certainly did not solve the problem: he, like everyone else, fell back on an intuitive definition. He does, however, have the merit of having abstained from false solutions, such as those of current authors.

We now note the order of the seven hypotheses. Carnot, like Newton, begins with the one that signals the divergence of modern mechanics from the medieval theory of impetus. In Carnot's statement, the expression "itself" may recall Aristotle, but in fact that expression ("if not by the work of other bodies") emphasizes that Carnot's physics is the physics of interaction, as was Leibniz's and D'Alembert's theories.

The principle of inertia states that being at rest and uniform rectilinear motion are equivalent. But what does equivalent mean? Newton's statement treats the two cases as if they were the same ("... at rest or in motion ..."). Carnot's, on the other hand, is more cautious and

---

8 M. Jammer, *Concepts of mass in classical and modern physics, Cambridge*, Harvard University press, 1961, chp. 11.

separates out two parallel but distinct statements. It does not therefore consider the transition from static to dynamic to be obvious or easy. After the first hypothesis, the other hypotheses articulate this equivalence in gradual steps. While the second hypothesis still concerns static situations, the third and fourth finally include the dynamic We conclude that Carnot's hypotheses, after the first one, represent a precise strategy for the transition from statics to dynamics.

We also note that all the aforementioned hypotheses are constructive because they are essentially experimental, except the fourth one which is considered by Carnot to be a mathematical convention.

2.15 The development of mechanical theory: the index of elasticity and the conservation of kinetic energy

From the seven hypotheses of Carnot it is possible to develop mechanical theory in two ways (as specified by the author at the end of the *Essai* and in the preface of the *Principes*): as a theory based on the artificial Newtonian concept of force (that is, reformulating the Newtonian theory withour abstractions and undue idealizations) or as "the science of the communication of movements" (which is the path chosen by Carnot himself), which studies the ways in which movement is transmitted from one body to another through mutual interactions.

To follow the first possibility we know that it would be necessary to extend the static definition of force to dynamics (as does Newton), to postulate as an axiom $F = ma$ and then proceed using differential calculus. However, Carnot believes that mechanics, understood in this way, has two fundamental defects: the "metaphysical and obscure" concept of force-cause and the use of infinitesimal analysis, at the time openly metaphysical. (A book by Lazare became famous in Europe for its radical critique of infinitesimal metaphysics; he reduced this type of calculation to a mere operational technique). Like Leibniz, Carnot also did not want to introduce infinitesimal calculus into the foundations of Physics.

Let us now see how Carnot develops mechanics in the second manner indicated by him (the "empirical" manner of his first work, *Essai*), that is, as "the science of the communication of movements". He starts from an equation (the “first fundamental equation”) valid for plastic bodies (although called "hard" by him), whose derivation is clear: the principle of action and reaction between bodies (of an isolated system) results in:

$$\sum m_i \vec{U}_i = 0$$

where m_i represents the mass of the i-th body, $\vec{U}_i$ the velocity it loses in the interaction i.e. the difference between the velocity $\vec{W}_i$ before the interaction and the velocity $\vec{V}_i$ after. Therefore the equation is equivalent to $\sum m_i \vec{W}_i = \sum m_i \vec{V}_i$. Now for completely plastic bodies the final velocity $\vec{V}_i$ is the same for every body, hence it can be written formally as:

$$\sum_i m_i \vec{U}_i \cdot \vec{V}_i = 0$$

Carnot is able to apply the above equality to all types of bodies, introducing the elasticity index *n*. Reasoning by analogy he says that on impact the products $m\vec{U}_i$ change by a factor that ranges from 1 (for completely plastic bodies) to 2 (for elastic bodies). Indeed, given that $\vec{V} = \vec{W} - \vec{U}$, the preceding equation can be written as:

$$\sum_i m_i\vec{W}_i \cdot \vec{U}_i - \sum_i m_i\vec{U}_i^{\,2} = 0$$

Now dividing $\vec{U}_i$ by *n* in accordance with the above reasoning, we have a new equation:

$$\sum_i m_i\vec{W}_i \cdot \left(\vec{U}_i / n\right) - \sum_i m_i\left(U_i^2 / n^2\right) = 0$$

or

$$n\sum_i m_i\vec{W}_i \cdot \vec{U}_i - \sum_i m_iU_i^2 = 0$$

But since $\vec{W} = \vec{U} + \vec{V}$:

$$\sum_i m_i\vec{U}_i \cdot \vec{V}_i + \sum_i m_iU_i^2 \frac{n-1}{n} = 0$$

Then if *n*=1 we have the previous equation, valid for plastic bodies. Instead for *n*=2 we have the case of elastic bodies. In fact,

$$2\sum_i m_i\vec{V}_i \cdot \vec{U}_i + \sum_i m_iU_i^2 = 0 \ ,$$

which, by virtue of the trigonometric relation $\vec{W}^2 = \vec{V}^2 + \vec{U}^2 + 2\vec{U} \times \vec{V}$, reduces to the well-known conservation of kinetic energy, characteristic of elastic bodies:

$$\sum m_iW_i^2 = \sum m_iV_i^2$$

2.16 Geometric motions. The other conservations

Then to replace differential equations a mathematical calculation tool is introduced; it is based on the concept of "geometric motion", whose clearest definition is given in the *Essai*: "a motion assigned to a system of bodies is geometric if it is such that the opposite movement is also possible". The introduction of this concept allows Carnot to establish the second fundamental equation, to be applied to a system of interacting bodies (by impact and by contact mediated by wires or rods). This is the crux of his theory:

$$\sum_i m_i\vec{U}_i \cdot \vec{u}_i = 0 \qquad (2.6)$$

where $\vec{u}$ is the velocity of any geometric motion, attributable to the system. In other words, the velocity is an indeterminate magnitude and any specification of it gives rise, in the relation (2.6), to an equation that is applicable to the system.

For example, we assign to u_i the same value u = const (we can do this, since a uniform translation of an entire system of bodies is certainly a geometric motion); we have:

$$\sum_i m_i\vec{u} \cdot \overrightarrow{U_i} = 0 \Rightarrow \vec{u}\sum_i m_i\overrightarrow{U_i} = 0$$

From this, because $\vec{u}$ is arbitrary, it follows that:

$$\sum_i m_i \overrightarrow{U_i} = 0 \; i.e. \sum_i m_i \overrightarrow{W_i} = \sum_i m_i \overrightarrow{V_i} \, .$$

This is precisely the conservation of the total momentum of the system.

Then we can assign another geometric motion, consisting in rotating the whole system around a fixed axis with angular velocity $\vec{\omega}$. In this case we have:

$$\sum_i m_i \overrightarrow{U_i} \cdot \vec{\omega} \times \vec{r_i} \qquad \left(\vec{u_i} = \vec{\omega} \times \vec{r_i} \right)$$

By the property of the mixed product, we have:

$$\sum_i m_i \vec{\omega} \cdot \vec{r_i} \times \overrightarrow{U_i} = 0$$

By the arbitrariness of ω we have:

$$\sum_i m_i \vec{r_i} \times \vec{U_i} = 0$$

Finally, since $\overrightarrow{U_i} = \overrightarrow{W_i} - \overrightarrow{V_i}$, we have:

$$\sum_i m_i \vec{r_i} \times \overrightarrow{W_i} = \sum_i m_i \vec{r_i} \times \overrightarrow{V_i} \, .$$

This last equation expresses the conservation of the moment of momentum.

> We could still attribute to u_i other values, but this would be pointless and lead to some equations already contained in the previous ones .....

These calculations by Lazare Carnot show the extent to which the line of development and the methods of his Mechanics are different from those of the Mechanics of Newton.

Lazare Carnot was proud to have introduced the new concept of geometric motion: its general purpose was to build "a new Science, intermediate between Geometry and Mechanics". Let us ask ourselves if the project declared by Carnot was realistic.

We have seen that by introducing particular geometric motions, he derives the fundamental equations of motion from equation (2.6). In particular, by assigning a uniform translational motion to the system, he derives the conservation of momentum. In doing so, he explicitly observes that the equation expressing this conservation is independent of the particular uniform translational motion assigned, i.e. the result is the same for all motions of that type. Similarly, the result of the introduction of a rotational geometric motion in equation (2.1) does not depend on the particular geometric motion used:

We note that Hermann Weyl gave this definition of symmetry: "A thing is symmetrical if it can be subjected to a certain operation and it remains exactly the same as before". In the light of this definition it can be said that under the transformation given by a geometric motion (as long as it is of the same type; for example, translational), the result of the second fundamental equation being the same manifests a symmetry. In other words, geometric motions lead to symmetries because we can easily verify that they form a group of transfomations: they are associative, commutative, and always possess an inverse (in that they are by definition reversible). We can therefore say that the subgroup of uniform translational motions leads, with the Carnot technique, to the conservation of momentum; while the subgroup of uniform rotatory motions leads to the conservation of moment of momentum. In addition, we know then that conservation of energy is equivalent to invariance with respect to temporal translations; we will see that he derives it in another way, but still from his fundamental equations. We conclude that Carnot essentially founded his Mechanics on the invariants of motion provided by the application of the geometric group of symmetries. Ultimately, Lazare Carnot was the first to introduce the theory of symmetries in theoretical physics, using simple mathematics without differential operations. No surprise if this discovery was for a long time ignored; in Mathematics the symmetries were introduced with a stroke of genius by Galois 50 years later, and that its result remained unknown for the following twenty years.

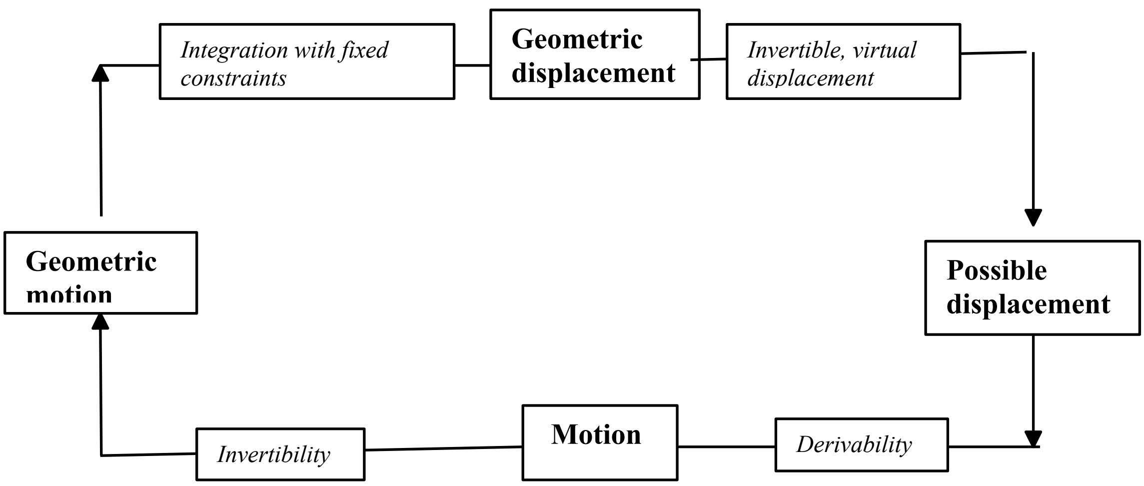
2.17 *Carnot's fundamental equation and the principle of virtual work*

Let us now compare the geometric motions, defined as invertible motions, with the virtual movements of today's mechanics. We will verify the validity of the opinion of Dugas and Gillispie, who consider the former an anticipation of the latter. First of all let us remember that a displacement is possible if it is compatible with the constraints considered fixed, virtual if it is compatible with the constraints even if in motion. Now, in order to compare the two different concepts, we need to translate geometric motions into displacements with $udt \rightarrow \delta s$, or the virtual displacements into motions with $\delta s/dt \rightarrow u$, (once again we find a limit operation that separates the two formulations of Mechanics). We will only consider the case of time-independent constraints.

Let us then begin by translating a possible displacements into a motion: it is sufficient to derive it, which is essentially always possible. The result is a motion that is possible but not always reversible (for example the motion of a sliding ring on a rotating bar). Therefore, to obtain a geometric motion, we must add the hypothesis of invertibility.

TABLE 2.2

THE TRANSITION FROM GEOMETRIC MOTIONS TO POSSIBLE DISPLACEMENTS AND VICE VERSA



Conversely, we pass from a geometric motion, which is invertible, to a geometric displacement. We need to integrate the geometric motion and thus obtain a geometric displacement, which by definition of the first will result in a virtual and invertible displacement. This last operation is essentially always feasible, if the hypothesis of fixed constraints is valid; otherwise the equations of the motion of the constraints will be considered, which can in general be complicated and difficult to integrate. At this point we can conclude *that for time-independent constraints a geometric displacement is equivalent to an invertible virtual displacement, but inversely a possible displacement only if it is invertible gives, by derivation, a geometric motion*.

With this partial equivalence, we wish to demonstrate, in the manner of the scientists of the 1700s (for whom , since all the functions were considered continuous curves), the derivation of the Carnot equation (2.1) from the principle of virtual work for a system of particles. Let us recall the textbook definition of this principle: *A system of particles is in equilibrium if, and only if, the total virtual work of active forces is zero; that is, if, and only if*

$$\sum_i \vec{F}_i^{(a)} \cdot \delta \vec{s}_i = 0 \qquad (2.7)$$

But let us recall that for invertible motions the work of the constraints is null (otherwise there would be free work):

$$\sum_i \vec{F}_i^{(v)} \cdot \delta \vec{s}_i = 0 \qquad (2.8)$$

Which expresses mathematically the impossibility of perpetual motion, hence:

$$\sum_i \vec{F}_i^{(a)} \cdot \delta \vec{s}_i + \sum_i \vec{F}_i^{(v)} \cdot \delta \vec{s}_i = \sum_i \vec{F}_i \cdot \delta \vec{s}_i = \sum_i m_i \frac{d\vec{V}_i}{dt} \cdot d\vec{s}_i = 0$$

We note that, from the point of view of constructive mathematics, the principle of virtual work has a formula that poses no problems, because it is a simple mathematical formula and its equality is to be understood as approximate equality, as it is for all experimental laws; there might be problems with the constraint equations, to the extent that a solution of them is sought without approximations and they may have double solutions, which however seems extraneous to the usual applications.

Let us go back to (2.8). Confusing, as was done in the 1700s, with (which is correct if the constraints are fixed and the forces are continuous), we have :

$$\sum_i m_i \Delta \vec{v}_i \cdot \frac{\Delta \vec{s}_i}{\Delta t} = \sum_i m_i \vec{U}_i \vec{u}_i \qquad (2.9)$$

where the last step applies if the virtual displacement, having become the motion $\Delta s / \Delta t$, is invertible.

Now let us proceed with the second fundamental equation of Lazare Carnot. If we consider the following relations:

$$\left\{ \begin{aligned} m_i \vec{U}_i &= m_i \Delta \vec{v}_i \\ \vec{U}_i &= \frac{\Delta \vec{s}_i}{\Delta t} \end{aligned} \right\}$$

(where $\Delta \vec{s}_i$ is the geometric displacement if the *i*-th point and Δt has to be specified),
we obtain from (2.9)

$$\sum_i m_i \Delta\vec{v}_i \times \frac{\Delta\vec{s}_i}{\Delta t} = 0 \qquad \text{i.e.} \quad \sum_i m_i \Delta\vec{v}_i \times \frac{\Delta\vec{s}_i}{\Delta t} = 0$$

Switching to infinitesimals we have:

$$\sum_i \vec{F}_i \cdot d\vec{s}_i = 0 \qquad (2.10)$$

This equation and (2.7) differ from each other because:

1) in (2.10) there is a geometric motion which gives a $d\vec{s}$ and not a virtual displacement $\delta\vec{s}_i$

2) $\vec{F}$ is the resultant force acting on the *i*-th particle, not its active component $\vec{F}^{(a)}$.

However, note that if a geometric displacement equals an invertible virtual displacement, the virtual work of the constraining force components is null. Furthermore, when constraints are present, the resulting force acting on the *i*-th particle is given by , where the second represents the constraining force acting on the *i*-th particle. Moreover, in applying the principle of virtual work, it is assumed that if every displacement is invertible (precisely the condition that derives from being geometric motion) then the virtual work of the constraining forces is null

$$\sum_i \vec{F}_i^{(v)} \cdot \delta\vec{s}_i = 0$$

Thus the equation (2.10), removing the latter from it, assumes the form of the principle of virtual works (2.7). The agreement is surprising, for we started from concepts that are quite different from the usual ones.

Carnot himself, in the preface to the *Principes*, explains this novelty:

> My theory could not be founded precisely on the principle of virtual velocities ..., which is not applicable, without modification, to the impact of bodies. I therefore start from a principle which is different but very analogous, or rather, which is this same principle of virtual velocities, but suitably extended. (L. Carnot 1803, p. x)

2.18 *The historical importance of the Lazare Carnot formulation*

Historically it is important to observe that Carnot's generalization of the principle of virtual work precedes the different one of Lagrange's *Mécanique Analytique*[9] by 6 years; however, it has not been remembered to the same extent as the latter, which is based on the infinitesimal analysis. Yet L. Carnot's Mechanics did not go unnoticed among the physicists of his time: this is demonstrated by citations in texts, some prestigious (those by Lagrange and Fourier, for example) and by the fact that it gave rise to the tradition of technical Physics in engineering.

The usual history of physics eliminates Lazare Carnot's contribution, which is charged with being too engineering-oriented, although his mechanics is aimed not so much at studying machines, but at maintaining the experimental character of the theory and the theoretical principles as well. Although it has a fundamentally experimental character, the theoretical

---

[9] J.-L. Lagrange, *Mécanique Analytique,* Paris: Desaint, 1788.

level reached is universal, without erring into metaphysics. In other words, the mechanical work of Lazare Carnot constituted a fruitful union of technology and theory, without subordinating the first to the second, but rather with the first demonstrating an innovative theoretical capacity in the second, at that time too tied to metaphysics. (His son, Sadi, did something similar shortly after with thermal machines and Thermodynamic theory).

In fact, in the history of mechanics, Lazare Carnot ranks as the one who brought together two mechanical traditions, artisanal-engineering (Bernoulli, Borda) and classical. Because of this convergence his mechanics must be seen both as a superior, and ultimately adequate, conception of artisanal mechanical techniques, and as a deepening of the theoretical heritage of Physics.

Because of this the mechanics of Carnot realizes an alternative to the conception of Physics of Descartes and above all of Newton, who idealized the principles of Physics according to a conception that has long dominated the history of physical theory. It was only in the 1900s, through two profound crises (the ether, quanta), that the foundations of Physics returned to Carnot's conception of both the theoretical principles and the mathematical (algebraic) techniques.

Furthermore, his experimental and anti-metaphysical outlook leads Lazare Carnot to a mechanical-mathematical relationship based on a controllable and secure mathematics, different from that (infinitesimal Analysis) which Newton established in *Principia* and which informed all the subsequent physical theory up to the XX century. For this reason we associate with Carnot an unexpected result in the relationship between physics and mathematics: his theory is compatible with the current constructive mathematics. In other words, Lazare Carnot's restriction of the potential of mathematics used in theoretical physics, allowed him to anticipate the constructive relationship between physics and mathematics. The Mechanics of Lazare Carnot presents a clear alternative choice to that of Newton: potential infinity (PI) rather than actual infinity (AI).

The table below clarifies the contrast between the two mechanical theories and highlights the divergencies existing between the various concepts of Newtonian Mechanics and those of Carnotian Mechanics.

TABLE 2.3

DIFFERENCES BETWEEN NEWTON AND L. CARNOT'S FORMULATIONS

SUBJECT	NEWTON	LAZARE CARNOT
Cultural value of the theory	Theory also philosophical	Completely experimental theory
Organization of the theory	By principles	Based on a fundamental problem
Space	Infinite and absolute	Delimited and relational
Time	Absolute	Finite variation in time
Bodies	As a set of material points without extension	Extended bodies, machines
Movement	Property of the body	Comunication of motion
Inertia	As perpetual motion	Impossibility of creating perpetual motion
Basic concept	Force-cause	Quantity of motion exchanged

Espression of interaction	Force, as a synthesis of all the influence of the external environment	Work
Mathematical conception	The same for the 'infinitely large for the infinitely small	Geometric motions relative to the geometric configuration of bodies
Central problem	$\vec{F} = m\vec{a}$	Laws of conservation on impact
Mathematical technique	Infinitesimals Differential equations	Geom. Motions. Vectorial calculus $\sum m\vec{U} \cdot \vec{u} = 0$
Mathematical problem	$Solutions\ \ of\ F = m\dfrac{d^2x}{dt^2}$	Invariances of $\sum m\vec{U} \cdot \vec{u} = 0$
Solutions	Trajectories from minus to plus infinity	Quantities conserved
Machine	Particular application of the theory	Universal subject of the theory
Capacity of machines	Possibility of an infinite power	Against the chimera infinite power

It should be noted that almost all the constituent elements of the theory have radical variations of meaning in the transition from one formulation to another. In particular, the organization of the theory in Newton is decidedly based on principles-axioms, from which to deduce all the laws of the theory. Instead, that of Carnot (excluding the edition of the *Principes* of 1803) is based on a *problem*, that of the collision of the bodies (as it is in Leibniz and D'Alembert), with the solution of which the whole of mechanics is constructed. Its purpose is to find a new scientific method that is able to solve this problem; that is, capable of determining quantities that are invariant despite the impact. He himself says: "there is therefore ... in every percussion or communication of movement a quantity which *is not altered* (remains the same) b*y* the impact." (*Essai* p. 47) Note that the sentence is doubly negated, as is any methodological principle. His theory is therefore not organization by principle-axioms as in Newtonian mechanics (OA), but is based on a universal problem.

Note that if an additional column were added for the corresponding concepts of Sadi Carnot's Thermodynamics, they would almost always be similar to the concepts of Lazare Carnot's mechanics. This shows that the birth of Thermodynamics was possible only because Lazare Carnot had formulated mechanics differently from Newton, who, on the other hand, had nothing to say to theorists constructing this new theory. It also shows that the difference between formulations of the same theory relating to the same field of phenomena can be more radical than the difference between formulations of different theories concerning different fields of phenomena. Ultimately, the premises of our theories are very strong, sometimes more so than is needed to describe physical phenomena adequately.

One might add that all of this underlines the considerable importance of concepts in Physics. They arise directly from the experiments but are formed by a large degree of intellectualization.

We have seen the great historical importance of Carnot's formulation of Mechanics. It must be admitted, however, that Carnot's formulation is not directly applicable in our day, because it is based on the idea of full space, on tensile and pressure forces alone without taking into account forces acting at a distance (both in the definitions, including that of geometric motion, and in mathematical development) and on the concept of continuous force as the limit of a series of impulses; this last idea is intuitively valid, but it would be better to formalize it, taking into account the different ways of founding Mathematics, a task that has not been carried out so far. It undoubtedly has great historical and cultural value, but today it should be adequately reformulated considering interactions at a distance, continuous forces, etc as well. Let us not forget, however, that three centuries of clarifications and improvements of Newton's mechanics have passed, while the further development of the Lazare Carnot's mechanics, which was forgotten, is just beginning.

However, all the above-mentioned defects do not detract from the importance of this formulation of mechanics for the history of physics and also for current theoretical physics.

2.1 The history of false demonstrations of the principle of virtual work

The preceding study re-evaluates the principle of virtual work, which is usually given a secondary role in texts of rational mechanics, as if it were a simple appendix of mechanical theory. Let us study the history of this principle.

Some of its simplified formulations date back to antiquity, to epochs before Aristotle. It was applied widely, first in the case of the lever and then for a system of heavy bodies, according to the version of the Torricelli principle (the centre of gravity of that system cannot rise). It was then mathematically formulated by Johann Bernoulli in 1717. At the end of the same century Lazare Carnot and Lagrange were the first to propose this principle as a new way of founding Mechanics. It was the first time in an experimental science, well constituted and systemized (by Newton and Euler), that the problem of finding new foundations in logico -mathematical terms without appealing to metaphysics (as Maupertuis did when he suggested the new principle of least action), was addressed. Now, for the mentality of the physicists of the time, who by then had internalized the AO as the only possible organization, replacing Newton's principles meant reworking the theory at the metaphysical level of its axioms, far higher than the single laws. Since the principle of virtual work originated, however, in the ongoing practice of artisans, it seemed of too humble origins to be comparable with the previous principles. It was therefore thought that it was necessary "to prove it", that is to derive it from some other elevated principles (similar to that of Maupertuis), or to derive it from experimental facts that were so simple as to appear self-evident.

All the great scientists of the time were engaged in this enterprise: Lagrange, Laplace, Poisson, Poinsot, Navier and Carnot himself. Leaving aside the metaphysical demonstrations, the other proofs tried to reduce the new principle to its simplest case, e.g. the sum of forces (law of the parallelogram) or the very ancient one of the lever. It was essentially a question of reducing the n variables of the principle to just two or even one. It was believed that this goal could be easily achieved by exploiting the technical innovation of the time. A few years earlier (1799) Mascheroni had shown that geometric constructions performed by means of ruler and compass, a pair that in the past had had great significance, even religious and philosophical, may be performed by means of the compass alone. Since then, theoretical importance also began to be given to geometric instruments which until then had not been used or had been used only with diffidence, that is, both linkages (rigid hinged rods, e.g. the pantograph) and tracer wheels, with which more results were obtained than those obtainable with ruler and compass.

Lagrange suggested a demonstration in which the forces of the system were connected to each other with a complex system of pulleys in order to reduce it to a single force. Fourier provided four "proofs" of the principle, the broadest of which uses a complicated three-dimensional linkage with the same aim as Lagrange: to connect all the forces together so that they're all derivable from one.

Each demonstration depended then on the schematization of the machines which was common then: they were treated theoretically as if they had no mass (as is done today for the mathematical pendulum wire; it was Lazare Carnot who established that machines should instead be treated as having mass, but he was not immediately listened to). It was therefore believed that by unifying all the forces into one (the first for example), it would be possible to determine with the same previous displacement all displacements caused by other forces. If instead we take into account the masses to be moved, then the total force which is distributed with the same initial forces on the points forming the system cannot give the same δs_1, which are now no longer known and should be calculated precisely using the formula of that principle of virtual work which must be demonstrated. In conclusion these demonstrations all are vicious circles.

There were decades of proofs that were not entirely conclusive even to the minds of contemporaries. Until, as Poinsot testifies, the problem was abandoned due to exhaustion. Only a few had the idea that this principle didn't need to be proved, but rather generalized, as Lagrange had done with his equations, or even (as we saw earlier) L. Carnot with his second fundamental equation.

The consequence of the inglorious end of the debate on the principle of virtual works was that Mechanics has never studied its foundations in depth. Rather it tried to resolve them by mythologizing them, to such an extent that an extra university exam is imposed on students, called "rational mechanics", which simply repeats in abstract form what is a physical theory that is already studied in General Physics. The myth that underlies the books on this subject is that Mechanics, by completing tying itself to infinitesimal analysis, clarifies its foundations as no other physical theory has ever done; in other words, it is believed that the fundamentals of Physics are all the clearer, the more they are abstracted from reality and mathematized by analysis (philosophically, this attitude is Platonic).

The subsequent historical development of theoretical physics (through the births of electromagnetism, thermodynamics, and then relativity and quantum mechanics) has completely discredited this idea. Of course it is in fact useful to develop Mechanics with Analysis (which can underline the mathematical isomorphism of the same differential equation with completely different physical situations; for example, it can give rise to the Hamilton-Jacobi equations, which are unique in providing a link between Newtonian mechanics and quantum mechanics; etc.); but this perspective did not give any suggestions for solving the conceptual problems that were historically crucial to arriving at the new theories of the nineteenth and twentieth centuries, and to understanding the Foundations of theoretical Physics.

2.20 The two lines of development of mechanics

This false foundational perspective was not only detrimental in the subsequent crisis of Physics, but even led to the widespread belief that the foundations of Physics could not be clarified any more than they were in Mechanics. This left uninvestigated many deep misconceptions in the fundamental principles of Mechanics (such as the ancient misconceptions of absolute space and time; or the modern one of the vicious circle between the three definitions of the principle of inertia, the system of inertia and the clock) and it left undecided whether Newton's principles are the most general possible (while admitting that they are technically less effective than the other principles), or whether there are alternative principles, as L. Carnot and Lagrange had begun to maintain. This false perspective led physicists to think that Newtonian Mechanics could include all the other formulations, even those of L. Carnot and Lagrange, which had threatened its monopoly of theory. It was therefore a common opinion of many theorists that the formulations of Mechanics are all equivalent.

This last point can also be clearly seen in current textbooks of rational Mechanics: which, even today, do not agree on the problem of whether or not it is necessary to prove the principle of virtual work (where "proving" can only mean deriving it from the Newtonian principles). In fact, some texts claim to give the "proof", which is what Appell, for example, does. At the crucial stage of the proof, however, he is forced to recognize a difficulty: he has to hypothesize the statement (2.8) that the work of the constraints cannot be positive. He admits that this statement can be proved only in some particular cases, but nevertheless he wants to make a general use of it, even if he does not have a rational argument to offer. (other books attempt incorrect proofs; for example, deducing it from the Lagrange equations, which, as we know, generalized it).

In reality, we know that constraint reactions are our inventions, introduced to justify the statics of bodies in the presence of constraints: it is we who attribute imaginary (or "phantom") forces to rigid bodies, i.e. constraint reactions, exactly equal and opposite to the applied forces. No instrument has directly measured a constrain reaction. But, after admitting these inventions of ours, we must exclude the possibility that they can do positive work (it would be like believing that it was a thought in the head of Jupiter that generated Minerva): if they could, it would be enough to extract it from constraints in an unlimited way, which would allow perpetual motion, which is absurd. But then we see that the statement from which the principle of virtual works derives is nothing but a version of the impossibility of perpetual motion; the latter statement was suggested and accepted as a methodological principle owing to fruitless attempts over millenia to disprove it, and, as such, cannot be demonstrated deductively from other principles, or be posited as an axiomatic principle from which to derive other truths by simple logical analysis.

There are also wise books on rational mechanics (eg Levi Civita and Amaldi; Sommerfeld) which declare that they refrain from giving a proof of it, rightly recognizing that the crucial statement on the work of constraints is equivalent to the impossibility of perpetual motion; but by doing so these authors seem to say that this fact concerns a grey area of theoretical physics, which is the boundary between genuine theory and empirical practice that has not yet been theorized (or mathematized by analysis). Some other texts suggest that the principle of virtual work is actually a method to solve problems, but they do not develop the above generic statement.

In conclusion, none of these authors make it clear whether this principle is independent of Newtonian principles or not. Yet simple reasoning easily recognizes it easily: since the Newtonian principles concern an isolated particle, or little more, then they cannot deal with systems of extended bodies, which include constraints (which is what Lagrange was the first to say);the formula (2.8) is an essential novelty. It is therefore easy to come to the conclusion that, in general, the principle of virtual work cannot be derived from Newtonian principles and that Euler had already made the maximum effort to extend the theoretical scheme of Newton's equations to discrete or continuous particle systems (in the latter case, valid for fluids, the continuum is unlimited and does not admit impassable boundaries which is what constraints are).

This great foundational problem, which persisted despite two hundred years of fruitless reflections on the foundations of Mechanics, becomes even clearer if it is recognized that there exist two ways of organizing a scientific theory; and that a principle can be either methodological or axiomatic. Indeed, if mechanical theory is conceived as a problem-based theory (PO) (whose fundamental problem is: how do the constraints react?), then the principle of virtual work is a methodological principle, as is the principle of the impossibility of perpetual motion. Consequently the usual treatment of textbooks of rational mechanics has confused what constitutes a problem with the corresponding positive statement (as does the famous phrase of the marquis de la Palisse: If you do not die, then you will live on), obtaining a version of the principle of virtual work which is a non-physical abstraction.

It is important to note that this great scientific debate on the foundations of mechanics occurred almost simultaneously and in a similar way to the other great debate on the foundations of the first scientific theory, Geometry. The debate on the elimination of the only "blemish" of Euclidean geometry had been going on since antiquity, namely that it was not possible to prove the fifth postulate by deriving it from the other four. The motivation seemed obvious since Euclid had introduced that postulate only at the 29th proposition (that is, as late as possible) and since its inverse proposition (obtained by exchanging the hypothesis for the thesis) was provable, which is the case for all the other theorems of Euclidean geometry. We know that this debate was resolved when finally (in the first half of the nineteenth century) some mathematicians (Gauss, Lobacevskij, Bolyai) decided to consider it a postulate that is independent of the others. As a result, they developed a new geometry. In this way began the revolution of the foundations of geometry and then of all mathematics.

Just as finally accepting the independence of the parallel postulate led on to a new History of Mathematics, in which the problem of the Foundations constituted the first problem, and to different ways of doing both Geometry and Mathematics in general, so, analogously, accepting the independence of the principle of virtual works from Newtonian principles, it would have led on to recognizing the radical novelty of L. Carnot’s and Lagrange’s Mechanics; and research in theoretical physics would have been attentive to the problem of the foundations (first of mechanics, and then of all physics) and a new history of physics would have begun, that of the plurality of lines of development. We have started to see that in the previous paragraphs, albeit after a delay of two centuries.

2.21 The Foundations of Physics

To recognize the Foundations of Physics, it is necessary to highlight an unconscious choice by the vast majority of theoretical physicists. Following Newton, they adopted infinitesimal analysis, whereas Leibniz wanted to exclude it from the foundations of a physical theory, as did L. Carnot and Sadi Carnot. In fact he latter were ignored. However, at the beginning of the twentieth century, the discovery of quanta forced physicists to take an interest in the discrete even in the most advanced theory. It is therefore clear that in the Foundations of Physics there is a choice regarding the type of Mathematics.

Another unconscious choice made by theoretical physicists following Newton (and Euclid) was to evaluate a theory as appropriate only if it was organized deductively, that is, founded on a few self-evident principles. from which the entire theory was deduced. Leibniz. L. Carnot and Sadi Carnot, on the other hand, had founded their theories on a problem, without principles. But this innovation was also ignored, although D'Alembert and L. Carnot had clearly indicated that there are two types of theoretical organization. Then, at the beginning of the 1900s, first Poincaré and Lorentz, then Einstein, reflecting on the theories built up by physics, distinguished two types, those deduced from a priori principles and those, like Thermodynamics, based on the impossibility of perpetual motion.[10]

Everything that has been said so far leads us to the conclusion that the Foundations of Physics *include two fundamental dichotomies, one regarding the type of infinity (PI or AI) each characterizing a particular type of Mathematics, and one regarding the type of organization of a theory, deductive or based on a problem. In every theory of the second type there is at least one crucial doubly negated proposition (i.e. a principle) which is not equivalent to the corresponding affirmative, implying non-classical logic. Thus the type of theoretical organization is characterized by the type of logic, which therefore distinguishes these theories with formal precision from AO theories. We can therefore treat both dichotomies formally, that is, the first in mathematical and the second in logical terms.*

TABLE 2.4

THE FOUR CHOICES EXPRESSED BOTH IN INTUITIVE AND FORMAL TERMS

	INFINITY		**ORGANIZATION**	
	AI	**PI**	**AO**	**PO**
PHILOSOPHICAL CONCEPT	*Actual infinity*	*Potential infinity*	*Aristotelian Organization*	*Problem-based organization*
FORMAL MATHEMATICAL SYSTEM	*Classical mathematics*	*Constructive mathematics*	*Classical logic*	*Non classical logic*

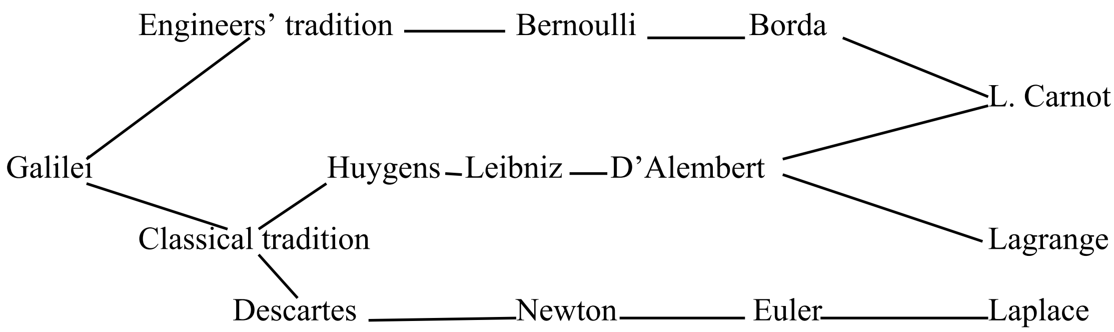
10 Flores, F. (1999). Einstein's Theory of Theories and Types of Theoretical Explanation. *International Studies in Philosophy of Science, 13,* 123-134. http://dx.doi.org/10.1080/02698599908573613

With these two dichotomies, a cultural division between the various foundational lines of Mechanics has been evident since the time of Galilei (who saw them but accepted neither AI nor AO) and by the end of the 1700s was manifested with theories based on the principle of virtual work. This division was, however, hidden and / or ignored.

We can therefore represent the history of Mechanics with the following table, which is far more comprehensive than the traditional one of a unilinear type.

TABLE 2.5

THE LINES OF DEVELOPMENT OF MECHANICAL THEORY



In the table above we have the lower line which from Newton onwards followed the two choices of infinitesimal analysis (IA) and of the entirely deductive organization of a theory (AO), while the upper lines up to the mechanics of L. Carnot represent the two choices of a simple mathematics (IP) and of a problem-based organization (PO). Lagrange, on the other hand, represents an intermediate position: he uses a new mathematical technique with AI, but it is aimed at solving any problem of Mechanics (PO).

CHAPTER 3

NEW HISTORY OF THE FOUNDATIONS OF STATISTICAL MECHANICS AND ELECTROMAGNETISM. APPLICATION OF THE TWO DICHOTOMIES TO THE OTHER CLASSICAL THEORIES

We do not have the space here to study the theory of statistical mechanics in detail. We will do it synthetically using two summary tables, which show that Newton's Mechanics, credited with having represented the current paradigm of theoretical physics at that time, actually did not dominate classical Physics, but rather it was almost dismissed by the new theories of the 19^{th} century.

3.1 History of the Foundations of Statistical Mechanics

On the contrary, Newtonian Mechanics proved to be inadequate very early on in the first attempts to formulate a kinetic theory of gases, which is usually considered a natural extension of Mechanics and an example of beginning to go beyond the usual formulation of thermodynamics, considered to be merely "phenomenological".

Since its inception, Newtonian Mechanics extended to include geometric optics, conceiving a light ray as a trajectory of material corpuscles, and Acoustics as formalized by the wave equation that generally describes the oscillations of matter.

It is little known that Newton had also begun to formulate the kinetic theory of gases according to a conceptual model that was consistent with his Mechanics. The central element was first of all the force-cause, that gravitational force which he considered universal, which would also have to explain the interactions between atoms, conceived as perfectly hard and imagined to be fixed. Despite the obvious limitations of this kinetic theory of gases, he was able to arrive at Boyle's and Mariotte's law PV = const. At that time this result seemed to be the proof of the correctness of his approach. Thus the tradition of Newtonian Mechanics suggested models and techniques for the study of every field of phenomena, all of which, together with Newton's authority, assured it a monopoly of theoretical physics until the end of the 1800s.

However, as early as the beginning of the 1700s, a follower of Leibniz, J. Bernoulli, envisaged founding a mechanical theory based on energy and on the principle of virtual work. He was unsuccessful, he did, however, lay the foundations of a kinetic theory of gases by imagining atoms in modern terms, that is, as moving mass-points that collide with each other elastically. He also succeeded in deriving Boyle's law from Mariotte's PV = cost, but his approach was ignored and forgotten.

Newton's theoretical schema could not be developed much further in the new field of phenomena. Some progress was achieved a century later, when Laplace and Poisson changed the gravitational forces into forces that were still central, but with a dependence on the variable radius relating to particular cases. Thus in some fields of study (for example capillarity) they did formulate several new and modern laws, but not a complete theory of gases. In the first half of the 1800s some amateur English scientists re-proposed concepts and formulas similar to Bernoulli's, without taking into account his previous theory. Their theoretical perspective was finally re-evaluated towards 1850, when the inception of Thermodynamics induced physicists to accept the conservation of energy and therefore to abandon ~~the~~ Newton's schematization of atoms as perfectly hard, replacing it with that of elastic and mobile bodies.

Finally in 1858, Maxwell founded the kinetic theory of gases on the basis of the laws of elastic impact already set forth by Leibniz a century and a half before. However, he did not cite him and demonstrated the said laws with geometric theorems, also using formulae for elastic bodies but, in homage to Newton, he sometimes called them hard bodies. The divorce from Newton's scientific legacy, especially in England, was long and tormented.

Reconsidering a previous table, which presented the radical variations of meaning in the basic concepts of Newton's Mechanics and that of L. Carnot, and applying it to the concepts of the completed theory of kinetic theory of gases, we note that the elements of the latter are completely different from those of Newton.

TABLE 3.1

THE FUNDAMENTAL CONCEPTS OF THE KINETIC THEORY OF GASES COMPARED WITH THOSE OF NEWTON

FUNDAMENTAL CONCEPTS OF THE NEWTONIAN THEORY	KINETIC THEORY OF GASES
Acceleration	Velocity
Force-cause	Random interaction through collisions
Example: Universal gravitation	Example: state equation for perfect gases
F=ma	Conservation laws in collisions.
Differential equations	Temporal averages
The motion of an isolated point-mass	The motion of many point-masses in a limited space
Hard atom	Elastic atom
Perfectly rigid, solid bodies	Perfect gases
The motion of a system in an infinite space	The motion of a system in a finite and limited space
Theory to be explained: Astronomy	Theory to be explained: Laws of Gases

In table 3.1 we see that the concepts of L. Carnot's mechanical theory are exactly the same as those that kinetic theorists had to follow in order to construct their theory of gases.

Therefore not only was the kinetic theory of gases not a development of Newtonian mechanics, but it even represented its historical defeat with respect to its Carnottian alternative.[1] This profound difference between the kinetic theory of gases and Newtonian mechanics was, however, perceived merely as a slight dissonance, justifiable by the fact that the new theory was highly speculative, since it concerned particles of matter that were merely hypothesized without experimental proof.

[1] A. Drago e P. Saiello, "La teoria cinetica dei gas: una sconfitta del meccanismo newtoniano", in A. Rossi (ed.): *Atti XIII Congr. Naz. St. Fisica,* Lecce: Conte, 1995, 295-308. A. Drago, "A rational reconstruction of the history of the kinetic theory of gas as founded on Leibniz-Carnot's formulation of mechanics", *Atti Fond. Ronchi*, 69 (2004) 365-387.

Boltzmann's attempt to explain thermodynamics with statistical mechanics came later, at the end of the century. He reinforced the basis of his theory by replacing Newtonian with Hamiltonian mechanics and in addition by introducing probability theory. In reality, these foundational operations constituted a complete departure from Newtonian mechanics, which is, however, still commonly recognized as the basis of his theory.

Furthermore, Table 3.2 alone (where Bernoulli and Laplace have been added for completeness) shows how in the subsequent Statistical Mechanics the fundamental concepts do not refer to those of Newtonian Mechanics: the first concept that is left out is precisely Newton’s basic concept, the force(-cause).

TABLE 3.2

THE FUNDAMENTAL CONCEPTS IN THE MAJOR THEORISTS OF STATISTICAL MECHANICS

CHARACTERISTIC	NEWTON	*Bernoulli*	*Laplace*	*Maxwell*	*Boltzmann*	*L.CARNOT*
Space	Absolute	***C***	***N***	***C***	***C***	Delimited and relational
Time	Absolute	***C***	***N***	***C***	***C***	Finite variation
Inertia	As perpetual motion	***N***	***N***	***N***	***N***	Impossibility of perpetual motion
Basic concept	Acceleration	***C***	***N***	***C***	***C***	Velocity
Expression of the interaction	Force-cause (metaphysical)	***C***	***N***	***C***	***C***	Work and energy
General mathematical problem	*F=ma*	***C***	***N***	***C***	***C***	The problem of finding the invariants
Mathematical techniques	Differential equations	***C***	***N***	***N***	***N***	Geometric motions
Solutions	All possible trajectories - ¥ £ *t* £ ¥	***C***	***N***	***C***	***C***	One specific motion, studied before and after impact

As a further proof of the departure of the new theory from Newtonian mechanics, consider the so-called virial theorem (with which the various forms of the laws of real gases were produced):

$$\Sigma(Xdx+Ydy+Zdz)=1/2\ \Sigma mv^2,$$

where X, Y, Z are the components of the force acting on the molecules, x, y, z the displacements, v the velocity (indices of summation over molecules and an overbar above each side of the equation to indicate the mean are implied). Clausius, who introduced this theorem, derived it from Newton’s equation applied to each particle of the gas; but he had to introduce the mean to take into account that it is impossible to track the displacements or accelerations of around 10^{24} molecules. Apart from having to consider the mean of each variable, it is a simple energy balance between work and kinetic energy; hence, it is essentially a theorem of conservation of energy, rather than Newton’s second principle. The essentially new character of this "virial theorem" also appears in its applications; it obtains global results, not attributable to trajectories or positions.

It was L. Carnot who had begun to propose the use of energy balances a century earlier in the mathematical treatment of input and output of a machine. According to this method of energy balance, it is sufficient to calculate the work done by forces on the totality of the molecules and the resulting kinetic effect of those forces on the same molecules, and then calculate the mean of each magnitude.

It can therefore be said that normally when one speaks of the history of Physics, "Mechanics" is identified with the Newtonian formulation, as if it were the only formulation. If, on the other hand, L. Carnot's formulation was kept in mind, we would have a very different view of the history of mechanics and of subsequent theories: we saw this in the cases of the kinetic theory of gases and statistical mechanics.

3.2 Brief summary of the history of electrical and magnetic theories: 1785-1860

The emergence of Electromagnetism was certainly problematic, since for the first time a theory had to be constructed starting from phenomena that are not perceivable by our senses (except for the most powerful electrical discharges). It was therefore a question of highlighting "phantom" phenomena, invisible to our intuition, which is based on our physiology. Every new phenomenon therefore had to be mediated by innovative measuring instruments so that the whole history of Electromagnetism has depended on the numerous inventions of new measuring devices. Furthermore, its phenomena, appropriately mediated by artificial instruments, appeared very strange because their conceptualizations was outside the traditional concepts of Mechanics known at the time; so much so that new mental categories (e.g. field of forces) were required to understand them. Furthermore, the philosophy of the time (late 1700s and later) experienced the well-known "post-Kant crisis", whereby the scientists of that time did not have the help of philosophers when dealing with an unknown field of mysterious phenomena.

It would take too long to retrace the entire history of Electromagnetism in detail; a table is presented below that summarizes the sequence of the most important advances (the table was unpublished elsewhere not because the historical facts indicated there are little known, but because it provides a rather inglorious image of the history of a scientific theory: not that of a reliable accumulation of new contributions in an almost pre-established order, but rather a more realistic one of multiple attempts, uncertainties and vicissitudes).[2]

TABLE 3.2

THEORETICAL ADVANCES UP TO MAXWELL'S THEORY

Date	**ELECTRICITY**	*MAGNETISM*
1785	Coulomb's law $F=kQq/r^2$	
1799	Voltaic pile	
1813	Poisson's equation for electrostatic potential $\nabla^2 V = -4\pi\rho$	
1820	Oersted's experiment	
1820	Biot and Savart's law: $dB = i\, d\vec{l} \times \vec{r} / r^3$	

[2] A. Drago: "Volta and the strange history of electromagnetism", in F. Bevilacqua and E.A. Giannetto (eds.): *Volta and the history of Electricity*, Hoepli, Milano, 2003, pp. 97-111.

1821		Faradayis concept of lines of magnetic force
1825	Ampère’s formalism for the interaction between magnetism and electricity in movement	
1826	Ohm’s law: $Ri = V$	
1831	Faraday’s law: *f.e.m.* = - $d\Phi(\vec{B})/dt$	
1832	Faraday’s law for electrolytic conductance $m = MQ/Fz$:	
1834		Lenz’s rule
1837	Faraday’s relative dielectric constant: $\varepsilon = c/c_0$	
1839	Gauss’s theorem: $\Phi_S(\vec{E}) = \Sigma q_i/\varepsilon_0$	
1843	Faraday’s experiment on the conservation of electric charge: $div\ i + d\rho/dt = 0$	
1845	Neumann’s potential function V and potential vector A	Faraday’s concept of “magnetic field”
1846	Weber’s force: $F = [1-(\dot{r}^2 - 2r\ddot{r})/2c^2]ee'/r^2$	
1847	Helmholtz: the energy of electric work: $eV=W$	
1848	Kirchhoff’s laws: $\Sigma i = 0$ (in a node); $\Sigma Ri = \Sigma$ *f.e.m.* (in a network)	
1850		Thomson: $div\ B = 0$; $rot\ H = 0$
1852	Thomson: Energy of the electric field $\propto \int \varepsilon E^2 dv$	$B = \mu H$ Energy of the magnetic field $\propto \int \mu H^2 dv$
1861	Maxwell’s introduction of the word “field ”	
1861-2	The four Maxwell equations for the electromagnetic field: $div\ E = \rho/\varepsilon$ $\quad rot\ E = -\ \partial B/\partial t$ $div\ B = 0$ $\quad rot\ B = \mu J + (1/c^2)\partial E/\partial t$ Wave equation: $\nabla^2 B = (1/c^2)\partial B^2/\partial t^2$; $\nabla^2 E = (1/c^2)\partial E^2/\partial t^2$; Velocity of light: $c = 1/\sqrt{\varepsilon\mu}$	

When Coulomb demonstrated the analogy between the law of attraction of charges and the law of universal gravitation, it seemed that Newtonian Mechanics had also advanced towards the irreversible conquest of this new field of mysterious phenomena: there was only one thing to add: with respect to the Newtonian schema of central forces, repulsive forces should also be considered.

Volta's innovations were revolutionary because they displayed hitherto completely unsuspected phenomena. First the electrophorus, which made it possible to create an electric charge apparently at will; then the battery and the electric circuit, which introduced a dynamics without inertia and without mechanical forces, characterized moreover by cyclicity and not linearity. Newtonian mechanics suggested the schematization of massive particles subject to forces; but //~~this and~~ the concept of Newtonian force itself did not play a decisive role here. However, there emerged what was certainly a novelty, that of repulsive interactions, unknown in Mechanics, which even suggested that electricity produced perpetual motion.

Indeed, with the discovery of the electrophorus Volta already expressed his conviction that he had produced perpetual motion in electricity, because, he argued, an inexhaustible amount of electric charge could be obtained (in reality, it was created by the work done by the operator who worked to bring the two plates of the electrophorus closer together). He again declared this belief when he invented the pile, which, working thanks to the homonymous effect (explained only at the beginning of 1900 as originating from the contact between two different metals), gave an induced electromotive force which seemed to him inexhaustible (even if he saw it diminish *(infinitive)* over time). The construction of the theory had to be established in the midst of these claims, departing from the Newtonian theoretical model that until then had constituted the certainty of theoretical physicists. Not surprisingly, when Ohm invented his very simple, but absolutely new law, his article was rejected by the major journals and published by a journal of pharmacology (1821).

Let us now examine the concepts on which the new theory was gradually built. As mentioned, it had to abandon Newtonian mechanism; this was in particular the case after Oersted's experiment, in which an interaction between the magnet and electric wires was demonstrated. This experiment was decisive in the abandonment of the dominant theoretical and mathematical model of central forces acting on single particles, which had recently been significantly developed by Laplace. Thus was established the concept that Faraday (who, being basically a chemist, did not want to use mathematics on principle) proposed as basic: the force field.

However the Newtonian mechanism regained importance through an unfounded appeal by Ampère to infinitesimal formalism: the introduction of the electric line element *dl* clearly constituted an idealization with respect to the real situation of a closed line of an electric circuit through which an electric current is travelling. However, this formalism allowed many new phenomena to be represented mathematically, suggesting the idea that once again this entire field of phenomena would be subsumed in a schema of a Newtonian kind, at least from a mathematical point of view. However, Faraday's later discoveries forced scientists to think of electricity and magnetism in an entirely new way and led him to consider these two new theories, that had emerged according to the Newtonian model of gravitational force, as a single unitary theory that included both electrical and magnetic phenomena interacting with one another, inconceivable for Newtonianism.

Following his important discoveries, Faraday's theoretical model seemed to have prevailed upon the rival ideas definitively. However, the mathematics of Newtonian mechanism re-established itself, when from the 1940s new phenomena and laws were interpreted according to a unified ~~schema/~~model. First Kelvin and then Maxwell used mathematics intensively in this new branch of phenomena. Maxwell interpreted electromagnetic phenomena according to a model based on the idea of mechanical vortices at a microscopical level; although today it appears bizarre, it allowed him to formulate his famous equations. This new mathematical formalism distorted Faraday's basic concepts, because it expressed them abstractly and in a different order from the one he himself had conceived; that is, it gave precedence to the local concepts of the neighborhood of each point over the global concepts of physical objects related to fields.

There is not enough space here to indicate the alternative theories that emerged before and after Maxwell; we need only mention the most important, those of Weber and Riemann, based on the effects produced by magnetic charges and dipoles, i.e. the macroscopic material supports of electromagnetic phenomena, without taking into account fields. These electromagnetic theories succeed~~s~~ in reproducing much of the Maxwellian theory, but not all of it: for example the displacement current, not being supported by matter, remains excluded. Other highly developed electromagnetic theories were those of Helmholtz and that of

Poincaré. It was the birth of relativity that put an end to the rivalry between the various theories, establishing the primacy of Maxwell's, the only one able to fit the new theory.

3.3 The theoretical oscillation of the new theory between Newtonian and Carnotian Mechanics

We note that in the historical development that led to Maxwell's Electromagnetism there were considerable oscillations between Newtonian and Carnotian concepts. It can be seen briefly in the following table.

TABLE 3.4

THE FUNDAMENTAL CONCEPTS OF THE PHYSICISTS OF ELECTROMAGNETISM COMPARED TO THOSE OF NEWTON OR L. CARNOT

NEWTON	**Coulomb**	**Oersted**	**Ampère**	**Faraday**	**Weber**	**Kelvin**	**Maxwell**	**Herman n**	***L. CARNOT***
Theory also philosophical	~	~	~	*C*		*N*			Theory completely experimental
Organization by principles	*N*			*C*	*N*	*N*	≈*N*	*C*	Organization on a universal problem
Infiite and absolute space	*N*			*C*		*N*	*N*	*C*	Limited and relative space
Absolute time	*N*			*C*		*N*	*N*	*N-C*	Time before-after
Mass-points	*N*	*C*	*N-C*	*C*	*N-C*	*N*	≈*N*	*C*	Extended bodies, machines
Movement as property of bodies	*N*	*C*	*N-C*	*C*	*N*	*N*	≈*N*	*C*	transmission of motion
Inertia as perpetual motion	*N*				*N*		*N*	*C*	Impossibility of perpetual motion
Basic concept: acceleration	*N*		*N-C*		*N*	*N*	≈*N*	*C*	Quantity of motion
Force-cause, synthesis of the interaction with the environment	*N*		*N*			*N*	≈*N*	*C*	Work and energy
Infinitesimal Math.	*N*		*N*	*C*	*N*	*N*	*N*	*C*	Geometric motions
F=ma	*N*		*N*	*C*	*N*	*N*	≈*N*	*C*	Laws of conservation
Differential equations	*N*		*N*	*C*	*N*	*N*	*N*	*N-C*	Principle of virtual work
Solutions: all trajectories		*C*				*N*		*C*	Invariants of motion
Machines applicatiions				*C*	*N-C*	*N*			Object of the theory
Infinite power of machines									Infinite power is a chimera, machine as universal subject of theory

Key: **N** = Newtonian; **C** = Carnotian; **N-C** = uncertain; **≈N** = almost Newtonian; **~** = indifferent; blank = to be studied.

From the above table we see that the effort to construct the new theory led to an opposition between two theoretical models of reality: on the one hand, the dominant one of Newtonian mechanism (which for example enjoyed a moment of considerable success when Ampère was able to introduce differential calculus to describe the repulsions or attractions between the elements of ~~an~~ electrical wire); and, on the other, the alternative conception, which led to the progressive evanescence of the concept of force-cause and, in contrast, the advancement of the concept of force field and the discretization of electricity.

Ultimately, Electromagnetism, until Maxwell's first work, appears to have been the product of a conflict between two different traditions in Physics, initiated respectively by the Newtonian and the Carnotian formulations of Mechanics. The development of Electromagnetism was therefore for theoretical physics of the time the continuation of the opposition already existing in the theory of Mechanics and had called into question the monopoly of Newtonian theory.

Let us now try to gain a better understanding of what is presented as the conclusion of this story, the Maxwell equations, through a critical analysis of his theory. We need only consider the concept of electric field, leaving aside the others. An analysis of the 'field' concept shows that its definition is partially operational. Indeed the only way we can measure a field is to use the formula $E=F/q$, which suggests introducing a test charge in the space permeated by the electric field; however, the said charge interacts with the charges that generate the field, modifying the initial undisturbed situation. There is a change when the exploratory charges are not point-like, but are distributed over a surface; in this case the interaction involves forces acting on the generating charges, which therefore shift from the original configuration, which has thus been lost. But even when the exploratory charge is small, it is reasonable to think that it modifies the field to be measured. An attempt is made to overcome this difficulty by defining E as the value obtained when an increasingly small exploratory charge is used, its value tending to zero. Obviously this limit is idealistic, because physically either the charge exists or does not exist. All of the above shows that in Maxwell the concept of field is idealized.

That the Mechanics of L. Carnot is the implicit matrix of Faraday's Electromagnetism is confirmed by a modern formulation[3] (shown in the last column of Table 3.5). It is capable of refounding Electromagnetism with a change only in the concepts of Mechanics: usually the physical notion of a force is "reified" by relating it to personal muscular effort, while here it is suggested that the physical notion of momentum e is “reified”. Thus phenomena are conceived in terms of flows of this momentum, according to the correspondence shown in the table below.

TABLE 3.5

COMPARISON OF ELECTRICAL THEORY WITH CARNOTIAN MECHANICS

MECHANICS	***ELECTRICIYÀ***
Physical magnitudes/quantities	
Momentum Ip_i	Electrical charge Q
Current of moment Ip_i	Current of electrical charge I_Q
Velocity v_i	Electrical potential V

3 Falk, G., Herrmann, F. and Schmid, G. B. (1983). 2Energy forms or energy carriers?, *Am. J. Phys.* 51, p. 1074-1077. F. Di Liberto ed al., *Innovazione nella didattica della fisica di base. La Fisica di Karlsruhe: risultati e prospettive*, Napoli: Loffredo. 2006. Includes my evaluation on this approach according to the two dichotomies.

Viscosity η	Conductivity σ
Mechanical resistance R_p	Electrical resistance R_{Or}
Mass of a body, m	Capacity of a charged body, C
Mechanical inductance $1/k$ (k of a spring)	Inductance I
Formulas	
Dp/dt for isolated bodies	DQ/dt for isolated electrical bodies
$Dp/dt + Ip = 0$	$DQ/dt + Iq = 0$
$Ip(1) = - Ip(2)$ for the flow of momentum between two bodies	$I_Q(1)= - I_Q(2)$ for the flow of charge between two bodies
$R_p = \Delta V/I^p$ (R_p momentum resistance)	$R_O = \Delta V/I_O$ Ohm's law
$Kx = I_p$ (x= displacement of the position of equilibrium from the end of the spring)	$N\Phi(B) = LI_Q$ (N=number of coils; $\Phi(B)$ flow of magnetic field)
$P=\Delta vxI_p$ (P power)	$P=\Delta VI_O$
$P=m\Delta v$	$Q=C\Delta V$

In conclusion, reasoning in terms of momentum, we find in both the theoretical attitudes of L. Carnot's Mechanics and Faraday's Electromagnetism an alternative to the theories of Newton and Maxwell, which are based on force as the main concept.

3.4 A new history of Electromagnetism

At the time of the birth of the Maxwell theory in 1864, the history books of Electromagnetism seemed to consider that the evolution of the theory had reached its conclusion. In reality they then had to admit that Maxwell had only hypothetized the displacement current (which is of decisive importance for the existence of electromagnetic waves) theoretically. Its existence was then demonstrated experimentally by Hertz in 1885. Moreover, Maxwell had not formulated the equations in a vacuum, but in matter. This led to the idea of the existence of a new substance, ether, distinct from known chemical substances. The history of the evolution of Electromagnetism is therefore actually longer. Above all the birth of the Lorentz transformations is dated incorrectly, being attributed to Einstein's theory of relativity (1905); it is, however, entirely classical, belonging, that is, to electromagnetic theory.

In the late 1990s Lorentz studied what happens when an electron is in accelerated motion, a very complicated problem, because the acceleration causes the charge to interact with the field. In the study of this problem Lorentz found a transformation for which Maxwell's equations were invariant. It was only several years later that such transformations were regarded as a group, called the Lorentz group. Ascribing this fact, which concerns Electromagnetic theory alone, to the birth of special Relativity means extending the "golden age" of classical Physics (considered essentially Newtonian) to 1905 and making modern physics (Einstein's Relativity) responsible for undermining the classical theoretical framework. This, however, not only places the conceptual and foundational origin of Relativity in the wrong period, but also does not recognize that the Lorentz group is an integral part of Electromagnetism whose invariance group is therefore essentially different from Newtonian mechanics.

Thus the history of Electromagnetism represents an essentially theoretical dispute. It was by no means resolved by Maxwell when he introduced the higher mathematics of differential equations, because the dispute is also present in the relationship of theory with mathematics.

First there was Coulomb's law, which has the Newtonian type of mathematics of gravitational law. Then the importance of mathematics diminished significantly to the point where it seemed absent in the new phenomena (see the case of Oersted's experiment, which at the time did not lend itself to any mathematical schematization, while Faraday, like other physicists, took no interest in mathematical descriptions, considering them an unnecessary addition to experimental description). However, Ampère succeeded in introducing a Newtonian type of formalism again and produced Electrodynamics. But in the same years and then later, Faraday carried out so many new and important experiments (also in order, among other things, to put electricity and magnetism on a par with each other as phenomena) that Ampère's advances were considered by most scientists to be completely formal with respect to true theory. Then Maxwell proposed – after a curious physical model in terms of vortices - a mathematical structure that at that time appeared complete for Electromagnetism as a whole. It seemed therefore that finally the mathematics of differential equations had definitively won the theoretical battle that had lasted for almost a century. However, Lorentz then obtained (albeit in a completely analytical manner) a transformation of the group of those equations. At that very time the Theory of groups developed in Mathematics according to three strands of research (the works of Galois on permutation group in the early 1930s rediscovered a decade later, Klein's theory of geometric groups, the Lie theory of groups of differential equations). This radical novelty with respect to differential equations justifies those who had previously (for example Faraday) wanted to dispense with the mathematics of time, almost as if they wanted to innovate it completely.

Of course, it seems to many historians that in theoretical physics we must necessarily speak above all of differential equations, which did in fact dominate the history of physics of the past. But we know that this is not true, both because, already a century before Maxwell's time, L. Carnot had introduced transformation groups into his theory of mechanics, and because there are no differential equations in the mathematics of thermodynamics. Moreover, bearing in mind the aforementioned battles within Electromagnetism regarding its relationship with mathematics, we are minded to think of the discovery of the Lorentz group as an essential aspect of the theory and therefore as its completion. Hence, considering the various contributions that historically constituted this theory, it should no longer be seen either as dominated by Maxwell's works or as a (tendentially) Newtonian theory, but also as a Carnotian theory, since with Faraday it was based on many concepts of that alternative formulation and also because it shares its specific mathematics, that of groups.

A similar dichotomy can also be noted in the manner of reasoning, either deductive or inductive, or linked to axiomatic principles from which to derive an AO theory, or linked to a search for methodological principles that support the discovery of new phenomena and ideas in a field that is essentially problem-based. The historical evolution of this dichotomy can be summarized briefly. With Coulomb the principle was equality: "electricity = magnetism"; because it seemed that both fields of phenomena could be traced back to central forces, attractive or repulsive. But then Volta created a new experimental apparatus (pile) that clearly showed that "It is <u>not</u> true that electricity = magnetism", because only with the former can circuits and currents be obtained. The two fields of phenomena seemed therefore to diverge from each other definitively. But then Oersted's experiments and Faraday's many works modified the previous certainties and led to a doubly negated proposition: "It is <u>not</u> true that electricity is <u>not</u> magnetism". But then, with the advent of Maxwell's mathematical equations, based on the mathematical symbol of equality, the theory assumed a perspective (shared by textbooks) similar to Newtonian mechanics, entirely affirmative and deductive, at the cost of obscuring the great creative work of Faraday. It seemed therefore that the theory restored equality (in the old manner of Coulomb): "electricity *is* magnetism". However, the equations for electrical and magnetic vectors express their dual roles, not their exact equality; consequently Faraday's doubly negated proposition remains the correct one and, moreover, the affirmative mode of reasoning is not only the one.

Also when we assume these two theoretical characteristics as the basis of a theory, mathematics (differential equations or groups) and principles (axiomatic or methodological), then we can note that the history of Electromagnetism truly exemplifies an oscillation between the two different possible polarities. Moreover, we note that, after more than century of such oscillation, a formulation was arrived at that was no longer compatible with Newtonian mechanics, since it included the Lorentz group. At this point the theory abandoned, with formal and conceptual precision, the theoretical monopoly of Newtonian mechanism. This is the true theoretical framework of the crisis around the concrete question as to the existence or non-existence of the ether.

We emphasize that the history of electromagnetism thus delineated is different from the usual one (usually nothing more than a chronicle of advances, but not very faithful to the oscillations and ultimately minimally interpretative), because we have not considered it according to the usual interpretative categories, based on the introduction of: 1) the concept of field understood by Faraday as a step-by-step action in matter versus Newton's concept action at a distance (infinite) also in the void, and 2) the mathematics of differential equations, considered as the solution to all problems. Instead we have compared the history of electromagnetism with two previous formulations of mechanics: L. Carnot's formulation and Newton's, which diverge from each other by virtue of their different mathematics (transformation groups as opposed to differential equations), their different logic and their different concepts of space and time (not absolute, but relative).

It can therefore be said that the two formulations of Mechanics have incommensurable foundations not only because of their contrasting contents, but also because they have caused oscillations of the history of electromagnetic theory between their polarities. On the other hand, the history of the new theory represents an exceptional intellectual and scientific effort to understand how to construct a new theory dealing with an incommensurability, that of the two rival mechanical theories. This fact justifies, on the one hand, the great effort made by the theorists of electromagnetism and, on the other, the haste of historians, by nature pacifiers of profound theoretical conflicts, to close down the problems encountered by Newtonianism once Maxwell had introduced a mathematics (mathematical Analysis) which gave rise to an ostensibly complete formalism; actually, they disregard the true mathematics of Electromagnetism, the Lorentz group.

3.4 If… Faraday had been born after Maxwell

In literature, many scholars, without referring to alternative theoretical perspectives, have proposed a series of studies to reconstruct the history of electromagnetism in a more linear manner, that is, without following the chronological sequence, but emphasizing the conceptual development of the physical theory. Of course, they accept the risk of producing a history that goes beyond recorded events to the point of creating fictions. But it is also clear that we cannot refer solely to historical documents, because the process of growth of scientific knowledge is very complex and obscure and only a small part of it can be reconstructed from historical documents and from the biographies of scientists. So having on the one hand clarified that the history of Physics is not written with "ifs", the following "fable", suggested by two famous historians,[4] is, on the other hand, instructive, and shows that the real history of Physics contained some interesting alternatives.

We have already noted that Maxwell formulated his equations considering fields in matter (and only later in a vacuum); moreover, for him the phenomena described were in absolute space. His (modernized) equations are (in a suitable c.g.s. system of units of measurement and with vector signs implicit):

$$div\, E = \rho \qquad\qquad div\, B = 0$$

$$rot\, E = -1/c^2\, \partial B/\partial t \qquad\qquad rot\, B = j + \partial E/\partial t$$

We have already seen that Maxwell arrived at his laws in a strange manner, on the basis of mechanical models of vortices. He could, however, have arrived at the local laws of the first line by induction from experimental laws. This also applies to Ampère's local law in the second line. Let us move on to the last law, which contains the displacement current, which Maxwell derived from empirical considerations on a circuit containing a capacitor. Let us see how he might have obtained it with a mathematical derivation. If he had applied his type of mathematics (of gradients, divergences, etc.) consistently, that is, if he had applied the divergence operator to the last law of the preceding four, but with the displacement current (that is, the last addend) removed, he would have obtained

$$div\, rot\, B = div\, j;$$

where the first term is null, by virtue of the properties of the operators, while the term on the right is not. To annul it, he would have had to remember the continuity equation for the electric charge ($div\, j + \partial\rho/\partial t = 0$) and therefore link it to the first equation $\rho = div\, E$; this would have suggested to him that, in order to annul $div\, J$ it would have sufficed to add $\partial E/\partial t$ in the last equation (that is, what will then be called the displacement current); he would then have obtained

$$div\, rot\, B = div\, j + \partial div\, E/\partial t = div\, j + \partial\rho/\partial t,$$

where now the last of the 3 equated terms is also null by virtue of the continuity equation.

Having reconstructed Maxwell's equations according to a clear theoretical standpoint, we now see the implications of this historical alternative for the birth of special Relativity. We note that if there were no Faraday term in the third equation ($-1/c^2\, \partial B/\partial t$) Maxwell's equations would still form a complete set of equations, since, by assigning the charge and current

---

4 M. Jammer e J. Stachel, "If Maxwell Had Worked between Ampere and Faraday: An Historical Fable with a Pedagogical Moral." *American Journal of Physics*, 48 (1980) 1, 5-7.

sources, they determine the magnetic H and electric E vector fields, even when these depend on time. Indeed, since now $rot\ E = 0$ as well as $div\ B = 0$, we can introduce, as we do in electrostatics and magnetostatics, the scalar potential ψ and the vector potential A (for which $div\ A = 0$ holds):

$$E = - grad\ \psi \qquad\qquad B = rot\ A$$

Hence the first equation takes the form of the Poisson equation

$$div\ (- grad\ \psi) = \rho \qquad i.e.\ \nabla^2\psi = -\rho\ ;$$

while, introducing in the fourth equation the vector potential A

$$rot\ rot\ A \equiv grad\ div\ A - \nabla^2 A \equiv - \nabla^2 A = j + \partial/\partial t\ (-grad\ \psi)$$

Thus, given ρ and j, the two Poisson equations determine completely A and ψ. This is analogous to how in the gravitational field the Laplacian equation $\nabla^2 V = -\sigma$, with σ the mass density, determines the potential V. Therefore in this theory of an electromagnetism without the Faraday term, the electric field and the magnetic field do not have independent degrees of freedom; they are only tools to describe instantaneous interactions between charges and currents. There are therefore, it should be noted, no electromagnetic waves (which give interactions at speed c). We can therefore say that this theory, unable to describe electromagnetic waves, is not electro*dynamic*.

Now (in 1973 two authors,[5] discovered that) these equations are invariant with respect to the Galilean transformations. Verifying it is a little complicated (we need to consider the transformation of carriers, as well as operators). But the final result can be obtained in a rough but simple way by simply applying the correspondence principle to all Maxwell equations written in the above form; these, with the third equation, depend on c. This dependence disappears when $c \to \infty$. We conclude that without Faraday's term Maxwell's equations give the actions at a distance of classical mechanics and are therefore invariant with respect to Galileo's transformations (but not the Lorentz transformations, which we know hold for equations with that term).

Maxwell might have noticed it if he had paid attention to the scholars's interest in transformation groups that existed from the second half of the 1800s, both in crystallography (from 1784) and in geometry (F. Klein's Erlangen programme in 1871, according to which each geometry is characterized by its specific transformation group. He could therefore have shown that without Faraday's term his equations are invariant in the Galilei group. This result would have been very comforting, because it would have confirmed the validity of that same Galilean invariance group that was valid for Newton's mechanics. This would have represented the closing of the circle, associating the new Electromagnetism with the old Newtonian paradigm and concluding that, even if electrical and magnetic phenomena are completely different from mechanical phenomena, the mathematical formalism of the new physical theory is substantially the same.

---

5 M. Le-Bellac and J.-M. Levy Leblond, "Galileian Electromagnetism", *Nuovo Cimento,* 14 (1973) 217-233.

But let us continue the fable. If in this imaginary historical situation Faraday (remember that almost all of Faraday's ideas (tab. 3.4) are Carnotian) had come after Maxwell's reduced equations and had demonstrated the phenomenon of electromagnetic induction and therefore had introduced his term into the Maxwell equations, there would have been a crisis in theoretical physics. Physicists would have had to decide whether to try to reject Faraday's experiments (which even in the history that actually occurred were not easily accepted), or to accept Maxwell's equations but completed with the new term. In this second case they would have noticed that the principle of Galilean relativity could no longer hold true, precisely because the group of transformations of the new equations (Lorentz) is incompatible with that of Newtonian mechanics (Galilei). Electromagnetic actions could no longer be understood as actions at a distance, as they were at that time, and the idea of electromagnetic waves which do not need a medium (ether) to propagate in would have been born. This would have immediately called into question the existence of the ether, with all that follows, and led to the birth of special Relativity.

Instead, in actual history, in order to get closer and closer to the Newtonian paradigm, Maxwell's Electromagnetism became trapped in absolute space, a concept that led to belief in the existence of the ether. This Electromagnetism therefore did not even establish the Galilean relativity (which some had already noticed[6]) of classical mechanics. The ultimate consequence of all that was the well-known fiasco from which physicists were saved precisely because Maxwell's equations contain in addition the Faraday term of electromagnetic induction, electromagnetic waves and the theory of light (optics); in short, electrodynamics, not just an electromagnetic theory.

In actual history we also note that when Einstein introduced his theory of relativity he was perspicacious. His work of 1905 started precisely from electromagnetic induction: the description of the phenomenon presented a dissymmetry between the two cases of relative motion (in the case in which the magnet moves, an electric field is created at every point in space, so that in the whole of space, and also in the coil, an (electromotive force) e.m.f. is generated with the consequent current, while in the case in which the coil moves, the lines of force intercepted by the coil only create in it an e.m.f., which generates the current in that circuit).

Thus the point of departure was once again Faraday's induction.

> In this sense, and not historically, we can say that Faraday is the father of the theory of special relativity, because [in Maxwell's equations] it is the law of induction that creates the conflict between electrodynamics and the principle of Galilean relativity, a conflict which was resolved by Einstein only in 1905.

The history of special relativity might therefore have begun much earlier, if physicists had realized the importance of the group of transformations to which a physical theory is subject. Furthermore, the birth of Relativity would not have occurred by way of the resounding fiasco that was the ether, but through the progressive and natural evolution of the problems of theoretical physics, which would have gone through a crisis anyway, that generated by the post-dated Faraday, but this crisis would have been the result of growth, not due to an error.

---

6 For example the book on mechanics by Jules Violle (A. Drago e A. Pirolo; "Analisi del testo di Violle: il libro di meccanica di Einstein", *Giornale di Fisica*, lu. sett. (1999), pp. 129-141.

Even more radically, we can say that the entire history of Electromagnetsim depends on the different choices of the two theories; while that of Maxwell, which is of the Newtonian type, is essentially an AO (to the extent that it reasserts absolute space and time, and thereby insists on the existence of the ether), the other theory that indicates the Faraday term is essentially based on the Lorentz group and therefore raises the problem (PO) of compatibility with the previous Newtonian theory, leading in turn to the birth of Relativity through the problem of reconciling these opposite choices. In terms of concepts alone, if scientists with the alternative view of electromagnetism, like Faraday, had kept in mind that in the (then ignored) Mechanics of L. Carnot space is relative and that it was therefore necessary to pay attention (as did Leibniz) to the invariance group (Galilean), they would have immediately established the premises for the alternative historical development outlined by the above "fable" and so also for a less traumatic birth of Relativity.

CHAPTER 4

NEW HISTORY OF THE FOUNDATIONS OF MODERN PHYSICS: SPECIAL RELATIVITY. THE EMERGENCE OF THE TWO ALTERATIVE CHOICES ON THE DICHOTOMIES

4.1 Which choice of organization in relativity theory?

We note that so far we have interpreted classical physics through two dichotomies (one regarding the type of infinity and therefore the type of Mathematics) and one regarding the type of organization of the theory (therefore the type of Logic). We have derived a history of physical theories which goes beyond that of the dominant paradigm of Newtonian mechanics in that it appears to consist in several lines of development, each theory represented by different formulations. In other words, we identified a new foundational structure and renewed a history of physics that had appeared unilinear and almost deterministic in its incessant expansion.

Now we must verify whether the two dichotomies are valid for interpreting the history of physics during and after the crucial revolution that occurred in the early 1900s. When dealing with modern physics we encounter a greater problem because its history is not at all clear (despite the considerable efficiency of its formalisms), if only because the research orientation of Relativity does not appear to be compatible with that of quantum mechanics. Why did the two revolutions, each unifying theoretical physics, give rise to two decidedly different theories?

Here we will go beyond consideration of the efficiency of their formalisms (which in the following will be assumed only to the extent necessary for introducing their main concepts) and study these new theories in order to clarify their fundamental aspects and thus find their respective foundational structures.

It is clear that here we must face a crucial historical problem, resulting from the crisis and subsequent "cultural revolution" of the beginning of the 1900s. The revolutionary events pose the question: are the two dichotomies found in classical physics still valid in modern physics? We will verify not only that they are, but even that they actually explain the crisis and the cultural revolution. We will proceed with the usual method: we will study each of the two new theories to identify the choices of the usual formulations; if we do not succeed in doing this we will look for the alternative PI and PO formulations of these theories, even at the cost of reconstructing the two theories.

Let us start with the first modern physical theory: Relativity. First of all it must be remembered that it consists of two parts, special and general. Special Relativity (1905) is more supported by experimental evidence than General Relativity, but it is however an inadequate theory, because it does not deal with accelerated motions, which constitute an essential component of the kinematics and the dynamics of mechanics.

In order to remedy this essential deficiency, about twelve years after its birth Einstein began to formulate a new theory, which he called General Relativity. Its dynamics was reduced to kinematics: here masses affect the curvature of space, which is traversed by other bodies (provided they are small) along geodetic lines (of minimum distance). However Einstein's work and that of many other scientists who followed this field of research are highy debatable: the theory of General relativity can boast of having explained some effects (such as the precession of the perihelion of Mercury) but these can also be explained by other non-linear theories (which is what General Relativity is). A current assessment is that Einstein's initial approach is certainly not the most suitable and that in any case today's theory is still incomplete.

In the following we will consider only special Relativity. According to our interpretative categories it is very clear which choice of organization is rejected by this theory: the choice of AO. One first reason for this choice is the well-known rejection of absolute space, the concept with which theoretical physicists had subjectively materialized AO in previous centuries.[1] The new concept of space, developed by means of experiments that make possible a definition of simultaneity, has nothing of the self-evident character that Aristotle required of the principles of a scientific theory; indeed, it is a difficult concept, because in the end it is composed of two basic concepts: length and time.

A second reason for the rejection of AO is that the theory has no principles from which to deduce the whole set of laws. In fact, special relativity speaks of two "principles", the first of which states that the speed of light cannot be exceeded. But this is not a principle for obtaining deductions; it is clearly a principle of impossibility. Einstein wrote that he conceived it analogously with the principle of the impossibility of perpetual motion in thermodynamics; we know well that in thermodynamics this principle has the character of a methodological principle because it's the starting point of an essentially PO theory, in that it is based on the problem of the efficiency of thermal machines. This principle is therefore methodological, it is precisely thanks to it that we can try to reorganize space and theory.

The second principle of special relativity is, precisely, that of relativity, that is, of invariance in the form of physical laws for changes in inertial reference systems. Once again it is not self-evident or certain. It is a methodological principle and only indicates the beginning of that method which is necessary for us to understand each other when we speak of Physics from different, but hopefully equivalent, reference systems. One corroboration of the absence of true self-evident principles in special Relativity is the inconclusive result of more than a century of efforts to obtain its own axiomatics (it was possible to obtain the Lorentz group from a set of hypotheses, such as isotropy and the homogeneity of space, or from the law of propagation of electromagnetic waves, but always without succeeding in determining a constant or a parameter that remains to be specified).[2]

---

[1] It can easily be seen in the series of classical physicists who constructed a theory; those who had the idea of absolute space chose AO, those who founded PO theories had a relative concept of space. A. Drago, "La storia del concetto di spazio quale rivelatrice delle scelte fondamentali di una teoria fisica: I. La correlazione tra lo spazio e la organizzazione della teoria", in F. Bevilacqua (ed.): *Atti VII Congr. Naz. Storia Fisica,* Padova (1986), 113-118.

[2] M. Strauss, *Modern Physics and its philosophy,* Dordrecht: Reidel, 1972, 133-137. N. Mermin, "Relativity without light", *Am. J. Phys*., 52 (1984), 119-123.

Einstein himself gives us three proofs to indicate that the choice of his theory is PO. The first is in his paper of 1905, the very foundation of the theory; first there is the discussion, based on rulers and clocks, of the concept of simultaneity and, further on, the new version of all the old concepts of physics. This is a typical methodological discussion, by no means the beginning of an axiomatization.

The second proof is also found in the original article of 1905[3], where he formulated the problems of his research by means of doubly negated sentences, which we know to be typical of a PO. Let us see.

The problem of the asymmetry of the formulas of classical electromagnetism for Faraday's law of induction: "... asymmetries that do not seem to be inherent in phenomena". Here the negation is followed by a dubitative verb; the exact phrase would be: "... it is not true that they are not inherent...".[4]

The ether problem: "Examples of this type lead to the conjecture that the concept of absolute rest does not correspond to the properties of phenomena ...". Here the negation concerns the concept of non-relative rest, the relative concept being the only concept that can be verified by an experimental physicist.

The problem of the compatibility of the two principles: "... another postulate, which only apparently is irreconcilable with the first ...". The phrase with the "apparently" is expressed more precisely by: "It is not true that the two principles are irreconcilable.

The problem of the constancy of the speed of light: "... a precise speed *c*, which is independent of the state of motion of the emitting body". Here one word accumulates the two negations; the dependence of motion on the emitter is a negative property for the physicist, because it requires an explanatory theory.

The problem of the definition of simultaneity:

> Therefore we see that we cannot attribute an *absolute* meaning to the concept of simultaneity, but that two events that, examined by a system of coordinates, are simultaneous, can no longer be interpreted as simultaneous events when they are examined by a system that is in motion relative to that system.

In the first sentence, the word that Einstein puts in italics, "absolute", must be replaced by "not relative", for the reason already mentioned. The last sentence repeats the same concept, but using the words "in motion", instead of "not at rest".

Einstein concludes: "We assume this definition of simultaneity is free from [= without] all contradictions".

The third proof is the suggestion, in Einstein's epistemological writings, of the existence of two types of theories: 'constructive' (e.g. statistical mechanics), which start from a priori hypotheses concerning the constitution of reality (which is thus decomposed into its elements

---

3 A. Einstein (1905), "On the electrodynamics of moving bodies", *Annalen der Physik*, 17, 891-921 https://einsteinpapers.press.princeton.edu/vol2-doc/311.

4 For an analysis of all the double negatives in the text see A. Drago's work: "La teoria delle relatività di Einstein del 1905 esaminata secondo il modello di organizzazione basata su un problema", in E. Giannetto, G. Giannini e M. Toscano (edd.), *Relatività, Quanti, Caos e altre rivoluzioni della Fisica*, Atti del XXVII Congr. Naz. di Storia della Fisica e della Astronomia, Bergamo 2007, Rimini: Guaraldi, 2010, 215-224.

and then recomposed), then verified in their empirical consequences; and theories of 'principle' , based on a general phenomenological principle (such as thermodynamics, based on the principle of the impossibility of perpetual motion.[5]) It is clear that, even if his words ("constructive" and "principle") are deceptive, this distinction is an approximation to the precise one between AO and PO; it is proved by the examples of the theories indicated by Einstein (AO Boltzmann's statistical mechanics and PO Thermodynamics). In particular, it is important for us that Einstein explicitly states that his theory of special relativity is a "theory of principle". In fact, we immediately see that the theory is clearly PO, since it is centred on the universal problem of how to understand the same physical law, though communicating from different reference systems.

However, if we broaden the theory, we see that *the organization of general relativity changes completely, becoming AO*. It begins in the traditional manner of the physical theories of Descartes and Newton: physics derives from geometry. In other words, geometry, even if modified as Riemannian space with variable curvature, is still understood by Einstein as the self-evident mathematics on which to build, as an extension, Mechanics. Hence the theory first posits Mathematics-Geometry, understood as an "a priori" governing reality, and then, as its physical extensions, the empirical laws of reality.

We must conclude that *Einsteinian relativity*, *as a whole, has an essential ambivalence about the choice of the organization of the theory*, while special Relativity alone is clearly PO.

4.2 Which choice of Mathematics in relativity theory?

Which mathematics? The one used by Einstein does not have actual infinity (AI) idealizations, although it is traditional mathematics, which also uses differential equations and expressions, which, in classical physics, indicated the actual infinity (AI) of infinitesimals and their combinations in equations. But in modern physics it is not necessarily true that they are sufficient proof for AI, since constructive mathematics, emerging at the beginning of the 20th century, is able to recover even differential equations with PI alone. If we examine the paper on special relativity taking this into account it is not clear if there is a choice concerning mathematics.

But we note that with this theory he essentially changed the group of invariance of the fundamental equations of Mechanics: from Galileo's group of transformations to that of Lorentz. Since symmetry groups constitute a typical algebraic technique of the constructive mathematical tradition (PI, as in the theory of L. Carnot), it is to be expected that the theory is based on PI. In fact in the aforementioned article Einstein uses a differential equation to

---

[5] In fact Einstein had been preceded by H. Poincaré, *La Science et l'Hypothèse*, Flammarion, Paris,. 1902, chp. "Optique et Electricité". H. Poincaré, *La Valeur de la Science,* Flammarion, Paris, (1905), chp. VII. See Flores, F. (1999). Einstein's Theory of Theories and Types of Theoretical Explanation. *International Studies in Philosophy of Science, 13,* 123-134.

derive the Lorentz transformations; but this derivation is wrong, because (as was immediately pointed out to him) he adds the speed of the body (or system) v to the speed of light c, against his own postulate. Then he also makes use of differential calculus when he verifies that Maxwell's equations are invariant to those transformations; but since derivatives, even partial, are constructive operations, there is nothing that does not belong to constructive mathematics. But all this is admittedly a qualitative analysis and might leave doubts. Therefore we will give a *precise* answer *regarding the choice of the type of Mathematics* later on.

In General Relativity there is no restriction on the use of traditional analysis, since it starts from the infinitesimal differential expression of space traversed, ds^2. This involves choosing the idealizations of AI. Moreover, there is no precise group of transformations underlying the theory; therefore the typical technique of PI is also lost. We conclude that *the choice of mathematics by Relativity as a whole is not clear, and in any case the choice of General relativity is that of Special relativity*.

4.3 Special Relativity and L. Carnot's Mechanics

We can take up the problem again in another way: let us leave Einstein's paper out of consideration and reformulate special Relativity strictly following the PO choice, which appeared to be his clearest choice.

We have already stated the fundamental problem of the theory. The following is the principle that seems most characteristic: $v \leq c$. In order to be a methodological principle, it should be expressed with a doubly negated sentence. In fact it can be said: "It is <u>impossible</u> to measure the velocity of a material body that is <u>not</u> less than that of light, c" (which exactly states the experimental situation and also explains the unusual sign for less than or equal to.) It is now a matter of using this methodological principle to first clarify the relationship between physical theory and mathematics, so that the end result is the Lorentz invariance group.

If we observe this principle we note that it concerns the velocities and not the positions or the trajectories of the bodies. Therefore, if it is to develop a new method that solves the problem of the invariance of physical laws, it must refer to the space of velocities, not to the (Newtonian) space of positions. So we do not have to concern ourselves so much with what change the concept of time or the concept of simultaneity in relation to positions undergoes (as Einstein did, in this sense remaining tied to the typically Newtonian approach); we need only reason in the space of velocities.

We can do this because we are already familiar with a theory founded and developed exclusively in the space of velocities, namely L. Carnot's Mechanics. In fact, this formulation is precisely the one that can be linked to relativistic mechanics: if from the formalism of the second we want to return to classical mechanics by executing the step $c \to \infty$, the interaction allowed in Relativity, that of contact, corresponds to, for example, L. Carnot's formulation, not Newton's, in which the admitted interaction (for example, gravitational) is instantaneous at a distance.[6]

---

[6] H. Goldstein, *Classical Mechanics*. New York: Addison Wesley, 1980, p. 332.

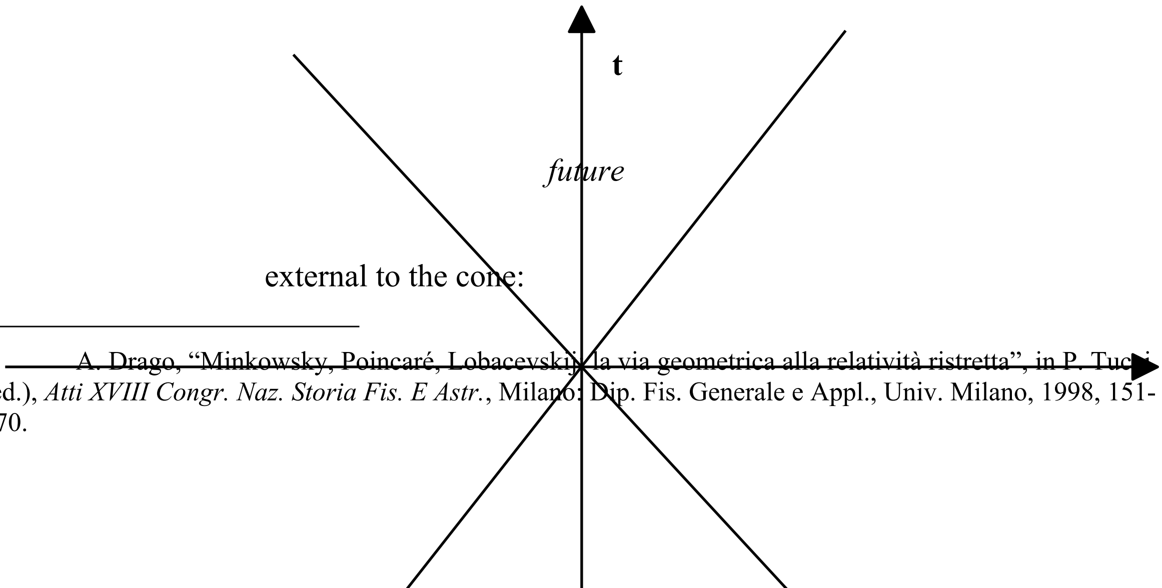
4.4 The choice of Geometry in special Relativity

L. Carnot's theory of mechanics suggests that the crux of a PO theory is its relationship to mathematics, in particular with geometry. Shortly after L. Carnot Non-Euclidean geometries and many other geometries emerged. How many geometries are there then to choose from?

Poincaré in a study of 1887 (unfortunately ignored by later geometers) established with three different methods (quadratic surfaces, transformation groups and a system of axioms including all the Euclidean and non-Euclidean Geometries) that there are four main Geometries: Euclidean (represented in the simplest way with a flat surface), elliptic (whose simplest quadratic surface is ellipsoidal, as is approximately the surface of the earth), hyperbolic (which can be represented by the surface of a hyperboloid of two sheets: for example the surface of two facing conical electrodes) and a fourth geometry (represented by a hyperboloid of one sheet: the surface of a stylized cup). The last one has very different properties from the others, because in it geometric figures (and therefore also solids) are not preserved in the displacements, which is equivalent to the radius of curvature not being the same at every point of its space. Poincaré was surprised that the geometers left the fourth geometry out of consideration. Actually, Helmholtz's prejudice that the geometry of theoretical physics should leave rigid bodies invariant under translations was commonly shared. It is clear that this property holds true in Mechanics, certainly not in both Thermodynamics and Electromagnetism: a gas and light, which are not rigid bodies, do not satisfy this property.

Now we will show that the four Geometries correspond to the four pairs of choices concerning the two dichotomies. It should be formally decided, establishing which type of Mathematics and which type of Logic are the foundation of each of them. Instead we will decide more rapidly and simply at the subjective level, by means of two characteristic geometric elements which substitute for the two fundamental dichotomies: the straight line, which represents the typical geometric process of construction of a infinite line-segment, substituting for the choice of infinity; and the radius of curvature, which, since it represents the typical geometric magnitude that is characteristic of the globality of space, substitutes for the organization of the theory.[7]

Fig. 4.1: The light cone



[7] A. Drago, "Minkowsky, Poincaré, Lobacevskij: la via geometrica alla relatività ristretta", in P. Tucci (ed.), *Atti XVIII Congr. Naz. Storia Fis. E Astr.*, Milano: Dip. Fis. Generale e Appl., Univ. Milano, 1998, 151-170.

non-real events **x**

past

It is clear that Euclidean geometry (represented by a plane) has an infinite radius of curvature, while the lines are the well-known ones, which can only have one parallel, because, being straight, they also indicate with absolute precision their point at infinity. Elliptical geometry (represented by an ellipsoid) has a real and positive radius of curvature; the lines are the maximum circles and therefore are periodic; since each pair of them intersect at two points, there are no parallel lines. On the other hand, in hyperbolic geometry (represented by a hyperboloid with two sheets) the radius of curvature is imaginary and finite; the lines are infinitely extendable, but the point at infinity of one end-point is not the same as that at the other end-point; therefore there are two parallel lines, one on the left and the other on the right of the given line (and many others between the two). The fourth geometry (representable on an ellipsoid with one sheet) has a different radius of curvature depending on the region of space.Indeed, Poincaré's fourth geometry is none other than the geometry of Minkowsky's space (the mathematician who, without mentioning Poincaré's study, proposed this space as the most suitable for special Relativity). If then we give the classical representation of this space with the light-cone, we note the following properties: on the points of the light cone ($ds^2 = c^2dt^2$) the radius of curvature is infinite, on points at spatial distance (i.e. inside the light cone, $ds^2 < c^2dt^2$) the radius of curvature is negative, as in hyperbolic geometry, while it is positive in the time zone (i.e. outside the light cone, $ds^2 > c^2dt^2$). Similarly, the lines are different depending on the region.

In summary, we have the following table, where for each geometry the quadratic form which represents it is indicated in parentheses.

TABLE 4.1

CLASSIFICATION OF THE FOUR GEOMETRIES ACCORDING TO THEIR GEOMETRIC CHARACTERISTICS AND THE TWO DICHOTOMIES.

	Real radius of curvature	***Unreal radius of curvature***	
Actual infinite straight line	**Parabolic or Euclidean G.** *(paraboloid elliptic)*	**G. of Poincaré-Minkowsky** *(Hyperboloid of one sheet)*	ACTUAL INFINITY
Finite straight line or potential infinity	**Elliptical or Riemannian G.** *(ellipsoid)*	**Hyperbolic or Lobachevskian G.** *(Hyperboloid of two sheets)*	POTENTIAL INFINITY
	ARISTOTELIAN ORGANIZATION	PROBLEM-BASED ORGANIZATION	

The table shows that each particular Geometry corresponds to a pair of choices concerning the two dichotomies (note that this link is not deterministic, given that the Geometries are precise mathematical theories, while the pairs of choices indicate scientific-philosophical conceptions. For this reason the fourth geometry, for example, is to be understood, not as the last possible geometry, but as the group of less structured geometries, which can be replaced, depending on the case in question, by one of the other less structured geometries).

However it is a fact that today the most applied Geometries in the practice of theoretical Physics number just four, the fourth being Minkowsky's for Relativity. This confirms once again the ability of the two dichotomies to represent the foundations of science, in this case of that scientific theory, Geometry, which is the oldest and most important for the relationship between physics and mathematics.

We note that Newton's Mechanics, AI & AO, was based on Euclidean geometry, also AI & AO. If in special relativity we consider the principle $v \leq c$ as a methodological principle, it immediately tells us to impose c as the maximum speed limit: this requires a geometry with a non-infinite radius of curvature. In fact, special Relativity abandoned absolute space and with it Euclidean geometry which has an infinite radius. So either the radius of curvature must be taken as either finite and real (elliptical geometry), in which case however light rays, travelling along the maximum circles, would return from behind, which is physically absurd; or the radius of curvature is finite; and then only two cases remain: the radius of curvature is either imaginary (hyperbolic geometry) or is not constant (Minkowsky's geometry).

4.5 Hyperbolic geometry and special relativity

As it is taught in universities special Relativity is associated with Minkowsky's Geometry, which is AI & PO, while hyperbolic geometry is PI & PO; their different choices again regard the type of infinity: this ambiguity recalls the problem we saw in par. 4.3: which Mathematics in special Relativity? It needs to be clarified with an appropriate study.

It is a historical fact that Lobachevsky himself, when he invented hyperbolic geometry, immediately thought of another Mechanics, which some Russians then developed tentatively. However, immediately after Einstein's work, in 1909 Sommerfeld reformulated special relativity in the space of velocities (without realizing that he was using Lobachevsky's space). In this space the metric is given by

$$dv^2 = dv_x^2 + dv_y^2 + dv_z^2;$$

that is, a completely positive quadratic form. This, along with the limitation $v \leq c$, represents hyperbolic geometry in three-dimensional space, according to what is usually called Klein's model. In this space the translations (in v) are precisely the Lorentz transformations.

If we now consider light rays, we have for their velocities c: $c^2 = dx^2/dt^2$ (with x a spatial 3-vector), that is, $dx^2 = c^2dt^2$, from which $dx^2 - c^2dt^2 = 0$; we thus obtain the metric:

$$ds^2 = dx^2 + dy^2 + dz^2 - c^2dt^2.$$

This is a degenerate metric because the signs of its addends are not all positive. It gives the geometry of the light cone (generated by actually infinite lines); it is divided into distinct parts; the outside of the cone (temporal part $dx^2<c^2dt^2$ of the non-real events, because their velocity is greater than that of light) and the inside (spatial part $dx^2>c^2dt^2$ of the real events which the plane t=0 -divides into past and future. It is evident that here the curvature is not constant, the distances can be null even between distant points (those of the light cone) and the straight lines can be either actual infinity (on the light cone), or potential infinity (in the spatial part, where the "straight lines" are represented by hyperboles given by $1/x=cost$). It is the pseudoeuclidean metric of Minkowsky space, which in fact is none other than the fourth geometry already indicated by Poincaré in 1887.

Thus we note that the complication that space usually represents for special Relativity depends entirely on whether we want to represent it as a space of positions, as in Newton's Mechanics. Since, however, as we have seen, "true" space is that of velocity, then the space of position is a derived notion. It is necessary to conceive it by adding time, so that we measure two quantities together, as Einstein does at the beginning of his paper: the spatial distances travelled with rulers and the travel times with clocks; and with this both Einstein and the usual approach consider first the concept of simultaneity, which unites space and time. For this reason then it is necessary to use the unusual pseudo-Euclidean metric of Minkowsky space, which is not at all easy to deal with; in addition Physics must restrict its attention exclusively to the significant part of all of that space, the inside of the light cone.

But if we interpret the minus sign of the metric as an i^2, and then make the transition to complex space (whose mathematics will become habitual in quantum mechanics), we make the metric of that space entirely positive again.

If we posit $dx_4=icdt$, then $ds^2+i^2c^2dt^2=0$, that is: $\Sigma dx_i^2=0$ (i= 1,…4).

Then the speed of a point (for convenience, let us consider the case of a single spatial coordinate) is given by

$$v = \frac{dx}{dt} = ic\frac{dx}{d(ict)} = ic\frac{dx_1}{dx_4} = ic\tan\varphi = c\tanh\psi$$

where j is the angle of the incremental relation; the angle ψ is interesting because in Minkowsky space it is the rotation angle with respect to the coordinate axis. Note that relativistic mathematics is therefore that of Trigonometry, the same that characterized L Carnot's Mechanics, which did without infinitesimal analysis. Moreover, according to the relativistic principle $v \le c$, the space of velocities of special Relativity has an essential novelty with respect to classical Mechanics: it has a unit of measurement which is fundamental; just as there is one in the trigonometric circle, where all lengths are measured with reference to the radius of the circle, obtaining as a result pure numbers for the arcs and chords. In the same way, in this space of velocities we can measure all the velocities with the unit of measurement c (or its sub-multiples), thus obtaining pure numbers.

The usual formulas follow.[8] In particular, the formula for the addition of velocities:

---

8 See chp.. III of the excellent text W. Pauli: *Theory of Relativity*, London: Pergamon P.,1981.

$$v_1 + v_2 = c \tanh \psi_1 + c \tanh \psi_2$$

is now the simple addition of angles according to the hyperbolic tangent formula: $c\tanh(\psi_1+\psi_2)$. This works very well for defining the invariance group, which is simply additive in the definition parameter. This shows therefore that special Relativity can actually be developed entirely in hyperbolic Trigonometry and can therefore also be taught in high schools, in order to extend/develop a mathematical theory (Trigonometry) which today is somewhat neglected, because apparently superseded by modern mathematics. Ultimately, special relativity needs only potential infinity, PI.

Note that ψ is an angle and at the same time measures a velocity. It is called rapidity to distinguish it from classical velocity.[9] In special relativity there are three relativistic procedures for measuring velocity. Suppose we are on a train and travelling at a speed close to that of light. We can measure the train speed in the usual way, that is with a clock placed in the reference frame at rest (or looking from the window at the time elapsed on the clock of the station that we pass through) and with the measurement of the distance travelled, carried out still in the same system : this is classic velocity v. We can then measure the times again in the same way, but by measuring the distances on the train (e.g. by counting the sound of the jolts of the wheels on the sleepers) so that we obtain another concept of velocity, which is what compose the definition of relativistic momentum, $v\gamma$ (with ■ the well-known relativistic factor) . Finally we can also measure the time in the system in motion, for example as on a spaceship: measuring the acceleration function which it's subjected to and integrating it in one's own time τ: $\psi=\int a(\tau)d\tau$. We obtain therefore a new concept of velocity ψ, which is called rapidity. (Note this triplication of a classical concept, which represents an example of radical variation of meaning of a physical concept in the transition from classical to relativistic mechanics).

Ultimately, the most appropriate space for special relativity is not so much the space of Poincaré-Minkowsky, but that of Lobacevsky, that of velocities. Notice that this space is characterized by the PO & PI choices, while Minkowsky's is characterized by the PO & AI choices, that is, the latter has preserved a mathematical link (AI) with the Newtonian formulation. It is precisely that link that Einstein's 1905 article maintains with his initial discussion on simultaneity, a concept involving the two initial concepts for Newtonian mechanics, space and time which represent AO and AI; indeed it is clear that he had first of all to introduce a new invariance group, which certainly does not represent Newton's mathematics with AI.

This partially conservative approach of Einstein appears even more evident if we refer to the subsequent General Relativity, in which Geometry is once again that of positions and moreover has a variable curvature depending on the presence of the masses. This focus on space coordinates shifts attention to the analytical formalism for calculation of the curvature and ascribes lesser importance to the problem of eventually choosing the type of Geometry suitable for the particular field of phenomena studied and therefore the particular invariance group.

---

9 J.M. Levy-Leblond: "Speed(s)", *Am. J. Phys.*, 48 (1980) 345-347.

4.6 How to base special Relativity on L. Carnot's Mechanics

The theory of special relativity has from its outset a crucial and universal problem: that of "reconciling" (as Einstein says) the invariance group of electromagnetism (i.e. the Lorentz group) with mechanical theory, which instead (if we abandon Newton's absolute space and time) is invariant with respect to Galilei's transformations (which it might be better to call Huygens-Leibniz's: they made it a fundamental tenet of their theory of mechanics). It can also be said that the problem is that of extending the Euclidean invariance group by keeping what we have seen to be more important in the previous formulations of mechanical theory; that is, not the typical concepts of the Newtonian formulation (for example the equation $f = ma$ or even the concept of force itself, which is not invariant in relativity, or the usual concepts of space and time); but rather the classical conservations (over time): that of energy and that of momentum, those that are typical of the L. Carnot formulation.

This problem can be expressed with a doubly negated sentence: "It is <u>impossible</u> that the preservation (in time) of classical physics is <u>changed</u> in extending mechanical theory to a new group of spatial transformations" (which does not mean that they "remain equal ", because then there would be no generalization).

The problem can be solved by introducing changes of the reference frames.[10] If u is the speed of the moving reference frame~~s~~, then in the two frames we have:

$$\sum E(v_i) - \sum E(v_i') = 0 \qquad\qquad \sum p(v_i) - \sum p(v_i') = 0$$
$$\sum E(v_i + u) - \sum E(v_i' + u) = 0 \qquad\qquad \sum p(v_i + u) - \sum p(v_i' + u) = 0\,.$$[11]

It is clear, however, that in the transition to special relativity we can hope that the form of the laws of invariance (in time) remains the same only if we modify the definitions of the energy functions *E(v)* and momentum functions *p(v)*.

Suppose that, as usually happens in theoretical physics, the two new functions p and E are analytic. Expanding them in the Taylor series we note that the new functions would give an infinity of invariants, as many as are the terms of the development in series, unless we link together the different derivatives for each series. The simplest solution is to impose a cross-dependency:

$$\frac{dE}{dv} = kp \qquad\qquad \frac{dp}{dv} = kE$$

so that only two invariants remain, as in the case of classical Mechanics.

Let us now operate on those two equations in order to define the form of the new functions energy E (v) and momentum p (v). Deriving a second time, the previous relations give rise to a second order differential equation,

---

10 J.M. Levy-Leblond, "What is so "special" about "relativity?", in A. Janner (ed.), *Group theoretical method*; Berlin: Springer, 1976, 617-627.

11 Note that the velocity indexes could also be different, to take into account the creation and annihilation of particles.

$$\frac{d^2E}{dv^2}=k\frac{dp}{dv}=k^2E \qquad\qquad \frac{d^2p}{dv^2}=k\frac{dE}{dv}=k^2p$$

the solutions of which are:

$$E=E_0\sinh(kv) \qquad\qquad p=p_0\cosh(kv)$$

We know that c is a limit velocity, so that we can change the notations in this way: $k=1/c$, which gives $v=c\psi$. Dividing the two, we then obtain the relation $v=c\,\text{tangh}\psi$.

Now we will physically interpret the angle ψ, or rapidity. We have already seen that it can also be understood as the proper acceleration of the body integrated in its proper time, or as the angle of rotation of Minkowski space.

Now it is easy to see that the composition of the velocities of two reference frames is expressed precisely by the *tanh* of the sum of the velocities. This transforms the Lorentz group of the transformations of the theory into an additive group, i.e. based on the simplest mathematical operation. It is then easy to derive the transformations of any other magnitude.

We note that we are in the space of velocities, that is in hyperbolic geometry. Carnot's Mechanics could also be considered in this geometry. It depends on the Principle of virtual works applied to impact, an instantaneous phenomenon that occurs at a precise point. Now, since the 1700s many physicists had wondered what would have remained unchanged if the parallels axiom had been dropped. It is clear that forces, understood as vectors applied to a point, are invariant, but not their composition when they are applied to different points, given that the distance between those points then comes into play. In the case of Carnot's Mechanics there is no distance between the forces (or rather, between the momenta) of impacting bodies. Its Geometry is therefore absolute, that is, independent of the parallel postulate and therefore including Euclidean and hyperbolic geometry. Thus, if Carnot's mechanics had been extended to Relativity by restricting the generality of its absolute space to hyperbolic velocity space, it would have been possible to avoid the conceptual trauma of the transition from Newtonian Mechanics to Relativity formulated in the complicated Minkowski space.[12]

Finally, we note that we have obtained this formulation with a simple mathematical formalism (with the exception of the above two first order differential equations first order, which however can also be considered as simple finite difference equations). We can conclude that the whole theory, here reformulated as an PO theory, is certainly within PI mathematics.

The "revolutionary" novelty of Einstein's article did not therefore consist so much in the concept of space-time or even in the new concept of simultaneity, as in having introduced, through a new theory, the foundational choices that were alternatives to those of Newton's Mechanics and having expressed them almost clearly.

---

12 A comparison between the treatment of relativistic impact and that of L. Carnot in F. Scarpa, "Lazare Carnot e la Relatività ristretta", *Il Giornale di Fisica*, **43** (2002) pp. 205-212.

CHAPTER 5

NEW HISTORY OF THE FOUNDATIONS OF MODERN PHYSICS: QUANTUM MECHANICS

5.1 Einstein's "most revolutionary" scientific paper: the 1905 article on quanta

Einstein dedicated considerable attention to the foundations of physical theories; in 1924 he wrote: «From the earliest years all my scientific efforts have been directed at understanding the foundations of Physics»[1]. In particular, before Einstein the general relationship between Mathematics and Physics was ignored by theoretical physicists, except for brief considerations by Mach and Poincaré. The main subject of Einstein's paper is the problem of the relationship between physics and mathematics.

In the introduction to the paper[2], rethinking the previous physical theories, he emphasizes that in theoretical physics there is a conflict in the use of two different types of mathematical formalisms, namely "continuous spatial functions" (which allow the "subdivision" [for example of a wave] in arbitrarily small parts ") and a "finite mathematics" to describe discrete material points (for example, in the theory of gases):

> A profound formal distinction exists between the theoretical concepts which physicists have formed regarding gases and other ponderable bodies and the Maxwellian theory of electromagnetic processes in so-called empty space. While we consider the state of a body to be completely determined by the positions and velocities of a very large, yet finite, number of atoms and electrons, we make use of continuous spatial functions to describe the electromagnetic state of a given volume, and a finite number of parameters cannot be regarded as sufficient for the complete determination of such a state. According to the Maxwellian theory, [for instance, the] energy is to be considered a continuous spatial function in the case of all purely electromagnetic phenomena including light, while the energy of a ponderable object should, according to the present conceptions of physicists, be represented as a sum carried over [a finite number of] the atoms and electrons. [Indeed] The energy of a ponderable body cannot be subdivided into arbitrarily many or arbitrarily small parts, while the energy of a beam of light from a point source (according to the Maxwellian theory of light or, more generally, according to any wave theory) is continuously spread over an ever increasing volume.

In this passage Einstein evaluates discreteness not as a property of a physical system, i.e. particles rather than waves, but as a characteristic of Mathematics relating to the foundations of theoretical physics. He does not consider discrete mathematics instrumentally or provisionally, with a view to arriving inevitably at a continuous mathematics. Instead, he is dealing with the fundamental problem of how one can choose between the discrete and the continuous in theoretical physics.

His hypothesis of a discrete electromagnetism was an "absolutely heretical proposal"[3], which he therefore presented timidly:

---

[1] Quoted by Klein M.J. (1980), "No firm foundations: Einstein and the early quantum theory", in H. Woolf (ed.), *Some Strangeness in the Proportion. A Centennal Symposium to Celebrate the Achievements of Albert Einstein*, Addison-Wesley, Reading, 161-185, p. 167.

[2] Einstein A. (1905), "Ueber einen die Erzeugung der Verwandlung des Lichtes betreffenden heuristisch Gesichtpunkt", *Ann. der Physik*, 17, 132-148; reprinted in Stachel J. (ed.) (1989), *Collected Papers of Albert Einstein*, Princeton U.P., Princeton, vol. 2, 149-165.

[3] Klein M.J., (1980), op cit., p. 167.

> The wave theory of light, which operates with continuous spatial functions, has worked well in the representation of purely optical phenomena and will probably never be replaced by another theory. It should be kept in mind, however, that optical observations refer to time averages rather than instantaneous values. In spite of the complete experimental confirmation of the theory as applied to diffraction, reflection, refraction, dispersion, etc., it is still conceivable that the theory of light which operates with continuous spatial functions may lead with experience to contradictions when it is applied to the phenomena of emission and transformation of light [where discreteness is at play].

What are the consequences of this "profound formal distinction" concerning the use of two different mathematical formalisms in theoretical physics? Einstein's words above lead one to suspect that there are some electromagnetic phenomena - of interaction of radiation with matter - that cannot be explained by means of continuum mathematics; some basic physical concepts, therefore, may be conceived as discrete (energy quanta). Moreover, Einstein suggests as an alternative a theoretical point of view - that is to say a discontinuous Electromagnetism - capable of interpreting certain new phenomena:

> It seems to me that the observations associated with blackbody radiation, fluorescence, the production of cathode rays by ultraviolet light, and other related phenomena connected with the emission or transformation of light are more readily understood if one assumes that the energy of light is discontinuously distributed in space. In accordance with the assumption to be considered here, the energy of a light ray spreading out from a point source is not continuously distributed over an increasing space but consists of a finite number of energy quanta which are localized at points in space, which move without dividing, and which can only be produced and absorbed as complete units.

In paragraph 1 Einstein begins by posing a problem: is continuous electromagnetic theory valid for all phenomena? He responds by applying it to a particular physical situation: a gas of electrons and molecules plus radiation, all interacting with each other. His description by means of Maxwell's continuous formulae of electromagnetism leads to a surprising mathematical result. The approach that aims to unite Maxwell's Electromagnetism, which has a continuous mathematics, with the energy equipartition theorem, based on the discrete schematization of statistical mechanics, leads to a divergence in the calculations (an integral whose value is infinite) and therefore to "an undefined energy distribution", which is absurd, in both experimental and theoretical terms. The answer to the above question is therefore negative.

To discover a discrete theory of electromagnetic phenomena, he then turns to black-body radiation; previously Planck had found a formula that was valid experimentally, but it had been based on the unprecedented idea of energy quanta, which, however, he considered only formally. Einstein, on the other hand, wants to reason directly about experimental facts. He shows first of all that it is possible to deduce a typical fact of discrete reality, the Avogadro number, 6.10^{23}, with Planck's formula alone, which was based on that quantum idea. So he can conclude that Planck's approach is theoretically valid; it remains to consolidate it with experimental facts.

At this point he uses thermodynamics and statistical mechanics to theorize about radiation. His arguments lead to an analogy (which, remember, is a double negation: "It is not true that it is not ..."): the energy of radiation behaves, even at the microphysical level, in accordance with the law $S = S(V)$ of an ideal gas. From this analogy Einstein proceeds to the affirmative hypothesis according to which light is composed of independent quanta of energy.

From this affirmative predicate he makes some deductions in order to compare them with the experimental data relating to three effects, of which the first is the photoelectric effect; he for the first time obtains precise interpretations of these phenomena. Therefore, his work gives to the concept of quantum, considered in Planck's writings to be a mathematical hypothesis only, the character of a directly verifiable physical concept.

The article was indeed revolutionary, a shock to the scientific community, as it provided a solid, directly experimental, basis for a physical concept that involved a revolutionary change in the whole of theoretical physics.

The aforementioned historian of Physics observed that since then

> [The] existence of this distinction [in the type of mathematics] ... marked a fundamental inhomogeneity in the foundations of physics.... [4]
> [Or rather,] a dichotomy in the foundations of Physics ...[5]

Note that the word 'dichotomy' suggests a dilemma existing between two different approaches to the mathematical description of nature. In other words, Einstein explicitly showed that in theoretical physics as a whole there is a dichotomy regarding the type of mathematics that a theorist should follow in constructing his theory.

We may ask whether Einstein's writing is consistent with his choice of “discrete” mathematics. To answer one has to examine his presentation with today's constructive mathematics, that is, according to the strict PI choice. In fact, his calculations contain derivatives and integrals on uniformly continuous functions, such as those of thermodynamics, thus conforming to constructive mathematics. However in paragraph 4 of the article (to obtain the function $S = S(V)$ for the radiation) he calculates a maximum on a set of functions by means of an exact equality with 0, which, as we know, is not constructive. However, the first translator of the paper into English, the well-known theoretical physicist Pais, already felt the need to suggest, for reasons of "simplicity", an alternative reasoning of a thermodynamic nature[6] that is completely constructive. This indicates that the constructiveness of mathematics is not an abstraction for theoretical physics, but is suggested by the operationality of physics itself.

In conclusion, Einstein, *even if unaware of constructive mathematics, conformed his calculations to this type of Mathematics, provided that one replaces a technique of Einstein’s with the one suggested by Pais*.

Yet in those years

> "... very few were willing to follow him in accepting the startling idea of light quanta on the strength of deductions that were based on the [dubious] statistical interpretation of the second law of thermodynamics.”[7]

In 1917, Millikan's paper reporting the experiments he had carried out on the photoelectric effect, which definitely confirmed the quantum hypothesis, added:

> “I shall not attempt to present the basis for such an assumption, for, as a matter of fact, it had almost none at that time.”[8]

---

[4] Klein M.J. (1970), *Paul Ehrenfest. The Making of a Theoretical Physicist*, North-Holland, Amsterdam, p. 241.
[5] Klein M.J. (1980), op. cit., p. 167.
[6] Pais A. (1979), “Einstein and Quantum theory”, *Rev. Mod. Phys.*, 41, 863-913, p. 873 col. I. He suggests a typical textbook computation for obtaining $S=S(V)$ for a perfect gas; by operating on uniformly continuous functions, his differential operations are all constructive.
[7] Klein M.J. (1963), “Einstein’s First Paper on Quanta”, *The Natural Philospher*, Blainsdell Co., New York, vol. 2, 59-86, p. 61.
[8] Millikan R.A. (1917), *The Electron*, Chicago U.P. Chicago, 238.

However, though slowly,

> "it was primarily the photoelectric effect [as interpreted by Einstein's paper] to which physicists referred as an irrefutable demonstration of the existence of photons and which thus played an important part in the conceptual development of quantum mechanics."[9]

Let us now consider the type of organization of Einstein's paper. In sect. 2.21 it was recalled that Einstein, together with his contemporaries Poincaré and Lorentz, indicated a dichotomy regarding the type of organization of a physical theory. Although neither he nor the others were able to characterize it formally, he, wanting to frame a theory to solve a problem, had to reason (regardless of whether he was aware of it or not) not according to an axiomatic-deductive organization but according to the model of a PO theory.

Indeed the title of the paper already characterized his theory as "heuristic", by which is commonly understood the absence of principles-axioms from which deductions are derived. In fact he demonstrates the existence of light quanta without using the formal apparatus of his previous publications on statistical mechanics and Maxwell's electromagnetic theory. The point is illustrated by a philosopher of science:

> The structure of this paper of Einstein's does not fit in easily with popular twentieth-century conceptions of scientific method. The traditional way of forcing it into the hypothetical-deductive mould is to avoid any mention of the first two-thirds of the paper…[leaving] the impression that Einstein merely postulated the existence of photon in order to explain the photo-electric effect[10].

Some historians have also questioned the type of logical reasoning of the text:

> Commentators often characterize the arguments of the paper as creating an analogy [read: a theoretically little relevant result] between radiation and classical ideal gas of material particles. But to Einstein, the connection was more than a weak argument, as is an analogy; it was a "far reaching formal relationship"[11].
>
> "But [asks the historian Klein] how seriously was one to take this conclusion? Did it really amount to anything more than an analogy, with the "as if" the essential phrase in its proposition?" According to him, "The conclusion [of the first part of the paper] … was to be taken seriously, and Einstein immediately exploited this "suggestion" as to the nature of radiation, tenuous as it might (and did) seem to others, pressing it in directions that might yield experimentally verifiable consequences." [12]

But, what "wide-ranging formal relationship" could Einstein's "proposal" be? The question raises the more general question of the model of organization of a scientific theory. Careful reading of the text shows a different organization, the PO type, from the deductive one. Indeed, half-way through the "Introduction" a fundamental problem is presented: is the continuous electromagnetic theory valid in all physical situations? Can the mathematics of this consolidated theory lead to contradictions?

Furthermore, in only five pages of the first methodological part, 55 doubly negated propositions (DPNs) can be recognized and it can be shown that their sequence is in itself sufficient to provide the logical thread of Einstein's presentation.

An even more important verification that it is an PO theory is the use of the only kind of proof that is appropriate for the conclusion of an argument based on PDN, i.e. a *reductio ad*

---

[9] Jammer M. (1966), *The conceptual development of Quantum Mechanics*, Mc Graw-Hill, New York, p. 33.
[10] Dorling J. (1971), "Einstein's introduction to photons: Argument by analogy or deduction from phenomena?", *Brit. J. Phil. Sci*., 22, 1-8, p. 7.op. cit., p. 6.
[11] Klevard P.A. (2008), *A Fresh Approch to Quantum Mechanics and Relativity*, Aventine P., p. 14.
[12] Klein M.J. (1963), op. cit., 70-71.

absurdum proof. We find it in two cases in which it is concluded that the failure of Maxwell's theory is not impossible in a particular case: when light is treated using discrete mathematics.

Furthermore we must observe that Einstein reasoned about the microscopic world, unknown at that time and about which it was quite possible to imagine absurd things, if referred to the macroscopic world, so that he could not use this kind of demonstration more than twice. Therefore the universal theoretical conclusion is not deduced from a *reductio ad absurdum* proof, but from an analogy, which in fact summarizes a DNP: "It is not true that it is not ..." (also later in laboriously constructing the theory of quantum mechanics, only analogies made it possible to introduce into the theoretical physics the innovations concerning the new microscopic world, for example the Bohr atom, analogous to the solar system, or Dirac's analogy for introducing the quantization of the Hamiltonian). Thus Einstein could not have suggested a more precise separation from deductive reasoning than his reasoning by means of intuitionistic logic.

In conclusion, all the above facts show that *Einstein not only chose an alternative model of organization of his theory, but also adapted his inductive reasoning to the non-classical reasoning typical of a PO theory.*

Thus, of all the writings of Einstein the article of 1905 on quanta was *the "most revolutionary" because in it he explicitly expressed the dichotomy concerning the type of Mathematics; in particular he conceived of discreteness (actualizing PI) as an alternative to traditional continuity (actualizing AI); moreover with respect to the other dichotomy, he chose the alternative organization (PO), giving it almost full expression. These two choices were alternatives to those of Newton's Mechanics.*

Two studies by the historian Klein intuited this revolutionary nature of Einstein's article:

> Einstein was well aware that all of this marked the beginning of a new era in physics, and he indicated that awareness by referring to his work in the title of his paper as offering "a heuristic viewpoint". He saw that thoroughgoing changes in the foundations were needed…[13]
>
> His deepest concern, expressed in the opening sentences of his paper, was the very foundation of science.[14]

The episode of the birth of quanta in Einstein's paper calls for deep reflection

We have seen in the above that of classical physical theories the only one that declared its foundational choices was that of Lazare Carnot. This theory was, however, obscured and all the other formulations that differed from classical mechanics, that might have suggested the choices by mutual comparison, were treated as its specific variants. On the other hand, the emergence of Einstein's two theories, which are of much greater importance than the introduction or modification of some concepts (quanta, space, time, etc.), posed with unprecedented clarity the foundational problems of the choices. In fact the previous physical theories (L. Carnot's mechanics and thermodynamics) are unsatisfactory, the first because of a confused use of non-classical logic, making it necessary to interpret and also reconstruct the development of its Mathematics, the second because it was based on a subsequently discarded hypothesis (caloric). Moreover, Einstein's other theory, special relativity, only

---

[13] Klein M.J. (1961), "Max Planck and the Beginning of Quantum Theory", *Arch. Hist. Exact Sci.*, 1, 459-479, p. 477.

[14] Klein M.J. (1979), "Einstein and the development of quantum physics", in French A.P. (ed.) *Einstein. A Centenary Volume*, Harvard U.P., Cambridge, 133-151, p. 135.

partially conformed (in the first pages) to a PO model and the use of constructive mathematics. Thus in the history of physics the paper on quanta presented for the first time a valid theory that is, in a substantially adequate manner, based on the two alternative choices.

Unfortunately the crisis of the early 20th century in the Foundations of Physics was so profound that theoretical physicists lacked the tools to solve the fundamental problems of the crisis and put the construction of new theories on a sound footing. They needed new notions (which they laboriously constructed), new techniques (transformation groups emerged in theoretical physics in the ‘60s) and above all the formalizations of the two alternative choices, constructive mathematics (born in the ‘60s) and intuitionistic logic (whose recognition as a respectable logic came equally late).

No surprise if, after presenting his "revolutionary" theory, which to some historians today seems "miraculous", Einstein did not persist in exhibiting the fundamental choices: he lacked the suitable formal tools to deal with the new problems directly. For this reason he did not found his further theory (General Relativity) operationally nor did he devote attention to it, when in the 1920s some mathematicians - including H. Weyl, who was well known to Einstein – began to define constructive mathematics.

5.2 Chronology of the birth of quantum mechanics

> It was unfortunate that the systematic development of quantum mechanics had to deal first with the problem of the quantization of the energy of electromagnetic harmonic vibrations. There can be no doubt that the first quantum theorists would have saved themselves much intellectual effort ... if there had been a conceptually less convoluted theme at the beginnings of the theory ... But *post jacturam quis non sapit?* [15]

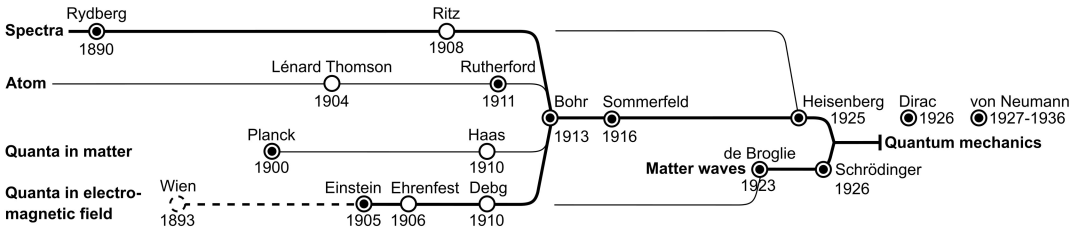
The history of quantum mechanics is complex. The fundamental cause of this complexity is the intellectual conversion that the physicists had to undergo in order to accept a sequence of concepts that were radically different from the classical ones. There was therefore no linear development, but an accumulation of new and shocking facts, which were only later clarified by a unitary overall picture. The theoretical upheaval had frustrated many physicists, who began to believe either that a limit to human knowledge had been reached (W.H. Bragg wrote "Physicists use the wave theory of light on odd days and the corpuscular on even days"), or that it was impossible find a physical solution to its apparently contradictory problems. One example of this frustration is Schroedinger, who at the age of 35, shortly before arriving at his famous equation, was so discouraged that he was considering retiring. This period of frustration lasted two decades.

Without going into details, here we give the fundamental stages of the emergence of quantum mechanics, first with a graph showing the main lines of development of the history of its birth and then with a table that presents the sequence of its main authors and their respective advances.

---

[15] M. Jammer, *The Conceptual Development of Quantum Mechanics*, New York: Mc Graw-Hill, 1966, p. 42.

GRAPH OF THE EVOLUTIONARY LINES OF THE BIRTH OF QUANTUM MECHANICS



This graph highlights a series of facts that indicate the complexity of this history.

1) the accumulation of discoveries follows multiple lines of development, which only eventually converge in a single theory;

2) the development was not linear, - for example Rutherford's model of the atom as analogous to the solar system was very important since it provided an initial idea of the structure of the atom (still used by chemists), but was completely misleading for subsequent research, both because the uncertainty principle prohibits drawing trajectories (which would indicate p and q together with absolute precision), and because the planetary system does not have the stability of real atoms (electrons in circular accelerated motion do not radiate, which would result in their losing energy, causing them to fall almost immediately onto the nucleus).

3) the final theory was preceded by two completely different syntheses: the Mechanics of the Heisenberg matrix and Schroedinger's Wave Mechanics. Only then was their equivalence "proved", starting from the harmonic oscillator alone, after which, with Dirac and von Neumann, its most general formulation was demonstrated in Hilbert space;

4) it was therefore only after several theoretical steps, including a general riformulation of the initial theory, that a definitive conceptual and mathematical framework of this theory was established;

5) the price of this new framework, however, was very high: a large mathematical formalism (a space of square-integrable functions) and a considerable abstraction from reality (the fundamental physical quantity is the amplitude of probability expressed in the field of imaginary numbers; Einstein and other researchers never accepted quantum mechanics: "God does not play dice", i.e. theoretical physics, which directly concerns reality, cannot be based on abstract probabilistic concepts).

It should however be noted that quantum mechanics is certainly the best verified physical theory ever to have been conceived in the History of Physics; there is no experiment that can disprove it. Recently, subtle theoretical clarifications, such as Bell's inequalities, have further led to confirm this theory with experimental tests, contradicting suggestions by Einstein and others that hidden variables should be sought that might restore the theory to the determinism that we find in classical mechanics.

TABLE 5.1

SEQUENCE OF MAIN AUTHORS AND THEIR RESPECTIVE ADVANCES

1890	Rydberg	Formula for atomic spectra
1900	Planck	Black body formula: the quantum hypothesis
1905	Einstein	Formula for the photoelectric effect and for the specific heat of

		solids: quanta are real
1911	Rutherford	Planetary model of the atom
1913	Bohr	Interpretation of atomic spectra. Principle of correspondence.
1916	Sommerfeld	Quantization of periodic motions. Stability of the atom
1923	Compton	Corpuscular effect of waves
1924	De Broglie Born Heisenberg Slater	Wavelength relationship with momentum Probability waves
1925	Heisenberg-Born-Jordan	Matrix mechanics (only observables, no trajectory)
1926	Schroedinger Dirac	Wave equation (materials): interpretation of the hydrogen atom. Equivalence of its wave mechanics with matrix mechanics Algebraic commutation relationships
1927	Heisenberg All, apart from Einstein and some others.	Uncertainty principle Copenhagen interpretation
1927-32	Von Neumann	Mathematical formulation in Hilbert space. Paradox of quantum mechanics
1936	Birkhoff e von Neumann	The logic of Quantum Mechanics is non-classical

5.3 Which choices are made by quantum mechanics?

It can be argued that quantum mechanics is conceptually twofold, in that it unites two aspects of reality: that of corpuscles, traditionally interpreted mechanically à la Newton (AI and AO) and that of waves, interpreted à la Huygens, through the construction of the wave front and by involution of the circles of radius $v\Delta t$ of the wave front; (PI and PO). Because of this duality, the theoretical structure of quantum mechanics is essentially new with respect to classical physics; and the interpretative method of the two dichotomies seems to underline it.

But this concerns only the conceptual aspect of this theory. Let us go even deeper and try to identify the choices of the new theory. Does quantum mechanics choose potential infinity (PI)? A naive answer is yes, because it is based on quanta and many of its variables are discrete. However it also uses continuous variables. One might think that they can be considered in the PI continuum: yet Hilbert's space certainly makes use of AI without any limitation.

From the history of quantum theory we know that this ambiguity regarding infinity was manifested by the first two clearly distinct formulations: that of Heisenberg's algebraic "matrix Mechanics" and that of the "Wave Mechanics" of Schrödinger's differential equation. Today it is generally believed that this discrete - continuous contrast, underlined by the two original formulations, has been resolved with Dirac's function δ (which is a function $f(x)$ null on the whole of the x-axis, except for one point where it is positive but the value of the integral of this function on the whole of the axis, is equal to 1) and with concepts that certainly include AI. Ultimately, the naive affirmative answer seems to be contradicted by the theoretical developments of the theory.

Let us now consider the other question: does quantum mechanics choose PO? Once again there is a naive answer: yes. Indeed, quantum mechanics is essentially new with respect to classical theories, in that it affirms the <u>impossibility</u> of measuring two conjugate variables simultaneously with non-relative precision. It therefore has a fundamental problem wrongly called "principle" (Heisenberg's uncertainty). This is not an axiom, since it is not at all self-evident, but rather a methodological principle. It is because of this that we introduce the probability amplitude ψ and that we construct the theory (or method) that ultimately limits the measurement of the conjugate quantities, no longer admitting absolute precision for both.

If, however, we look into the question more deeply, we note that the answers given above concern only a part of the theory: its connection with experiments, while the remainder is constructed on Schrödinger's differential equation, (the only one taught at university). This equation is posited as a self-evident axiom-principle on the strength of the great tradition accumulated by mathematical physics, based on differential equations. In fact Schrödinger's differential equation generates deductively all the statements of quantum mechanics (for the simplest cases); this is all the more evident when one uses the Hilbert space formalism, whose choice is AO. We must conclude that quantum mechanics is also ambiguous in its organization of the theory.

On the other hand, it can be argued that Wave Mechanics makes the opposite choices to those of Newtonian theory: the probability amplitude of ψ has nothing of the self-evidence typical of principles (indeed the wave-particle dualism has removed quantum mechanics from all intuitive self-evidence). It is true that ψ leads to the deducibility of the theory, a typical characteristic of axiomatic organization (AO), but only as a mathematical technique for the solution of differential equations in the fairly restricted cases of a mathematically "tractable" potential. Furthermore, the Mechanics of waves, while using actual infinity (AI), must always include some essentially discrete variables in its mathematical formalism, and furthermore, it must impose some mathematically unclear limitations on its functions which are (restricted to square-integrable functions under self-adjoint operators; the set of functions thus defined is not invariant under differential operations.) Restrictions on the set of functions are chosen to obviate the obvious fact: classical AI mathematics is too powerful for the practice of theoretical physicists dealing with discrete quanta.

In conclusion, the two fundamental choices remain undecided.

Most theoretical physicists would agree with this assessment. Von Neumann pointed out, objectively and mathematically, a similar contrast between two parts of the theory: the Schroedinger equation, essentially reversible and linear, which accounts for the undisturbed temporal evolution of the system and the "collapse" of this representation of Quantum mechanics caused by measurement operations (quantum jumps, wave packet reduction), typically irreversible and non-linear. In fact, it is clear that the aforementioned branches of Quantum mechanics have opposite choices: one AI and AO, the other PI and PO. Since, as discussed above, these two branches are based on respectively two incompatible fundamental pairs of choices these two parts of the theory are not reconcilable. In fact they are only technically linked, although there is a clear logical division between measurement and theory (precisely as we have seen with regard to the first principle of Thermodynamics).

Current quantum Mechanics is therefore built on the fundamental theoretical contrast that was indicated by von Neumann and here it has been underlined through the diversity of the fundamental choices of its two main parts as well as in the historical fact that its first formulation, which we owe to Heisenberg and Born, was based on the PI-PO choices, the said formulation being based on the problem (PO) of the incommensurability of the measurements of conjugate quantities and using finite matrices (PI). But this formulation encountered a difficulty: von Neumann showed that, in some cases where it is necessary to use infinite matrices, they do not correspond well to unlimited operators.

The problem of a general refounding of quantum mechanics was addressed by one of the major scientists of the twentieth century: Hermann Weyl. In 1918, he planned to found mathematics anew using AI as little as possible (only one quantifier on decidable numbers). Although he failed to complete his programme, he consistently adhered to this "elementary" mathematics, and in 1928 he published the first textbook on quantum mechanics, which proposed the refounding of all modern physics (relativity and quantum mechanics through a single mathematical formalism based on *transformation groups*, introduced by him into theoretical physics thanks to the techniques of his elementary mathematics. However, his programme was blocked by Neumann's aforementioned proof.

We thus see that in the history of quantum mechanics there was a spontaneous theoretical effort to formulate quantum mechanics according to the choices PI and PO, even if it did not arrive at a complete alternative formulation. With today's knowledge of PI mathematics, the attempt could be reconsidered. But in these notes we cannot deal with theoretical problems of this nature. It will suffice to note that the great physicist-mathematician Weyl came close to an innovative PI-PO formulation.

5.4 Non-classical logic, double negations and organization of the theory

Now we will clarify a basic problem: the PO choice and the kind of logic that belongs to quantum Mechanics. The last line of table 5.1 recalls a very important result that was accepted fifteen years after its discovery (and is culpably ignored in universities): in 1936 Birkhoff and von Neumann showed that the logic of quantum Mechanics is not classical logic, which suggests that the teaching of physics has ignored the link between physics and logic (as important as the link between physics and mathematics), to the point that logic is not taught and it is taken for granted that physicists inevitably reason only using classical logic.

The above discovery of non-classical logic in quantum mechanics can be explained with an example. The uncertainty principle says that two conjugated quantities (such as position and momentum) of the same system cannot be measured at the same time with absolute precision, but with the precision given at most by $\Delta p \Delta q = h/2\pi$ where Δ is the standard deviation, q the position and p the momentum, h the Planck constant. Let us suppose we have measured the position q of an electron as located in the interval $(-L, +L)$ and that we then measure p with the greatest possible precision, that is, so as to satisfy the equality above. Δp will therefore be equal to $h/2\pi 2L$. We can translate this fact into the language of propositional calculus, using the two propositions:

A={the electron has position q between $-L$ e $+L$} and

B={the electron has a momentum p measured with uncertainty $h/2\pi 2L$}.

We are now saying that $A\wedge B$ holds. But we can subdivide the interval ($-L,+L$) into two parts, ($-L$,0) and (0, $+L$) and therefore consider the proposition A as composed of two:

A'={ the electron has position q between $-L$ and 0} and

A''={ the electron has position q between 0 and $+L$)}.

Thus the same proposition as before $A\wedge B$ becomes $(A'\vee A'')\wedge B$.

Now in classical propositional calculus the law of distributivity applies; therefore from the last formula we can obtain $(A'\wedge B)\ \vee\ (A''\wedge B)$. In quantum Mechanics, however, this compound proposition has no physical sense, because each of the two propositions, $(A'\wedge B)$ and $(A''\wedge B)$, does not satisfy the principle of uncertainty, since for each one we have $\Delta p\Delta q=L\ h/2\pi 2L=\ h/2\pi 2$ which is less than the limit $h/2\pi$!

From this point of view quantum mechanics is the first theory to force us to recognize a non-classical logic in theoretical physics.

However, the study by Birkhoff and von Neumann does not directly suggest a new logical structure. What is obtained by them must be incorporated into an algebraic structure, which can vary with a certain margin of discretion. Many studies have appeared, but after 80 years they have not led to decisive progress (often this field of study is referred to as a "labyrinth").

In establishing which logical law is violated by quantum mechanics, distributivity is usually indicated, although several authors disagree, arguing that it is the law of double negation. This can be maintained on the basis of what has been presented so far regarding the PO theories.

Let us ask how a structured theory proceeds using doubly negated propositions. From a comparative examination of theories of this type we see that an argument composed of doubly negated sentences can only be concluded with a *reductio ad absurdum* proof, that states $\neg\neg\ T$. These groups of sentences form a unit of reasoning, whose conclusion works as the methodological principle for the next unit. The last of the n reasoning cycles demonstrates a universal thesis $\neg\neg\ UT$ (as for example the universal proposition of S. Carnot's theorem: "It is <u>not</u> true that the efficiency of reversible thermal machines is <u>less</u> than that of irreversible thermal machines"; this S. Carnot's theorem holds for *all* thermal machines. Lobachevsky also reasons in this way, when he proves that the new hypothesis of two parallels holds for all lines and triangles of space). At this point all possible evidence has been accumulated to translate the merely plausible conclusion, $\neg\neg\ UT$, into the statement UT. As do both S. Carnot and Lobachevsky after deriving the universal thesis: S. Carnot reasons as if the efficiency of thermal machines was the maximum; and Lobachevsky reasons as if the hypothesis of two parallels were valid.

In the 1905 paper on quanta Einstein, starting from an absurdity (classical Electromagnetism applied to a special case leads to an infinity), reasoned with doubly negated sentences and then concludes with an analogy, which constitutes a double negation ("It is <u>not</u>

true that it is not ...). Here he substantially followed the pattern of non-classical logic that holds in a PO.

If we apply this schema to Einstein's original 1905 paper on Relativity, we note that he correctly poses problems through doubly negated sentences but does not manage to formulate a *reductio ad absurdum* proof. Indeed the paper contains errors, both in logic and mathematics (it has already been said that he sums v with c precisely at the point where, after the discussion on simultaneity, the *reductio ad absurdum* proof should begin).

If we apply this kind of development to quantum Mechanics we should discover an alternative formulation already established in PO and hence using doubly negated sentences, or at least one paper by the founders of that period that corresponds to this type of formulation. They do not seem to exist, so in the following we will proceed by reconstructing the theory independently.

5.5 Problem-based organization and non-classical logic in quantum theory

Let us therefore see whether quantum mechanics can also be formulated according to a PO that uses non-classical logic through doubly negated sentences.
First we recognize its problem: it is clearly that of measurability. This problem can be expressed with a sentence that is doubly negated, as is the case in every PO. In fact we can correctly say that "It is impossible to measure p and q simultaneously with absolute precision", where p and q are two conjugate variables. (The complementarity between the intuitive descriptions of quantum phenomena can also be formulated with a doubly negated sentence: "It is not true that a wave is not a corpuscle; and vice versa"). To overcome this problem, the PO theory proposes a new experimental and mathematical method, with which to restore a physical reality, at first sight seemingly unknowable, in which measurements and calculations with the maximum possible precision are possible. Let us therefore translate this into a sentence: " It is not true that p and q cannot be measured simultaneously with relative precision".

An interpretation of quantum mechanics of this type posits the indeterminacy relations as fundamental. The other characteristic aspect, the evolution of the unperturbed system, does indeed belong to the untestable world. With this interpretation we choose to emphasize the experimental part of quantum mechanics, as well as the clear gulf between quantum and classical phenomena. This clarifies why quanta necessitated a new type of theory: it is a question of describing phenomena that are unknowable using classical methods.

It is useful to adopt this type of theoretical organization if only because we are dealing with a problem on the border between Physics and Metaphysics: in quantum mechanics we need to understand how to make a measurement when the system is inevitably influenced by the operation of measurement itself, which consequently is no longer able to provide a direct correspondence between theory and reality. Given the inadequacy of classical experimental solutions, the theory must establish a new scientific method. This great problem, and hence the necessity of a PO theory explains the inevitability of Birkhoff and von Neumann's discovery in 1936 of a non-classical logic in quantum Mechanics. However, up to now this result did not lead to identifying a particular non-classical logical theory; from our point of view here, it alludes only to intuitionist logic, which is specific to PO theory.

It can also be explained why non-classical logic has become essential in quantum mechanics. If in Thermodynamics we consider the doubly negated statement that poses the main problem, we note that it concerns two quantities, work and heat, which do not define the state of the system. Therefore the theory, which develops from the concept of state (that is, in *P, V, T* and *S*), does not contain in its descriptive structure the element of non-classical logic constituted by the doubly negated proposition concerning work and heat ("It is not true that heat is not work"). Thus the logic of the central part of the thermodynamic theory may still be classical. If we then consider classical mechanics and base it on the doubly negated sentence of the principle of inertia ("It is not true that the state of rest is not equal to the state of motion of the system") we note that this concerns only one state variable, velocity v. Now, every physical variable is already subject to the imprecision of measurement and therefore originally involves a doubly negated statement $\neg\neg A$: "It is not true that $x = 0$ is not equal to $x \neq 0$". In traditional theoretical physics the problem has remained obscure because, inter alia, of the use of infinitesimals, which blurred the distinction between equality and inequality. From 1870, infinitesimals were eliminated and a "rigorous" theory of real numbers was given, but even in this new theory of continuum the precision of measurements was idealized metaphysically as infinite. Therefore it was thought that ideally it was always possible to decide whether $x = 0$ or not (leaving it to experimental physicists the task of squaring this ideality with experimental results). Thus theoretical physicists have operated under the assumption, not of the previous doubly negated proposition, $\neg\neg A$, but of the idealizing one, $\neg A = ((x = 0) \neq (x \neq 0))$, which claims to be able to determine with precision the approximations to zero or not, in order to be able to speak of π and other irrationals as if they were experimentally ascertainable numbers. Thus theoretical physicists also conceived the values null or constant of concept that constitutes the essential part of the 1st principle, velocity. In this way the (Newtonian) theory also appeared to be independent of non-classical logic.

In quantum theory, on the other hand, the doubly negated statement concerns (both variables of) the description of the state of the system. That is, it is inherent in the theory's ability to describe reality. Thus, the theory of quantum mechanics includes double negation as its essential element; which makes the logic of the whole theory non-classical. In other words, in the other theories the doubly negated sentences could be placed in a lateral position with respect to the part of the theory that could be deductively reformulated, or they could be concealed with a minimum of idealism. On the other hand, in quantum mechanics a sentence of this type concerns precisely the relation between the measurements of the crucial theoretical variables, which define the state of the system.

Ultimately we recognize that the formulation of Quantum Mechanics corresponding to the PO is Heisenberg's formulation, in that it is based on the uncertainty principle, understood as the first methodological principle. We could proceed with the reconstruction of the theory according to the PO (reasoning cycles and final universal thesis), but we choose to stop here to avoid treating groups of symmetries according to constructive mathematics, which would greatly complicate this study.

What we have seen so far is enough to make clear that quantum Mechanics came closer to the formal structure of the foundations of a scientific theory more than any other theory, but without going so far as to specify them. This point is seen clearly if we represent the growth of the theory with respect to its conceptual increments in five-year time periods as shown in the following graph. On the one hand, the graph shows the formidable collective intellectual progress that led to the formulation of quantum Mechanics starting from zero (the single concept of quantum). On the other, it shows that, after the first formulations made it possible to reduce all the problems to problems of mathematical calculation, the foundations of the theory eluded careful consideration.

Table 5.2: THE MAIN STEPS IN THE CONSTRUCTION OF QM ACCORDING TO THE TWO DICHOTOMIES

1930					Birkhoff e von Neumann: LNC von Neumann: Hilbert space AI AO	von Neumann: Paradox (Inc.)
1925			Schroedinger: Wave mechanics (AI and AO) Proof of equivalence	Copenhagen interpretation	Dirac: formalism AI Heisenberg: Mech. Matrices PI and PO	Weyl: Formu-lation using transformation groups, Inc
1920				De Broglie Compton: Wave (e.m.) and corpuscle (mech.): VRS		
1915				Bohr: principle of correspondence		
1910			Bohr: Model of the atom Rutherford: Model of the atom			
1905	Planck: Quanta in the black body VRS Einstein: Quanta in the photoel- effect. VRS					

1900		Planck: Quanta (VRS)				
	Concepts	***Formulae***	***Theories***	***Theories in conflict***	***Choices***	***MTS***

Legenda: AI= Infinity; IP= Potential Infinity; AO= Axiomatic Organization; PO= Problem-based Organization; NCL = non Classical logic; Inc= Incommensurability; RVM= Radical variation in meaning.
The authors were not fully aware of the fundamental aspects that are placed in brackets.

The novelty of non-classical logic in quantum mechanics appears even more relevant considering that even philosophers of science are not yet aware of its necessity. Let us take for example K. Popper. Although he studied non–classical logics intensively and although he expressed a new, now famous, philosophy using doubly negated sentences ("science is <u>fallible</u> [due to <u>negative</u> experiments]", "science is <u>falsifiable</u> [due to <u>negative</u> experiments]") he himself did not realize that he was doing so, perhaps because of the use of the abbreviations of those propositions. He did not therefore recognize the fundamental role of non-classical logic, but rather he stated several times that the only valid logic is classical.

5.6 Conclusions on the modernity of Physics

Owing to the profound crisis of the early years of 20th Century, theoretical physics were led to question the very foundations of physics and to consider some radical alternatives to the previous theories. But paradoxically from what we have seen regarding the two modern theories in question, the “modernity” of Physics consists first of all in a greater foundational obscurity, recognizable in each of the two theories of the first half of the 20th century; textbooks generally do not clarify their foundations (which are divergent), but tries to present them deductively Moreover, each of them, wanting to bring together all the previous theoretical experience of the past,is irreducible to the other: the Lorentz transformations cannot be applied to both indeterminacy relations, which are non-linear relations, and quantum leaps.

If we try to grasp the positive aspect of modern physics, then we see that it constitutes a great effort both to go beyond the level of mere technicalities to which the Physics of the late nineteenth century had descended, and to recover the foundations of Physics; that is, it introduces a self-reflexive relationship with reality. For example, in Relativity we have the essential concept of a frame so that the theory speaks of both the observer and the object observed; in quantum mechanics the uncertainty principle formalizes the influence of the observer on the object observed.

Even more profoundly, the two theories introduce fundamental novelties with respect to classical physics: they led to no longer consider AO and AI (of the mathematics of idealized continuum) as the only possibilities; at their inception the two theories, in fact, manifestly rejected them. Unfortunately these theories did not make manifest the alternative foundations: the possibility of basing physical theories on PO and on PI. After Einstein’s special relativity and theory of quanta and Heisenberg’s matrix mechanics the effort to find an alternative in the foundations was not followed through to completion (we have already mentioned this point for Relativity; for quantum mechanics see Tab. 5.2). In particular, quantum physicists were not been able to achieve an understanding of the conflict between the first two formulations of quantum mechanics and were satisfied with choosing the culturally least traumatic one, Schroedinger's. They subsequently dealt with all the questions abstractly from a mathematical point of view and with a considerable degree of formalism, oblivious to the level of abstraction they had introduced (to the extent that the contradiction in the fundamentals of the theory between the evolution of the undisturbed system and measurement was barely noticed). They did not therefore reach an independent clarification of the foundations of Physics, merely remaining in a position of rejection of the past (never very well characterized: the Newtonian paradigm, mechanism, determinism ...). Instead the two dichotomies clearly show the incommensurabilities of modern theories, or rather their intrinsic incompatibility that can no longer be concealed with summary judgments (like those adopted in classical physics: for example, calling L. Carnot's mechanics "technical" and Thermodynamics "phenomenological"). Despite their many and very important successive advances, physicists have not yet been able to recognize these incommensurabilities

Overall, modern physics represents a colossal effort to clarify the foundations of science, despite the contemporary crisis of the philosophy of knowledge and the lack of any similar cultural experience to refer to. Here we have tried to complete that cultural effort, availing ourselves of that research into the Foundations of Logic and the Foundations of Mathematics developed at the same time as the intellectual effort of physicists, to whose misfortune, however, the said research was completed decades later than the completion of their new theories: the two dichotomies were clearly formalized in mathematics in 1967 and in logic in the 1970s, i.e. several decades after the birth of both new physical theories.

CHAPTER 6

NEW HISTORIOGRAPHY OF PHYSICS ACCORDING TO THE TWO DICHOTOMIES

6.1 Koyré's and Kuhn's Histories of Physics

The history of science does not explain why physical theory came into being in the West[1]; nor why it occurred in the seventeenth century, after a very long preparation that can be traced back to the Greeks, if not to previous populations. Its emergence was the work of several geniuses (Descartes, Galilei, Newton) interacting (and even in conflict) with each other and with society.

Nor is the history of eighteenth-century science simple. It was then that mechanics incorporated more and more fields of phenomena, to the point of achieving a monopoly of physical theory. Then in the nineteenth century new theories emerged: Physical optics, Thermodynamics, Electromagnetism; all theories essentially different from Mechanics. And later statistical mechanics was also a considerable novelty, not referring to the concepts of absolute space, absolute time, force-cause and the differential equations of Newtonian mechanics. The result was a greater pluralism of theories. In the early 1900s, a major crisis shook the Foundations of traditional Theoretical Physics. It led to the emergence of two theories that differed completely from those of the past to the point that they divided modern from classical physics, but were also different from each other. The evolution of the history of physics, as shown in the previous chapters, is therefore complex and is also surprisingly characterized by traumatic events, which would not have been expected a priori of an enterprise assumed to be based only on experimental truths and on the most advanced rationality.

How then should the history of physics be written? Until about 1930 it was practically no more than an accumulation of data, a chronicle. Then, once enough historical descriptions had been accumulated and the facts and situations of the past made known, it began to interpret particular periods.

The main problem of the history of science is, of course, the birth of science itself. We are not referring so much to why Science did not arise in the East but in the West, or why it was developed by men and not by women. These are fascinating but too general problems compared to the first crucial problem: with which fundamental concepts did science emerge? The problem is in itself formidable, if only because of the duration of the period in which it took place (at least a century) and the many divergent interpretations of the works of the three main founding scientists, Descartes, Galilei, Newton. One very important answer is that of Alexander Koyré which we will examine below.

---

[1] I suggested an answer in the paper: "Il ruolo centrale di Nicola Cusano nella nascita della scienza moderna", in M. Toscano, G. Giannini, E. Giannetto (edd.), *Intorno a Galileo: La storia della fisica e il punto di svolta Galileiano*, Rimini: Guaraldi, 2012, 17-25.

Every historiography depends on a philosophy, as can easily be seen by reading any serious history book. The traditional history of science has distinguished the various philosophies into two groups: philosophies that explain the history of science through variables that are all internal to science itself (or, if it is preferred, to the mind of the scientist) and philosophies that emphasize the decisive role of some (or many) social factors. Clearly the position of science with respect to this division is suggested by the particular conception one has of the nature of the subject of study, science. Science is usually seen as the culmination of intellectual progress in recent centuries and, as such, as the activity capable, by itself, of projecting us into the future. This conception therefore considers science as an intellectual construct whose birth and historical evolution are not influenced by external conditioning, having its own internal logic of development, determined by its techniques and its methods (usually seen as objective and eternal), applied to the problems suggested by nature. Its history is therefore studied as if it were an absolute with respect to society.

This conception may be accused of being typical of an idealization of the nature of Science and of blindness to what links it to the culture and the social organization of the particular time in which it develops. Indeed, many maintain that not only over a very long time span, at least 500 years, science depends on society but also that its short-term evolution, with its single crises of growth and its individual successes, depends on the society of the time.

The last hundred years of Western philosophical life have been dominated by the conflict between idealist philosophy (neo-Kantians, neo-Hegelians, etc.) supporting internalism in the historiography of science and a social philosophy (in particular Marxism) that promoted externalism. It is therefore not surprising that historians of science have split into two distinct groups. Koyré clearly belongs to the internalist historians (and is accused by others of being an idealist).

Koyré was Russian, emigrated first to Germany and then to France in the early twentieth century. In Germany he followed the lessons of the philosopher E. Husserl[2], who knew mathematics well and had written about it and logic. His thinking was dominated by the importance of the concept of infinity in Science. Koyré was a humanist and became known for the first time for works of philosophy, but his love for Science led him to devote much of his life to the historical problem of its birth. Koyré followed Husserl; his first scientific essay concerns the paradoxes of infinity, whose importance had been shown to him by Husserl. It was precisely because he favoured an idea which then appeared to be philosophical, infinity, that his standpoint was classified as that of an internalist historian and it is actually true that he focused his attention exclusively on historical factors within Science and never to factors suggesting its dependence on the social context.

His method was clearly at odds with that of the externalists who in his time studied the same subject. The thesis of these historians is that the emergence of science is due to the artisans who collectively had produced so many inventions that cooperatively they were able

---

[2] He is the author of: *The crisis of the European sciences and transcendental phenomenology*, Evanston, IL: Northwestern Univ Pr, 1970. The first chapter is the only text of philosophy whose reading is useful for deeper analysis of our theme.

to achieve the intellectual leap from the philosophy of medieval science to the experimental science of the 1600s. The artisanal construction of boats with a perfectly hydrodynamic form or the manufacture of unique violins (Stradivari, Amati, etc.) would be proof of the almost unlimited capacity for improvement of this type of work. For these historians the crucial philosopher of reference of that time is Francis Bacon, while Leonardo da Vinci is emblematic of the quintessential scientist.

Koyré was of the opposite opinion, as expressed in a paper entitled *From the world of approximation to the universe of precision*[3]. In it he wanted to underline how any artisanal work, however advanced it may be, is always immersed in the approximation of measurements (both sensorial and instrumental). Measurements, remaining inaccurate even after innumerable improvements, could not have given rise to that qualitative leap that generated science, that leap to the perfection of calculations, concepts and theories that makes theoretical science an intellectual activity based on absolute accuracy. The birth of science for Koyré was inexplicable without this conceptual factor: *it is the idea of infinite accuracy, and therefore the idea of infinity, which Koyré believed gave rise to modern science*. This is well expressed by the title of his most famous book.

In chapter 3 we saw that the whole history of mechanics has depended on the type of mathematical infinity that physics has adopted. Mechanics was developed using a mathematics founded on actual infinity; but, after a century, it was reformulated (by L. Carnot) with a mathematics without this infinity. We have also seen that the historical development of the rest of Physics was also determined by the introduction, through mathematics, of either one or the other type of infinity. Koyré's idea is therefore confirmed and specified by the importance of the two types of infinity (PI and AI) for the history of the Physics subsequent to that studied by him.

Since the 1960s the contribution of T.S. Kuhn, a physicist who preferred to devote himself to the history of science as a young man, has become of great importance. He offered a completely new vision, also with respect to Koyré, first with acute interpretations of important historical cases (S. Carnot, the Copernican revolution, etc.) and then with a book that covers the whole history of classical physics. This historical vision enjoyed much greater success than that of Koyré. It popularized the concept of "paradigm" and definitively introduced the concept of "revolution" into the history of science; today these two concepts are familiar to and used frequently by scientists of all disciplines (e.g. sociology, psychology). With his work, the history of science became a recognized and appreciated academic discipline.

Note that Kuhn's book has a very attractive title: *The structure of scientific revolutions*. In years of great social changes and revolutionary (Marxist) ideals the introduction of the word "revolution" in the realm of certainty (which was how Science was considered) was, in addition to being a considerable novelty, a genuine renewal of an academic sector (history of science) which is of great importance for cultural progress. Kuhn's title promised to highlight the revolutions that had taken place in Science; and, once crucial moments of reference had been determined, it was hoped that the evolution of that enormous human construction which

[3] In *Études d'histoire de la pensée philosophique,* Paris : Gallimard, 1961.

is Science could be completely clarified. Moreover the word "structure" suggested the idea that Kuhn had succeeded in determining the very foundations of Science, its supporting framework, thus concluding the centuries-old effort of research by philosophers and scientists developing new philosophies spontaneously.

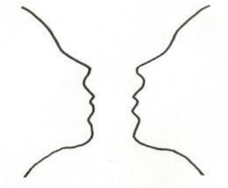
> Few books have created such a profound impression, especially in the author's lifetime. Today it is difficult to think of an academic discipline, both humanistic and scientific, that has not been influenced by it.

Kuhn's book does not deal with the birth of modern physics (nor even Koyré's interpretation of the birth of science, as if it were no longer necessary to discuss it), rather it has the more ambitious goal of interpreting the history of the whole of classical physics. Indeed, the book's illustration of the history of physics ends before the crisis of Physics in the early 1900s. It covers a much longer period of time (three centuries) than that considered by Koyré (about a century).

Concentrating the history of such a long and complex period into 200 pages means abandoning the history of individual theories in order to study the history of the totality of theories, seen in their temporal succession and in their conceptual and hopefully foundational relationships. This is the qualitative leap that Kuhn actually makes with respect to Koyré and other previous historians. (with this leap, his book, more than any other historical account, approached, but did not achieve, the idea that the set of so many theories has a "structure", as the book's title indicates). In only about 200 pages. Kuhn chose as the main subject of study not the history of science, as one would expect, but which is the best schema for writing history of science; that is, he presented examples of historical cases and historical assessments in order to support a historiographic schema with which to interpret the history of the classical physical theories. The exposition of this schema is accomplished intelligently by referring to a series of historical cases which, taken together, represent a historical sequence of theories. The result is that the book is at the same time a book of history and a book of historical interpretation, but much more the second.

Kuhn's schema sees the evolution of the history of science as being interrupted by revolutions. These separate long historical periods, in each of which scientists carry on ordinary science, that is they solve puzzles, while all following the same *paradigm* (set of techniques, undisputed fundamental concepts, common assumptions, cognitive presuppositions) applied systematically to all possible situations, in a not very creative endeavour. The paradigm is transmitted to scientists mainly by textbooks of that period, all similar. In periods of paradigmatic science, the judgment of validity of a scientific result is not given only by the experimental confirmation, but by the consensus of a group of authoritative scientists, the scientists community, who share (more or less consciously) the same paradigm; their judgment weighs much more than that of any single leading scientist.

At a certain point, however, an *anomaly* appears, in the sense of a problem that cannot be solved by means of the usual paradigm. This leads to a period of crisis in Science. The crisis is overcome with a paradigm shift, suggested perhaps by some genius. A drastic change then takes place in the perception of the world by the scientific community. Kuhn suggests an analogy with Gestalt, which is that particular type of visual phenomenon where two completely different things can be seen in the same figure, but never simultaneously. Two of such figures are classics: one depicts a chalice, which can also be seen as a pair of profiles of human faces looking at each other; the other is the figure of a black bat against a yellow background, which can also be seen as a toothed mouth.



The Gestalt change occurring in the scientific vison of the scientific community leads to a period of a new paradigm, during which scientists resume solving "crossword-problems". But between the previous and subsequent paradigm there is, as indicated by the unexplained Gestalt phenomenon, an *incommensurability*, as evidenced by the radical variations of meaning that basic scientific concepts undergo. These variations give rise to a profound difficulty in the communication between the two paradigms, and also to difficulties in the work of the historian of science who must strive to enter into the Science of a historical period that is not his own. Moreover, Kuhn says that, due to incommensurability, not all the scientific results obtained by the previous paradigm are conserved (non-cumulativity of scientific results).

On the validity and limitations of Kuhn's interpretation, a broad debate that is still going on has developed since the book's inception. The debate oscillates between, on the one hand, the accusation of irrationalism aimed at Kuhn's thesis that, during revolutions, science changes (through a Gestalt) in an irrational manner because it is not explained by the scientists themselves, and, on the other, recognition of his decisive contribution to the history of the Science of the 1900s, which until then had remained completely undervalued

The first analytical study of Kuhn's work was that by Mastermann, who investigated the contextual meanings of the word "paradigm", the main category of the historical interpretation of Kuhn's book. The result is that in Kuhn's work this word has at least twenty-two different meanings (the contextual analysis of the other key word, i.e. "revolution", did not enjoy a better fate). This study draws attention to the fact that Kuhn's work is written in a fascinating manner but is actually more a work project (ingenious, but methodologically incomplete) than what the author had claimed it was, that is, a new stable interpretation of the history of classical physics and therefore of most of modern science.

Moreover, another contradictory aspect can be observed. Kuhn does not see revolutions beyond the two well-known ones: that of the birth of Science and that of the early 1900s (which however he did not deal with). According to him, the birth of Thermodynamics and that of Chemistry did not constitute revolutions of the Newtonian paradigm. In fact, the book does not present these two theories as constituting a new paradigm to undermine the first one and hence creating a Gestalt. In particular, Kuhn states that even the birth of Chemistry (considered by many, including Lavoisier himself, a true scientific revolution), was not a real revolution since it was caused by "supra-mechanical" factors (that remain mysterious in the book); that is – in the book there is no other way of understanding it – the birth of chemistry was brought about through the influence of Mechanics (obviously Newton's). Mechanics appears therefore to loom large over the whole history of classical Physics, in contrast to the title itself of the book, which promised to present a plurality of revolutions in the history of Science. Thus we have a paradox: a book that wants to revolutionize the history of science and moreover proposes to introduce into it the concept of scientific revolution, actually never describes a revolution.

The existence of fundamental problems in Kuhn's interpretation became evident to all historians when, after his famous work, he tried to apply his interpretative schema to the crucial case of what everyone considers a true, profound and still unexplained crisis of physical science, a crisis that, with the introduction of the quantum, had necessarily to be explained as a revolution in theoretical physics. But his book on quanta never used the words "anomaly", "paradigm" and "revolution". Rather he argued (in any way intelligently) about the priority of discovery of quanta (attributing it to Einstein rather than Planck). Indeed, throughout his life he failed to describe a scientific revolution by means of his categories (apart from that of physics in the nascent state, the Copernican).

6.2 Interpretation through the two dichotomies of the historiographies of Koyré and Kuhn

> Newtonian mechanics was a paradigm for most scientists during the greater part of the last two centuries and is therefore the paradigm of the paradigms for historians of science… one may recall that until around 1900 it was assumed that any problem in physics could be solved, at least "in principle", by applying Newton's law of mechanics; it was only necessary determine forces and the mechanical properties of the parts of the system and then compute a solution for the appropriate set of differential equations. Since this was the most successful theory in any science, theorists [of even all other branches of sciences] tried to imitate it; but that meant adopting what was thought to be Newton's philosophy of nature as well as his scientific method. If we think that the sciences have now rejected the Newtonian paradigm, we can nevertheless use this historical case to understand what it means to be dominated by a paradigm. (Brush 1976, p. 21)
>
> … In other, more mathematical words, the historical space [of the historiography of science] is insufficiently defined. I use the word 'space' deliberately, for I regard the historian as working in a space in a modern mathematician's sense: a multi-dimensional region, whose dimensions are determined by the historical and historiographic factors which he brings to his studies. Historical figures are like mass-points; influences between them are like forces of attraction and repulsion, and more general influences resemble fields. A community is thus a collection of mass-points, usually in some sort of equilibrium, but vulnerable to substantial disturbance. I find the analogy useful, although I do not take it further and play Cauchy, for example, and try to set up the differential equations to represent the phenomenon. Cliometrics has not yet advanced so far. (Grattan-Guinness, vol. 1, p. 6)

Let us rethink the interpretation of the birth of science.

Koyre's interpretative model, compelling because it sees the conflicts between different scientists, concerns the period in which Newtonian mechanics was emerging (while geometric optics can be devalued both as a mathematically simplistic theory, and as a special case of the mechanical theory according to Newton's hypothesis that light is composed of material quanta). Its history sees a progressive effort to give birth to a "new science" that reached its conclusion with the "complete" theory of Newtonian mechanics.

But we know now that the Newtonian theory established the AO and AI choices. Considering all the works of Galilei, Descartes, and those of "minor" scientists to be absorbed by the Newtonian theory, Koyré left out of consideration an important part of Galileo's problems (for example the reflection on the two dichotomies, the antagonistic opposition between Descartes and Newton, Newton's inability to formulate optics within his theoretical schema, the contributions of Huygens and all the anticipations of science that can be linked to non-Newtonian choices). However, his descriptive picture holds substantially well, one reason being that it corresponds to the dominant Newtonian mentality. This partial result is already a great merit for a historian.

But Koyré's greatest merit is that he chose to clarify that part of Science which, although not experimental, constitutes its metaphysical presupposition. He courageously pointed out the metaphysics that in his opinion determined the emergence of science, believing that he had found it in Galilei, Descartes and Newton. He expressed it in almost all of his writings. It concerns Mathematics and, essentially and even more clearly, space, as the title of his most famous book (*From the closed world to the infinite universe*) says through the words world and universe. This is one of his many statements about the metaphysics of the science of that period:

> The scientific revolution of the 17th century, the era of the birth of modern science, in itself presents a very complicated history. But since I have dealt with it in a series of works, I will allow myself to be brief. I therefore characterize it as having the following aspects:
>
> a) *dissolution of the Cosmos,* that is, substitution of the finite and hierarchically ordered world, of Aristotle and the Middle Ages, with the infinite Universe.
>
> b) *geometrization of space,* that is, replacing the concrete space (set of places) of Aristotle with the abstract space of Euclidean geometry, now considered as real .... [4]

If we try to reconstruct the origin of these categories, it can be argued that Koyré received from Husserl the guiding idea that the metaphysics of science originated in its mathematization and, in particular, in the concept of infinity, as expressed at least in aspect a). But when Koyré went on to study the scientists of the 1600s, he found that its philosophical nature alone made it highly controversial and unpopular with scientists, who instead, as happens in Copernicus, Galilei, Descartes and Newton, were much more at ease with typically physical concepts, such as the universe and space. The concept of space also had the metaphorical meaning of an organization of all the constituent objects, such as for example, the organization given to the points of the space by the Cartesian axes, which emerged precisely in the period studied by Koyré. In this sense the choice of space corresponds to the choice of the organization of the theory itself and replaces it. In fact, in the physical theories emerging at that time, namely those of Descartes and Newton, there is the same type of space and the same organization of the theory, AO.

Having established this connection between apparently very different concepts, it is easy to interpret the other words of Koyré's statements; for him the word "dissolution" indicates the concept of "negative modification" of the organization, that is the rejection of problem-based organization of ancient theory, based on the problem of explaining the world around us.

He also makes it clear that in the title of his most famous book the word "Cosmos", as understood by the ancients, is to be understood as "finite Cosmos", i.e. as PI. This is also rejected by modern science at its inception. In fact, with Descartes modern scientists chose the

---

4 A. Koyré, *Etudes de l'Histoire de la Pensée Philosophique*, A. Colin, Paris, 1961, p. 258

mathematization of space, which he calls "geometrization"; however, we are well aware that the most advanced mathematics of the time, rather than geometry (Euclidean or Cartesian analytical geometry), was infinitesimal analysis. Koyré also knew it, but he could not indicate it with a word like "analysisation" or "infinitesimalization". He therefore referred to a concept relating to the mathematics that is the most intuitive and of *dominio commune* commonly known. This makes it clear that Koyré's "geometrization" is really not so much about Descartes' analytic geometry, but Newton's analysis, with which the AI choice was made.

Ultimately, Koyré intuitively conceived all the choices given by the two dichotomies and with these intuitive words he characterized the transition from ancient to modern science in a historically very suggestive way, representing it as a struggle between intuitive concepts. This is appealing to readers and at the same time suited to the science of the historical period studied, since in fact these concepts actually represent the pair of choices AI and AO as against PI and PO.

Some historians have tried to extend the Koyreian programme to the Physics of the time following its emergence, but the results are much less interpretative than those of Koyré. Kuhn, believing that he had learned the Koyreian lesson on the birth of science, in fact described the Physics that followed the birth of modern science as an unstoppable progression of Newtonian Mechanics until 1900. Despite having announced "revolutions" in the plural, he disregarded the entirely new theories that had emerged before 1900 (e.g. it is a historical fact that the birth of classical Chemistry occurred as a clash between Lavoisier and the ideas of chemistry inherited from the classical mechanics of Newton; in this case Kuhn presented only a transient conflict between Priestly's chemical ideas and Lavoisier's theory. In this sense, Kuhn's historiography does not represent a real advance in the clarification of the Foundations, which Koyré's historiography had intuitively begun; rather, it represents a cultural operation which conserves the fundamental preconception: Newtonian mechanism and its undisputed scientific progress.

Table 6.1: *CONCEPTUAL CORRESPONDENCES BETWEEN KUHN, PLANCK AND NEWTONIAN MECHANICS*

KUHN	PLANCK	NEWTON
Paradigm	Image of the world	Null force
Normal science	Typical activities of scientists	Velocità costante
Anomaly	New facts	Non null force
Revolution	Profound reconstruction	Acceleration
Incommensurability	Incompatibility	$v_2 \neq v_1$

Table 6.1 shows a surprising correspondence between the key concepts of Kuhn and those of Newtonian mechanics, which he considers a paradigm. But even Planck, when he represented that history of theoretical physics of which he was the protagonist (quantum crisis), used keywords that are similar to those used subsequently by Kuhn. The table explains Kuhn's success: he used categories that were quite acceptable to all modern scholars of science, because they were associated with those of Newton's dynamics. In other words, the Newtonian concepts, conceived to interpret dynamics in Physics, were elevated by Kuhn to historiographical categories to interpret the history of the period in which the Newtonian

paradigm dominated. The operation of elevating scientific ideas to interpretative categories of the history of science was indeed a considerable novelty and will) then be theorized by two other important historians, Brush and Grattan-Guinness (see the heading of this paragraph). Given the obvious consonance between the two levels considered by Kuhn (that of the paradigmatic theory and that of the history of the paradigmatic period), his historiography was certainly able to explain much of the period considered and therefore his work as a historian was truly productive.

But such surrogate Kuhnian concepts do not relate to physical concepts (space, time, force) or mathematical concepts (line, radius of curvature, derivative and integral), but to notions of the humanistic sciences: AO has become a social organization, that of the most authoritative scientists, that is, “the scientific community”; whilst IA, which in mathematics is represented principally by the infinitesimal, relates to psychology: in the sense that IA, which cannot be understood in terms of real numbers, has become a mental act which is operationally incomprehensible, namely, the psychological phenomenon of *Gestalt.*

It is remarkable that, like Koyré’s, his historiography can also be summarized in subjective terms with the phrase: "the Gestalt-switches [AI] suffered by the organization of scientists [AO]"; which in fact captures the culminating events of scientific revolutions. Then, changing the little-known concept of *Gestalt* to the widespread and appealing concept of "revolution", we find “the revolutions suffered by the scientific community” which is the famous title of his book. Note that these two main concepts of Kuhn’s – Gestalt and organization of scientists – represent AI and AO through respectively psychological and sociological concepts applied to large human groups. They implicitly suggest – according Kuhn’s neo-positivist standpoint - that the metaphysical part of Science could be ignored and that another branch of knowledge, the humanities, should be referred to.

In conclusion: *the "new historiography" of Koyré and Kuhn is characterized by having united, unconsciously but effectively, the history and the Foundations of Science. Each of them accomplished an original translation of the fundamental choices of a scientific theory into appropriate and suggestive historical-scientific notions, with which they were able to interpret a specific field of study of the history of Physics more deeply than in the past.*

It should be noted that these historians who attempted to interpret the history of science in a new way can be very satisfied with the results of their work. They succeeded in offering the first solutions to a task that was not specifically of their competence, but rather of philosophers of knowledge, i.e. to understand the Foundations of Science at least in partially. In fact, they came much closer to this understanding than previous philosophers, despite the fact that the latter had accumulated many theories and reflections and had had much more time at their disposal to understand what the interpretative categories of science were.

6.3 General picture of the History of Physics

> [In Philosophy and in the historiography of Science] An increasingly deeper and safer [material] ground was sought

[empirical facts, mathematical techniques] so as to lay there the foundations; while there is no firm thing but the *firmament*, as the word itself says.
Lanza del Vasto, 1943

We have previously seen that the foundations of a scientific theory consist of two dichotomies that we have found by comparing the various formulations of the theories of classical physics and we have verified in the two main theories of modern physics. We can therefore analyse the entire history of Physics in terms of these two dichotomies, hopefully in greater depth than Koyré's and Kuhn's analysis.

The pairs of choices regarding the two dichotomies indicated above give four possibilities for constructing the foundations of a theory. We will call them models of a scientific theory (MST). In Table 6.2 we find the Newtonian, Carnotian, Cartesian and Lagrangian MSTs, named after the most famous scientists who can represent them through their main physical theories (although Descartes' geometric optics was of little importance for Physics after the first half of the nineteenth century).

Now let us assign each classical physical theory and each formulation to a precise MST. We note that the most well-known formulations of physical theories are distributed almost equally among the four MSTs: this is indicative of the fact that the inventiveness of physicists has already intuitively explored all the potential of MTSs with specific theories. And conversely, the categories used are reasonable and appropriate to the variety existing in theoretical physics.

TABLE 6.2

THE FOUR MODELS OF A SCIENTIFIC THEORY AND THE CLASSICAL THEORIES THAT REPRESENT THEM

	***ARISTOTELIAN ORGANIZATION* (AO)**	***PROBLEM-BASED ORGANIZATION* (PO)**
***ACTUAL INFINITY* (AI)**	**I° MST (Newtonian)**: Newtonian mechanics (1687), mechanicistic optics (1700), Maxwell's electromagnetics (1862)	**III MST (Lagrangian)**: Maupertuis' mechanics and variational formulations in general, Lagrangian mechanics (1788), statistical mechanics (1890)
***POTENTIAL INFINITY* (PI)**	**II° MST (Cartesian)**: geometric optics (1630), L. Carnot's mechanics (*Principes* 1803), classical thermodynamics (1851), physical chemistry (1878)	**IV° MST (Carnotian)**: L. Carnot's mechanics (*Essai*, 1783), S. Carnot's thermodynamics (1824), classical chemistry (1866)

If we consider all the theories that in the first half of the 19th century were alternatives to the Newtonian MST, we can see that their founders were only partially aware of making fundamental choices. None were able to characterize them (with the exception of Lazare Carnot, whose ideas were not followed up). This suggests that the history of science lived by scientists, that is the "*subjectively understood* history", was certainly different from the "history that *effectively occurred over time*", that of the fundamental choices actually made by the new theories, i.e. the effective evolution of the foundations of physics (independently of the knowledge that scientists had of it); and obviously, it was also different from the "history *objectively presented with facts*", that which is codified in the manuals through only experimental results and the resulting mathematical laws. To highlight these three different aspects of physical theories and the resulting mathematical laws, let us contrast them, but, for the sake of simplicity, only with respect to the two most important MSTs

TABLE 6.3

THE TWO MAIN MODELS OF A SCIENTIFIC THEORY

(A few centuries)	*(One century)*	*(A generation)*
Effective science (as scientific geniuses determine it through the two dichotomies)	**Subjective science** (as scientists think by means of surrogate concepts)	**Objective science** (as it is taught, by means of techniques and objective concepts)
NEWTONIAN MST ***(AO + AO)***	'*Dissolution of the finite Cosmos and geometrization of space'*	Classical logic Analytical method Infinitesimal analysis (eg: 2^{nd} order differential equation)
CARNOTIAN MST ***(OP + IP)***	*'Evanescence of force-cause and discretization of matter'*[5]	Non-classical logic Synthetic method Symmetries (or cycle, eg S. Carnot's in thermodynamics)

Legenda: MST = Model of a Scientific Theory; AO = Aristotelian Organization; PO = Problem-based Organization; AI = Actual Infinity; PI = Potential Infinity.

In teaching methodology objective physics is the physics of examinations, subjective physics is that expressed in the intuitive representations of facts, laws and principles that teachers indicate to the students to help them understand the subject. The effective presentation of theoretical physics concerns only one foundation (the AI and AO choices of the Newtonian MST), at the cost of obscuring the cultural value of Physics.

The clarification obtained allows us to propose a general interpretative schema of the whole history of Science, whichever MTS a theory belongs to.

TABLE 6.4: *THE EVOLUTION OF PHYSICAL THEORIES THROUGH FOUNDATIONAL CHOICES AND THEIR SURROGATES (CONCEPTS AND THEORIES)*

---

[5] This couple of propositions are derived by interpreting Koyré's through the two Newton's choices and then inventing two similar propositions based on the opposite choices. The paper A. Drago, "The several categories suggested for the "new historiography of science": An interpretative foundational viewpoint, *Episternologia*, 24 (2001) pp. 42-81 interpret through the two dichotomies all the historical categories suggested by the numerous historians of science.

	1630	1687	1782	1870	1905	1925	1982
CHOICES AND REPRESENTATIVE THEORIES	AO Geometric optics	**AO** Mech. Newton ? Optics, Acoustics	**AO** Mec. Newton ? Optics, Acoustics	AO Math-.Phys Electromagn. Carath. Thermod Statistical mechan	AO Math-.Phys Electromagn. Carath. Thermod Statistical mechan	AO Statisticl Mechan Quantum Mechan. General Relat. Quantum Chem.	AO Field Theory General Relat. Lagrangian
		AI Mech. Newton ? Optics, Acoustics	**AI** Mech. Newton ? Optics, Acoustics, Lagrange's Mechanics	**AI** Mech. Newton Electromagnet.	AI Math-Phys. Electromagnet.	**AI** Quantum Mech. Gen. Relativ..	**AI** Lagrangian and Hamiltonian Theory of Fields
			PO Mech of L.Carnot and Lagrange Chemistry Therm.S.Carnot	PO Chemistry Physical Chemistry	PO Special relativity. Physical Chemistry	PO Special relativity. Mecc. Matrices	PO Mech. L.Carnot Chemistry Informat.theory
	PI Geometric optics		**PI** Mech.L.Carnot Chemistry Therm.S.Carnot	**PI** Thermo-dynamics Chemistry Physical Chemistry	**PI** Thermo-dynamics Physical Chemistry Statistical mech	**PI** Thermo-dynamics Special Rel. Mech Matrices	**PI** L. Carnot mech. Inform. theory. Reticular theories
FOUNDATIONAL CHOICES	*Two choices*	*Two new choices*	*Pluralism of choices*	*Newton's MST dominates over the other choices*	*Four choices or 'confusion'*	*Restoration of Newtonian MST, New MSTs?*	*Pluralism of MSTs*
FOUNDATIONAL THEORIES	*GEOMETRY*	*MECHANICS*	*MECHANICS VS. CHEMISTRY*	*MECHANICS VS. THERMO-DYNAMICS*	*STATISTICAL MECH. & RELATIVITY & ELECTRO-MAGNETIM*	*QUANTUM MECH VS. RELATIVITY*	*MANY DISPARATE THEORIES*
FOUNDATIONAL CONCEPTS	*Light ray*	*abs. space abs. time abs. force-cause*	*abs. space abs. time abs. force-cause vs matter, work*	*force, field, ether vs. entropy*	*particle & wave vs. ? particles and system*	*particle & wave & ? particles (system)*	*Many basic concepts: particles, quarks, strings, etc.*

Key: AO = Aristotelian Organization; PO = Problem-based Organization; AI = Actual Infinity ; PI = Potential Infinity; → : inclusion. In bold the dominant choices.

It should be noted that physicists conceived these theories through those subjective concepts that they considered to constitute the foundations of physics. It is for this reason that the concepts of space, time, force and matter have assumed so much importance in the history of Physics and for the same reason seem so important in teaching physics; but their role has been exaggerated, because they refer only tenuously to the Foundations. of Physics.

For completeness we give an analogous brief history of Mathematics (excluding the history of the XX century, which usually does not interest physicists).

Table 6.5: THE EVOLUTION OF MATHEMATICAL THEORIES THROUGH FOUNDATIONAL CHOICES AND THEIR SURROGATES (CONCEPTS AND THEORIES)-

	540 BC	300 BC	1300	1620	1687	1778	1870
CHOICES AND REPRESENTATIVE THEORIES	(AO) (everything is number)	**AO** Eucl.Geom	**AO** Eucl.Geom.	**AO** Eucl.Geom., Geom. analit.	(AO) Euclidean Geom	(AO) Euclidean Geom	AO Cantor's set Theory, Frege's Logicism
				(AI) (Transc.numbers, series, Algebra)	Infinitesimal analysis	**AI** Infinitesimal analysis	**AI** Analysis ε-δ, Cantor's Set Theory
			(PO) (Algebra)	(PO) (Algebra)	(PO) (Algebra)	**PO** Geom. L. Carnot and Lobachevski Analysis Lagrange, L. Carnot and Lobachevski	**PO** Erlangen Progr. Klein
	(PI) Finite cosmos	(PI) Ruler and compass	(PI) Curves ruler and comp. Algebra	PI Analyt. geom. Algebra	**PI** Synthetic geometry	**PI** Finite analysis Lagrange, L. Carnot, Lobachevski Geom. Of synthetic geom..	**PI** Finite groups, Kronecker fields
FDT. CHOICES		*one*	*choice*	*Two choices*	*Three choices*	*All choices*	*Two dominant choices*
FDT. THEORIES	*(Arithmet.)*	*Geometry*			*Infinitesimal analysis*		*Set theory*
FDT. CONCEPTS		*Ruler and compass*	*Ruler and compass, elem. Algebra*		*Infinitesimals and Ruler and compass*	*analytical method vs. synthetic method*	*Rigour vs. Intuition*

Key: as in the previous table.

Note how in the history of mathematics the emergence of specific theories has given rise, one by one, to the individual foundational choices; it was only at the end of the nineteenth century that there emerged a theory, set theory, with a pair of strong choices, AI and AO. On the other hand, the previous tab. 6.4 shows that in Physics this pair of choices was established almost immediately by Newtonian Mechanics. In this foundational sense, Physics was able to emerge because it had been prepared by the millennial effort of preceding theories of Mathematics. Then Newton's Mechanics assumed the strong choices (AI and AO) which for a long time (almost two centuries) became paradigmatic for the whole of science, including Mathematics, where any of its theories had this couple of choices.

6.4 New definitions of paradigm, incommensurability and scientific revolutions

Comparing the concept of MST with Kuhn's "paradigm" we note that the first concept makes precise what was inaccurate in the second. Conversely, "paradigm" adds only one characteristic to the concept of MST: that of being a dominant MST with respect to others (as represented by the theories of its time, to the point of obscuring them all and being considered the only possible foundation. The clear example of an MST that was a paradigm in history of Physics for at least two centuries is the Newtonian MST.

Furthermore, the concept of the *incommensurability of different theories that refer to the same phenomena*, introduced, only intuitively by both the philosopher of science Feyerabend and the historian Kuhn, is very important here. They applied the concept above all to the theory of Newtonian mechanics, on the one hand, and special relativity or quantum mechanics, on the other, but their definition of it was approximate and not supported by a precise analysis of these historical examples. There are still therefore divergent interpretations of the importance of this concept: is it crucial for understanding the history and the foundations of science, or misleading in that it questions the very rationality of science? Can two theorists, espousing different incommensurable theories, communicate with each other? Is the evolution of science in the period of incommensurable paradigms irrational, given that the scientific community is not immediately aware of the Gestalt change?

We can now, in the light of the previous clarifications of the two fundamental dichotomies, explain this concept.

Neither theorists saw problems of incommensurability in classical physics, whereas our previous study leads us to consider various pairs of classical theories that are radically different. In modern physics also we saw that every new theory has pairs of formulations that are very different.

A moment of reflection convinces us that we cannot use PI and AI together in the same theory: the second (infinity) in fact cancels the problems of undecidability indicated only by the first. The examples given with the study of the previous theories are sufficient to indicate the linguistic and conceptual difficulties a theoretical physicist encounters when he does not take into account the fundamental choice of the type of Mathematics (let us recall, for example, the concept of reversibility). (In the past the problem has been circumvented, for example, by taking the "not unequal" of the formulas with infinitesimals to mean "exactly equal"; or using the word "equivalent" for the same purpose without defining it; or imposing, instead of choosing PI, mathematical limitations to AI mathematics (such as the limitation to continuous functions with first and second derivatives, limitations that do not give a a set which is closed under differential operations).

Previously it was suggested that each physical theory makes a choice between two types of organization, Aristotelian (AO) and problem-based (PO), which correspond to two types of logic, also in opposition to each other. Classical logic and non-classical logic cannot be used together, because a double negation either is or is not equivalent to the corresponding affirmative. Until now, every scientific theory has been abstractly conceived as conforming to classical logic, so that the theoretical physicist was not able, through reflection, to identify the cause of his difficulties concerning the differences between the logics which occur e.g. in thermodynamics. And the textbooks refer (without even mentioning the fact) physical theory to classical logic, thus overlooking the incompatibility between the two logics (at least as far as double negations are concerned) and committing a serious imprecision of language, thus confusing the mind of the student who wants to acquire a full understanding of the subject studied.

Also in this context recourse has often been made to convenient solutions, such as classifying non-classical logic as "deviant", taking that which is "<u>in-variant</u>" as "constant", speaking only of "principles" without ever distinguishing between those that are axiomatic and those that are methodological. It is this art of suggesting "linguistic" solutions that falsely represents the intellectual work on the part of the physicist as a work of genius, that is to say, it causes him to depart from a regular and straightforward mental path according to *a particular type of logic and by means of clearly defined concepts*, in order to be able instead to join, through opportunistic words, mentally idealizing theoretical abstractions together with the operational reality of experiments, nuanced concepts together with rigorous mathematical ideas (exclusively of classical logic).

On the basis of the dichotomies we now give a precise definition of incommensurability: *two scientific theories that concern the same field of phenomena are incommensurable when they are: 1) systematically organized, 2) mathematized, 3) different as regards at least one of the two fundamental choices.*

Condition 1) generalizes to theories that which was the incommensurability of the ancient Greeks between magnitudes; condition 2) restricts theories to mathematized theories in order to avoid comparisons between informal theories and impossible to compare with precision; condition 3) generalizes the ancient Greek definition of incommensurability between magnitudes in two respects: just as the Greeks called incommensurable two magnitudes whose relationship was not expressed in the Mathematics of that time, so two theories are incommensurable when their types of Mathematics are different. Moreover, since in modern science the concept of incommensurability refers to theories, it also concerns the organization of the theory, as well as its logic. Thus the Greek word for “incommensurable” as applied to magnitudes has been generalized to modern scientific theories which are characterized by their pair of choices regarding the two dichotomies.

The new definition suggests that, just as the historical development of Greek science was limited by the incommensurability between magnitudes, since the Greeks did not want to develop studies on incommensurable numbers, so the development of Western science was essentially limited by the incommensurability between scientific theories (e.g. that which led to the imposition of Newtonian Mechanics on Electromagnetism, or that which prevented the understanding of S. Carnot's Thermodynamics for 25 years (as we saw in Chapter 2); with the difference that for the Science of the Greeks the limitation was conscious and deliberate, while for Western science the limitation remained unknown and was therefore experienced as an incomprehensible incompleteness of a project of omniscience (always revived and always belied by new crises) of the Newtonian paradigm. The origin of this project can be attributed to Newton's *Opticks*, when he outlined with 31 Queries the programme to derive all the knowledge of the world from mechanical theory (including that of the new moral laws founded on the laws of the mechanics of Creation, to which men would have to be subject).

The incommensurability between theories was also the reason for the inconclusive discussion over three centuries about the Foundations of Physics as well as of the dramatic crisis of the early 1900s, because at that time it was inconceivable to admit choices other than the Newtonian (and Einstein who did so, was misunderstood for a long time, and then later his theory of relativity was received in technical terms only). In short, the history of physics suggests numerous cases of incommensurability and consequent misunderstandings and closed-mindedness. This historical misunderstanding does not however mean the impossibility of understanding another paradigm. Incommensurability requires a *c*areful work for connecting the basic notions of a paradigm with respect the basic notions of the other n order to obtain a translation between the two paradigms

Looking at the previous tables, one can define precisely *the concept of scientific revolution*, poorly defined by Kuhn (even though he dedicated a book to it). On the basis of the pairs of choices it can be characterized in several ways that range from those which give more importance to subjective science to those which give more importance to effective science.

In the subjective understanding of Science, by scientific revolution we can even mean the birth of *a single new concept* (for example, the infinitesimal that provoked opposition by the mathematicians of the time); while in the understanding science through the basic choices, the revolution is more profoundly defined as *a change in the dominant MST*. Since this second concept refers to the Foundations, it involves the change in a whole series of concepts, which therefore undergo radical changes of meaning (take for example, how the concept of space changes with the Einsteinian revolution). (Note that according to this definition "revolution" does not necessarily imply the suppression of the concepts and theories and the rival MST; this new concept of revolution exists within a pluralism of MST's in which a dominant MST, that is, a paradigm, is also overshadowed by the others, but not suppressed).

According to this last *effective* definition of "revolution", only the following four revolutions are obtained (in italics) as occurred in past history of physics:

1. *birth of the Newtonian MST* (1687)
2. *birth of the Carnotian MST* (1782)
3. *counter-revolution of the suppression of the latter* (1815)
4. *emergence of a plurality of MSTs* (1858)
5. *rebirth of the Carnotian MST* (1905)

Note that these scientific revolutions take place in the history of science almost contemporaneously with the great political revolutions of Western society: the English "glorious Revolution" (1685), the bourgeois revolution (the American (1783) and French revolutions (1789), and the Russian revolution (1917). This externalist linkage suggests that collective human history is interconnected, albeit at a deep level and not with direct material links.

Now let us think back to the history of Physics of the 20th century in the light of what has been said so far. At the beginning of the 1900s the paradigm of classical mechanics was resoundingly disavowed by the theory of special relativity (which is a PO and PI theory) in particular on basic concepts (space, time) and experimental facts (non-existence of ether). Furthermore, the discovery of the existence of quanta (PI) required a new theory, which differed radically from the Newtonian paradigm, if only with regard to the set of permissible values for physical quantities, which had also become discrete.

The current opinion among physicists is that classical Physics developed, from Descartes and Newton onwards, mainly according to the Newtonian paradigm (this opinion ignores classical physical theories that deviated too much from it). Moreover, The "modernity" of Physics consists in having become aware that radical changes were needed in the classical paradigm. But we can remark that the changes that actually took place above all through the technical novelties (probability theory, Hilbert space, tensorial calculus, etc.); but physicists achieved only a little part of the knowledge of the Foundations, that illustrated in the above. It is precisely for this reason that "modernity", which has undermined and gone beyond the paradigm of classical physics, has not found a precise definition among historians and the views of physicists regarding the current situation are still uncertain. Has the Newtonian paradigm been definitively superseded? And if so, what is the new paradigm, if relativity is not compatible with quantum mechanics? Or has the Newtonian paradigm essentially never been changed? Or are we now without a paradigm and without Foundations?

However, with the fundamental choices, we have seen that revolutions involving theories relying on new, non-Newtonian choices had already occurred several times during the period of classical physics. Yet they were reabsorbed, either due to the cultural repression by a dominant paradigm of any alternative, or with reassuring compromises (e.g. declaring that the differences between the different mechanics formulations were technical or covering the radical variations in meaning of concepts with linguistic tricks. By the end of the 19th century the embers glowed under the ashes, because the fundamental, non-Newtonian choices had already manifested themselves in various theories (Thermodynamics, Faraday's electromagnetism, kinetic theory of gases, etc.). But unsuccessfully. When Einstein disavowed the concepts of absolute space and time and showed the reality of quanta (which his 1905 paper clearly states), it was eventually apparent that he had departed from the Newtonian choices AO and AI: AO, because the concepts that previously served as principles for deductions in the Newtonian AO theory were now lacking, AI because the discreteness of quanta cannot admit actual infinity. Thus theoretical physics finally became "modern", as is rightly said, which is to say that by departing from the Newtonian paradigm on the basis of considerable evidence (including that of mathematics), Einstein manifestly suggested that there are other foundational choices in theoretical physics (although the subsequent development of theoretical physics either absorbed Einstein's novelties in traditional, idealistic formalisms or ignored them.

We conclude by adding the following comparison of the awareness on the part of physicists, mathematicians and logicians of the foundations of their respective sciences.

Physicists lost sight of the Foundations along the way, when they failed to evaluate the revolutionary novelty of the theories based on the principle of virtual work, those of L. Carnot and Lagrange; and furthermore they did not clarify the foundations fully to the extent of identifying them, when (special relativity and quanta) had called them into question and got closer to them than ever before.

Mathematicians have the merit of having recognized a crisis in the Foundations in the emergence of non-Euclidean geometries. But then, when in the '900 the radical differences among those Geometries might have brought to light the Foundations, the vast majority of working mathematicians followed Bourbaki, who rejected the discussion of the foundational problems by making the efficientism of operating on three special structures (order, topological and algebraic) the basis of mathematics. The minority of mathematicians who had been attentive to the Foundations since the beginning of the 20th century considered Mathematics as a science whose foundations had been clarified by Hilbert's choice of AO alone (rendered more powerful as an axiomatic) and then accepted the framework created by set theory (based on choices AO and AI), obviously taking it for granted that the only logic is classical (they barred the way to non-classical logic, even considering it to be a departure from rationality). The search for the Foundations *of Mathematics* in this abstract cultural environment divided them from physicists and meant that they had little influence on the development of new theories (for example, computer theory, genetics, artificial intelligence). On the other hand, those isolated mathematicians who at the beginning of the twentieth century had courageously begun a search leading to the identification of the alternative choices in mathematics, after some decades after sought compromises in order to co-exist with the dominant mathematics. Fortunately in the 60s an operative and efficient mathematics based on only PI, constructive mathematics, was obtained.

After the failure of Hilbert's program owing to Goedel's theorem, *mathematical logicians* being sceptical about the utility of non-classical logic, isolated themselves in their increasingly abstract formalisms of classical logic, without seeking contact with the reality of current scientific research (with the exception of their involvement in computer programming in the last few decades and their discovery that non-classical logic can also be useful there). Hence, they lost a connection with experimental science which might have given their science a clearer and more fecund role among the various sciences.

CHAPTER 7

NEW DIDACTICS OF MATHEMATICS AND PHYSICS

7.1 The formula for the current teaching of Physics and going beyond its past revisions

Here I will suggest a comprehensive rethinking of physics teaching methodology high schools.

Let us start with the question: why is statistical mechanics taught and not fluid dynamics? Why electromagnetism and not physical optics ~~more broadly~~? Ultimately, to what does the choice of the particular theories currently taught correspond? Of course, there is a criterion of technical and conceptual simplicity that strongly conditions the choice from among the many theories. But it is an objective fact that currently four precise theories have been chosen: Optics, Newtonian Mechanics, Thermodynamics and Electromagnetism; there is also Acoustics, which is however effectively assimilated into Mechanics, or, rather we should say, serves as a bridge between Mechanics and Electromagnetism.

The number four might lead one to suspect that the radical differences between the above four theories have lead to an awareness in didactics, in its striving to make Physics as comprehensible as possible, of the Foundations of Physics. If this were the case, then it should be possible to present the four theories according to the two dichotomies. This is in fact the case, as the following table shows.[1]

TAB. 7.1 THE FOUR PHYSICAL THEORIES IN THE TEACHING OF PHYSICS IN SCHOOLS OF PHYSICS, SEEN ACCORDING TO THE TWO DICHOTOMIES

	AO	**PO**	
AI	**Newton's Mechanics**	**Electricity and Magnetism**	*trajectory, line of force*
PI	**Geometric optics**	**Thermody-namics**	*distanzces, processes*
	absolute space, ref. frame	*field, system*	

Indeed geometric optics is taught with principles (AO) using algebraic techniques (PI). Newton's Mechanics is presented with certain principles, which are also emphasized (AO) while leaving the task of using AI mathematics to university studies; however, starting from the principle of inertia it proceeds with AI idealizations. In the teaching of Electromagnetism the concept of field is used: as it is presented, it is an obvious idealization (the AI of the infinite smallness of the "exploratory charge" in order not to disturb the field to be measured); whereas this theory is presented in the context of the important problem of linking electrical and magnetic phenomena within a theoretical organization that is suited to the level of studies (PO). The teaching of Thermodynamics clearly uses simple mathematical techniques (PI) and,

[1] A. Drago: "The paradigmatic schema of the teaching of Physics: the attempt to unify the four theories", *Giornale di Fisica*, 45 (2004) 3, 173-191.

although it proposes principles, which are presented as AO but are actually methodological principles, the problem-based perspective (PO) is maintained in the introduction of the concept of entropy and the reasoning in general concerning the efficiency of thermal machines. Acoustics, as was said previously, can act as a link between Mechanics and Electromagnetism, providing an aid to the better understanding of the latter and thus reinforcing the upper part of the table, where the more "authoritative" theories are located (An analogous schema can be indicated for university teaching, with, clockwise, again Newtonian Mechanics, then Quantum Mechanics, Thermodynamics and Relativity).

We note then that physics teaching in high schools has two great merits: it teaches theories and not just experimental facts or merely isolated laws, or simply experimental techniques, but also teaches those theories that, taken together, approach the Foundations of Science, even if it is not declared or even noticed by teachers. It should also be noted that managing four theories, instead of one, as was the case during the "good times" of mechanism, is a complex problem; and four is more than two or three and involves a non-trivial level of complexity.

Although it is clear that there are differences in the various theories in terms of approach and reasoning, the teaching methodology tries to attribute them all to the diversity of the phenomena presented. Moreover it presents both the method of physics and the mathematics used as unitary. It is what we find in every textbook: the presentation of “physics” rather than “physical theories” and the attempt to unify everything. The striving of didactics towards unity results from nostalgia for the "good old times" of the single paradigm or, in other words, the rejection of the essential diversity of theories.

However, this striving has a limit. Although Electromagnetism and Geometric Optics can be brought closer to Mechanics in certain of their aspects, this is not at all possible with Thermodynamics, which therefore appears a "strange" theory ~~in the context~~. To remedy this, in the usual teaching of physics, it is neglected as an "immature" theory, although a few hasty final lessons of the school year are set aside for it. This didactic choice confirms what appears clear from our previous analysis: the choices of Thermodynamics are essentially different from those of Newtonianism, which is dominant in teachers’ minds. As a result the incommensurability of the two theories gives rise to a cultural breakdown in communication. Teachers, being aware that the two theories are incompatible with each other, sacrifice the least prestigious one in an attempt to maintain the unity of Physics. However, we know that in terms of choices the incommensurability with Thermodynamics exists, and cannot be overcome. This renders the above didactic approach unreasonable.

The teaching of Physics is indeed exceptional because it wants to present as unitary a set of theories that instead form a complex picture in several aspects. Furthermore its search for unity is exceptional with respect to the teaching of other subjects: for example, the teaching of Philosophy or Italian Literature presents a multiplicity and has no difficulty in declaring it to the student, given that the current humanistic culture has never had a paradigm.

This way of teaching Physics^, that presents a complex picture of different theories in conformity with a striving towards unity, presents itself as "scientific". Consequently it cannot be questioned by students and must be accepted a priori as a necessary condition of understanding the subject.

But then when the student has assimilated the contents of the teaching he has received, he realizes that this striving cannot actually reach its goal of the unity of the subject studied, so that the student either accepts that such a goal can't be reached, which has to be regarded as anti-scientific, or else becomes convinced of his/her incapacity to grasp all the meanings of Physics. In this the teaching of Physics tends to create either a myth (the striving for scientific unity) or a disillusioned disaffection. This feature is once again exceptional and perhaps unique, but we know that it is based on the attempt to put together what is incommensurable.

It is usually said that the teaching of physics is difficult. Previously we have seen various reasons for this statement. But now we see the central reason: the impossible search for a lost unity due to the incommensurability between theories.

Taking into account this clarification of the essential intellectual content of the teaching of Physics, we see that the measures implemented in past decades in teaching methodology could not have (and have not) achieved significant results. Indeed, active pedagogy, the change of roles in learning and quasi-commercial attractions (figures, colours, comics, special effects, etc.) remain irrelevant to the central problem. Not even the cognitive sciences, which are the product of common knowledge, can say much about incommensurability (apart from Gestalt, which however remains unresolved). They are therefore limited to a series of didactic aids and miss its central point.

How to overcome the difficulty then? What previously seemed negative, i.e. the impossibility since the 1800s of reaching a lost unity, seems positive if considered from the point of view of the pluralism of choices (and therefore of theories). We have seen that for some time now Physics has irreversibly reached a pluralism, precisely defined through four theories that are incommensurable. What is being proposed then is to enable the student to participate in the intellectual difficulty that is encountered when approaching the four theories in the teaching-learning process. In other words, to explain, through the four representative theories, the incommensurability and / or the two dichotomies on which it depends. In this sense this book has tried to suggest ways to approach and explain the dichotomies and fundamental choices.

If there is a didactic difficulty here it is that in current theoretical physics it is not easy to give an example of a clearly PI and PO theory as a counterpoint to Newton's Mechanics: the difficulty of Sadi Carnot's Thermodynamics is that it is based on an outdated caloric theory. We could, on the other hand, introduce the Mechanics of Leibniz and L. Carnot, even if it might be more costly to the image of what Mechanics is, given that it ascribes more importance to energy, to which less attention is usually given in the teaching of physics.

In reality there is a theory that well represents the PI and PO choices and is already taught: Chemistry (provided it is taught without assuming the planetary model of the atom). Chemistry, taught according to the two choices mentioned above, alternatives to those of Newton, would have a very important theoretical function (as opposed to the almost servile one of a simple practice, as it is usually presented) of bringing out the two alternative choices, making it possible to clarify the Foundations of other physical theories on the basis of the same PI and PO choices and, moreover, it would manifest more than do other theories the cultural fact that in the past these choices were not adequately recognized. But Chemistry was separated from Physics without the possibility of claiming that its themes ~~did not~~ belong to

Physics. The stronger reason was that it is too different from the Newtonian MST, that is, its choices are manifestly PI and PO.

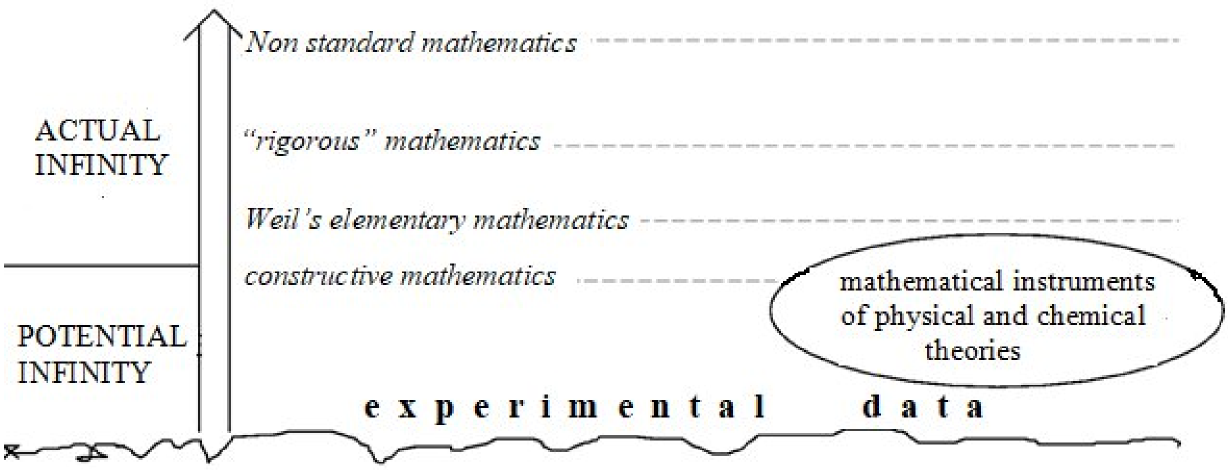
7.2 The transition to actual infinity in the history of theoretical physics

> It Is impossible to discuss realism without drawing in the empirical sciences A truly realistic Mathematics should be constructed hand in hand with Physics, as a branch of the theoretical construction of
> the one real world; and, like Physics, should maintain the same sober and cautious attitude towards hypothetic extensions of the foundations as is exhibited by physics. H. Weyl 1918

Previously many innovations in the conception of mathematics were presented in order to generate ideas for a new didactic model. It has already been suggested that the four geometries can be taught. Let us now deal with its main difficulty, namely its inability to teach the most sensational advance of modern mathematics, infinitesimal analysis, which moreover constitutes the mathematical tool of much of physics. Here we propose a way to emancipate the teaching of mathematics in high school from that position of inferiority that makes it seem an undertaking without interest. In reality, Calculus can be taught in a formally correct manner according to a very simple method, capable of obtaining all the results that are usually taught at University with the ε-δ limit technique. This will be achieved by refining the dichotomy regarding the type of mathematics, so far considered only through two opposite choices. We will now consider an additional choice, midway between constructive and classical mathematics.

TABLE 7.2

THE POSSIBLE RELATIONSHIPS BETWEEN THE DIFFERENT MATHEMATICS
AND THE DIFFERENT FORMULATIONS OF A PHYSICAL THEORY



To introduce the subject, let us ask ourselves the questions: how could Physics, which since Galilei's time, has clearly been an experimental and therefore operational science, accept the use of metaphysical concepts and the use of actual infinity in mathematics? And why was this fact not seen as an undue development, or in any case as an imposition?

There are philosophical and historical answers based on cultural influences from outside science (philosophical Platonism, the idealism of the ruling aristocracy). But there must also be answers from within science itself and it is only these that can explain three centuries of failure to acknowledge these problems. In the following we will consider one answer.

Mathematical logicians have studied the hierarchy of non-recursive functions (which are similar to constructive functions). Of course among them there is an infinity of degrees above recursivity, but it is interesting that the degree immediately above, having various noteworthy features, is that of "Elementary Mathematics" by H. Weyl,[2] which he proposed in 1918 in order to obtain a greater adaptation of Mathematics to Physics, in which continuous functions are defined by extrapolation from a series of rational numbers (just as the theoretical physicist extrapolates, from a series of measurements of experimental physics with rational values, a law represented by a function defined on real numbers in the mathematical continuum).

Elementary mathematics has various properties; the following three seem the most interesting, above all because they distinguish it with precision from the less powerful constructive mathematics.

1) On sets of infinite elements of concrete and countable objects only one logical quantification is admitted. Therefore we can speak of totality, for example the totality of integers, or that of rational numbers, but not that of the real numbers (each of these numbers is already an infinity, e.g. of digits) and even less the totality of real functions.

2) Existential quantification can also be represented mathematically, as usual on constructed or constructible entities. It is therefore possible to consider maxima and minima zeros of functions, etc.

3) The greatest lower bound (and the least upper bound) can be represented, in order to be able to find a convergent sequence; like all extremal values, these are not obtainable by constructive mathematics.

It is important to note that the above shows that *everything that is geometrically intuitive* can be obtained[3] (On the other hand, it should be remembered that constructive mathematics cannot express all geometric intuitions concerning the continuum: it cannot find the exact maxima of a function, nor the exact intermediate points between two given values, nor the uniform continuity of a continuous function defined on a closed and limited interval).

From the study of the evolution of physical theory we can deduce what motivated theoretical physics to use a mathematics that is more powerful than constructive mathematics, at least as powerful as Weyl's: it was the use of geometric intuition. There is a precise episode in the history of mathematics that concerns this transition: Cavalieri and Torricelli (~1640) tried to invent an infinitesimal calculus by subdividing a plane figure by means of a number

---

2 H. Weyl, *Il continuo* (orig. 1918), Napoli:,Bibliopolis, La sua matematica è stata chiarita da S. Feferman, “Weyl vindicatus. *Das Kontinuum* sixty years later”, in C. Cellucci, G. Sambin (edd.): *Atti SILFS*, CLUEB, Bologna, 1987, 1, 59-93. e anche in *In the Light of Logic,* Oxford: Oxford U.P., 1998, cap. 13.

3 A. Drago: “L'intuizione geometrica come chiave di volta della matematica della scienza moderna. Cavalieri e la matematica di Weyl”, in B. D'Amore, C. Pellegrino (eds.), *Atti del conv. per i 65 anni di F. Speranza*, Bologna, 1997, 40-44.

of parallel segments (called by them "lines"), each one being set in proportion with a segment of a figure of a known area (for example the segments of a parallelogram with those of a rectangle) and then proceeding to consider the measurement of *omnes* (in Latin; all) the segments, even if these are infinite and their totality does not therefore correspond to the usual definition of magnitude (according to the ancient rules of Eudoxus). This method implements the very first property of Weyl's mathematics: the use of a quantifier ("all") on rational quantities. A legacy of their notation that indicated *omnes* with an uppercase italic O, is the current integral sign ∫.[4]

They achieved important results with their method and this type of calculus established itself throughout Europe. They could integrate figures with variable contours, even of infinite extension (hyperbolas, spirals, etc.); Torricelli in particular obtained an astonishing result: he succeeded in calculating the volume of hyperboloid obtained by rotating the hyperbola around the *y* axis, truncated at the base, but extending upwards to infinity: the result was finite despite the fact that the figure was unlimited (With our methods, if the hyperboloid is centred on the *y* axis, it is sufficient to calculate the integral of the product of the base area $\pi x2$ multiplied by an infinitesimal height dy; the latter is also $d(1/x) = - dx/x^2$; therefore $\int \pi x^2/x^2 dx = \int \pi dx$, which, between $x=0$ and e.g. $x=1$, gives π).

The new method of calculus was based on geometric intuition (either of totalities or of single points) and was used systematically by the two Italian mathematicians. For example, Cavalieri's well-known theorem states that: "Given a curve and its secant, there is a point on the curve where the tangent to the curve is parallel to the secant." The main justification for this theorem is given by the geometric intuition of the existence of this point. Here the existential quantifier, i.e. the property 2) of Weyl Mathematics comes into play. Cavalieri and Torricelli received many criticisms for relying on geometric intuition. To defend themselves, they maintained that ancient geometers had also used the intuition of mathematical objects that were as ideal as "all lines", or ~~the intuition of~~ the single point.

The most important result of the calculation is the theorem of inversion between the two operations of (called in modern terms) integration and differentiation. Torricelli demonstrated (again by means of a technique that belongs to Weyl's mathematics) this result in a manuscript unfortunately only posthumously discovered in 1900. In other words he, together with Cavalieri, had already founded differential calculus, but had done it on the basis of the AI as it is suggested by geometric intuition (or Weyl mathematics), not a free use of AI, as Leibniz and Newton did forty years later with Infinitesimal Analysis. Thus this type of calculation dates back to 1640 and not to 1687.

It should be noted that, at the time of this historic mathematical development, the two mathematicians also proposed a specific version of the principle of inertia. With the new method of indivisibles Torricelli obtained so many exciting results that he allowed himself less rigour than Cavalieri. He considered legitimate the step from a sequence given only intuitively to the "final elements". It was in his opinion "the privilege of the surveyor"; in fact it corresponds to the use of the g.l.b or the l.u.b of Weyl's Mathematics.

---

[4] K. Andersen: "Cavalieri's method of indivisibles", *Archive for History of Exact Sciences*, 31 (1985) 4, 291–367. E. Giusti: *Bonaventura Cavalieri and the Theory of Indivisibles*, Bologna: Zanichelli, 1980.

Torricelli also considered physical situations obtained in this way to be valid. In particular, he discussed whether the supports of the two plates of a scale with equal arms are parallel or not. He concluded that they are not (which is clear if we recall that on Earth all vertical lines meet at the centre of the Earth). He argued, however, that the real situation is the ideal one that would occur if the balance were further and further away from the Earth until in the "final" condition the supports would be perfectly parallel. Recall that even Galilei had presented an ideal experiment of the same type: a ball rolling down an inclined plane (on the left) from a certain height and then ascending to the same height up an inclined plane (on the right), but whose inclination can be varied at will; the experiment could be extended to the extreme case of the inclined plane on the right being horizontal and the ball then proceeding to infinity. For Galilei this experiment was merely thinkable, not something which could happen in reality.

Analogously to the aforementioned case of the weight scale, Torricelli conceived the principle of inertia with the following period (albeit incidental, during a demonstration of a property of parabolic motion):

> It is clear that, without the attraction of gravity, the moving object would proceed with a straight and equable motion along the line with direction AB.[5]

Twelve years before him (and Descartes, when he enunciated the current version of the principle of inertia) Cavalieri had also expressed the same idea (although in a more articulated and complex form) of a rectilinear and uniform motion to infinity, provided that one abstracts from gravity. To understand fully these versions of the principle, it is necessary to consider that until then weight was attributed to the nature of the body: it could be diminished (for example by means of an inclined plane), but not cancelled; in other words, there was no distinction between weight, $P = mg$, and mass, m; conceiving a body without gravity was a conceptual leap, like conceiving the perfect parallelism of the arms of the weight scale, in Torricelli's previous example. We can therefore say that Cavalieri was actually the first inventor of the principle of inertia, formulated according to the mathematics of Weyl, and not Descartes (whose version, written in 1644 and taken up by Newton in the *Principia* of 1687, is based on many notions appealing to AI, as we have already seen in section 2.8).

It should be noted that in those years also in Optics magnitudes with infinite values were beginning to be encountered, again generated by geometric intuition: the geometric concept of light beam extended far beyond the physical measurements of that time, inevitably suggesting the idea of its path at infinity. The mathematical representation of the effects of lenses and mirrors (e.g. the formula $1/p + 1/q = 1/f$) also inevitably led to considering points at infinity, completely intuitive geometrically and physically no more than plausible. Indeed, Descartes' Optics (which was the first theory in Physics) introduced trigonometric functions and transcendent curves in general, which do not belong to the curves obtainable with ruler and compass, then considered by the geometric tradition and by Descartes himself, as the only "safe curves".

All this fits very well with what one of the most profound scholars of the Foundations of Mathematics had concluded, L. Brunschvigc:

---

[5] E. Torricelli: *De Motu Projectorum*, 1644, l. II, p. 156.

It is with Geometry that the consideration of the [actual] infinite was introduced into positive science, and this [fact] was anything but a fortuitous event.

In fact, Geometry has intuition as its main intellectual tool:

Geometric intuition owes its fruitfulness to the intellectual dynamism that is implicit in it and that actually makes it *trans intuitive*.[6]

At the time of Cavalieri there were some scholars (Guldino, Hobbes) who opposed this not fully justified extension to a mathematics that contained ideal elements, but they were ignored by many scientists who were enthusiastic about the new results. Once actual infinity was introduced with the insufficient but reassuring justification of the intuition suggested by Geometry (the scientific theory that underlied all the science of the time), scientists had no difficulty later in moving to unrestricted AI, as happened in the Mathematics of Newton's Mechanical Theory and Lagrange's Mechanics.

In the formalization of the variational principles we find not only the characteristic extremants of Weyl's mathematics, but also metaphysical elements of a philosophical nature (saving God's work, the principle of perfection, the idea that an event represents the most direct path between two points in a generic space, etc.). As a result, in theoretical physics, they remained a practice limited to certain restricted and not well justified cases (even the Dirichelet principle remained imprecise until the end of the nineteenth century). However, they were then strengthened, especially by Lagrange, with AI mathematics, used unrestrictedly. Thus in theoretical physics the use of a mathematics that was more powerful than constructive mathematics derived also from variational principles thanks to which it was established definitively.

Conversely, there is a reason why, at the end of the 1700s, Chemistry and Thermodynamics limited themselves to a reduced Mathematics. Already in Cavalieri Geometry had suggested the passage to actual infinity in Optics and in Mechanics. In the period of the emergence of the two experimental theories, with Poncelet, Geometry openly introduced points at infinity as having the same status as points with finite coordinates. Given that the two experimental theories clearly wanted to exclude the use of AI, they cautiously used a Mathematics that was reduced as far as possible, to the point of excluding Geometry, since geometric intuition suggested the passage to the AI of Weyl's mathematics.

This strategy of separating from the dominant theoretical physics and its AI mathematics has continued until today. In the twentieth century, for example, information theory and computability theory, being without actual infinity, are also without geometry (nor do they introduce any concept of space).

This historical reconstruction provides an answer to a possible question by a high school student: why is there no Geometry in Thermodynamics and Chemistry, which emerged two millenia after Euclidean Geometry and a century after Mechanics, which was based on analytic geometry? Furthermore, this reconstruction indicates how physicists and mathematicians have in fact made choices based on the type of mathematics, but have never been able to make them explicit, while at the same time these choices decided the foundations of science itself. In the light of this it does not seem surprising that, after a long period of

6 Léon Brunschvicg, *Les étapes de la philosophie mathématique*, Paris: PUF, 1947, p. 129.

choices made unconsciously, there occurred at the beginning of the 1900s a crisis (quanta) concerning the relationship between physics and mathematics and a crisis (absolute or relative space) concerning the relationship between Physics and Geometry. And there was also a crisis in Mathematics as even there the same alternative choices concerning its Foundations came to the surface.

7.3 New didactics of analysis using Cavalieri-Weyl mathematics

All this shows how Cavalieri's differential calculus and Weyl's mathematics played a crucial role in the history of Mathematics and Physics. Introducing this Mathematics into High School teaching would therefore not be reductive or limiting, but, on the contrary, it would allow for a historical understanding of that tension which, as we have seen in the previous chapters, has pervaded the relationship between Mathematics and Physics throughout the entire history of Physics.

There is currently an intense debate on the teaching of differential calculus at university; traditional teaching is considered oppressive, complex, and disconnected from reality. Here we suggest the use of intuition alone to simplify various difficult proofs, in particular that, crucial for teaching calculus, of the mean value theorem (which is in fact equivalent to Cavalieri's theorem).

Let us recall that in the nineteenth century Cauchy's studies initiated the reformulation of Analysis based on the concept of limit, conceived "rigorously" with the definitional technique of ε-δ. Today it is clear that the "rigour" of this reform was aimed at excluding the idea of the infinitesimal (and therefore departing from non-standard analysis). However, this reform, through the quantifiers "$\forall\varepsilon, \exists\delta\ldots$"^, maintained the link with actual infinity, surpassing in power Weyl's mathematics.

For teaching purposes it should be noted that at the same time as this reform, Cauchy discovered important discontinuous functions; since then the discontinuous functions and pathological functions (like that of Dirichelet) have proliferated. Throughout the nineteenth century, mathematicians competed in seeking ever more general definitions of basic operations (derivative and integral), in the hope of *being able to apply them to any function*, even if pathological. After 150 years we must conclude that this program has clearly failed: the definitions of those operations, however they are generalized, always exclude most of the functions (whether pathological and not).

In fact, physicists and engineers work with a narrow sub-set of functions, continuous with first (and sometimes second) derivative continuous (it should be noted that this class is not closed under differentiation and is therefore mathematically ill-defined).

We have therefore to ask ourselves why even today teaching leads the student to understand Analysis in terms of its old nineteenth-century claim: to define the two basic operations in such a way as to achieve complete generality in the application to any function. This program is also ineffective with respect to the totality of the functions that the student encounters; moreover the resulting teaching approach is complicated; it is linked to one particular Mathematics among different possible ones and above all it uses AI without declaring it. The most natural idea would be to abandon that old unfinished programme.

The simplest and most natural methodological proposal is to base analysis on an a priori restricted class of functions, the most useful ones. One might consider the class of *functions that are continuous*, or comprised of continuous sub-functions which is the class which mathematicians worked with until 1800 and which still covers most of the requirements of Physics. (In the history of science this proposal corresponds to the founding of Analysis by L. Carnot). Today it corresponds to *constructive mathematics* (which moreover succeeds in obtaining all the regular solutions of differential equations).

Note that with this foundation we remain within calculability, without metaphysical leaps to actual infinity or to its surrogate concepts. Moreover, problems of undecidability are dealt with and this constitutes a considerable cultural novelty. We can explain much of the theory of computers (the halting theorem, ...) with them, whereas these problems are lost in all the most powerful types of Mathematics.

However, with constructive mathematics, we obtain only approximations of the extremum points. To overcome this limitation and to enable students to acquire the use of AI (in history it was nevertheless decisive for the development of Science), we might base ourselves on the Mathematics of Cavalieri and Torricelli, which presents AI in the simplest and most intuitive way possible, on the basis of the experience of a thousand-year-old mathematical theory such as Geometry.

A first proposal is therefore to base the teaching of higher mathematics on geometric intuition, an extremely useful mental faculty that every student learns to develop in school geometry courses, but which is usually only used implicitly.[7]

Moreover, in this methodological choice we can add the history of this mathematical theory as well as a tribute to those two mathematicians, now unjustly ignored.

Finally, recognizing the value of geometric intuition would have the additional advantage of reaffirming that pedagogical attitude that was characteristic of teaching of Mathematics in Italy (and in particular of Geometry) compared to the usual perspectives in other Countries (rationalist in France, empiricist in England, practical in Germany).

Now let us retrace the methodology of the teaching of analysis. Here following the method of Cavalieri and Torricelli, there are several proposals reinterpretable with Weyl Mathematics. Let us start with the definition of a real number. We discover with surprise that Dedekind's well-known definition is exactly what Cavalieri was able to propose and Weyl's Mathematics now proposes: the existence of a point of separation in an ordered series of rational numbers. The said definition loses therefore the mysterious aspect of a statement about an infinity, because its explanation emerges from its personal and historical origin (intuition) and its rationality (existential quantifier).

We could then proceed with the definition of the operations of differentiation and integration. In fact, all mathematicians before Leibniz and Newton began with the second operation. It can be argued that this is more intuitive and greatly simplifies understanding of

---

7 A. Drago, "Una formulazione alla Cavalieri della analisi matematica: una didattica all'italiana, basata sull'intuizione geometrica", in L. Corso et al. (eds.): *Matematica é/e Cultura*, Bologna: Pitagora, 1999, 137-148.

those operations if one starts with the calculation of areas, (as Cavalieri and Torricelli did and as it is done in Thermodynamics, which uses the dQ /T integral first), rather than that of tangents. But to conform to the more modern custom we also begin with the operation of differentiation.

In Weyl's mathematics we can assert the existence of the limit point, thus ensuring that maxima and minima on a curve can be obtained, as well as conferring on them the geometric intuition of the specific points that represent them. Moreover it can be stated with certainty that the incremental ratios give intervals whose limit point exists, in that it can be intuited geometrically and identified the derivative to the curve at this point. This brings us to the concept of derivative of a function.

The fundamental theorems of analysis can be demonstrated (today we know: rigorously) and without difficulty with this type of intuition. Cavalieri's famous theorem can be demonstrated. The theorem of *de l'Hopital*: $\lim f(x)/g(x)=\lim f'(x)/g'(x)$, is demonstrated as a result.

In Cavalieri's intuitive terms the concept of integral is the simplest one; the *omnes lineae* are very intuitive for the students and can be understood correctly in accordance with a concept of mathematical Logic, without any difficulty. The idea of indefinite integration was obtained for the first time by Torricelli as was the aforementioned inverse theorem. The areas they calculated on the most varied curves (such as spirals) is a fascinating subject of study. In what is proposed here, geometric intuition turns the tables on the tradition of Analysis.

But there is also another *proposal, closer to the tradition of Cauchy's "rigorous" analysis*. One can have recourse to a foundation based on the concept of limit, always to be defined with the ε-δ technique. It should be noted that this definitional technique does not characterize rigorous mathematics, because it simply expresses an approximation process. It can be made constructive by requiring that all approximation numbers are rational or decidable and that δ is actually calculated. However, in constructive mathematics this technique generally does not obtain a limit point, but only an approximation interval. By means of the existential quantifier of Weyl's mathematics and/or Cavalieri's geometric intuition we can, on the other hand, state that this limit point exists (whereas in classical mathematics, or rigorous mathematics, it is stated as an idealized value, and moreover that mathematics does not specify the sequence of numbers, which can also be composed of non-calculable real numbers, and conceives as real the final point which instead can only be hypothesized, without being calculable). With this concept of the Weyl limit it is possible to retrace the whole of normal Analysis.

According to recent studies of mathematical logic (which go under the name of *Reverse mathematics*[8]), this Mathematics is able to demonstrate again rigorously all the usual theorems, including that of Heine-Borel, that of Dini on the inverse function and that of Peano on the solution of the first degree differential equation. In other words, this is all compatible with Cauchy's usual rigorous analysis, but only the Zermelo axiom, or equally powerful

---

8 S. G. Simpson, *Subsystems of Second-Order Arithmetic*, Berlin: Springer, 1999. A.G. Marcone, "Equivalenze tra teoremi: il programma di ricerca della reverse mathematics", *La Matematica nella Società e nella Cultura*, Bologna, 2 (2009)1,101-126. See also the table in K. Tanaka, "Weak axioms of determinacy and subsystem of Analysis Π (Σ_0^2 games)", *Ann. Pure Applied Logic*, **52** (1991), 181- 193.

axioms, are excluded.

By explicating the concepts of derivative and integral with facility, their use is greatly extended and made familiar to the student, in such a way that the correspondences between the most common functions (polynomial, trigonometric, etc.) and their derivatives and integrals (this is also done at University) are demonstrated. With these definitions we can advance further. We are assured by the mathematics of Cavalieri-Weyl that our geometrical intuition is not in error when it suggests the solution of a differential equation. We can then propose solve the easiest differential equations, which are very commonly used and fundamental for the study of Physics and Biology. e.g. $f'(x) = f(x)$, which can be solved as

$$dx = df/f\,; \qquad \int dx = \int df/f; \qquad x = \log f(x) + c; \qquad f(x) = kA^x;$$

where the solution of the integral $\int df/f$ can be given, which is neither more nor less than how it is done at University: as a correspondence with the function $x = \log f(x) + c$, verifiable by replacing f in the differential equation; the second differential equation, $f''(x) = f(x)$, which gives trigonometric expressions (real or complex), can be studied in the same way.

Ultimately, with the teaching of this Mathematics, the time and effort devoted to formalisms, the reasons for which today remain obscure to the student, are reduced and the cultural value of mathematics is enhanced. Note in particular that a new teaching of Physics is also possible with this Italian manner of teaching Analysis. Because the *principle of inertia* can be taught very simply in Cavalieri's or Torricelli's version. In this way the teacher can accurately establish *the moment at which actual infinity entered theoretical physics*, occurring not in a lateral manner (like one of the values of the variables in the formulas of geometric optics), but in the structure of the theory, through the principle of inertia: 1632.

Thus the concise words of Koyré acquire full meaning.

> Galilei explains the real with the ideal [i.e. with a reachable limit; e.g. the absence of friction] ... [Cavalieri, Torricelli,] Descartes and Newton explain the real with the impossible [non-operational abstractions] ... Galilei does not.[9]

7.4 -Consequences for (the (teaching of) Philosophy of Science: Leibniz

Let us now outline briefly the history of Western philosophy, which should have understood and perhaps prepared the emergence of modern science and then its development.

In fact, Leibniz had already come close to a conception of the Foundations of Science, declaring that human reason encounters two labyrinths: one concerns the infinite (actual and potential) and corresponds precisely to the dichotomy between AI and PI. The other labyrinth concerns Law and Freedom; with these terms Leibniz expressed the subjective perception of the two possible types of organization (moreover, theoretical): it is then a matter of choosing either the organization that involves mandatory rules (of deduction) or the one that allows an unrestricted search for the solution of a (theoretical) problem. They correspond to the two possibilities of AO and PO.

---

9 A. Koyré, *Etudes Galiléennes* (orig. 1966), Paris : Hermann, 1986, p. 276.

Leibniz defines the second labyrinth less well. We note, however, that he also stressed that *there are two principles of human reason*: that of non-contradiction and that of sufficient reason. We have already seen that the second principle is expressed by a doubly negated sentence of non-classical logic; these two principles say therefore that human reason can follow one of two logics, thereby also hinting at a formal definition of the above poorly defined dichotomy.

Although he had planned to express the Foundations of Science as a "science of sciences", he did not go beyond declaring these two labyrinths and these two principles and did not indicate how to choose between them. However we can conclude that Leibniz *had intuited the Foundations of Science*.

7.5 -Consequences for (the (teaching of) Philosophy of Science: Kant:

Let us then draw the conclusions of this for the (teaching of) Kant’s philosophy.

Philosophical textbooks present it as the summit of Western philosophy, because he reconciled the two currents that emerged and developed in opposition to each other, Empiricism and the Rationalism of innate ideas. But *Kant’s interpretation of the Foundations of Science was erroneous.*

The basis of his system is the "demonstration" of the idea of several earlier philosophers that the human mind cannot achieve knowledge of the "thing in itself". He presents four empirical theses and, in contraposition, the corresponding four rationalist antitheses. He thus obtains four dilemmas. It should be noted that the first two concern infinity, either only potential or only actual; while the second two concern the organization (of the world), either necessary or free. They therefore follow Leibniz’s two labyrinths, except that each of these is applied to two cases of reality. Ultimately, the four Kantian dilemmas reproduce the two dichotomies, but in the form of a thesis.

KANT’S ANTINOMIES, LEIBNIZ’S TWO LABYRINTHS, THE TWO DICHOTOMIES

	‘Mathematical antinomies’		*‘Dialectical antinomies’*	
Kant’s four dilemmas	Finite or infinite cosmos, in time and in space	Divisibility of matter: finite or infinite	Law or freedom	Esistence or not of a necessary Entity
Leibniz’s two labyrinths	Actual or potential infinity		Law or freedom	
The two dichotomies	*Infinity*		*Organization*	
	Actual or potential		either deductive or problem-based	

For every thesis T he gives a theorem that proves it and likewise for every corresponding antithesis A.

It is very important to note that all the proofs are *reductio ad absurdum*, so each of them concludes with a doubly negated proposition, i.e. $\neg\neg T$ for the thesis and $\neg\neg A$ for the antithesis. It should also be noted that Kant, in agreement with his times, thought that the only logic was classical logic, which makes it possible always to change a doubly negated sentence into its affirmative. Therefore at the conclusion of the *reductio ad absurdum* proof of the empiricist thesis $T = \neg A$, he uses classical logic and has $\neg\neg T = T$ which however through Kant's construction of the theses and the antitheses is $= \neg A$. He then concludes that the four dilemmas had thereby become four contradictions (these are called "antinomies" when they are inherent in the system of thought) between each T and the corresponding antithesis A.

However, today we know that there are two types of logic: classical and non-classical. We have therefore, at the end of the *reductio ad absurdum* proof of the empiricist thesis T, which uses non-classical logic, a double final negation $\neg\neg T$, which cannot be changed into the affirmative. The two conclusions of the *reductio ad absurdum* proofs, of the empirical thesis $\neg\neg T$ and of the antithesis $A = \neg T$, differ from each other by a negation, but do not give a contradiction, because classical negation is different from non-classical negation; that is, the empiricist conclusion is not the (classical) negation of the rationalist conclusion; rather, there is no possible comparison between the two conclusions. According to the two types of logic, Kant's two proofs for each dilemma represent two ways of studying a problem: the inductive one of empirical philosophy, the deductive one of rationalist philosophy.

We conclude *that the contents of Kant's two dilemmas on infinity revisit the labyrinth concerning the two types of infinities, while the two other dilemma revisit~~s~~ the other labyrinth, relating to organization. Kant's theorems, rather than giving antinomies, manifest the incommensurability between the two types of mathematical Logic, namely Leibniz's two principles of human reason.*

Let us now consider the other glory attributed by textbooks to Kant's philosophy of knowledge: judgments. Let us compare them with the two dichotomies.

From Kant's definitions it is quite clear that "analytic" corresponds to AO and that "synthetic" corresponds (also due to its stated opposition to the first) to PO. To clarify the ill-defined 'a priori' and 'a posteriori' let us refer to the mathematics of his time. The first clearly represents the infinitesimal, which, not having at the time a mathematical definition, was given a priori; and the second represents the obtaining of a result (for example with a ruler and a compass)).

The following table compares Kant's concepts with the choices of the two dichotomies and with the corresponding theories representing the four models of scientific theory:

KANT JUDGEMENTS AND THE FOUR PAIRS OF CHOICES

Judgements according to Kant	**Pairs of choices**	**Corresponding physical theories**
Rationalism: A priori analytic	AO&AI	Newton's mechanics
Empiricism: A posteriori synthetic	PO&AO	Thermodynamics and Chemistry
Kantian 'Invention': A priori synthetic	PO&AI	Lagrange's mechanics
Lacking in Kant: A posteriori analytic	AO&PI:	Descartes' geometric optics

It can be seen that Kant misinterprets the judgments. He leaves one aside (the last one in the table), the one linked to geometric optics. This is explained by the fact that that theory no longer existed at his time, since Newton had included the theory in his Mechanics. Furthermore, Kant rightly according to the idealism of his time, attributes an idealist (non empirical) foundation to Newton's Mechanics (let us recall absolute space, absolute time and force conceived as a metaphysical cause). Moreover Kant considers the a priori synthetic judgment, but believes that empiricism is incapable of formulating theories. Yet Thermodynamics and Chemistry succeeded in theorizing concerning respectively all thermal machines and all chemical reactions. Finally Kant believes he has invented a new judgment, which echoes the most advanced mechanical theory of time, Lagrange's (based on infinitesimal analysis, IA, and the resolution of an extremum problem, OP).

Globally, he gives importance, on the one hand, to Newton's mechanics (deriving from it the a priori transcendental categories of space, time and causality) and on the other, implicitly, to the theory that will dominate theoretical physics in the nineteenth century, a period in which the previous theory will prove increasingly limited and open to criticism. This merely echoing of the main line of progress in theoretical physics makes him appear one of the most advanced philosophers of knowledge.[10]

7.6 -Consequences for (the teaching of) Philosophy of Science in general:

Let us obtain some general consequences for the teaching of philosophy.

Kant's Philosophy generated four paradoxes in the Philosophy of Western Science:

1st paradox: the current unjustified philosophical dominance of Kantianism in the Philosophy of Science.

> Despite a fundamental lack of clarity regarding the epistemology of his 'critical philosophy' and the profound attacks by idealists, realists and naturalists, Kantianism still influences almost all of the current (non-positivist) Philosophy of Science, as well as the formalist and intuitionist Philosophies of Mathematics.[11]

We know that Kant believed that Aristotelian logic had no possibility of further development and he knew little about infinitesimal analysis, despite the fact that he lived a century after its invention. His conception of knowledge was disavowed by subsequent developments in: Logic, Geometry, Physics and Chemistry. For these reasons and for what we have seen in the previous paragraph his contribution to the philosophy of knowledge should be drastically re-evaluated.

2nd paradox: the alleged "Kantian reconciliation" was followed by divergences among philosophers and among scientists, and also between the two groups.

Several scientists (for example, Comte, Mach, Enriques, ...) constructed "spontaneous" philosophies of knowledge, but none of them confirmed Kantianism or found agreement amongst other philosophers. Moreover, the developments in science, which had already moved away from Kantiansm in the nineteenth century, diverged even further from it in the

---

[10] Ernst Mach was an exception. He did not, however, in the end diverge from the theoretical tradition. While criticizing in depth the Newtonian concepts of force, absolute space and absolute time, he did not recognize any theory as an alternative to Newtonian Mechanics: neither his outline of a theory reconstructed on his own criticisms of Newton, nor that of Lagrange, which he considered simply more effective than that of Newton in solving problems. E. Mach: The Science of Mechanics (1883), Open Court, Lasalle, pt. IV, chp. III.

[11] R. Bhaskar, "Kant's philosophy of science", in W.F. Bynum, E.J. Brone, R. Porter (eds.), *Dictionary of the History of Science*, London: Macmillan, 1981, p. 223.

twentieth century (think of the unprecedented philosophical problems of quantum mechanics). (The last scientist to connect with Kant was Brouwer, who in 1905 founded mathematical intuitionism on time understood as an a priori transcendental category; but no scientist followed him in this foundation).

3° paradox; *the devaluation of the philosophy of Leibniz, a profound philosopher and highy advanced scientist.*

Leibniz was *the last philosopher who knew and extended all formal science* (Logic, Geometry, Mathematics, and Mechanics), initiating that alternative to Newton's, later completed by L. Carnot[12]. He was criticized radically by Kant who moreover knew him principally through his follower Wolff. He did not know all of Leibniz's writings and already interpreted them idiosyncratically, for example reducing the principle of sufficient reason to a statement of classical logic. Furthermore, Leibniz was ridiculed as a *pig-headed metaphysician* (in particular by the non-scientist and non-philosopher, which is what Voltaire was!), Although he always presented Physics and Metaphysics as two parallel levels, of which the first is based on the <u>impossibility</u> of "motion which does <u>not</u> end" (and therefore on non-classical Logic, PO) and on operational principles (PI), without Infinitesimal Analysis (ai). Ultimately, the last true philosopher-scientist was disparaged with facile prejudice.

4° paradox: *incommensurability, understood by Leibniz in terms of the idea of labyrinths, was distorted by Kant, who also claimed to have formalized them; rather he taken them down a blind alley*.

With the idea of the labyrinths Leibniz had perceived the incommensurability between theories that are different in the types of logic and / or types of mathematics they employ. Kant has the merit of having formalized them logically, but he then misinterpreted their meaning because he believed that only classical logic existed. He thus concluded that he had found contradictions that constituted an insurmontable barrier to the mind. He then sought to elevate to preconditions of human knowledge transcendental categories), which did not correspond to the subsequent developments of science. The misunderstanding has not been corrected, despite the many reformulators of Kant's ideas (Cassirer, Cohen, etc.). As a result modern Western philosophy is marked by having encountered two labyrinths, i.e. incommensurability, and the inability to deal with them. It has already been said that while the ancient Greeks consciously halted before their incommensurabilities, two thousand years later Western philosophers came up against the new incommensurabilities without even recognizing them as such.

Today therefore a radical change is needed in the way we consider the history of the philosophy of knowledge and this begins with a re-evaluation of the merits of what Leibniz suggested. And with regard to the teaching of Philosophy of Science we can conclude by stating that *the current* approach of the teaching of modern Philosophy, which devalues the scientist Leibniz with facile prejudices, considers its history as a progression centred on the Kantian system, whose idealism (of the transcendental categories) prepares the philosophies of successive idealists (in particular the philosophy of Hegel, considered by most textbooks as an advance towards the future). It is necessary to abandon this conception of the history of Philosophy, in order to reclaim its relationship with the actual historical development of scientific knowledge in accordance with Leibniz' ideas concerning its Foundations. With this

12 L. Carnot. *Essai sur les Machines en général*, Defay, Dijon, 1873.

reform of the history of Philosophy, it will thus be possible for the Philosophy teacher to have a well-defined relationship of collaboration with teachers of the sciences.

APPENDIX

CHRONOLOGICAL LIST OF THE FIVE MAIN GROUP OF SCIENTISTS
WITH THEIR DATES OF BIRTH AND DEATH

Ancient period PYTHAGORAS (VI B.C.) ZENON (V B.C.) PLATO (429-348 B.C.) ARISTOTLE (384-322 B.C.) EUCLID (∼300 B.C.) ARCHIMEDES (287-212 B.C.)	*Birth of modern science* COPERNICO (1473-1543) GALILEI (1564-1642) DESCARTES (1596-1650) HUYGENS (1629-1695) NEWTON (1642-1727) LEIBNIZ (1646-1717) BERKELEY (1685-1753) EULER (1707-1783)	
Around the birth of French revolution D'ALEMBERT (1717-1783) COULOMB (1736-1806) LAGRANGE (1738-1816) LAVOISIER (1743-1794) LAPLACE (1749-1827) L. CARNOT (1752-1822) DALTON (1766-1844) AMPERE (1775-1836) GAUSS (1777-1855) CAUCHY (1789-1857) FRESNEL (1788-1827) FARADAY (1791-1867) LOBACEVSKY (1792-1856) S. CARNOT (1797-1832) GALOIS (1803-1834) HAMILTON (1806-1865)	*Second half of 19th Century* WEIERSTRASS (1815-1897) HELMHOLTZ (1821-1894) CLAUSIUS (1822-1888) KELVIN (1824-1907) RIEMANN (1826-1866) DEDEKIND (1831-1916) MAXWELL 1831-1879) MENDELEIEFF (1834-1907) MACH (1838-1916) BOLTZMANN (1844-1906) CANTOR (1845-1918) POINCARE' (1854-1912) HERTZ (1857-1894) LORENTZ (1853-1928)	*First half oh 20th Century* PLANCK (1858-1947) HILBERT (1862-1943) RUSSELL (1872-1970) EINSTEIN (1879-1955) BROUWER (1881-1966) SCHROEDINGER (1887-1961) WEYL (1885-1955) HEISENBERG (1901-1976) von NEUMANN (1903-1957)