ÉVORA STUDIES IN THE PHILOSOPHY AND HISTORY OF SCIENCE

Volume 1

IN MEMORIAM HERMÍNIO MARTINS

calei
dosc
ópio

TÍTULO
Évora Studies in the Philosophy and History of Science
Volume 1 – In Memorian Hermínio Martins

EDITOR
João Príncipe

CO-EDITOR
Jorge Leandro Rosa

DESIGN E PAGINAÇÃO
Nuno Pacheco Silva

ISBN
978-989-658-

DEPÓSITO LEGAL
______/15

EDIÇÃO

calei
d○sc
ópio

CALEIDOSCÓPIO – EDIÇÃO E ARTES GRÁFICAS, SA
RUA DE ESTRASBURGO, 26, R/C DTO.
2605-756 CASAL DE CAMBRA
TELEF. (+351) 21 981 79 60
FAX (+351) 21 981 79 55
caleidoscopio@caleidoscopio.pt
www.caleidoscopio.pt

CONTENTS

WHY SOME PHYSICAL THEORIES SHOULD NEVER DIE

Olivier Darrigol

With a proper dose of amused aloofness, Hermínio Martins recently reported that a few actors of the "Big Data" movement had declared the end of theory. In their view, the collection of a vast amount of data on everything you may think about, their statistical inductive analysis by computer means, and proper imaging techniques should allow us to predict any phenomenon of interest to our satisfaction. As Martins perceives, what is here at stake is the notion of explanatory depth: How much of the world do we understand if we can predict without theory? Can efficient prediction truly occur without a prior idea of the comprehensibility of the world?[1]

The recent pronouncements of the end of theory may perhaps be seen as the natural conclusion of a historical process that began with the demise of Cartesian rationalism. Schematically, the Newtonian denial of the rational necessity of René Descartes's world led to a moderate empiricism in which rational a priori arguments still had a say. In particular, Immanuel Kant's followers assumed that Euclidean space, absolute time, and Newton's laws of mechanics were necessary conditions of experience and were therefore immune to experimental refutation. In the early twentieth century, Einstein's relativity theory discredited this non-metaphysical sort of rationalism. From then on, empiricism has reigned in various forms (logical empiricism, linguistic and semantic approaches to physical theory, etc.), although there were a few important attempts to adjust Kant's doctrine to the new states of affairs. Among the scientists themselves, it is often heard

1. H. Martins, "Images and Imaging in science", in this volume, pp. 128-129.

that theories come and go, that they have only limited value, and are nothing but instruments of research. In the same vein, today's theorists tend to avoid the word "theory" and to describe their work as "model" making. The Big Data enthusiasts go only one step further in the same direction by entrusting their computers with implicit modeling by some sort of Bayesian adjustment of computational algorithms to the empirical world.

Should we agree that Big Data extremism completely replaces explanation with prediction? Maybe not. This vision could indeed be compared to an impoverished sort of Kantianism, with its own conditions for the possibility of experience. Remember that for Kant, these conditions imply a representative faculty of the mind, which he calls intuition, and a legislative faculty, which he calls the understanding. Intuition involves space as the external form of sensibility and time as the internal form of sensibility. The understanding involves a number of categories including substance, cause, quantity, and quality. Experience usually involves the application of the categories to representations of phenomena in our intuition. This application requires a way of representing the categories themselves in our intuition, which Kant calls schematism. Similarly, the Big Data vision involves a representative faculty called imaging, and a legislative faculty called computing. The digitalization of images and the graphic representation of computing results play the role of Kant's schematism.

This admittedly vague analogy between Kant's transcendental philosophy and the Big Data philosophy (if we may call it so) suggests a very simple and greatly stable theory of explanation: something is understood when it can be expressed through images and computing algorithms, and the rest is psychology. The problem with that theory of explanation is its extreme poverty: It does not include some of the most commonplace modes of explanation; it does not articulate its own variety of schematism; it cannot really do without our best received theories and it does not care to explain the success of these theories. The two first points have frequently been made by critiques of Big Data extremism and they need not be rehearsed here. Let me only recall that the proper use of imaging and the proper association of an algorithm with an external process requires all sorts of structural, cognitive considerations that are not expressed in the simple data-images credo; and the success of Big Data analysis crucially depends on these considerations.

My main concern, in the present essay, is to develop the third kind of criticism of the Big Data philosophy: that it cannot do without our best physical theories. If the ambitions of Big Data analysis were limited to predicting which will be the next most popular movie it could perhaps stick to its rudimentary epistemology. These ambitions are much higher, however. The Big Data fans want to send any good old theoretical analysis to the dustbins of history. For instance, we would not need fluid mechanics to predict the behavior of a plane in the air; we would just have to program computers to establish the correlations between a huge amount of data regarding observed flows at different moment in different circumstances. The Dataists would presumably admit that today's aeronautic engineers still rely on fluid mechanics. But they would claim that this is not necessary: in their view, any reliance on theory should ultimately be avoided since history has taught us that any theory someday becomes obsolete.

This alleged obsolescence of all of our former theories is only a fiction of a naive history of science. Deeper history involves the ways in which some very old theories still play a role in today's science, and it points to the rationality of perpetually relying on some of our best theories. In the late nineteenth-century, there already was a skeptical, anti-theory current that fed on the perishability of all theories. Henri Poincaré famously responded that our best theories never quite perished because their success implied their containing true relations (*rapports vrais*) that would remain (approximately) valid in any future theory. Although the ontology with which we originally dressed a theory may not resist the erosion of time, the theory contains a structure that will forever remain valid in the experimental domain in which the predictions of the theory have been verified. For example, a structure extracted from Newton's mechanics is still abundantly used by physics and engineers to describe the motion of macroscopic bodies, even though Newton's underlying concept of space and time are now obsolete.[2]

This is not the only way old theories survive alleged scientific revolutions. They become parts of the newer theories, which I call modules. The latter process has long been underappreciated, for the following reasons. Whenever physicists adopt a new theory and compare it to its predecessors, they tend to emphasize its higher unity, its broader scope, and its finer

2. H. Poincaré, *La science et l'hypothèse* (Paris, 1902), Chap. 10. Cf. Darrigol, «Diversité et harmonie de la physique mathématique dans les préfaces de Henri Poincaré,» in Jean-Claude Pont et al. (eds.), *Pour comprendre le XIX*[e] *: Histoire et philosophie des sciences à la fin du siècle* (Florence: Olschi, 2007), 221–240; João Príncipe, «Sources et nature de la philosophie de la physique d'Henri Poincaré,» *Philosophia scientiae*, 16 (2012), 197–222.

resolution. These merits concern the theory as a whole, not any substructure of it. Philosophers of science also favor a holistic view of (physical) theory, for various reasons. One is Duhemian holism, which denies the possibility of separately testing the various components of a theory as a consequence of the theory-ladenness of experiment. Another is Quinean holism, which unfolds consequences of Willard Van Orman Quine's demolishment of the distinction between analytic and synthetic judgments. Still another is the now dominant view of physical theories: the semantic view according to which they are families of mathematical models (set-theoretical constructs) enjoying a partial isomorphism with concrete devices and processes. In this last view, the holism comes from the vagueness of the "isomorphism" between the formal and the concrete, which falls from a cognitive heaven without implying any substructure of the theory.[3]

The holistic view of physical theory does not resist careful inspection of the way theories actually function. Every advanced theory essentially depends on internal or external links with other theories that have a different (usually smaller) domain of application. This is what I call the modular structure of physical theory. This structure turns out to play an essential role in the construction, application, comparison, and communication of theories. In particular, the concrete application of new theories essentially depends on their modular connection with earlier theories that we already know how to apply. This view is as remote as we could imagine from the idea that any theory is doomed to perish. The Big Data anti-theory sect could still argue that they do not need intertheoretical relations since they do not need theories at all. In their view, by renouncing theories one renounces only a mode of explanation of nature, not the prediction of its behavior. Against this view, it will be argued that the arguments for the necessity of a persisting modular function of older theories derive from pragmatic criteria of predictive efficiency and not from any preconceived idea of explanation.[4]

That is not to say that the concept of explanation is useless. That is just

3. On the linguistic and semantic views, cf., e. g., Frederick Suppe, *The structure of scientific theories* (Urbana, 1974); Marion Vorms, *Qu'est-ce qu'une théorie scientifique?* (Paris, 2011). The structuralist variety of the semantic view, as introduced by Joseph Sneed and his followers, brings some useful substructure and has similarities with the modular view proposed below. On Sneedian structuralism, cf. Roberto Torretti, *Creative understanding: Philosophical reflections on physics* (Chicago, 1980), Chap. 3.
4. Cf. Darrigol, "The modular structure of physical theories," *Synthese*, 162 (2008), 195–223.

to say that pragmatic efficiency and explanatory depth are not as distinct notions as we would spontaneously think. For instance, a theory that has a simpler, more unified structure better explains the world in a Kantian regulative sense; and it is also a more powerful theory in practice because it reduces the number and variety of steps required for effective prediction. There is more. We may define broad conditions of comprehensibility of the world, for instance causality, space and time measurement, or decoupling of scales, and investigate the consequences of these conditions. It turns out that some of our most fundamental theories, mechanics, thermodynamics, electrodynamics, general relativity, and quantum mechanics can be derived in this manner. The possibility of this sort of arguments depends on the modular structure of physical theory, which enable us to conceive the ideal measurements and the intertheoretical relations on which they rely.[5]

Thus a proper notion of the comprehensibility of the world enables us to severely restrict the spectrum of possible theories. This is pragmatically important not only because imperfect comprehensibility arguments have had heuristic import in a few historical cases, but also because in a perfected form they tell us which theories are likely to forever persist as modules of future theories and they also give us a strategy for devising new theories by relaxing some of the conditions of comprehensibility. The comprehensibility criteria are relative to a limited domain of investigation. No matter how natural they may seem at a given time of history, they typically fail when the domain of investigation is excessively extended. Consequently, comprehensibility arguments are not purely rational arguments. They rely on very broad, refutable assumptions on how we may comprehend the world. Their force does not rest on absolute necessity but on the simplicity and naturalness of the premises. Here we seem to be regressing to some sort of psychologism. We are not in reality because the premises have to do with our basic ability to live, predict, and act consistently in this world.

The present essay is an attempt to justify the necessity of physical theories in a twofold manner: by showing the pragmatic necessity of old and new theories on the basis of their modular structure, and by showing the derivability of some of our most powerful theories from simple assumptions about the comprehensibility of the world. My arguments presuppose

5. Cf. Darrigol, *Physics and necessity: Rationalist pursuits from the Cartesian past to the quantum present* (Oxford, 2014).

a general definition of a physical theory that integrates the way physicists actually conceive the application of a theory. This will be given in section 1. Section 2 defines modules and modular relations, and gives a few examples. Section 3 shows how the definitions of the two former sections enable us to conceive comprehensibility arguments for the necessity of some theories. The implied comprehensibility conditions share some characteristics of the relativized a priori that a few neo-Kantian philosophers have introduced since the beginning of the twentieth century. A comparison between these two varieties of attenuated rationalism is given in section 4, written with the benefit of Thomas Ryckman's insightful discussion of Hans Reichenbach's and Ernst Cassirer's neo-Kantian reading of general relativity, and in reaction to Michael Friedman's stimulating revival of Reichenbach's constitutive principles. In conclusion, I wrap up the arguments in favor of the indefinite persistency of a few of our best theories, against the contingentist, relativist, and anti-intellectualist views of which the Big Data ideology is an extreme case.[6]

THEORIES DEFINED

Not much can be said on the nature, role, and necessity of physical theories without a sufficiently precise generic definition. Two possibilities come to mind: to borrow the now typical physics textbook definition of a theory as a mathematical formalism equipped with rules of interpretation, or to rely on today's philosophers preferred definition of theories as a family of models (set-theoretical constructs) endomorphic to concrete physical systems or processes. This will not do because these definitions do not give sufficient guidance in the conception of experiments to which the theory can be applied. The semanticists' notion of endomorphism says strictly nothing in this regard, and the rules of interpretation found in physics textbooks are only a fragment of what is needed to concretely apply the theory. Typically, physics students learn how to apply a theory by being taught a series of paradigmatic applications, both *in abstracto* and in the laboratory. The following definition has been obtained by inspecting the most important physical

6. References to secondary literature are here reduced to a minimum. More relevant bibliography is cited in Darrigol, refs. 4–5.

theories and the ways in which they are concretely applied.

A physical theory is defined by four components:

a) a *symbolic universe* in which systems, states, transformations, and evolutions are defined by means of various magnitudes based on powers of **R** (or **C**) and on derived functional spaces and algebras.

b) *theoretical laws* that restrict the behavior of systems in the symbolic universe.

c) *interpretive schemes* that relate the symbolic universe to idealized experiments.

d) methods of *approximation* that enable us to derive the consequences that the theoretical laws have on the interpretive schemes.

In order to flesh out this definition, let us consider the simple example of the classical mechanics of a system of a finite number of mass points. The symbolic universe consists in systems defined by a number N of mass points, in states defined by the spatial configuration of the particles (N vectors in a three-dimensional Euclidean space), in evolutions that give this configuration as a twice differentiable function of a real time parameter, in the list of masses of the particles (N real positive constants), and in an unspecified force function that determines the forces acting on the particles for a given configuration and a given time. The basic law is Newton's law relating force, mass, and acceleration. Interpretive schemes may vary. A first possibility is that the scheme consists in a given system (choice of N) and the description of ideal procedures for measuring spatial configuration, time, forces, and masses. Then an idealized experiment may consist in the verification of the motion predicted by the theory for given initial positions and velocities and for a properly selected frame of reference. Or the scheme may involve mass, position, and time measurement only, allowing idealized experiments in which the forces are determined as a function of the configuration. Or else, the scheme may involve mass, position, and time measurements and the choice of a specific force function, allowing idealized experiments in which the motion predicted by the theory is verified for given initial conditions.

This first example suggests a more precise definition of an interpretive scheme as the choice of a given system in the symbolic universe together with a list of characteristic quantities that satisfy the following three properties:

1) They are selected among or derived from the (symbolic) quantities that are attached to this system.
2) At least for some of them, ideal measuring procedures are known.
3) The laws of the symbolic universe imply relations among them.

The characteristic quantities of a given interpretive scheme are divided into measured quantities and more theoretical quantities. Only the former quantities are measured in experiments based on the scheme. The theoretical quantities either are the unknowns that the experiments aim to determine, or their value is taken from empirical laws established by preliminary experiments.

Before further comment, let us consider the more difficult example of quantum mechanics. There the symbolic universe involves a Hilbert space of infinitely many dimensions, operators representing physical quantities, and a few real-number parameters such as time, mass, charge, and external fields. The two basic laws are Schrödinger's equation and the law giving the statistical distribution of a given quantity for a given state. Interpretive schemes involve the various quantities attached to the particles and fields and the parameters. The laws of the symbolic universe imply statistical correlations between the quantities for given values of the parameters. The complexity of the symbolic universe and of the interpretive schemes varies with the type of system considered (single particle in external fields, several interacting particles, quantum fields).

In this example, it is obvious that the interpretive schemes do not spontaneously derive from the symbolic universe because there is no direct correspondence between the symbolic state vectors and the measured quantities. In the classical case, one may be tempted to believe the contrary because one can easily imagine an approximate concrete counterpart of the symbolic universe. This would be a mistake, because the symbolic quantities never have a direct concrete counterpart. Their concrete implementation requires ideal measurement procedures that are not completely definable within the symbolic universe of the theory. Most mechanical experiments require position and time measurements, and these require, besides the laws of the symbolic universe, a notion of inertial frame and the means to concretely realize length measurements.

In general, the set of interpretive schemes associated with a theory varies in the course of time. Some schemes are there from the beginning of the theory, as they are associated with its invention. Others came at later stages of the evolution of the theory when it is applied in a more precise or a more extensive manner. In this process, some purely symbolic quantities may be promoted to the schematic level. For instance, late nineteenth-century studies of gas discharge and cathode rays provided experimental access to the invisible motions assumed in the electron theories of Hendrik Lorentz, Joseph Larmor, and Emil Wiechert.

A last remark on the present definition of theories is that all the structures it employs are defined mathematically. In this respect, it agrees with the semantic view. The main difference is that it contains evolving substructures, the interpretive schemes, that enable us to conceive blueprints of concrete experiments. In a vague way, we may understand this power of the schemes as a consequence of their being generated all along the history of applications of the theory. But we are still in the dark regarding the precise way in which physicists articulate the relation between symbols, schemes, and experiments. This is where the notion of modules become indispensable.

MODULES

Any advanced theory contains or is constitutionally related to other theories with different domains of application. The latter theories are said to be *modules* of the former. Modules occur in the symbolic universe, in the interpretive schemes, and in limits of these schemes. Since by definition they are themselves theories, they also contain modules, submodules, and so forth until the most elementary modules are reached. There are (at least) five sorts of modules. In reductionist theories such as the mechanical ether theories of the nineteenth century, there is a *reducing module* diverted from its original domain to build the symbolic universe of another domain. In many theories, the symbolic universe also appeals to *defining modules* that define some of the basic quantities. For instance, mechanics is a defining module of thermodynamics because it serves to define the basic concepts of pressure and energy. There are *schematic modules* that occur at the level of interpretive schemes and serve to describe the relevant measurements. These may belong to the symbolic universe, as is the case for pressure in the schemes of a thermodynamic gas system; or they may require additional modules as is the case for position and momentum in the schemes of one-particle quantum mechanics. There are *specializing modules* that are exact substitutes of a theory for subclasses of schemes under certain conditions. For instance, electrostatics is a specializing module of electrodynamics. Lastly, there are *approximating modules* that can be obtained by taking the limit of the theory for a given subclass of schemes. For instance, geometrical optics is an approximating module of wave optics. These categories are not mutually exclusive: for example, a schematic module can also be a defining module or an approximating module.

Thus we see that there are diverse ways in which the full exposition of a given theory calls for other theories. My choice of the word "module" is intended to convey metaphorically this diversity as well as the fact that the same theory can be a module of a number of different theories. For instance, classical mechanics is a module of electrodynamics, thermodynamics, quantum mechanics, general relativity, etc.; and it can be so in different ways. At any given time, any non-trivial theory has a modular structure, namely: it includes a number of modules of the above-defined kinds.

The modular structure of a theory is not unique and invariable. It depends on a number of factors: the conception we have of this theory, the type of experience that is conceivable at a given period of time, the degree of elaboration of the theory, etc. As an example of the first factor, for some nineteenth-century physicists mechanics was a reducing module of electrodynamics; for phenomenologists it was only a defining module; for believers in the electromagnetic worldview, it was a schematic module. As an example of the second factor, approximating modules for the description of stochastic processes appeared in statistical mechanics only after the development of relevant experiments. As an example of the third factor, the boundary-layer approximating module of hydrodynamics appeared only at a late stage of its evolution, even though it concerned an old domain of experience.

This ambiguity and variability of modular structure may explain why philosophers of physics have paid little or no attention to it. This structure seems to elude any formal, rigorous epistemology. It seems too fleeting and too vague to embody the epistemic virtues that philosophers wish to find in physical theories. Against these appearances, I will now argue that modular structure is essential to the application of theories, to their comparison, to their construction, and to their communication. These four aspects of theorizing activity will thus appear to be intimately related to each other. Moreover, modular structure will acquire some sort of necessity: without it physical theories would remain paper theories. By extension, we come to the conclusion that no genuine knowledge about the physical world is possible without theories that function as modules for the generator of knowledge, whatever be this generator (human individual, human collective, computer, or hybrid) and however big the data may be.

APPLICATION

The symbolic universe of a theory never applies directly to a concrete situation. The application is mediated through interpretive schemes that describe ideal devices and quantitative properties of these devices. In order to build a concrete counterpart of a scheme, we must know the correspondence between ideal device and real device, as well as concrete operations that yield the measured quantities. In any advanced theory this correspondence obtains in a piecewise manner, through the modules involved in the scheme. In the earlier said metaphor, the schemes are blueprints, and the modules help us select the materials for realizing them. The most superficial observer of a modern test of a theory cannot fail noticing the contrast between the simplicity of the theoretical statement to be tested and the complexity of the experimental setting. What enables physicists to make sense of this complexity is, for the most part, the modular structure of schemes.

The modules enable us to exploit the competence we have already acquired in applying the modular theories. This application may involve sub-modules and their schemes, and so forth until the concrete operations become so basic that their description can be expressed in ordinary language. Take the relatively simple case of mechanics. The schemes involve a geometric module, which one already knows how to realize by means of surveying with rigid rods (for example). This knowledge is essential in building the apparatus and realizing the relevant measurements. Other useful modules may be kinematics and statics.

This is not to say that modules are all we need for the realization of schemes. Non-theoretical knowledge is also needed on the part of the experimenter, and external theories may be involved in the functioning of the measuring apparatus. This complicating circumstance does not make modules less useful. On the contrary, it brings us to appreciate two additional virtues of modules. Firstly, the non-theoretical knowledge implied in the application of a given theory can be exploited in the application of any other theory that contains this theory as a module. Secondly, when two theories share the same module, the applications of one theory may benefit from the other theory in the measurement of modular quantities. For instance, electronics can be used in building galvanometers, relativistic mechanics in building oscilloscopes, and optics in measuring distances.

The modular structure of theories also affects the discussion of their refutability. The freedom in defining the schemes and the tacit knowledge involved in their concrete realization seem to leave plenty of room for protecting theories from refutation. In reality, the modular structure severely limits the protective strategies because it restricts the form of the schemes and because it tends to confine tacit knowledge in the application of well-understood modules. As far as experimental error and reasoning lapses can be avoided, the accommodation of adverse experimental results is made difficult. Surely there still is some sort of Lakatosian protective belt: as long as no better alternative theory is available, physicists prefer to modify the symbolic universe or the non-modular components of the schemes. But the modules themselves usually remain untouched. Duhemian holism, or unrestricted "open-endedness," do not occur in the actual practice of physics. The modular structure of theories conveys to them much more rigidity in their adaptation to the empirical world than some historians and some Big Data enthusiasts would have it.

COMPARISON

The comparison of two theories obviously requires a non-vanishing intersection of their domains of application. In my terminology, this means that the two theories should share the same subset of interpretive schemes. More exactly, the characteristic quantities for a subset of schemes in one theory should be the same for a subset of schemes of the other theory. This can only be the case if the characteristic quantities are defined through modules that belong to both theories. Once this condition is met, the predictions of the two theories are said to agree if and only if the laws of the two theories imply the same relations between the schematic quantities in the compared subsets of schemes. The physicists' practice of comparison always involves schematic quantities defined by shared modules. Radical incommensurability is only a philosophical fiction.

In particular, there are crucial experiments that enable physicists to decide between two competing theories. The crucial character of an experiment requires the sharing of the modules involved in its scheme, as well as the exclusion of ad hoc modifications of the symbolic universe. A famous example is that of the experiment that François Arago performed at the

Paris Academy of Sciences in answer to an objection to Augustin Fresnel's diffraction theory. A supporter of Newton's older theory, Siméon Denis Poisson, had noted that according to Fresnel's theory there should be a bright point of light in the middle of the shadow cast by a disc. The corpuscular theory, even in a version including deflections of the rays by the rims of the disc, could not possibly yield this bright point. Arago's experiment confirmed the prediction of the wave theory. The experimental setup only involved geometric and primitive photometric modules that both theories shared. Their predictions were clear-cut, with no tolerable tampering on their symbolic universe.[7]

In rare cases, the two compared theories do not share basic defining modules such as Euclidean geometry or mechanics. This happens for instance when the predictions of classical and relativistic electron dynamics are compared, or when the predictions of Newton's theory of gravitation are compared with those of general relativity. It would seem that in such cases the shared interpretive schemes could only involve pre-spatial and pre-mechanical observations about the coincidence of two small material objects or the emission and reception of light flashes. This very limited conception of interpretive schemes may in principle allow the comparison of the two theories, for it permits an idealized coordination between theory and simple concrete procedures. In practice, however, physicists never work on a tabula rasa devoid of Euclidean theory, Newtonian mechanics, and other pre-relativistic theories. Comparative schemes involve approximate, local use of these older theories in a complex manner that would deserve systematic study. At any rate, the astronomical tests of general relativity all involve earth-based or satellite-based instruments whose internal design requires earlier accepted geometry and optics, even though the tested spacetime relations are essentially non-Euclidean and non-Minkowskian.

The comparison between two theories may lead to the approximate inclusion of one theory into the other, also called reduction. In this case, the

---

7. Cf. Jed Buchwald, *The rise of the wave theory of light: Optical theory and experiment in the early nineteenth century* (Chicago, 1989), appendix. This experiment did not immediately persuade Poisson to give up the corpuscular theory. The reason is that the wave theory of light, which worked so well in the explanation of interference and diffraction, had not yet been proven to contain geometrical optics as an approximation. It is only after Poisson had such a proof in hands that he gave up the corpuscular theory. More generally, the lack of decision between two incompatible theories may come from insufficient development of their modular structure.

schemes of the reduced theory must correspond to a subset of those of the more general theory. The sharing of schematic modules is trivial, since by our definition of modules, the reduced theory is itself a module of the general theory. What is less trivial is the necessity of defining the schemes of the reduced theory. In a common misconception, the reduction of a theory to another is regarded as a mere limiting process involving a characteristic parameter of the more general theory (for instance *c* in relativistic mechanics, *h* in quantum mechanics) and some correspondence between the theoretical quantities of the two theories. In reality, one must introduce the schemes that define the domain of the reduced theory. Limits performed in the symbolic universe alone are ambiguous and lack definite empirical applicability.[8]

CONSTRUCTION

From history we learn that theory construction is a very complex process, depending on diverse resources both internal and external to the investigated domain. This complexity of what Hans Reichenbach called the "context of discovery" has often discouraged philosophers from finding any rationality in it. Yet a closer analysis of the practice of modern theoretical physics shows that the construction of theories is highly constrained and that at some stages it may proceed almost automatically, as if the plan were known in advance. Well-known constraints in theory construction are experimental laws and general principles such as the conservation of energy or the principle of least action. Less appreciated is the fact that the construction of a new theory always relies on earlier theories in specifiable ways. In other words, some anticipation of the modular structure of a theory efficiently guides its construction.[9]

Most generally, theory construction depends on defining modules whose validity is assumed from the start. For example, the construction of Newtonian mechanics presupposed the module of Euclidean geometry;

---

8. A striking example of this ambiguity is that of Galilean electrodynamics as an approximation to relativistic electrodynamics: cf. Michel Le Bellac and Jean-Marc Lévy-Leblond, "Galilean electromagnetism," *Nuovo cimento*, B14 (1973), 217–233.
9. Jürgen Renn's powerful notion of "integration of knowledge" through "mental models" implicitly involves such anticipation of modular structure. See Jochen Büttner, Jürgen Renn, and Matthias Schemmel, "Exploring the limits of classical physics: Planck, Einstein, and the structure of a scientific revolution," *Studies in history and philosophy of modern physics*, 34 (2003), 37–59.

and the construction of electrodynamics presupposed the module of mechanics (at least in the definition of forces). Such defining modules occur both in the symbolic universe and in the interpretive schemes. They sustain our theoretical imagination in a concrete manner, in direct connection with measurement possibilities.[10]

In a less universal and less concrete mode, theory construction may rely on reducing modules, as was for instance the case in Maxwell's first derivation of his electromagnetic field equations. The analogy between magnetic phenomena and the rotational motion of a substance inspired Thomson's and Maxwell's idea that the electromagnetic ether could be a connected system with internal rotations to be identified with the magnetic field. The consistent development of this idea led to Maxwell's equations. Although Maxwell suppressed the mechanical model in the final version of his theory, he retained a broader principle of Lagrangian structure. This is only one example of a historical process in which a reducing module evolves into a general principle of a more abstract nature. Our theories are full of such vestiges of past modular reductions.[11]

In the development of his mechanical model of the ether, Maxwell was also guided by his desire to integrate electromagnetic, electrostatic, and optical modules in the same theory. In this case modular structure played a double role: in founding a reductionist strategy, and in bringing together different partial theories as modules of a new theory. To sum up, reduction and unification are modes of theory construction that explicitly depend on modular structure. There are two kinds of reduction of a theory to another: one in which the second theory is a reducing module of the second, and another in which the first theory becomes an approximating module of the second. The unification of two or more theories is a process following which these theories end up being approximating or specializing modules of the same theory.

Theory construction also depends on the important modular constraint that the new theory should contain earlier successful theories as approximations. In our terminology, the earlier theories should be approximating modules of the newer theory. This constraint is usually called

10. The defining modules thus share some virtues of Michael Friedman's constitutive principles, which will be discussed in section 4.
11. Cf. Buchwald 1985; Siegel 1991.

a *correspondence principle*, in reverence to Bohr's endeavor to construct quantum theory in a way that ensured the asymptotic validity of classical electrodynamics. In combination with some assumed symmetries or some general postulates, such a principle may completely define the sought-after theory. This happened in the case of relativistic dynamics and in the case of quantum mechanics. In Bohr's conception of the latter theory, the classical module is important not only in the construction of the symbolic universe but also in the definition of the interpretive schemes. Indeed for Bohr any measurement ultimately relies on classical modules.[12]

The pursuit of modularity does not always bring progress. In some cases, theories that had long been used as defining or reducing modules must be thrown away or relegated to the humbler modular role of approximation. For instance, most nineteenth-century physicists regarded mechanical reduction as a legitimate and accessible aim for the whole of physics. They were blinded by the success of early reductions of this kind. Toward the end of the century, the pragmatist or positivist convictions of a few physicists confined mechanics to the more modest function of a defining module. The downgrading of classical mechanics went on at the beginning of the twentieth century, when it appeared to be an approximating module of a more fundamental relativistic or quantum mechanics. In this process, even the defining modules of Euclidean geometry and Galilean kinematics came under attack. They were ultimately replaced by and became approximating modules of the pseudo-Riemannian geometry of general relativity theory.

The lesson to be drawn from this evolution is that the modular structure of a theory should never be regarded as definitive. The most we can say is that any theory that has been successful in a given domain of physics is likely to remain, after adequate purification or reformulation, a module of future, more general theories. But its modular function may evolve in time. As we saw, classical mechanics once played the role of a defining or a reducing module. It remains a defining module in useful macroscopic theories. But it is only an approximating module for the most fundamental theories such as general relativity or quantum field theory. Theories all

12. On Bohr's views, cf. Catherine Chevalley, «Complémentarité et langage dans l'interprétation de Copenhague», *Revue d'histoire des sciences et des techniques*, 38 (1985), 251–292; «Le dessin et la couleur», introduction to Niels Bohr, *Physique atomique et connaissance humaine* (Paris, 1991), 19–140.

have modular structure. But the modular structures of successive theories bear partial resemblance only. Any stiffening of our modular habits could become an obstacle to further progress.

COMMUNICATION

The modular structure of theories is essential to successful communication among practitioners belonging to different social groups. Physicists who belong to different local subcultures may adhere to different theories of the same domain. As Poincaré and Ludwig Boltzmann forcefully argued, this cultural diversity is usually beneficial to science, because it favors the exploration of a greater variety of symbolic universes and thus increases chances to find the one that best fits the widest domain. It can be so only if communication is possible between the different subcultures. Maximal communication, in which the physicists of one subculture perfectly understand the theories of the other, almost never occurs. It is not even to be wished, because it would interfere with the creative energy of each group. More commonly, the two groups communicate through interpretive or descriptive schemes that involve shared modules only. The shared modules are in part given, or they are constructed by a few "bilingual" individuals who labor for the easy communicability of science.[13]

An instructive example is that of electrodynamics in the nineteenth century. British physicists favored a field-based approach; German physicists favored direct action at a distance. Yet these two communities were able to benefit from each other's results and to compare the predictions of their theories. In part, this was possible because of their inheriting mechanical, electrostatic, magnetostatic, and electrodynamic modules from the same French sources (Coulomb, Poisson, Ampère). For the rest, William Thomson played a crucial role in designing modular concepts that could be used equally well by physicists and engineers of any country. For example, he defined the electric potential through the mechanical concept of energy, independently of any deeper interpretation in the competing symbolic universes. As a consequence or as a motivation, electrometers and other electrical apparatus could be traded between the two cultures, because the

13. On Boltzmann's pluralism, cf. Nadine de Courtenay, *Science et philosophie de Ludwig Boltzmann. La liberté des images par les signes*. Thèse de doctorat (Université de Paris 4, 1999).

modules necessary to their use were made available on both sides.[14]

As was just mentioned in the Thomson case, modules are also essential in the communication between physicists and engineers. Engineers almost never have a detailed knowledge of the deeper theories through which physicists would understand some aspects of their practice. Yet they are constantly benefiting from these theories because they master the modules that are sufficient for their own purposes. More generally, division of work necessitates modular communication between groups who have unequal access to deeper theory. In some domains of modern physics such as particle physics, there are separate subcultures of theorists, experimenters, and instrument makers. As Peter Galison has argued, the necessary communication between these various groups leads to the formation of "trading zones," that is, virtual places of exchange in which the various protagonists can benefit from each other's competences without ever acquiring all of them. Theoretical modules play a crucial role in this sort of trade. In some cases, physicists forge the modules just for this purpose. This does not mean, however, that modules only are an arbitrary product of a social consensus formed in the trading zone. The structure they reflect is an inherent structure of the embedding theories, and it becomes part of our ultimate understanding of these theories.[15]

Modularity is also important in the communication and understanding of theories within the same subculture of physicists. Physics courses and textbooks are divided into chapters that often correspond to approximating or specializing modules of the theory to be taught. For instance, a textbook of electrodynamics typically has chapters of electrostatics, electrokinetics, quasi-stationary electrodynamics, and electromagnetic radiation. Within each chapter, exemplars are given of interpretive schemes for which the consequences of the laws of the relevant modules can be fully worked out. Eventually, approximation methods are taught for dealing with systems that somewhat depart from these exemplars. The importance of exemplars in learning physics has already been emphasized by several authors including Thomas Kuhn, Ronald Giere, and Nancy Cartwright. What I want to emphasize is that exemplars usually concern approximating or specializing

14. Cf. Crosbie Smith and NortonWise, *Energy and empire: A biographical study of Lord Kelvin* (Cambridge: 1989).
15. Galison, *Image and logic: A material culture of microphysics* (Chicago, 1997), Chap. 9.

modules of a theory rather than the whole theory, and that their treatment almost always involves defining and schematic modules that the students have already learned in other contexts.

Modularity may also be illustrative. Its purpose then is to feed the intuition. Many British physicists of the nineteenth century believed that a theory could not be properly understood without illustrating some of its parts by other well-understood theories. They relied on *illustrating modules*, namely: reducing mechanical modules that worked for limited classes of interpretive schemes of the global theory. For instance, in "On Faraday's lines of forces," Maxwell illustrated the electrostatic, magnetostatic, and electrokinetic modules of electrodynamics by means of the mechanics of resisted flow in a porous medium. He thus insufflated some life into the dry symbols of potential theory. His British contemporaries similarly liked to flesh out the equations of their theories by attaching them to partial reducing modules. This sort of fictitious concreteness is part of any understanding of theories. Besides its pedagogical virtue, it eases the mental associations through which a theory can evolve and fuse with other theories. Concrete illustrations are not there to replace theory, as today's imaging fans would have it; they are there to enliven theory.[16]

COMPREHENSIBILITY AND NECESSITY

In the definition of theories developed in the two previous sections, a symbolic universe and its laws are first given. This universe and attached modules enable us to conceive interpretive schemes which involve ideal measurements and function as blueprints of experimental setups. The laws of the symbolic universe imply relations between the characteristic quantities of each interpretive scheme, and these relations may be verified experimentally. This conception of theories allows us to formulate the following question: in a given domain of physics, could we infer the nature of the interpretive schemes by idealizing some concrete conditions of comprehensibility of this domain and then infer the symbolic universe and its law

16. Cf. Wise, "The mutual embrace of electricity and magnetism," *Science*, 203 (1979), 1310–1318. Jordi Cat, "On Understanding: Maxwell on the methods of illustration and scientific metaphor," *Studies in the history and philosophy of modern physics*, 32 (2001), 395–441.

from these conditions? In other words, can we move from the conditions of possibility of experience in a given domain to the theory of this domain? On the one hand, this sounds like a foolish ambition, for history teaches us that the results of experiments – not solely our ability to conceive them – are needed to design new theories. On the other hand, the question is not so farfetched because in some cases, such as Einstein's discovery of relativity theory, a priori conditions of measurability did play a significant role in the genesis of the theory. Moreover, there are cases in which already known theories have been a posteriori shown to derive from simple conditions of comprehensibility. Let us first consider two such cases, physical geometry, and statics.

PHYSICAL GEOMETRY

One of the simplest physical theories, if not the simplest, is Euclidean geometry understood as the geometry that correctly predicts the properties of concrete figures at a reasonable human scale. Let us first see how this theory fits our general definition of theories. The systems of the symbolic universe are the subsets of a three-dimensional real affine space in which a distance is given in the mathematical sense (a positive real symmetric function of two points that vanishes if and only if the two points are equal and that satisfies the triangular inequality). The basic law of the universe may be taken to be the existence of a positive non-degenerate bilinear form from which the distance derives. In more concrete terms, there exist systems of coordinates (x, y, z) in the affine space such that the distance of a point from the origin is simply given by the Pythagorean formula $\sqrt{x^2 + y^2 + z^2}$. We may now define interpretive schemes in which the point systems are figures made of straight lines (defined by the linear structure of the vector space) and circles in order to arrive at traditional synthetic geometry. The characteristic quantities of these schemes are angles and lengths (one could add surfaces and volumes). The ideal measurement of a length is given by the transport of an invariant unit and its subunits thanks to the positive isometries (translations and rotations) of the Euclidean space, and the ideal measurement of an angle is given by the measure of the length of the associated arc of a unit circle. The Pythagorean law implies relations between the angles and the sides of a given figure. For instance, in the triangle ABC the length of BC is a well-defined function of the lengths AB and AC and the angle (AB, AC). A typical geometric experiment would be to measure the three segments and the angle and to verify the theoretical relation between them.

All of this is a bit artificial and the empirical verification of triangular relations seems to be in need of some explanation. An axiomatic, neo-Euclidean reformulation of the theory would not help much, because we would not know the rationale behind the axioms from which Pythagoras's theorem and the triangular relations would now derive.

An alternative route, open by Hermann Helmholtz in the late 1860s, consists in inquiring about the meaning of space measurement before any relevant mathematical structure is given. Following Helmholtz, let us assume that geometry is about measuring distances by means of some gauge. For instance, we may count the minimum number of steps needed to go from one point to another; or, better, we may do the same with a rigid rod. The success and non-ambiguity of this procedure entails the following assumptions for the class of rigid rods:

1) For any two rods, if an extremity of the first rod is kept in contact with an extremity of the second, the other extremity of the first rod can be brought in contact with at most one point of the second (no plasticity or elasticity).
2) If coincidence can be obtained in one place and at one time between a pair of points of one rod and a pair of points of the other, this coincidence will be possible at any other place and time, no matter how variously and differently the two rods have traveled before meeting again (free mobility and stability).

This definition does not entail any prior concept of distance. It permits a direct empirical test of rigidity and free mobility. Of course, there are infinitely many classes of rigid rods according to this definition. Rods obtained by subjecting the rods of a given class to a dilation that depends only on their location will form another class of rigid rods, whatever be the dilation law. For instance, in a thermostatic universe with heterogeneous temperature, iron rods and copper rods define two distinct classes of approximately rigid rods.[17]

Once congruence has been defined by means of a class of rigid rods, the distance between two points (that is, two small objects) can be measured

17. Hermann Helmholtz, „Über die Thatsachen, die der Geometrie zum Grunde legen," Königliche Gesellschaft der Wissenschaften von der Georg-August-Universität zu Göttingen, *Nachrichten* (1868), 193–221. The following is a very free reconstruction of Helmholtz's argument.

by means of chains of unit rods. At the precision of the unit, the distance is given by the minimal number of links of a chain joining the two points. This distance measurement can be refined by using smaller and smaller unit rods. In common practice, a sequence of subunits is used such that the lengths of two consecutive subunits differ by a factor ten (for instance). The outcome of the measurement is a decimal number whose last digit corresponds to the last subunit whose extremities can be distinguished.

So far, we have considered concrete objects and operations that can be realized in an approximate manner only. We may now leave the empirical world and take our flight to a mathematical set-theoretical world in which the properties (1) and (2) of rigid rods hold exactly, and the sequence of subunits can be pursued indefinitely. The usual sets of natural, rational, and real numbers can thus be engendered in harmony with the geometer's needs. Whether or not geometry truly motivated the historical introduction of these mathematical constructs, it is important to recognize that any theory of measurement requires these constructs or similar non-standard ones.

We now know how to measure distances with arbitrary precision. Suppose there exist three points A, B, and C whose mutual distances are found to be invariable. We know by experience that except for singular cases the location of any fourth point within a sufficiently small domain is determined by its distances from these three points. In other words the location is determined by three coordinates. Moreover, the distance of a variable point M from a fixed point O varies linearly under small increments of its coordinates, except when M is originally at O. In the latter case, the variation cannot be linear since this would allow the distance between M and O to vanish without their coordinates being equal. This variation must nonetheless be a homogenous function of first degree of the coordinate increments, because for a sufficiently small unit of length, a (reasonable) unit change implies a multiplication of all measured distances by the same constant.

These conclusions are only valid to a certain approximation, given by the precision of the distance measurement. Again we may jump to the ideal, mathematical level in which coordinates are sharply defined as real numbers. At this level, the distance OM should be a differentiable function of OA, OB, and OC whenever M differs from O. This implies the differentiability of any change of coordinates resulting from a different choice of the reference points A, B, and C. The resulting mathematical concept is that of a differentiable manifold in Bernhard Riemann's original sense, namely, endowed with a metric that is not necessarily of the locally Euclidean form.

In order to further restrict the form of the metric, we need some additional condition. By experience we know that the position of a point M with respect to three rigidly connected reference points A, B, and C is completely determined by its distance from these three points. This implies that the distance between two points M' and M" is a function of their distances from the reference points. This function of course depends on the choice of the reference points. However, by experience we know that it only depends on the mutual distances of these reference points, as long as none of the involved distances is exceedingly large. This fact can be regarded as a precise expression of the homogeneity of space over moderate distances, since it means that the same relations between all measured distances can be used in surveys performed in a not too large domain.

This local homogeneity implies the existence of rigid bodies in the following sense: the distances between any number of points remain the same when their distances to three points A, B, C and the distances between these three points are kept constant. In other words, there exist transformations that preserve the mutual distances of any number of points. The rigid bodies defined in this manner enjoy free mobility, since the choice of three new reference points A', B', C' such that AB = A'B', AC = A'C', BC = B'C' involves six degrees of freedom (nine coordinates minus three constraints).

The assumption of freely mobile rigid bodies allows us to define an angle as a rigid connection of two straight segments of arbitrary length with a common extremity. The addition or difference of two angles is defined by making these two angles share one of their sides in the same plane, and taking the angle made by the two remaining sides. The straight angle is the angle that makes a flat angle (a single straight line) when added to itself.

Call d the length of the hypotenuse of a rectangular triangle, x the length of one of its shorter sides, and α the (positive or negative) angle between this side and the hypotenuse. The intersection between two lines being unique (at least locally), and there being only one line perpendicular to a given line through a given point, the length d is a function of α and x only. This function vanishes for $x = 0$, and it must be differentiable with respect to x because the space manifold is differentiable. Therefore, d is a linear function of x for small triangles. This means that the ratio between one of the shorter sides of a small rectangular triangle and its hypotenuse is completely determined by the angle that they make.

Owing to the free mobility of rigid bodies, this ratio cannot be altered by rigid displacement of the triangle. Now take a look at figure 1. The former theorem implies AH/OA = OA/AB as well as BH/OB = OB/AB. Since AB = AH+BH, this leads to the relation

$$AB^2 = OA^2 + OB^2,$$

which is Pythagoras' theorem. As is well known, Euclid's premises support the same proof. The reason is that his "common notions" contain the assumption of freely mobile rigid bodies, and his postulates turn the local validity of the theorem into a global one.

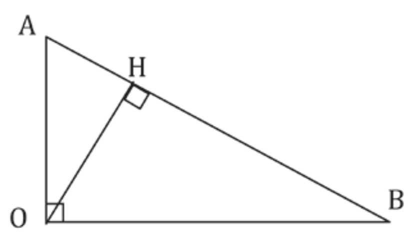

Fig. 1: Figure for a proof of Pythagoras' theorem.

In a two-dimensional space, the validity of Pythagoras' theorem for small triangles would imply the validity of Euclidean geometry at a sufficiently small scale and the Riemannian character of space at large scale. Other considerations are needed to extend this result to three dimensions. They are found in Helmholtz's original memoir. For our present concern, the two-dimensional case is sufficient to illustrate the power of a simple consideration of measurability.[18]

To sum up, the existence of a class of (small) freely mobile rigid *rods* leads to a consistent surveying technique. Although the choice of this class is largely conventional, it must meet certain empirical conditions such as the criteria (1) and (2). Together with the observed smoothness and local homogeneity of the space manifold, the existence of freely mobile rigid rods leads to the existence of freely mobile, approximately rigid *bodies*, from which the Riemannian character of physical geometry follows. The further determination of this geometry depends on the class of rigid bodies

18. See Darrigol, ref. 5, Chap. 4.

that has been conventionally adopted. For a given convention, systematic surveys determine the metric of the Riemannian manifold.

This approach involves a gradual sharpening of elementary spatial notions. Ultimately, it leads to a highly idealized theory in which the thickness of material points, errors in the appreciation of congruence, the imperfect rigidity of rods, and the rationality of the numbers yielded by actual measurements are ignored to permit exact relations and deductions. The outcome can be considered as a purely mathematical theory, as an abstract set-theoretical construct whose properties no longer depend on experience.

However, the empirical motivation of this idealized geometry explains its success as a physical theory. The most important assumption, that of freely mobile rigid bodies, depends on a kind of invariance assessed in our experience of the world: the fact that congruence among a certain type of objects is constrained and reproducible. In concrete form, this assumption is the basis of our physical concept of space. In idealized from, it is the basis of the Riemannian concept of space. Although mathematicians are free to imagine more general concepts of space, the locally Euclidean property is essential to the measurability of physical space.

STATICS

Statics is the theory that gives the conditions of equilibrium of simple mechanical systems called "connected systems," made of levers, pulleys, threads etc. in permanent rolling or sliding contact and subjected to a given set of forces. The symbolic universe of this theory involves rigid bodies, inextensible threads, and incompressible fluids whose possible configurations are defined by means of an Euclidean geometric module and constrained by the contact conditions. It also involves a vector space of forces. The fundamental law or this universe is the *principle of virtual works*:

> If $\mathbf{F}_\alpha$ denotes the force acting on the material point of the system, the system is in equilibrium if and only if $\sum \mathbf{F}_\alpha \cdot \delta\mathbf{r}_\alpha = 0$ for any possible displacement $\delta\mathbf{r}_\alpha$ of the material points (called *virtual displacement*).

The forces involved in this statement do not include the internal contact forces used in elementary expositions of statics. They may include internal forces of elastic, gravitational, or electric origin. An interpretive scheme of the theory is a choice of a system together with characteristic quantities that are geometric configuration parameters and applied forces. The ideal measurement of the geometric parameters is dictated by the geometric defining module of the theory, and the ideal measurement of forces is given by the equilibrium condition of a simple mechanical system used as a comparator of forces. For example, we may use the pulley-thread device of fig. 2. The force to be measured is applied to one end of the thread and it is balanced by a force that is produced by multiples of a unit force and its subunits applied at the other end of the thread (concretely the latter forces may be marked measuring weights on a pan suspended to the thread). The direction of the measured force is given by the direction of the thread that it pulls, and its intensity is given by the numbers of balancing units and subunits as a trivial consequence of the principle of virtual works. A possible experiment under this scheme involves the measurement of the forces acting on the system and the verification of the condition of equilibrium.

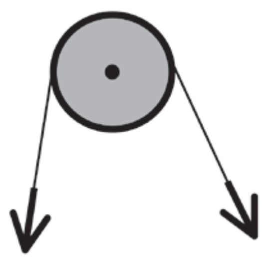

Fig. 2: The pulley-thread comparator of forces.

Just as we did in the case of geometry, we may now try to reverse the logic and to infer the symbolic universe and its laws by idealization of the concrete measurement conditions. In this approach, the balancing of forces through a pulley-thread system concretely defines their direction and their intensity, without prior mathematical concept of force. This concrete definition leads to the real vector space of forces through the idealization of indefinitely precise measurement, just as concrete length determination leads to real-number lengths. The consideration of concrete, well-built connected systems leads to the following ideal conditions:

1. For any infinitesimal change of configuration compatible with the constraints (virtual displacement) in which the position of the material point α changes by $\delta\mathbf{r}_\alpha$, the opposite change - $\delta\mathbf{r}_\alpha$ is also possible.
2. Arbitrarily small forces acting in the direction of mutually compatible displacements suffice to break equilibrium.

The first condition implies the permanence of contacts between rigid bodies. The second excludes solid friction (caused by roughness, for example). Following an improved version of a reasoning of 1798 by Joseph Louis Lagrange, it will now be proved that the principle of virtual velocities derives from the impossibility of perpetual motion combined with these two conditions and with the pulley-thread definition of force.

The basic idea is to synthesize the forces $\mathbf{F}_\alpha$ through a set of tackles, a single rope running through them, and a weight. The simple tackle of fig. 3 yields the force $2F$ under the tension F of the rope. Indeed if the force acting on the axis of the pulley differed from $2F$, a perpetual motion could be generated by connecting this axis and the two ends of the rope to the same rigid frame. Similarly, the triple tackle of fig. 4 yields the traction $4F$, and so forth. As the intensities F_α can all be regarded as even multiples $2N_\alpha F$ of the same small intensity F (with a precision increasing with the smallness of F), they can be generated by a properly arranged system of tackles through which the same rope runs (see fig. 5). The tension F of the rope is produced by a weight W.[19]

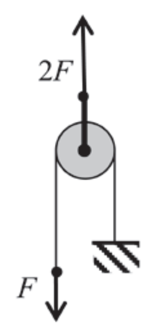

Fig. 3: Simple tackle.

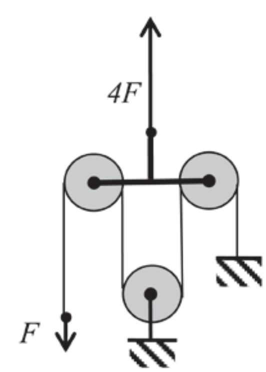

Fig. 4: Triple tackle.

19. Joseph Louis Lagrange, «Sur le principe des vitesses virtuelles», *Journal de l'École Polytechnique*, vol. 2 , *cahier* 5 (1798), 115–118. In a concrete connected system on earth, the $\mathbf{F}_\alpha$'s would include the weight of the various components of the system, so that Lagrange's construction can only be an imaginary one (the more so because W itself is subjected to gravitation).

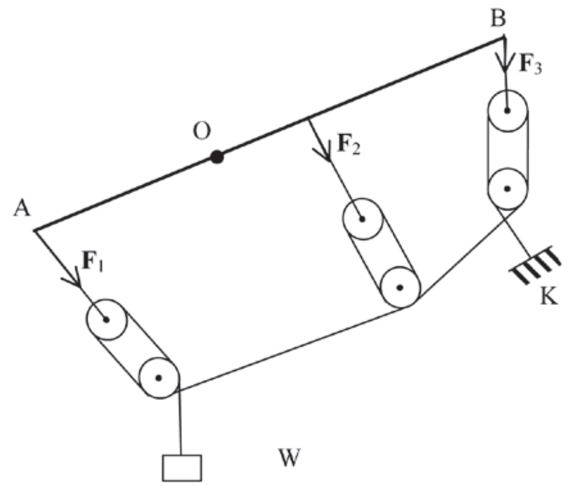

Fig. 5: Lagrange's contraption for a proof of the principle of virtual velocities in the case of a lever. The forces $\mathbf{F}_1$, $\mathbf{F}_2$, $\mathbf{F}_3$, acting on the lever AOB, are produced by three rigidly held tackles through which a single rope runs from the anchor K to the suspended weight W.

The virtual displacement $\delta\mathbf{r}_\alpha$ of the material point α on which the force $\mathbf{F}_\alpha$ is acting induces a shift $2N_\alpha(\mathbf{F}_\alpha / F_\alpha) \cdot \delta\mathbf{r}_\alpha$ of the rope, as an obvious consequence of the makeup of the tackles. The resulting shift of the end of the rope is $F^{-1}\sum \mathbf{F}_\alpha \cdot \delta\mathbf{r}_\alpha$. If there exists a virtual displacement such that $\sum \mathbf{F}_\alpha \cdot \delta\mathbf{r}_\alpha \neq 0$, this displacement or the opposite displacement (warranted by the property (1) of connected systems) is such that $\sum \mathbf{F}_\alpha \cdot \delta\mathbf{r}_\alpha > 0$. The weight W therefore pulls the rope in the direction of a possible displacement. According to the property (2) of connected systems, the rope must move no matter how small this weight is. Therefore, the system is not in equilibrium. By contraposition, the virtual work $\sum \mathbf{F}_\alpha \cdot \delta\mathbf{r}_\alpha$ must vanish for the system to be in equilibrium.

Reciprocally, the vanishing of $\sum \mathbf{F}_\alpha \cdot \delta\mathbf{r}_\alpha$ for any virtual displacement implies equilibrium. We will prove this *ad absurdum*. Suppose that the system is not in equilibrium under the forces $\mathbf{F}_\alpha$. Then equilibrium can be restored by applying additional forces $\mathbf{X}_\alpha$ directed against the initial displacements $d\mathbf{r}_\alpha$ of the material points α. Otherwise, the same weight W and the same resulting displacements $d\mathbf{r}_\alpha$ could be used to lift arbitrary heavy weights through a pulley-rope mechanism and perpetual motion would be possible. On the one hand, the counterbalancing forces $\mathbf{X}_\alpha$ verify the inequality $\sum \mathbf{X}_\alpha \cdot d\mathbf{r}_\alpha < 0$. On the other hand, the restored equilibrium require $\sum (\mathbf{F}_\alpha + \mathbf{X}_\alpha) \cdot d\mathbf{r}_\alpha = 0$. These two relations contradict the vanishing of $\sum \mathbf{F}_\alpha \cdot d\mathbf{r}_\alpha$.

Thus we see that a very powerful principle of statics simply results from idealized definitions of forces and mechanical systems and from the impossibility of perpetual motion. The definitions serve to define the domain of the theory, and they involve considerations of measurability directly in the case of forces and indirectly in the case of spatial relations. Measurability evidently is a condition for the quantitative comprehension of the world. The impossibility of perpetual motion also has to do with the comprehensibility of the world, because it implies a weak kind of causality: the impossibility of spontaneous changes (such as the rise of a weight) without any compensatory change in the environment.

KINDS OF COMPREHENSIBILITY

The arguments given above are rational inasmuch as they carry the conviction that the only conceivable geometry at a moderate scale is Euclidean geometry, and the only theory of mechanical equilibrium is the one based on the principle of virtual work. They are not purely rational because the premises of the reasoning are empirically fallible. For instance, the measurement of space through the congruence of rigid bodies or the measurement of force through a pulley-thread mechanism may cease to make sense when the scale of measurement is too small or too large. In fact, we know from general relativity that purely spatial measurement only makes sense at sufficiently small scale (there cannot be any extended rigid body or frame), and we know from quantum theory that the laws of classical mechanics fail at sufficiently small scale. So all we can prove is that simple conditions of comprehensibility that seem natural *in a given domain of experience* may completely determine the theory appropriate to this domain.

With the same caveat, we may demonstrate the necessity of several important theories. The necessity of Newtonian mechanics derives from the measurability of space and forces, Galilean relativity, a causality principle that relates any alteration of motion to the action of forces, and the *secular principle* that requires motion at the macroscopic scale to be independent of microscopic fluctuation of the applied forces. The necessity of thermodynamics derives from the impossibility of perpetual motion and from the uniqueness of thermodynamic equilibrium. The necessity of relativistic mechanics derives from the measurability of space and time, the relativity principle, and the correspondence with Newtonian mechanics for moderate velocities. The necessity of the pseudo-Riemannian spacetime of general relativity derives from the measurability of space and time in an optically controlled manner. Or the broader Weylian structure of spacetime can be derived from geodesy based on light rays and free-falling particles. The necessity of our main classical field theories, including electromagnetic theory and general relativity, derives from the *Faradayan principle* that the field action can only depend on properties that can be tested by point-like particles. The necessity of the Hilbert-space structure of quantum mechanics derives from natural assumptions about the statistical correlations of measurements performed on an individual system. In simple cases, the associated dynamics derives from the principle of correspondence.[20]

Altogether, we see that the necessity arguments involve three kinds of comprehensibility. Firstly, we may require the measurability of some basic quantities. A broad consequence of this requirement is the relevance of mathematical analysis in the formulation of physical theories. More detailed consequences depend on the type of quantity and on the way in which the measurement process is idealized. Space measurement by rigid bodies leads to the locally Euclidean character of space; time measurement by inertial motion, together with space measurement, to the locally Minkowskian structure of spacetime; space and time measurement by light signals and free-falling particles leads to a Weyl spacetime. Field measurement by point-like particles leads to the accepted classical field theories. Considerations of measurability often go hand in hand with a requirement of objectivity: measurements performed by different observers or with different conventions should be interrelated in a consistent manner. The principle of relativity expresses this sort of objectivity.

The second kind of comprehensibility rests on varieties of causality. The broadest variety is the stability of statistical correlations between measurements performed on the same system. This is the one admitted for quantum systems and leading to the Hilbert space structure of states when combined

20. See Darrigol, ref. 5.

with natural assumptions on the type of correlations. In the classical case, we require a stricter kind of causality according to which the same cause creates exactly the same effect in similar circumstances. In addition, we may require secular average effects at a given scale to be unaffected when the causes fluctuate at a finer scale: this is the principle of secularity, which can be used together with the previous principle in a derivation of Newton's second law. We may also assume the weaker kind of causality implied in the principle of the impossibility of perpetual motion. This principle not only contributes to a proof of the necessity of classical mechanics, but it also helps justifying the energy principle and the first principle of thermodynamics without appealing to mechanical reduction. The uniqueness of thermodynamic equilibrium may also be seen as a kind of causality since it requires the uniqueness of the macrostate of a system under given macroscopic circumstances. The second principle of thermodynamics, expressed as the impossibility of spontaneous heat flow from a cold body to a hot body, derives from the uniqueness of equilibrium if we regard the state of equal temperatures as an equilibrium state.

The third kind of comprehensibility rests on the applicability of correspondence principles. In this case, the necessity of some features of a theory T is derived from its agreement with a theory T' known to be (approximately) true in a restricted domain of experience. This agreement implies the existence of sub-theories that are approximations of the theory T, as well as the identity or the equivalence of one of these sub-theories with the theory T'. In combination with other arguments, a correspondence argument can be used to show the necessity of a theory. The strength of the demonstration of necessity depends on the quality of the other arguments and on the necessity of the restricted theory. The latter may be established empirically, or it may itself be derived by necessity arguments. Both circumstances are met in the case of non-relativistic mechanics as a correspondence-basis of relativistic mechanics, or in the case of classical mechanics as a correspondence-basis for quantum mechanics.

As was announced at the beginning of this section the formulation and the deployment of comprehensibility arguments requires three features of the definition of theories propounded in this essay: the interpretive schemes, the schematic modules, and the approximating modules. Idealized measurability conditions are proto-theoretical schemes that imply definitional modules. Correspondence arguments rely on the notion of approximating modules. Although causality principles may be formulated without much theoretical substructure, their effective deployment requires measurement and correspondence principles that require substructure.

We thus reach a highly structured conception of the means required to efficiently predict physical phenomena. The least structured conception with which we started this essay is the Big Data ideal of combining the massive collection of data with computer-based algorithmic bootstrapping. In a somewhat more structured conception, theories are declared indispensable but they are essentially seen as mathematical structures and any interpretive substructure is ignored. In the next degree of structuration, interpretive schemes and modular structure are deemed necessary. In the ultimate degree, some important theories and their modules are shown to derive from broad considerations of comprehensibility.

KANT'S GHOST

The comprehensibility arguments of the previous section are based on a priori conditions for the comprehensibility of the world. These conditions sound similar to Kantian a priori conditions for the possibility of experience, except that they vary in time and depend on the domain of experience. Kant's great service, in the context of eighteenth-century philosophy, was to offer a via media between David Hume's skepticism and the idealism of René Descartes and Gottfried Wilhelm Leibniz, in compliance with the success of Newtonian physics. In Kant's transcendental philosophy, the mind plays an active role in constituting the object of knowledge. It does so through a representative faculty called intuition and through a legislative faculty called understanding. Time is the pure form of internal intuition, and space the pure form of external intuition. Arithmetic and (Euclidean) geometry, which then were the basis of all known mathematics, derive from the injection of a category of understanding, quantity, into intuition. The laws of Newtonian mechanics (conservation of mass, inertia, equality of action and reaction), which many then believed to be the general foundation of physics, derive from applying three other categories (substance, causality, community, which are the subcategories of relation) to matter reduced to a continuous distribution of centers of force.[21]

21. Immanuel Kant, *Critik der reinen Vernunft* (Riga, 1781); *Prolegomena zu einer jeden*

In the nineteenth century, it became clear that Euclidean geometry was not the only conceivable geometry and that mathematics should be defined independently of any concept of space and time. This strong blow to Kant's doctrine did not discourage attempts to salvage some of it. In his memoirs on the foundation of geometry, Helmholtz argued that Kant's form of external intuition could be preserved if it was restricted to the general idea of space as a continuous and uniform manifold. The Riemannian structure and the constant curvature of this manifold derived from the measurability of space by freely mobile rigid bodies, which could loosely be regarded as a rule of the understanding. Experience was needed only in the determination of the value of the curvature. Poincaré proceeded differently: he gave up Kant's idea of a passive intuition and he made the general notion of space depend on the concept of Lie group, which he regarded as a synthetic a priori "form of the understanding." Poincaré regarded the choice of the group as conventional because geometrical laws were never tested independently of mechanical laws (ruling the deformation of bodies or the propagation of light) and the latter laws could always be adjusted to fit a conventionally given geometry. A regulative principle of simplicity induced Poincaré to maintain Euclidean space and Galilean spacetime in face of relativistic challenges.[22]

Further blows to Kant's doctrine came with Einstein's relativity theory. Special relativity mixed up the two forms of intuition that Kant had separated, and it downgraded Newtonian mechanics to an approximation of a deeper theory in which Newtonian mass and momentum were no longer conserved. General relativity undermined the distinction between inertial force and external force and brought the metric properties of spacetime to depend on the distribution of matter. Not only the idea of space and time as

künftigen Metaphysik die als Wissenschaft wird auftreten können (Riga, 1783); *Metaphysische Anfangsgründe der Naturwissenschaft* (Riga, 1786). Cf. Michael Friedman, *Foundations of spacetime theories: Relativistic physics and philosophy of science* (Princeton, 1983).

22. Cf. Howard Stein, "Some philosophical prehistory of general relativity," in John Earman, Clark Glymour, and John Stachel (eds.), *Foundations of space-time theories. Minnesota studies in the philosophy of science*, vol. 8, (Minneapolis, 1977), 3–49; David Hyder, *The determinate world: Kant and Helmholtz on the physical meaning of geometry* (Berlin, 2009); Michel Paty, *Einstein philosophe : la physique comme pratique philosophique* (Paris, 1993), 250–263; Gerhard Heinzmann, "The foundations of geometry and the concept of motion: Helmholtz and Poincaré," *Science in context*, 14 (2001), 457–470; Robert DiSalle, *Understanding space-time: The philosophical development of physics from Newton to Einstein* (Cambridge, 2006).

pre-structured metric frames had to be given up, but the very idea of space and time as a stage for phenomena had to be given up.

Altogether, something had gone deeply wrong in Kant's doctrine. Was it the distinction between two separate faculties of the mind (intuition and the understanding)? Was it the schematism that bridged the two faculties? Was it the table of categories of the understanding? Was it the strict distinction between object-defining "constitutive principles" and theory-guiding "regulative principles"? Or was it a bit of all of that?

REICHENBACH'S PRINCIPLES OF COORDINATION

In 1920, one year after Eddington's solar eclipse expedition had confirmed Einstein's prediction of the gravitational deflection of light, the young Hans Reichenbach published a brave tentative to rescue what he took to be the essence of Kant's project: the idea that the object of knowledge is constituted by the human mind. In Kant's a priori, Reichenbach argued, one should distinguish between two aspects: apodictic certainty, and constitutive power. The first aspect had to go, but the second was as needed as ever. Within the mess of sensations, the mind had to impose some order in a manner that ordinary perception could not by itself suggest.[23]

The central concept of Reichenbach's new theory of knowledge was coordination (*Zuordnung*). Although Reichenbach borrowed this word from Moritz Schlick's empiricist theory of a one-to-one set-theoretical correspondence between theoretical concepts and physical reality, he believed that the epistemological concept of coordination essentially differed from the set-theoretical concept because the coordinated elements of physical reality were not defined before the coordination. These elements were defined by the coordination. In order to be successful, the coordination had to be univocal, surely not in the set-theoretical sense of the word (which presupposes the target elements to be predefined), but in the following empirically testable sense: the value of any measurable quantity must be the same whatever be the data used for its determination.[24]

At that stage of his reasoning, Reichenbach introduces the principles of

23. Hans Reichenbach, *Relativitätstheorie und Erkenntnis apriori* (Berlin, 1920). Cf. Thomas Ryckman, *The reign of relativity: Philosophy in physics 1915–1925* (Oxford, 2005), 28–39.
24. Reichenbach, ref. 22, Chap. 4, p. 43: „*Eindeutigkeit* heißt für die Erkenntniszuordnung, daß eine physikalische Zustandsgröße bei ihrer Bestimmung aus *verschiedenen Erfahrungsdaten* durch *dieselbe Messungszahl* wiedergegeben ist."

coordination (*Zuordnungsprinzipe*) as the principles that make the coordination univocal. These principles constitute the object of the theory as they define the mathematical form of physical quantities and the manner of combining them. They do not by themselves determine the theory; in addition, Reichenbach appeals to laws of combination (*Verknüpfungsaxiome* or *Verknüpfungsgesetze*) that relate different physical quantities in an empirically testable manner (once the principles of coordination are given). For instance, Euclidean geometry and the vector character of forces are principles of coordination in classical mechanics, and a specific law of force is a law of combination. In general relativity, the differential manifold and the rules of tensor calculus on this manifold are principles of coordination, and Einstein's equations relating the (derivatives of) the metric tensor with the energy-momentum tensor are laws of combination. In the latter theory, the coordination depends on an arbitrary choice of coordinates and does not require a fixed given metric. On the one hand, this arbitrariness shows the necessity of a subjective form in the physical description. On the other hand, it shows that there are equivalent coordination frameworks. These are equally univocal and they are related by differentiable coordinate transformations. The invariants of the theory under these transformations define the objective content of reality (*den objektiven Gehalt der Wirklichkeit*) according to Reichenbach.[25]

As history teaches us, the principles of coordination do not share the apodictic certainty of Kant's a priori. Radically new theories such as Einstein's two theories of relativity require new principles of coordination. Future theories may require still different principles of coordination, as Reichenbach inferred from Weyl's contemporary proposal of a variable gauge. Owing to the constitutive value of these principles, any such change implies a new mode of constituting the object of knowledge. In each such change Reichenbach sees a closer and closer approximation to reality.[26]

To sum up, from Kant Reichenbach retains the idea of constitutive principles that define the object of knowledge. He departs from Kant by allowing these principles to vary in the history of physics. However, he admits

25. Reichenbach, ref. 23, pp. 51–52 (*Zuordnung*/*Verknüpfung*), p. 86 (*subjektiv*/*objektiv*): „So ist es offenbar nicht in dem Charakter der Wirklichkeit begründet, daß wir sie durch Koordinaten beschreiben, sondern dies ist die subjektive Form, die es unserer Vernunft erst möglich macht, die Beschreibung zu vollziehen."
26. Reichenbach, ref. 22, Chap. 7.

the absolute meta-principle of the univocal character of the coordination provided by the constitutive principles. Although he rejects Kant's definition of intuition as a passive, pre-structured theater of our representations, he maintains a notion of space and time coordinates as a subjective form of description from which the true objects are extracted by a theory of invariants. At the end of his book, he proposes to replace the Kantian "analysis of reason" with the procedure of extracting invariants from the subjective form of description. This procedure would replace Kant's transcendental deduction of categories, and coordination would presumably take the place of Kant's schematism: where Kant sees the application of concepts to sensible experience, Reichenbach sees the extraction of invariants from equivalent coordinations.[27]

There are some evident weaknesses in Reichenbach's notion of coordination: It is not clear how the coordination between the mathematical formalism and the theory and empirical reality is effectively done; it is not clear how the principles of coordination should be chosen and how they permit univocal coordination; and it is not clear how the univocal character of the coordination can be tested without knowing what the measured quantities and the measurements should be in Reichenbach's imagined tests. This may be why he later gave up this notion and turned to an empiricist-conventionalist philosophy in which constitutive principles no longer had a place.

CASSIRER'S RULES OF UNDERSTANDING

In 1921, Reichenbach's former Berlin teacher Ernst Cassirer published his own conciliation of Kant's system with Einstein's general relativity. In some important respects, Cassirer's views agree with Reichenbach's. For Cassirer too, the most important legacy of Kantianism is the insight that the objects of knowledge are constituted by the mind. In this constitution, the focus on invariants with respect to various reference frames is essential. The constitutive principles, which include Cassirer's rules of the understanding

27. Reichenbach, ref. 22, p. 88: „Das Verfahren, durch Transformationsformeln den objektiven Sinn einer physikalischen Aussage von der Subjektiven Form der Beschreibung zu eliminieren, ist, indem es indirekt diese subjektive Form charakterisiert, an Stelle der Kantischen Analyse der Vernunft getreten. Es ist allerdings ein sehr viel komplizierteres Verfahren als Kants Versuch einer direkten Formulierung, und die Kantische Kategorientafel muß neben dem modernen invarianten-theoretischen Verfahren primitiv Erscheinen."

(*Regeln des Verstandes*), forms of thought (*Denkformen*), or ordering forms (*Ordnungsformen*) are not immutable. In major theoretical breakthroughs such as Galilean mechanics or Einstein's relativity, they undergo a fundamental revision: they are "living and moving forms" (*lebendige und bewegliche Formen*).[28]

There are significant differences, however. Although Reichenbach clearly sees that constitutive principles have to do with procedures of measurement (since they changed when space and time measurement were conceived differently), he defines coordination *in abstracto* and uses measurement only as a test for the univocal character of the coordination. In contrast, Cassirer (like Ernst Mach and Max Planck) regards measurement as a basic precondition of any scientific knowledge of the physical world. Of course he does not understand measurement in a naive operational manner. On the contrary, he asserts the unavoidable ideal component in the definition of any measurement and he regards the critical analysis of this component as the main task of transcendental philosophy.[29]

A second characteristic of Cassirer's philosophy of knowledge is its being based on the basic rule of the understanding from which Kant deduces his table of categories: the synthetic unity of apperception. For Cassirer as for Reichenbach, raw sensations are completely amorphous and the object of knowledge needs to be constituted by the mind. For Reichenbach, this is done by univocal coordination between mathematical structures and physical reality. For Cassirer, this is done by synthesizing the perceptually diverse in a systematic unity. This rule of understanding precedes any other constitutive principles and it is not subject to revision. Cassirer calls it coordination (*Zuordnung*). His coordination, unlike Reichenbach's, is no to be understood as a relation between a theoretical representation and physical reality. Rather, it is the demand of synthetic unity that presides over any such representation.[30]

A related characteristic of Cassirer's approach is the de-substantification or de-reification of the object of knowledge as advocated in his *Substance and function* of 1910. In his view, substances, things, images, and Kant's

28. Ernst Cassirer, *Zur Einstein'schen Relativitätstheorie. Erkenntnistheoretische Betrachtungen* (Berlin, 1921), pp. 82 (*Regel des Verstandes*), 88 (*Denkformen*), 58 (*Ordnungsform*), 87 (*lebendige und bewegliche Formen*). Cf. Ryckman, ref. 22, pp. 39–46.
29. Cassirer, ref. 28, pp. 14, 75.
30. Cassirer, ref. 28, pp. 41, 84–85.

intuition all are naive or temporary products of an unfinished synthesis. Thorough synthetic unity dissolves them into functional systems of relations. In the relativity book of 1921, we learn that relativity theory has produced just the desired dissolution in the case of space and time by attending to the preconditions of their measurement. Here is how it goes.[31]

For Kant, space is the expression of synthetic unity with regard to coexistence and proximity, and time is the expression of synthetic unity with regard to succession. Space and time are not objects of perception, they are intellectual preconditions for the constitution of objects of perception. There is no absolute space and there is only a generic (Euclidean) space defined with respect to an arbitrary reference system (moving or not). The synthetic unity of space implies that the same distance measurements apply in any system. Similarly, the synthetic unity of time implies that time measurements are the same in any reference frame. The problem with this view, Cassirer tells us, is that it reifies distinctions that are contingent presuppositions of measurement in Newtonian physics. The absence of effects of the earth's motion on optical phenomena and the isotropy of the propagation of light in one frame suggest the "heuristic maxim" of the relativity principle and the "rule of understanding" of the constant velocity of light. Consequently, the separate unities of time and space are lost but the higher synthetic unity of Minkowskian spacetime emerges through the Lorentz transformations that connect the time and space measurements in different reference frames. In Cassirer's assessment, an arbitrarily reified distinction has thus been turned into a more unified system of functional relations.[32]

There still is, in special relativity, a reification of a privileged class of reference systems. The next de-reification occurs thanks to a new "prescription for the formation of concepts," the equivalence between gravitation and acceleration. This equivalence prohibits the existence of global rigid frames and makes space and time measurement a necessarily local affair. A remaining precondition for this measurement is the four-dimensional continuous manifold of events, which Cassirer regards as the ultimate expression of the synthetic unity of space and time. A widely enlarged group of

31. Cassirer, *Substanzbegriff und Funktionsbegriff: Untersuchungen über die Grundfragen der Erkenntniskritik* (Berlin, 1910).
32. Cassirer, ref. 28, pp. 81–86. By *Regel des Verstandes*, Cassirer means „ein Grundsatz , der den Verstand in der Deutung der Erfahrungen hypothetisch als Norm der Forschung gebraucht" (p. 82).

transformations, the coordinate changes of the manifold captures the new conquered unity. That is not to say that the metric properties of space and time become accessory. The differential manifold structure is only there as a precondition for the expression of these properties. The error of Newton and Kant was not their requiring both the manifold and metric properties of space and time, it was their regarding a specific metric as inherent in the pure intuition of space and time. In general relativity, this is no longer possible since the metric properties depend on the distribution of matter. The first demand of synthetic unity is the "coordination under the viewpoint of coexistence and proximity or under the viewpoint of succession" and it only involves the differential manifold, not the metric structure.[33]

Cassirer and Reichenbach both see three stages in the evolution of our concepts of space and time, corresponding to Newtonian mechanic, special relativity, and general relativity; and they characterize each stage by its constitutive principles. With some extrapolation and modernization of their identification of these principles, we might say that the Galilean group, the Lorentz group, and the group of diffeomorphisms respectively constitute the object of Newtonian mechanics, special relativity, and general relativity. For Reichenbach, these groups play a double role: they warrant the univocality of the coordination between the mathematical apparatus of the theory and physical reality, and they interconnect equivalent coordinations. For Cassirer, these groups express the synthetic unity of apperception with regard to the relative nearness of events, the better with the less metric background.

We may now return to comparing the ways in which Reichenbach and Cassirer depart from Kant. They both retain notions of space and time as our most basic way of coordinating phenomena, and they both reject Kant's interpretation of these notions in terms or a passive, pre-structured faculty of the mind. In Reichenbach's case, the rejection results from the need of relativizing coordination in two manners: by making the principles of coordination depend on the empirical domain, and by recognizing the possibility of equivalent coordinations. In Cassirer's case, this rejection is the natural consequence of the Marburg neo-Kantian doctrine that

33. Cassirer, ref. 28, p. 106 (*Vorschrift für unsere physikalische Begriffsbildung*), 85 (*Das Zuordnen unter dem Gesichtspunkt des Beisammen und des Nebeneinander oder unter dem Gesichtspunkt des Nacheinander*).

sensibility and understanding have a common origin in the synthetic unity of apperception. Synthetic unity is a higher regulative principle without which perception and knowledge would be inconceivable. This principle generates the concepts of space and time when combined with more provisional principles that have both a constitutive and a constructive function. Coordination is nothing but synthetic unity, and any attempt to define it by reference to a non-conceptual entity, be it Kant's intuition or Reichenbach's reality, is doomed to fall into psychologism.

FRIEDMAN'S RELATIVIZED A PRIORI

In recent years the Stanford-based philosopher Michael Friedman has revived Reichenbach's early neo-Kantianism under the "relativized a priori" label. Friedman's main target his Quinean holism, which in his view misses the most essential aspects of the structure and evolution of modern theoretical physics. In his opinion, Quine's objection to any distinction between the formal and empirical components of a theory only applies to the linguistic, Carnapian view of physical theories in which this distinction is meant to be purely logical. This objection does not apply to the distinction between "constitutive principles" and "properly empirical laws," which Friedman takes to be a central feature of any advanced physical theory. Like Reichenbach's principles of coordination, Friedman's constitutive principles serve to constitute the object of scientific knowledge but they do not have the apodictic certainty of Kant's a priori. They may undergo radical changes in revolutionary circumstances. More precisely, Friedman defines his constitutive principles as basic preconditions for the mathematical formulation and the empirical application of a theory. These principles are not refutable because a refuting experiment could not be conceived without them. Reasons for rejecting them can only be pragmatic or meta-theoretical.[34]

Like Reichenbach and Cassirer, Friedman tries to capture some rationality in the transition from one set of constitutive principles to the next and he has the regulating idea of an asymptotic convergence toward stable principles. For Reichenbach, intertheoretical approximation and consistency between the constitutive principles and a higher principle of normal

34. Friedman, *Dynamics of reason: The 1999 Kant lectures at Stanford University* (Stanford, 2001), pp. 40–41 (anti-Quine); pp. 20, 37–40, 76–80 (constitutive principles).

induction control such transitions. For Cassirer, synthetic unity by means of increasing de-substantification and de-reification is the rational motor of change. For Friedman, it is a philosophical meta-framework that physicists find in contemporary philosophical writings.

Friedman's choice of constitutive principles in the three standard examples of Newtonian physics, special relativity, and general relativity somewhat differ from Reichenbach's and Cassirer's choices. For Newtonian physics, the constitutive principles are Euclidean geometry and Newton's laws of motion, in conformity with Kant's doctrine. The laws of motion, in Friedman's view, are by themselves void of empirical content and so is too Newton's law of gravitation by itself because force and acceleration only acquire physical meaning through the laws of motion. For special relativity, Friedman's constitutive principles are the light principle, the relativity principle, and the mathematics needed to develop the consequences of these principles. For general relativity, the relevant principles are the Riemannian manifold structure, the light principle (used locally), and the equivalence principle (understood as the statement that free-falling particles follow geodesics of the Riemannian manifold); Einstein's relation between the Riemann curvature tensor and the energy-momentum tensor is regarded as a "properly empirical law," whose content cannot be expressed and tested without prior formulation of the constitutive principles.[35]

One might object that some Friedman's constitutive principles, notwithstanding with Friedman's claim of irrefutability, are or contain empirical laws. Friedman anticipates this objection and replies in a Poincarean manner: some constitutive principles do have antecedents that were empirical laws but they have been "elevated" to a higher status in which they become conventions for the construction of the new theory. Let us see how it works in the three standard examples. In the case of Newtonian physics, the laws of motion do have empirical content. For instance, the law of inertia implies the testable existence of a reference system in which all free particles travel in straight lines and travel proportional distances in equal times (granted that global synchronization is possible); and the law of acceleration can be tested by comparing the observed motion in an inertial frame with the static measure of the force. Friedman would reply that the derived constitutive principles differ from empirical laws by an element of decision or convention

35. Friedman, ref. 33, pp. 77 (mechanics), 79–80 (relativity)

that makes them rigid preconditions for the description of any mechanical behavior and for the expression of further empirical laws such as the law of gravitation. At least this is Friedman's reply to the similar difficulty for the light principle and for the relativity principle in the case of special relativity, and for the equivalence principle in the case of general relativity.[36]

Another difficulty concerns the rationality of changes in the systems of constitutive principles. As Friedman has these changes depend on the intellectual context of the time (especially the philosophical debates), he introduces an element of historical contingency that seems incompatible with a rationalist idea of scientific progress. Friedman counters this objections in two different ways: by insisting on the rational demand that the earlier system should in some sense be an approximation of the earlier one (as Reichenbach had earlier done), by describing the inner logic of each intellectual context, and by showing a natural evolution of each of these contexts from Kant's original transcendentalism.[37]

FROM CONSTITUTIVE PRINCIPLES TO COMPREHENSIBILITY CRITERIA

For those who dare to pronounce the end of theory, the history of transcendental philosophy from Kant to Friedman teaches an important lesson: No object of knowledge is reduced to the mere conjunction of observational data; the object must be constituted by the mind before any empirical investigation becomes possible. As philosophers are very fond of chairs and tables, let us recall the classic argument that when we see a chair we see nothing but an amorphous and changing set of points of diverse luminous intensities and colors unless we already have a concept of chair through which we recognize regularities in this set of points. Kant and his followers simply extrapolate this observation: just as there is no perception without concepts, there is no advanced knowledge of the physical world without constitutive principles.

The problem with this view is the difficulty of a precise and effective characterization of the constitutive principles. Kant's original notion was

36. Friedman, ref. 33, pp. 86–91. On Poincaré and principles, cf. Príncipe, ref. 2, and his contribution to this volume.
37. Friedman, "Einstein, Kant, and the relativized *a priori*," in Michel Bitbol, Pierre Kerszberg, and Jean Petitot 2009, *Constituting objectivity: Transcendental perspectives on modern physics* (Berlin, 2009), 253–267.

too tied to Newtonian physics to survive the later evolution of physics. The more flexible and more evolvable notions of Reichenbach, Cassirer, and Friedman seem more apt to capture the constitutive component of modern theories. On the one hand, Reichenbach and Friedman define the constitutive principles as the means through which we connect a mathematical structure with physical phenomena. On the other hand, Cassirer defines the constitutive (or regulative) principles as the means through which we satisfy the basic demand of synthetic unity in diverse contexts of measurement. The first view has the advantage of addressing the basic and yet notoriously inscrutable difference between a mathematical theory and a physical theory, and the defect of relying on an ill-defined concept of coordination. The second has the advantage of being based on a universally acceptable demand of unifying synthesis and measurability, and the defect of leaving measurement and physical interpretation in the dark.

Let us look more closely at the coordination between theory and the physical world in the first view. For Reichenbach, constitutive principles are the warrants of the univocality of coordination. As Cassirer saw, this notion is problematic because it cannot make sense without some preconceived idea of what is being coordinated in the physical world, an idea that seems to lead either to naive realism or to psychologism.[38] Friedman avoids Reichenbach's notion of univocality and instead propounds an intra-theoretical characterization of constitutive principles as the component of the theory that is needed to express and test the properly empirical laws of the theory. This leaves us with a number of questions: How selective is this characterization? Where do the constitutive principles come from? Do they have some sort of necessity or are they merely convenient conventions? How do they connect the mathematical formalism to the world of experience?

The first question is about the legitimacy of Friedman's distinction between constitutive principles and properly empirical laws. According to Friedman, some constitutive principles such as the light principle or the equivalence principle do have empirical content, but unlike ordinary empirical laws they are regarded as preconditions for the expression of any other empirical law. The problem with this view is that it seems to rely on a subjective "decision." For instance in Newtonian physics, why could not

38. Cf. Ryckman, ref. 22, p. 27.

we regard both the acceleration law and the law of gravitation as properly empirical laws in a constitutive framework defined by Euclidean geometry and Newton's first and third law? Pace Kant, this was the view of Daniel Bernoulli and of later textbook writers who asserted the empirical character of the second law. In order to justify a constitutive role of this law, we would have to demonstrate that it is in some sense more necessary than the law of gravitation. Neither Kant nor Friedman do so.

The same difficulty occurs in the context of general relativity. Here Friedman regards the geodetic principle (according to which free-falling particles follow geodesics of the spacetime manifold) as constitutive and Einstein's relation between the curvature tensor and the energy-momentum tensor of matter as properly empirical. Why could not we also regard this relation as constitutive, and confine the properly empirical in the expression of the energy-momentum tensor? This seems to be the more natural choice for physicists accustomed to regard the principle of least action as constitutive, because this principle together with general covariance and with plausible simplicity assumptions lead to both the geodetic principle and the Einstein field equations. Unless further arguments are provided in favor of its defense, Friedman's notion of constitutive principles seems dangerously close to the Quinean psycho-cognitive notion of "entrenched principles," which is precisely what Friedman wanted to avoid.

The second question is about the origin of the constitutive principles. Some of the principles, for instance Euclidean geometry in Newtonian physics or the pseudo-Riemannian manifold structure in general relativity are mathematical preconditions for the formulation of the theory and Friedman regards them as external mathematical resources on which physicists rely when they have philosophical reasons to do so. The rest of the constitutive principles, Friedman tells us, are the "coordinating principles" obtained by elevating empirical laws to a higher constitutive status. Not every empirical law is a candidate for this elevation, only laws that have sufficient generality and that appear to be problematic in a proper philosophical meta-framework. As Wolfgang Goethe wrote, "The highest art in intellectual life and in worldly life consists in turning the problem into a postulate that enables us to get through." [39]

39. Goethe to Zelter, cited in Cassirer 1921, pp. 30–31: „Die größte Kunst in Lehr- und Weltleben besteht darin, das Problem in ein Postulat zu verwandeln, damit kommt man durch."

This sounds fine except that there may be other ways to get through. For instance, in the case of special relativity it became clear to Poincaré, in 1900, that the velocity of light as measured by moving observers would be the same as in the ether frame and that, as a consequence, the time measured by moving observers who synchronize their clocks by optical means would differ from the time measured in the ether frame. In 1905 this insight led Poincaré to a version of the theory of relativity that was empirically equivalent to Einstein's slightly later theory, and yet Poincaré refused to regard the light principle (constancy of the velocity of light in the ether frame) and the relativity principle as constitutive with respect to the definition of space and time. In general, the underdetermination of theories by experimental data leads to several equivalent options in solving the same difficulties, and these options have different constitutive principles. The choice between these options is largely contingent. This remark also answers my third question about the necessity of the constitutive principles. They are necessary only within a given, philosophical meta-framework. In physics narrowly considered, they are conventions the suitability of which is a matter of convenience.

The fourth and last question is about the manner in which the constitutive principles connect the mathematical apparatus of a theory with the world of experience. The relevant principles are Friedman's "coordinating principles." Are these principles truly sufficient to determine the applications of the theory? The textbook definition of the major physical theories suggests so much, because this definition typically involves a "mathematical formalism" and a few "rules of interpretation" which play a role somewhat similar to Friedman's coordinating principles. Also, the theories Friedman discusses, which are Newtonian mechanics and relativity theory, seem to provide for their own interpretive resources, unlike other theories such as electromagnetic theory, statistical mechanics, fluid mechanics, or quantum mechanics. In my terminology, their interpretation involves a relatively small amount of modular structure.

Yet a closer analysis of the physicists' practice suggests that constitutive principles do not suffice to define the application of theories. Again, it is not by contemplating the textbook definition of a theory that physicists learn how to apply it; it is by applying the theory to a series of exemplars given in any good textbook. That is not only for pedagogical reasons. Knowledge is needed that is not contained in the bare rules of interpretation, even in the simple cases of classical mechanics and relativity theory.

Consider the case of general relativity. Any effect involving concrete clocks (for instance the gravitational redshift of spectral lines or the gravitational slowing down of atomic clocks) requires the strong equivalence principle according to which the laws of physics in a free-falling, non-rotating local reference frame are the same as in the absence of gravitation. This principle is not contained in the weak principle that Friedman takes to be the basic coordinating principle (in addition to the light principle). Friedman would perhaps have no objection to accepting the strong principle as one of the constitutive principles. Note, however, that this principle differs from Friedman's other coordinating principles by implicitly involving phenomena (and theories) that do not belong to the official domain of the theory under consideration.

Admittedly, there are many concrete consequences of general relativity that do not principally involve extra-theoretical time gauges. But even in such cases more is needed than just the light principle and the equivalence principle to interpret relevant experiments. Optical instruments and goniometric techniques are used. Their implying electromagnetic theory is not much of a problem, because additional constitutive principles could be introduced for the electromagnetic sector of the theory. What is more problematic is the fact that the global theory, as long as it is defined only by its constitutive principles and its general laws does not provide the means to *conceive* the experimental setups through which it is applied. For this purpose we rely on previous theoretical and practical knowledge that remains regionally valid. One might retort that this knowledge is implicitly contained in the new theory because the older theories are regional approximations of the new theory in some operationally meaningful sense. This sort of reductionism is a will-o'-the-wisp because we would not consider the regional approximations without having previous knowledge of the relevant region of experience and because, even if we chanced to consider these approximations for purely formal reasons, we would thus access only the formal apparatus of the earlier theories, not the associated laboratory practice. Moreover, the reductions would require the entire theory, not only the constitutive principles, so that the allegedly non-constitutive, empirical component of the theory indirectly plays a coordinating role in defining local conditions of measurement.

We thus see that the distinction between coordinating principles and properly empirical laws is problematic. Yet the general idea that the object of inquiry may be differently constituted in different physical theories seems intuitively sound: for instance, anyone would agree that the Newtonian theory of gravitation and Einstein's general relativity differ at the very deep level of space and time concepts and that this difference is not the result of an empirical induction. What can we do to save the relativized a priori from the dilemmas of coordination?

In sections 1, I have proposed a definition of physical theory that involves a symbolic universe, its laws, and interpretive schemes. The laws only serve to restrict the possible states and evolutions of the symbolic universe. They do not yet have any concrete value and they could very well be included in the definition of the symbolic universe. Their extraction from this definition is a matter of convenience. The interpretive schemes, not any law or principle, are responsible for the coordination between formalism and concrete experiments. Unlike Friedman's coordinating principles, the interpretive schemes are not an invariable component of the theory: they vary with the history of its applications. The symbolic universe and its laws control their form without dictating it. As argued in section 2, their definition and their deployment requires a modular structure of the theory that itself evolves in the history of applications of the theory. The interpretive schemes are blueprints for possible experiments in which the correlation between various modularly-defined quantities is tested or exploited.

Being variable and partly contingent, the interpretive schemes cannot replace Friedman's coordinating principles in defining a constitutive a priori component of the theory. In section 3, however, we saw that coordination by means of interpretive schemes and modules enables comprehensibility arguments from which the structure of a given theory may derive in a given domain of experience. The comprehensibility criteria offer a workable definition of the relativized a priori. They indeed proceed from very broad and natural assumptions on the expected or desired regularity of the world, and they vary when the domain of inquiry is changed or extended. As we saw these assumptions include forms of causality, measurability, and correspondence. They are not components of the theory (as Friedman's constitutive principles would be), and they do not originate in the enigmatic character of former empirical laws (as Friedman's coordinating principles do). They are regulative principles for the construction of the theory and may be regarded as avatars of Cassirer's higher regulative principle of synthetic unity. The emphasis on measurability is also Cassirer's, although Cassirer lacked the modular means do discuss the implications of measurability in a given domain before knowing the theory of this domain.

As for the correspondence idea, it is already important in Reichenbach's and Friedman's rationalizations of the transition from one system of constitutive principles to the next.

Regarding this last point, there is an interesting contrast between comprehensibility arguments and constitutive principles. For the most, Reichenbach and Friedman first obtain their constitutive principles by inspection of a standard sample of theories (Newtonian physics, special relativity, and general relativity); and then they introduce intertheoretical considerations of correspondence in order to rationalize changes of the constitutive principles. Friedman compares these changes to Thomas Kuhn's changes of paradigms: they involve global, holistic changes of our way of conceptualizing the world and they challenge our ability to interpret these changes rationally. In the approach defended in this essay, the intertheoretical relations of modular structure come first as a most basic requirement of applicability and communicability of theoretical knowledge; then this structure is used to express comprehensibility criteria that show the necessity of some our best theories. The relevant picture of scientific change crucially differs from Kuhn's idealized historiography: it involves much substructure within Kuhn's alleged theoretical wholes, and much modular continuity during Kuhn's alleged revolutions despite discontinuity in basic theoretical concepts and principles.[40]

Thanks to the modular substructure, comprehensibility criteria can be expressed precisely and mathematically in given domains of physics. In favorable cases, they uniquely determine the theory or the principles of this domain. The claim would be enormous if it implied our ability to discover physical theories without consulting experiments. In reality the comprehensibility criteria, no matter how natural they may seem at a given time of history and in a given domain of experience, were hard conquests of empirically motivated inquiries and they are likely to be relaxed when the domain of experience is enlarged to include extreme scales. For instance, the basic criterion of quantitative description through measurable quantities was not a general criterion of physics until the second half of

40. On the weaknesses of the Kuhnian picture of scientific change, cf. Martins, "The Kuhnian 'revolution' and its implications for sociology," in Albert Hanson, Thomas Rossiter, and Stein Rokkan (eds.), *Imagination and precision in the social sciences: Essays in honor of Peter Nettl* (London, 1972), 13-58, on 20, 24-25, 35 (in this volume: 20, 23-25, 34); Galison, ref. 14, pp. 781–802.

the nineteenth century; and the classical expression of causality by strict correlation had to be to downgraded to a merely statistical causality in the quantum domain. All we can say is that some of our best theories, no matter how empirically and culturally conditioned their genesis was, can retrospectively be deduced from simple requirements for the comprehensibility of the world. These requirements are refutable since their consequences are so. Simple though they are, their necessity remains domain-dependent.

CONCLUSIONS

Elementary considerations on the nature of ordinary perception lead to the conclusion that every knowledge is necessarily conceptual. The economy of knowledge further requires our ability to integrate previous synthetic knowledge. At the advanced stage of knowledge production that we call theory, we need to preserve and integrate earlier successful theories as modules of the new theories. In physics, any advanced theory is based on a mathematical formalism for reasons that have to do with quantitative measurement and with the homogeneity of elementary phenomena. This raises the question of the coordination of the mathematical formalism with concrete phenomena. The answer requires an evolving class of interpretive schemes that function as blueprints of experimental setups and rely on the modular structure of the theory. In favorable cases, simple notions of causality, measurability, and modularity enable us to infer the theory from associated constraints on its interpretive schemes. Hence the necessity of physical theories not only includes the indispensability of earlier successful theories, but it may also include the uniqueness of the theory that fits a domain defined by a given type of comprehensibility.

ACKNOWLEGMENT

I am grateful to João Príncipe for stimulating discussions.

SUMMARY

It has recently become fashionable to dismiss theories as antiquated instruments of knowledge and to advertize more powerful and more universal engines of automated, computer-assisted prediction. Against this view, I argue a modular conception of knowledge in which the best physical theories of the past will forever play a significant role in our harnessing of nature, no matter how insufficient these theories turn out to be in extreme regions of experience. I further allege that these theories derive from natural (though refutable) preconditions for the comprehensibility of the domain to which they apply, there being no sharp distinction between comprehending and harnessing. I compare this constitutive claim with various attempts to relativize Kant's a priori.