Janus–faced physics: On Hilbert's 6th Problem.

N. C. A. da Costa
F. A. Doria*

Institute for Advanced Studies, University of São Paulo.
Av. Prof. Luciano Gualberto, trav. J, 374.
05655–010 São Paulo SP Brazil.

Academia Brasileira de Filosofia
R. do Riachuelo, 303
20230–011 Rio de Janeiro RJ Brazil

ncacosta@usp.br
fadoria@filosofia.org.br
fadoria@gmail.com
fadoria@pep.ufrj.br

Version 6.6
April 22, 2007

*Address after Feb. 1, 2007: Fuzzy Sets Lab, Production Engineering Program, COPPE–UFRJ, P.O.Box 68507, 21945–972 Rio RJ Brazil.

If geometry is to serve as a model for the treatment of physical axioms, we shall try first by a small number of axioms to include as large a class as possible of physical phenomena, and then by adjoining new axioms to arrive gradually at the more special theories. ... The mathematician will have also to take account not only of those theories coming near to reality, but also, as in geometry, of all logically possible theories. We must be always alert to obtain a complete survey of all conclusions derivable from the system of axioms assumed.

D. Hilbert, 1900.

Dedicated to
Greg Chaitin
on his 60th birthday.

1 Prologue

[1]

Why axiomatize a scientific theory? Doesn't the extra rigor required carry with itself an unwanted burden that hinders the understanding of the theory's concepts?

We axiomatize a theory not only to better understand its inner workings but also in order to obtain metatheorems about that theory. We will therefore be interested in, say, proving that a given axiomatic treatment for some physical theory is incomplete (that is, the system exhibits the incompleteness phenomenon), among other things. As a follow–up, we would also like to obtain examples, if any, of physically meaningful statements within that theory that are formally independent of its axioms.

Out of the authors' previous work [15, 16, 17, 18] we describe here a technique that allows for such an axiomatization. Its main guidelines are:

- First, the mathematical setting of the theory is clarified, and everything is formulated within the usual standards of mathematical rigor.

[1] *Abstract.* We argue that physics can be seen as a Janus–faced discipline whose theories may be conceived as a kind of computational device, as suggested by Chaitin, which is then complemented by a conceptual play in the sense that we elaborate here. We take our cue from Hilbert's 6th Problem (the axiomatization of physics) and present an axiomatic unified treatment for classical physics (classical mechanics, first–quantized quantum mechanics, electromagnetism, general relativity and gauge field theory) based on what we call Suppes predicates. We then obtain several undecidability and incompleteness results for the axiomatic systems which are developed out of a simple set of concepts where we have cradled significant portions of physics and show that they imply the undecidability of many interesting questions in physics and the incompleteness of the corresponding formalized theories. We give chaos theory as an example. Those results point towards the conceptual depth of the 'game of physics.' Our presentation follows several results that we have obtained in the last two decades.

- We then formulate those results within an adequate axiomatic framework, according to the prescriptions we present in this paper.

We may also be interested in semantic constructions. As the required step to obtain those results, we show here how to embed a significant portion of classical physics within a standard axiomatic set theory such as the Zermelo–Fraenkel system together with the Axiom of Choice (ZFC set theory). By classical physics we mean classical mechanics as seen through the analytico–canonical (Lagrangian and Hamiltonian) formalism; electromagnetic theory; Dirac's theory of the electron — the Schrödinger theory of the electron is obtained through a limiting procedure; general relativity; classical field theories and gauge field theories in particular.

Then, it is possible to examine different models for that axiom system and to look for sentences that are true or not depending on their interpretation — and that hopefully have corresponding different physical interpretations.

It is obvious that the crucial idea is the rather loose concept of "physically meaningful sentence." We will not try to define such a concept. However we presume (or at least hope) that our main examples somehow satisfy that criterion, as they deal with objects defined within physical theories, and consider problems formulated within the usual intuitively understood mathematical constructions of physics. Chaitin says that mathematics is random at its core. We show here that mathematics and the mathematically–based sciences are also pervaded by undecidability and by high–degree versions of incompleteness when axiomatized.

2 Hilbert's 6th Problem

When we discuss the possibility of giving physics an axiomatic treatment we delve into an old and important question about physical theories [12, 68]. The sixth problem in Hilbert's celebrated list of mathematical problems sketches its desirable contours [34]:

> *The Mathematical Treatment of the Axioms of Physics.*
>
> The investigations on the foundations of geometry suggest the problem: *to treat in the same manner, by means of axioms, those physical sciences in which mathematics plays an important part; in the first rank are the theory of probability and mechanics.*
>
> As to the axioms of the theory of probabilities, it seems to me to be desirable that their logical investigation be accompanied by a rigorous and satisfactory development of the method of mean values in mathematical physics, and in particular in the kinetic theory of gases.
>
> Important investigations by physicists on the foundations of mechanics are at hand; I refer to the writings of Mach..., Hertz..., Boltzmann..., and Volkman... It is therefore very desirable that

the discussion of the foundations of mechanics be taken up by mathematicians also. Thus Boltzmann's work on the principles of mechanics suggests the problem of developing mathematically the limiting processes, those merely indicated, which lead from the atomistic view to the laws of continua. Conversely one might try to derive the laws of motion of rigid bodies by a limiting process from a system of axioms depending upon the idea of continuously varying conditions on a material filling all space continuously, these conditions being defined by parameters. For the question as to the equivalence of different systems of axioms is always of great theoretical interest.

If geometry is to serve as a model for the treatment of physical axioms, we shall try first by a small number of axioms to include as large a class as possible of physical phenomena, and then by adjoining new axioms to arrive gradually at the more special theories. At the same time Lie's principle of subdivision can perhaps be derived from the profound theory of infinite transformation groups. The mathematician will have also to take account not only of those theories coming near to reality, but also, as in geometry, of all logically possible theories. We must be always alert to obtain a complete survey of all conclusions derivable from the system of axioms assumed.

Further, the mathematician has the duty to test exactly in each instance whether the new axioms are compatible with the previous ones. The physicist, as his theories develop, often finds himself forced by the results of his experiments to make new hypotheses, while he depends, with respect to the compatibility of the new hypotheses with the old axioms, solely upon these experiments or upon a certain physical intuition, a practice which in the rigorously logical building up of a theory is not admissible. The desired proof of the compatibility of all assumptions seems to me also of importance, because the effort to obtain such a proof always forces us most effectively to an exact formulation of the axioms.

3 A review of axiomatization techniques

There are two widely known basic procedures to axiomatize a mathematically–based theory, such as theories in physics, economics or the ecology of competing species:

- We either use a syntactical approach à la Bourbaki [45]; or
- We follow a semantic approach, with the help of what we have called Suppes predicates.

Both methods are, in effect, essentially equivalent (see [8, 14, 62]). Mathematical structures and the Suppes set–theoretical predicates are of course essentially formulated in the language of set theory.

Our goal is to follow the programme sketched in Hilbert's 6th Problem. We propose here an axiomatic treatment for physics that encompasses the whole of classical physics — classical mechanics and classical field theories — plus first–quantized theories like Dirac's theory of the electron and its non–relativistic counterpart, Schrödinger's theory.

Hilbert stresses that we have to take into consideration:

> [...] not only [...] those theories coming near reality, but also, as in geometry, [...] all logically possible theories.

So we will delve here with incompleteness and other metamathematical phenomena in the realm of physics.[2]

Models

Notice that we deal here with different concepts for which we loosely employ the same word "model":

- As interpretation of a formal language, as in the domain which is usually called "theory of models." This is the usual meaning for model.
- As in "model of, or interpretation for a domain of the physical world." This is the customary informal meaning of the word for physicists, perhaps a synonym for "picture," or "description."

Context will make clear the sense we require for "model."

Physics: from informality to mathematical rigor

Our main goal here is to axiomatize portions of physics within the framework of axiomatic set theory. It is therefore interesting and worthwhile to follow the main historical development of formal treatments of physics in three central domains, namely electromagnetic theory, gravitation theory, and classical mechanics. We will point out in an informal vein how mathematical rigor evolved together with the conceptual development and clarification of those domains. The last step, that of a rigorous axiomatization, will be treated below.

The main point in our exposition is: physics, both classical and quantum, is here seen as an outcome, or as an extension of classical mechanics.[3] The Lagrangian and Hamiltonian formalisms, for systems of particles and then for

[2]Results that derive from the use of metamathematical techniques which are among other things applied to the mathematics that underlie physics had already been obtained by Maitland–Wright [49] in the early 1970s. That author investigated some aspects of the development of Solovay's mathematics [58], i. e., of a forcing model of set theory where a weakened version of the axiom of choice holds, as well as the axiom "every subset of the real line is Lebesgue–measurable." Among other interesting results, it is shown that the theory of Hilbert spaces based on that model does not coincide with the classical version. Explorations of forcing models in physics can be found in [6, 7, 46].

[3]This is the actual way most courses in theoretical physics are taught.

fields, are seen as a basic, underlying construct that specializes to the several theories considered. A course in theoretical physics usually starts from an exposition of the Lagrangian and Hamiltonian (the so–called analytico–canonical) formalisms, show how they lead to a general formal treatment of field theories, and then one applies those formalisms to electromagnetic theory, to Schrödinger's quantum mechanics — which is obtained out of geometrical optics and the eikonal equation, which in turn arise from Hamilton–Jacobi theory — and gravitation and gauge fields, which grow out of the techniques used in the formalism of electromagnetic theory. Here we use a variant of this approach.

Electromagnetism

The first conceptually unified view of electromagnetic theory is given in Maxwell's treatise, dated 1873 (for a facsimile of the 1891 edition see [50]).

Maxwell's treatment was given a more homogeneous, more compact notation by J. Willard Gibbs, and a sort of renewed presentation of Maxwell's main conceptual lines appears in the treatise by Sir James Jeans (1925, [38]). Next step is Stratton's textbook with its well–known list of difficult problems [59], and then Jackson's book, still the main textbook in the 1970s and 1980s [37].

When one looks at the way electromagnetic theory is presented in these books one sees that:

- The mathematical framework is calculus — the so–called advanced calculus, plus some knowledge of ordinary and partial differential equations — and linear algebra.
- Presentation of the theory's kernel becomes more and more compact; its climax is the use of covariant notation for the Maxwell equations.

  However covariant notation only appears as a development out of the set of Maxwell equations in the traditional Gibbsian "gradient–divergence–rotational" vector notation.

So, the main trend observed in the presentation of electromagnetic theory is: the field equations for electromagnetic theory are in each case summarized as a small set of coordinate–independent equations with a very synthetic notation system. When we need to do actual computations, we fall back into the framework of classical, 19th–century analysis, since for particular cases (actual, real–world, situations), the field equations open up in general to complicated, quite cumbersome differential equations to be solved by mostly traditional techniques.

A good reference for the early history of electromagnetism (even if its views of the subject matter are pretty heterodoxical) is O'Rahilly's tract [52].

General relativity and gravitation

The field equations for gravitation we use today, that is, the Einstein field equations, are already born in a compact, coordinate–independent form (1915/1916)

[30]. We find in Einstein's original presentation an explicit striving for a different kind of unification, that of a *conceptual* unification of all domains of physics. An unified formalism at that moment meant that one derived all different fields from a single, unified, fundamental field. That basic field then "naturally" splits up into the several component fields, very much like, or in the search of an analogy to, the situation uncovered by Maxwell in electromagnetism, where the electric field and the magnetic field are different concrete aspects of the same underlying unified electromagnetic field.

This trend starts with Weyl's theory [67] in 1918 just after Einstein's introduction in 1915 of his gravitation theory, and culminates in Einstein's beautiful, elegant, but physically unsound unified theory of the non–symmetric field (1946, see [29]). Weyl's ideas lead to developments that appear in the treatise by Corson (1953, [13]), and which arrive at the gauge field equations, or Yang–Mills equations (1954), which were for the first time examined in depth by Utiyama in 1956 [65].

An apparently different approach appears in the Kaluza–Klein unified theories. Originally unpromising and clumsy–looking, the blueprint for these theories goes back to Kaluza (1921) and then to Klein (1926, [63]). In its original form, the Kaluza–Klein theory is basically the same as Einstein's gravitation theory over a 5–dimensional manifold, with several artificial–looking constraints placed on the fifth dimension; that extra dimension is associated to the electromagnetic field.

The unpleasantness of having to deal with extraneous conditions that do not arise out of the theory itself was elegantly avoided when A. Trautmann in the late 1960s and then later Y. M. Cho, in 1975 [11], showed that the usual family of Kaluza–Klein–like theories arises out of a simile of Einstein's theory over a principal fiber bundle on spacetime with a semi–simple Lie group G as the fiber. Einstein's Lagrangian density over the principal fiber bundle endowed with its natural metric tensor splits up as Einstein's usual gravitational Lagrangian density with the so–called cosmological term plus an interacting gauge field Lagrangian density; depending on the group G one gets electromagnetic theory, isospin theory, and so on. The cosmological constant arises in the Cho–Trautmann model out of the Lie group's structure constants, and thus gives a possible geometrical meaning to its interpretation as dark energy.

Here, conceptual unification and formal unification go hand in hand, but in order to do so we must add some higher–order objects (principal fiber bundles and the associated spaces, plus connections and connection forms) to get our more compact, unified treatment of gravitation together with gauge fields, which subsume the electromagnetic field. We are but a step away from a rigorous axiomatic treatment.

Classical mechanics

The first efforts towards an unification of mechanics are to be found in Lagrange's *Traité de Mécanique Analytique* (1811) and in Hamilton's results.

But one may see Hertz as the author of the first unified, mathematically well–developed presentation of classical mechanics in the late 1800s, in a nearly contemporary mathematical language. His last book, *The Principles of Mechanics*, published in 1894, advances many ideas that will later resurface not just in 20th century analytical mechanics, but also in general relativity [33]. Half a century later, in 1949, we have two major developments in the field: C. Lanczos publishes *The Variational Principles of Mechanics*, a brilliant mathematical essay [42] that for the first time presents classical mechanics from the unified viewpoint of differential geometry and Riemannian geometry. Concepts like kinetic energy or Coriolis force are made into geometrical constructs (respectively, Riemannian metric and affine connection); several formal parallels between mechanical formalism and that of general relativity are established. However the style of Lanczos' essay is still that of late 19th century and early 20th century mathematics, and is very much influenced by the traditional, tensor–oriented, over a local coordinate domain, presentations of general relativity.

New and (loosely speaking) higher–order mathematical constructs appear when Steenrod's results on fiber bundles and Ehresmann's concepts of connection and connection forms on principal fiber bundles are gradually applied to mechanics; those concepts go back to the late 1930s and early 1940s, and make their way into the mathematical formulations of mechanics in the late 1950s. Folklore has that the use of symplectic geometry in mechanics first arose in 1960 when a major unnamed mathematician[4] circulated a letter among colleagues which formulated Hamiltonian mechanics as a theory of flows over symplectic manifolds, that is, a Hamiltonian flow is a flow that keeps invariant the symplectic form on a given symplectic manifold. The symplectic manifold was the old phase space; invariance of the symplectic form directly led to Hamilton's equations, to Liouville's theorem on the incompressibility of the phase fluid, and to the well–known Poincaré integrals — and here the advantage of a compact formalism was made clear, as the old, computational, very cumbersome proof for the Poincaré invariants was substituted for an elegant two–line, strictly geometrical proof.

High points in this direction are Sternberg's lectures (1964, [60]), MacLane's monograph (1968, [47]) and then the Abraham–Marsden treatise, *Foundations of Mechanics* [1]. Again one had at that moment a physical theory fully placed within the domain of a rigorous (albeit intuitive) mathematical framework, as in the case of electromagnetism, gauge field theory and general relativity. So, the path was open for an axiomatic treatment.

From classical to quantum mechanics

Quantum mechanics has always been snugly cradled in the classical theory, at least when considered by theoretical and mathematical physicists, far from the cloudy popular misconceptions that have surrounded the domain since its inception in the late 1920s. The Bohr–Sommerfeld quantization conditions in the first, "old," quantum theory, arise from the well–known Delaunay condi-

[4]Said to be Richard Palais.

tions in celestial mechanics; so much for the old quantum theory. The new, or Schrödinger–Heisenberg–Dirac quantum mechanics is nearly empirical in its inception [66], but when Schrödinger and Dirac appear on stage [23] we clearly see that the theory's conceptual roots and formalism arise out of classical mechanics. Schrödinger's wave equation is a kind of reinterpretation of the eikonal equation in geometrical optics, which in turn is a consequence of the Hamilton–Jacobi equation; the Dirac commutators and Heisenberg's motion equations are new avatars of well–known equations in the classical theory that involve Poisson brackets. We can also look at the motion equations:

$$\frac{dG}{dt} = \frac{\partial G}{\partial t} + \{H, G\}$$

as the definition of a partial connection given by the Hamiltonian H on a manifold [40].

A surprising technical development stems from the efforts by Wightman to place quantum mechanics and the second–quantization theories on a firm mathematical ground. The starting point here was von Neumann's view in the early thirties that quantum mechanics was a linear dynamical theory of operators on some Hilbert space. The Murray and von Neumann theory of what we now know as von Neumann algebras (1936), later expanded to the theory of C^* algebras, allowed a group of researchers to frame several quantum–theoretic constructions in a purely algebraic way. Its realization in actual situations is given by a quantum state that induces a particular representation for the system (representation is here taken as the meaning used in group theory). This is the so–called Gelfand–Naimark–Segal construction [31].

The C^* algebra approach covers many aspects of quantum field theory, and is again framed within a rigorous, albeit intuitive mathematical background. It also exhibits some metamathematical phenomena, since the existence of some very general representations for C^* algebras are dependent of the full axiom of choice.

To sum it up: physics has strived for conceptual unification during the 20th century. This unification was attained in the domains we just described through a least–effort principle (Hamilton's Principle) applied to some kind of basic field, the Lagrangian or Lagrangian density, from which all known fields should be derived.

Most of physics is already placed on a firm mathematical ground, so that a strict axiomatic treatment of the main physical theories is possible. Still, there are mathematically uncertain procedures which are part of the everyday activity of the theoretical physicist, like Feynmann integration — but in this particular example we can take Feynmann's technique as an algorithm for the generation of a series of Feynmann diagrams, that is, as a strictly symbolic computational procedure. Other theoretical physics constructs that do not have a clear mathematical formulation (e.g. Boltzmann's H–theorem) can perhaps be approached in a similar way, as when we obtain formal series expansions out of the entropy integral, while one waits for a sound mathematical formulation for it.

4 Structures, species of structures, models

We introduce here the concept of a mathematical structure. Presentation will be mostly informal for the sake of clarity, and it is quite easy to develop the ideas introduced in a rigorous way. However we will go slightly beyond what is strictly required in this essay for the sake of completeness.

Mathematical structures are usually introduced either within the framework of set theory, or within higher–order logic, that is, type theory. Both presentations turn out to be essentially equivalent; professional mathematicians (see Bourbaki [8]) and model–theorists favor the first approach, while some logicians like Russell and Carnap explored the second way [9, 10, 55].

We follow here the set–theoretic approach. We will define mathematical structures within set theory, that is to say, mathematical structures will be conceived here as set–theoretic constructs.

Mathematical structures, species of mathematical structures

Very roughly: the structure is the arena where the game is to be played; the species of structures tells us the rules of the game.

To sum it up before we start and to give a first idea of the concepts we deal with here:

- A *mathematical structure* is finite sequence of sets, the *basic sets*, and some other sets which are obtained from the basic sets through a finite number of applications of two operations, the power set operation and the Cartesian product.

- A *species of mathematical structures* is a set–theoretic predicate which is the conjunction of two parts: first, a description of the structures we shall deal with, and second, the axioms that those structures must satisfy.

We suppose that we are working within some standard system of set theory, such as, say, Zermelo–Fraenkel set theory (ZF) [8, 32].

Suppes predicates

This presentation is semi–formal. We leave aside several important concepts like transportability [8, 14] for the sake of clarity in the exposition. The usual concept of species of mathematica structures as introduced by Bourbaki [8] is a syntactical one. One of the authors formulated [14] a semantical version of it which is fully equivalent to the original construction. That second notion was called a *Suppes predicate* in the reference.

Loosely speaking, it can be described as follows. Let $\mathcal{L}_{\mathrm{ZF}}$ be the language of ZF set theory. We construct a predicate

$$P(S, X_0, X_1, \ldots, X_n),$$

that is to say, a formula of $\mathcal{L}_{\mathrm{ZF}}$ that defines a particular kind S of structures based on the sets $X_1, X_2, \ldots$. Predicate P is given by the conjunction of two pieces:

- First piece, $P_1(S, X_0, X_1, \ldots, X_n)$, shows how structure S is built out of the basic sets $X_0, X_1, \ldots, X_n$.
- Second piece, $P_2(S, X_0, X_1, \ldots, X_n)$, is the conjunction of the axioms that we wish S to satisfy.

We get:

$$P(S, X_0, X_1, \ldots, X_n) \leftrightarrow_{\mathrm{Def}} P_1(S, X_0, X_1, \ldots, X_n) \wedge P_2(S, X_0, X_1, \ldots, X_n).$$

Here $P(S, X_0, X_1, \ldots, X_n)$ is called *a species of structures on the basic sets*

$$X_0, \ldots, X_n,$$

and the predicate:

$$\exists X_0, X_1, \ldots, X_n \, P(S, X_0, X_1, \ldots, X_n)$$

is called the class of structures that corresponds to P.

5 Axiomatization in mathematics

The preceding construction sketches the required background for our formal treatment in this essay. It shows the way we will fit usual mathematical concepts like those of group or topological space within a formal framework like ZF set theory. We will in general assume that those formalizations have been done as in our examples; if required, each structure and species of structures we deal with can easily be made explicit (even if with some trouble). It is in general enough to know that we can axiomatize a theory of our interest with a Suppes predicate.

An axiomatic theory starts out of some primitive (undefined) concepts and out of a set of primitive propositions, the theory's axioms or postulates. Other concepts are obtained by definition from the primitive concepts and from defined concepts; theorems of the theory are derived by proof mechanisms out of the axioms.

Given the set–theoretic viewpoint for our axiomatization, primitive concepts are sets which are related through the axioms. We adopt here the views expressed with the help of Suppes' slogan [62], slightly modified to suit our presentation:

> *To axiomatize a theory in mathematics, or in the mathematics–based sciences, is to define a set–theoretic predicate, that is to say, a species of structures.*

More precisely, to axiomatize an informal theory is to exhibit a species of structures so that:

- The primitive terms of the theory are the basic sets and the primitive relations of the species of structures.

- The theorems of the theory are the logical consequences of the species of structures, whose primitive elements are replaced by the corresponding primitive terms of the theory.

  Proofs are made within set theory.

Let P be the set–theoretic predicate that describes our theory. The structures which are models of P are a family $\mathcal{F}_P$ that may be taken to determine P. Therefore, in order to study the theory $\mathcal{T}_P$ given by P we can either proceed syntactically out of P, or semantically, out of $\mathcal{F}_P$.

Roughly, the syntactical and semantical approaches are complementary: given $\mathcal{F}_P$ we can recover P, and vice–versa. As we define P in set theory, and as set theory can be taken as a fully axiomatic theory [8, 39], then the theory of P–structures, which is the theory of $\mathcal{T}_P$ can also be formulated (and formalized) within set theory.

Therefore, any mathematical theory of the kind considered in physics as we have described it in this paper can in principle be formalized. That is to say: given any mathematical argument, such as the proof of any major theorem within mathematics, from Euclid's proofs to the prime–number distribution theorems, to the Malgrange preparation theorem or to the ergodic theorem, anything in mathematics can be formalized with the present techniques, either within ZF set theory or within one of its extensions.

What is an empirical theory?

We may identify theories in an obvious way with sets of sentences, to which we add an empirical counterpart, that is to say, observational terms, protocol sentences, and the like. Also, we may assert that a theory is a family of structures (models). Finally, according to our viewpoint a theory may also be identified to a triple $\langle \mathcal{S}, D, R \rangle$, where $\mathcal{S}$ is a species of structures (Suppes predicates), D the set of domains of applications, and R the rules that relate $\mathcal{S}$ to R [15, 16].

These views are not mutually incompatible.

We mainly adopt this third viewpoint. We will take here an empirical theory to be a species of structures plus a domain D and the set of rules R that relate the sentences to D. However we must distinguish the two steps required in the axiomatization of an empirical theory:

- Construction of the theory's Suppes predicate;

- Characterization of D and R, a procedure that depends on the science where we find that theory.

6 Suppes predicates for classical field theories in physics

The usual formal[5] treatment for physics in s general setting goes as follows: one writes down a Lagrangian or a Lagrangian density for the phenomena we are interested at, and then use the variational principle as a kind of algorithmic procedure to derive the Euler–Lagrange equations, which give us the dynamics of the system. The variational principle also allows us to obtain a conservation–law, symmetry dependent, interpretation of interaction as in the case of the introduction of gauge fields out of symmetry conditions imposed on some field [13, 65].

We take here a slightly different approach. We describe the arena where physics happens — phase space, spacetime, fibered spaces — and add the dynamics through a Dirac–like equation.

Our results are not intended as a complete, all–encompassing, axiomatics for the whole of physics: there are many interesting areas in physics with uncertain mathematical procedures at the moment, such as statistical mechanics or quantum field theory, and the present techniques do not seem to be adequate for them. But we may confidently say that our axiomatization covers the whole of classical mechanics, classical field theory and first–quantized quantum mechanics.

We follow the usual mathematical notation in this subsection. In particular, Suppes predicates are written in a more familiar but essentially equivalent way.

The species of structures of essentially all classical physical theories can be formulated as particular dynamical systems derived from the triple $P = \langle X, G, \mu \rangle$, where X is a topological space, G is a topological group, and μ is a measure on a set of finite rank over $X \cup G$ and it is easy to put it in the form of a species of structures.

Thus we can say that the mathematical structures of physics arise out of the geometry of a topological space X. More precisely, physical objects are (roughly) the elements of X that:

- Exhibit invariance properties with respect to the action of G.

  (Actually the main species of structures in "classical" theories can be obtained out of two objects, a differentiable finite–dimensional real Hausdorff manifold M and a finite–dimensional Lie group G.)

- Are "generic" with respect to the measure μ for X.

[5] We will proceed in an informal way, and leave to the archetypical interested reader the toil and trouble of translating everything that we have done into a fully formal, rigorous, treatment of our presentation.

(This means, we deal with objects of probability 1. So, we only deal with "typical" objects, not the "exceptional" ones. This condition isn't always used, we must note, but anyway measure μ allows us to identify the exceptional situations in any construction.)

Let's now give all due details:

Definition 6.1 *The species of structures of a* **classical physical theory** *is given by the* 9*–tuple*

$$\Sigma = \langle M, G, P, \mathcal{F}, \mathcal{A}, \mathcal{I}, \mathcal{G}, B, \nabla\varphi = \iota\rangle,$$

which is thus described:

1. **The Ground Structures**. $\langle M, G\rangle$*, where* M *is a finite–dimensional real differentiable manifold and* G *is a finite–dimensional Lie group.*

2. **The Intermediate Sets**. *A fixed principal fiber bundle* $P(M, G)$ *over* M *with* G *as its fiber plus several associated tensor and exterior bundles.*

3. **The Derived Field Spaces**. *Potential space* $\mathcal{A}$*, field space* $\mathcal{F}$ *and the current or source space* $\mathcal{I}$*.* $\mathcal{A}$*,* $\mathcal{F}$ *and* $\mathcal{I}$ *are spaces (in general, manifolds) of cross–sections of the bundles that appear as intermediate sets in our construction.*

4. **Axiomatic Restrictions on the Fields**. *The dynamical rule* $\nabla\varphi = \iota$ *and the relation* $\varphi = d(\alpha)\alpha$ *between a field* $\varphi \in \mathcal{F}$ *and its potential* $\alpha \in \mathcal{A}$*, together with the corresponding boundary conditions* B*. Here* $d(\alpha)$ *denotes a covariant exterior derivative with respect to the connection form* α*, and* ∇ *a covariant Dirac–like operator.*

5. **The Symmetry Group**. $\mathcal{G} \subseteq \mathrm{Diff}(M) \otimes \mathcal{G}'$*, where* $\mathrm{Diff}(M)$ *is the group of diffeomorphisms of* M *and* $\mathcal{G}'$ *the group of gauge transformations of the principal bundle* P*.*

6. **The Space of Physically Distinguishable Fields**. *If* $\mathcal{K}$ *is one of the* $\mathcal{F}$*,* $\mathcal{A}$ *or* $\mathcal{I}$ *field manifolds, then the space of physically distinct fields is* $\mathcal{K}/\mathcal{G}$*.* $\square$

(In more sophisticated analyses we must replace our concept of theory for a more refined one. Actually in the theory of science we proceed as in the practice of science itself by the means of better and better approximations. However for the goals of the present work our concept of empirical theory is enough.)

What we understand as the classical portion of physics up to the level of first–quantized theories easily fits into the previous scheme. We discuss in detail several examples: Maxwellian theory, Hamiltonian mechanics, general relativity and classical gauge field theory. Then, what is given an abstract form will receive its usual empirical translation.

Maxwell's electromagnetic theory

Let $M = \mathsf{R}^4$, with its standard differentiable structure. Let us endow M with the Cartesian coordination induced from its product structure, and let $\eta = \mathrm{diag}(-1, +1, +1, +1)$ be the symmetric constant metric Minkowskian tensor on M.

Then M is Minkowski spacetime, the physical arena where we do special relativity theory. As it is well–known, out of the linear transformations that keep invariant tensor η we obtain the well–known relativistic contraction and dilation phenomena.

We use standard physics notation. If the $F_{\mu\nu}(x)$ are components of the electromagnetic field, that is, a differentiable covariant 2–tensor field on M, $\mu, \nu = 0, 1, 2, 3$, then Maxwell's equations are:

$$\partial_\mu F^{\mu\nu} = j^\nu,$$

$$\partial_\mu F_{\nu\rho} + \partial_\rho F_{\mu\nu} + \partial_\nu F_{\rho\mu} = 0.$$

The contravariant vectorfield whose components are given by the set of four smooth functions $j^\mu(x)$ on M is the current that serves as source for Maxwell's field $F_{\mu\nu}$. (We allow piecewise differentiable functions to account for shock–wave like solutions.)

It is known that Maxwell's equations are equivalent to the Dirac–like set

$$\nabla\varphi = \iota,$$

where

$$\varphi = (1/2)F_{\mu\nu}\gamma^{\mu\nu},$$

and

$$\iota = j_\mu\gamma^\mu,$$

$$\nabla = \gamma^\rho\partial_\rho,$$

(where the $\{\gamma^\mu : \mu = 0, 1, 2, 3\}$ are the Dirac gamma matrices with respect to η, that is, they satisfy the anticommutation rules $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma\mu = 2\eta^{\mu\nu}$). Those equation systems are to be understood together with boundary conditions that specify a particular field tensor $F_{\mu\nu}$ "out of" the source j^ν [25].

The symmetry group of the Maxwell field equations is the Lorentz–Poincaré group that acts on Minkowski space M and in an induced way on objects defined over M. However since we are interested in *complex* solutions for the Maxwell system, we must find a reasonable way of introducing complex objects in our formulation. One may formalize the Maxwellian system as a gauge field. We sketch the usual formulation: again we start from $M = \langle\mathsf{R}^4, \eta\rangle$, and construct the trivial circle bundle $P = M \times S^1$ over M, since Maxwell's field is the gauge field of the circle group S^1 (usually written in that respect as $U(1)$). We form the set $\mathcal{E}$ of bundles associated to P whose fibers are finite–dimensional vectorspaces. The set of physical fields in our theory is obtained out of some of the bundles in $\mathcal{E}$: the set of electromagnetic field tensors is a set of cross–sections of the bundle $F = \Lambda^2 \otimes s^1(M)$ of all s^1–valued 2–forms on M, where s^1 is the group's Lie algebra. To be more precise, the set of all electromagnetic fields is $\mathcal{F} \subset C^k(F)$, if we are dealing with C^k cross–sections (actually a submanifold in the usual C^k topology due to the closure condition $dF = 0$).

Finally we have two group actions on $\mathcal{F}$: the first one is the Lorentz–Poincaré action L which is part of the action of diffeomorphisms of M; then we have the (here trivial) action of the group $\mathcal{G}'$ of gauge transformations of P when acting on the field manifold $\mathcal{F}$. As it is well known, its action is *not* trivial in the non–Abelian case. Anyway it always has a nontrivial action on the space $\mathcal{A}$ of all gauge potentials for the fields in $\mathcal{F}$. Therefore we take as our symmetry group $\mathcal{G}$ the product $L \otimes \mathcal{G}'$ of the (allowed) symmetries of M and the symmetries of the principal bundle P.

We must also add the spaces $\mathcal{A}$ of potentials and of currents, $\mathcal{I}$, as structures derived from M and S^1. Both spaces have the same underlying topological structure; they differ in the way the group $\mathcal{G}'$ of gauge transformations acts upon them. We obtain $I = \Lambda^1 \otimes s^1(M)$ and $\mathcal{A} = \mathcal{I} = C^k(I)$. Notice that $\mathcal{I}/\mathcal{G}' = \mathcal{I}$ while $\mathcal{A}/\mathcal{G}' \neq \mathcal{A}$.

Therefore we can say that the 9–tuple

$$\langle M, S^1, P, \mathcal{F}, \mathcal{A}, \mathcal{G}, \mathcal{I}, B, \nabla\varphi = \iota \rangle$$

where M is Minkowski space, and B is a set of boundary conditions for our field equations $\nabla\varphi = \iota$, represents the species of mathematical structures of a Maxwellian electromagnetic field, where P, $\mathcal{F}$ and $\mathcal{G}$ are derived from M and S^1. The Dirac–like equation

$$\nabla\varphi = \iota$$

should be seen as an axiomatic restriction on our objects; the boundary conditions B are (i) a set of derived species of structures from M and S^1, since, as we are dealing with Cauchy conditions, we must specify a local or global spacelike hipersurface C in M to which (ii) we add sentences of the form $\forall x \in C\, f(x) = f_0(x)$, where f_0 is a set of (fixed) functions and the f are adequate restrictions of the field functions and equations to C.

Hamiltonian mechanics

Hamiltonian mechanics is here seen as the dynamics of the "Hamiltonian fluid" [1, 3, 42]. Our ground structure for mechanics starts out of basic sets which are a $2n$–dimensional real smooth manifold, and the real symplectic group $\mathrm{Sp}(2n, \mathsf{R})$. Phase spaces in Hamiltonian mechanics are symplectic manifolds: even–dimensional manifolds like M endowed with a symplectic form, that is, a nondegenerate closed 2–form Ω on M. The imposition of that form can be seen as the choice of a reduction of the linear bundle $L(M)$ to a fixed principal bundle $P(M, \mathrm{Sp}(2n, \mathsf{R}))$; however given one such reduction it doesn't automatically follow that the induced 2–form on M is a closed form.

All other objects are constructed in about the same way as in the preceding example. However we must show that we still have here a Dirac–like equation as the dynamical axiom for the species of structures of mechanics. Hamilton's equations are

$$i_X \Omega = -dh,$$

where i_X denotes the interior product with respect to the vectorfield X over M, and h is the Hamiltonian function. That equation is (locally, at least) equivalent to:

$$L_X \Omega = 0,$$

or

$$d(i_X \Omega) = 0,$$

where L_X is the Lie derivative with respect to X. The condition $d\varphi = 0$, with $\varphi = i_X \Omega$, is the degenerate Dirac–like equation for Hamiltonian mechanics. We don't get a full Dirac–like operator $\nabla \neq d$ because M, seen as a symplectic manifold, doesn't have a canonical metrical structure, so that we cannot define (through the Hodge dual) a canonical divergence δ dual to d. The group that acts on M with its symplectic form is the group of canonical transformations; it is a subgroup of the group of diffeomorphisms of M so that symplectic forms are mapped onto symplectic forms under a canonical transformation. We can take as "potential space" the space of all Hamiltonians on M (which is a rather simple function space), and as "field space" the space of all "Hamiltonian fields" of the form $i_X \Omega$.

Interpretations are immediate here: h is the system's Hamiltonian, which (given some simple conditions) can be seen as the system's total energy. Invariance of the symplectic form by the Lie derivative with respect to a Hamiltonian flow is equivalent both to Poincaré's integral invariant theorem and to Liouville's theorem — just as a flavor of the way our treatment handles well–known concepts and results in mechanics.

General relativity

General relativity is a theory of gravitation that interpretes this basic force as originated in the pseudo–Riemannian structure of spacetime. That is to say: in general relativity we start from a spacetime manifold (a 4–dimensional, real, adequately smooth manifold) which is endowed with an pseudo–Riemannian metric tensor. Gravitational effects originate in that tensor.

Given any 4–dimensional, noncompact, real, differentiable manifold M, we can endow it with an infinite set of different, nonequivalent pseudo–Riemannian metric tensors with a Lorentzian signature (that is, $- + ++$). That set is uncountable and has the power of the continuum. (By nonequivalent metric tensors we mean the following: form the set of all such metric tensors and factor it by the group of diffeomorphisms of M; we get a set that has the cardinality of the continuum. Each element of the quotient set is a different gravitational field for M.)

Therefore, neither the underlying structure of M as a topological manifold, nor its differentiable structure determines a particular pseudo–Riemannian metric tensor, that is, a specific gravitational field. From the strictly geometrical viewpoint, when we choose a particular metric tensor g of Lorentzian signature, we determine a g–dependent reduction of the general linear tensor bundle over M to one of its pseudo–orthogonal bundles. The relation

$g \mapsto g$–dependent reduction of the linear bundle to a pseudo–orthogonal bundle

is 1–1.

We now follow our recipe:

- We take as basic sets a 4–dimensional real differentiable manifold of class C^k, $1 \leq k \leq +\infty$, and the Lorentz pseudo–orthogonal group $O(3,1)$.
- We form the principal linear bundle $L(M)$ over M; that structure is solely derived from M, as it arises from the covariance properties of the tangent bundle over M. From $L(M)$ we fix a reduction of the bundle group $L(M) \to P(M, O(3,1))$, where $P(M, O(3,1))$ is the principal fiber bundle over M with the $O(3,1)$ group as its fiber.

  Those will be our derived sets. We therefore inductively define a Lorentzian metric tensor g on M, and get the couple $\langle M, g\rangle$, which is spacetime.

  (Notice that the general relativity spacetime arises quite naturally out of the interplay between the theory's "general covariance" aspects, which appear in $L(M)$, and — as we will see in the next section — its "gauge–theoretic features, which are clear in $P(M, O(3,1))$.)
- Field spaces are:
  - The first is the set (actually a manifold, with a natural differentiable structure) of all pseudo–Riemannian metric tensors, $\mathcal{M} \subset C^k(\odot^2 T_*(M))$, where $C^k(\odot^2 T_*(M))$ is the bundle of all C^k symmetric covariant 2–tensors over M.
  - Also out of M and out of adequate associated bundles we get $\mathcal{A}$, the bundle of all Christoffel connections over M, and $\mathcal{F}$, the bundle of all Riemann–Christoffel curvature tensors over M.
- We need the space of source fields, $\mathcal{I}$, that includes energy–momentum tensors, and arise out of adequate associated tensor bundles over M.
- $\mathcal{G}$ is the group of C^k–diffeomorphisms of M.
- If $\mathcal{K}$ is any of the field spaces above, then $\mathcal{K}/\mathcal{G}$ is the space of physically distinct fields.
- Finally the dynamics are given by Einstein's equations (there is also a Dirac–like formulation for those, first proposed by R. Penrose in 1960 as a neutrino–like equation; see [24]).

The quotient $\mathcal{K}/\mathcal{G}$ is the way we distinguish concrete, physically diverse, fields, as for covariant theories one has that any two fields related by an element of $\mathcal{G}$ "are" the "same" field.

Classical gauge fields

The mathematics of classical gauge fields can be found in [5, 65]. We follow here the preceding examples, and in particular the treatment of general relativity:

- The basic sets are a spacetime $\langle M, g\rangle$, and a finite dimensional, semi–simple, compact Lie group G.
- The derived set is a fixed principal bundle $P(M, G)$ over M with G as the fiber.
- The group of gauge transformations $\mathcal{G}$ is the subgroup of all diffeomorphisms of $P(M, G)$ that reduce to a diffeomorphism on M and to the group action on the fiber.
- If $\ell(G)$ is the Lie algebra of G, we get:
  - Connection–form space, or the space of potentials, noted $\mathcal{A}$, is the space of all C^k–cross sections of the bundle of $\ell(G)$–valued 1–forms on M.
  - Curvature space, or the space of fields $\mathcal{F}$, is the space of all C^k–cross sections of $\ell(G)$–valued 2–forms on M, such that $F \in \mathcal{F}$ is the field with potential $A \in \mathcal{A}$.
  - Source space $\mathcal{I}$ coincides with $\mathcal{A}$, but is acted upon in a different way by the group $\mathcal{G}$ of gauge transformations. (Currents in $\mathcal{I}$ are *tensorial* 1–forms, while gauge–potentials in $\mathcal{A}$ are transformed via an inhomogeneous transformation.)
- The space of physically different fields is $\mathcal{K}/\mathcal{G}$, where $\mathcal{K}$ is any of the above field spaces.
- Dynamics are given by the usual gauge–field equations, which are a nonlinear version of the electromagnetic field equations. There is also a Dirac–like equation for gauge fields [27].

To sum it up: with the help of the schema presented at the beginning of the section, we can say that the structure of a physical theory is an ordered pair $\langle \mathcal{F}, \mathcal{G}\rangle$, where $\mathcal{F}$ is an infinite–dimensional space of fields, and $\mathcal{G}$ is an infinite–dimensional group that acts upon field space. To get the Suppes predicate we must add the information about the dynamical equations $D(\phi) = 0, \phi \in \mathcal{F}$, for the fields ϕ.

Notice that general relativity can be seen as a kind of degenerate gauge field theory, more precisely a gauge theory of the $O(3, 1)$ group.

Quantum theory of the electron

The Dirac electron theory (and the general theory of particles with any spin) can be easily formalized according to the preceding schemata. One uses as geometrical background the setting for special relativity; dynamics is given either by

Dirac's equation or Weyl's equation, for the case of zero–mass particles. Higher spin fields are dealt with the help either of the Bargmann–Wigner equations or their algebraic counterpart [25]. The Schrödinger equation is obtained from the Dirac set out of a — loosely speaking — 'standard' limiting procedure, which can be formally represented by the addition of new axioms to the corresponding Suppes predicate.

General field theory

Sometimes one may wish to discuss field theory in a very general, motion–equation independent, way. We then use as geometrical background the construction of Minkowski space and take as dynamical axioms the field–theoretic Euler–Lagrange equations, or, as we've said, we can take the variational principle as a formal algorithm to derive the dynamics of the system.

Other domains of science

We can extend the preceding techniques to several scientific domains. For example, the bulk of economics, as presented in Samuelson's *Foundations of Economic Analysis* [56], or some specific results, such as the Nash equilibrium theorem [18], easily fit within our construction — we can find in a straightforward way a Suppes predicate for results in mathematical economics [17]. The same goes with mathematical biology [44].

Summing it up

We have proceeded from start with a specific goal in mind: we wished to follow Hilbert's programme in his 6th Problem, that is, we proposed an axiomatiation of physics that allows us to explore many interesting mathematical consequences of those theories.

We now wish to obtain examples of Gödel sentences — undecidable sentences — within the axiomatic versions of those theories, and in a more general cadre, we wish to see the effects and consequences of metamathematical results and techniques when applied to those theories, or to their axiomatic versions.

7 Generalized incompleteness

Preliminary

For concepts from logic see [51]. We use: $\neg$, "not," $\vee$, "or," $\wedge$, "and," $\rightarrow$, "if... then...," $\leftrightarrow$, "if and only if," $\exists x$, "there is a x," $\forall x$, "for every x." $P(x)$ is a formula with x free; it roughly means "x has property P." Finally $T \vdash \xi$ means T proves ξ, or ξ is a theorem of T. ω is the set of natural numbers, $\omega = \{0, 1, 2, \ldots\}$.

We deal here mainly with algorithmic functions. These are given by their programs coded in Gödel numbers e [54]. We will sometimes use Turing machines (noted by sans–serif letters with the Gödel number as index M_e) or partial recursive functions, noted $\{e\}$.

We start from arithmetic, specifically from Peano Arithmetic, noted PA [51]. Its language includes variables $x, y, \ldots$, two constants, $\mathbf{0}$ and $\mathbf{1}$, the equality sign $=$, and two operation signs, $+, \times$. We will also require Russell's ι symbol [41]. $\iota_x\, P(x)$ is, roughly, the x such that $P(x)$.

The *standard interpretation* for PA is: the variables $x, y, \ldots$ range over the natural numbers, and $\mathbf{0}$ and $\mathbf{1}$ are seen as, respectively, zero and one. PA is strong enough to formally include most of Turing machine theory [18, 19]. Recall that a Turing machine is given by its Gödel number, which recursively codes the machine's program. Rigorously, for PA, we have:

Definition 7.1 *A Turing machine of Gödel number e operating on x with output y, $\{e\}(x) = y$ is* **representable** *in* PA *if there is a formula $F_e(x, y)$ in the language of our arithmetic theory so that:*

1. $\mathrm{PA} \vdash\ F_e(x, y) \wedge F_e(x, z) \to y = z$, *and*
2. *For natural numbers a, b, $\{e\}(a) = b$ if and only if* $\mathrm{PA} \vdash\ F_e(a, b)$. $\square$

Then we have the representation theorem for partial recursive functions in PA:

Proposition 7.2 *Every Turing machine is representable in Peano Arithmetic. Moreover there is an effective procedure that allows us to obtain F_e from the Gödel number e.* $\square$

Remark 7.3 We mainly consider here theories that are *arithmetically sound*, that is, which have a model with standard arithmetic for its arithmetical segment. $\square$

A first example of generalized incompleteness

The example we now give shows that incompleteness is a pervasive phenomenon, from an arithmetic theory like PA and upwards, that is, it affects all theories that contain enough arithmetic, have a model where arithmetic is standard, and have a recursively enumerable set of theorems.

Suppose that our theory S has Russell's description symbol ι [41]. Let P be a predicate symbol so that for closed terms $\xi, zeta$ so that $S \vdash \xi \neq \zeta$, $S \vdash\ P(\xi)$ and $S \vdash\ \neg P(\zeta)$ (we call such P, nontrivial predicates). Then, for the term:

$$\eta = \iota_x[(x = \xi \wedge \alpha) \vee (x = \zeta \wedge \neg\alpha)],$$

where α is an undecidable sentence in S:

Proposition 7.4 $S \nvdash\ P(\eta)$ *and* $S \nvdash\ \neg P(\eta)$. $\square$

This shows that incompleteness is found everywhere within theories like S.

Remark 7.5 From now on we will consider theories S, T, like the one characterized above. $\Box$

Our main tool here will be an explicit expression for the Halting Function, that is, the function that settles the halting problem [54]. We have shown elsewhere that it can be constructed within the language of classical analysis. We have originally used the Richardson transforms (see details and references in [18]) in order to obtain an explicit expression for the Halting Function, but they are not essential in our construction. We start from a strengthening of Proposition 7.2:

Proposition 7.6 *If $\{e\}(a) = b$, for natural numbers a, b, then we can algorithmically construct a polynomial p_e over the natural numbers so that $\{e\}(a) = b \leftrightarrow \exists x_1, x_2, \ldots, x_k \in \omega\, p_e(a, b, x_1, x_2, \ldots, x_k) = 0$.* $\Box$

Proposition 7.7 *$a \in R_e$, where R_e is a recursively enumerable set, if and only if there are e and p so that $\exists x_1, x_2, \ldots, x_k \in \omega\, (p_e(a, x_1, x_2, \ldots, x_k) = 0)$.* $\Box$

Our results derive from the preceding propositions.

The Halting Function

One of the main results in Alan Turing's great 1937 paper, "On computable numbers, with an application to the Entscheidungsproblem" [64], is a proof of the algorithmic unsolvability of a version of the halting problem: given an arbitrary Turing machine of Gödel number e, for input x, there is no algorithm that decides whether $\{e\}(x)$ stops and outputs something, or enters an infinite loop.

Remark 7.8 Let $\mathsf{M}_m(a)\downarrow$ mean: "Turing machine of Gödel number m stops over input a and gives some output." Similarly $\mathsf{M}_m(a)\uparrow$ means, "Turing machine of Gödel number m enters an infinite loop over input a." Then we can define the halting function θ:

- $\theta(m, a) = 1$ if and only if $\mathsf{M}_m(a)\downarrow$.
- $\theta(m, a) = 0$ if and only if $\mathsf{M}_m(a)\uparrow$.

$\theta(m, a)$ is the halting function for M_m over input a. $\Box$

θ isn't algorithmic, of course [54, 64], that is, there is no Turing machine that computes it.

Then, if σ is the sign function, $\sigma(\pm x) = \pm 1$ and $\sigma(0) = 0$:

Expressions for the Halting Function

Proposition 7.9 (The Halting Function.) *The halting function $\theta(n, q)$ is explicitly given by:*

$$\theta(n, q) = \sigma(G_{n,q}),$$

$$G_{n,q} = \int_{-\infty}^{+\infty} C_{n,q}(x) e^{-x^2} dx,$$

$$C_{m,q}(x) = |F_{m,q}(x) - 1| - (F_{m,q}(x) - 1).$$

$$F_{n,q}(x) = \kappa_P p_{n,q}. \ \square$$

Here $p_{n,q}$ is the two–parameter universal Diophantine polynomial and κ_P an adequate Richardson transform.

The succession of definite integrals

$$K(m) = \int_{-\infty}^{+\infty} \frac{C(m, x) e^{-x^2}}{1 + C(m, x)} dx,$$

also gives us the Halting Function:

$$\theta(m, x) = \theta(\langle m, x \rangle) = \theta(m) = \sigma(\frac{K_m}{1 + K_m}). \ \square$$

Remark 7.10 There is an expression for the Halting Function even within a simple extension of PA. Let $p(n, \mathbf{x})$ be a 1–parameter universal polynomial; $\mathbf{x}$ abbreviates $x_1, \ldots, x_p$. Then either $p^2(n, \mathbf{x}) \geq 1$, for all $\mathbf{x} \in \omega^p$, or there are $\mathbf{x}$ in ω^p such that $p^2(n, \mathbf{x}) = 0$ sometimes. As $\sigma(x)$ when restricted to ω is primitive recursive, we may define a function $\psi(n, \mathbf{x}) = 1 - \sigma p^2(n, \mathbf{x})$ such that:

- Either for all $\mathbf{x} \in \omega^p$, $\psi(n, \mathbf{x}) = 0$;
- Or there are $\mathbf{x} \in \omega^p$ so that $\psi(n, \mathbf{x}) = 1$ sometimes.

Thus the Halting Function can be represented as:

$$\theta(n) = \sigma[\sum_{\tau^q(\mathbf{x})} \frac{\psi(n, \mathbf{x})}{\tau^q(\mathbf{x})!}],$$

where $\tau^q(\mathbf{x})$ denotes the positive integer given out of $\mathbf{x}$ by the pairing function τ: if τ^q maps q–tuples of positive integers onto single positive integers, $\tau^{q+1} = \tau(x, \tau^q(\mathbf{x}))$. $\square$

Undecidability and incompleteness

Our main undecidability (and the related incompleteness) results stem from the following:

Lemma 7.11 *There is a Diophantine set D so that*

$$m \in D \leftrightarrow \exists x_1, \ldots, x_n \in \omega\, p(m, x_1, \ldots, x_n) = 0,$$

p a Diophantine polynomial, and D is recursively enumerable but not recursive. $\square$

Corollary 7.12 *For an arbitrary $m \in \omega$ there is no general decision procedure to check whether $p(m, x_1, \ldots) = 0$ has a solution in the positive integers.* $\square$

Main undecidability and incompleteness result

Therefore, given such a p, and $F = \kappa_P(p)$, where κ_P is an adequate Richardson transform:

Corollary 7.13 *For an arbitrary $m \in \omega$ there is no general decision procedure to check whether, for F and G adequate real–defined and real–valued functions:*

1. *There are real numbers $x_1, \ldots, x_n$ such that $F(m, x_1, \ldots, x_n) = 0$;*
2. *There is a real number x so that $G(m, x) < 1$;*
3. *Whether we have $\forall x \in \mathsf{R}\ \theta(m, x) = 0$ or $\forall x \in \mathsf{R}\ \theta(m, x) = 1$ over the reals.*
4. *Whether for an arbitrary $f(m, x)$ we have $f(m, x) \equiv \theta(m, x)$.*

Proof: From the preceding results. The last undecidability statement follows from the third one. $\square$

We conclude with a first, quite all–encompassing, result. Let $\mathcal{B}$ be a sufficiently large algebra of functions and let $P(x)$ be a nontrivial predicate. If ξ is any word in that language, we write $\|\xi\|$ for its complexity, as measured by the number of letters from ZFC's alphabet in ξ. Also we define the *complexity of a proof* $C_{\mathrm{ZFC}}(\xi)$ of ξ in the language of ZFC to be the minimum length that a deduction of ξ from the ZFC axioms can have, as measured by the total number of letters in the expressions that belong to the proof.

Remark 7.14 Recall that theory $T \supset$ PA is *arithmetically sound* if T has a model where its arithmetic is standard. $\square$

Proposition 7.15 *If* ZFC *is arithmetically sound, then:*

1. *There is an $h \in \mathcal{B}$ so that neither* ZFC $\not\vdash \neg P(h)$ *nor* ZFC $\not\vdash P(h)$*, but* $\mathbf{N} \models P(h)$*, where* $\mathbf{N}$ *makes* ZFC *arithmetically sound.*

2. *There is a denumerable set of functions $h_m(x) \in \mathcal{B}$, $m \in \omega$, such that there is no general decision procedure to ascertain, for an arbitrary m, whether $P(h_m)$ or $\neg P(h_m)$ is provable in* ZFC.
3. *Given the set $K = \{m : \mathrm{ZFC} \vdash \phi(\widehat{m})\}$, and given an arbitrary total recursive function $g : \omega \to \omega$, there is an infinite number of values for m so that $C_{\mathrm{ZFC}}(P(\widehat{m})) > g(\|P(\widehat{m})\|)$.*

Proof: Let θ be as above. Let f_0, g_0 satisfy our conditions on P, that is, $\mathrm{ZFC} \vdash P(f_0)$ and $\mathrm{ZFC} \vdash \neg P(g_0)$. Then define:

$$h(m, x) = \theta(m, x) f_0 + (1 - \theta(m, x)) g_0.$$

This settles (2). Now let us specify θ so that the corresponding Diophantine equation $p = 0$ is never solvable in the standard model for arithmetic, while that fact cannot be proved in ZFC. We then form, for such an indicator function,

$$h = \theta f_0 + (1 - \theta) g_0.$$

This settles (1). Finally, for (3), we notice that as K is recursively enumerable but not recursive, it satisfies the conditions in the Gödel–Ehrenfeucht–Mycielski theorem about the length of proofs. $\square$

8 Higher degrees

Our main result in this section is:

Proposition 8.1 *If T is arithmetically sound then we can explicitly and algorithmically construct in the language $\mathcal{L}_T$ of T an expression for the characteristic function of a subset of ω of degree $\mathbf{0}''$.*

Remark 8.2 That expression depends on recursive functions defined on ω and on elementary real–defined and real–valued functions plus the absolute value function, a quotient and an integration, or perhaps an infinite sum, as in the case of the β and θ functions associated to the halting problem. $\square$

Proof: We could simply use Theorem 9-II in [54] (p. 132). However for the sake of clarity we give a detailed albeit informal proof. Actually the degree of the set described by the characteristic function whose expression we are going to obtain will depend on the fixed oracle set A; so, our construction is a more general one.

Let us now review a few concepts. Let $A \subset \omega$ be a fixed infinite subset of the integers:

Definition 8.3 *The* **jump of** *A is noted A'; $A' = \{x : \phi_x^A(x) \downarrow\}$, where ϕ_x^A is the A–partial recursive algorithm of index x.* $\square$

In order to make things self–contained, we review here some ideas about A–partial recursive functions.

From Turing machines to oracle Turing machines

1. An oracle Turing machine ϕ_x^A with oracle A can be visualized as a two–tape machine where tape 1 is the usual computation tape, while tape 2 contains a listing of A. When the machine enters the oracle state s_0, it searches tape 2 for an answer to a question of the form "does $w \in A$?" Only finitely many such questions are asked during a converging computation; we can separate the positive and negative answers into two disjoint finite sets $D_u(A)$ and $D_v^*(A)$ with (respectively) the positive and negative answers for those questions; notice that $D_u \subset A$, while $D_v^* \subset \omega - A$. We can view those sets as ordered k– and k^*–ples; u and v are recursive codings for them [54]. The $D_u(A)$ and $D_v^*(A)$ sets can be coded as follows: only finitely many elements of A are queried during an actual converging computation with input y; if k' is the highest integer queried during one such computation, and if $d_A \subset c_A$ is an initial segment of the characteristic function c_A, we take as a standby for D and D^* the initial segment d_A where the length $l(d_A) = k' + 1$.

   We can effectively list all oracle machines with respect to a fixed A, so that, given a particular machine we can compute its index (or Gödel number) x, and given x we can recover the corresponding machine.

2. Given an A–partial recursive function ϕ_x^A, we form the oracle Turing machine that computes it. We then do the computation $\phi_x^A(y) = z$ that outputs z. The initial segment $d_{y,A}$ is obtained during the computation.

3. The oracle machine is equivalent to an ordinary two–tape Turing machine that takes as input $\langle y, d_{y,A} \rangle$; y is written on tape 1 while $d_{y,A}$ is written on tape 2. When this new machine enters state s_0 it proceeds as the oracle machine. (For an ordinary computation, no converging computation enters s_0, and $d_{y,A}$ is empty.)

4. The two–tape Turing machine can be made equivalent to a one–tape machine, where some adequate coding places on the single tape all the information about $\langle y, d_{y,A} \rangle$. When this third machine enters s_0 it scans $d_{y,A}$.

5. We can finally use the standard map τ that codes n–ples 1–1 onto ω and add to the preceding machine a Turing machine that decodes the single natural number $\tau(\langle y, d_{y,A} \rangle)$ into its components before proceeding to the computation.

Let w be the index for that last machine; we note the machine ϕ_w.

If x is the index for ϕ_x^A, we write $w = \rho(x)$, where ρ is the effective 1–1 procedure above described that maps indices for oracle machines into indices for Turing machines. Therefore,

$$\phi_x^A(y) = \phi_{\rho(x)}(\langle y, d_{y,A} \rangle).$$

Now let us note the universal polynomial $p(n, q, x_1, \ldots, x_n)$. We can define the jump of A as follows:

$$A' = \{\rho(z) : \exists x_1, \ldots, x_n \in \omega\, p(\rho(z), \langle z, d_{z,A}\rangle, x_1, \ldots, x_n) = 0\}.$$

With the help of the Richardson map described above, we can now form a function modelled after the θ function that settles the Halting Problem; it is the desired characteristic function:

$$c_{\emptyset'}(x) = \theta(\rho(x), \langle x, d_{x,\emptyset'}\rangle).$$

(Actually we have proved more; we have obtained

$$c_{A'}(x) = \theta(\rho(x), \langle x, d_{x,A}\rangle),$$

with reference to an arbitrary $A \subset \omega$.)

Finally, we write $\theta^{(2)}(x) = c_{\emptyset''}(x)$. $\Box$

We recall [54]:

Definition 8.4 *The complete Turing degrees $\mathbf{0}, \mathbf{0}', \mathbf{0}'', \ldots, \mathbf{0}^{(p)}, \ldots$, $p < \omega$, are Turing equivalence classes generated by the sets $\emptyset, \emptyset', \emptyset'', \ldots, \emptyset^{(p)}, \ldots$.* $\Box$

Now let $\mathbf{0}^{(n)}$ be the n–th complete Turing degree in the arithmetical hierarchy. Let $\tau(n, q) = m$ be the pairing function in recursive function theory [54]. For $\theta(m) = \theta(\tau(n, q))$, we have:

Corollary 8.5 (Complete Degrees.) *If T is arithmetically sound, for all $p \in \omega$ the expressions $\theta^p(m)$ explicitly constructed below represent characteristic functions in the complete degrees $\mathbf{0}^{(p)}$.*

Proof: From Proposition 8.1,

$$\begin{cases} \theta^{(0)} = c_{\emptyset}(m) = 0, \\ \theta^{(1)}(m) = c_{\emptyset'}(m) = \theta(m), \\ \theta^{(n)}(m) = c_{\emptyset^{(n)}}(m), \end{cases}$$

for c_A as in Proposition 8.1. $\Box$

Incompleteness theorems

We now state and prove several incompleteness results about axiomatized versions of arithmetic with a classical first–order language, a recursive vocabulary and a recursively enumerable set of axioms; say, Peano Arithmetic (PA). These results are of course also valid for extensions of it T with the same language and recursively enumerable set of axioms.

We suppose, as already stated, that PA $\subset T$ means that there is an interpretation of PA in T.

The next results will be needed when we consider our main examples. We recall that "$\dot{-}$" — the truncated sum — is a primitive recursive operation on ω:

- For $a > b$, $a \stackrel{\bullet}{-} b = a - b$.
- For $a < b$, $a \stackrel{\bullet}{-} b = 0$.

In the next result, Z is the set of integers. The starting point is the following consequence of a well–known result which we now quote: let $\mathbf{N}$ be a model, $\mathbf{N} \models T$, and $\mathbf{N}$ makes T arithmetically sound. Then:

Proposition 8.6 *If T is arithmetically sound, then we can algorithmically construct a polynomial expression $q(x_1, \ldots, x_n)$ over Z such that $\mathbf{M} \models \forall x_1, \ldots, x_n \in \omega\, q(x_1, \ldots, x_n) > 0\}$, but*

$$T \not\vdash \forall x_1, \ldots, x_n \in \omega\, q(x_1, \ldots, x_n) > 0$$

and

$$T \not\vdash \exists x_1, \ldots, x_n \in \omega\, q(x_1, \ldots, x_n) = 0.$$

Proof: Let $\xi \in \mathcal{L}_T$ be an undecidable sentence obtained for T with the help of Gödel's diagonalization; let n_ξ be its Gödel number and let m_T be the Gödel coding of proof techniques in T (of the Turing machine that enumerates all the theorems of T). For an universal polynomial $p(m, q, x_1, \ldots, x_n)$ we have:

$$q(x_1, \ldots, x_n) = (p(m_T, n_\xi, x_1, \ldots, x_n))^2. \ \Box$$

Corollary 8.7 *If* PA *is consistent then we can find within it a polynomial p as in Proposition 8.6.* $\Box$

We can also state and prove a weaker version of Proposition 8.6:

Proposition 8.8 *If T is arithmetically sound, there is a polynomial expression over Z $p(x_1, \ldots, x_n)$ such that $\mathbf{N} \models \forall x_1, \ldots, x_n \in \omega\, p(x_1, \ldots, x_n) > 0$, while*

$$T \not\vdash \forall x_1, \ldots, x_n \in \omega\, p(x_1, \ldots, x_n) > 0$$

and

$$T \not\vdash \exists x_1, \ldots, x_n \in \omega\, p(x_1, \ldots, x_n) = 0.$$

Proof: See [21]. If $p(m, x_1, \ldots, x_n)$, $m = \tau\langle q, r\rangle$, is an universal polynomial with τ being Cantor's pairing function [54], then $\{m : \exists x_1 \ldots \in \omega\, p(m, x_1, \ldots) = 0\}$ is recursively enumerable but not recursive. Therefore there must be an m_0 such that $\forall x_1 \ldots \in \omega\, (p(m_0, x_1, \ldots))^2 > 0$. (This is actually a version of Post's original argument for the proof of Gödel's theorem [53].) $\Box$

Proposition 8.9 *If* PA *is consistent and* $\mathbf{N} \models$ PA *is standard, and if P is nontrivial then there is a term–expression $\zeta \in \mathcal{L}_{\mathrm{PA}}$ such that $\mathbf{N} \models P(\zeta)$ while* PA $\not\vdash P(\zeta)$ *and* PA $\not\vdash \neg P(\zeta)$.

Proof: Put $\zeta = \xi + r(x_1, \ldots, x_n)\nu$, for $r = 1 \stackrel{\bullet}{-} (q + 1)$, q as in Proposition 8.6 (or as p in Proposition 8.8). $\Box$

Remark 8.10 Therefore every nontrivial arithmetical P in theories from formalized arithmetic upwards turns out to be undecidable. We can generalize that result to encompass other theories T that include arithmetic; see below. $\Box$

9 The θ function and the arithmetical hierarchy

We now give alternative proofs for well–known results about the arithmetical hierarchy that will lead to other incompleteness results. Recall;

Definition 9.1 *The sentences* $\xi, \zeta \in \mathcal{L}_T$ *are* **demonstrably equivalent** *if and only if* $T \vdash \xi \leftrightarrow \zeta$. $\square$

Definition 9.2 *The sentence* $\xi \in \mathcal{L}_T$ *is* **arithmetically expressible** *if and only if there is an arithmetic sentence* ζ *such that* $T \vdash \xi \leftrightarrow \zeta$. $\square$

Then, for $\mathbf{N} \models T$, a model that makes it arithmetically sound,

Proposition 9.3 *If* T *is arithmetically sound, then for every* $m \in \omega$ *there is a sentence* $\xi \in T$ *such that* $\mathbf{M} \models \xi$ *while for no* $k \leq n$ *there is a* Σ_k *sentence in* PA *demonstrably equivalent to* ξ.

Proof: The usual proof for PA is given in Rogers [54], p. 321. However we give here a slightly modified argument that imitates Proposition 8.8. First notice that

$$\emptyset^{(m+1)} = \{x : \phi_x^{\emptyset^{(m)}}(x)\}$$

is recursively enumerable but not recursive in $\emptyset^{(m)}$. Therefore, $\overline{\emptyset^{(m+1)}}$ isn't recursively enumerable in $\emptyset^{(m)}$, but contains a proper $\emptyset^{(m)}$–recursively enumerable set. Let's take a closer look at those sets.

We first need a lemma: form the theory $T^{(m+1)}$ whose axioms are those for T plus a denumerably infinite set of statements of the form "$n_0 \in \emptyset^{(n)}$," "$n_1 \in \emptyset^{(m)}$,"..., that describe $\emptyset^{(m)}$. Of course this theory doesn't have a recursively enumerable set of theorems. Then,

Lemma 9.4 *If* $T^{(n+1)}$ *is arithmetically sound, then* $\phi_x^{\emptyset^{(m)}}(x) \downarrow$ *if and only if*

$$T^{(m+1)} \vdash \exists x_1, \ldots, x_n \in \omega\, p(\rho(z), \langle z, d_{y,\emptyset^{(m)}} \rangle, x_1, \ldots, x_n) = 0.$$

Proof: Similar to the proof in the non–relativized case; see [48], p. 126 ff. $\square$

Therefore we have that the oracle machines $\phi_x^{\emptyset^{(m)}}(x) \downarrow$ if and only if

$$T^{(m+1)} \vdash \exists x_1, \ldots, x_n \in \omega\, p(\rho(z), \langle z, d_{y,\emptyset^{(m)}} \rangle, x_1, \ldots, x_n) = 0.$$

However, since $\overline{\emptyset^{(m+1)}}$ isn't recursively enumerable in $\emptyset^{(m)}$ then there will be an index $m_0(\emptyset^{(m)}) = \langle \rho(z), \langle z, d_{y,\emptyset^{(m)}} \rangle \rangle$ such that

$$\mathbf{N} \models \forall x_1, \ldots, x_n [p(m_0, x_1, \ldots, x_n)]^2 > 0,$$

while it cannot be proved neither disproved within $T^{(m+1)}$. It is therefore demonstrably equivalent to a Π_{m+1} assertion. $\square$

Now let $q(m_0(\emptyset^{(m)}), x_1, \ldots) = p(m_0(\emptyset^{(m)}), x_1, \ldots))^2$ be as in Proposition 9.3. Then:

Corollary 9.5 *If T is arithmetically sound, then for:*

$$\beta^{(m+1)} = \sigma(G(m_0(\emptyset^{(n)})),$$

$$G(m_0(\emptyset^{(n)})) = \int_{-\infty}^{+\infty} \frac{C(m_0(\emptyset^{(n)}), x)e^{-x^2}}{1 + C(m_0(\emptyset^{(n)}), x)} dx,$$

$$C(m_0(\emptyset^{(n)}), x) = \lambda q(m_0(\emptyset^{(n)}), x_1, \ldots, x_r),$$

$\mathbf{N} \models \beta^{(m+1)} = 0$ *but for all* $n \leq m+1$, $\neg\{T^{(n)} \vdash \beta^{(m+1)} = 0\}$ *and* $\neg\{T^{(n)} \vdash \neg(\beta^{(m+1)} = 0)\}$. $\square$

We have used here a variant of the construction of θ and β which first appeared in [16]. Then,

Corollary 9.6 *If T is arithmetically sound and if $\mathcal{L}_T$ contains expressions for the $\theta^{(m)}$ functions as given in Proposition 8.5, then for any nontrivial arithmetical predicate P there is a $\zeta \in \mathcal{L}_T$ such that the assertion $P(\zeta)$ is T–demonstrably equivalent to and T–arithmetically expressible as a Π_{m+1} assertion, but not equivalent to and expressible as any assertion with a lower rank in the arithmetic hierarchy.*

Proof: As in the proof of Proposition 8.9, we write:

$$\zeta = \xi + [1 \overset{\bullet}{-} (p(m_0(\emptyset^m), x_1, \ldots, x_n) + 1)]\nu,$$

where $p(\ldots)$ is as in Proposition 9.3. $\square$

Remark 9.7 Rogers discusses the rank within the arithmetical hierarchy of well–known open mathematical problems ([54], p. 322), such as Fermat's Conjecture—which in its usual formulation is demonstrably equivalent to a Π_1^0 problem,[6] or unsettled questions such as Riemann's Hypothesis, which is also stated as a Π_1^0 problem. On the other hand, the $P < NP$ hypothesis in computer science is formulated as a Π_2^0 sentence that can be made equivalent to an intuitive Π_1^0 sentence, while its negation, the $P = NP$ conjecture, can be formalized as a Π_1^0 sentence within Peano Arithmetic [19].

Rogers conjectures that our mathematical imagination cannot handle more that four or five alternations of quantifiers. However the preceding result shows that *any* arithmetical nontrivial property within T can give rise to intractable problems of arbitrarily high rank.

We stress the need for the extension $T \supset \mathrm{PA}$, since otherwise we wouldn't be able to find an expression for the characteristic function of a set with a high rank in the arithmetical hierarchy within our formal language. $\square$

An extension of the preceding result is:

[6] The question of whether Wiles' proof can be fully formalized within ZFC is still open, and so, while we know that Fermat's Theorem is true of the standard integers, we don't know which minimum axiomatic resources are required for its proof.

Corollary 9.8 *If T is arithmetically sound then, for any nontrivial P there is a $\zeta \in \mathcal{L}_T$ such that $P(\zeta)$ is arithmetically expressible,* $\mathbf{N} \models P(\zeta)$ *but only demonstrably equivalent to a Π^0_{n+1} assertion and not to a lower one in the hierarchy.*

Proof: Put

$$\zeta = \xi + \beta^{(m+1)}\nu,$$

where one uses Corollary 9.5. □

Beyond arithmetic

We recall:

Definition 9.9

$$\emptyset^{(\omega)} = \{\langle x, y\rangle : x \in \emptyset^{(y)}\},$$

for $x, y \in \omega$. □

Then:

Definition 9.10

$$\theta^{(\omega)}(m) = c_{\emptyset^{(\omega)}}(m),$$

where $c_{\emptyset^{(\omega)}}(m)$ is obtained as in Proposition 8.1. □

Still,

Definition 9.11

$$\emptyset^{(\omega+1)} = (\emptyset^{(\omega)})'.$$

□

Corollary 9.12 $\mathbf{0}^{(\omega+1)}$ *is the degree of $\emptyset^{(\omega+1)}$.* □

Corollary 9.13 $\theta^{(\omega+1)}(m)$ *is the characteristic function of a nonarithmetic subset of ω of degree $\mathbf{0}^{(\omega+1)}$.* □

Corollary 9.14 *If T is arithmetically sound, then for:*

$$\beta^{(\omega+1)} = \sigma(G(m_0(\emptyset^{(\omega)})),$$

$$G(m_0(\emptyset^{(\omega)})) = \int_{-\infty}^{+\infty} \frac{C(m_0(\emptyset^{(\omega)}), x)e^{-x^2}}{1 + C(m_0(\emptyset^{(\omega)}), x)} dx,$$

$$C(m_0(\emptyset^{(\omega)}), x) = \lambda q(m_0(\emptyset^{(\omega)}), x_1, \ldots, x_r),$$

$\mathbf{N} \models \beta^{(\omega+1)} = 0$ *but* $T \not\vdash \beta^{(\omega+1)} = 0$ *and* $T \not\vdash \neg(\beta^{(\omega+1)} = 0)$. □

Proposition 9.15 *If T is arithmetically sound then given any nontrivial predicate P:*

1. *There is a family of terms $\zeta_m \in \mathcal{L}_T$ such that there is no general algorithm to check, for every $m \in \omega$, whether or not $P(\zeta_m)$.*

2. *There is a term $\zeta \in \mathcal{L}_T$ such that $\mathbf{M} \models P(\zeta)$ while $T \not\vdash P(\zeta)$ and $T \not\vdash \neg P(\zeta)$.*

3. *Neither the ζ_m nor ζ are arithmetically expressible.*

Proof: We take:

1. $\zeta_m = x\theta^{(\omega+1)}(m) + (1 - \theta^{(\omega+1)}(m))y$.

2. $\zeta = x + y\beta^{(\omega+1)}$.

3. Neither $\theta^{(\omega+1)}(m)$ nor $\beta^{(\omega+1)}$ are arithmetically expressible. $\square$

Remark 9.16 We have thus produced out of every nontrivial predicate in T intractable problems that cannot be reduced to arithmetic problems. Actually there are infinitely many such problems for every ordinal α, as we ascend the set of infinite ordinals in T. Also, the general nonarithmetic undecidable statement $P(\zeta)$ has been obtained without the help of any kind of forcing construction. $\square$

For the way one proceeds with those extensions we refer the reader to references on the hyperarithmetical hierarchy [4, 22, 54].

10 First applications: mechanics and chaos theory

The search for some algorithmic procedure or at least for some reasonable criterion that would distinguish chaotic systems from non–chaotic ones was the original motivation that led to the results presented in this work. After striving for a short time to get one such criterion, the authors wondered around 1985 whether the question wasn't in fact algorithmically undecidable, perhaps due to the complexity of the behavior of systems that exhibit chaotic behavior, even if such systems are described by rather simple systems of equations.

We later saw that the metamathematical phenomena we were looking for had a different origin.

The original intuition was, roughly, that if a simple description encapsulates an involved behavior, then we would perhaps lack the tools to check for specific properties of the system like chaos, since the system's complexity might — perhaps — exceed by far the available tools for its analysis, whatever that might mean. However as it turned out the undecidability and incompleteness of chaotic dynamical systems turned out to stem from a totally different aspect of the question: as we have shown in detail, it is essentially a *linguistic* phenomenon, that is, it depends on the tools that we have within formal systems to handle expressions for the objects in those systems.

(When we say "the undecidability and incompleteness of dynamical systems," we are making an abuse of language: more precisely, we mean the undecidability or incompleteness properties of the formal theory of those systems developed with the help of the language of classical analysis.)

Undecidability in classical mechanics

Let's go straight to the point and ask three questions in order to give a good example:

1. Given a Hamiltonian h, do we have an algorithm that tells us whether the associated Hamiltonian dynamical system X_h can be integrated by quadratures?

2. Given a Hamiltonian h such that X_h can be integrated by quadratures, can we algorithmically find a canonical transformation that will do the trick?

3. Can we algorithmically check whether an arbitrary set of functions is a set of first integrals for a Hamiltonian system?

The answer to those questions is, no. Proof follows from the techniques developed in the previous sections [16].

Chaos theory is undecidable and incomplete

We finally reach the question that originally motivated our quest. Let X be a smooth vectorfield on a differentiable manifold M. Can we algorithmically check whether X has some kind of chaotic behavior — in any of the usual meanings for that word, that is, given an arbitrary vectorfield X, can we algorithmically decide whether X is chaotic?

This problem was explicitly discussed by M. Hirsch [35] when he makes some remarks about the Lorenz system of equations [43]:

> (. . .) By computer simulation Lorenz found that trajectories seem to wander back and forth between two particular stationary states, in a random, unpredictable way. Trajectories which start out very close together eventually diverge, with no relationship between long run behaviors.
>
> But this type of chaotic behavior has not been proved. As far as I am aware, practically nothing has been proved about this particular system (. . .)
>
> *A major challenge to mathematicians is to determine which dynamical systems are chaotic and which are not. Ideally one should be able to tell from the form of the differential equations.*

Emphasis is ours. Therefore we can ask:

> *Is there a general algorithmic criterion so that, if we specify some formal definition for chaos in a dynamical system, we can determine whether an arbitrary expression for a dynamical system satisfies that definition?*

Again the answer is, no.

Let M be a differentiable manifold, and let $U \subset M$ be a starshaped open domain. As always, we suppose that ZFC is arithmetically sound, and that our results happen within ZFC.

Remark 10.1 For the next proposition we need two specific results:

- The construction of a Hamiltonian system with a Smale horseshoe [36].
- The fact that some geodesic flows are Bernouilli flows.

  That is, those geodesic flows have a decidedly random behavior.

We assert:

Proposition 10.2 *There is no general algorithmic procedure to check:*

1. *Whether an arbitrary vectorfield X over U is ergodic.*
2. *If $\dim M \geq 4$, whether an arbitrary vectorfield X over U has a Smale horseshoe.*
3. *If M is compact, real, two–dimensional, of class C^3 and has a constant negative curvature, whether an arbitrary X is a Bernouillian flow.*

Proof: As above. For the first two assertions, let K_0 be a constant vectorfield on U, and let Y be ergodic, or have a Smale horseshoe (for an explicit example, see [36]). Now let θ be the "yes–no" function, or Halting Function. Then

$$Z_m = \Theta_m K_0 + (1 - \Theta_m)Y,$$

where

$$\Theta_m(x_1, \ldots, x_n) = \theta_m(x_1),$$

has the same smoothness properties of Y, as both K_0 and θ are constant functions. It is an undecidable, countable family of vectorfields, as in the preceding results.

For the third assertion, we know that geodesic flows on such an M are Bernouillian. Then, if X is one such flow, we write

$$Z_m = \theta(m)X.$$

Again we cannot in general decide whether we have a trivial zero field or a Bernouillian flow. □

We conclude with an incompleteness theorem:

Proposition 10.3 *If T contains an axiomatization of dynamical system theory then there is an expression X in the language of T so that*

$$T \vdash T \textit{ is a dynamical system}$$

while

$$T \nvdash X \textit{ is chaotic}$$

and

$$T \nvdash \neg(X \textit{ is chaotic}),$$

for any nontrivial predicate in the language of T that formally characterizes chaotic dynamical systems. $\square$

That is to say, chaos theory is undecidable and incomplete (in its axiomatic version), no matter what nontrivial definition that we get for chaos.

11 Janus–faced physics

Theoretical physics has two faces. On one side, it allows us to do computations, to quantitatively obtain data that describe and predict the behavior of real–world systems. On the other side it allows us to imagine the inner workings of the phenomena out of the mathematical tools used in their description. This is the 'conceptual game' we mentioned before; we believe that it clarifies and complements Chaitin's vision of physical theories as algorithmic devices. The plethora of incompleteness results we've offered is of course a parallel to his vision of randomness as deeply ingrained into mathematics.

Let's consider one aspect of our work as an example. We have axiomatized physics with axiom systems where the dynamical rule is given by Dirac–like equations, instead of the more commonplace variational principles. Dirac–like equations are today an important tool in differential geometry [57], where they are used in the classification of bundles over 4–dimensional manifolds (and 4–dimensional differential manifolds are our current formal depiction of space-times). They appear in K–theory and in modern theories of cohomology. When one uses Dirac–like equations in the axiomatization of physical theories one wishes to stress the wide–ranging relations that such a mathematical object has within today's geometry. Dirac–like equations are a kind of crossroad–concept in today's mathematics. They lead to manifold routes, many of them still unexplored.

We've briefly mentioned 4–dimensional differential manifolds. The problem of the physical meaning of exotic, "fake" spacetimes, if any, is still wide open. We've, again briefly, mentioned it before [15], when we pointed out that there is an extra difficulty in that question: once we have uncountable many exotic differentiable structures for some fixed adequate topological 4–manifold, we will have uncountable many *set theoretically generic* differentiable structures for that manifold in adequate models for our theory. What is their meaning?

We don't know. Our axiomatization for classical and first–quantized physics opens up large vistas towards unknown, totally new, landscapes. Such is its *raison d'être.*

Acknowledgments

The authors wish to thank their institutions for support. N. C. A. da Costa acknowledges support from CNPq, Philosophy Section; F. A. Doria thanks C. A. Cosenza and S. Fuchs for their invitation to join respectively the Fuzzy Sets Research Group and the Philosophy of Science Group at the Production Engineering Program, COPPE–UFRJ.

References

[1] R. Abraham, J. Marsden, *Foundations of Mechanics*, 2nd. ed., Addison–Wesley (1978).

[2] S. Albeverio, J. E. Fenstad, R. Högh–Krohn, T. Lindström, *Nonstandard Methods in Stochastic Analysis and Mathematical Physics*, Academic (1986).

[3] V. I. Arnol'd, *Les Méthodes Mathématiques de la Mécanique Classique*, Mir, Moscow (1976).

[4] C. J. Ash and J. Knight, *Computable Structures and the Hyperarithmetical Hierarchy*, Elsevier (2000).

[5] M. F. Atiyah, *Geometry of Yang–Mills Fields*, Lezioni Fermiane, Pisa (1979).

[6] P. A. Benioff, "Models of Zermelo Frankel set theory as carriers for the mathematics of physics. I," *J. Math. Phys.* **32**, 618 (1976).

[7] P. A. Benioff, "Models of Zermelo Frankel set theory as carriers for the mathematics of physics. II," *J. Math. Phys.* **32**, 629 (1976).

[8] N. Bourbaki, *Set Theory*, Hermann and Addison–Wesley (1968).

[9] R. Carnap, *The Logical Syntax of Language*, Routledge and Kegan Paul (1949).

[10] R. Carnap, *Introduction to Symbolic Logic and its Applications*, Dover (1958).

[11] Y. M. Cho, "Higher–dimensional unifications of gravitation and gauge theories," *J. Math. Physics* **16**, 2029 (1975).

[12] L. Corry, "David Hilbert and the axiomatization of physics (1894–1905)," *Arch. Hist. Exact Sciences* **51**, 83 (1997).

[13] E. M. Corson, *Introduction to Tensors, Spinors and Relativistic Wave–Equations*, Blackie & Sons. (1953).

[14] N. C. A. da Costa and R. Chuaqui, "On Suppes' set–theoretical predicates," *Erkenntnis* **29** 95 (1988).

[15] N. C. A. da Costa and F. A. Doria, "A Suppes predicate for general relativity and set–theoretically generic spacetimes," *Int. J. Theoretical Phys.* **29**, 935 (1990).

[16] N. C. A. da Costa and F. A. Doria, "Undecidability and incompleteness in classical mechanics," *Int. J. Theoretical Physics* **30**, 1041 (1991).

[17] N. C. A. da Costa and F. A. Doria, "Suppes predicates and the construction of unsolvable Problems in the axiomatized sciences," P. Humphreys, ed., *Patrick Suppes, Scientific Philosopher*, II, 151–191 Kluwer (1994).

[18] N. C. A. da Costa and F. A. Doria, "Computing the Future," in K. V. Velupillai, ed., *Computability, Complexity and Constructivity in Economic Analysis*, Blackwell (2005).

[19] N. C. A. da Costa, F. A. Doria and E. Bir, "On the metamathematics of P *vs.* NP," to appear in *Appl. Math. Computation* (2007).

[20] N. C. A. da Costa and S. French, *Science and Partial Truth*, Oxford (2003).

[21] M. Davis, "Hilbert's Tenth Problem is unsolvable," *Amer. Math. Monthly* **80**, 233 (1973).

[22] M. Davis, *Computability and Unsolvability*, Dover (1982).

[23] P. A. M. Dirac, *The Principles of Quantum Mechanics*, Oxford U. P. (1967).

[24] F. A. Doria, "A Dirac–like equation for the gravitational field," *Lett Nuovo Cimento* **14**, 480 (1975).

[25] F. A. Doria, "A Lagrangian formulation for noninteracting high–spin fields," *J. Math. Phys.* **18**, 564 (1977).

[26] F. A. Doria, "Informal and formal mathematics," to appear in *Synthèse* (2007).

[27] F. A. Doria, A. F. Furtado do Amaral, S. M. Abrahão, "A Dirac–like equation for gauge fields," *Progr. theor. Phys* **75**, 1440 (1986).

[28] F. A. Doria and J. F. Costa, eds., Special issue on hypercomputation, *Applied Math. Computation* **178** (2006).

[29] A. Einstein, *The Meaning of Relativity*, Methuen (1967).

[30] A. Einstein and others, *The Principle of Relativity*, Dover s/d.

[31] G. Emch, *Algebraic Methods in Statistical Mechanics and Quantum Field Theory*, Wiley (1972).

[32] P. R. Halmos, *Naive Set Theory*, Van Nostrand (1960).

[33] H. Hertz, *The Principles of Mechanics*, transl. by D. E. Jones and J. T. Walley, Dover (1956).

[34] D. Hilbert, "Mathematical Problems," in F. E. Browder, ed., *Mathematical Developments Arising from the Hilbert Problems,* Proc. Symp. Pure Math. **28**, AMS (1976).

[35] M. Hirsch, "The Chaos of Dynamical Systems," in P. Fischer and W. R. Smith, eds., *Chaos, Fractals and Dynamics*, Marcel Dekker (1985).

[36] P. J. Holmes and J. Marsden, "Horseshoes in perturbations of Hamiltonian systems with 2 degrees of freedom," *Commun. Math. Phys.* **82**, 523 (1982).

[37] J. D. Jackson, *Classical Electrodynamics*, John Wiley (1962).

[38] J. H. Jeans, *Mathematical Theory of Electricity and Magnetism*, Cambridge U. P. (1925).

[39] T. Jech, *Set Theory*, Academic Press (1978).

[40] F. W. Kamber and P. Tondeur, *Foliated Bundles and Characteristic Classes*, Lecture Notes in Mathematics # 493, Springer (1975).

[41] G. T. Kneebone, *Mathematical Logic*, Van Nostrand (1963).

[42] C. Lanczos, *The Variational Principles of Mechanics*, U. of Toronto Press (1977).

[43] E. Lorenz, "Deterministic nonperiodic flow," *J. Atmos. Sci.* **20**, 130 (1963).

[44] A. J. Lotka, *Elements of Mathematical Biology*, Dover (1956).

[45] G. Ludwig, *Les Structures de Base d'une Théorie Physique*, Springer (1990).

[46] C. A. Lungarzo, "Superposition of States in Quantum Logic from a Set–Theoretical Point of View," in A. I. Arruda, N. C. A. da Costa and R. Chuaqui, eds., *Mathematical Logic—Proceedings of the First Brazilian Conference*, Lecture Notes in Pure and Applied Math., Marcel Dekker (1976).

[47] S. MacLane, *Geometrical Mechanics*, monograph, Dept. of Mathematics, U. of Chicago (1968).

[48] M. Machtey and P. Young, *An Introduction to the General Theory of Algorithms*, North–Holland (1979).

[49] J. D. Maitland Wright, "All operators on Hilbert space are bounded," *Bull. of the American Math. Soc.* **79**, 1247 (1973).

[50] J. C. Maxwell, *A Treatise on Electricity and Magnetism*, I and II, Dover (1954).

[51] E. Mendelson, *Introduction to Mathematical Logic*, 4th ed., Chapman & Hall (1997).

[52] A. O'Rahilly, *Electromagnetic Theory*, I and II, Dover (1965).

[53] E. Post, "Recursively enumerable sets of positive integers and their decision problems," *Bull. AMS* **50**, 284 (1944).

[54] H. Rogers Jr., *Theory of Recursive Functions and Effective Computability*, McGraw–Hill (1967).

[55] B. Russell, *The Analysis of Matter*, Dover (1954).

[56] P. A. Samuelson, *Foundations of Economic Analysis*, Atheneum (1967).

[57] A. Scorpan, *The Wild World of 4–Manifolds*, AMS (2005).

[58] R. M. Solovay, "A model of set theory in which every set of reals is Lebesgue measurable," *Ann. Math,* **92**, 1 (1970).

[59] J. A. Stratton, *Electromagnetic Theory*, McGraw–Hill (1941).

[60] S. Sternberg, *Lectures on Differential Geometry*, Benjamin (1964).

[61] P. Suppes, *Introduction to Logic*, Van Nostrand (1957).

[62] P. Suppes, *Representation and Invariance of Scientific Structures*, CSLI, Stanford University (2002).

[63] M. A. Tonnelat, *Les Théories Unitaires de l'Électromagnétisme et de la Gravitation*, Gauthier–Villars (1965).

[64] A. M. Turing, "On computable numbers, with an application to the Entscheidungsproblem," *Proc. London Math. Society* **50**, 230 (1937).

[65] R. Utiyama, "Invariant theoretical interpretation of interaction," *Phys. Review* **101**, 1597 (1956).

[66] B. L. van der Waerden, ed., *Sources of Quantum Mechanics*, Dover (1968).

[67] H. Weyl, *Space Time Matter*, Dover (1952).

[68] A. S. Wightman, "Hilbert's Sixth Problem: Mathematical Treatment of the Axioms of Physics," in F. E. Browder, ed., *Mathematical Developments Arising from the Hilbert Problems*, Proc. Symp. Pure Math. **28**, AMS (1976).