Mathematical Undecidability,

Quantum Nonlocality

and the Question of the Existence of God

Edited by:

A. Driessen and A. Suarez

With contributions of:

F.T. Arecchi
J.S. Bell
F. Cacace
G.J. Chaitin
P. Davies
A. Driessen
A.M. Fox
J. Laeuffer
P. Pliska
H.C. Reichel
J.M. Schins
A. Suarez

Submitted to:
Kluwer Academic Publisher,

Available: August 1997
now available at Springer

Mathematical Undecidability, Quantum Nonlocality and the Question of the Existence of God

Table of content

Preface

On January 22, 1990, the late John Bell held at CERN (European Laboratory for Particle Physics), Geneva a seminar organized by the Center of Quantum Philosophy, that at this time was a association of scientists interested in the interpretation of quantum mechanics. In this seminar Bell presented once again his famous theorem. Thereafter a discussion took place in which not only physical but also highly speculative epistemological and philosophical questions were vividly debated. The list of topics included: assumption of free will in Bell's theorem, the understanding of mind, the relationship between the mathematical and the physical world, the existence of unobservable causes and the limits of human knowledge in mathematics and physics.

Encouraged by this stimulating discussion some of the participants decided to found an Institute for Interdisciplinary Studies (IIS) to promote philosophical and interdisciplinary reflection on the advances of science. Meanwhile the IIS has associated its activities with the Fondation du Léman, a Swiss foundation registered in Geneva. With its activities the IIS intends to strengthen the unity between the professional activities in science and the reflection on fundamental philosophical questions. In addition the interdisciplinary approach is expected to give a contribution to the progress of science and the socio-economic development. At present three working groups are active within the IIS, i.e.:
- The Center for Quantum Philosophy
- The Wealth Creation and Sustainable Development Group
- The Neural Science Group

Since the talk by John Bell, and encouraged by him in those months before his unexpected death, the Center for Quantum Philosophy of the IIS has organized a number of seminars at CERN and promoted several symposiums and seminars in collaboration with different European University Institutes and Foundations[1,2]. During the holidays around the New Year of 1993 a group of scientists and University students met in the Italian Alps for a symposium on Mathematical Undecidability, Quantum Nonlocality and the Question of the Existence of God. Each day of this meeting had its introductory lectures, followed by a vivid and informal discussion. Being present at all presentations and discussions the editors could observe a more than usual interest in the issues dealt with. Especially the manifold of disciplines represented by the attendees, like philosophy, physics, chemistry, mathematics and information science gave rise to unusual observations and remarks.

The unceasing interest in the discussion on the relationship between Undecidability, Nonlocality and the Existence of God led the editors to the idea to prepare the present publication based mainly on the contributions of the above mentioned symposium. They are aware that the book cannot claim for any completeness. The experience, however, of the vivid exchange of thoughts in the different meetings justifies a preliminary summary of ideas. It is hoped that with this edition further debates and discussions will be strongly stimulated.

The title chosen is intentionally quite provoking. How can mathematics and quantum physics be brought in connection with God? The introductory chapter intends to show the connection. The body of the book is distributed in three parts. The first discusses the nature of mathematical knowledge, complexity and undecidability. Thereafter physics and nonlocality

[1] H. Thomas, editor, *Naturherrschaft*, Busse Seewald, Herford, Germany, 1991
[2] H.-Ch. Reichel, and E. Prat de la Riba, editors, *Naturwissenschaft und Weltbild*, Hölder-Pichler-Tempsky, Vienna, 1992.

plays a central role. At the end general aspects of science and meta-science are dealt with and brought in relation with the existence of God.

The block about mathematics starts with the contribution of Hans-Christian Reichel, from the Institute of Mathematics of the University of Vienna, about the recent developments in mathematics and the relation with the philosophy of mathematics. He discusses several concepts, like meaning, exactness, chaos, incompleteness, experimental mathematics, which ask for a deeper philosophical reflection. Gregory J. Chaitin from IBM Research Division, New York, one of the leading researchers on incompleteness and undecidability, explains why there are fundamental limits to what we will ever be able to understand. Moreover he establishes the somewhat astonishing parallelism between mathematics and statistical mechanics, and associates Gödel's incompleteness theorem with quantum mechanics. The next contribution is written by Filippo Cacace, an information scientist from the University of Napoli, who examines the relation between undecidability and the universality of a given mathematical problem. His analysis of the implications of the Turing theorem corroborates the existence of limits in man's knowledge of physical reality. Antoine Suarez, physicist and philosopher of the Institute for Interdisciplinary Studies in Zurich, proves in a simple way that in arithmetic there always will be solvable unsolved problems and raises the question whether such problems refer not to the existence of a superior intelligence outside human mind. In any case, the Kantian view that mathematics is an *a priori* mode of man's thinking has to be given up. Juleon M. Schins, physicist from the University of Twente, finishes the first part with a discussion of Roger Penrose's interpretation of the Gödel and Turing theorems[3].

The second part on physics and nonlocality starts with a contribution of F. Tito Arecchi, Director of the National Optics Laboratory at the University of Florence. He discusses the role of the three C's - catastrophe, chaos and complexity - for science and includes philosophical considerations. Thereafter the transcript is given of the above mentioned lecture by John Bell at CERN on indeterminism and nonlocality. It includes also a selection of the informal one-hour discussion. In the following three contributions, written by physicists, the arguments leading to the Bell inequalities are further explained and an overview of experimental work is presented. Paul Pliska, physicist of the Institute for Interdisciplinary Studies and patent attorney in Zurich, explains in a for non-physicists understandable language the nonlocality theorems by Bell, Greenberger-Horne-Zeilinger (GHZ), and Hardy. In addition he gives some comments on the relationship between the principles of free experimentation and nonlocality. Mark Fox from the Department of Physics in Oxford, reviews recent experiments on nonlocality. He emphasizes that one should make a distinction between Einstein's local-realistic view that rejects faster-than-light influences in nature, and the realism that admits a physical reality that is not exclusively a construction of the human mind. Antoine Suarez presents the principles of a nonlocal causal model, in which one is not led to assume an absolute time or a universal order of succession, and describes possible experiments where the prediction of this theory would be in contrast to quantum mechanics. Juleon M. Schins finishes this section with a discussion of the recent book of Bernard D'Espagnat on Veiled Reality[4], who in the context of quantum mechanics proposes an intermediary position between conventional realism and radical idealism.

The third part of the book deals with several specific philosophical aspects and draws some conclusions. Jacques Laeuffer, Expert-Engineer for G.E. Medical Systems in Paris, discusses Laplace's and Comte's assumption that one day human beings will enjoy an absolute and total

[3] R. Penrose, *The Emperor's New Mind,* Oxford University press (1989), *Shadows of the Mind,* Oxford University Press (1994).
[4] B.D'Espagnat, *Veiled Reality,* Addison-Wesley, Reading, Mass. (1995)

knowledge of the world. He also includes an analysis of the consequences of this scientistic or positivistic mentality for the knowledge of the existence of God. What follows then is the Templeton Prize Address 1995 of Paul Davies on *Physics and the Mind of God*, who asks the burning question, where do the laws of physics come from? Arguments sustaining that these laws are our laws, not originated by nature, he considers as *arrant nonsense*. in his view the universe is *a coherent, rational, elegant and harmonious expression of a deep and purposeful meaning*. Alfred Driessen, from the Department of Physics of the University of Twente and member of the board of the IIS, evaluates the bestseller of Stephen Hawking, *A brief History of Time*[5], which is *also a book about God...or perhaps about the absence of God*[6]. He asks whether the impossibility to demonstrate the beginning of the world in time, as stated by Hawking, leads to the impossibility to demonstrate the existence of God. The final remarks by Alfred Driessen and Antoine Suarez summarize the main stream of argumentation in this book. By evaluations of recent results in mathematics and physics man becomes aware of fundamental limits in knowing and controlling reality. The plenitude and power of reality seems to exceed the capacities humanity will ever possess. In this way science opens the road to a reality which people in general call God.

The book would not be finished without the willingness of the different authors and their contributions to the discussion, which helped to crystallize the ideas on this highly fascinating subject. The authors would like to thank J.M. Schins and A. M. Fox for their critical reading of some of the manuscripts. Kluwer Academic Publishers are gratefully acknowledged for their interest and stimulation for this edition.

Enschede, 26 July 1996

A. Driessen and A. Suarez

[5] S. Hawking, *A brief history of time*, Bantam Books, New York, (1988).
[6] C. Sagan in the Introduction of ref. 5

Chapter 0

INTRODUCTION

Alfred Driessen
Department of Applied Physics, University of Twente, Enschede, The Netherlands

and

Antoine Suarez
Institute for Interdisciplinary Studies, P.O. Box 304, 8044 Zurich, Switzerland

1. Reality is intelligible and man's knowledge and power limited

The title of the present book suggests that scientific results obtained in mathematics and quantum physics can be in some way related to the question of the existence of God. This seems possible to us, because it is our conviction that reality in all its dimensions is intelligible. The really impressive progress in science and technology demonstrates that we can trust our intellect, and that nature is not offering us a collection of meaningless absurdities. Having achieved an amazing control over nature man is tempted to declare that "hard" science (i.e. the knowledge about what is observable and controllable) is the only reliable knowledge.[7]

It is remarkable, however, that present day science itself is inviting us to cross the boundary of what we can observe and control, and to enter the domain of meta-science. The same logic used in science seems to show also man's inherent limitations in power and knowledge. And that is what we first of all intend to show with results taken from mathematics and quantum physics:

- Mathematical Undecidability refers to the result that man will never have a universal method to solve any mathematical problem. In arithmetic there always will be unsolved, solvable problems.
- Quantum Nonlocality refers to the discovery that certain phenomena in nature seem to imply the existence of correlations based on faster-than-light influences. These influences, however, are not accessible to manipulation by man for use in, for example, faster than light communication.

In the various contributions pieces of a puzzle are offered, which suggest that there exists more than the world of phenomena around us. The results discussed point to intelligent and unobservable causes governing the world. One is led to perceive the shade of a reality which many people would call God. In order to avoid confusion it is important from the very beginning to give a clear definition of what is meant by God in the title of this book.

2. What is it that people call God?

[7] This attitude represents the positivistic form of skepticism. Already the founding fathers of metaphysics, Plato and Aristotle, have been confronted with the skepticism of sophists, who denied the possibility of finding the truth. Aristotle argued that the fundamental statement of philosophical skepticism, that it is not possible to affirm anything truly, is already an absolute statement not demonstrable by science. By accepting this statement as an absolute truth, one is therefore in contradiction with its meaning. (cf. Aristotle, *Metaphysics*, Book XI - Part 5). Plato and Aristotle's endeavor may encourage also present day men to search for truth beyond the observable and controllable.

When analyzing what people call God, one encounters in general two basic elements. First, there is an event or phenomenon which one tries to understand, to explain or eventually to control. Second, there is the awareness of one's own fundamental limits of understanding and controlling that specific phenomenon. These two elements lead people to think about the existence of God. It is in a certain sense the application of the principle of causality: if there is an effect, then there should be a cause. More specifically, if there is an effect which exceeds human knowledge and power by a large amount, then there should be a cause superior to man. One may say that the first idea of God is that of a being who knows also those things we don't know and has the power to do things we cannot do.

This view appears very clear in a conversation between the physicists Paul Davies and Richard Feynman. Referring to the laws of physics, Davies comments: *at one time people used to believe that God explained the universe. It seems now that these laws of physics are almost playing the role of God - that they're omnipotent and omniscient*. And Feynman answers: *On the contrary. God was always invented to explain mystery. God is always invented to explain those things that you do not understand. Now when you finally discover how something works, you get some laws which you're taking away from God; you don't need him anymore. But you need him for the other mysteries... God is always associated with those things that you do not understand. Therefore I don't think the laws can be considered to be like God because they have been figured out.*[8]

The postulate that man will one day, by means of science, reach a complete knowledge and control of nature and history was the heart of the materialistic worldview which arose in the XIX century [9] and up to now has exerted great influence. As Feynman suggests, the human mind is led to conclude the existence of an omniscient and omnipotent being, because this is a necessary assumption to explain those things that now are beyond human understanding and power. Consequently, if one assumes that one day man will be capable of explaining all mysteries, then one would not need God anymore. On the contrary, if it is possible to demonstrate that at any time in the future there will be always things beyond human understanding - *mysteries* - then one has to conclude that human knowledge and power is limited and inferior to divine knowledge and power. Therefore, any position arguing that human beings are capable, in principle, of reaching complete knowledge and control of the world, becomes also a position arguing against the existence of God.

3. The classical proofs of the existence of God and the classical criticism of these proofs

Any proof of the existence of God implicitly refers to unobservable causes moving the world from beyond the realm of space and time, or assumes that there is an order in the world that does not arise from the human mind. The first is, for instance, the case of the proof of the unmoved mover that Aristotle gives in his Physics.[10] He firstly states that movement in a circle as observed in astronomy is eternal and then concludes that there should be an eternal first unmoved mover who is the cause of the movement of the heavenly bodies. As the basis for his proof Aristotle clearly searched for a phenomenon that cannot be explained by a

[8] P.C.W. Davies and J. Brown, *Superstrings: A Theory of Everything?* Cambridge University Press, 1988, p.208-209.
[9] see for example J. Laeuffer, *Scientism and scientific knowledge of things and God,* this volume, chapter XII.
[10] Aristotle, *Physics*, VIII, 9-10.

temporal chain of causes alone.[11] However, the choice of the movement of the heavenly bodies was not a lucky one, as the arrival of Galileo's and Newton's mechanics showed.

Thomas Aquinas, in his fifth way, states that the natural bodies lacking intelligence move not fortuitously but towards an end, and concludes that some intelligent being exists by whom these bodies are directed and governed.[12] Paul Davies reaches, on the basis of today's physics, basically the same conclusion, as he states that the universe is *a coherent rational and harmonious expression of a deep and purposeful meaning.*[13] Notice that from the point of view of today's physics this argument is stronger than that of movement by Aristotle: In his 5th way Aquinas takes the appearance of correlations in nature as the very indication for the presence of causality and consequently of a cause.

A drastic change in the appreciation of the classical proofs of the existence of God could be observed after the discovery in the 16th century that the laws of nature are written in a mathematical language. This discovery weakened the arguments of Aristotle and Aquinas, at least when one adds the assumption that mathematical truth is a man-made construction, and that behind the phenomena there is no physical reality which escapes observation. On the basis of these assumptions one may correctly conclude the following: If the cause of each phenomenon can be observed, and what is observed can be described in terms of mathematical information (algorithms or sequences of bits), and mathematics is a construction of the human mind, then the laws behind the phenomena and even the phenomena themselves are a man-made construction. Physical reality reduces to man-made virtual reality, and there is no reason why we, human beings, will not actually know and control the future of the world.

The postulates that the laws of nature originate from human modes of thinking, and that every physical phenomenon can be explained by observable causes, are the heart of Kant's criticism of the proofs of the existence of God.[14] They are also characteristic of the positions of Laplace and Comte.[3] The summary of modern atheism is this: If one considers thinking as an activity in which God does not play any role at all, at the end one considers the exterior world as a domain in which God does not play any role either. The belief that God is not present in the thinking methods with which we become capable of mastering nature, is in practice widely accepted. For many there is a separation and intellectual incompatibility between scientific activities and religious life. The Dutch Nobel prize-winner, Simon van der Meer expressed this as follows: *As a physicist, you have to have a split personality to be still able to believe in a god.*[15]

4. The aim of this book

Without any doubt the discovery that the phenomena can be described in mathematical terms has been of great benefit to the scientific, economic and cultural development of mankind. The above mentioned assumptions, however, that mathematics is a man-made construction and that observable things originate exclusively from observable things, are not so obvious at all. In the field of philosophy many arguments can be presented against it, but also new insight obtained in scientific research gives evidence for its disputable character.

---

[11] see, e.g., A. Driessen, *The question of the existence of God in the book of Stephen Hawking: A brief history of time*, this volume, chapter XIV.
[12] Thomas Aquinas, *Summa Theologiae* I. q.2, a.3, see also Aristotle, *Fragments*. Dialogues F 10 R^3.
[13] P. Davies *Physics and the Mind of God*, this volume, chapter XIII.
[14] I. Kant, *Kritik der reinen Vernunft*, B 647-B 667
[15] S. van der Meer, news-paper interview, NRC-Handelsblad, Amsterdam, 18-4-1987.

In this book, recent mathematical theorems are discussed, which show that man never will reach complete mathematical knowledge. Also experimental evidence is presented that physical reality will always partially remain veiled to man, inaccessible to his control. It is intended to provide, in the various contributions, the pieces of a puzzle which restore the possibility of a natural, intellectual access to the existence of an omniscient and omnipotent being.

One should bear in mind, that the argumentation of this book is based exclusively on noncontradictory thinking and observations from physical reality. When we therefore consider God, it is the reality behind the phenomena which people assume when they realize their fundamental limits in knowing and doing. But as said above, this principle that explains the mysteries we cannot explain, is essential to the idea of God as First Cause or First Mover, already encountered, for example, in pre-Christian Greek philosophy.

In this use of the term *God* we do not necessarily include God as the Creator or first temporal cause. In our opinion, the idea of Creation as a beginning of the world in time is not something that appears to the mind as a necessary conclusion. In this respect it is interesting to note that in most of Greek philosophy, including Aristotle, the notion of a Creator is absent. Moreover, in the history of Christianity no formal declaration has been given that it is possible to prove by rational means the beginning of the world in time. Also philosophers like Aquinas, who delivered proofs of the existence of God, denied the possibility of a rational demonstration of the beginning of the world in time.[16] Regarding the widely spread view of the Big Bang as the beginning of the Universe, we would like to emphasize that scenarios for the very early Universe have been proposed, in which the Big-Bang does not mark any beginning but is an event occurring in time,[17] or where the character of a singularity has been removed.[18]

In comparison with the great monotheistic religions, the principle we refer to as God or even the God of the philosophers is conceptually still quite undeveloped. It is comparable to the notion of a Beethoven symphony for a person deaf from birth, who knows musical theory and can read the score. He gets real knowledge, but misses almost completely the full richness of the musical experience. Why then do we make the effort and go through all of this detailed reasoning? It is because of an optimistic vision on the capacities of human beings. Man with his intellectual effort is able to know the existence of an unobservable reality, which he already encounters deep in his heart. We think that the reflections presented in this book may contribute to showing that science itself can become a road leading to God. And this may undoubtedly promote the unity between man's scientific and professional activity, and his religious life.

[16] Thomas Aquinas, *Summa Theologiae*, I, q. 46, a.2.
[17] G. Veneziano, *String Cosmology: Basic ideas and general results*, CERN-TH/95-254, September 1995
[18] Cf. S. Hawking, *A brief history of time,* Bantam Books, New York (1988).

Chapter I

How can or should the recent developments in mathematics influence the philosophy of mathematics?

Hans-Christian Reichel
Institute for Mathematics, University of Vienna

The following contribution is reproduced with the kind permission of the Institute Vienna Circle[19].

1. Introduction

Mathematics and philosophy not only have common roots, they have also continually influenced each other and shared close connections for at least 2500 years. Over and over, at decisive moments in the history of philosophy, mathematical insights and results served as guideposts for the development of philosophy. One need think only of Plato, of Pascal, of Descartes, of Wittgenstein!(See e.g. [30]).

The theme addressed here is different, however. It is on philosophy of mathematics. Here we find *ontological* and *epistemological* problems of mathematics in the foreground: What is mathematics about? What is the nature of the objects it studies? What kind of being, of existence are shared by mathematical entities? Are its concepts and methods *discovered* or *invented*? And what kind of *knowledge* does mathematics provide?

Many other problems are included in philosophy of mathematics, but the problem of the *foundations* of mathematics plays a central role. Indeed, this topic became a principle issue of philosophy of mathematics after the foundation of set theory by Georg Cantor, who also discovered paradoxes in it (think of Bertrand Russell in this connection!). Perhaps even more essential for this development was the flourishing of logic by the second half of the 19th century: in this connection we must mention primarily Bolzano, Husserl, Boole and Frege, among others.

It was above all the purging from logic of psychologistic elements, i.e. its *formalization*, which culminated in the axiomatization of arithmetic and geometry. New areas of research arose through this axiomatic ways of seeing and thinking. What we now call *formalism* seems to rule all mathematics (one need only consider some of Hilbert's problems of 1900).

As a principle result of this development of mathematics in the second half of the 19th century, *all of mathematics* was thought to be reducible to logic (we refer to Whitehead and Russell's *Principia Mathematica*). But opposed to this was what we summarize under the titles of intuitionism and constructivism (Brouwer, H. Weyl, Heyting, van Dalen, Bishop. Bridges, Beeson); (cf. [5], [2], among many others).

Indeed, we may say, in retrospect, that philosophy of mathematics, until late into this century was entirely confined in foundational debate to Platonism. formalism logicism, intuitionism and constructivism. And although this took place in the most variegated manner,

[19] Yearbook of the Institute Vienna Circle 3/95, Ed. F. Stadler, Kluwer Academic Publ., Dordrecht 1995.

the debate seems to have exhausted itself somewhat, above all in that no appreciable answers of philosophers have been forthcoming to the challenge of Kurt Gödel and other more recent developments in mathematics (see, e.g., [31]).

In the second half of our century, there have been no really important debates or new innovations in the philosophy of mathematics. Of course there are several important studies: Lorenzen, Stegmüller, Lakatos, among others in Europe, MacLane, Davis, Hersh, Tymoczko and others in the USA, such as the recently published collection by White, as well as works by Kac, Rota and Schwartz (see the References!)

Nevertheless: *The present impasse in mathematical philosophy is the aftermath of the great period of foundationist controversies from Frege and Russell through Brouwer, Hilbert and Gödel. What is needed now is a new beginning,* as Reuben Hersh recently put it.

Let me now formulate in a few words which problems should be asked by philosophy of mathematics today; and let me name just a few milestones and new trends in mathematics out of which I think quite new developments will emerge for philosophy. One of the questions in the philosophy of mathematics which must be asked anew is the question after the meaning of *mathematical sentences* and *theories*. I want to support this immediately with some examples.

2. New trends relevant for the philosophy of mathematics

2.1 Independence results of set theory

In the first place, I would like to mention the *independence results of set theory* with their deep and manifold influence on other parts of mathematics and its foundations. Gödel showed in 1938, with the help of his constructibility axiom V = L, that the continuum hypothesis and the axiom of choice are consistent with the classical axioms of set theory.

However, Paul Cohen showed in 1963 that the negation of the continuum hypothesis is consistent with the Zermelo-Fraenkel axioms as well. His "forcing" method (together with Boolean-valued model theory) provided several additional independence results since then- consider, e.g., Martin's axiom, the "club" axiom and other axioms of transfinite combinatorics, an area with many developments and surprises.

Above all, we think of the existence of large cardinals independent of ZF, as well as other existence propositions with all their consequences for the theorems of topology, algebra and measure theory. Thus there exist models of set theory in which the cardinality of the continuum can take practically arbitrary values. All of these developments represent results which certainly have philosophical consequences in the sense mentioned above, and even though they may not immediately effect concrete applications, they certainly effect analysis and its applications and will hence influence all of mathematics.

We must mention above all the fact that Gödel's axiom of constructibility is not fully sufficient for set theory, thus making a stronger framework plausible and necessary. Furthermore, since Gödel’s incompleteness proof of 193l showed that not only arithmetic but also set theory can never completely captured within the framework of a single axiom system, nor can even their consistency ever be established internally, we may conclude that the entire edifice of classical mathematics cannot be completely comprised in any formal scheme. This fact certainly presents one of the prime challenges to a modern philosophy of mathematics: it must not merely treat set theory, the classical foundation of mathematics, but also the role of

symbolism, of the meanings of mathematical expressions anew. And it must do this against the background of further developments in computer mathematics - we shall return to this below.

Finally, it should be mentioned that the first independence result originates in the previous century already: this was of course that of the axiom of parallels of Euclidean geometry. Indeed, discussions continually reignite on this topic even up to the present day, even though our understanding of the independence of various mathematical propositions from classical assumptions has only really matured to its full philosophical dimension in the sixties and seventies of the present century.

2.2 Application of mathematics in science and profession

Another challenge for philosophy of mathematics is suggested by a topic just touched on: today's enormous range of *applications of mathematics* in the practical domain, in particular in all sciences and professions. The techniques applied here are based throughout on so-called mathematical models, with all their problems inherent in as well as outside of mathematics (see [31]). To consider only one example, we know that in a *continuous representation* of processes using one or two differential equations with, say, two parameters, chaotic solutions cannot arise - but in discrete representations using difference equations they can! The purely mathematical explanation for this situation presents no "technical" difficulty, typically following the theorem of Poincaré and Bendixon or the scenario of Feigenbaum, but a philosophical problem most certainly remains!

The same holds for the issue of whether mathematical models have any epistemological/explanatory value. Do mathematical models have a genuine heuristic and epistemological content for the underlying process studied, or for nature as such (whatever we may mean by this word)? Or must we view mathematical models exclusively according to Wittgenstein's demand in the *Tractatus* that *All explanation must go and give place to description*?[20] What is required is, in may opinion. a "philosophy" of *metaphors*, or put another way: of *mathematical models* - of course not in the sense of model theory in logic. More directly still, we need a philosophy of modern *applied and applications-oriented mathematics*.

Typical of mathematical modeling and its applications is of course *non-standard analysis*, which has taken on considerable importance in the last decades again, with its infinitesimally small and infinitely large "objects", which proceed from the hyperreal numbers and their applications in practically all branches of analysis and physics. But let us now turn to the next point!

2.3 Chaos theory

What we have just claimed of mathematical models holds - perhaps with even more force, as controversies of the most recent past demonstrates - of the entire area of *chaos* theory, i.e. of the theory of dynamic systems and their applications. Its philosophical foundation is very much an open question, and we find many problems hardly dealt with, for example the issue of *local vs. global causality* (cf. [9]; [4] and its bibliography). The most notable phenomenon here is the property of systems highly sensitive to changes in boundary conditions: *similar causes do not always have similar effects!* The "weak law of causality" is thus highly

[20] See [31]. Consider in this connection for example the models of catastrophe theory of the sixties and seventies, as well as many other examples!

debatable! The slightest rounding error, the slightest variation of initial conditions for a process can send it into dramatically different trajectories which can hardly be dealt with computationally. In this sense, Laplace's world-view needs amplification, to say the least. Strictly deterministic processes, speaking globally, can therefore behave in ways which can only be treated stochastically, among other surprises. It follows, for example, that experiments - even in the natural sciences - need in principle *not necessarily be repeatable at will*, since exactly identical initial conditions can rarely or never be reproduced, and system behavior at the points studied could always lie within a chaotic interval. Concerning chaos, one may refer in general to people such as Peitgen and the so-called Bremen school; on the other hand we have the extensive discussions triggered by the well-known and extremely detailed recent polemical article of Klaus Steffen [33]. There is a lot of work to do by way of a deeper philosophical foundation of non-linear, in particular chaotic, processes. Chaotic systems, strange attractors and other intrinsically *pure mathematical* concepts with all their genuine applications, but also trivializations and pseudo-applications in the natural sciences, economies, biology, medicine and other areas, are discussed by everyone nowadays, but often much too superficially. (Motto: "Not all that jerks and moves a little is a chaotic system".)

On the other- hand, as is well known, we do not even have a uniform and so-to-speak canonical definition of the concept of chaos, apart from the strong dependency of chaotic trajectories on their particular initial values.

2.4 Experimental mathematics

With this observation I land at my penultimate theme: the advance of so-called *experimental mathematics*, which can bring many changes in the classical image mathematics had of itself - although I do not want to exaggerate this. If one understands as an experiment - according to a definition of Braun and Rademacher - the systematic procedure of winning reproducible empirical data, this is relatively new for mathematics. Of course, mathematicians have always "experimented" in the broadest sense; thought experiments have always played an essential role in mathematical activities and have always served as starting points for mathematical theorizing. However, (especially computer) experiments have started transcending the *context of discovery* and must already be reckoned to the *context of justification* - to use terminology introduced by Reichenbach. Think of, for example, the theory of manifolds, of the topology of $\mathbf{R}^n$, of several results in number theory. Chaitin [8] quotes pertinent concrete examples, and he appeals to Boltzmann encouraging mathematics to behave a little like physicists. In the natural sciences, it is considered quite proper to make use of experiments in the context of justification. In classical mathematics, only proofs may be used in this way. The essential point of mathematical theories has always been, not conceptualization, but *argumentation*, i.e. *proof*.

A *proof in the classical sense* has three features: (1) A proof is convincing, (2) it is surveyable, (3) it is formalizable.

With respect to these features, the largely completed acceptance of computers in mathematics does indeed bring changes: as you know, there are - after many attempts to use computers to prove conjectures - many mathematical propositions for which "only" *computer proofs* are available! The proof of the four-color problem of 1976 is one of the best-known examples (features 1 and 2 are at issue here).

Other propositions and theories can no longer be surveyed by a single human being at all, because they are too ramified - think of, e.g., the classification of finite Abelian groups, which

probably no one except the recently deceased mathematician Gorenstein ever went through. The problem of the *complexity* of mathematical theories will have serious consequences, despite their generally accepted validity - not least because of the by now manifold, sometimes not completely surveyable, computer proofs (e.g. [9]).

Now that we have mathematical propositions which are computationally and in their content so complex that no single human being, not to say each of us, can entirely survey, the problem arises: can we accept things, can we work with things which no one of us can completely reproduce? The personal relation between man and mathematics itself is at stake! ([11], [23])

In this connection a corresponding re-evaluation is required of results which are shown - once again by computer - to be merely *very probable* rather than *true*. I am thinking here of *probabilistic proofs* and methods, such as the decomposition of large numbers into prime factors (or pseudo prime numbers), as this is nowadays done in coding theory. Coding theory and information theory are themselves prime candidates for the development of a new philosophy of mathematics.

Under "experimental mathematics" as such, we mean the large-scale employment of computer graphics, or the employment of parallel processing by way of generating new conjectures, maybe even *results*, in topology, in higher-dimensional geometry, number theory, and many other areas. Above all, once again, we seem to have here a shift from the context of discovery to the context of justification. However, to arrive at deeper mathematical propositions in this way requires an even greater resort to mathematical practice as a foundation than ever!

Proving theorems is only *one way* of doing mathematics. At least as important examples of mathematical activity are scientific computation; experiments, sketches and commenting of algorithms; manipulations in computer graphics; imaginative intuitive procedures in problem solving and heuristics; and so on! Philosophy of mathematics has yet to include them! *All* types of mathematical practice have their own laws of *rigor*, but they have yet to be discussed in philosophy - in contrast to so-called classical problems of proof and foundations.

Working with computers raises another issue relevant for philosophy: that of the *internal arithmetic* of employed machines and programs, and the associated reliability and robustness of results. This means: if one runs certain procedures, e.g. an iteration, on different machines, or if one executes them within different programs, different results could be obtained - above all when the iteration or differential equation involved lies in a chaotic domain, and the computation makes even the slightest rounding-off or approximation errors. However, such errors are unavoidable in principle, and they are even typical for the so-called internal arithmetic of the machine or the program - without even being explicit! Probably all of us know the example where, for given x_0 and the iteration:

$$x_{n+1} = 3.95\, x_n\, (1 - x_n)$$

the calculation of x_{100} yields considerably different values, all depending on what machines and what programs ran the iteration. (We may recognize here the effects of the Verhulst dynamics in the chaotic parameter-domain.) A similar example is that of executing the following *arithmetically equivalent* iterations with one and the same machine:

$$x_{n+1} = 4\, x_n\, (1 - x_n) \qquad \text{and} \qquad x_{n+1} = 4x_n - 4x_n^2.$$

The values for x_{65} already differ at the first decimal-place when calculating with ARTARIST: for $x_n = 0.1$, $x_{65} = 0.981299$ and $x_{65} = 1.006784$, respectively.

Despite all of these numerous and well-known examples, the employment of computers nonetheless loses none of its persuasiveness and importance. In sum: *computer proofs, computer graphics,* fractal geometry, artificial intelligence of course, virtual reality, so-called internal arithmetic of various computer techniques, the *employment of computers* at all, and associated with them *experimental mathematics* with all its new methods and problems, all these call for *new approaches in philosophy* of mathematics and logic.

2.5 Algorithmic information theory

A fifth and last item, again a recent development in mathematics, which in my opinion should receive philosophical attention, is *algorithmic information theory*, as it has been developed and described in the last decade by Gregory Chaitin. We have here to do with the definition and computation of the degree of complexity of 0-1 sequences and - more generally - with the complexity of axiom systems and of any systems encodable by numbers. A 0-1 sequence is called a "*random sequence*" whenever it cannot be algorithmically abbreviated, i.e. whenever the same number of bits are required to encode its shortest description as for its explicit display([6, 8]).

We thereby have achieved an important step, founded on work of Kolmogoroff and Martin-Löf, in mathematically explicating the concept of randomness and making it workable - in a literally computable manner.

The *complexity of a sentence*, encoded in some 0-1 sequence, is accordingly the length of the smallest computer program capable of writing out the sentence explicitly. In this sense, entire axiom systems will accordingly also have some numerical complexity value *m*. Starting with this definition, Chaitin presented Gödel's incompleteness theorem in an "algorithmic guise" by proving, back in 1974 already, that within a given formal system H - say, with complexity *k* - , it is not possible to prove that the complexity of any given 0-1 sequence exceeds some bound *m*, which is ultimately determined by *k*, i.e. by the system in which we calculate. It therefore follows that we cannot calculate the *degree of randomness* of sufficiently large number sequences, in particular not that of number sequences exceeding the constant *m* which characterized the complexity of the "working system" *H*.

But this result yields decisive consequences for the *philosophy of all natural sciences*. Consider an example from biology:

...1101100010011111101110100000100101110110111010110011010101011000010000001111101100001010011011000011001101011110010000001000001111110011011010000100000011000001010011111001010101010011001101111011010001000110010110111011001111010111110100100010000111010000. . .

The above shows a section of the genetic code of the virus MS2, the nucleotide sequence of the replicase gene (β-component)[17]. Our question is: is this 0-1 sequence random or can it be generated by some perhaps hidden principle?

Such questions naturally have a great impact on our whole world view: Is the evolution of life *random* or is it based on some *law*? (Jacques Monod vs. Teilhard du Chardin!) The only

answer which mathematics is prepared to give has just been indicated: the hypothesis of *randomness* is *unprovable* in principle, and conversely the *teleological* thesis is *irrefutable* in principle. If such statements have no philosophical importance, then I don 't know any important philosophical statements at all!

Similar consequences may be derived with respect to the mathematical notion of randomness in general, or more precisely, in connection with random sequences in mathematics. Thus, one can show for example that many different random sequences occurring in the course of mathematical work cannot be distinguished from each other in the sense discussed above. One of the most famous examples arose on the periphery of Hilbert's 10th problem, which was solved in 1970 by Matyashevich. Roughly speaking, he demonstrated the equivalence of Hilbert's 10th problem to the *halting problem* of Turing machines, which was already shown to be unsolvable by Turing (see e.g. [8, 7]).

By way of a certain modification of Matyashevich's proof, Chaitin was able to show that, for a suitably built-up system of diophantine equations, the question whether they have finitely or infinitely many solutions has, roughly speaking, a *random answer* - although the answer is apparently determinate!

Now I do not want to expand on these problems at this time; nevertheless, a new set of issues for philosophy of mathematics has been indicated. Recently a five-day course at the University of Maine was given on similar questions relating algorithmic information theory to the "limits of mathematics" - thus the course title.

3. Summary and conclusions

I now conclude my introductory sketch with a kind of summary:

(1) The original problem inquiring anew after the *meaning of mathematical propositions* now seems somewhat clearer. What is the *meaning* of mathematical propositions? - here we should hold to Wittgenstein's dictum: *Philosophy describes which propositions have a meaning*. With respect to applications, we may ask: what do mathematical sentences say; do mathematical models have any epistemic value? (What we learn from mathematics is just as unanswerable as "What do we learn from Tolstoy's ‘War and Peace’”!)

(2) What can we say, after all the remarks above, about the *exactness* of experiments and their results? Is the very concept of exactness itself problematic?

(3) Mathematics is the technique, which has been refined over the centuries, of solving problems. What are the questions arising from the necessary *symbolizing* of this, e.g. by way of formalizing when constructing theories, or when studying new non-linear phenomena (i.e. "chaotic" ones)? What are the problems arising from the use of computers, from "experimental mathematics", etc.?

(4) In the same connection we may put the question once again of the *meaning*, of our *understanding*, of the *self-assessment* of mathematics. In connection with computer proofs, probabilistic proofs and experimental mathematics, the question of the *truth value* of propositions arises which may perhaps have been obtained by experiment, by a probabilistic *simulation*. What is acknowledged today by the mathematical community as a proof and what is not?

(5) The actual *activity of the working mathematician* must be more strongly involved, in addition to an evaluation of his results - which in the classical sense are supposed, after all, to be independent of the *method used to obtain them*.

(6) In connection with methods of proof, a brief remark: after Gödel's proof of incompleteness, the issue of self-reference has become "presentable" in mathematical society. To what extent can calculi say anything about themselves, to what extent not?[21]

(7) Finally, we ought to discuss issues raised by Imre Lakatos, who calls present-day mathematics a quasi-empirical science (see [18]). This issue may also be more clearly understood in the course of the discussion above. "Quasi- empirical" means - as a first approximation, in my opinion - the balance peculiarly present in mathematics between empirical-experimental procedures on the one hand (which after all is typical for many mathematical activities), and "transcendental certainty" on the other. (“Transcendental” means here: beyond any empirical domain, beyond all personal experience or experimental confirmation.) The question of balance is a cardinal problem for the philosophy of recent mathematics!

This concludes my proposals. l originally set myself the task of presenting a framework, marked off by several examples from recent mathematics, within which new problems for the philosophy of mathematics arise and according to which the "foundational debate" needs to be extended. This was of course not at all intended to be exhaustive.

Thus, the entire question of the "*social context*" in which mathematics stands today has to be philosophically examined - for the topic of "mathematics and the public interest" clearly takes on a completely different interest today than it did just a few decades ago. (Just consider, e.g. the ethical, social and legal problems arising from mathematical advances in the area of coding theory alone, which has many practical ramifications concerning security, e.g. in restricting the use of atomic weapons.) Similar consequences hold in the area of "mathematics and schools", i.e. for mathematical *didactics*, and many other themes, which have become drastically changed since the first decades of this century (see [23], [11] and others).

Only one thing remains an absolute fixed point: the beginning and indeed the center of any philosophy of mathematics must lie in "*mathematical practice*", the work of the "working mathematician", mathematical experience.

It is the practice of mathematics that provides philosophy with its data, its problems and solutions. At the turn of the century it seemed as if foundationalism could capture the essence of mathematical practice. But in the last half of the century, foundational research and ordinary mathematical practice have evolved along quite different lines. To revive the philosophy of mathematics, we must return to the present work and mainstreams and from there continue the threads that always begin in the historical depths of our science.[22]

References

[1] Augarten, S.: *Bit by Bit An Illustrated History of Computers*, New York: Ticknor and Fields, 1984.

[21] Gödel enabled Peano arithmetic, so to speak, to talk about itself with the help of his Gödel numbers, as we call them today

[22] Quoted with slight revisions from [35].

[2] Beeson, M.J.: *Foundations of Constructive Mathematics: Metamathematical Studies*, Berlin, Heidelberg, New York: Springer, 1985.
[3] Benacerraf, P. and Putnam, H.: *Philosophy of Mathematics: Selected Readings*, 2nd edition, Cambridge:Cambridge University Press, 1984.
[4] Bigalke, H.-G.: *Chaostheorie: Eine neue Herausforderung für die Mathematikdidaktik*, Zentralblatt für die Didaktik der Mathematik 6 (1993), p. 197-208. [Good survey and bibliography].
[5] Bishop, E. and Bridges, D.: *Constructive Analysis*, Berlin, Heidelberg, New York: Springer, 1985.
[6] Chaitin, G.: *Algorithmic Information* Theory, Cambridge: Cambridge University Press, 1987.
[7] Chaitin, G.: *Information, Randomness, and Incompleteness Papers on Algorithmic Information Theory*, Singapore: World Scientific, 1987
[8] Chaitin G., *Number and randomness: Algorithmic Information Theory - new results on the foundation of mathematics*, this volume, chapter II; see also *Zahlen und Zufall - Algorithmische Informationstheorie, neueste Resultate*, in [26], p. 30-44.
[9] Davies, P.: *Prinzip Chaos*, Munich: Goldmann Verlag, 1988.
[10] Davis, P.J. and Hersh, R.: *The Mathematical Experience*, Basel & Boston: Birkhäuser, 1981.
[11] Fischer, R. and Malle, G.: *Mensch und Mathematik*, Mannheim, Vienna, Zurich: Bibliographisches Institut, 1985.
[12] Fritsch, R.: *Der Vierfarbensatz*, Mannheim, Vienna, Zurich: Bibliographisches Institut, 1993.
[13] Gödel, K.: *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme*, I", Monatshefte für Mathematik und Physik 38 (1931), p. 173 198.
[14] Van Heijenoort, J.: *From Frege to Gödel A Source Book in Logic*, Cambridge MA: Harvard University Press, 1967.
[15] Hintikka, J.: *The Philosophy of Mathematics*, Oxford & New York: Oxford University Press, 1969.
[16] Kac, M., Rota, G.-C. and Schwartz, J.T.*: Discrete Thoughts: Essays on Mathematics, Science and Philosophy*, Basel, Boston: Birkhäuser, 1992.
[17] Küppers, B.O.: *Zufall oder Planmässigkeit; Erkenntnistheoretische Aspekte der biologischen Informationsentstehung*, Biologie in unserer Zeit,13 (1983), p. 109 119.
[18] Lakatos, I.: *Mathematik, empirische Wissenschaft und Erkenntnistheorie*, Braunschweig: Vieweg, 1982.
[19] Lamber, J.: *Are the Traditional Philosophies of Mathematics Really Incompatible?*, Mathematical Intelligencer 16 (1) (1994), p. 56-62.
[20] Lorenzen, P.: *Konstruktive Wissenschaftstheorie*, Frankfurt: Suhrkamp Taschenbuch Wissenschaft, 1974.
[21] MacLane, S.: *Mathematics, Form and Function*, Berlin, Heidelberg, New York: Springer, 1986.
[22] Oeser, E. and Bonet, E. (eds.): Das Realismusproblem, Vienna: Edition S, Österreichische Staatsdruckerei, 1988.
[23] Otte, M.: *Das Formale, das Soziale und das Subjektive*, Frankfurt: Suhrkamp, 1994.
[24] Peitgen, H.-O., Jürgens, H. and Saupe, D.: *Fractals for the Classroom*, Berlin, Heidelberg, New York: Springer, 1992.
[25] Peitgen, H.-O. and Richter, P.H.: *The Beauty of Fractals*, Berlin, Heidelberg, New York: Springer, 1987.
[26] Prat, E. und Reichel, H.-C. (eds.): Naturwissenschaft und Weltbild, Vienna: Verlag Hölder-Pichler-Tempsky, 1992.
[27] Putnam, H.: *Philosophical Papers V*, second edition, Cambridge: Cambridge

University Press, 1987.
[28] Radbruch, K.: *Die Mathematik in den Geisteswissenschaften*, Göttingen: Vandenhoeck und Ruprecht, 1987.
[29] Reichel, H.-C.: *Zum Realismusproblem mathematischer Begriffe*, in [22]
[30] Reichel, H.-C.: *Mathematik als Paradigma und Propädeutik der Philosophie*, Mathematische Semesterberichte 37 (1990), p. 180-215.
[31] Reichel, H.-C.: Mathematik und Weltbild seit Kurt Gödel, in [26], p. 9-29.
[32] Reichel, H.-C.: *Bildung durch Mathematik*, in H. Köhler und K. Rottel (eds.*): Mehr Allgemeinbildung im Mathermatikunterricht*. Bildungsraum Schule 2, Buxheim-Eichstätt: Polygon, 1993, p. 113 126.
[33] Steffen, K.: *Chaos, Fraktale und das Bild der Mathematik in der Öffentlichkeit*, Mitteilungen der Deutschen Mathematiker-Vereinigung, 1/94, p. 25-46.
[34] Stegmüller, W.: Hauptströmungen der Gegenwartsphilosophie, Stuttgart: Kröner, 1969.
[35] Tymoczko, T.: New Directions in the Philosophy of Mathematics, Basel, Boston: Birkhäuser, 1986.
[36] White, A.M. (ed.): *Essays in Humanistic Mathematics*, Washington: Mathematical Association of America, 1993
[37] Ziegler, R.: *Mathematik und Geisteswissenschaft*, Dornbach: Philosophischer Anthroposophischer Verlag, 1992.

Chapter II

Number and Randomness

Algorithmic Information Theory — New Results on the Foundations of Mathematics *

Gregory J. Chaitin
IBM Research Division, New York
E-mail: chaitin@watson.ibm.com

1. Introduction

It is a great pleasure for me to be speaking today here in Vienna. It's a particularly great pleasure for me to be here because Vienna is where the great work of Gödel and Boltzmann was done, and their work is a necessary prerequisite for my own ideas. Of course the connection with Gödel was explained in Prof. Reichel's beautiful lecture [1]. What may be a bit of a surprise is the name of Boltzmann. So let me talk a little bit about Boltzmann and the connection with my own work on randomness in mathematics.

You see, randomness in mathematics sounds impossible. If anything, mathematics is where there is least randomness, where there is most certainty and order and pattern and structure in ideas. Well, if you go back to Boltzmann's work, Boltzmann also put together two concepts which seem contradictory and invented an important new field, statistical mechanics. I remember as a student reading those two words "statistical mechanics," and thinking how is it possible - aren't these contradictory notions? Something mechanical is like a machine, predictable. What does statistics have to do with mechanics? These seem to be two widely separate ideas. Of course it took great intellectual courage on Boltzmann's part to apply statistical methods in mechanics, which he did with enormous success.

Statistical mechanics now is a fundamental part of physics. One forgets how controversial Boltzmann's ideas were when they were first proposed, and how courageous and imaginative he was. Boltzmann's work in many ways is closely connected to my work and to Gödel's work, which may be a little surprising. I'm trying to understand Gödel's great incompleteness theorem, I'm obsessed with that. I believe that the full meaning of Gödel's result can be obtained by taking Boltzmann's ideas and applying them to mathematics and to mathematical logic. In other words, I propose a thermodynamical approach, a statistical-mechanics approach, to understanding the foundations of mathematics, to understanding the limitations and possibilities of mathematical reasoning. Thermodynamics and statistical mechanics talk about what can be accomplished by machines, by heat engines, by steam engines, by physical systems. My approach to understanding the full implications of Gödel's work is mathematically analogous to the ideas of thermodynamics and Boltzmann and statistical mechanics. You might say, not completely seriously, that what I'm proposing is "thermodynamical epistemology!"

---

* Lecture given Tuesday 15 January 1991 in the Technical University of Vienna at a meeting on "Naturwissenschaft und Weltbild," immediately following a contribution by Prof. Hans-Christian Reichel on *Mathematik und Weltbild seit Kurt Gödel* (ref. 1). The lecture has been published in German language in H.C. Reichel and E. Prat de la Riba (Eds.) *Naturwissenschaft und Weltbild*, Hölder-Pichler-Tempsky, Vienna (1992), pp. 30-44.

2. Gödel's theorem

What led me to all this? Well, I was absolutely fascinated by Gödel's theorem. It seemed to me that this had to be the most profound result, the most mysterious result, in mathematics. And I think that a key question that one should ask when one reads Gödel's enormously surprising result, is, well, how seriously should one take it. It's clearly an enormously startling and unexpected result, but consider the mathematician working on normal mathematical questions. What is the meaning of Gödel for daily work in mathematics? That's the question I'd like to ask.

Gödel explicitly constructed an arithmetical assertion that is true but not provable within the system of *Principia Mathematica* of Russell and Whitehead. It's a very strange assertion. It's an enormously clever assertion: It says of itself, "I'm unprovable!" This is not the kind of assertion that one normally is interested in as a working mathematician. But of course a great part of Gödel's genius was to take such a bizarre question very seriously and also to clothe it as an arithmetical question. With the years this has led to the work on Hilbert's tenth problem, which is an even more straight-forward arithmetical incompleteness result inspired by Gödel's fundamental path-breaking work.

Let me make my question more explicit. There are many problems in the theory of numbers that are very simple to state. Are there an infinity of twin primes, primes that are two odd numbers separated by one even number? That question goes back a long way. A question which goes back to the ancient Greeks is, are there infinitely many even perfect numbers, and are there any odd perfect numbers? Is it possible that the reason that these results have not been proven is because they are unprovable from the usual axioms? Is the significance of Gödel's incompleteness theorem that these results, which no mathematician has been able to prove, but which they believe in, should be taken as new axioms? In other words, how pervasive, how common, is the incompleteness phenomenon?

If I have a mathematical conjecture or hypothesis, and I work for a week unsuccessfully trying to prove it, I certainly do not have the right to say, "Well obviously, invoking Gödel's incompleteness theorem, it's not my fault: Normal mathematical reasoning cannot prove this - we must add it as a new axiom!" This extreme clearly is *not* justified.

When Gödel produced his great work, many important mathematicians like Hermann Weyl and John von Neumann took it as a personal blow. Their faith in mathematical reasoning was severely questioned. Hermann Weyl said it had a negative effect on his enthusiasm for doing mathematics. Of course it takes enormous enthusiasm to do good research, because it's so difficult. With time, however, people have gone to the other extreme, saying that in practice incompleteness has nothing to do with normal, every-day mathematics. So I think it's a very serious question to ask, "How common is incompleteness and unprovability?" Is it a very bizarre pathological case, or is it pervasive and quite common? Because if it is, perhaps we *should* be doing mathematics quite differently.

One extreme would be experimental number theory, to do number theory as if it were physics, where one looks for conjectures by playing with prime numbers with a computer. For example, a physicist would say that the Riemann ζ hypothesis is amply justified by experiment, because many calculations have been done, and none contradicts it. It has to do with where the zeros of a function called the Riemann ζ function are. Up to now all the zeros are where Riemann said they were, on a certain line in the complex plane. This conjecture has

rich consequences. It explains a lot of empirically verified properties of the distribution of prime numbers. So it's a very useful conjecture. Now in physics, to go from Newtonian physics to relativity theory, to go from relativity theory to quantum mechanics, one adds new axioms. One needs new axioms to understand new fields of human experience.

In mathematics one doesn't normally think of doing this. But a physicist would say that the Riemann hypothesis should be taken as a new axiom because it's so rich and fertile in consequences. Of course, a physicist has to be prepared to throw away a theory and say that even though it looked good, in fact it's contradicted by further experience. Mathematicians don't like to be put in that position. These are very difficult questions: How should one do mathematics? Should number theory be considered an experimental science like physics? Or should we forget about Gödel's result in our everyday work as mathematicians? There are many possibilities in this spectrum.

3. A thermodynamical approach

I think these are very difficult questions. I think it will take many years and many people to understand this fully. But let me tell you my tentative conclusion based on my "thermodynamical" approach. It's really an information-theoretic approach: The work of Boltzmann on statistical mechanics is closely connected intellectually with the work of Shannon on information theory and with my own work on *algorithmic* information theory. There's a clear evolutionary history connecting these ideas.

My approach is to measure how much information there is in a set of axioms, to measure how much information there is in a theorem. In certain circumstances I can show that if you have five pounds of axioms, only five pounds, but here is a ten-pound theorem, well this theorem is too big, it weighs too much to get from only five pounds of axioms. Of course, I actually use an information-theoretic measure related to the Boltzmann entropy concept. Boltzmann would recognize some of the formulas in my papers, amazingly enough, because the interpretation is quite different: it involves computers and program size. But some of the formulas are identical. In fact, I like to use H for the same reason that Shannon used H, in honor of the Boltzmann H function, the H function dear to the heart of statistical physicists. (Of course, there's also a Hamiltonian H function, which is something else.)

The incompleteness phenomenon that Gödel discovered seems very natural from my information-theoretic point of view. You see, there is no self-reference. Gödel's incredibly clever proof skirts very close to paradox. I was fascinated by it. I was also very disturbed by it as a child when I started thinking about all this. If one measures information, then it seems natural to think, that if you want to get more information out, sometimes you have to put more information in. A physicist would say that it's natural that if one wants to encompass a wider range of mathematical experience, one needs to add additional axioms. To a physicist that doesn't seem outrageous. To a mathematician it's quite questionable and controversial.

So the point of view of algorithmic information theory suggests that what Gödel found is not an isolated singularity. The information - theoretic point of view suggests that Gödel's incompleteness phenomenon is very natural, pervasive and widespread. If this is true, perhaps we *should* be doing mathematics a little bit differently and a little bit more like physics is done. Physicists always seem very pleased when I say this, and mathematicians don't seem at all pleased. These are very difficult questions. I'm proposing this point of view, but by no means is it established. I think that one needs to study all this a lot more.

4. Gödel's incompleteness theorem and the Heisenberg uncertainty principle

In summary, let me tell a story from ten years ago, from 1979, which was the centenary of Einstein's birth. There were many meetings around the world celebrating this occasion. And at one of them in New York I met a well-known physicist, John Wheeler. I went up to Wheeler and I asked him, "Prof. Wheeler, do you think there's a connection between Gödel's incompleteness theorem and the Heisenberg uncertainty principle?" Actually, I'd heard that he *did,* so I asked him, "What connection do you think there is between Gödel's incompleteness theorem and Heisenberg's uncertainty principle?"

This is what Wheeler answered. He said, "Well, one day I was at the Institute for Advanced Study, and I went to Gödel's office, and there was Gödel...". I think Wheeler said that it was winter and Gödel had an electric heater and had his legs wrapped in a blanket. Wheeler said, "I went to Gödel, and I asked him, 'Prof. Gödel, what connection do you see between your incompleteness theorem and Heisenberg's uncertainty principle?' ". I believe that Wheeler exaggerated a little bit now. He said, "And Gödel got angry and threw me out of his office!" Wheeler blamed Einstein for this. He said that Einstein had brain-washed Gödel against quantum mechanics and against Heisenberg's uncertainty principle!

In print I recently saw a for-the-record version of this anecdote, [2] l which probably is closer to the truth but is less dramatic. It said, not that Wheeler was thrown out of Gödel's office, but that Gödel simply did not want to talk about it since he shared Einstein's disapproval of quantum mechanics and uncertainty in physics. Wheeler and Gödel then talked about other topics in the philosophy of physics, and about cosmology. There is some little-known work of Gödel connected with general relativity [3], some very interesting work, about universes where the past and the future is a loop, and you can travel into your past by going around. That's called a Gödel universe. It's a little-known piece of work that shows the stamp of Gödel's originality and profundity.

Okay, so what was the final conclusion of all this? I went up to Wheeler at this Einstein centenary meeting, and I asked him this question. Wheeler told me that he asked Gödel the same question, and Gödel didn't answer Wheeler's question, and Wheeler never answered my question! So I'm going to answer it! I'll tell you what I think the connection really is between Gödel's incompleteness theorem and Heisenberg's uncertainty principle. To answer the question I want to make it a broader question. I would like to tell you what I think the connection is between incompleteness and *physics.*

I think that at the deepest level the implication of Gödel's incompleteness theorem is as I said before that mathematics should be pursued more in the spirit of physics, that that's the connection. I see some negative reactions from the audience! Which doesn't surprise me! Of course this is a difficult question and it's quite controversial. But that's what my work using an information-theoretic approach to Gödel suggests to me.

Number theory has in fact been pursued to a certain extent in the spirit of an experimental science. One could almost imagine a journal of experimental number theory. For example, there are papers published by number theorists which are, mathematicians say, "modulo the Riemann hypothesis." That is to say, they're taking the Riemann hypothesis as an axiom, but instead of calling it a new axiom they're calling it a hypothesis. There are many examples of how this information-theoretic point of view yields incompleteness results. I think the most interesting one is my recent work on randomness in arithmetic, which I haven't really referred to yet in my talk.

5. Randomness in arithmetic [4]

A fundamental question that many of us wonder about, especially as teenagers—that's an age particularly well-suited for fundamental questions—is the question, "To what extent can the universe be comprehended by the human mind?" Is the universe ordered? Is there chaos and randomness? Are there limits in principle to what we will ever be able to understand? Hilbert stated very beautifully that he didn't believe that there were limits to what the human mind could accomplish in mathematics. He believed that every question could be resolved: either shown to be true or false. We might not be able to ever do it, but he believed that in principle it was possible. Any clear mathematical question would have a clear resolution via a mathematical proof. Of course, Gödel showed that this is not the case.

But it's really a more general question. Can the universe be comprehended, the physical universe as well as the universe of mathematical experience? That's a broader question. To what extent can all this be comprehended by the human mind? We know that it cannot be completely comprehended because of Gödel's work. But is there some way of getting a feeling for how much can be comprehended? Again it boils down to that.

When I was a student at the university, I totally believed in science. But my faith in science was tried by the work I had to do in experimental physics laboratories. The experiments were difficult. It was hard for me to get good results. I'm sure some of you are excellent experimentalists. There are people who have a natural talent for doing physics experiments like there are people who have a natural talent for growing flowers. But for me, the physics laboratory was a difficult experience and I began to marvel that scientists had been able to create modern science in spite of the fact that Nature does not give a clear answer to questions that we ask in the laboratory. It's very difficult to get a clear answer from Nature as to how the world works.

So I asked myself, what is it that is the most convincing evidence, in our normal daily experience, that the universe can be comprehended, that there is law and order and predictability rather than chaos and arbitrary things which cannot be predicted and cannot be comprehended? In my experience I would say that what most convinces me in science and predictability and the comprehensibility of the universe is, you'll laugh, *the computer!*

I've done calculations which involved billions (10^9) of successive operations each of which had to be accurately derived from the preceding ones. Billions of steps each of which depended on the preceding ones. I had ways of suspecting or predicting the final result or some characteristic of it, and it worked! It's really rather amazing. Of course, it doesn't always work, because the machine breaks down, or the programmer makes a mistake. But it works a lot of the time. And if one runs a program several times one usually gets the same answers. It's really amazing when one thinks how many steps the machine is doing and how this chain of causal events *is* predictable and *is* understandable. That's the job of the computer engineer, to find physical principles that are as predictable as possible, that give him a physical way to model the predictability of mathematics. Because computers are actually mathematical machines, that is what they really are. At least a mathematician might say that.

So the computer is a wonderful example of predictability and a case where the physical behavior of a big chunk of the universe is very understandable and very predictable and follows definite laws. I don't know the detailed laws of how a transistor works. But the overall behavior of the system is amazingly comprehensible and predictable. Otherwise one would not use computers. They would be absolutely useless.

Now it may seem strange that starting with the computer one can construct what I believe to be a very dramatic example of randomness. This is an idea I got from the work of Turing, which in turn was inspired by the work of Gödel, both of which of course were responses to questions that Hilbert asked. Turing asks, can one decide if a computer program will ever halt, if it will ever stop running? Turing took Cantor's diagonal argument from set theory and used it to show that there is no mechanical procedure for deciding if a computer program will ever halt. Well, if one makes a small change in this, in Turing's theorem that the halting *problem* is undecidable, one gets my result that the halting *probability* is algorithmically random or irreducible mathematical information. It's a mathematical pun!

The problem with this theorem is of course that in doing everyday mathematics one does not worry about halting probabilities or halting problems. So I had the same problem that Gödel had when he was thinking about mathematical assertions which assert of themselves that they're unprovable. My problem was how to take this bizarre notion of a halting probability and convert it into an arithmetical assertion. It turns out that one *can* do this: One can exhibit a way to toss a coin with whole numbers, with the integers, which are the bedrock of mathematics. I can show that in some areas of arithmetic there is complete randomness!

6. An example

Don't misunderstand. I was interviewed on a BBC TV program. A lot of people in England think I said that 2 + 2 is sometimes 4, sometimes 5, and sometimes 3, and they think it's very funny! When I say that there is randomness in arithmetic I'm certainly not saying that 2 + 2 is sometimes 3 and sometimes 5. It's not that kind of randomness. That is where mathematics is as certain and as black and white as possible, with none of the uncertainties of physics.

To get complete randomness takes two steps. The first step was really taken by Turing and is equivalent to Hilbert's tenth problem posed in 1900. One doesn't ask if 2 + 2 = 4 (we know the answer!). One asks if an algebraic equation involving only whole numbers, integers, has a solution or not. Matijasevic showed in 1970 that this problem, Hilbert's tenth problem, is equivalent to Turing's theorem that the halting problem is undecidable: Given a computer program one can construct a diophantine equation (an algebraic equation in whole numbers) that has a solution if and only if the given computer program halts. Conversely, given a diophantine equation, an algebraic equation involving only whole numbers, one can construct a computer program that halts if and only if the given diophantine equation has a solution.

This theorem was proven by Matijasevic in 1970, but intellectually it can be traced directly back to the 1931 incompleteness theorem of Gödel. There were a number of people involved in getting this dramatic 1970 result. It may be viewed as Gödel's original 1931 result restated in much simpler arithmetical terms. Unfortunately it turns out that this doesn't give complete randomness; it only gives partial randomness.

I'll now speak information-theoretically. Consider N cases of Hilbert's tenth problem. You ask, does the equation have a solution or not for N different equations? The worst would be if that were N bits of information, because each answer is independent. It turns out that it is only order of $\log_2 N$ bits of information, because the answers are not at all independent. That's very easy to see, but I can't go into it. So what does one do to get completely independent mathematical facts in elementary arithmetic? It's very simple. One goes a step farther: Instead of taking the halting problem and making it into the question of whether a diophantine

equation has a solution or not, one takes my halting probability, and makes it into the question of whether a diophantine equation has a finite or an infinite number of solutions.

If the equations are constructed properly, whether they have a finite or an infinite number of solutions is completely random. In fact, a single equation with a parameter will do. One takes the parameter to be 1, 2, 3, 4, 5, ... and one gets a series of derived equations from the original equation by fixing the value of the parameter. For each of these derived equations one asks: "Is there a finite or an infinite number of solutions?" I can construct this equation in such a way that the answers to this question are independent irreducible mathematical facts. So that is how you use arithmetic to toss a coin, to give you randomness. By the way, this equation turns out to be about 200 pages long and has 17,000 variables, and it's fun to calculate it. But one doesn't do it by hand! One does it with a computer. A computer is essential to be able to exhibit this equation. It is an infinite series of equations really, each of which has a different value of the parameter. We ask whether each of the equations has a finite or an infinite number of solutions. Exactly what does it mean to say that these are irreducible mathematical facts?

Well, how does one reduce mathematical facts? To *axioms*, to postulates! And the inverse of the reduction is to prove a theorem, I mean, to expand axioms into theorems. The traditional notion of mathematics is that a small finite set of axioms can give us all of mathematics, all mathematical truths. That was the pre-Gödel notion that Hilbert believed in. So in a sense what we're doing is we're compressing a lot of mathematical facts enormously, into a small set of axioms. Or actually, we're expanding a finite set of axioms into individual mathematical facts.

I'm asserting that I've constructed irreducible mathematical facts. What does this mean? It means that you cannot shrink them any more, you cannot squeeze them into axioms. In fact, that these are irreducible mathematical assertions means that essentially the only way to prove them is if we directly take each individual assertion that we wish to prove as an axiom! That's cheating! Yes, one can always prove an assertion by putting the assertion itself as a new axiom, but then we're not using reasoning. Picking new axioms is not deduction; it's the kind of thing that physicists worry about. It is surprising that we can have an infinite number of independent mathematical facts that can only be proven by taking them as axioms. But if we think about coin tossing this is not at all surprising. You see, the notion of independent coin tosses is *exactly* like that.

Each time one tosses a fair coin, whether the outcome of that particular toss is head or tails, tells us absolutely nothing about the outcome of any future toss, and absolutely nothing about the outcome of any previous toss. That's how casinos make money: There is no way to predict from what has happened at a roulette wheel what is going to happen. Well, there is if the roulette wheel isn't balanced, and of course the casino works hard to make sure that the roulette wheel is working properly.

Let's go back to coin tossing, to the notion that a series of tosses has no structure. Even if one knew all the even results, it wouldn't help us predict any of the odd results. Even if one knew the first thousand tosses, that wouldn't help us predict the thousand-first toss. Well, it's the same with using my equation to get randomness. Even if somehow one were told for all the even cases, whether there are a finite or an infinite number of solutions, this would be absolutely no help in getting the odd cases. Even if one were told the first thousand cases, whether there are a finite or an infinite number of solutions, it would be no help in getting the thousand-first case. In fact I don't see how one could ever get *any* of the cases. Because there

is absolutely no structure or pattern, and as I said these are irreducible mathematical facts. Essentially the only way to prove them is to directly assume them, which is not using reasoning at all.

So we've gone a long way in less than a hundred years: From Hilbert's conviction that every mathematical problem can be settled decisively by mathematical reasoning, to Gödel's surprising discovery that any finite set of axioms for elementary arithmetic is incomplete, to a new extreme, areas of arithmetic where reasoning is totally impotent and totally irrelevant. Some people were depressed by Gödel's result. You might say, "This is all rather upsetting; should I switch fields and stop studying mathematics?" I certainly don't think you should!

You see, even though there is no pattern or structure in the question of whether *individual* cases of my equation have a finite or an infinite number of solutions, one *can* deal with it statistically: It turns out that in half the cases there's a finite number of solutions, and in half the cases there's an infinite number of solutions. It's exactly like coin tosses, independent fair coin tosses. One can use statistical methods and prove theorems about the statistical patterns and properties of the answers to the question, which cannot be answered in each particular case, of whether there are a finite or an infinite number of solutions.

Let me repeat that the answers have a very simple statistical structure, that of independent tosses of a fair coin. So half the cases are heads and half are tails, one-fourth are a head followed by a head, one-fourth a head followed by a tail, one-fourth tail-head, one-fourth tail-tail, and so on for larger blocks and all the other statistical properties that one would like.

7. Concluding remarks

This kind of situation is not new; it's happened before, in physics. In quantum mechanics the Schrödinger equation shows this very clearly. The Schrödinger equation does not directly predict how a physical system will behave. The Schrödinger Ψ function is only a probability. We can solve the Schrödinger equation to determine the *probability* that a physical system will behave in a certain way. The equation does not tell us what the system will do, it tells us the probability that it will do certain things. In the 1920's and 1930's, this was very controversial, and Einstein hated it. He said, "God doesn't play dice" But as you all know and as Prof. Reichel [1] explained, in recent times this lack of predictability has spread outside quantum mechanics. It turns out that even classical physics, Newtonian physics, contains unpredictability and randomness.

This is the field of non-linear dynamics or “deterministic chaos”. It occurs in situations where small changes can produce big effects, in non-linear situations, very unstable situations, like the weather. It turns out that the weather is unpredictable, even in principle, as Prof. Casti discusses in his forthcoming book [5]. He studies the question of predictability and comprehensibility in a very broad context, including mathematics, the weather, and economics.

So it begins to look now like randomness is a unifying principle. We not only see it in quantum mechanics and classical physics, but even in pure mathematics, in elementary number theory. As I said before, I don't think that this should be viewed pessimistically. What it suggests to me, is that pure mathematics has much closer ties with physics than one suspected. Perhaps Plato's universe of mathematical ideas and the physical universe that we live in when we're not doing mathematics, perhaps these are closer to each other than has hitherto been su spected. Thank you.

[1] Hans-Christian Reichel, *Mathematik und Weltbild seit Kurt Gödel*, in H.C. Reichel and E.Prat de la Riba (Eds.), *Naturwissenschaft und Weltbild*,, Hölder-Pichler-Tempsky, Vienna (1992), pp.9-29, see also H.C. Reichel, this volume, chapter I.
[2] Jeremy Bernstein, Quantum Profiles, Princeton University Press, 1991, pp. 140-141.
[3] Kurt Gödel, *Collected works*, volume II, Publications 1938-1974, edited by S. Feferman et al., Oxford University Press, New York, 1990, pp. 189-216.
[4] See also: Gregory J. Chaitin, *Randomness in Arithmetic,* Scientific American (1988), pp. 52-57.
[5] John L. Casti, *Searching for Certainty—What Scientists Can Know About the Future,* William Morrow, New York, 1991

Suggestions for further reading:

- Gregory J. Chaitin*, Randomness in arithmetic and the decline and fall of reductionism in pure mathematics,* in J. Cornwell, *Nature's Imagination,* Oxford University Press, 1995.
- Gregory J. Chaitin, *The Berry paradox*, Complexity, Vol. 1, No. 1, 1995.
- Gregory J. Chaitin, *A new version of algorithmic information theory*, Complexity, in press.

Chapter III

Meaning, Reality and Algorithms: Implications of the Turing Theorem

Filippo Cacace
Università di Napoli "Federico II",Via Crispi, 112, 80121 - Napoli
e-mail: cacace@nadis.dis.unina.it

1 Introduction

The theorems of Kurt Gödel and Alan Turing are likely to be remembered as the two most important achievements in mathematics of the twentieth century. Their far-reaching consequences are valid for any formal system and in particular for arithmetic and any advanced mathematical theory. Their formulation in the thirties has changed completely the objectives of the research for the foundations of mathematical theories. The task Hilbert had indicated, namely the reduction of mathematics to logic through the process of axiomatisation and the formulation of a method to answer mechanically any mathematical question, has been proved to be unfeasible.

The most remarkable feature of the theorems of Gödel and Turing is their generality. From this point of view they can be considered similar to the principles of thermodynamics, and their practical consequences are not far from having the same relevance.

For all these reasons much discussion has been devoted to the interpretation of these theorems, that is, to understand their philosophical consequences. The aim of the present paper is to contribute to an answer of the three following questions:

(i) Which implications has the Turing theorem for the methods of computer science?
(ii) Why is the Turing theorem true, that is, where are the deep reasons that justify its validity?
(iii) Which consequences has the Turing theorem on the structure of reality?

In this paper we will not discuss, whether the human mind can be simulated more or less perfectly by machines. The answer to this question would require to answer three sub-questions: whether machines are constrained to have an algorithmical behaviour or not, whether human mind has an algorithmic behaviour or not, and, in the case machines and humans have the same behaviour, whether they have the same computing power or not. Our task is different: we want to investigate the limitations of "exact'', that is, "formal'' reasoning, on which modern science has been based. It is doubtful that computers will ever have the intelligence of human beings. Sure instead, is that in order to equal them, they have to give up their "mechanical'' features.

The paper is organised as follows. In Section 2 we argue that computer science is modelled on the so-called methodological reductionism. This is an epistemological scheme already used in physics and in most of the other experimental sciences. In Section 3 we introduce the Turing theorem in order to show its implications on methodological reductionism. In Section 4 we use the Turing theorem and the Church thesis to deal with the problem of *recursive decidability* for general formal systems, that is, the problem of determining mechanically whether a certain proposition is correct in a given formal system. We analyse the reasons that make formal systems recursively undecidable. In Section 5 we describe some philosophical implications of the Turing theorem.

2 Reductionism

2.1 Methodological reductionism: from Newton to artificial intelligence

The method on which most of modern science is based since the seventeenth century can be summarised as follows:

- From reality one extracts some quantitative aspects of certain objects (for example speed, position, or weight), which are relevant to solve a certain problem.
- From the observations one extrapolates some formal relation between the quantitative aspects. Relations are usually expressed in a mathematical formalism.
- By manipulating these relations one obtains new results for the quantitative aspects that then are transferred back to reality. In this way the original problem can be solved or one is able to follow the evolution of the system knowing its configuration at a certain moment.

This approach is visualised by Figure 1.

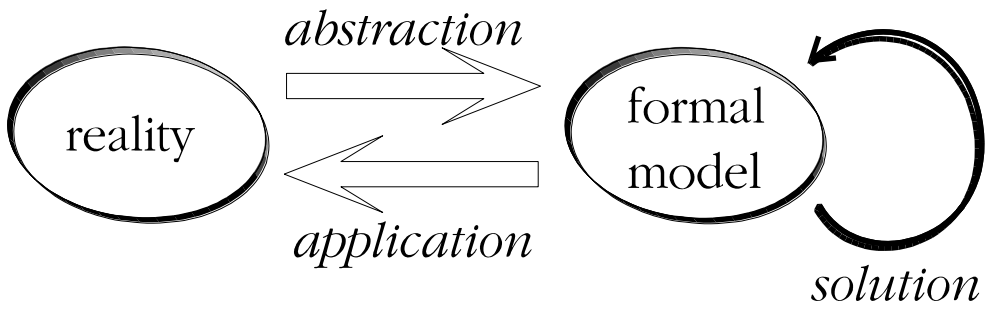

Fig.1 - Methodological reductionism

This methodological pattern has found wide applications in physics, chemistry and many other different areas of modern science. It has been so successful that it generated a philosophical vision of sciences known as methodological reductionism [1]. It consists of the claim that the "reduction'' of the reality to relations between its measurable aspects is the only correct, useful and meaningful way of understanding and analysing it.

Due to its success, the process of "reduction'' has become a general scheme for dealing with any kind of problems, even in areas different from experimental sciences. The method can be re-formulated as follows:

- Given a certain portion of reality and a problem to solve, one starts by extracting a formal representation of reality, that is appropriate for solving the particular problem.
- One solves the "formalised version'' of the problem, by exploiting the rules that the formal language provides.
- One then goes back to reality in order to apply the formal solution to the real problem one intends to solve.

Consider for example the problem, how one can determine the evolution of the population in a particular region. One starts by selecting the parameters of interest, like age, sex, etc. Thereafter one studies the historical evolution of these parameters in order to find, by statistical means, the parameters of the stochastic process that represent the evolution of the population. Based on this statistical model one finally can formulate future prospects.

The same approach can also be used in other areas. For example, if one can describe a language by means of a set of rules that produce all the meaningful phrases of the language, one may use the same set of rules to give an answer to each question formulated in that language.

The method of methodological reductionism is based on an "isomorphism'' between a certain aspect of the reality and the formal (abstract) model that we have created. In the example given above, the real population and the stochastic process are isomorphic with respect to the variables we want to compute. Because of this isomorphism the solution of the formal model can be considered as a solution for the real problem.

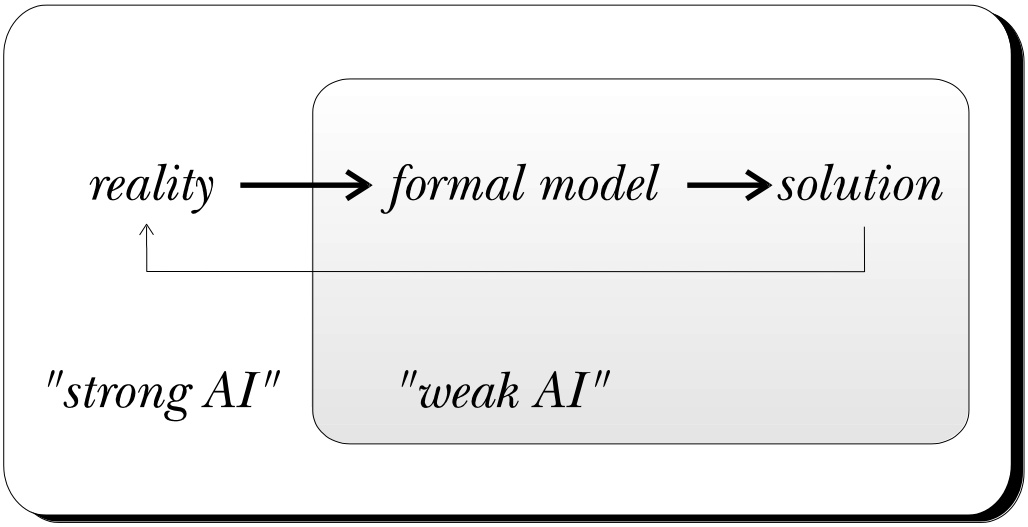

Fig. 2 - Reality and abstract models

Considering Artificial Intelligence (AI) one could represent the method by the scheme depicted in Figure 2. With the introduction of computers, that are able to perform symbolic as well as numerical computations, the supporters of AI claim that the mechanisation of this method would generate machines with an autonomous capability of solving problems.

In particular, given a certain problem the machine should be able to:

(I) generate a formal description of the problem in a suitable internal language;

(ii) manipulate this description in order to produce an algorithm that represents the solution to the problem.

We define this attempt as *strong AI*. "Intelligent'' computers, according to the vision of strong AI, should be able to extract from reality its underlying abstract model and then use it to solve different classes of problems.

When the first task is not mechanised and the formal model is the input for the computer, we have the so-called *weak AI* approach. In this case machines have only the task of manipulating the formal model and find out the correct algorithm to solve a particular problem.

2.2 Reductionism and computer science

It is important to emphasise that reductionism has wide diffusion in the whole area of Computer Science. In particular, research in the areas of Software Engineering and Database Systems, that are the two most relevant parts of Computer Science from a commercial point of view, is heavily influenced by the weak AI approach.

Research in the areas that require specification languages, logic languages and declarative languages is explicitly (but perhaps not consciously) aiming at the following. One attempts to build systems that accept a formal description of some part of the reality in a given language and produce a solving algorithm as their output.

To illustrate the point, one may consider declarative languages. The aim of declarative programming is to provide the user with languages that allow him to describe the problem he wants to solve, without the need of specifying *how* to solve it. The actual solution is produced by the compiler, or interpreter, of the language.

Consider the classical example of determining the set of descendants given the parenthood relationship. The "procedural'' solution of such a problem could consist in the following algorithm (let *P* be the person whose descendants we want to compute):

1. Assign the empty value to S_P, the set of descendants of *P*.
2. Perform the union between S_P and the set of children of *P* and assign the result to S_P.
3. Compute the set *C* of children of persons contained in S_P.
4. If C is a subset of S_P, then go to 7.
5. Perform the union between *C* and S_P and assign the result to S_P.
6. Go to 3.
7. The algorithm terminates and S_P is the result.

The "declarative" solution of the same problem can be expressed through the following pair of logic clauses (*parent* is the relation containing pairs of parent-child):

ancestor(X, Y) ✶ *parent(X, Y)*
ancestor(X, Z) ✶ *parent(X, Y), ancestor(Y, Z)*

Descendants of P are obtained by issuing the query*:-ancestor (P, X)* against the above pair of logic clauses. This pair constitutes a program written in a logic language (in the example, *Prolog* [3]). The program does not contain a solution to the problem, but only its logical description. The evaluation of the rules in order to produce a computation and the answer to the queries is performed by the Prolog interpreter.

The same approach has been followed in the language SQL [13], a *de facto* standard for commercial Relational Databases. SQL has not the computational power of Prolog (it cannot express *recursion*), but it is even more declarative. SQL is used to express queries on a database represented as a set of tables. The query "determine the set of employees who earn more than 5000 US dollars'' is expressed as:

```
SELECT name FROM employees WHERE salary > 5000
```

The SQL compiler and optimiser deals with these queries in order to produce an "execution plan''. In this plan a sequence of operations is performed to access files and to manipulate data that implements the correct answer to the query.

Also in the area of Software Engineering many efforts have been focused on the attempt of building specification languages (usually using logic or algebraic styles). These

languages are used to express requirements for a certain problem. The task of implementing these requirements in an executable program is left to the system, thus relieving the application programmer from most of the work and increasing the productivity of the software factory.

3 The Turing theorem

In this paper we will not analyse the proof of the Turing theorem [10,11] but will point out some of its consequences. We just recall that Turing introduced the formalism of the Turing machines, as a mean to express algorithms. Actually, any known mechanical procedure of solution can be implemented on a Turing machine. The hypothesis that Turing machines can express exactly the class of what we denote by the word "algorithm'' is known as the *Church thesis*.

On a given input, a given Turing machine might halt to generate a result or continue to compute forever. The simplest formulation of the Turing theorem is perhaps the following: **there is no algorithm to decide whether the generic Turing machine will halt or not on a generic input**.

3.1 Consequences of the Turing theorem

The problem of the halt for Turing machines is a perfectly defined mathematical problem. The Turing theorem therefore states that there are mathematical questions that admit no mechanical solution, thus giving a negative answer to the *Entscheidungsproblem* of Hilbert [12]. The evidence that there are problems that cannot be solved by means of algorithms has a particular relevance of its own.

However, there is more to say on the "technical'' consequences of the Turing theorem. First of all, it is important to note that it does not imply that there are certain problems that cannot be solved at all. It simply states that there are problems that have no *mechanical* (in the sense of *algorithmic*) solution.

The most important consequence is based on the observation that the problem of the halt for generic Turing machines is equivalent to the solution of the generic arithmetic problem. That is, if we produce an algorithm that can determine whether a generic Turing machine will halt or not, we could modify it in order to answer to generic questions about, for example, natural numbers.

Therefore from the Turing theorem follows *that there is no algorithmic way to solve a generic mathematical question* (that is, demonstrate a property, a theorem, etc.). The answer to each mathematical question requires "its own'' solution that cannot be generated through a general algorithm, procedure or program.

3.2 Criticising the reductionism

In the previous section we explained that both the "strong'' and "weak'' AI are based on the reductionist process. This process assumes that is possible to abstract from reality a formal model *isomorphic* to it, and that thereafter the "real'' problem can be solved within the formal scheme.

Unfortunately as we have just shown, the Turing theorem states that there is no algorithmic method for solving generic problems in the formal scheme. Even if there are classes of problems that are computable, the strength of the Turing theorem is to show that also for simple classes, like that of arithmetic properties, the uncomputability problem arises. This implies that the second pass of the reductionist scheme is not possible. There is, namely, no general way to solve the formalised version of a problem by exploiting the rules that the formal system gives us. In other words, even if we find a formal system isomorphic to a part of reality, this would not allow us to find a general solution to the class of problems that can be formulated for that part of reality. This does not imply that no problem can be solved by mechanical means, but that there will always be a great majority of problems for which we have no solution.

It is now evident why the approach of both strong and weak AI is unfeasible, since the two have in common the assumption that is in principle possible to solve any class of formalised problems. In the next section, we will consider the reasons of this fact analysing the relationship between algorithms and meaning. Incidentally, this will provide us more arguments against the strong AI approach.

4. Meaning and algorithms

As we pointed out in the previous section, no formal definition can be given for the notion of an algorithm. However, if we accept the Church Thesis, as we do throughout this paper[23], the set of algorithms becomes equivalent to the set of different programs for Turing machines.

In order to highlight the relationship between meaning and algorithms we start this section by considering formal systems and the Gödel theorem. Our conclusion will be that the notion of meaning is important in order to distinguish between *demonstrability* and *truth*. We then show that even for formal systems where demonstrability and truth coincide, it is in general impossible to find an algorithmic way to determine when a proposition is provable. Thereafter we present arguments that computability is a stronger property than demonstrability. We show that it is impossible to solve classes of problems because no algorithm can give a meaning to the formal description of a problem. Finally we conclude that in order to solve problems it is necessary to have an understanding of their meaning. This understanding, however, can not be obtained by algorithmic means.

4.1 Formal systems, proofs and truth

In the following we focus our attention on abstract problems that have no immediate reference to an external reality. In particular, we consider systems of logical axioms and problems related to the question, whether a certain proposition is true or not. In this case, an algorithm is a method that has a logic formula as its input and produces on the basis of the axioms a demonstration or a refutation of that logic formula.

Any logical theory (for example, formal theory of numbers, set theory, etc.) is based on a set of axioms, from which *theorems* are derived by means of *inference rules*. Certain axioms are common to any logic theory: they are called *logical axioms*. To the contrary,

[23]The Church Thesis can be formulated in different equivalent ways. One possible formulation is that the formalised notion of "recursive functions'' corresponds to the intuitive notion of a problem that is actually solvable. Since the expressive power of Turing machines and recursive functions is the same, this formulation of the Church Thesis implies that actually solvable problems are those which are algorithmically computable by Turing machines. In the sequel we use indifferently the terms "recursive" and "algorithmic"

axioms that are specific for the particular theory are called *proper axioms*. From logical axioms one can derive only logical truths, that is, truths that are independent of the particular theory. The logical theory based only on logical axioms (and therefore containing no proper axioms) is called *First Order Calculus* [10].

An interpretation of a logical theory results in an attribution of a *domain* for the variables, of *relations* for the predicative symbols, and of *operations* for the function symbols of the theory. For example, a certain interpretation for the formal theory of numbers uses the set ***N*** of natural numbers for the variables, gives the usual meaning to the predicative symbols =, <, etc. and the function symbols +, -, *, etc. Different interpretations can be given to the same theory. The standard interpretation of a logical theory is the intended interpretation for that theory (for example, arithmetics in the case of the formal theory of numbers). The intuitive notion of *truth* can be rigorously based on interpretations. A logical formula is called *logically valid* when it is true in any interpretation. A formula is called *satisfiable* when it is true in at least one interpretation. Logically valid formulas are similar to *tautologies* of propositional calculus.

A **completeness theorem** due to Gödel [6] states that in every First Order Calculus *the theorems are exactly the logically valid logical formulas*. That is, the notion of demonstrability (derivability from axioms) and truth (satisfaction of the interpretation) *coincide*. Truth can be reached through both, syntactic means, starting from axioms, and semantic means, starting from interpretations.

It is well known that this situation does not hold anymore for general logical theories. For example, for a formal system *S* that formalises the theory of numbers, the **Gödel uncompleteness theorem** [7] states that *if S is consistent, a logical formula exists in S that is true but is not provable nor refutable*. The Gödel theorem holds for any recursively axiomatisable extension of *S*, and therefore for any "useful'' formal system.

It should be noted that the Gödel theorem marks a difference between the notion of demonstrability and truth. It shows that for sufficiently complex formal systems there is a logical formula that is not provable (nor refutable) starting from the axioms, but *is true within the* standard *interpretation* of the formal system. In the case of *S*, this means that there are arithmetic properties that are true but are not provable from the axioms. That means, interpretation is needed to establish the truth of a logical formula in a logical theory. To put things differently, we could say that in First Order Calculus the interpretation is completely contained in the formalisation, while for generic logical theories a reference to external reality is needed. Therefore, aspects of reality on which logical theories are modelled (that is, arithmetics, mathematics, etc.) are not amenable to be completely formalised.

Let us call *external meaning* the portion of the standard interpretation of a logical theory that is not derivable from the axioms. The external meaning is the un-formalised part of the standard interpretation.

In the next subsections we consider the problem of relating computability, demonstrability and meaning.

4.2 The Church theorem

For First Order Calculus systems the standard interpretation is the set of logically valid formulas, and since any theorem in First Order Calculus is a logically valid formula, (and vice-versa) there is no external meaning: all the logical formulas derivable from axioms are true in any interpretation.

In this purely "logical'' context we could hope that it is possible to determine by means of an algorithm, whether a logical formula is logically valid, that is, whether it is a theorem of First Order Calculus. Unfortunately this is not feasible, as the following argument shows.

The **Church Theorem** [2] states that First Order Calculus is recursively undecidable. Or in the formulation of Davis et al. [4]: There are axiomatisable theories that are not recursive. If we accept the Church Thesis, one has to conclude that First Order Calculus is actually undecidable, that is, *it is not possible to determine algorithmically whether a logical formula is logically valid*. One should bear in mind that logically valid formulas are the set of purely logical propositions that are true in any theory.

This leads to the following conclusions:

- Non computable problems arise not only in mathematics but even in purely logic theories like First Order Calculus. This is somehow surprising, since First Order Calculus can be formalised by a complete theory, while mathematics can not.
- Computability is a weaker notion than demonstrability. We could establish a sort of hierarchy of classes of questions: there are classes of questions that admit a yes/no answer but cannot be proved; classes of questions that can be proved but whose demonstrations is not computable through a general algorithm; classes of questions whose demonstrations are computable through a general algorithm.
- Non computability is not directly related to the presence of an external meaning. Even theories that are complete, and therefore do not need an interpretation to determine the truth of their axioms, are not decidable through the use of algorithms.

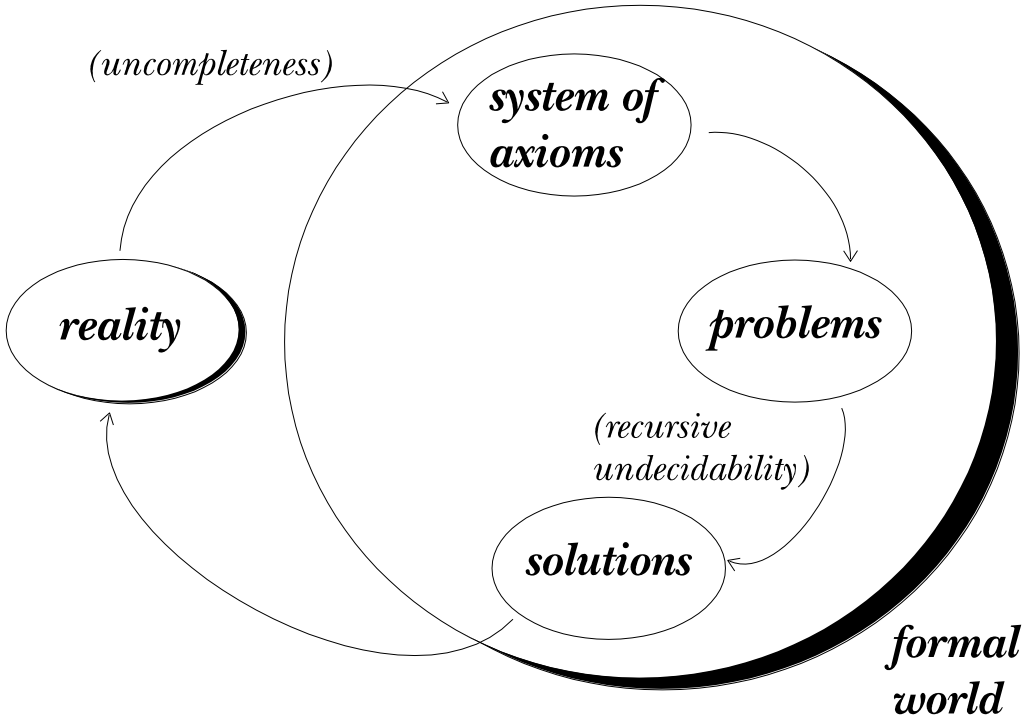

Fig. 3 - Uncompleteness and undecidability for formal theories

The consequences of the Gödel and Church theorems are illustrated in Figure 3. In order to solve problems in a completely mechanical way, it is necessary to follow the cyclic pattern as depicted in the figure. One starts from 'reality' (that is, a given portion of reality, for example mathematics), goes through axioms and the formal version of problems and the

production of demonstrations. Unfortunately, the two most critical edges of the pattern raise difficulties that cannot be circumvented. The axiomatisation of reality causes a loss of completeness (due to the Gödel theorem), and the formalisation of problems makes them recursively (that is, algorithmically) undecidable (due to the Church theorem).

One should note that theories that are both complete and decidable do exist. The monadic predicative calculus is a good example of a theory for which an algorithm exists that allows to determine whether a logical formula is a theorem. It is a subset of First Order Calculus. Sets of Horn clauses are another example of logic theories (subsets of logical theory) for which a solving procedure exists [9]. In particular, it is often possible to determine a subset of a theory of interest that is algorithmically decidable.

However, for theories that are sufficiently general to be interesting it is simply not possible to follow the pattern by any mechanical means.

4.3 Computability and meaning

Why can a formal system, in spite of being complete, be recursively undecidable? How is it possible that in a theory where a proof exists for each true proposition, this demonstration can not achieved by algorithmic means?

In our view the deep reason behind the Turing and Church theorems lies, as for the Gödel theorem, in the concept of ‘meaning’. Schematically, the situation can be depicted as in Figure 4.

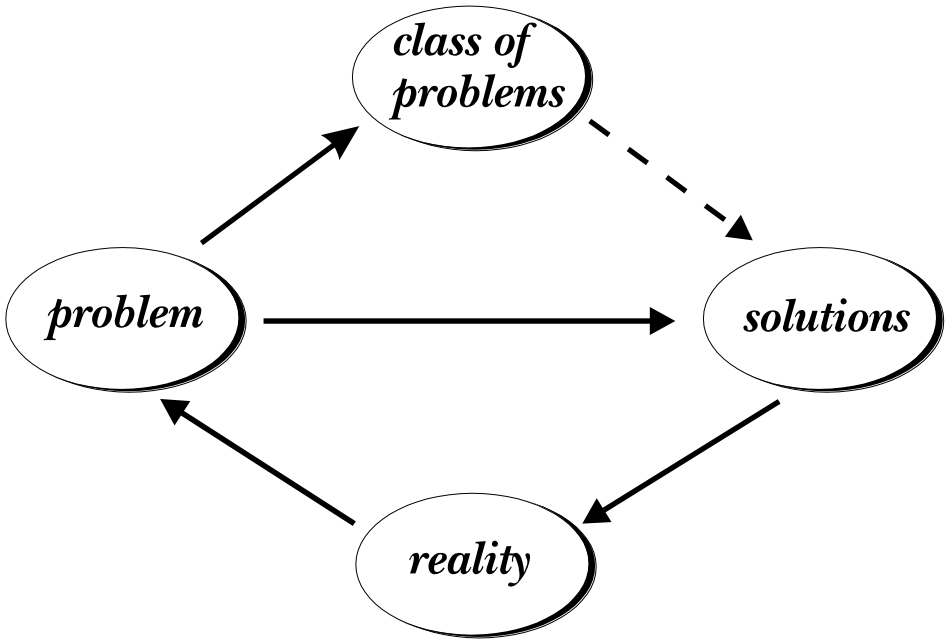

Fig. 4 - Algorithmically solving a class of problems

For each solvable problem a solution exists. In order to actually solve a class of problems we have to formalise the problems and the solutions. Despite the fact, however, that it is in principle possible to go from each problem to its solution, it is in many cases impossible to go from the class of problems and the formalised version of a problem to its solution.

To understand the issue consider the famous last theorem of Fermat, that states that for $n \geq 3$ it is impossible to find three integers x, y, z such that $x^n + y^n = z^n$. This problem (or class of problems) can be formalised by the following logical formula:

$$\forall\, x\, y\, z\, n \;\; ((\, n > 2\,) \;\; (x^n + y^n \neq z^n\,)\,)$$

Since the logical formula is *closed*, it is certainly either true or false. Moreover, Fermat's last theorem can now be considered demonstrated.

To the contrary, it is *in principle impossible* to find out how to demonstrate a generic property of natural numbers. Algorithmic methods are suited for specific problems, but not for classes of problems. An algorithm for a class of problem should accept as it's input the description of the problem. The deep reason behind the impossibility of the existence of such an algorithm seems to be the fact that it should be able to give a meaning to this description in order to solve the problem.

In other words, when the class of problems becomes too broad the "generic'' algorithm is not able to give a meaning to the description of the problem. One may say that the formal description is not sufficient to determine completely the problem, and that a reference to external reality is needed. This is similar to the problem that originates the Gödel Theorem. When a theory becomes too powerful, it is not possible to formalise it completely by means of a set of axioms, and a reference to external reality (the *interpretation*) is needed in order to give a truth value to logic formulas: in both cases, the negative result seems to be generated from the impossibility to accommodate within the formal scheme any non-trivial portion of reality.

5 Reality and intelligibility

In this section we focus our attention on some philosophical consequences of the Gödel and Turing theorems.

5.1 Infinite Intelligibility

The first conclusion one may draw is that truth can be known only in part; the domain of the problems is infinite, but the domain of solutions increases only at a finite speed. The metaphor we propose is that of an infinite tape we go across at a finite speed. There is no way to access randomly a given position of the tape, but it must be followed from the beginning. The situation is analogous with what happens for space travels. Since the speed has a maximum value (the speed of light) and distances are virtually unlimited, there are places that never will be reached. There will be truths that one never will know.

The important point is that reality contains infinitely many independent truths. Independent means that there is no method allowing us to discover all the truths.

5.2 Limits of knowledge

The epistemological implication of Turing theorem is that man should recognise the limits of his possibility to know the physical reality. The positivistic myth [8] of a complete power on nature through experimental science is not feasible even in the line of principle: "God was always invented to explain mystery. God was always invented to explain those things that you do not understand. Now when you finally discover how something works, you get some laws

which you're taking away from God; you don't need him anymore. *But you need him for the other mysteries*.'' (R. Feynman).

The pretended use of science as the final answer to most basic problems rises serious problems in the light of the fundamental limits in knowledge. The modern illusion that the complete knowledge of the physical world will explain man the why of its existence reveals itself as an illusion.

5.3 The causes of truth

If no formal system of axioms can represent completely even a small portion of reality, one is led to conclude that there is no immanent cause of truth in the world other than the reality itself. Reality obviously cannot justify its intelligible content.

Therefore one is led to the situation in which reality is the only source of truth, but it does not contain the causes of this truth. So, either the intelligibility of reality has no cause at all, or its cause is outside the (sensible) universe.

6 Conclusions

Two different approaches to the interpretation of results of modern science exist.

On one side there are those who consider scientific developments philosophically irrelevant, and philosophical interference with science an irritating abuse. They argue that science and philosophy are two domains whose independence must be vigorously protected. They emphasise that most of classical metaphysics was derived in spite of the wrong cosmological assumptions prevailing at that time, while the born of modern science required cutting its links with philosophy. Gilson [5], for instance, who considered the different philosophical positions of famous modern physicists, states that there is a constant confusion of language when one attempts to use arguments of a certain epistemological domain to deduce conclusions in a different context.

On the other side, a deeper knowledge of the laws of reality is a better starting point for philosophy. Modern minds are more inclined to trust a philosophical truth that has an evident experimental counterpart. This is not a confusion of domains, but, more simply a philosophical approach based on technical results instead exclusively on common evidence. This approach is difficult, since it requires a competence in different fields. It may, however, be the only way to overcome the conviction that modern science and philosophy are in conflict with each other.

Our analysis of the Turing theorem reveals something that we already intuitively knew. Not only our explanations of reality do not give us the justification of its being but also no formal, or mechanical, scheme is a good way to extend our understanding of reality. They are just "short-hand notations'' reproducing properties of parts of the real world, but neither the immediate nor the ultimate cause of truth.

It is however important, in our view, to see how this intuition of common sense is confirmed with great evidence by reflecting on results of modern logic. After results like the Turing theorem it is still possible, but less easy, to draw erroneous philosophical conclusions from scientific advances, because the nature and the limits of mechanical reasoning can now be clearly seen. A more serious awareness of the philosophical implications of technical results by the scientific community seems to be of the greatest relevance for the development not only of philosophy but also of science.

References

[1] Fortunato Tito Arecchi, and Iva Arecchi: *I simboli e la realtà*, Jaka Book, Milano, 1990.

[2] Alonzo Church: *A note on the Entscheidungsproblem*, JSC, **1**, 40-41; Correction, *ibid*., 101-102, 1936.

[3] W.F. Clocksin, C.S. Mellish: *Programming in Prolog*, Springer Verlag, Heidelberg, 1981.

[4] Martin D. Davis, Ron Sigal, and Elaine J. Weyuker: *Computability, Complexity, and Languages,* Academic Press, San Diego, 2nd edition (1994), p. 415

[5] Etienne Gilson: *God and Philosophy,* Yale University Press, New Haven, 1941.

[6] Kurt Gödel: *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I,* Mh. Math. Phys., **38**, 173-198, 1931.

[8] Jacques Laeuffer: *Scientism and scientific knowledge of things and God,* this volume, chapter XII.

[7] Kurt Gödel: *On undecidable Propositions of Formal Mathematical Systems*, Princeton, 1934.

[9] John Lloyd: *Foundations of Logic Programming*, Springer-Verlag, 1984.

[10] Elliott Mendelson: *Introduction to Mathematical Logic*, D. Van Nostrand Company, Princeton, NJ, 1964.

[11] Roger Penrose: *The Emperors' New Mind*, Oxford University Press, 1987.

[12] Alan Turing: *On Computable Numbers with an Application to the Entscheidungsproblem*, Proc. London Math. Soc., **42**, 230-65, **43**, 544-46, 1936-37.

[13] John D. Ullman: *Principles of Database and Knowledge Base Systems*, Computer Science Press, Potomac, MD, 1988.

Chapter IV

The Limits of Mathematical Reasoning
In arithmetic there will always be unsolved solvable problems

Antoine Suarez
Institute for Interdisciplinary Studies, P.O. Box 304, 8044 Zurich, Switzerland
E-mail: suarez@leman.ch

Abstract
The *Undecidability* theorem is presented and some related philosophical issues are discussed. In particular it is argued that the solutions to *today's unsolved solvable* problems, since not contained in any human intelligence, have to be contained in another intelligent principle, somewhere else.

1. Introduction
In a celebrated contribution on „Mathematical problems", at the Second International Mathematicians Congress (Paris, 8. August 1900), David Hilbert speaks about the conviction that our mind is capable of answering all questions it asks. Hilbert raises the question of whether this axiom is characteristic only of mathematical thinking, or whether it is a general and essential law of our mind. He shares the conviction that there is an answer to every mathematical question, and that we can find it through pure thinking: „in Mathematics there is no Ignorabimus." [24]

Assuming there is an answer to every mathematical problem, we are lead to the question: Will we have in the future a universal method which allows us to solve any mathematical problem simply by calculation? Will the day come in which the solution to a given problem becomes only a matter of computing time? This question is the so called *Entscheidungsproblem* [25].

Work by K. Gödel (1931), A. Turing (1936-7), A. Church (1935-6), E. Post (1936) and others, led to the astonishing result that the answer to Hilbert's *Entscheidungsproblem* is 'no'. We will never have a universal computer program capable of answering any well-posed question on numbers [26]. This result is usually referred to as the *Undecidability* theorem.

In Section 2 of this article we introduce the idea of an *unsolved solvable problem*. In Section 3 we state that in arithmetic there always will be an *unsolved solvable problem* (*Undecidability* theorem), and give a particularly simple proof of this theorem in an appendix. In Sections 4 and 5 we discuss some philosophical consequences of this theorem.

2. Mathematical unsolved solvable problems
Consider the set of natural numbers, *N*: {0, 1, 2, 3,...*n*,...}. Different properties, such as *even* number, *odd* number, *prime* number, etc. can be defined by a *finite* number of instructions. For instance:

Even numbers

A natural number is *even* if it is divisible by 2. This definition can be translated into a computer program, capable of deciding whether a given number is even or not. Consequently,

---

24. D. Hilbert, *Gesammelte Abhandlungen*, III. Band, Berlin: Springer, 1935, p.297-298.
25. D. Hilbert and W. Ackermann, *Grundzüge der theoretischen Logik*, 3.§1, Berlin: Springer, 1928.
26. All these basic papers are collected in M. Davis (ed.), *The Undecidable*, New York: Raven Press, 1965.

a computer working with this program and receiving as input the set *N* will yield as output a set containing only even numbers:

$$\{0, 2, 4, 6, 8, 10, ...\}$$

We can now ask, whether the computer will ever stop printing a new even number, i.e. whether there is a number *n* such that there is no even number greater than *n*. The answer is obviously 'no', because for each given *n*, the number 2*n* is both even and greater than *n*.

The precise meaning of the statement that there are infinitely many numbers having a certain property *P*, is that for each arbitrarily large *n*, it is always possible to find a number greater than *n* with property *P*.

Prime numbers

A *prime* number is a natural number, other than 0 and 1, that is divisible only by 1 and itself. Since the property can be defined in terms of a finite number of instructions it can also be translated into a computer program, producing as output a series of prime numbers:

$$\{2, 3, 5, 7, 11, 13 ...\}$$

The question whether after a certain time the computer will stop printing prime numbers (i.e. the question whether there is a largest prime number) is less trivial than the question about how many even numbers there are. However the answer is known since the Greek mathematician Euclid:

Assume the computer has printed the prime numbers:

$$2, 3, 5, 7, 11, 13, ... p$$

and consider the number:

$$p' = 2\times3\times5\times7\times11\times13\times...\times p + 1$$

None of the printed numbers is a divisor of *p'*. Consequently, either there is a prime number greater than *p* that is divisor of *p'*, or *p'* itself is a prime number. In any case the computer will print another prime number. Therefore, we conclude that there is no largest prime number, or in other words that there are infinitely many of them.

The question of whether there are infinitely many prime numbers is now a solved problem. Before Euclid it was a *solvable unsolved problem*.

Perfect numbers

A number is *perfect*, if it is the sum of all its factors except itself. For example the numbers 6, 28, 496 are perfect:

$$6 = 1+2+3$$

$$28 = 1+2+4+7+14$$

$$496 = 1+2+4+8+16+31+62+124+248$$

Now consider the questions:

$$Q_1: \text{Is 8128 a perfect number?}$$

Q_2: Is there a largest perfect number?

Someone who is not familiar with the topic of perfect numbers will not be able to give an immediate answer, to either Q_1 or Q_2, at the very moment he understands the questions. Nevertheless everybody feels he has the capacity to give the answer to Q_1. Even if I am not yet aware of the answer, the answer exists, it is either 'yes' or 'no', and I know how to find it: first search for all factors of 8128, second make the sum of them, third compare the sum to the number. Let us calculate:

$$1+2+4+8+16+32+64+127+254+508+1016+2032+4064 = 8128$$

Now I am aware that the answer to Q_1 is YES. Since I had the method to calculate it the answer to Q_1 not only existed, but it was implicitly a content of my mind, even before I did the calculation. It was *a priori* in my modes of thinking.

Now consider question Q_2. Once again, since the definition of perfect number is equivalent to a finite number of instructions, it can be translated into a computer program printing perfect numbers in increasing order:

$$\{6, 28, 496, 8128, 35550336, ... \}$$

At the moment, however, nobody can say whether this set has a largest element or not, nobody knows, neither the answer to Q_2, nor the method to find it. As Hilbert suggests, both the answer and the method exist. The answer is either 'yes' or 'no'. But unlike the question about the largest prime number, nobody has yet found a method to find the answer. This answer is knowable, but unknown, and Q_2 is to date an *unsolved, solvable problem*.

Twin primes

Two natural numbers *n* and *n*+2, both primes, are called *twin primes*. The pairs: (3,5), (5,7), (11,13), (17,19), (29,31), (41,43) are such twin primes. Accordingly, a computer program can be written that prints, in increasing order, all primes which have a prime twin:

$$\{3, 5, 7, 11, 13, 17, 19, 29, 31, 41, 43, ... \}$$

The question whether this set has a finite number of elements, cannot yet be answered.

Goldbach's conjecture

Christian Goldbach in 1742 put forward the conjecture that each even number, greater than 2, is the sum of two prime numbers:

$$4 = 2 + 2$$
$$6 = 3 + 3$$
$$8 = 3 + 5$$
$$10 = 5 + 5$$
$$12 = 5 + 7$$

This conjecture has not yet been proven or refuted.

Suppose a computer is programmed to print all even numbers till an even number appears that is not the sum of two prime numbers, and to stop, if such an even number appears. Up to now, we cannot say whether this computer will ever stop.

Fermat's prime numbers

The prime numbers of the form: $2^{2^n}+1$, are called *Fermat's prime numbers*. Up to now, computers have printed only 5 such numbers:

$$\{3, 5, 17, 257, 65537, ?\}$$

Nobody knows whether there are more of them.

3. In arithmetic there will always be unsolved solvable problems

Unsolved solvable problems challenge our creativity and are highly stimulating for research. However they let us feel the mystery, and it is tempting to ask for a method, as Hilbert did, that would allow us to solve any such problem in the same way as we have answered Q_1: merely through computing. In effect, in mathematics it is usual to search for general methods to establish, for instance, not only whether a particular equation can be solved or not, but moreover, whether any equation of a specific family of equations can be solved or not. A magnificent example of this way of thinking is the work that led to the proof of Fermat's last theorem [27].

Consider the family of all computer programs printing different sets of natural numbers in increasing order. To this family belong all the programs referred to in the preceding section. It is reasonable to assume that for some of them it will be decided, sooner or later, whether the output has a largest number or not. Will it be possible in the future to find a method which allows us to make a list of all the programs of this family which print a set of numbers without end?

The answer to this question exists, and it is 'no'. We prove this theorem in the Appendix. The proof is a particularly simple proof of the *Undecidability* result referred to in the Introduction. Consequently the day will never come in which man acquires the power to decide for any arithmetical property, whether there are finitely or infinitely many natural numbers having the property. We are led to the conclusion that there will always be unsolved solvable problems. Undecidability states the essential *incompleteness* of human knowledge in arithmetic.

The link between a problem and its solution is usually not apparent. Problems that at first look rather innocent (as for instance Fermat's last theorem), may require highly creative work over centuries to reach a solution. Hilbert's dream was precisely to replace creativity by calculation, so that at the very moment we perceive the problem we can be sure of having the solution. *Undecidability* states the impossibility of this. Creative work cannot be replaced by some technique, procedure or system described beforehand by the mind.

4. Philosophical consequences: Mathematics is not a man's mode of thinking

According to Kant, the world I know, results from my modes of thinking: not that my knowledge follows from reality, but reality follows from my knowledge [28]. I can only know, what my mind *a priori* constructs. This view seems to conflict with the theorem established in section 3, for it actually states that arithmetical truths cannot be considered as contained in man's modes of thinking, even *in principle*.

At this point, it is important to emphasize that Hilbert with his formalistic program, did not intend a reduction of mathematical truths to *a priori* contents of the human consciousness. Hilbert severely criticized Kant's view of mathematics as *a priori* knowledge:

---

27. A. Wiles, and R. Taylor and A. Wiles, *Annals of Mathematics*, **141**, No. 3, (1995).
28. I. Kant, *Kritik der reinen Vernunft* (Vorrede zur zweiten Auflage 1787, B X-XVI), Hamburg: Felix Meiner, 1956, p. 16-20.

„Kant has exaggerated too much the role and the importance of the a priori", regarding arithmetic, as well as geometry and physics [29]. According to Hilbert, the mathematical *a priori* rather remains limited to some basic *insights* which are at the beginning of all thinking and experience.

Nevertheless, one should also acknowledge that Kant's position remains a valid alternative, if one believes in the positive solution to the *Entscheidungsproblem*: the existence of a universal decision method. Answers to arithmetical problems are not apples which can be plucked from a tree. In our opinion Kant is right in considering them as a content of some intelligence. If man were capable of solving through mere computing the infiniteness question for any of the sets of numbers printed by computer programs, then it could be said that all arithmetical truth is actually contained in man's mind. In this case the Kantian view would prevail. The proof that man's mathematical knowledge will remain intrinsically *incomplete* for ever (*Undecidability* theorem), means the exact negation of Kant's epistemological postulate[30].

For a certain natural number to have a *twin prime* or not, to be *perfect* or not, to be a *Fermat prime* or not, is not a matter of choice or chance, but a matter of fact, here and now. The reason for this is that the properties referred to are defined through a *finite* number of instructions. Accordingly, the numbers having one of these properties exist at this very moment,[31] and therefore, whether each of these specific sets has a largest element or not, is a matter of fact here and now, it is something that may not yet be known but is certainly knowable.

This implies that the natural numbers with their properties constitute a mathematical universe which has its own existence, its own reality, independently of any human mind **operation**, in much the same way as, say, the existence of the sun is not due to a man-made construction. However in case of the mathematical reality it is clear that it necessarily needs a mind to be in, to exist. As Kant and Hilbert assume mathematical truths have to be considered as contained in some intelligence, without a mind there is no mathematical world. If this intelligence cannot be the human one, mathematical truths have to be contained in another intelligent principle, somewhere else.

5. „Logging in" to God

What are the alternatives to Kant's view after the *Undecidability* theorem? In our opinion the answer of Richard Feynman quoted in the introductory chapter to this book, indicates the road. „God is always associated with those things you do not understand". The intelligence which explains the mysteries is often called God. In former primitive times people attributed to God things that they did not understand, without taking care to ask whether these things could in the future become understandable through scientific knowledge. Our position now is quite different: We know through scientific knowledge that there will always be something we will never know, and this is the totality of mathematical truth. With

---

29. D. Hilbert, *Gesammelte Abhandlungen*, III. Band, Berlin: Springer, 1935, p. 383-385.

30. Certainly the problem of the *complexity* of mathematical proofs (see the contributions of H.-Ch. Reichel - this volume, chapter I - and F.T. Arecchi - this volume, chapter VI) also challenges Kant's view: indeed it is hard to accept that some mathematical result is *a priori* in my mind, even if I know the method of reaching it, if the proof is so complex that I cannot survey it; a „simple" decomposition into prime factors can become so *intractable* when the number is large, that one cannot go beyond results that are very probable rather than true. Nevertheless *intractability* of mathematical problems does not reveal any incompleteness *in principle*.

31. Such numbers are referred to as *computable* (A. M. Turing, On computable numbers, *Proceedings of The London Mathematical Society*, **42**, 230 (1937)). Notice that irrational numbers as π and *e* are c*omputable*, for they are real numbers whose expressions as a decimal are generated by a finite number of instructions or a computer program. The set of the computable numbers is enumerable. A set of real numbers containing elements which are not computable, as for instance the *continuum*, is not enumerable. We are restricting our considerations to the *computable* numbers.

other words, we are „scientifically“ led to the conclusion that the condition of Feynman's statement regarding God is always fulfilled. It is therefore reasonable to reckon with God.

The God we arrive at with our speculation, is the omniscient intelligence, who grasps at once all possible mathematical knowledge. God is the being who does not need to search to find the link between a problem and its solution. God is infinitely creative, in his mind there is no difference between question and answer. The very thrust of the *incompleteness* argument is this: mathematical truth is either possessed totally at once, or it will never be totally possessed. And if we will never possess it totally, then we will always have need of God's mind.

Undecidability suggests that the thinking process and internal dialogue involved in trying to grapple with unsolved mathematical puzzles is a kind of dialogue with the omniscient intelligence, an attempt to gain access to knowledge already existing in his mind. To ask for example whether there are finitely or infinitely many perfect numbers or twin primes is, in some way, to try to „log in“ to this superior mind. And to find the answer means to have grasped a part of the infinite truth which is contained in this mind, the mind of God.

APPENDIX: Proof of the theorem that in arithmetic there will always be *unsolved solvable problems*

Consider a computer program P_i which prints certain natural numbers in increasing order, and let S_i be the set of numbers generated by P_i. Since each P_i is a finite-bit string of the type:

[10011010111100000001....001],

it is possible to establish the list of all P_i.

Theorem: There is no program U which allows one to decide, $\forall\ S_i$, whether S_i has a finite or an infinite number of elements.

Proof: Let us assume that U exists. Then it is possible to establish the list of all sets S_i that have an infinite number of elements:

$$
\begin{array}{ll}
S_1: & \{a_{11}, a_{12}, a_{13}....a_{1n},....\} \\
S_2: & \{a_{21}, a_{22}, a_{23}....a_{2n},....\} \\
S_3: & \{a_{31}, a_{32}, a_{33}....a_{3n},....\} \\
\vdots & \qquad\qquad\qquad\quad \vdots \\
S_n: & \{a_{n1}, a_{n2}, a_{n3}....a_{nn},....\} \\
\vdots & \qquad\qquad\qquad\quad \vdots
\end{array}
$$

Let us now consider the set S: $\{b_1, b_2, b_3,...b_n,....\}$, defined in the following way:

	consequently, $a_{11} \notin S$
$b_1 = a_{11}+1$	
$b_2 = a_{22}+1$, if $a_{22} \geq b_1$	since $a_{21} < a_{22} < b_2$,
$b_2 = b_1+1$, if $a_{22} < b_1$	either $a_{21} \notin S$, or $a_{22} \notin S$
⋮ ⋮	
$b_n = a_{nn}+1$, if $a_{nn} \geq b_{n-1}$	since $a_{n1} < a_{n2} < ... < a_{nn} < b_n$,
$b_n = b_{n-1}+1$, if $a_{nn} < b_{n-1}$	there is at least one element
⋮ ⋮	a_{ni}, $i \leq n$, such that $a_{ni} \notin S$

- On the one hand, by definition, *S* is a list of natural numbers printed in increasing order, generated by a program (a finite number of instructions), and contains an infinite number of elements.

Therefore *S is a set of the list* $S_1, S_2, ... S_n$....

- On the other hand,

$\forall\ S_n\ \exists\ a_{ni}$, such that $a_{ni} \notin S$, i.e. $\forall\ S_n, S \neq S_n$.

Therefore *S is not a set of the list* $S_1, S_2, ... S_n$....

The only way to escape the contradiction is *to reject the assumption that U exists*. q.e.d.

Chapter V

Mathematics: A Pointer To An Independent Reality
Penrose's interpretation of the Gödel and Turing theorems

Juleon M. Schins
Dept. of Applied Physics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands,
E-mail J.M.Schins@tn.utwente.nl

INTRODUCTION

It is an irony of history that Alan Turing actually offered the main ingredient for an elegant proof of the non-algorithmic nature of human understanding [1]. Turing has always been convinced that human intellect, though (sometimes) more powerful than computers, is not essentially superior [2]. Yet his argument is not very convincing, and does not stand the more recent critique of Roger Penrose [3].

In this contribution we discuss the views of both Turing and Penrose, and see how the reasoning of Penrose has immediate bearing on the main theme of this book: by asserting the existence of an independent, external reality, essentially superior to human mind, Penrose offers the main ingredient for a philosophical proof of the existence of God.

ALAN TURING

In his paper of 1937 Turing considers digital computers with an infinite capacity [1]. An idealized machine, called the universal Turing machine, embodies the full potentiality of algorithmic (i.e. mechanical) procedures. It is supposed to give a number as output, when given as input a pair of numbers, representing the program itself and the program input, respectively. Specific conventions assure that every possible program has its own number, and *vice versa*.

Turing's theorem states that there exists no universal algorithm for deciding whether or not the Turing machine is going to stop (that means, yield a definite answer) when fed with an arbitrary input pair of numbers. As pointed out by Turing himself, a direct consequence of this theorem is that there exists no universal algorithm for answering all mathematical problems belonging to some broad, but well-defined class. This holds because all mathematical problems can be codified, and made amenable to processing by Turing machines. In Penrose's words: *The question of whether or not a particular Turing machine stops is a perfectly well-defined piece of mathematics (and we have already seen that, conversely, various significant mathematical questions can be phrased as the stopping of Turing machines). Thus, by showing that no algorithm exists for deciding the question of the stopping of Turing machines, Turing showed (as had Church, using his own rather different type of approach) that there can be no general algorithm for deciding mathematical questions. Hilbert's Entscheidungsproblem has no solution!*

This is not to say that in any individual case we may not be able to decide the truth, or otherwise, of some particular mathematical question; or decide whether or not some given Turing machine will stop. By the exercise of ingenuity, or even of just common sense, we may be able to decide such a question in a given case. (For example, if a Turing machine's instruction list contains no STOP order, or contains only STOP orders, then common sense alone is sufficient to tell us whether or not it will stop!) But there is no one algorithm that works for all mathematical questions, nor for all Turing machines and all numbers on which they might act [3, p 63].

KURT GÖDEL

Hilbert's *Entscheidungsproblem* (German for decision problem), first formulated in 1900, concerned the existence of a general algorithmic procedure for resolving mathematical questions [4]. Obviously, Turing's theorem represents a negative answer to Hilbert's problem, although it was not the first one: in 1931 the Austrian Kurt Gödel abruptly put an end to Hilbert's optimism by proving that *in any consistent and 'sufficiently powerful' axiomatic system, propositions can be formulated which are true but undecidable* [5].

Some of the above terms require elucidation. 'Undecidable' propositions are syntactically correct propositions that are not provable nor disprovable on the basis of the system's axioms and rules of procedure (that is, not provable nor disprovable 'from within'). 'True' propositions are those propositions whose negation is contradictory either with the axioms themselves (contradiction on the *formal* level) or with their philosophical grounds (contradiction on the level of the *meaning* of the axioms). Finally, a 'sufficiently powerful' axiomatic system must contain at least arithmetic and logic, extended with what is called the μ-operation (which yields the smallest number satisfying some property).

In order to prove that any consistent and 'powerful' axiomatic system contains propositions which are at the time true and undecidable, Gödel codified arithmetic and explicitly wrote down a true but undecidable proposition. To 'codify arithmetic' means to replace the axioms, propositions, proofs, and the individual rules of procedure, by numbers and arithmetical operations. This way, every axiom corresponds to a specific number, and so does every proposition or proof. Whether a given number corresponds to a proof, to just a proposition, or to plain nonsense, can be known either by inspection of the ' decoded number' (that is, of the proposition corresponding to the number), or directly, by checking some arithmetic property of that number (like, for example, its divisibility by some other number).

An ingenious calculation leads Gödel to write down what is now known as the Gödel number. Of course, like any other number in the context of Gödel's work, this special number corresponds to a well-defined sequence of formal symbols. For the present purpose, we shall not write out the Gödel number, nor its decoded counterpart (the Gödel proposition in formal language), but we shall simply state its meaning in words. Here it is: *there exists no proof for the Gödel proposition*.

The analysis of this Gödel proposition proceeds as follows: if the axioms and rules of procedure of Gödel's system are consistent, and no errors occurred in decoding the Gödel number, then there can indeed be no proof of the Gödel proposition. For if there were such a proof, the Gödel proposition would be false as an arithmetical proposition, which implies that the present formal system would allow false statements to be proved. This cannot be the case in a consistent axiomatic system, by which it is clearly established that there exists no formal proof of the Gödel proposition. But this is exactly what the Gödel proposition states itself! Hence, it does so truly, and it must necessarily be a true statement. This way Gödel succeeded in proving that every sound formal system allows for a true but undecidable proposition.

Note that the very existence of a proposition which is undecidable (i.e. it cannot be proved nor disproved by formal reasoning) but true (i.e. its negation implies an absurdity) forces one to admit the existence of what Penrose calls 'a level of meaning', as a level that essentially exceeds the level of formalism: *The point of view that one can dispense with the meaning of mathematical statements, regarding them as nothing but strings of symbols in some formal mathematical system, is the mathematical standpoint of formalism. Some people like this idea, whereby mathematics becomes a kind of 'meaningless game'. It is not an idea that appeals to me, however. It is indeed 'meaning'—not blind algorithmic computation—that gives mathematics its substance* [3, p 105].

CHARACTERISTICS OF MATHEMATICAL TRUTH

A direct consequence of the work of Gödel is that arithmetic shall never be exhausted, and of Turing more generally, that no single 'computing' discipline shall ever be exhausted. Indeed, whatever formal system one considers, it is always possible to formulate a truth that is formally undecidable. Since every truth can be formalized (this is what all science is about) it is in principle possible to adapt the old formal system such that its Gödel proposition becomes decidable. However, this adapted formal system contains a new undecidable truth! Since this is an endless process, there exists an infinite number of Gödel propositions, implying that mathematics is inexhaustible.

In a paper dating from 1939, Turing considers a formal system where the Gödel proposition of some primary system (for example, Gödel's arithmetic) is included as an axiom [6]: this is the most direct way to eliminate undecidable truths. All the subsequent Gödel propositions are recursively included as well. This procedure of recursive inclusion of Gödel propositions can be repeated again and again. Evidently, one shall never reach completeness, and for a quite simple reason. Whatever hyper-recursive formalism one proposes, it always will be formalism, nothing more, nothing less. And every formalism has inexorably its proper Gödel proposition. Penrose comments: *However, this does to some degree beg the question of how we actually decide whether a proposition is true or false. The critical issue, at each stage, is to see how to code the adjoining of an infinite family of Gödel propositions into providing a single additional axiom (or finite number of axioms). This requires that our infinite family can be systematized in some algorithmic way. To be sure that such a systematization correctly does what it is supposed to do, we shall need to employ insights from outside the system—just as we did in order to see that the Gödel proposition was a true proposition in the first place. It is these insights that cannot be systematized—and, indeed, must lie outside any algorithmic action!* [3, p110].

A funny consequence of the fact that mathematics (and all other scientific disciplines) is inexhaustible is that our knowledge can progress indefinitely, although our relative knowledge (i.e. considered as some fraction of all the knowable) is always zero! In this light mathematical truth appears as something which is *certainly not man-made*, but it unfolds an independent, external, and inexhaustibly rich world. But to encompass all of mathematical truth would mean to possess the infinite, which is impossible. Hence, that part of reality contemplated by mathematics is simultaneously accessible and external to man. Penrose formulates it as follows: *The notion of mathematical truth goes beyond the whole concept of formalism. There is something absolute and 'God-given' about mathematical truth. (...) Any particular formal system has a provisional and 'man-made' quality about it. Such systems indeed have very valuable roles to play in mathematical discussions, but they can supply only a partial (or approximate) guide to truth. Real mathematical truth goes beyond mere man-made constructions*. [3, p 112].

Penrose argues that mathematical truth has something absolute about it, in the sense that mathematical achievements appear as a progressively growing edifice, whereby ancient building blocks keep their original value as new blocks are being added. The absoluteness of the statement 'there is no largest prime number' is definitive, or at least of a different order than absoluteness in physics, where new theories sometimes render old concepts and procedures inadequate, or only approximately correct.

An aspect of mathematical truth that Penrose discusses less in detail is that no mathematical system is self-evident, nor self-supporting: its meaningfulness and consistency come from outside the system, and can be grasped by human mind. This is the message conveyed by another interesting number that Gödel wrote down. Translated into words, that number would read*: the consistency of the axioms is undecidable*. As a consequence, there exists not a single arithmetic system, however simple or however complex, that allows one to establish the consistency from within, that is, by using the rules of procedure and the axioms.

Stated in other words, the certainty that an axiomatic system is consistent depends on criteria which are external to the mathematical system, external even to mathematics.

All these features can be understood on assuming that mathematics describes an invisible, *independently existing* object, which is accessible to human mind, though not its 'creature'. Penrose clearly opts for this world view when discussing three important currents presently available: in most of the above citations Penrose rejects formalism; he considers intuitionism plainly absurd; only Platonism provides a world view in the light of which the Gödel and Turing results can be philosophically understood.

MIND VERSUS COMPUTER

One might wonder, after all this evidence, what Turing's views are concerning the computer-mind relationship. In his 1950 paper Turing tackles this question explicitly [2]. He discusses a well-known argument according to which the Gödel and Turing theorems imply that mind cannot operate purely algorithmically:

The short answer to this argument is that although it is established that there are limitations to the powers of any particular machine, it has only been stated, without any sort of proof, that no such limitations apply to the human intellect. But I do not think this view can be dismissed quite so lightly. Whenever one of these machines is asked the appropriate critical question, and gives a definite answer, we know that this answer must be wrong, and this gives us a certain feeling of superiority. Is this feeling illusory? It is no doubt quite genuine, but I do not think too much importance should be attached to it. We too often give wrong answers to questions ourselves to be justified in being very pleased at such evidence of fallibility on the part of the machines. Further, our superiority can only be felt on such an occasion in relation to the one machine over which we have scored our petty triumph. There should be no question of triumphing simultaneously over all machines. In short, then, there might be men cleverer than any given machine, but then again there might be other machines cleverer again, and so on.

What is intriguing about this passage is that Turing merely compares the *performance* of human intellect with that of machines. The revolution brought about by Gödel, and generalized by Turing, is reduced by Turing himself to a merely technical question of computing speed. It cannot be stressed enough that the question at stake is not about computing speed, but about essential superiority. A machine can be essentially superior to human mind only if there exists an operation that the machine can do and that human mind cannot do. As long as machines are designed by human mind, the conception of an essentially superior machine seems quite improbable. And what about the reverse: is it really possible to prove, starting from the Gödel and Turing theorems, that human mind is essentially superior to machine?

Penrose's answer to Turing, concerning the non-algorithmic nature of mathematical insight, deserves extensive citation: *We must first consider the possibility that different mathematicians use inequivalent algorithms to decide truth. However, it is one of the most striking features of mathematics (perhaps almost alone among the disciplines) that the truth of propositions can actually be settled by abstract argument! A mathematical argument that convinces one mathematician—provided that it contains no error—will also convince another, as soon as the argument has been fully grasped. This also applies to the Gödel-type propositions. If the first mathematician is prepared to accept all the axioms and rules of procedure of a particular formal system as giving only true propositions, then he must also be prepared to accept its Gödel proposition as describing a true proposition. It would be exactly the same for a second mathematician. The point is that the arguments establishing mathematical truth are communicable* [3, p 417].

The first ingredient of Penrose's argument is the objective character of mathematics. If mathematics does not contemplate some kind of object, with properties independent from any observer, one cannot understand why mathematical arguments can be settled at all. Mathematics would stop to be a scientific discipline, and turn to some democratic exercise, without any historical continuity. Penrose's word 'communicable' should be understood referring to its convincing power, rather than to the possibility of communicating personal taste.

He goes on: *Thus we are not talking about various obscure algorithms that might happen to be running around in different particular mathematician's heads. We are talking about one universally employed formal system which is equivalent to all the different mathematicians' algorithms for judging mathematical truth. Now this putative 'universal' system, or algorithm, cannot ever be known as the one that we mathematicians use to decide truth! For if it were, then we could construct its Gödel proposition and know that to be a mathematical truth also. Thus, we are driven to the conclusion that the algorithm that mathematicians actually use to decide mathematical truth is so complicated or obscure that its very validity can never be known to us* [3, p 417]. The second ingredient is the assumption that mathematical truth be established algorithmically. With these two ingredients, and a thorough understanding of the mathematical and philosophical consequences of the Gödel theorem, Penrose remarkably proves an inconsistency: *knowing* the truth-establishing algorithm, the mathematician also *could* construct its Gödel proposition, and know that to be a mathematical truth also, though *not* on the basis of that truth-establishing algorithm, but of some additional truth-establishing source! The conclusion is now exactly the negation of the assumption. Ergo, human mind does not establish mathematical truth algorithmically.

Penrose finishes as follows: *To my thinking, this is as blatant a reductio ad absurdum as we can hope to achieve, short of an actual mathematical proof! The message should be clear. Mathematical truth is not something that we ascertain merely by use of an algorithm. I believe, also, that our consciousness is a crucial ingredient in our comprehension of mathematical truth. We must 'see' the truth of a mathematical argument to be convinced of its validity. This 'seeing' is the very essence of consciousness. It must be present whenever we directly perceive mathematical truth. When we convince ourselves of the validity of Gödel's theorem we not only 'see' it, but by so doing we reveal the very non-algorithmic nature of the 'seeing' process itself* [3, p 418].

In his latest book Penrose dedicates 200 pages (roughly half the volume) to refuting objections raised against his interpretation of the Gödel theorem. [7] Evidently, this matter is not at all easy to grasp. An immediate benefit resulting from the discussion is the fact that Penrose has now defined his position much more accurately. For example, he analyzes in great detail the possibility that human understanding proceeds on the basis of sound/unsound, single/multiple, knowable/unknowable, provable/unprovable algorithms (one of the combinations being a single, sound algorithm that can be known by man, though such that it would not be possible to prove that human mind indeed decides the truth of mathematical propositions using that algorithm). Moreover, Penrose considers whether (pseudo-)randomness and chaos can provide the non-computable aspect of human understanding. In all these cases however, Penrose's analyses follow basically the same line of reasoning as presented above, and his main conclusions are unaltered.

CONCLUSION

It is an extremely important fact that Penrose succeeds in lifting the Gödel discussion from a purely pragmatic level to a philosophical one. This allows him to conclude that mathematical truth goes beyond mere man-made constructions. Formalists consider the pragmatic level to be the only existing one: they acknowledge the validity of the Gödel statement, but consider it of no use for further mathematical progress. And whatever is not

useful is not true. This intentional myopia seems an easy way out, but it does not allow for a consistent understanding of mathematical facts.

The philosophical level of discussion, as proposed by Penrose, is in general poorly appreciated, since philosophy is hardly considered a scientific discipline. As a matter of fact, a brief look at history shows a wild variety of philosophical currents, lacking that characteristic of scientific disciplines, which is the gradual settling of controversies in the course of time.

Penrose suggests to turn the argument around: without introducing *philosophical* concepts like 'meaning' and 'existence', it is not possible to understand the *mathematical* results of Gödel and Turing properly. These mathematical results firmly establish the existence of something that is unlimited and absolute, fully rational and independent of human mind. This 'something' can evidently not be a creation of the human mind, nor can it be its necessary emanation. Penrose's main concern is to explain how the human brain, which is a physical system, can show non-computational activity, while all the presently known physical laws describe computable relations (apart from the randomness involved in quantum-mechanical measurement theory). Hence he proposes in which direction to investigate in order to find the hitherto unknown non-computable theory. In this paper our concern is more of philosophical nature: what is the origin of that mathematical world, of which Penrose emphatically claims that it *exists*?

The reference to God is normally considered as personal and not scientific. Penrose is not an exception: he asserts that the acceptance of religious doctrine necessarily implies a non-scientific approach, for example, to the problem of explaining awareness. [cf. 7, p12] This is not necessarily true. Certainly not for Christianity, which provided the very cradle for systematic scientific activity. It cannot be denied that the question for the origin of the mathematical world is a licit and scientific one. Penrose's interpretation of the Gödel theorem showed this mathematical world to be inexhaustible (man can know each truth separately, but he can not know them all, since there is an infinity of truths), absolute and independent of the human mind (if not one has a hard time to explain intersubjectivity, the *conditio sine qua non* for any scientific work), and eternal (physical laws or 'constants' may change in the course of time, but not the very measures of that change). What kind of object would represent a more convincing pointer to God, than this Platonic realm of mathematical entities?

To be annoyed because Gödel does not allow for complete dominion of mathematics, is an achievement, worthy of a hard worker. To recognize the invisible hand of some infinitely higher intelligence, requires more effort, worthy only of the finest philosophical spirit.

References

1. A.M. Turing, *Proc. Lond. Math. Soc*. **42** (1936) 230 and 43 (1937) 544
2. A.M. Turing, *Mind* **59** (1950) 433
3. R. Penrose, in *The Emperor's New Mind*, Oxford University Press (1989)
4. D. Hilbert (tenth problem), Intern. Congr. Math., Paris (1900)
5. K. Gödel, *Monatshefte fur Math. und Phys.* **38** (1931) 173
6. A.M. Turing, *Proc. Lond. Math. Soc*. **45** (1939) 161
7. R. Penrose, in *Shadows of the Mind*, Oxford University Press (1994)

Chapter VI

A critical approach to complexity and self organization*

F.t. Arecchi
Universitá di Firenze and Istituto Nazionale di Ottica, Florence, Italy

1. Introduction

The debate about the role of the 3 C's (catastrophe, chaos, and complexity) in physics has become a popular subject of conversation even among lay people. In order to put some order in a matter which has rather important implications, I shall introduce the necessary definitions, recurring, as far as possible, to heuristic pictures. First of all, I show that the 3C's stem from nonlinear dynamics. They are strongly bound to the concept of organization (misnamed as self-organization), as well as to the strategies of recognition and decision in a cognitive system.

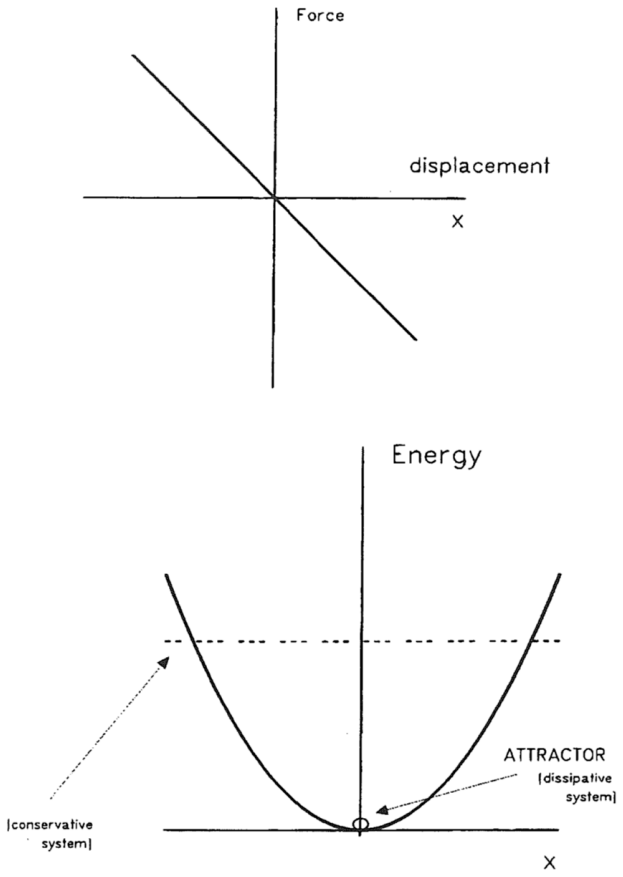

Figure 1. Linear dynamics. The straight line is force *versus* position. The associated potential energy has a parabolic shape. For a conservative system the total energy, E_0 is conserved, and the potential energy tranforms into kinetic energy, with a corresponding velocity increase; the velocity reaches a maximum at the bottom of the valley and is then reutilized to climb the other side. In the dissipative case, the velocity is lost and the motion stops at the bottom. This is then called "attractor".

* Lecture presented at the *Fourth European Conference on Science and Theology*, Rome, March 1992 and published in *Studies in Science & Theology* 1993, 1 *Origins, time & complexity*, Eds. G. v. Coyne, S.J., K. Schmitz-Moorman and C. Wassermann.

As a result of the 3C's the physical sciences have undergone a crisis analogous to that which exploded a few decades ago in mathematics with questions of undecidability and intractability. To cope with this crisis implies rediscussing the role and power of human knowledge, and this will have anthropological and ontological consequences. The main consequence is that the scientific program has fundamental implications for our world views. In fact, two main theses will conclude this study, namely: (i) in contrast to a Turing machine, a biological cognitive system does not get trapped in undecidable or intractable problems, but it decides, and does it within a reasonably short time; (ii) this happens because it does not obey fixed rules, but it continuously readjusts itself, adapting to the intrinsic evolution of the event under observation; in other words, knowledge is intentional, since it consists in a continuous *adaequatio* of the mind to the observed reality. In particular, human knowledge provides not only for an efficient survival strategy, but (in this it is unique among living beings) it includes a reflection on the structures, the natural forms, which have shaped our cognitive rules. Thus, a hierarchy of orders emerges as an intrinsic property of reality, not as superposed by the transcendental activity of the human mind, which imposes its own "values". This objective assessment of the "degrees of perfection" is the working tool upon which the fourth way of Thomas Aquinas is based. Thus, it becomes the trail through which the human mind hints at a Creator.

2. BIFURCATIONS AND ORGANIZATION

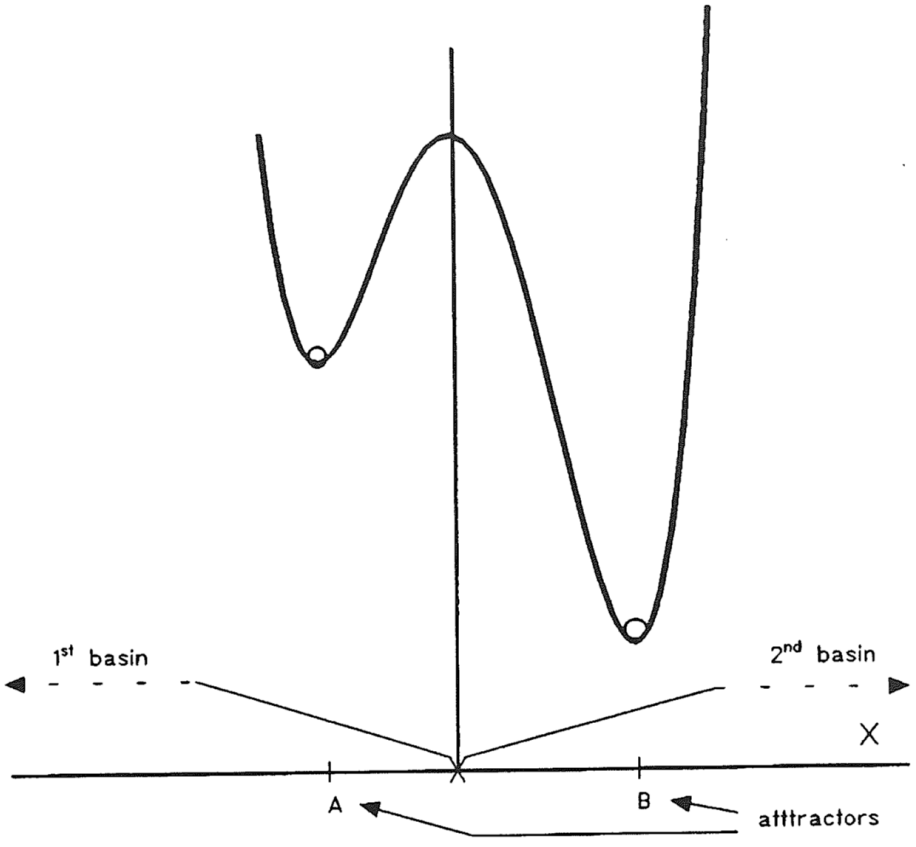

Figure 2. Nonlinear dynamics: example of an energy landscape with two attractors.

Most dynamical models until the late 1800's were linear, that is, with a straight line diagram of force versus position, and consequent parabolic behavior for the potential energy (Fig. 1). Such is the situation of Hooke's elastic spring (*ut tensio sic vis*). In conservative systems, as

the system evolves, what is lost in potential energy is gained in kinetic energy and the motion persists forever. If, on the contrary, there is a strong viscous damping (dissipative systems), the kinetic energy is released to the surrounding medium, and the falling body stops at the minimum of energy, $x = 0$. We call "attractor" this final point and "basin of attraction" all initial positions which merge in the bottom of this energy landscape. Thus dissipative linear dynamics has no surprises. From wherever the point starts, it will end up in the bottom.

If by changing a "control parameter" we distort the energy landscape, as, e.g., in Figure 2, we enter the realm of nonlinear dynamics.

Linear dynamics is universal in the sense that *any* initial condition (x in $-\infty, +\infty$) will converge to the unique attractor. There are instead many nonlinear dynamics characterized by different attractor numbers and positions, as in the example of Figure 2. Figure 3 shows how the energy landscape changes for different λ. At a critical point λ_c the single valley is replaced by two separate valleys; this is a qualitative change. If we represent the position, $\bar{x}$, of the attractors *versus* λ, the change at λ_c appears as a "bifurcation", as shown by the solid line of Figure 3b. The dashed line represents the hilltop.

We have shown the simplest of all bifurcations. Accounting for geometric constraints, one can classify all possible shapes of bifurcations. This was called "catastrophe theory" by R. Thom, but I rather prefer the less conspicuous term of bifurcations.

The previous figures refer to a simple system, that is, a marble sliding along the energy landscape. The ambitious program of statistical mechanics is to

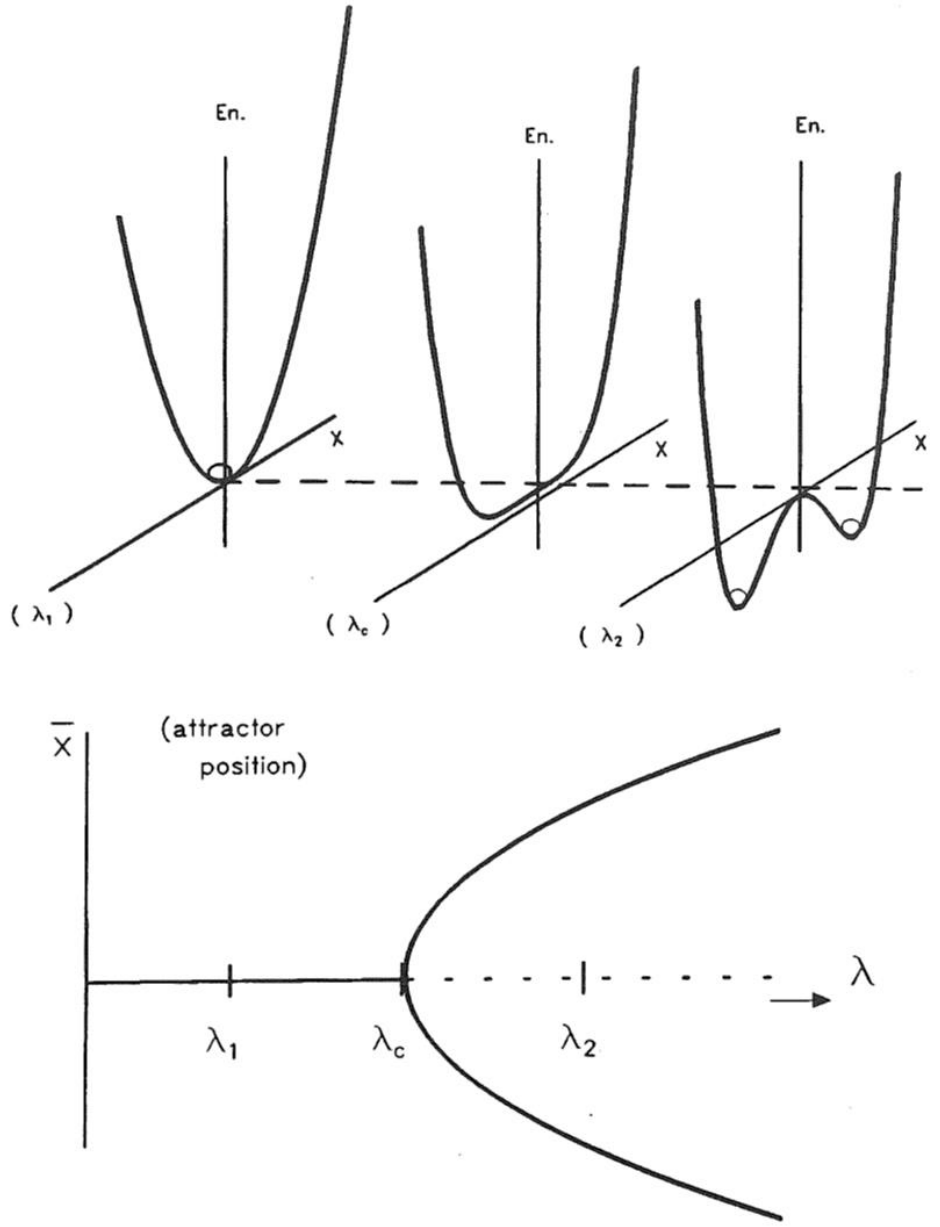

Figure 3. (a) Modification of the energy landscape for a varying control parameter λ. (b) The positions of the valley bottoms (solid lines) and of the hilltop (dashed line) are plotted *versus* λ. λ_c = bifurcation value. For each λ the bottom is a stable fixed point, the hilltop is an unstable fixed point.

extend this description from a single object, as an atom or a molecule, to a large collection of objects behaving as a single body. As far as the component particles do not interact, or interact by linear forces, like the atoms of an elastic chain, each of them is embedded in an energy landscape as that of Figure 1 with just one attractor and no bifurcations. It is immaterial whether we have 1 or 10^{22} atoms (such is the number of a cubic centimeter of solid matter), since the single object is representative of all the atoms. That is why

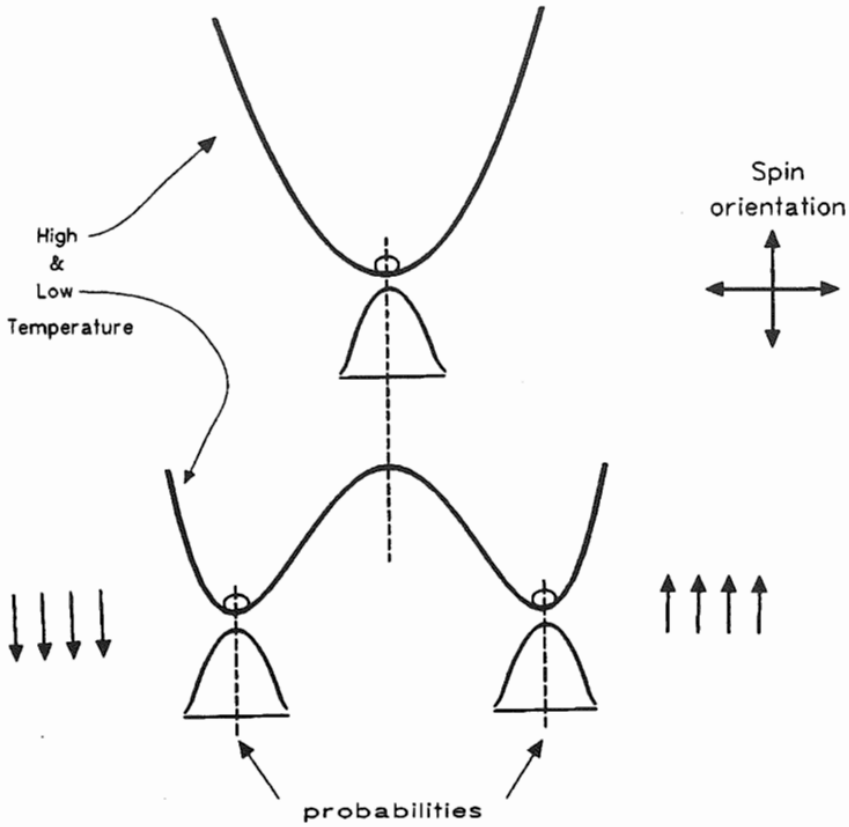

Figure 4. Collective model of a magnetic system at high temperature (paramagnetic phase) and at low temperature (ferromagnetic phase). The probability curves around the energy minima denote the role of thermal fluctuations.

an ideal gas has no qualitative changes, no "phase transitions". If instead we consider nonlinear forces, a parameter change can induce bifurcations. We are no longer able to provide a "global" description which holds everywhere, but we attempt a "local" description, the so called Landau model, which applies the above ideas to a collection of many interacting particles as if they were a single body. Such is indeed the approximate behavior close to a phase transition.

Consider indeed a macroscopic magnet. Its magnetization is the sum of all the atomic contributions (the so called atomic spins). At low temperature they attract one another and hence they become parallel yielding a macroscopic sum. At high temperature the mutual attraction is contrasted by thermal agitation which scrambles the single spin orientations. Thus they compensate one against the other, giving zero total sum.

The overall magnetization, that we may call the order parameter, acts as the position of the single particle of Figure 3. At high temperature it is attracted by the zero value, at low temperature it is attracted by two macroscopic values of opposite sign. Indeed the symmetric bifurcation does not favor either choice. To point out that this is a collective and not just a single particle descrip-

tion, we sketch also some bell-like curves which give the residual spread of magnetization around each attractor due to thermal fluctuations (Figure 4). If we repeat the same experiment many times, we pick up samples out of those probability distributions. Let us consider the one at high temperature. Since it is so symmetric around the zero value, as we cool down the system we have equal probability of landing with magnetization up or down.

The presence of a tiny external field $H \neq 0$ induces a "symmetry breaking" which makes the two branches no longer equivalent (Figure 5a). Thus an external perturbation is responsible for selection. Repeat this selection many times, through a sequence of bifurcations (Figure 5b). The system reaches a state of "organization", that is, it assumes just one among a large number of possible configurations. Some people call this "self-organization", to stress that the bifurcations depend on the inner dynamic relations among the system components. I find that word not appropriate, because a specific path (dashed line of Figure 5b) is the result of many selections imposed by external perturbations. Thus, it was not sufficient to have that number of components (atoms, biomolecules, etc.), but it was crucial to put them in that specific "ecological niche" which provided the appropriate selective push (symmetry breaking) at each bifurcation. For this reason I propose the term "hetero-organization" as more appropriate than "self-organization".

True self-organization takes place only when the system is not influenced by boundary perturbations and it depends only on its inner dynamics. This occurs at the so called thermodynamic limit, when the system has a homogeneous composition over an infinite range. In such a case we no longer have a peculiar final state solution, but all states will be reached in different trials, with a probability curve which reflects the dynamical constraints. This limit procedure may provide pedagogical matter for laboratory experiments, but it does not help in understanding the unique world in which we live, and how it was molded in the course of evolution. For such a purpose, only if we were able to describe each system with its ecological niche, and the niche of the niche, and so on, would we fulfil this endeavour. Thus, this program requires the feasibility of a global description and we have just learned that the Landau models have only a local value, close to a bifurcation.

3. SYNERGETICS, PATTERN FORMATION, AND PATTERN RECOGNITION

The Landau model arises from heuristic considerations. However, in some cases we are able to trace the whole passage from microscopic dynamics to a macroscopic description of the Landau type. This occurs for lasers, fluids, and chemical reactions, for which we can build a macroscopic model almost from first principles. Such an area of investigation was called "Synergetics" by Haken [1], to denote an identity of behaviors based only on dynamical arguments, independent of the physical nature of the problem.

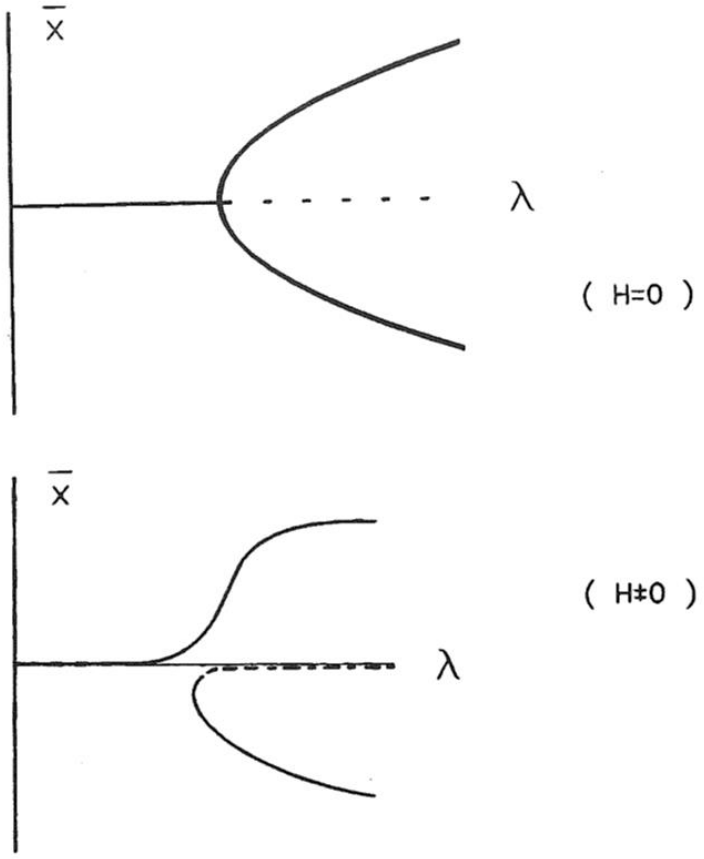

Figure 5. (a) In ideally symmetric conditions ($H = 0$) a temperature decrease as in Figure 4, yields equal probabilities of aligning the magnets upward or downward. An external field ($H \neq 0$) breaks the symmetry making one branch more preferable than the other. The gap (minimal separation) between the two branches must be larger than the fluctuations.

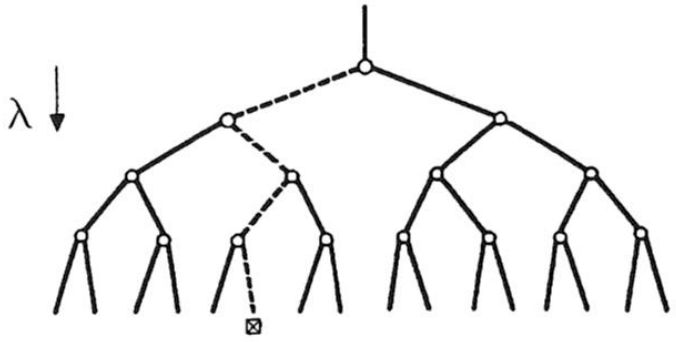

Figure 5. (b) If we have a long chain of bifurcations (the control parameter here points downward) and an external perturbation provides a selection at each bifurcation, the final state is unique, even though the dynamics allow for a very large number of final states: this is the meaning of organization.

In fact, only under strong assumptions can we trace all the way from microscopic to macroscopic, and, in general, we must rather resort to a heuristic of the Landau type, whereby the model is suggested by macroscopic symmetries rather than by individual behaviors. In the case of the magnet, we saw that the transition is explained in terms of competition between thermal agitation and spin interaction and that it is controlled by temperature. Analogous heuristic models without direct reference to a microscopic behavior were applied by A. Turing in 1952 to morphogenesis, that is to pattern formation in chemical reactions or embryonic development. In this case the control parameter is not the temperature, but the amount of two chemical species, respectively the activator and the inhibitor. To obtain an interesting dynamic of patterns the space and time scales of activator and inhibitor must be widely different, otherwise the relative change in their populations does not provide bifurcations. Similar competition, which reminds one of the struggle for life of two competing biological species (Volterra–Lotka population models), explains the convective instability of a fluid or the onset of the laser action.

The difference of time scales between the two competing processes has been called by Haken [1] "slaving" of the fast variables, which adapt themselves to the conditions imposed by the slow variables. These latter ones which survive over a long time scale, are the order parameters, as the magnetization in the example of the previous section.

Models of cognitive systems based on the properties of morphogenetic systems have been recently considered. They consist of coupled "neurons", the so called neural networks, with two time scales. A neural network with fixed coupling parameters adjusts itself quickly to an external input (STM = short time memory). Yet in order to be plastic, that is adaptable to broad classes of inputs, the coupling parameters are not fixed once and forever, but they are readjusted upon the collective state of the network (some mean, as the sum of the spin values in the above case of the magnet) over long time scales (LTM = long time memory).

The above model of parallel computation with adjustable rules is much more sophisticated than the early approaches of artificial intelligence in the 1950s based on sequential computation and fixed rules. If we consider them as fair prototypes of the cognitive strategy of a living system, it appears that recognizing signifies readjusting the parameters until the network state matches the input. Older approaches (Hopfield models), with rules assigned by learning sessions preceding the recognition (Hebbian rules), even though parallel, had a limited amount of memory content, because they could only project any external input on an internal archive of prelearned patterns which is limited by the size of the network. This new adaptable network is equivalent to a network of virtually infinite size (for a preliminary approach, see Basti [2]).

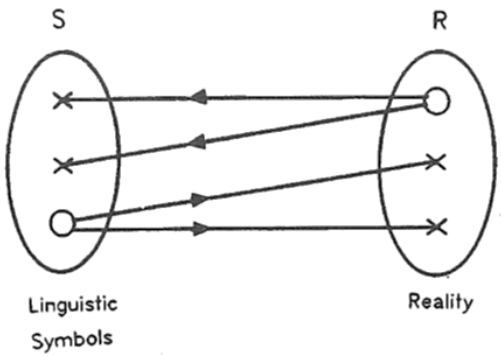

Figure 6. Mapping of reality into the symbols-words of ordinary language.

4. THE LANGUAGE OF SCIENCE AND ITS LIMITATIONS

The language of physics was introduced by Galileo as a way out of the ambiguities typical of ordinary language (Figure 6). Rather than trying to catch the "essence", that is, to formulate words which fully describe the "nature" of an event X, physics limits itself to assign some quantities related to ("quantitative affections of") X, e.g. the length of X, the weight of X, etc.

As a consequence, the approach of the observer to reality R is mediated by some measuring apparatus M (Figure 7). The new language in the space of symbols S is made of symbols-numbers which are the outputs of M and whose "semantics" depends upon the operations performed by M, while the "syntaxis" connecting those numbers is the mathematics. To complete physics as a formal language we fix a set of primitive relations among the numbers that we take as the "physical laws" and that become the "axioms" of the language. Whether those relations are extracted from reality by trial and error (guessed and then tested) (Galileo, Newton) or whether they are a transcendental human activity ordering otherwise raw data (Kant, Cassirer [3]) is a matter of philosophical debate, which for the time being, may be considered a metascientific question, not a scientific one. Scientists, however, are generally realist in the sense that they take as a strong evidence what "common sense" suggests to them before any linguistic formulation.

Three centuries of success in the research in physics have suggested three pretensions, or dogmas:

i. *Decidability*. Once a suitable set of symbols has been collected and the corresponding relations (laws) have been formulated, all the rest is a matter of straight-forward mathematical derivations.

ii. *Predictability*. Once the equations of a problem are known, given a set of initial conditions, the future of that problem is univocally fixed, since the solution of a differential equation with assigned initial conditions is unique.

iii. *Explainability*. Rather than postulating different sets of laws for any level of organization, apply Occam's razor (*Entia non sunt multipli-*

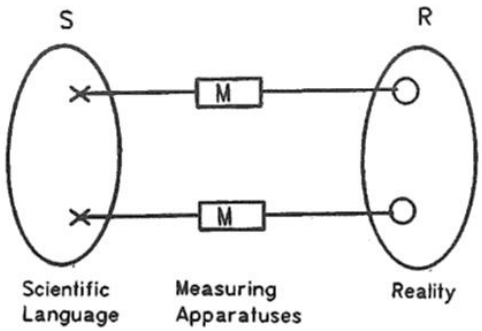

Figure 7. Mapping of reality into the symbolsnumbers of the scientific language, through the mediation of measuring apparatus, M.

canda praeter necessitatem). Hence split any complex object into its components, find the elementary laws connecting the components, and reconstruct the behavior of the whole object.

The above three dogmas have found "in principle" limitations from within the same scientific language. The three limitations are respectively:

i. Gödel undecidability theorem (1931)
ii. Deterministic chaos, i.e. sensitive dependence on initial conditions (Poincaré 1890, Lorenz 1963, etc.).
iii. Complexity, i.e. emergence of new information in a large system, not included in the separate constituents (Kolmogorov, Chaitin [4], Bennett [5], Atlan [6], etc. from 1960).

We consider these three limitations as new paradigms of science which introduce crises in the very foundations of mathematics (i) and experimental physics (ii and iii). We wish to show that these crises are healthy. On the one hand they provide new impetus to scientific investigation, on the other they destroy the pretension of science to be the only relevant language to describe reality (scientism), so that science becomes open to interdisciplinary relations with other languages.

4.1. *Undecidability*

Once we have organized some basic data in a set of statements, we have presumed to be able to do science in a deductive way, taking those statements as axioms and deducing all possible consequences. The Gödel theorem hampers such a pretension. It says that, given a set of axioms, there is eventually a statement which is true because it is expressed by the rules of formal language, but for which we cannot prove that it is true nor that it is false without falling into a paradox. We must consider such a lack of uniqueness as a limit to the reliability of a long list of grammatical constructions. Thus, we must get out of the realm of linguistic symbols and search in the real world for further relevant input. Truth can no longer be considered in a self consistent

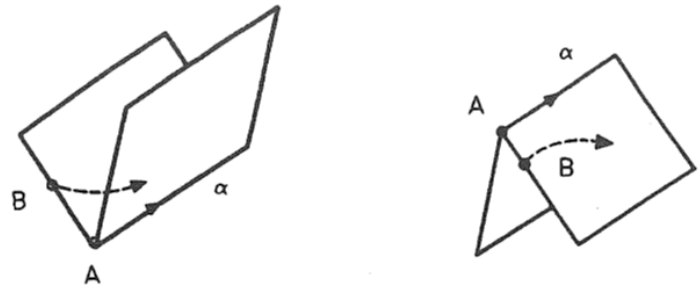

Figure 8. Energy landscape around a dynamical trajectory. (a) Valley floor means regular motion, that is, convergence of nearby trajectories; (b) ridge of a hill means irregular motion, i.e., rapid divergence of nearby trajectories. This is chaos.

way as the correctness of a formal procedure, but it must be given a semantic value with reference to R.

4.2. *Chaos*

Determinism was based on the uniqueness of the solutions to dynamical equations, once initial conditions have been assigned. Although a trajectory may be unique in starting from certain initial conditions, it is enough for it to have a minimal uncertainty to lose the predictability of its future path. Now these minimal uncertainties are intrinsic to the method of measurement itself. In coding events by numbers we can assign with accuracy only the rational numbers, but by far the majority of numbers are irrational, such as the square root of 2, that is, they consist of an unlimited sequence of digits. As infinity can neither be encompassed using our systems of measurement nor recorded in our memory, the truncated version of an infinite number introduces a tiny initial uncertainty, the effects of which become enormous when we try to extend our prediction beyond a certain time. We can express in a diagram (Figure 8) how it is possible to deviate from the "unique" path. Let us compare two equal paths but with different surrounding "landscapes": the first on a valley floor, the second on the ridge of a hill. The initial "exact" position A gives the required path; a slightly wrong position B gives a path which in the first case converges toward the correct one, but in the second diverges away from it.

Even simple physical systems like Poincare's problem of three bodies have critical paths which run along a ridge and can give rise to deterministic chaos. Similar effects, called deterministic chaos, can be observed nowadays in various different situations: chemical reactions, the motion of fluids, lasers, cardiac rhythms, the movements of asteroids, economic and social trends, etc.

Why did it take two centuries to discover that paths are not always stable (at the bottom of a valley), but are often unstable (on the ridge of a hill)? The fact is that physicists have limited their considerations mainly to stable equilibria and have then only examined small perturbations around these. Now these

tiny movements obey linear dynamics which always produces trajectories with valley floors (see Figure 1), i.e., they assure future predictability. Thus, the simplified models used to study nature excluded certain pathologies. Nonlinear dynamics is, however, the way in which nature normally behaves.

4.3. *Complexity: A Heuristic Approach*

Entering nonlinearity means to discover complexity. The fact that one's starting point is never a geometric point from which there emerges a single line into the future, but, in general, a small blob from which there fan out lines in all directions, can be seen as a case of dynamic complexity, i.e., as an infringement of the simplicity requirement which forms the basis of the Galileian method.

In addition to the dynamic complexity of deterministic chaos we can see the emergence of structural complexity which consists of the impossibility of satisfactorily describing a large object by reducing it to an interplay of its component parts with their elementary laws. In this sense, complexity is associated with descriptions, rather than being an intrinsic property of objects. A classification of the processes of description should then provide different complexity measures and we shall face this task in the next section.

The two problem areas sketched above, that of deterministic chaos and that of complexity, are beginning to be translated into quantitative parameters. It is precisely the limits on predictability which impose a continual introduction of information if we are to be able to continue to make predictions about the future. The speed with which information is used up is indicated by a parameter K (after the mathematician A.N. Kolmogorov). Moreover, tentative attempts are being made to portray complexity with a parameter C, which indicates the cost of a computer program capable of realising a predefined complex objective. Let us consider three sequences of letters of our alphabet: the first one is random, the second is regular, that is, a chain of letters repeating in the same order, and the third one is part of a poem, or any other literary text.

Let us agree to define as complexity the cost of the computer program which enables us to realize one of the three sequences. In the first case the program is simple because it can be associated with the outcome of a random number generator. For the second case the instruction is very short. In the third case there exists no program which would be shorter than the poem itself. This last sequence will be called complex, the other two simple.

In conclusion, we can classify the limits which physics imposes on itself at present with a diagram CK (Figure 9). Note the significance of the diagram. Let us start with the horizontal axis: $K = 0$ means predictable, highly ordered systems. A very high K means that the information disappears very rapidly as in the systems with maximum entropy. While the horizontal axis tells us how things develop in time, the structure is represented by the complexity C.

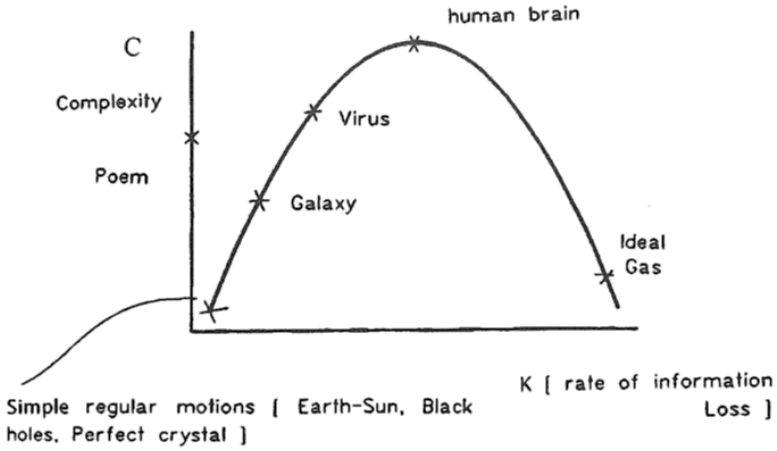

Figure 9. Qualitative classification of natural events in a diagram C (complexity) K (chaos).

Objects in traditional physics, from the Newtonian system to Black Holes, are relatively simple. By contrast complex systems can be regarded as those systems in which the objectives to be achieved are in mutual conflict, so that if one is achieved, the others have to be relinquished. A typical example is our own immune system.

We can say, in anthropomorphic terms, that complex systems are those in which a choice has to be performed, and hence a unique description cannot exist. Once we have recognized that the reductionist's dream leads to complexity, the only way out is to give up on having a unique set of elementary laws. In respecting the various, irreducible levels of the description of reality, physics is on the one hand respecting the scientific organization of other areas of science (from biology to sociology) without trying any longer to reduce these to "applied physics"; on the other hand it is rehabilitating certain aspects of the Aristotelian "organicism" for which "a house is not the sum of its bricks and beams inasmuch as the architect's plan is an integral part of it".

5. COMPLEXITY: ATTEMPTS AT FORMAL DEFINITIONS

After the heuristic approach of Section 4, here we try to establish a quantitative indicator, C, of complexity. This has been done by many people, but all definitions proposed thus far are subject to some kind of criticism.

5.1. *Probabilistic Definition*

What is common to all the phenomena of organization is the emergence of spacetime structures which seem "new" with respect to what is known of the constituents as shown in Figure 5b. Let us see how this was expressed by Anderson [7]:

The behavior of large and complex aggregates of elementary particles, it turns out, is not to be understood in terms of a simple extrapolation of the properties of a few particles. Instead, at each level of complexity entirely new properties appear, and the understanding of the new behaviors requires research which I think is as fundamental in its nature as any other. That is, it seems to me, that one may array the sciences roughly linearly in a hierarchy, according to the idea: The elementary entities of science X obey the laws of science Y.

X	Y
manybody physics	Elementary particle physics
chemistry	manybody physics
molecular biology	chemistry
cell biology	molecular biology
*	*
*	*
psychology	physiology
social sciences	psychology

But this hierarchy does not imply that science X is "just applied Y". At each stage entirely new laws, concepts, and generalizations are necessary, requiring inspiration and creativity to just as great a degree as in the previous one. Psychology is not applied biology, nor is biology applied chemistry.

Has a higher level to be described by different laws or is organization a recipe to guess the higher level of organization once we know the rules at a lower level? Such a program seems to be contained in the following statement by Atlan [6]:

> ... We cannot assume that the finality of natural organization has been set up by some consciousness either from the outside or from the inside. In either case, this would amount to resorting to God as an explanatory principle ... and no one knows what God's purpose was ... in setting up an organism or an ecosystem, as opposed to a manmade machine which has been designed with a given known purpose in mind. This is why various attempts have been made to understand mechanisms of self-organization by which non-purposeful systems, not goal oriented from the outside, can organize themselves in such a way that the meaning of information is an emerging property of the dynamics of the system.

My fundamental criticism is that such a program is based on a subjective attribution of a purpose, a "value", by the observer. Such is the case of the attempt by Kauffman to attribute an "emergent" complexity to an otherwise simple Boolean model.

Let me explain my criticism in detail. The emergence of meaning, in the sense of Atlan, is based on a comparison between the global information, I_t, of a compound system and the sum $\Sigma_i I_i$ of the information, I_i, of the individual components. If $I_t > \Sigma I_i$, the difference, which is the mutual information (M.I.) among the components, is the complexity C, that is,

$$C = \text{M.I.} = I_t - \Sigma I_i.$$

In the case of a written text, M.I. is related to the word positions in the phrase, while I_i is the information provided by the dictionary for each single word. Of course, a random collection of words or a collection put in some conventional order (e.g. alphabetical) has $C = 0$. Now information, according to Shannon,

is related to the "surprise" of a message, and hence to the inverse of its probability p, by a logarithmic dependence

$$I = \langle \log 1/p \rangle = -\Sigma p_i \log p_i .$$

Here the pointed brackets denote an average over the ensemble of possible situations whereby the event may occur, and this average is the sum of $\log 1/p = -\log p_i$ with each term weighted by its own probability p_i of occurrence.

The drawback of this approach is that assignment of a probability measure is subjective, related to the "value" we attribute to an object. In communication theory, where we worry only about information loss in a communication channel, it is sufficient to choose a conventional probability measure. When we refer to a piece of the real world, information has a semantic connotation, and the probabilities must be assigned after a prescientific recognition which has attributed different weights to different events, depending on their use.

An apparently objective definition is the one based on "frequencies" by Venn and von Mises. We partition the space of events by small equal boxes, and attribute to the ith box a probability given by the number of times, N_i, the event falls into it, divided by the total number N of events

$$N_i = N_i/N.$$

While this definition is helpful to classify trajectories in chaotic dynamics, it is useless for classifying forms of objects. For instance, if we take a chip of a few cubic centimeters of glass and shape it as a Venetian artistic cup, the above criterion applied to the space occupied by the amorphous chip and by the formed cup yields the same information, provided that the glassy material has the same consistency. So, information provides criteria just for detecting sponge-like (fractal) structures in the interior of the glass. Only by attributing a higher weight ("value") to those boxes of the partition along the artist's contour do we succeed in attributing higher information to the cup.

5.2. *Computational Definition*

In addition to the semantic problems related to a probability measure, the previous definition requires knowledge of an ensemble of systems, in order to assign probabilities as relative frequencies of occurrence. It seems reasonable that we should be able to describe a system as ordered or disordered irrespective of how much we know about it. What is needed is a measure of complexity that refers to individual states, rather than to ensembles. Mathematicians have proposed such a measure using a new branch of mathematics known as algorithmic complexity theory.

The algorithmic complexity of a state is the length (measured in bits of information) of the shortest computer program that can describe the state. Suppose that a particular string of information looks like 101010101010... We

would assign this state a low algorithmic complexity because we can recover it by a very short algorithm, namely: "Print 10 n times". This regular binary array can be regarded as the arithmetic equivalent of a crystal. We cannot, however, generally display an arbitrary sequence of ones and zeros using a short algorithm. It can be proven that almost all sequences cannot be reproduced by algorithms significantly shorter than themselves. That is, the algorithm contains almost the same information as the sequence itself. In attempting to generate such a sequence, therefore, we can do little better than simply display a copy of the sequence. Only rarely is a sequence "algorithmically compressible", i.e., it can be generated by an algorithm containing less information than the sequence itself.

This notion provides a formal definition of randomness: a random sequence is one that cannot be algorithmically compressed. This satisfies our intuitive expectation that random sequences are devoid of all patterns; the existence of any pattern would imply a more compact description because we could write a short computer program to specify that pattern. We cannot tell just by looking, however, whether a given sequence is random. Indeed, Chaitin [4] has shown, in an extension to Gödel's incompleteness theorem, that one can never prove a given string to be random. On the other hand, one can show that a string is nonrandom, simply by discovering a short algorithm to generate it.

It is in the realm of complex systems, and especially biology, that algorithmic complexity theory promises to provide major conceptual advances. Chaitin [4] has proposed a rigorous definition of the degree of organizational complexity of a system. Algorithmic complexity is not a suitable measure on its own, because it would assign a high complexity to a random state of a gas, for example. A gas is disordered, but it is not organized. What is needed is a definition that assigns low organisational complexity to systems that are either highly ordered or highly disordered (see Figure 6).

Bennett [5] has proposed a different definition of organization, called the *logical depth* of a system. This is the "difficulty of generating a description of the system from the shortest algorithm". A system with high organizational complexity, such as a living organism, would require a long and elaborate computation to describe it. This reflects the long and elaborate sequence of steps in the evolution of the organism. On the other hand, we can describe a crystal by a short computation from a simple algorithm. But what about a random gas? As we have seen, a short program cannot generate an algorithmically random state at all. We can do little better than take a description of the actual state, translate it into computer language, and use it as the algorithm: "Print that". This algorithim is quite short and it implies that a gas, like a crystal, has low *logical depth*.

Since *logical depth* is, roughly speaking, identified with the time required to compute the message from its minimal algorithmic description, we here again find a difficulty: in the Gödel sense the choice of the minimal description cannot be decided!

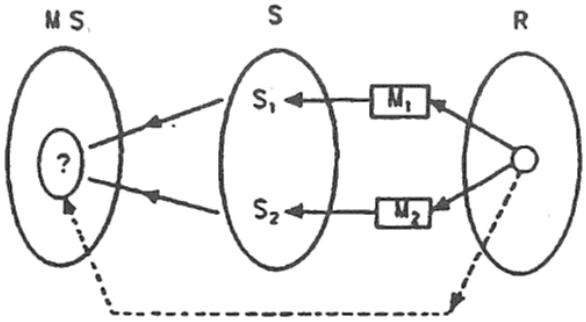

Figure 10. Mapping of a piece of reality, R, into the space of scientific symbols, S, through different models associated with different measuring procedures, M. Successive mapping of S into MS, in order to compare different theories referring to the same reality. Complexity means different irreducible models.

By identifying *logical depth* with the complexity of a physical entity, Bennett [5] appeals to the computational view of physical processes, in which physical processes are viewed as equations for computing specified by the laws of nature. The solar system can, in this view, be seen as an analogue computer which solves Newton's equations. We may begin with a very elementary set of rules or algorithms to do such a computation (the "minimal algorithmic description"), like Newton's laws for the solar system or the rules of molecular combination in the case of living systems. The *logical depth* of an object, its complexity, is measured by how long it takes a computer to simulate the full development of that object beginning with the elementary algorithm and taking no short cuts. Complexity, in this sense, is a measure of how hard it is to put something together starting from elementary pieces. In Section 6 we will return to this point of view, and we will see its limitations.

5.3. *Colloquial Definition*

From the difficulties we have encountered trying probabilistic and computational definitions of C, we have realized that a phenomenon is complex when it cannot be described by a unique model. We could then colloquially define C as the number of irreducible models of that event.

Once again, we come across a Gödel difficulty: we cannot decide how many irreducible models of a given situation the formal apparatus (symbols and laws) of physics provides. Furthermore, is there any way of comparing these models, or are they incommensurable as Kuhn maintains? The question requires a metasymbolic space MS (Figure 10) where we can compare the different models classified in S, depending on different sets of measuring apparatus, M_i, applied to the piece of the reality R under observation.

If the problems of the confrontation of different models is not set at the level MS, it may require a meta-meta-level and so on *ad infinitum*! The comparison can be made by introducing extrascientific criteria, but where do

they come from? If they are imposed as "values", then our scientific endeavour is contaminated by a highly subjective procedure.

On the other hand we should expect that MS considerations would be directly motivated by reality, through a pre-linguistic (pre-formalized) act of knowledge (dashed line from R to MS in Figure 10). We shall deal with this matter in Section 7.

6. PHYSICS AND COMPUTATION: UNDECIDABILITY AND INTRACTABILITY

This section is based on a paper by Wolfram [8], who analyses the limits to the descriptive and predictive power of physics. There is a close correspondence between physical processes and computations. Theoretical models describe physical processes by computations that transform initial data according to algorithms representing physical laws.

The behavior of a physical system may always be calculated by simulating explicitly each step in its evolution. Much of physics has, however, been concerned with devising shorter methods of calculation that reproduce the outcome without tracing each step. Such short cuts can be made if the computations used in the calculation are more sophisticated than those that the physical system can itself perform. Any computations must, however, be carried out on a computer. But the computer is itself an example of a physical system. And it can determine the outcome of its own evolution only by explicitly following its full course. No short cut is possible. Such computational irreducibility occurs whenever a physical system can act as a computer.

Computational reducibility may well be the exception rather than the rule. Most physical questions may be answerable only through irreducible amounts of computation. Those that concern idealized limits of infinite time, volume, or numerical precision can require arbitrarily long computations, and so be formally undecidable. Universal computers are as powerful in their computational capabilities as any physically realizable system can be, so that they can simulate any physical system [Church–Turing thesis]. This is the case if in all physical systems there is a finite density of information which can be transmitted only at a finite rate in a finite-dimensional space. No physically implementable procedure could then cut short a computationally irreducible process.

Let the amount of information specifying an instance of a problem be n. One may then distinguish several classes of problems. The first, denoted P, consists of those problems such as arithmetical ones which take a time polynomial in n. The second, denoted PSPACE, consists of those that can be solved with polynomial storage capacity, but may require exponential time, and so are in practice effectively intractable. A final class of problems, denoted NP, consists in identifying among an exponentially large collection of objects, those with some particular, easily testable property. A computer

that could follow arbitrarily many computational paths in parallel could solve such problems in polynomial time. For actual computers that allow many paths but only in a bounded way, we suspect that no general polynomial time is possible.

This suggests a possible interpretation of mental behavior as a computing strategy which readjusts its paths depending on the context, that is in an adaptive fashion, so that their number is virtually unlimited. This may be what philosophers call intentionality of knowledge. Such an interpretation will be developed in the next section.

The structure of a system need not be complicated for its behavior to be highly complex, corresponding to a complicated computation. Computational irreducibility may thus be widespread even among systems with simple construction. Cellular automata (CA) provide an example. A CA consists of a lattice of sites, each with k possible values, and each updated in time steps by a deterministic rule depending on a neighborhood of R sites. CA serve as discrete approximations to partial differential equations, and provide models for a wide variety of natural systems.

One may ask whether the pattern generated by evolution with a CA rule from a particular seed will grow forever, or whether it will eventually die out. This is analogous to the problem as to whether a computer with a particular input will ever "halt", that is, stop without an additional command; this "halting" problem was shown by Turing to be equivalent to the Gödel theorem on decidability of a formal theory. If the evolution is computationally irreducible, then an arbitrarily long computation may be needed to answer this question. One may determine by explicit simulation whether the pattern dies out after any specified number of steps, but there is no upper bound on the time needed to find out its ultimate fate. Simple criteria may be given for particular cases, but computational irreducibility implies, in general, that no short cut is possible. The infinite-time limiting behavior is formally undecidable. No finite mathematical or computational process can reproduce the infinite CA evolution. The fate of a pattern in a CA with a finite total number of sites N can always be determined in at most k^N steps.

Irreducible computations may be required not only to determine the outcome of evolution through time, but also to find possible arrangements of a system in space. Quantum and statistical mechanics involve sums over possibly infinite sets of configurations in systems. To derive finite formulas one must use finite specifications for these sets. But it may not be possible to decide whether two finite specifications yield equivalent configurations. Similarly, it is not possible to decide what is the simplest such model that describes a given set of empirical data.

7. MAN AND COMPUTERS. TOWARDS AN ONTOLOGY OF COMPLEXITY

In European cultures there is a period called Humanism, which stands as a miracle between the Medieval fundamentalism of a theological approach to reality, which enslaved any other human endeavour, and Modernity, whose dawn is marked by radical divergences of opinion as well as bloody religious wars. In the few decades between 1440 and 1510 people like L.B. Alberti, Marsilio Ficino, Erasmus and Thomas More realized a marriage between Christian revelation and classical philosophy. Old texts were rediscovered and read in their original languages. It was compulsory to study and compare Latin, Greek and Hebrew texts. This philological effort put in evidence the ambiguities of any language as indicated in Figure 6, paving the way towards the successive Galilean formulation of a simplified unique language (Figure 7). A synthesis of this period is represented by a 1486 text, "*De Hominis dignitate*" by G. Pico della Mirandola. In this "manifesto" for Humanism, the Creator endows man with liberty saying:

> I have given to you, Adam, neither a determined place nor a proper aspect nor any other prerogative, because that place, that aspect, that prerogative that you wish, by your choice and advice you achieve and maintain.

This text provides two characteristics of the human being: liberty of decision, together with the confidence of being part of a design.

Modernity appears as an anthropological bifurcation with two paths: either man has liberty, but he is the outcome of chance, not part of a design (nihilism); or man is part of an overall determined natural plan, but he is just a machine, a biological computer without any free will (mechanism). Both paths seem to be the result of the impact of science upon philosophical thinking. Such a strong relationship between science and philosophy may appear at variance with the proclaimed gap between the two cultures. In fact such a gap arose only in the middle 1800s with positivism. H. Spencer gave an extensive definition of science as a collection of tools which increase our power of investigation by coping with the limitations of our senses. This definition implies a neutral science which does not take issue on philosophical problems. As a result, any attempt at interdisciplinary relations reduces to a "concordism" between the results of science (e.g. Big Bang, evolution) and a given "Weltanschauung". I am strongly against such a view, and share the viewpoint of Cassirer [3] that science in an intensive sense is the "how" and "why" reality looks as we describe it. Intensive science interacts strongly with philosophy by a critical analysis of its own linguistic procedures, rather than by a concordism in the use of its own findings. Indeed, Galileo and Newton, as well as the other scientists of the 1600s considered themselves to be philosophers.

The crisis of scientific language, outlined in Section 4, has challenged the neopositivistic pretension that science is the only language able to speak about reality. But would the result be a pure scepticism? In other words, do

the limitations of Section 4 destroy trust in the power of science and hence make us despair about our ability to grasp reality (nihilism)?

Not at all! In fact we have seen in many instances that any cognitive strategy adapts successfully to the chain of changes of a complex event (see again Figure 5b not only as a morphogenetic strategy but also as a cognitive strategy). The criticism of Section 4 means in fact that science is an *open* language which has to readjust itself whenever this is required by reality, in order to avoid undecidable or intractable questions.

It has been proposed (Webb [9] and Hofstadter [10]) that procedures by finitude, that is, by the handling of a finite number of possibilities, avoids Gödel undecidibility. Thus we might foresee a science and a human knowledge as well, based on finitude, and hence not limited by Gödel. In fact, the discussion of Section 6 has shown that quantum and statistical physics are not limited to procedures by finitude; furthermore, intractability, that is, a prohibitively long processing time, is peculiar of most problems.

On the contrary, we experience that man decides, and decides within a short time. To use the computer metaphor, man does not act as a "braid" [10] of self-referential routines implying one another, and hence trapped by undecidable questions, but he rather changes his rules, adapting them to the external bifurcations of an organizing reality (Figure 5b). Since this is my main thesis, as stated in the Introduction, it may be worthwhile to give more details related to my research on chaotic systems (see e.g. Arecchi and Arecchi [11] and Arecchi [12]). In Section 3, speaking of pattern recognition, it was mentioned that neural models of the past decade considered the brain as an extended dynamical system whose energy landscape has many valleys corresponding to stable fixed points. By a suitable codification each of those points represents a useful pattern. Recognition of a new pattern means providing energy to escape from one attractor (old pattern) and enter another one (new pattern to be memorized). This requires a given amount of energy and time.

On the other hand, physiologists (see Skarda and Freeman [13]) have recently shown evidence that the electrical activity of neurons is chaotic in the absence of an input, and that it regularizes in the presence of a stimulus. Notice that, as to information, deterministic chaos is not as hopeless as thermodynamic disorder. In this latter case not even a Maxwell demon could make any further program out of the system, since he needs the same amount of information he wants to handle (Gabor [14]). In the former case, even though the trajectory slides downhill (Figure 6), loss of information takes some time (the corresponding rate we called K). Thus, there is room either to stabilize chaos (Ott et al. [15]) or to synchronize another system to a chaotic one (Pecora and Carroll [16]). In the presence of chaos the Maxwell demon is efficient!

Combining the physiological evidence with the synchronization strategy suggests a new cognitive paradigm: the most efficient neural network has chaotic activity, and it synchronizes itself to an external pattern. Synchro-

nization means that the rules are not fixed, as in a Turing machine (which gets stuck in undecidable or intractable questions), nor are they arbitrarily modified by the observer (pure subjectivism of the trascendental orderings as proposed by Kant), but they are motivated by the observed reality. Thus, any symmetry breaking in the organization tree of an external reality induces a corresponding rule modification in the recognizing system.

One could object that such a superiority with respect to a Turing machine is common to any biological cognitive system (even the immune systems which build antibodies against a virus!). This is true, but, as Blaise Pascal said, we are the only living system which does not limit knowledge to a survival strategy, but we make a challenge of the fact that many levels of order exist (the different layers of a tree such as Figure 5b), and hence we build an ontology of complexity. In other words, man is not the measure of all things, as Protagoras claimed, but he is rather measured by things. I conclude with a citation from Thomas Aquinas (In I Sent., XIX, V, ii):

Therefore our science does not measure things, but is measured by things, as Aristotle said in the X book of Metaphysics.

Università di Firenze and
Istituto Nazionale di Ottica
Florence, Italy

REFERENCES

1. Haken, H., *Synergetics: An Introduction*, 3d ed., Springer-Verlag, Berlin, 1983.
2. Basti, G., *Il rapporto mente-corpo nella filosofia e nella scienza*, Edizioni Studio Domenicano, Bologna, 1991.
3. Cassirer, E., *Determinism and Indeterminism in Modern Physics*, Yale University Press, New Haven, 1956.
4. Chaitin, G.J., *Algorithmic Information Theory*, Cambridge University Press, Cambridge, UK, 1987.
5. Bennett, C.H., "Dissipation, Information, Computational Complexity and the Definition of Organization" in D. Pines (ed.), *Emerging Syntheses in Science*, pp. 215–234, Addison Wesley, New York, 1988.
6. Atlan, H., "Self Creation of Meaning", *Physica Scripta* 36, p. 563, 1987.
7. Anderson, P.W., "More is Different", *Science* 177, p. 393, 1972.
8. Wolfram, S., "Undecidability and Intractability in Theoretical Physics", *Physical Review Letters* 54, p. 735, 1985.
9. Webb, J.C., *Mechanism, Mentalism and Metamathematics: An Essay on Finitism*, D. Reidel, Dordrecht, 1980.
10. Hofstadter, D.R., *Gödel, Escher and Bach: An Eternal Golden Braid*, Basic Books, New York, 1979.
11. Arecchi, F.T. and Arecchi, I., *I simboli e la realtá*, Jaca Book, Milano, 1990.
12. Arecchi, F.T., "Why Science Is an Open System Implying a Meta-Science", in C. Wasserman, R. Kirkby and B. Rordorf (eds.), *The Science and Theology of Information*, pp. 108–118, Editions Labor et Fides, Genève, 1992.

13. Skarda, C. and Freeman, W., "How Brains Make Chaos in Order to Make Sense of the World", *Behavioral and Brain Sciences* 2(10), pp. 161–195, 1987.
14. Gabor, D., "Light and Information", in E. Wolf (ed.), *Progress in Optics*, Vol. 1, North-Holland, Amsterdam, 1964.
15. Ott, E., Grebogi, C. and Yorke, J., "Controlling Chaos", *Physical Review Letters* 64, p. 1196, 1990.
16. Pecora, L.M. and Carroll, T.L., "Synchronization in Chaotic Systems", *Physical Review Letters* 64, p. 821, 1990.

Chapter VII

Indeterminism and Nonlocality

John S. Bell

The following contribution is the transcript of a talk presented on 22 January at CERN on invitation of the Center for Quantum Philosophy, Geneva. The text is literally taken from the video-recording of this quite informal colloquium. It has the spontaneity and originality of the spoken word, but misses the final touch and the polishing of a written contribution. The late Prof. Bell could not revise the text himself. On the other hand the editors are not so sure that changes by them would improve the document. Instead we ask our readers to apologise the imperfections, which in some cases are also due to the not always perfect voice recording of the session. We would like to thank R.W. Nowak who did an excellent job in transcribing the text.

Introductory remarks

by Antoine. Suarez,
Center for Quantum Philosophy Geneva

About indeterminism we all as students have heard a lot. On the other hand there are still best-sellers of famous physicists that do not include the term nonlocality in the index. Every man who reads the rigorous book of John Bell *Speakable and Unspeakable in Quantum Mechanics*[1] wants to ask the author "Are both indeterminism and nonlocality necessary for a coherent physical theory or only one of these principles - and in this case which? - or neither of them." We are very grateful to Professor Bell for having agreed to discuss with us today this question. The talk will be registered - those who are interested in a copy of the tape can request it at the address of the Center. Please Professor Bell.

The Talk

1. Introduction

Imagine that this is a metal grid made of some very tough metal, and I am shooting bullets at it. Sometimes the bullet will hit the metal and be stopped, and sometimes it will go through the hole. Now imagine that that is made on a microscopic scale, so that I cannot see with my eyes or even with my instruments that it is made like that, that some parts are transparent and some parts are opaque. And what would happen when I shoot my little bullets would be sometimes they would go through, and sometimes not, and I would have a situation of **unpredictability**[2]. My experiment would sometimes give one result 'pass', and sometimes another result 'stop'. And a classical physicist would understand this situation by saying that the unpredictability arises from the **lack of control**. That although I think I am doing the same experiment each time, I am not doing the same experiment each time, because there are little things which I am not reproducing sufficiently accurately from one time to the next. And for a

[1] J.S. Bell:, *Speakable and unspeakable in quantum mechanics,* Cambridge University Press, Cambridge (1987)
[2] Words written down by Prof. Bell on the board are typed with bold characters

classical physicist, or let's say nineteenth century physicist, unpredictability simply meant lack of control.

So the unpredictability was in no way synonymous with **indeterminism**. For that classical physicist the unpredictability was a reflection of human weakness, and nature might very well be perfectly deterministic.

Now we have a situation a bit like that when we think of the photons streaming out from this lamp and coming up here, and being reflected towards the blackboard. If I put in some absorbing material, not all the photons get through. And if you did the experiment one photon at a time, you would find that sometimes the result of the experiment would be 'pass', and sometimes it would be 'stop'. And it would be quite unpredictable which was the case. A classical physicist would say: 'Aha, we don't get the same result every time because we don't know everything, and we don't control everything. There is something out of control in this experiment'.

But the founding fathers of quantum mechanics did not say that, they said we are confronted here with the case of **indeterminism**. Now, when I read history it always amuses me that the founding fathers could come so confidently to that conclusion. Up to then the tradition in physics was that if something was unpredictable, it was because it had not be fully controlled. And for me the natural assumption would be that the situation with photons is like this situation here.

And that was also the natural photon for Einstein, who said: '*God does not play dice'*. For him it was always an article of faith that the nature is lawful, and if we could not control it, that was our fault. And so he contemplated that the description given by quantum mechanics was **incomplete**. That it should be supplemented by hypothetical **hidden variables**, which we do not know just by inspection, but we will have to make theories about. And that when these hidden variables were found, at least conceptually, that we would again see that nature was perfectly lawful, that we would have **determinism**.

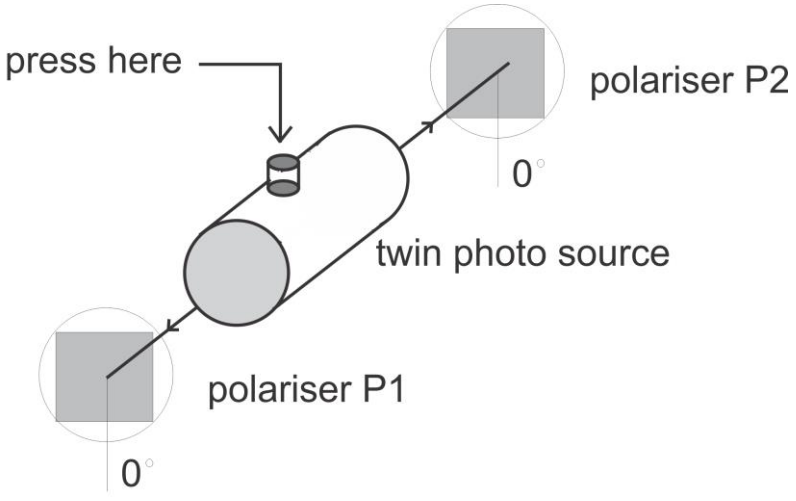

Fig. 1. Experimental set-up for an EPR experiment with twin photons. The polarisers are set parallel at an angle of 0° with respect to the vertical.

Now, the piece of absorber that they put there (see Fig. 1) is not just any old material, but it's actually polarising material. And the photons that get through a polarising material have the property that they cannot get through a second piece of the same material, if it is oriented at right angles to the first. Whereas if it is oriented parallel to the first, it gets through again.

So that is a property of photons, that they have polarisation, which can be detected by such pieces of material, and that plays a role in the sequel to what I have to say.

Now at first this was simply a working hypothesis of Einstein, and for a few other die-hard classical nineteenth century physicists. But in 1935 he invented an extremely powerful argument, for this position, based on another hypothesis which most people who have not met these phenomena before would accept; the hypothesis of no **action at a distance**, which is sometimes called **local causality** or just **locality**. And he said that there are situations where this hypothesis implies determinism. So in this argument determinism was no longer a hypothesis, but a theorem, but with locality as the axiom.

locality ⇒ determinism

2. The EPR Gedanken-experiment

Now the situation that he[3], and later on in a more simple example David Bohm, the situation which he paid attention to was this: You can make, in a way that I will not go into in any detail, twin photons. You can have a source, which when you press the button, emits a photon going into this diagram, and another photon coming out (see Fig. 1). And you can see whether these photons do or do not pass through pieces of polarisation detecting material. And the twin photon source has the character, that whenever the photon passes here its twin passes there, and whenever this one does not pass here, its twin does not pass there. That's a feature of this particular source. Such sources exists according to quantum mechanics, and have been built experimentally. And that is true no matter which way these things are oriented, provided they are oriented in the same way. And this is a very powerful argument against the idea that the passage of this photon through that polariser is a matter of pure chance. If it was determined by a die being cast over here, this was determined by a die being cast over there, how could they possibly always agree either to pass or not to pass at a given location.

The situation here is rather like one which is extremely interesting to biologists, when you have identical human twins. I believe that all babies are born with blue eyes, but some of them may turn brown, and you might think that was a matter of chance. Interesting to take bets on whether brown or blue was going to come. But it's not a matter of chance. We know that from the phenomenon of identical twins. But even if they are separated before their eyes acquire their final colour, if one turns brown so does the other, and if one turns blue, so does the other. And we don't think "Oh, what a mysterious long distance correlation", we say that's genetics.

And similarly here the reasonable assumption, for any reasonable person, is that we are concerned here with the case of genetics. That these photon unknown to us carry some small packets of information, the same information in the two cases, which dictates that when they meet the same circumstances they will behave in the same way. And that was the Einstein-Podolsky-Rosen (EPR) argument, from no action at a distance to determinism. If you didn't want to accept that hypothesis of determinism, I think you are obliged to accept that in some way things can agree at a distance, without any explanation. And that for Einstein would have been an objection about action at a distance.

3. The Bell inequality

[3] A. Einstein, B. Podolsky and N. Rosen, *Can quantum-mechanical description of physical reality can be considered complete?*, Phys. Rev. **47**, 777-780 (1935).

Now, for Einstein then, quantum mechanics was incomplete. The formalism that we are taught is not the whole story, the 'genes' are missing. It would be like biology without genes, where things just happen by chance, and it is quite mysterious that sometimes they happen in the same way to identical twins. He supposed therefore, that there must be some variables that we don't know about, which are not under our control, analogous to the genes. And they explained these correlations at a distance, without causation at a distance. Because the causation was transmitted from the source to these objects. Einstein said originally "*God does not play dice"*, but in this argument it was no longer an assertion by itself, it was an inference from the assumption of no action at a distance. And in later years, Einstein was much more concerned about the fact that quantum mechanics was not explicitly free from action at a distance, than he was concerned about determinism as such.

Now it is a great irony that this argument, which I think was an extremely powerful argument, and which the people of today should have accepted, that this argument boomeranged on him. And became an extremely powerful argument against his own position. Now that came about as follows:

We did that experiment, I described the results of that experiment with parallel polarisers (see Fig. 1). But of course you can also do it with off parallel polarisers. Now, let's start with the parallel case. And then suppose we turn this through thirty degrees. We will no longer get perfect agreement. Sometimes a photon will pass here, and its twin will not pass there, and vice versa. And quantum mechanics gives a formula for the degree of discordance that creeps in when you turn such a device. Now we could turn the other device instead, and there would be some discordance, and we could turn both of them, and there would be another discordance. Now there's a simple relation between those three situations that I have described, turning this one, turning that one or turning both, and it's the following.

Let N be the number of cases in which you have disagreement on the two sides, a yes and a no, or a no and a yes. When the two polarisers are both at zero degrees there is no discordance. It'll always be yes yes, or no no. Never yes no, or no yes.

$$N(0°,0°)=0 \qquad (1)$$

If I now consider the case where both are turned, one minus thirty degrees and the other plus thirty degrees, then it's easy to see, that on the genetic hypothesis, that this number must be less than or equal to the discordance if just one turned, plus the discordance if only the other turned.

$$N(+30°,-30°) \leq N(+30°,0°) + N(0°,-30°) \qquad (2)$$

And that is indeed easy to see because the programming of this device must be such that a yes, some yes's are replaced by no's, and some no's by yes's, when I turn it. Similarly when I turn that, some no's are replaced by yes's, and some yes's by no's. But now in this case every change in the programming introduces a discordance, because originally I had perfect concordance, yes yes and no no.

So every change here introduces a discordance, every change there introduces a discordance. But not every change here introduces a discordance, because there may be compensating changes in the program on the two sides, a yes may turn to a no, or the no may turn to a yes. And again you will have agreement, and therefore the amount of discordance

with the two movements is less than or equal to the amount of discordance with both movements.

Now, if you calculate that according to quantum mechanics, that just isn't so. I don't remember, I think I remember the exact numbers, that this turns out to be 3/8 as a fraction of the total number of trial, that 3/8 of them will be discarded. And here it is 2/8, and here it is 2/8. That's not right. It must be something like that, 1/8 and 1/8.

$$N(+30°,-30°) = 3/8$$
$$N(+30°,0°) = 1/8$$
$$N(0°,-30°) = 1/8$$

Therefore with Eq. (2)

$$3/8 \leq 1/8 + 1/8 \text{ !!!}$$

It just is a fact that quantum mechanical predictions and experiments, in so far as they have been done, do not agree with that inequality. And that's just a brutal fact of nature. The genetic hypothesis, which seems absolutely compelling for parallel devices, simply doesn't work for nonparallel devices. You can't get away with the genetic hypothesis, and therefore the Einsteinian argument fails. No action at a distance led you to determinism, in the case of parallel polarisers, but determinism, in the case of off parallel polarisers, leads you back to action at a distance:

no action on a distance (polarisers parallel) $\Rightarrow$ **determinism**

determinism (polarisers nonparallel) $\Rightarrow$ **action on a distance**

4. Action on a distance

Now, in my opinion, in answer to the question that you posed at the beginning, I don't know this phrase is too strong and active an assertion, I cannot say that action at a distance is required in physics. But I can say that you cannot get away with no action at a distance. You cannot separate off what happens in one place and what happens in another. Somehow they have to be described and explained jointly. Well, that's just the fact of the situation; the Einstein program fails, that's too bad for Einstein, but should we worry about that? So what?

Now, there are three replies to the question “So what?” One is that the whole idea of action at a distance is very repugnant to physicists. If I were speaking for an hour..., I would bombard you with quotations from Newton, and Einstein, and Bohr, and all the other great men, telling you how unthinkable it is that by doing something here, we can change the situation in a removed place. I think that the founding fathers of quantum mechanics did not so much need Einstein's arguments about the desirability of no action at a distance, as they looked away. The whole idea that, either there might be determinism, or action at a distance, was so repugnant to them that they looked away. Well that's tradition, and we have to learn in life sometimes to learn new traditions. And it might be that we have to learn to accept not so much action at a distance, but inadequacy of no action at a distance.

There are two more professional reasons for being discontented with the situation. Now one is relativity. According to relativity, the notion of simultaneity is relative. And events which are simultaneous for one observer are not simultaneous for another. So it does not make sense

for very distant situations, to say that one event has occurred before or after another. So if we allow the result at one of these experimental set-ups to depend on what an experimenter does at the other, we have a puzzle, because we would not like what he does here to have an effect there, before it is done here. But if I say that this is affecting that, I can find some observer for whom this comes after that. So if I set up a traditional causal model, which the cause effects are allowed to be nonlocal, in the sense of propagating instantaneously over large distances, in some frame of reference the cause will come before the effect.

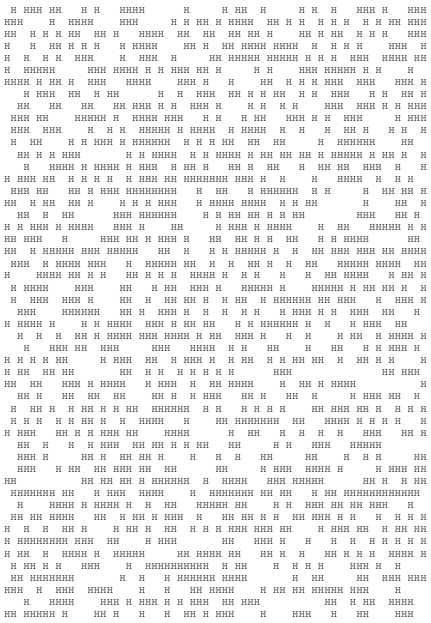

Fig. 2. Result of computer simulation of a random series of heads 'H' and tails (blank)

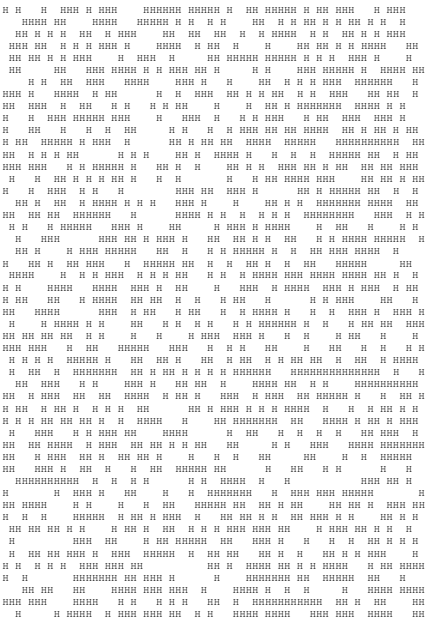

Fig. 3. Inverted display of Fig.2: heads (blank), tails 'H'. In some places the random code is changed.

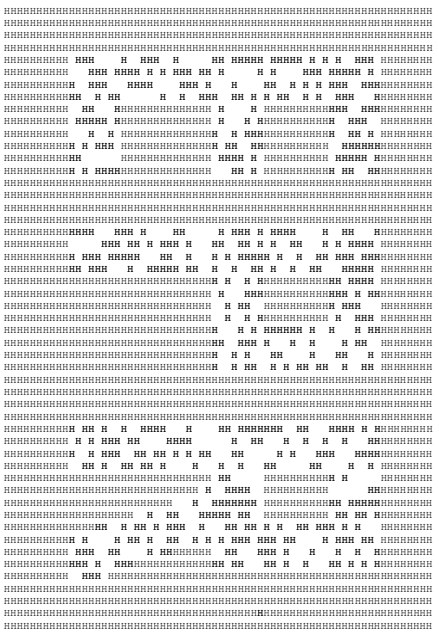

Fig. 4. Superposed images of Figs. 2 and 3 (after B. Julesz)

So we have to be a bit more subtle than that somehow. I have to find some way out of this situation, which allows something somehow to go from one place to another, very quickly, but without being in conflict with special relativity. And that has not been done. We have the statistical predictions of quantum mechanics, and they seem to be right. The correlations seem to cry out for an explanation, and we don't have one.

The other reason is no signals. It is a fact that I cannot use whatever this nonlocal connection is to send signals. When you look at what quantum mechanics predicts, it predicts so long as you look at just one side or other of this experiment, you will simply have no information about what is happening in the other place. No matter what that other fellow does with his equipment, you will not notice anything funny happening in your side. As an analogy of that, I could say, supposing we were tossing coins, I here and one of you people down here. And supposing I had the power to say that your coin will turn an extra time before it falls on the table. Now you are looking at your coins and you see heads tails heads tails. And you don't know when I have exercised my power to turn it once more, because you didn't know whether it was going to fall heads or tails. So we have the curious situation, that to explain the correlations between my results and yours, we have to invoke some such mysterious power. But it is one which I absolutely cannot use to send you a message. I got here a demonstration of that. This is a computer simulation of such an experiment in which people are calling heads and tails. And when it comes up heads I have written ‘H’, and for tails I have written blank, so that you can see it from where you're sitting. And they're a whole series of random heads and tails, you can see it there (Fig 2). Now at some point I exercised my power. My remote power to turn a head to a tail. And here is the result (Fig. 3). So somewhere in there I have done something to the random code. I exchanged heads for tails, but you absolutely cannot see that. This message, as far as you're concerned, is as meaningless as the other. It's only if you have two copies of that, that you can compare it, that you can get something (Fig. 4).

That curious situation has inspired a musical composition. There is a musical composition called `The Bell's theorem blues'. I'm not going to sing it, I'll say the words:

Doctor Bell say we’re connected,
He called me on the phone,
But if we’re really together baby,
How can I feel so all alone?

5. Conclusion

And that is the dilemma. We are led by analysing this situation to admit that in somehow distant things are connected, or at least not disconnected. And yet we do not feel that we are connected. So as a solution of this situation, I think we cannot just say ‘Ohoh, nature is not like that’. I think you must find a picture in which perfect correlations are natural, without implying determinism, because that leads you back to nonlocality. And also in this independence as far as our individual experiences goes, our independence of the rest of the world is also natural. So the connections have to be very subtle, and I have told you all that I know about them. Thank you.

Discussion

Note:

In the following, a selection of the 60 minutes discussion is given. Professor Bell's speech is presented in plain typeface, all other contributors to the discussion in *italic*. A new section is started for each person talking, with comments made by others presented in parentheses, '(....)'. The typeface rule is maintained even within parentheses. Speech which could not be understood is represented as ' ..... ', this often indicates more than one word. The sections are numbered and run up to 113. We reproduce a selection of sections, which we think are relevant for the subject of this book.

10....., *we have learnt, that there always will be questions in arithmetic that could be undecidable. That is unsolvable by mathematical reasoning. Nonlocality seems to suggest that there are also questions which could in principle be answered by way of mathematical reasoning, but that in fact cannot be answered because information about what is going on in regions far away from us is not available to use. What do you think? Is correct to think this way?*

11. I don't know, I mean you gave your thesis very briefly, and so I cannot judge .. whether you are correct or incorrect. But I do not myself think it is right to use the words 'Gödel theorem' in this discussion. I don't see a connection, except insofar as both subjects are very difficult to us. But, I think you see, with Gödel's theorem we have some kind of permanent boundary to the possibilities of systematic reasoning. But here, I think we have a temporary confusion. It's true that it is sixty years old, but on the scale of what I hope will be human existence, that's a very small time. I think the problems and puzzles we are dealing with here will be cleared up, and we will look back on them with the same kind of superiority, our descendants will look back on us with the same kind of superiority as we now are tempted to feel when we look at people in the late nineteenth century who worried about the ether. And Michelson-Morley .., the puzzles seemed insoluble to them. And came Einstein in nineteen five, and now every schoolboy learns it and feels .. superior to those old guys. Now, it's my feeling that all this action at a distance and no action at a distance business will go the same way. But someone will come up with the answer, with a reasonable way of looking at these things. If we are lucky it will be to some big new development like the theory of relativity. Maybe someone will just point out that we were being rather silly, and it won't lead to a big new development. But anyway, I believe the questions will be resolved. I do not think the questions raised by Gödel will be resolved.

12. *It seems that from what you have shown by that, the concept of an isolated system, is appealing. Nothing is isolated any more from the rest of the whole of the universe. In general I know that in classical physics you can think that influences from outside, the outside world, you can reduce to small as you want. Is that still the situation?*

13. Well, conceptually you could reduce them as small as you liked, but of course we can't, we can't declare that Jupiter does not exist. And so long as it does exist it will affect the orbit of the Earth, and even the way this piece of chalk falls a little little bit. The chalk falls only roughly the way Galileo said. And to some extent it feels the gravitational field of Jupiter, and of the Crab Nebula, and God knows what, and I can't switch those off.

14. *But how far are we allowed now to describe anything without taking the whole universe in to account?*

15. Well first of all, strictly speaking, you never were allowed to do that. It was always a rough procedure, that for rough work we can neglect much of the universe. But when we looked at what our equations said, Newton's equations include Jupiter's gravitational field acting on this bit of chalk. So that when we looked at what our equations said, we had to remember that it was ... And I think it's about the same. If you are interested only on a limited part of the world, in quantum mechanics you use the so called density matrix, which kind of averages over the rest of the world. And then the density matrix has an equation evolving for itself, and not mentioning the rest of the world. You still have to admit that if you wanted as much precision as the theory could possibly yield, you just have to put things ... What I think is novel is that Einstein gave us a way of switching off the rest of the world outside the light cone. You could say that yes the rest of the world is going to have an effect, but that effect will not arrive before light could propagate. So, that was a way of dividing the world into bits which are relevant, and bits which could not be relevant. And that we don't have any more.

16. *(That is a very important affirmation.) I'm sorry, does that mean that you say relativity and quantum mechanics are not compatible?*

17. No no I can't say that, because I think somebody will find a way of saying that they are compatible. But I haven't seen it yet. For me it's very hard to put them together, but I think somebody will put them together, and we'll just see that my imagination was too limited. Well, as the people in that department work at present, they are not coming to this question, because the superstring is still formulated within traditional quantum mechanics, and you still have the superposition principle which is maybe the root of all these things. But it could be that as they go further into that, they find that it just won't work along the traditional lines, and at some point they'll have to give up the superposition principle. .. in that direction, and occasionally see papers from over there called `How worm holes reduce the wave function.' Haha.

18. .. *that there are no hidden variables. ....*

19. Well I don't know that there are no hidden variables. What we said there was that hidden variables cannot restore locality. But there was still the consideration that I began with, that the classical physicist still thinks repeating the experiment gives different answers. He has not quite repeated the experiment. And you can have nonlocal determinism, nonlocal hidden variables. So I could imagine the situation where we do have hidden variables, and we do have again determinism, but not predictability. Because we won't be smart enough to get things with sufficient precision. We'd be able to gain that determinism, and it would be nonlocal. And we see that in some way distant things were connected, and yet not in a way that we could get hold of to send messages.

20. *But if there are hidden variables, somehow we are back in classical mechanics. ..*

21. Well it depends on what you call classical mechanics. You won't be back with the mechanics of Einstein. With the light cone neatly separating what is relevant and what is irrelevant. But you would be back with classical mechanics in the sense that you would have some equations, presumably integral rather than differential equations, and the equations would tell you what is going to happen. Whereas in quantum mechanics it's perfectly obscure whether we have a theory like that. In fact we don't have a theory like that.

22. *Well, .. the principle of free experimentation. (Yes.) That seems important for us. (Right.) Then Bell theorem holds. Is it possible to formulate this experiment in this way: In a world, in which there are human free actions, either action at a distance, in this sense, exist or we must accept that correlated events will not have any rational explanation*.

23. I think that's alright. (*Do you think?*) No, I would not sign that yet. (*laughter*) But certainly it sounds OK what you said. That's the kind of way I think also.

24 *Do you think it is important to say that Bell theorems has something to do with free experimentations*.

25 I eh, that's not what you said. (*Yeah, exactly*) .. Because you in your, in your (*Yeah*) what you said the free action was the hypothesis, and now you're trying to turn it into a theorem.

26. .... *I think. If one admits the principle of free experimentation, the Bell theorems holds. And my question is, is possible to formulate this result in this way. In a world in which there are human free actions, either action at a distance exists or we must accept that correlated events do not have any rational explanation*.

27. *What means free experimentation and free human action?*

28. *Well, free experimentation is the fact that there is, in this sense of what Professor Bell says in his book, no super determinism, in the kind that when I am arranging the setups, this action is not determined from the beginning of the world.*

29. That's correct. It comes into the analysis. When I turn one of these polarisers, I assume that I can consider the same hidden variables in the turned position or in the unturned position. But if my choice was itself determined by the hidden variables, the argument would fail at that point. So I have assumed that there is something outside my field which is quite free, which is not dictated by the parameters, the variables in the theory. Now, it is a fact that if I give that up I have no theory. But I myself do not think it's a very essential point, in the following sense that you know that there are, if you have random number generators on a computer, you could have a computer here generating random numbers and another one other there, with a different program, generating random numbers. Strictly speaking, they are determined what comes out. But for somebody doesn't know the program, it's unpredictable what comes out. And they are self determined, so that they have nothing to do with the hidden variables that are determining whether the photons do, or do not go through polarisers. So I think that that is enough freedom, the freedom of random number generators. But that's not a theorem, because when you set a random number generator going you have to pick the program, and maybe your choice of program will be determined by the hidden variables, (*laughter*) in your experimental setup. So you have to be very far fetched to make this the escape plan.

32. *I have a question .. . .. . Is it correct to say that indeterminism is not a necessary condition for a coherent quantum theory?*

33. I would agree with that, yes.

34. *OK. But quantum theory is not incompatible with indeterminism. Also, quantum theory is compatible with indeterminism.*

35. I believe that quantum theory is compatible both with determinism and with indeterminism. (*With both?*) With both.

36. *Depending on locality or nonlocality*.

37. That's right. If it's deterministic. Well I think it has to be nonlocal, whether it's deterministic or indeterministic. I think you're stuck with the nonlocality. I don't know any conception of locality which works with quantum mechanics. So I think we're stuck with nonlocality. Whether we're stuck with determinism or indeterminism is another question. I know that if you didn't worry about Lorentz invariance, you can make explicit deterministic models which agree with all the experiments. But they're not Lorentz invariant. You have an ether, in there, and that's hard to swallow. It may be that Lorentz invariance plus quantum mechanics is incompatible with determinism, but I don't know that. That's a possibility.

40. *Because, I mean, I would think maybe all the trouble come from special relativity, which to my mind is anyway only a frame that doesn't contain any physics. For instance if you say that, well, the action is at a distance is not possible. Well, you mention then, with relativity one gets into trouble, one had the trouble with the ether. Special relativity has abolished the ether, but what is producing the action now? So that you can get only by field theories, where you then have field quanta, which in case of gravitation you need the graviton, which means you need, you have to quantise also gravitation. So I think to solve the problem one probably has to go to general relativity and not stick to special. Maybe al.., I don't know, but could be that all the difficulties would disappear if we would forget special relativity.*

41. Well, I certainly agree with you that we should be thinking in terms of general relativity. I think all the difficulties may disappear, but not at once. (*laughter*) There are quite a number of people who have tried to combine general relativity with quantum theory, never mind about hidden variables. It is very difficult indeed.

42. *But I thought the big success of supergravitation..and superstrings is that they show that in principle there exist theories which are mathematically alright.*

43. Well again they have over sold their message to the general public, I would say. When you press them, you find that whereas they are announcing a theory of everything, they only have theories of S matrix elements. That's to say, bodies come in from infinity, and they say nothing until they go out again at infinity. (*laughter*) But we don't refer to infinity. And in order to make a serious theory, including gravity, it must not be an S matrix theory but a field theory, which discusses things at finite times and places. And they can't do it. They have tried very hard and given up. So string theory is not in the state where you could bring it into this discussion.

58. *There is a word .. of language .. which confused me. When .. there is a word interaction, action at a distance, or cause, causality. I think in general this involves a force or a transfer of energy, through a signal. And then you say that is no signal, there is information involved but not energy transfer. Is it right?*

59. There is no energy transfer and there is no information transfer either. That's why I am always embarrassed by the word action, and so I step back from asserting that there is action at a distance, and I say only that you cannot get away with locality. You cannot explain things by events in their neighbourhood. But, I am careful not to assert that there is action at a distance.

60. *.. everything interacts with everything .. but is not interact ...*

61. Well let's say. Let's say, let's change it and say that you cannot account for anything without taking into account everything. And then I have avoided the word interaction. It's true, the word action brings up in your mind force. And one does not want that image to arise, because I don't know of any sense that I could attach to the word force in this context.

66. *Well don't you see the problem already there in classical physics, without having to .. ?*

67. Well you see the problem that things are not isolated. But in classical physics the light cones do determine the separation, which is a status in classical physics that you do not have in quantum physics. As soon as you permit free actions, or effectively free actions, in classical theory their consequences are confined to the future light cone. And that's a concept that we seem unable to fit in to quantum theory.

90. *The next question, another question of more philosophical nature. The question of mind and consciousness. .. some months ago Wheeler, in a talk in Sweden, said to me that he's convinced ... the fundamental role played by the observer. And Wigner also insists very much that man is a central element of the theory. Because of the fact of nonlocality, it seems that mind, .., which constructs the observed world. If a mind constructs the observed world, this mind could not be alone the mind of the human observer.*

91. If two and two are five, what is three and three? I simply do not follow Wheeler, and Wigner, in saying that atomic physics involves the human mind. I see not the slightest trace of the human mind in atomic physics. And here, I am with Bohr rather than with others, although many times I am against Bohr. Bohr insisted very strongly that the only observer that he was concerned with was the inanimate apparatus. He .. insisted that you could not separate atomic processes from the apparatus that was used to amplify those processes onto our scale. But then he insisted that whether somebody was looking at the apparatus or not doesn't matter a bit.

92. *Yes, but he says that this is all right because different observer will always see the same. This conception objectivity is wrong. But I think he eh, his opinion is that the phenomenon occurs in the moment in which I am observing.*

93. That is not Bohr. (*That is not Bohr*.) No. That may be .., and it may be Von Neumann and Wigner and .. and .. and Wheeler, and a whole (*But not Bohr*.) host of people, but it is not Bohr. And I think that Bohr was the sound person on this. It's an enormous extrapolation from the situation in physics, to the idea that being and mind is involved. Now, I can see the motivation for that extrapolation, 'cause when you say the apparatus is playing an essential role, you are asking `Well, how do we divide the world into systems, and apparatus, and the rest?'. And it's clear that any such division is a very shifty thing. So people look for somewhere that a division could naturally be made, they say `Ah! Between matter and mind!' And so they postpone the so called reduction of the wave packet, which is an element of ordinary quantum mechanics. Instead of that occurring in the apparatus, they say let it occur in mind. But that's a conjecture, an extrapolation. There is no evidence in physics for that, and as far as physics is concerned, it can all end at the apparatus level.

Chapter VIII

Nonlocality and the Principle of Free Experimentation

Paul Pliska
Center for Quantum Philosophy, Zurich, Switzerland

Abstract

We derive one of Bell's inequalities, one possible form of Greenberger-Horne-Zeilinger (GHZ) perfect correlations, and Hardy's exclusion prediction in a very simple way, starting from simplified experimental situations and using very simple algorithms. These algorithms result from five assumptions: 1. real existence of observed events; 2. logic; 3. counterfactual reasoning; 4. free experimentation; 5. Einstein's local realism. We show that if (i) experiments violate Bell's inequalities, the GHZ perfect correlations, or Hardy's exclusion prediction, and (ii) free experimentation is accepted, then Einstein's local realism has to be refuted. Therefore, Bell's inequalities, the GHZ prefect correlations, or Hardy's exclusion prediction may be called „locality criteria". We also discuss the relevance of the free-experimentation assumption for the whole issue of nonlocal effects. Nonlocality suggests the existence of a principle acting beyond space and time, which, however, cannot be influenced or exploited by human beings for practical purposes.

1. Introduction

„Nonlocality" is today a catchword known not only to physicists but also to a public generally interested in science. Moreover, it is attracting the attention of philosophers since it seems to go beyond the scope of today's physics or even to lead to a new understanding of the world.

Nonlocality means that strictly correlated events not determined by any event in the past occur „simultaneously" in „distant" regions. More precisely: any of the correlated events lies beyond the light cone of the other ones. This astonishing property of nature has been brought to light by modern physics in the last three decades. It is astonishing and unsettling because it suggests that there are influences in nature which are faster than the speed of light - apparently in contradiction to A. Einstein's special relativity [1]. Physical systems in distant regions seem to be „entangled" by a spooky principle beyond space and time. This makes the physicists wonder just like children watching a puppet show where different puppets are moved by an invisible backstage actor.

Quantum mechanics, maybe today's most fundamental physical theory, was developed in the 1920's by physicists such as N. Bohr, W. Heisenberg, W. Pauli, P.A.M. Dirac, E. Schrödinger, M. Born, and others. A startling feature of quantum mechanics is that it is an indeterministic theory, i.e., it yields only statistical predictions or probabilities for future events. This was hard to accept for physicists trained in deterministic physics, considering all events to be strictly determined by the past. A profound, animated discussion about the interpretation of quantum mechanics started. One of the most famous contributions to this discussion was a paper [2] presented in 1935 by A. Einstein, B. Podolsky, and N. Rosen (*EPR*). EPR proposed certain plausible assumptions, among them one we call *„Einstein's local realism"*, and used them to show that certain perfect correlations predicted by quantum mechanics were baffling (*EPR paradox*); they concluded that quantum-mechanical states cannot in all situations be complete descriptions of physical reality.

In 1964, J.S. Bell [3] resumed the EPR paradox in a form adapted to a „gedankenexperiment“ presented in 1951 by D. Bohm [4]. J.S. Bell proved in his landmark paper [3] a theorem (*Bell's theorem*) which states that the assumptions proposed by EPR are incompatible with some predictions of quantum mechanics. The crucial point of J.S. Bell's proof is a family of inequalities (*Bell's inequalities*), which are statistical predictions about outcomes of many measurements on two particles, typically photons or particles with spin ½.

D.M. Greenberger, M.A. Horne, and A. Zeilinger (*GHZ*) demonstrated in 1989 the incompatibility of EPR's assumptions with quantum mechanics regarding only „perfect correlations“ of at least three particles in one event [5], rather than statistical correlations in many events. *Perfect correlations* are arrangements by which the result of the measurement on one particle can be predicted with certainty given the outcomes of measurements on the other particles of the system. The *GHZ theorem* is even stronger than Bell's theorem, since it concerns only perfect correlations and works without resorting to an inequality. In 1990, D.M. Greenberger, M.A. Horne, A. Shimony, and A. Zeilinger presented a new „gedankenexperiment“ [6] proving the GHZ theorem without referring to spin, using a three-particle interferometer.

The advantages of J.S. Bell's arrangement - involving only two particles - and of the GHZ gedankenexperiment - not using any inequalities - were combined by L. Hardy in 1993 [7]; L. Hardy's discovery was developed and extended by T.F. Jordan one year later [8]. L. Hardy and T.F. Jordan demonstrated the inconsistency between the EPR assumptions and quantum mechanics for two photons or particles with spin ½ without inequalities *(Hardy's theorem).* The inconsistency is demonstrated by one single event which, according to the EPR assumptions, should never occur (a prediction we call *„Hardy's exclusion prediction“*). However, that event is not arranged to happen with certainty; there is a nonzero probability, a statistical prediction, that it will happen. The experiment may have to be repeated several times before the event happens, bur once it does happen, the demonstration is complete.

Real experiments on Bell's inequalities have already been performed and are the topics of A.M. Fox' chapter in this book [9]. The most famous among these experiments are those of A. Aspect and coworkers [10-12]. They seem to confirm quantum mechanics and to refute Einstein's local realism. Experiments on the GHZ perfect correlations have been proposed [see, e.g., 6], but not yet to be performed to date. L. Hardy's and T.F. Jordan's propositions [7, 8] have promise for new experiments which might be particularly simple and clear. At the time of writing this chapter, the research group of L. Mandel from the University of Rochester are reporting preliminary experimental results supporting L. Hardy's and T.F. Jordan's propositions [cf. 9, 13].

In this chapter we show that if (i) the experiments really violate Bell's inequalities, the GHZ perfect correlations, or Hardy's exclusion prediction, and (ii) free experimentation is accepted, then Einstein's local realism has to be refuted. Our main purpose is to derive one of Bell's inequalities (Sec. 2), one form of the GHZ perfect correlations for a three-particle system (Sec. 3), and Hardy's exclusion prediction (Sec. 4) in a very simple way. „Very simple“ means that the derivations shall be also understandable to non-physicists. As far as possible, technical terms, physical notations, and mathematical equations are avoided. We start from simplified experimental situations and use very simple algorithms to derive the locality criteria. Readers skilled in physics will have to make allowances if the Sections 2 to 4 give the impression of brainteasers rather than derivations of physical theorems.

The algorithms used in the derivations are discussed in Sec. 5. A special emphasis is laid on the importance of *free experimentation* and of human free will to the EPR discussion and to modern physics. We also discuss some consequences of the experimental violation of the locality criteria. If free experimentation is rejected, then the whole issue on nonlocality can be solved superdeterministically, without any necessity of faster-than-light influences. If, however, free experimentation is accepted, then one also has to accept a principle acting beyond space and time, which, however, cannot be influenced or exploited by human beings for practical purposes.

2. Bell's Inequality

Two physicists A and B built up the experiment shown in Fig. 1. It consists of a central object S called the „source" and of two identical objects D_A and D_B called „detectors". The three components lie on a single straight line, both detectors D_A and D_B being separated by large but equal distances from the source S. Each detector is provided with a switch with three positions labeled by U (for „Up"), M (for „Middle"), and D (for „Down"), and with a printer. Each physicist is responsible for one detector, and can *arbitrarily* switch his detector to one of the three possible positions U, M, and D. Whenever a button on the source S is pressed, shortly thereafter the printers on both detectors D_A and D_B print an „output" (or „observation") on sheets of paper. For either switch position, the output consists of a „+" or a „-".

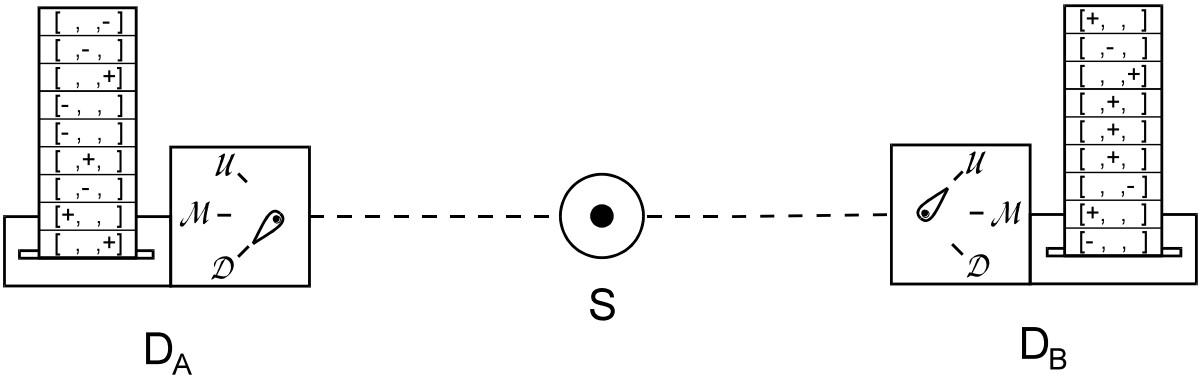

Figure 1. Simplified ficticious experiment of the Bell type. S, source; D_A and D_B, identical detectors. Each detector is provided with a switch with three positions U, M, and D, and with a printer.

In order to register the output and also the switch position of a single run, a printer uses the following notation:

```
   U   M   D      ← Switch position
[  +,      ,     ]  ← Output
```

In this example, the switch was positioned on U and the output was a „+". It is easy to see that there are 3×2 = 6 possible results.

The physicists put the results obtained in the two detectors for one single run together, using a notation explained in the following example. Assume that the source button is pressed, A sets his switch to position U and observes „+", and B sets his switch to position D and observes „-". The result of this run is noted as follows:

A's output: B's output:

U	M	D	U	M	D
+					-

Simple combinatorial analysis tells us that there are 6×6 = 36 possibilities for such compiled results. However, the physicists soon find out that the experiment never yields certain results and thus restricts the multitude of combinations. If they switch their detectors to the same positions, let's say, both to M, they always obtain the same outputs, e.g., two „-":

U	M	D	U	M	D
	-			-	

and never observe results such as

U	M	D	U	M	D
	-			+	

This experimental fact leads to the following conclusion: if both detector switches are in the same position, one obtains only results of the following six types:

U	M	D	U	M	D
+			+		
-			-		
	+			+	
	-			-	
		+			+
		-			-

We now introduce a particular theoretical assumption, namely *Einstein's local realism (or Einstein's local causality)*. According to A. Einstein's intuitive picture of phenomena, every time the button on the source is pressed, the source produces a pair of identical objects, e.g., identical „twin photons" which fly to the detectors and tell them what to print. Each photon can be characterized by a string of three bits corresponding to the outputs caused. For instance, the photon pair

U	M	D	U	M	D
+	-	-	+	-	-

will cause each detector to print a „+" if both detectors are switched to U, each detector to print a „-" if both are switched to M, and each detector to print a „-" if both are switched to D. It is as if the twin photons carried identical genes determining them to behave in the same way in relation to the same environmental circumstances. These genes are generally called „hidden variables"; in their original article [2], A. Einstein, B. Podolsky, and N. Rosen called them „elements of the physical reality". Assuming Einstein's local realism and bearing in mind the experimental results described above, we conclude that the source can produce $2^3 = 8$ possible kinds of photon pairs or identical twins:

	U	M	D	U	M	D
(a)	+	+	+	+	+	+
(b)	+	+	-	+	+	-
(e)	-	+	+	-	+	+
(f)	-	+	-	-	+	-

(c) [+ , - , + | + , - , +] (g) [- , - , + | - , - , +]
(d) [+ , - , - | + , - , -] (h) [- , - , - | - , - , -]

In a quantitative experiment, A switches his detector to U and B switches his to D. The experiment consists of a certain number of runs in which the physicists press the source button, thus producing a photon pair. In how many runs will A get the output „+“ and, simultaneously, B the output „-“? In the following we call this number of runs N(U+, D-). An examination of the eight possible kinds of photon pairs shows that the output in question is caused only when the source produces either a photon pair of the type (b) or one of the type (d); all other types of photon pairs cause different outputs. If we call the number of photon pairs of type (b) N_b and of type (d) N_d, we get the answer:

$$N(\mathrm{U+}, \mathrm{D-}) = N_b + N_d \ . \tag{1}$$

Analogously, we can put down the equations

$$N(\mathrm{U+}, \mathrm{M-}) = N_c + N_d \tag{2}$$

and

$$N(\mathrm{M+}, \mathrm{D-}) = N_b + N_f \ . \tag{3}$$

Our aim is now to give a relation between the three numbers on the left-hand sides of Eqs. (1) to (3). For this purpose, we first add the Eqs. (2) and (3):

$$N(\mathrm{U+}, \mathrm{M-}) + N(\mathrm{M+}, \mathrm{D-}) = N_c + N_d + N_b + N_f \ . \tag{4}$$

A comparison with Eq. (1) shows that the right-hand side of Eq. (4) contains N(U+, D-):

$$N(\mathrm{U+}, \mathrm{M-}) + N(\mathrm{M+}, \mathrm{D-}) = N_c + N(\mathrm{U+}, \mathrm{D-}) + N_f \ . \tag{5}$$

Of course, all numbers involved, such as N_c or N_f, are positive integers. Therefore, we conclude from Eq. (5): if Einstein's local realism is correct and all pairs produced by the source are detected, the following inequality must hold:

$$N(\mathrm{U+}, \mathrm{M-}) + N(\mathrm{M+}, \mathrm{D-}) \geq N(\mathrm{U+}, \mathrm{D-}) \ . \tag{6}$$

This is one form of *Bell's inequalities*. Evaluating the same experimental situation, but now with the aid of quantum mechanics, one gets the astonishing prediction that Bell’s inequalities will be violated. Even more astonishing is that experiments seem to violate Bell’s inequalities and to confirm quantum mechanics [cf. 9-12].

3. GHZ Perfect Correlations

Three physicists G, H, and Z built up the experiment shown in Fig. 2. It consists of a central object S called the „source“ and of three identical objects D_G, D_H, and D_Z, called „detectors“. The four components lie in a horizontal plane. The detectors D_G, D_H, and D_Z are separated by large but equal distances from the source S; the distances between any two detectors are equal, too. Each detector is provided with a switch with two positions labeled by U (for „Up“) and D (for „Down“), and with a printer. Each physicist is responsible for one detector, and can *arbitrarily* switch his detector to one of the two possible positions U and D. Whenever a

button on the source S is pressed, shortly thereafter the printers on all three detectors D_G, D_H, and D_Z print an „output“ (or „observation“) on sheets of paper. For either switch position, the output consists of a „+“ or a „-“.

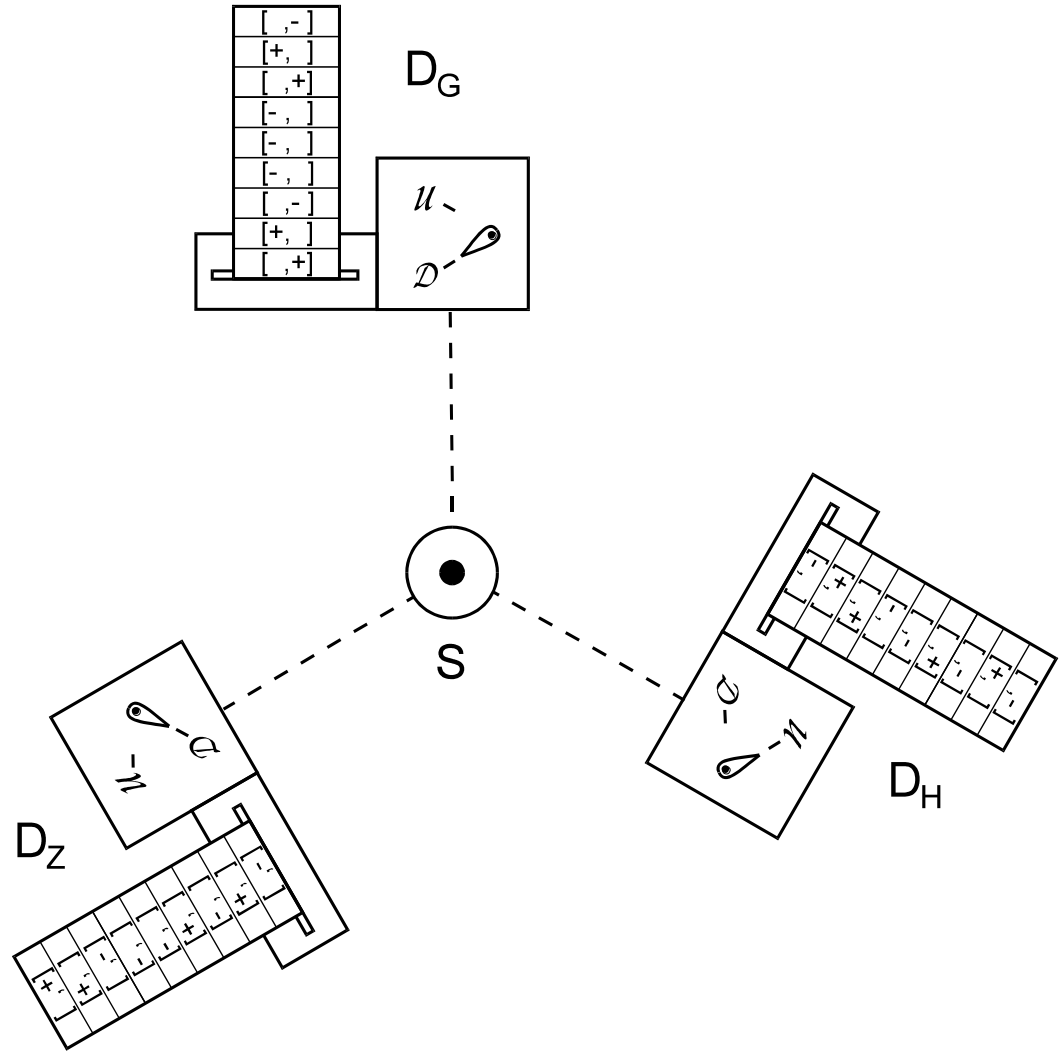

Figure 2. Simplified ficticious experiment of the GHZ type. S, source; D_G, D_H, and D_Z, identical detectors. Each detector is provided with a switch with two positions U and D, and with a printer.

In order to register the output and also the switch position of a single run, a printer uses the following notation:

```
  U   D      ← Switch position
[ + ,    ]   ← Output
```

In this example, the switch was positioned on U and the output was a „+“.

The physicists put the results obtained in the three detectors for one single run together, using a notation explained in the following example. Assume that the source button is pressed, G sets his switch to position U and observes „+“, H sets his switch to position D

and observes „-“, and Z sets his switch to position D and observes „-“. The result of this run is noted as follows:

```
        G's       H's       Z's
        output:output:    output:
          U   D   U   D   U   D
      [ + ,   |   ,  -  |   ,  - ]
        1   2   3   4   5   6        ← Position
```

The janitor of the laboratory is an interested person who wonders about the instruments and piles of paper in the laboratory. One evening, while cleaning, the janitor sees three piles of paper with experimental results accumulated after many runs. Each pile corresponds to a determined position of the detector switches:

Pile 1 contains results obtained when G switches to U, H to D, and Z to D.
Pile 2 contains results obtained when G switches to D, H to U, and Z to D.
Pile 3 contains results obtained when G switches to D, H to D, and Z to U.

Although the janitor is curious about the data, he thinks it would be pretentious for him to look for an insight into such a complicated thing without knowledge of any scientific theory. But then he notices on the wall a poster with a citation of somebody of whom G, H, and Z, as he repeatedly has observed, speak with reverence. The poster says,

> „... *the main object of physical science is not the provision of pictures, but is the formulation of laws governing phenomena and the application of these laws to the discovery of new phenomena. If a picture exists, so much the better; but whether a picture exists or not is a matter of only secondary importance.* P.A.M. Dirac“ [14].

Encouraged by these words, he carefully analyzes the results, and discovers the following regularities:

On pile 1 there are only results of the following four types:

```
          U   D   U   D   U   D
(1.1)  [ + ,   |   ,  + |   ,  + ]
(1.2)  [ + ,   |   ,  - |   ,  - ]
(1.3)  [ - ,   |   ,  + |   ,  - ]
(1.4)  [ - ,   |   ,  - |   ,  + ]
          1   2   3   4   5   6        ← Position
```

On pile 2 there are only results of the following four types:

```
          U   D   U   D   U   D
(2.1)  [   , + | + ,   |   ,  + ]
(2.2)  [   , + | - ,   |   ,  - ]
(2.3)  [   , - | + ,   |   ,  - ]
(2.4)  [   , - | - ,   |   ,  + ]
          1   2   3   4   5   6        ← Position
```

On pile 3 there are only results of the following four types:

	U	D	U	D	U	D	
(3.1)	[,	+	,	+	+ ,	]	
(3.2)	[,	+	,	-	- ,	]	
(3.3)	[,	-	,	+	- ,	]	
(3.4)	[,	-	,	-	+ ,	]	
	1	2	3	4	5	6	← Position

He has the strong feeling to be in the presence of those laws governing phenomena, of which the eminent man quoted on the wall speaks, and he seeks for formulations. He states the following three laws:

1. *Every time* G switches to U, H to D, and Z to D, the result is either (1.1), (1.2), (1.3), or (1.4).
2. *Every time* G switches to D, H to U, and Z to D, the result is either (2.1), (2.2), (2.3), or (2.4).
3. *Every time* G switches to D, H to D, and Z to U, the result is either (3.1), (3.2), (3.3), or (3.4).

But what about other switch combinations? Is it possible, starting from these three laws, to predict the result of a run in which, for instance, G switches to U, H to U, and Z also to U? P.A.M. Dirac says that such predictions are the main object of physics. Therefore, let us think a little bit ... Of course, it is possible to predict! It looks very much as if the source produced three objects with certain characteristics; these three objects fly to the detectors and tell them what to print. If we had a list or a table of all (not necessarily identical) „triplets“ the source can produce, we could predict the result for any switch combination. In order to figure out such a table, the janitor advances the following reasoning.

Let us consider a result of type (1.1). The position 2 in the bracket can be either „+“ or „-“. Let us assume it is „+“. Then, from case (2.1) it follows that position 3 in the bracket is „+“, and from case (3.1) it follows that position 5 in the bracket is also „+“. Therefore, a first possible „triplet“ is

	U	D	U	D	U	D
(a)	[+ ,	+	+ ,	+	+ ,	+]

Let us next assume position 2 in case (1.1) is „-“. Then, from case (2.4) it follows that position 3 in the bracket is „-“, and from case (3.3) it follows that position 5 in the bracket is „-“. This leads us to a second possible „triplet“:

	U	D	U	D	U	D
(b)	[+ ,	-	- ,	+	- ,	+]

(Note that the three objects of this „triplet“ are not identical.) Repeating this kind of reasoning, the janitor draws up the following table of „triplets“ for predicting results:

	U	D	U	D	U	D
(a)	[+ ,	+	+ ,	+	+ ,	+]
(b)	[+ ,	-	- ,	+	- ,	+]
(c)	[- ,	+	- ,	+	+ ,	-]
(d)	[- ,	-	+ ,	+	- ,	-]

	U	D	U	D	U	D
(e)	[- ,	+	+ ,	-	- ,	+]
(f)	[- ,	-	- ,	-	+ ,	+]
(g)	[+ ,	+	- ,	-	- ,	-]
(h)	[+ ,	-	+ ,	-	+ ,	-]

This table is complete, i.e., any „triplet“ not included in the table, such as

U	D	U	D	U	D
+ ,	+	+ ,	+	+ ,	-

cannot be produced by the source.

From the table, the janitor is able to predict: when G switches to U, H to U, and Z to U, they get one of the following four results:

U	D	U	D	U	D
+ ,		+ ,		+ ,	
+ ,		- ,		- ,	
- ,		+ ,		- ,	
- ,		- ,		+ ,	

i.e., in all results there must be an odd number (1 or 3) of plus signs. This prediction corresponds to one possible form of the *GHZ perfect correlations*.

The following day, he tells GHZ about his prediction and asks them for a test. GHZ say that according to a particular picture of phenomena called „quantum mechanics“, one should obtain one of the following four results:

U	D	U	D	U	D
+ ,		+ ,		- ,	
+ ,		- ,		+ ,	
- ,		+ ,		+ ,	
- ,		- ,		- ,	

i.e., in all results there must be an even number (0 or 2) of plus signs. Neither of these results predicted by quantum mechanics agree with those predicted by the janitor.

The physicists GHZ decide to start an experiment for testing whether nature violates the prediction of the janitor or that of quantum mechanics. At the time of writing this chapter, the results of the experiments are not yet available.

4. Hardy's Exclusion Prediction

Two physicists H and J have heard of GHZ's efforts for an experimental decision on the janitor's prediction on the one side, and quantum-mechanical prediction on the other side (cf. Sec. 3). They mull over an experiment which is simpler than that of their colleagues GHZ and finally build up the one shown in Fig. 3. Besides the „source“ S, it contains only two „detectors“ D_H and D_J - like the arrangement shown in Fig. 1 -, but each detector is provided with a switch with only two positions labeled by U (for „Up“) and D (for „Down“). Again, the two physicists can *arbitrarily* switch their detectors to one of the possible positions U and D. Whenever a button on the source S is pressed, shortly thereafter the printers on the two detectors D_H and D_J print an „output“ (or „observation“) on sheets of paper. For either switch position, the output consists of a „+“ or a „-“.

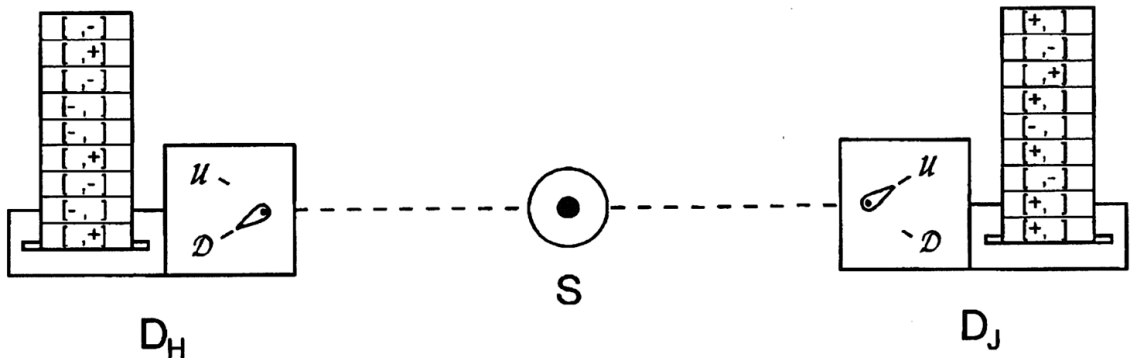

Figure 3. Simplified fictitious experiment of the Hardy type. S, source; D_H and D_J, identical detectors. Each detector is provided with a switch with two positions U and D, and with a printer.

In order to register the output and also the switch position of a single run, a printer uses the notation explained in Sec. 3. If, for instance, a switch of one detector is positioned on U and the output is a „+“, the printer notes

```
  U   D      ← Switch position
[ + ,   ]    ← Output
```

The results obtained in the two detectors for one single run are put together using the same notation as in Sections 2 and 3. In an example where the source button is pressed, H sets his switch to position U and observes „+“, and J sets his switch to position D and observes „-“, the result is noted as follows:

```
  H's      J's
  output:output:
  U   D   U   D
[ + ,   |   , - ]
  1   2   3   4      ← Position
```

In the same way as their colleagues GHZ, the physicists H and J perform many runs of experiments with various switch positions. They store their experimental results on three piles of paper, each pile corresponding to a determined position of the detector switches:

Pile 1 contains results obtained when H switches to U and J to U.
Pile 2 contains results obtained when H switches to U and J to D.
Pile 3 contains results obtained when H switches to D and J to U.

The physicists H and J analyze their results. They consult the janitor (who already has gained a certain fame), and with his help they discover the following regularities:

On pile 1 there are only results of the following three types:

```
      U   D   U   D
(1.1) [ + ,   | - ,   ]
(1.2) [ - ,   | + ,   ]
(1.3) [ - ,   | - ,   ]
```

1 2 3 4 ← Position

On pile 2 there are only results of the following three types:

	U	D	U	D
(2.1)	[+ ,		,	+]
(2.2)	[+ ,		,	-]
(2.3)	[- ,		,	-]

1 2 3 4 ← Position

On pile 3 there are only results of the following three types:

	U	D	U	D
(3.1)	[	, +	+ ,	]
(3.2)	[	, -	+ ,	]
(3.3)	[	, -	- ,	]

1 2 3 4 ← Position

On the base of these observed regularities, they state the following three laws:

1. *Every time* H switches to U and J to U, the result is either (1.1), (1.2), or (1.3), but never
   [+ , | + ,]

2. *Every time* H switches to U and J to D, the result is either (2.1), (2.2), or (2.3), but never
   [- , | , +]

3. *Every time* H switches to D and J to U, the result is either (3.1), (3.2), or (3.3), but never
   [, + | - ,]

What about the remaining switch combination, in which H switches to D and J also to D? It looks very much as if the source produced pairs of (not necessarily identical) particles, which fly to the detectors and tell them what to print. It would be desirable to draw up a table of all pairs the source can produce; with such a table, the results for all possible switch combinations could be predicted.

Let us start from a result of type (1.1). The position 2 in the bracket must be a „-“, according to case (3.3); law 3 precludes a „+“ on position 2. Position 4, on the other hand, can be „+“ or „-“ without violating any of the above laws. A „+“ on position 4 gives a first possible pair:

	U	D	U	D
(a)	[+ ,	-	- ,	+]

A „-“ on position 4 gives a second possible pair:

	U	D	U	D
(b)	[+ ,	-	- ,	-]

Repeating this kind of reasoning, the following table of pairs for predicting results can be drawn up:

	U	D	U	D
(a)	+	-	-	+
(b)	+	-	-	-
(c)	-	+	+	-

	U	D	U	D
(d)	-	-	+	-
(e)	-	-	-	-

This table is complete, i.e., any pair not included in the table, such as

U	D	U	D
+	+	+	+

cannot be produced by the source.

From the table, it is possible to predict: when H switches to D and J also to D, they get one of the following three results:

U	D	U	D
	+		-
	-		+
	-		-

In other words, they *never* get the result

U	D	U	D
	+		+

We call this prediction *Hardy's exclusion prediction*.

However, quantum mechanics predicts that in almost any experimental situation, there *is* a nonzero probability to obtain the result

U	D	U	D
	+		+

which is forbidden according to Hardy's exclusion prediction. In an „optimum" experiment, the probability is just over 9%, so it might be necessary to perform a few runs before this result is obtained. Such a relatively small probability should, however, not obscure the flagrant inconsistency between Hardy's exclusion prediction on the one side, and quantum mechanics on the other side.

The physicists H and J propose to start experiments for testing Hardy's exclusion prediction. At the time of writing this chapter, only preliminary results of the experiments are available [cf. 9, 13]. Astonishingly, they seem to violate Hardy's exclusion prediction and to support quantum mechanics.

5. Discussion

Our derivations of one of Bell's inequalities (Sec. 2), one possible form of the GHZ perfect correlations (Sec. 3), and Hardy's exclusion prediction (Sec. 4) start from simplified experimental situations with the characteristic feature that the variety of possible results is restricted; in other words, the detector outputs in an experiment are correlated. Based upon

this experimental fact, the procedures used for the derivations are simple algorithms which work by combining the following five assumptions [cf. 15-17].

1. *Real existence of observed events*. Assume that at a time t_0 a physicist A is in a region R_A and a physicist B is in a different region $R_B \neq R_A$. Physicist A observes at the time t_0 a certain detector output (e.g., „+“) and transmits this information to B by telephone. Thus, B receives the information at a later time $t_1 > t_0$. Physicist B has to assume that the information at the time t_0 possesses the same degree of reality as the human body of his colleague A at the time t_0. By assuming that the information acquires reality only at the time t_1, B would violate the principle of the real existence of observed events.

2. *Logic*. If a detector is arranged to print either a „+“ or a „-“ under certain conditions, and if the detector prints „+“ at a certain time t_0, then it cannot be assumed that the detector prints „-“ at the same time t_0.

3. *Counterfactual reasoning*. Assume that many runs of an experiment of the Bell type (such as described in Sec. 2) were performed, and that every run yielded perfect correlation; i.e., if the detector D_A switched to U printed „+“, then the detector D_B switched to U also printed „+“, and if the detector D_A switched to U printed „-“, then the detector D_B switched to U also printed „-“. Suppose that in a future run of the same experiment, D_A switched to U printed „+“ and D_B switched to M printed „-“. It is reasonable, then, to assume that if D_B had been switched to U, it would also have printed „+“ with a probability very close to one (let’s say, about as close to one as the probability for the sun to rise tomorrow). Briefly, counterfactual reasoning means using results of performed experiments to predict results of alternative, not performed experiments.

4. *Free experimentation*. Let us assume that each of two detectors, D_A in a region R_A and D_B in a different region $R_B \neq R_A$, is provided with a switch allowing three positions U, M, and D. During a run of an experiment, physicists in both regions R_A and R_B are free to change the positions of their switches arbitrarily and independently of each other. In other words, the physicists in region R_A choosing, for example, the arbitrary sequence U, M, D, M, M, M, ..., and the physicists in region R_B choosing the arbitrary sequence D, M, D, U, U, M, ... are not predetermined in their decisions.

5. *Einstein’s local realism (or Einstein’s local causality)*. The correlations in the experiments described, e.g., the fact that in the same run, the detectors D_A and D_B print the same sign in a Bell-type experiment, are already determined when the particles leave the source.

It is a fact that the predictions of quantum mechanics violate Bell’s inequalities and the GHZ perfect correlations. Now there still remains the possibility that quantum mechanics is wrong, even if it is regarded as our most fundamental physical theory [see, e.g., 18]. The only arbitrator can be the experiment. Several real experiments on Bell’s inequalities have been performed [9]. Although there is still some debate about their conclusions, they seem to violate Bell’s inequalities and to confirm quantum mechanics.

If the experiment, or in other words, nature violates Bell’s inequalities, then at least one of the five assumptions listed above is wrong. Which one is the most likely candidate? The assumptions 1 to 3 have been analyzed on that score, for example, by A. Zeilinger [15] and by A. Suarez [16, 17]. Their detailed analysis can be summarized as follows.

1. *Real existence of observed events:* A physicist B who thinks that the perfect correlations do not exist before bringing the outputs from the regions R_A and R_B together could as well

think that neither does another physicist A exist as long as he does not perceive his body. That is not only a radical nonrealism but rather an absolute solipsism.
2. *Logic:* To deny logic would simply mean to give up thinking.
3. *Counterfactual reasoning:* This is the foundation of experimental science, as well as of our everyday actions.

In our opinion, the assumption of *free experimentation* has so far not attracted the attention it deserves. Therefore, we focus on this point. First of all, we notice that free experimentation is a metaphysical principle. It is beyond the realm of physics to decide whether there is free will in the world or not. Nevertheless, the question of free will has its implications on physics.

Free will is the basic premise of every social order and legal system. Presenting, for instance, prizes and rewards for excellent scientific results, or protecting scientific publications and technical inventions as intellectual property, shows that even scientists take free will quite seriously to regulate their social behavior. The physicists of the 19th century had difficulties to incorporate free will into the classical, deterministic physics of their time. It was evident also for them that if free will exists, there has to be room for it in physics. Not until the twenties of the 20th century did it find its room - in quantum mechanics, which is a basically indeterministic physical theory. Today *there is room for free will in science*. This fact means that the basic premise of every social order and legal system is consistent with science. Paradoxically, the indeterministic physical theory with room for free will turned out to be much more powerful in describing reality than the deterministic one.

Free experimentation is especially important in the EPR and nonlocality discussion since in the Bell-type, GHZ-type, and Hardy-type experiments, freely adjustable apparatuses are essential. In the fictitious experiments of Sections 2, 3, and 4, the detector switches are freely adjustable by the physicists. In real experiments, the positions of these switches correspond, e.g., to certain discrete positions of polarizers.

The contradiction between Bell's inequalities, the GHZ perfect correlations, or Hardy's exclusion prediction on the one side, and quantum mechanics as well as, partly, experiments on the other side could in fact be removed by supposing a super-deterministic world [cf. 19]. In such a world, not only „inanimate" nature would run on a behind-the-scenes clockwork, according to the deterministic view commonly accepted by the 19th-century scientists. *„Super-deterministic"* means that even our behavior, including our belief that we are free to choose to do one experiment rather than another, is absolutely predetermined by some events in the past (let's say, the big bang). If everything is predetermined, including the experimenters' „decision" to choose certain switch positions in their experiment, the contradiction between Bell's inequalities and the experimental results disappears. The universe already „knows" long before the experiment what switch settings the experimenters will choose and what the measurement and its result will be.

J.S. Bell repeatedly [20-22] gave to understand that for him, super-determinism was no solution to the contradiction between his inequalities and quantum mechanics. „Apparently separate parts of the (super-deterministic) world would be deeply and conspiratorially entangled, and our apparent free will would be entangled with them", he wrote [20]. He thought this hypothesis would be „even more mind boggling than one in which causal chains go faster than light" [20].

Rather than giving up free will, J.S. Bell [cf. 22] considered the possibility of giving up Einstein's local realism, i.e., the fifth of the assumptions listed above. In fact, this is the point of the preceding discussion: if free experimentation is possible, and if a free experiment violates Bell's inequalities (and, eventually, the GHZ perfect correlations or Hardy's exclusion principle), then *Einstein's local realism is invalid*. From that point of view, Bell's inequalities, the GHZ perfect correlations, and Hardy's exclusion prediction can also be called *„locality criteria".* However, phenomena involving faster-than-light influences without any violation of these locality criteria are also possible [cf. 23].

Phenomena which violate Einstein's local realism are called *„nonlocal".* Nonlocality means that strictly correlated events not determined by any event in the past occur in distant regions, with each of the correlated events lying beyond the light cone of the other ones. Of course, no statement about nonlocality can be made if a signal traveling at the speed of light could carry information from one detector back to the source or to the other detector before a photon pair was produced or detected, respectively. If it is accepted that physicists can carry out their experiments freely, then *nonlocality and indeterminism are two characteristic features of nature*.

In our discussion, we have tacitly taken for granted that there is no faster-than-light communication between the correlated events, which follows from A. Einstein's special relativity [1]. Astonishingly, nonlocality seems to suggest that there are influences in nature which are faster than the speed of light. However, man cannot dispose of these influences, i.e., it is impossible to use them for superluminal signaling; thus, concerning this point, no contradiction with special relativity arises. A physical explanation of this fact is given by D.M. Greenberger, M.A. Horne, and A. Zeilinger in Ref. [24]. The authors of Ref. [24] also state that it is a deep mystery „why quantum theory, a specifically nonrelativistic theory, should conspire to be consistent with relativity in this way". The philosophical implications of the *dilemma that there is a lot of strange things going on we cannot make use of* [22] are analyzed by A. Suarez in Ref. [16]. The result of Suarez' analysis is his *„Non-controllability theorem":* „The cause of the correlated results appearing in the EPR experiments in distant regions cannot be any phenomenon, but must be an *unobservable* cause; the causal connection between these results must be *non-controllable for man*."

6. Conclusions

With the help of three fictitious experiments, we have derived one of Bell's inequalities, one possible form of the GHZ perfect correlations, and Hardy's exclusion prediction, respectively. For the derivations, we have used simple algorithms which work by combining five assumptions: 1. real existence of observed events; 2. logic; 3. counterfactual reasoning; 4. free experimentation; 5. Einstein's local realism. It is a fact that predictions of quantum mechanics and, partly, recent experiments violate Bell's inequalities, the GHZ perfect correlations, and Hardy's exclusion prediction. Therefore, at least one of the five assumptions must be wrong, namely Einstein's local realism postulating that the correlations in the experiments described are determined by the source. Bell's inequalities, the GHZ prefect correlations, and Hardy's exclusion prediction may be called „locality criteria".

We have discussed in more detail the principle of free experimentation. In today's physics, in which quantum mechanics plays a decisive role, there is room for the human free will. This is an astonishing although natural connection between physics and metaphysics.

Nonlocality suggests that there is a principle acting beyond space and time, i.e., an unobservable cause producing faster-than-light influences. However, these influences cannot be exploited by human beings for practical purposes.

References

[1] Einstein A., „Zur Elektrodynamik bewegter Körper“, *Annalen der Physik **17***, 891-921 (1905).

[2] Einstein A., Podolsky B., Rosen N., „Can quantum-mechanical description of physical reality be considered complete?“, *Phys. Rev. **47***, 777-780 (1935).

[3] Bell J.S., „On the Einstein-Podolsky-Rosen paradox“, *Physics **1***, 195-200 (1964); reprinted in: Bell J.S., *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press, Cambridge, 1987.

[4] Bohm D., *Quantum Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1951, pp. 614-623.

[5] Greenberger D.M., Horne M.A., Zeilinger A., „Going beyond Bell’s theorem“, in: *Bell’s Theorem, Quantum Theory, and Conceptions of the Universe*, ed. Kafatos M., Kluwer Academic, Dordrecht, The Netherlands, 1989, pp. 73-76.

[6] Greenberger D.M., Horne M.A., Shimony A., Zeilinger A., „Bell’s theorem without inequalities“, *Am. J. Phys. **58***, 1131-1143 (1990).

[7] Hardy L., „Nonlocality for two particles without inequalities for almost all entangled states“, *Phys. Rev. Lett. **71***, 1665-1668 (1993).

[8] Jordan T.F., „Testing Einstein-Podolsky-Rosen assumptions without inequalities with two photons or particles with spin ½“, *Phys. Rev. A **50***, 62-66 (1994).

[9] Fox A.M., „Optical tests of Bell’s inequality“, this volume, chapter IX.

[10] Aspect A., Grangier P., Roger G., „Experimental tests of realistic local theories via Bell’s theorem“, *Phys. Rev. Lett. **47***, 460-463 (1981).

[11] Aspect A., Grangier P., Roger G., „Experimental realization of Einstein-Podolsky-Rosen-Bohm *gedankenexperiment*: A new violation of Bell’s inequalities“, *Phys. Rev. Lett. **49***, 91-94 (1982).

[12] Aspect A., Dalibard J., Roger G., „Experimental test of Bell’s inequalities using time-varying analyzers“, *Phys. Rev. Lett. **49***, 1804-1807 (1982).

[13] J.R. Torgerson, D. Branning, C.H. Monken, and L. Mandel, *Phys. Lett. A*, **204**, 323 (1995).

[14] Dirac P.A.M., *The Principles of Quantum Mechanics*, 4th Ed., Clarendon Press, Oxford, 1958, p 10.

[15] Zeilinger A., „Physik und Wirklichkeit - Neuere Entwicklungen zum Einstein-Podolsky-Rosen Paradoxon“, in: *Naturwissenschaft und Weltbild*, eds. Reichel H.-C., Prat de la Riba E., Hölder-Pichler-Tempsky, Vienna, 1992, pp. 99-121.

[16] Suarez A., „Unentscheidbarkeit, Unbestimmtheit, Nicht-Lokalität. Gibt es unverfügbare Kausalverbindungen in der physikalischen Wirklichkeit?“, in: *Naturwissenschaft und Weltbild*, eds. Reichel H.-C., Prat de la Riba E., Hölder-Pichler-Tempsky, Vienna, 1992, pp. 223-264.

[17] Suarez A., *The EPR Correlations Cannot Be Explained by Controllable Causality*, reprint CQP-1116-91, Center for Quantum Philosophy, Zurich, 1991.

[18] Bell J.S., „Against measurement“, *Physics World*, 33-40 (1990).

[19] Gribbin J., „The man who proved Einstein was wrong“, *New Scientist **24***, 43-45 (1990).

[20] Bell J.S., *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press, Cambridge, 1987, p 154.

[21] Bell J.S., „Indeterminism and nonlocality“, this volume, chapter VII.

[22] Bell J.S., „‘Raffiniert ist Gott, aber boshaft ist Er nicht.’ (Einstein) - Wirklich nicht?“, in: *Naturherrschaft*, ed. Thomas H., Busse Seewald, Herford, Germany, 1991, pp. 277-325.

[23] Suarez A., “Nonlocal phenomena: physical explanation and philosophical implications“, this book, chapter X.

[24] Greenberger D.M., Horne M.A., Zeilinger A., „Multiparticle interferometry and the superposition principle“, *Physics Today*, 22-29 (1993).

Chapter IX

Optical Tests of Bell's Theorem

A.M. Fox*
Department of Physics, University of Oxford
and
Grandpont House, Oxford,

Abstract

This article reviews recent work using optical experiments to test Bell's theorem. The article begins by discussing the difference between the local hidden variable and Copenhagen interpretations of quantum experiments as illustrated by the Einstein-Podolsky-Rosen paradox. Then Bell's theorem is introduced, and some of the more recent experiments to test it are described. A brief review of the critical literature on these experiments is given. Finally, the implications of Bell's theorem are discussed in the context of Einstein's defence of realism.

1. Introduction

The formulation of quantum mechanics began in 1926 with the seminal work of Schrödinger and Heisenberg. In the years following their papers, the new calculational techniques were applied to a great variety of physical problems with an extraordinary degree of success. By any measure, quantum mechanics is an extremely successful theory: its predictions agree with experimental results to a very high degree of numerical accuracy. Yet despite the undeniable calculational precision of quantum mechanics, the debate still rages about what quantum theory tells us about the nature of matter at the microscopic scale. Physics students run into this problem every year when they first encounter quantum mechanics. Those who wish to do well in their exams tend to forget about the issue and concentrate on learning how to solve the equations, while experienced tutors tend to side-step awkward questions like "what does it all mean ?".

The heart of the problem is in understanding what measurements tell us about the microscopic world. Scientists try to devise experiments that are able to determine the properties of the system under investigation without significantly disturbing it in the process. Consider, for example, a classical experiment in which we wish to measure the speed of a moving object. The speed of the object has a well-defined value whether we measure it or not, and the aim of the experiment is to determine what that value is. The measurement could be made by a number of techniques. Two obvious ones are to measure the time taken for the object to move a known distance, or to use the Doppler effect. We expect that the answers we get will not depend on the choice of technique. If there are differences, we would have to examine in detail the way the measurements were made, and we would expect to be able to attribute the discrepancies to systematic experimental errors. This type of experimental approach does not carry over directly to the quantum world, where experiments are made on single particles rather than large systems. In the quantum case it is sometimes found that the measuring apparatus irreversibly affects the system that is being tested. The interaction

* address for correspondence: Clarendon Laboratory, Parks Road, Oxford OX1 3PU, U.K.,
E-mail: mark.fox@physics.ox.ac.uk

between the apparatus and the particle appears to be a random process, so that there is no way of extrapolating back to the properties the particle possessed before the measurement was made. This is a very different situation to the classical experiment, where the deterministic laws would always allow us to work back to the starting conditions, at least in principle.

The relationship between the values measured in an experiment and the underlying properties of the system under test is a crucial one in physics, and it is therefore not surprising that there has been considerable controversy over how to interpret the results of quantum experiments. Two contrasting interpretations were developed in the thirties, principally by Niels Bohr and Albert Einstein. The approach of Bohr and co-workers has come to be known as the Copenhagen interpretation, and it is generally regarded as the orthodox position. In the Copenhagen approach, the state of the particle prior to the observation is described by a wave function. Knowledge of this wave function allows us to calculate the probability of the possible results of a given experiment, but the results obtained cannot be considered as the consequence of pre-existing properties of the system and the apparatus. The wave function is probabilistic before the measurement, and only becomes determined once the measurement has been made. The measurement process is therefore called wave function *collapse*. The interpretation of Einstein and others is called the local hidden variables (LHV) approach[4]. In this approach it is supposed that the particle has well-defined properties prior to measurement, in much the same way as a classical particle has. These properties are quantified by mathematical variables, and the results obtained are the consequence of interactions between these pre-existing properties and the pre-existing properties of the measurement apparatus. The variables are *local* because they apply to a specific particle at a particular point in spacetime, and they are *hidden* because they refer to a hidden property of the particle that we have no way of measuring with the techniques available at the present time.

The difference between the two approaches was brought to the fore by a paper entitled "Can quantum-mechanical description of physical reality be considered complete" published in 1935 by Einstein, Podolsky and Rosen [2]. In that paper, commonly known as the EPR paper, the authors discuss the properties of a system which splits into two well-separated systems, in such a way that measurements on one of the separated systems pre-determines the result of the measurements on the other one. In 1951 David Bohm gave an example of an EPR system that might easily be tested in the laboratory [3]. Bohm's version of the EPR experiment (the EPRB experiment) involved the measurement of the spins of the two separated atoms that originate from the disintegration of a diatomic molecule. Conservation laws demand that the sum of the spins must be zero, and so if we measure the spin of one atom, we automatically know the spin of the other one without measuring it. The quantum explanation of the EPR experiment appeared to Einstein as "incomplete": he could not understand how a measurement of the properties of a particle at one point in space could instantly determine the properties of another which is well separated from it. This amounts to *action at a distance*, a concept that was rejected at the turn of the century with the advent of relativity.

For many years the debate between the two approaches was carried out at the philosophical level. However, the situation changed in 1964 with the publication of Bell's theorem [4]. In his paper, John Bell proved that testable contradictions could arise between the quantum mechanical and LHV approaches. In particular, he gave a theoretical analysis of a refined version of the EPRB experiment, and showed that if the particle possessed local

[4] The question of whether Einstein truly was a proponent of hidden variable theories is discussed in ref.[1].

hidden variables, an inequality (called Bell's inequality) would always be obeyed, but that this would not necessarily be the case in the quantum mechanical approach. Since 1964, many experiments have been performed attempting to test whether Bell's inequality holds or not, and at the present time, there is very strong experimental evidence that nature violates Bell's inequality. This does not automatically mean that the debate between the Copenhagen interpretation and realistic views assuming well-defined values prior to measurement has been settled. Rather, it shows that if a realistic interpretation of quantum mechanical experiments is to be given, then it must be *non-local*, and this is why Einstein's *local* hidden variable theories do not predict the right result. There is, in fact, one very well-known example of a hidden variables theory which has a non-local character, namely David Bohm's pilot wave theory [5]. It should be noted, however, that Bohm's model is not compatible with the relativity of spacetime, and that it leads one to assume a quantum aether [6].

For many physicists, the results and conclusions of the experiments testing Bell's theorem are highly disturbing: "Anybody who's not bothered by Bell's theorem has to have rocks in his head" [7]. If the experiments had given a positive result in favour of Bell's inequality, the interpretation would have been easy to reconcile with the rest of our knowledge of the physical world. As it stands, we must either accept a very strange concept (non-locality) or we must look for flaws either in the argumentation of Bell or in the experimental method. The aim of this article is to explain the difference between the local and non-local theories, and to discuss how it relates to philosophical issues. In §2 I explain how the EPR paradox arises. In §3 the detailed form of Bell's theorem is discussed, together with recent optical experiments that test it. Bell's theorem is introduced in §3.1. Then in §3.2 the experiments of A. Aspect and co-workers [8,9,10] are described, while §3.3 reviews some of the more recent work. These two sections are more technical than other parts of the article, and could be passed over by a reader without a background in physical sciences. §3.4 reviews some of the literature that discusses the possible weaknesses in the arguments used, and in §3.5 I discuss recent developments in which Bell's theorem has been stated and tested without using an inequality. §4 discusses the interpretation of the experiments, and the philosophical overtones, and in §5 I draw the article to its conclusions. An Appendix is included for the benefit of readers who are not fully familiar with the properties of polarised light, which are important for the optical experiments described.

2. The Einstein-Podolsky-Rosen paradox and nonlocality

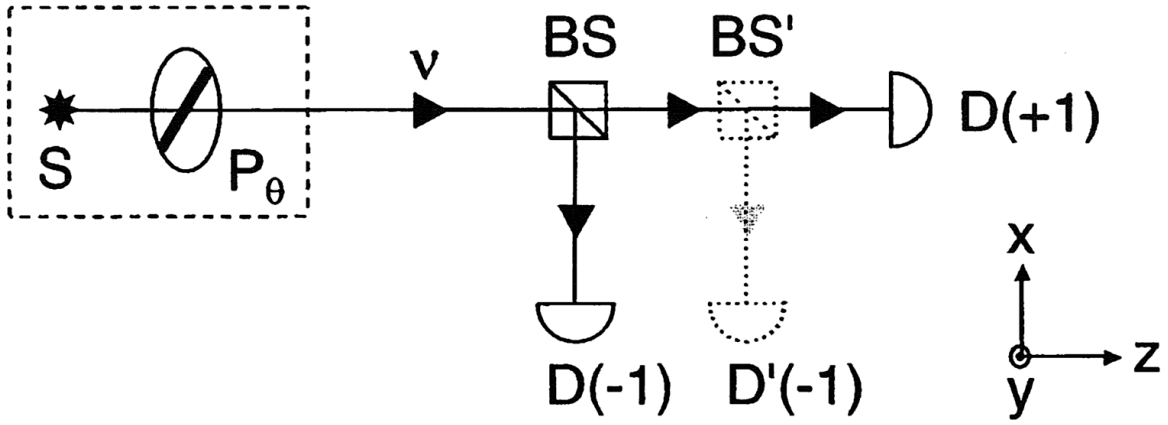

Figure 1
Apparatus for the measurement of photon polarisation using a polarising beam-splitter *BS* and two single photon counting detectors *D(+1)* and *D(–1)*. *S* is a variable intensity light source, and P_θ is a removable polaroid. By inserting an additional polarising beam-splitter *BS′*

between *BS* and *D(+1)*, one can check that all the photons transmitted by *BS* are horizontally polarised.

In order to explain how the EPR paradox arises, it is first necessary to understand how the LHV and Copenhagen approaches differ. This difference is well illustrated by considering the measurement of photon polarisation performed by the apparatus shown in Fig.1 [11]. The experiment consists of a light source *S*, a polarising beam-splitter *BS*, and two detectors, *D(+1)* and *D(–1)*. The detectors are sensitive enough so as to be able to register the arrival of single photons. We define axes such that the light is propagating in the z-direction before incidence on *BS*. The experiment has two possible results, which we denote ±1, corresponding respectively to counts on the detectors *D(±1)*. If detector *D(+1)* fires, we know that the photon has been transmitted by the beam-splitter, and if *D(–1)* fires, it has been reflected. The source is chosen so that it emits unpolarised light, and can be turned into a source of linearly polarised light by adding the polaroid P_θ before the beam-splitter.

Consider first what happens when linearly polarised light is incident on *BS*. Classically, the transmissivity and reflectivity of *BS* are $\sin^2\theta$ and $\cos^2\theta$ respectively, where θ is the angle between the polarisation and the y-axis, which is defined by the axis of the beam-splitter (see Appendix). In particular, if $\theta = 0°$ (vertically polarised light), all the light is reflected, whereas if $\theta = 90°$ (horizontally polarised light), all the light is transmitted. If the intensity of the source is turned down, we can eventually reach a situation where the individual photons can be detected one by one. This is the quantum regime. Since an individual photon cannot be split in two, we find that we either get a +1 result or a –1 result. If the source is horizontally polarised, we obtain +1, and if it is vertically polarised, we obtain –1. If we have a polarisation angle θ different from 0°or 90°, we obtain random results, except that the overall probability of +1 is $\sin^2\theta$ and that of –1 is $\cos^2\theta$. Here we are using the word "probability" to mean the average of a large number of events registered by the detectors. We can perform additional measurements to check what the polarisation of the photon is after it has passed through *BS*. For example, this can be done by placing a second beam splitter *BS′* with another detector *D′(–1)* between *BS* and *D(+1)*. In this case it is found that *D′(–1)* never fires. This implies that all the photons transmitted by *BS* are horizontally polarised. In the same way we can check that all the photons reflected are vertically polarised. Thus it appears that the beam splitter destroys the initial polarisation θ, except in the special cases when $\theta = 0°$ or $\theta = 90°$. We know the polarisation state before and after the measurement, and in general these are different. This is a manifestation of the fact that the act of making a measurement on a quantum system irreversibly alters the state of the system studied once a detection event has occurred.

Now contrast the case when P_θ is removed and unpolarised light is incident on *BS*. Classically, the light has random polarisation, and so half is transmitted and the other half reflected (see Appendix). In the quantum case, we find that we get random results, with +1 occurring with probability 50% and likewise for –1. As before, we can check that all the photons transmitted by *BS* are horizontally polarised, and that the reflected photons are vertically polarised. When we analyse the results, the situation is a little different from the case of the linearly polarised light source. This is because we now only know the polarisation state after the measurement is made. Before the beam-splitter, we know nothing about the polarisation of the individual photons other than that the average polarisation of many photons must be random.

Let us now analyse the results following the Copenhagen and the LHV interpretations. In the Copenhagen school we represent the states of the photon by wave functions. Let the wave functions ψ_v and ψ_h represent the states of vertically and horizontally polarised photons respectively. If detector *D(+1)* fires, then we know that the photon is horizontally polarised after passing through *BS* and thus that its wave function after *BS* is ψ_h. Likewise, if *D(–1)* fires then the wave function must be ψ_v afters *BS* . According to the superposition principle, we must describe the state of each individual photon before reaching *BS* as a superposition wave function $\Psi$of the form:

$$\Psi = \left(a_{+1}\psi_h + a_{-1}\psi_v\right). \qquad (1)$$

The act of measurement "collapses" the wave function from the superposition state Ψ into either the state of a vertical photon ψ_v, or into the state of horizontal photon ψ_h. This is a probabilistic process governed purely by chance, and the probabilities are given by $|a_{+1}|^2$ for the horizontal photon, and $|a_{-1}|^2$ for the vertical one. We can reproduce the results for the case of the polarised light source by letting $a_{+1} = \sin\theta$ and $a_{-1} = \cos\theta$. In the case of unpolarised light we simply set $a_{+1} = a_{-1} = 1/\sqrt{2}$.

The approach using hidden variables would be quite different. This school assumes that the results of the experiment can be deduced from unknown variables which determine the outcome according to deterministic (i.e. not probabilistic) laws. We do not have to know what these variables are, nor the details of the laws that govern the outcome of the experiment[5]. An analogy can be made with the example of tossing a coin. This appears to be a purely chance process, but we know in principle that if we were to know all the classical variables of the problem (e.g. the initial orientation of the coin, the forces applied to it, etc.) we could calculate the result from Newton's laws. This analogy is carried over to the quantum case. For example, we might conjecture that each individual photon has a well-defined polarisation state prior to the measurement. In this case, the hidden variable would be the angle θ of the polarisation vector relative to the y-axis. We then propose a new deterministic law for the transmission of individual photons through polarizers. One possibility is that the beam splitter rotates the polarisation of the photon until it is either horizontal or vertical depending on whether |θ| is greater or less than 45° respectively, and then the photon is reflected or transmitted according to the usual rules of polarising beam-splitters. This explains the outcome for the polarised light source. In the case of the unpolarised light source we simply say that the source emits randomly polarised photons, so that half will on average be polarised with |θ| < 45°, and the other half with |θ| > 45°.

In both schools of thought we are making suppositions about the state of the system before a measurement is made. We only have certain knowledge of the photon polarisation *after* the measurement has been made, and we know that in general the measuring act alters the polarisation except in the special cases of θ = 0° and θ = 90°. Moreover, the outcome of the experiment depends on how we measure it. For example, the angle θ is only defined relative to the axes of the beam-splitter, so that if we rotate *BS* while keeping the source

[5] Hodgson has argued that we should not exclude the possibility that physicists in the future may devise new experiments which reveal the underlying variables and laws, quoting from J.J. Thomson that "the immeasurable of today may be the measurable of tomorrow"[12]. Some of the hidden variable theories are in fact very well developed and can reproduce many key "quantum" results without invoking quantum mechanics at all. One example is Bohm's pilot wave theory [5]. Another is the theory of stochastic electrodynamics [13].

polaroid P_θ fixed, we will get different results. The argument is about what the polarisation state of the photon is *before* the measurement is made, and how the measuring apparatus alters the polarisation. In the Copenhagen interpretation, the state prior to the measurement is the superposition state of Eq.(1). We cannot say that the photon is either vertical or horizontal: it is neither, or perhaps both at the same time. We have an indeterminate quantum world which only becomes solid once measurements (whose outcome is probabilistic) are made. In the LHV approach, by contrast, the initial state is well-defined and the measuring process is governed by rigidly deterministic laws in which "God does not play dice", as Einstein expressed it.

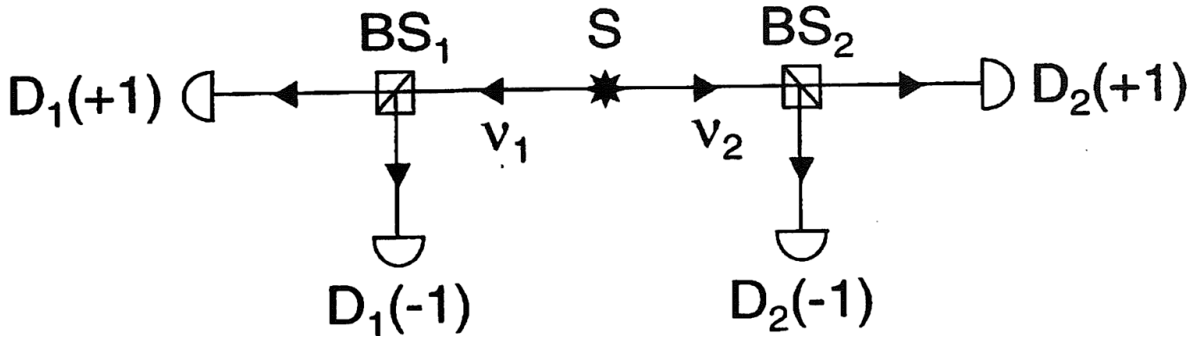

Figure 2
Apparatus for an Einstein-Podolsky-Rosen-Bohm (EPRB) experiment using correlated photon pairs emitted by atomic cascades in a source *S*. The two correlated photons ν_1 and ν_2 propagate in opposite directions and are measured by polarising beam-splitter/detector pairs similar to those of Fig.1.

The concept of locality entered the debate about the interpretation of quantum mechanics after Einstein, Podolsky and Rosen tried to prove that quantum mechanics was "incomplete" in the EPR paper. This can be seen by considering the EPRB experiment shown schematically in Fig.2. This is a modified form of the experiment of Fig.1 in which we now have a source *S* that emits photons in pairs rather than individually. The photons pairs (labelled ν_1 and ν_2) propagate in opposite directions. For convenience, we will say that ν_1 goes to the left, and ν_2 goes to the right. The polarisation of the two photons can be individually measured by identical beam-splitter arrangements similar to Fig.1, and the two pairs of detectors are well separated from each other. The subtlety in the experiment occurs when we choose *S* so that it emits *correlated* photon pairs. Correlated photon pair sources have the following properties:

1. The polarisation of either ν_1 or ν_2 measured independently of the other is random.
2. The polarisation of the pair of photons is perfectly correlated; that is, if $D_1(+1)$ fires, then $D_2(+1)$ always fires, and if $D_1(-1)$ fires, then $D_2(-1)$ always fires.

The second property is a consequence of internal conservation laws of the source which are discussed in more detail in §3.2.

If we adopt the LHV approach, the two properties do not pose any problems. We would suppose that the two photons possess a hidden variable that pre-determines the correlated results. For example, we could postulate that the polarisation angle θ is the same for ν_1 and ν_2 (due to the internal conservation laws), but random from pair to pair. Then if we apply the hidden variable rule proposed above for the behaviour of single photons at a polarising beam-splitter, we can easily explain the two properties. Property (1) is readily explained because θ is random from pair to pair, and we thus have the case of the unpolarised

source. Property (2) follows immediately from the fact that ν_1 and ν_2 have identical polarisations.

If we now analyse the experiment in the Copenhagen approach, we run into the issue of non-locality when thinking about property (2). According to the principle of complementarity, the photons are in superposition states before they are incident on the beam-splitters. The wave function is collapsed when the measurement is made. The problem is that the measurement on ν_1 seems to pre-determine the results obtained for ν_2, i.e. the act of collapsing the wave function of ν_1 also collapses the wave function of ν_2. Now there is nothing to stop us using detectors with a fast response time τ and to separate the beam-splitters by a distance further than $c\tau$, where c is the velocity of light. This means the "back-action" mechanism that enables the measurement on ν_1 to collapse the wave function of ν_2 must be faster than the speed of light. The theory of relativity tells us that it is impossible to send physical signals faster than c, and so this means that the back-action interaction appears to defy a well-established law.

The conclusion of the Copenhagen interpretation of the EPRB experiment is that "separated particles influence each other even when there is no known interaction between them" [14]. The formalism of quantum theory incorporates this non-locality by describing the correlated photons in a non-local "entangled state" such that they are never truly separated. The objection to this is a simple one, namely that it smacks of action at a distance, in which a cause at a given point in space produces effects elsewhere in space instantaneously. This concept was rejected long ago, and Einstein described the action at a distance implied by non-locality as "spooky" [15].

It is clear from this discussion that the LHV approach certainly seems to give the more sensible interpretation of the EPRB experiment. The concept of non-locality which is forced on us by the quantum mechanical interpretation seems to be very foreign, and looks like action at a distance. Einstein thought that the implausibility of non-locality was sufficient reason to disprove the Copenhagen interpretation. The question arises whether there is any experiment we can do which can give further insight into the issue. The answer is the experiments to test Bell's theorem, which are described in the next section.

3. Bell's theorem and its experimental verification

3.1 Introduction to the Bell theorem

In his classic paper of 1964, Bell discussed a gedanken experiment similar to the EPRB experiment, but with an important alteration. He considered the possibility of measuring the properties of correlated particles in which the axes of the measuring apparatus are not parallel. A schematic diagram of the optical version of this gedanken experiment is given in Fig. 3. As in §2, the apparatus consists of a source of correlated photon pairs (ν_1 and ν_2) which propagate in opposite directions, and the polarisation of each photon is measured with a polarising beam-splitter and two photodetectors. However, we now study what happens when the axes $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ of the two beam-splitters are different from each other. We keep $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ perpendicular to the z-axis ($\vec{\mathbf{z}}$), but allow them to make different angles α and β with the y-axis ($\vec{\mathbf{y}}$). If we consider the special case in which $\alpha = \beta$, we return to the situation of Fig. 2,

namely the EPRB experiment[6]. The experiment has four possible results, which are characterised by their respective probabilities:

$P_{++}(\alpha,\beta)$ is the probability that $D_1(+1)$ fires and $D_2(+1)$ fires,

$P_{+-}(\alpha,\beta)$ is the probability that $D_1(+1)$ fires and $D_2(-1)$ fires,

$P_{-+}(\alpha,\beta)$ is the probability that $D_1(-1)$ fires and $D_2(+1)$ fires,

$P_{--}(\alpha,\beta)$ is the probability that $D_1(-1)$ fires and $D_2(-1)$ fires.

As before, probabilities are defined as the average of a large number of experiments.

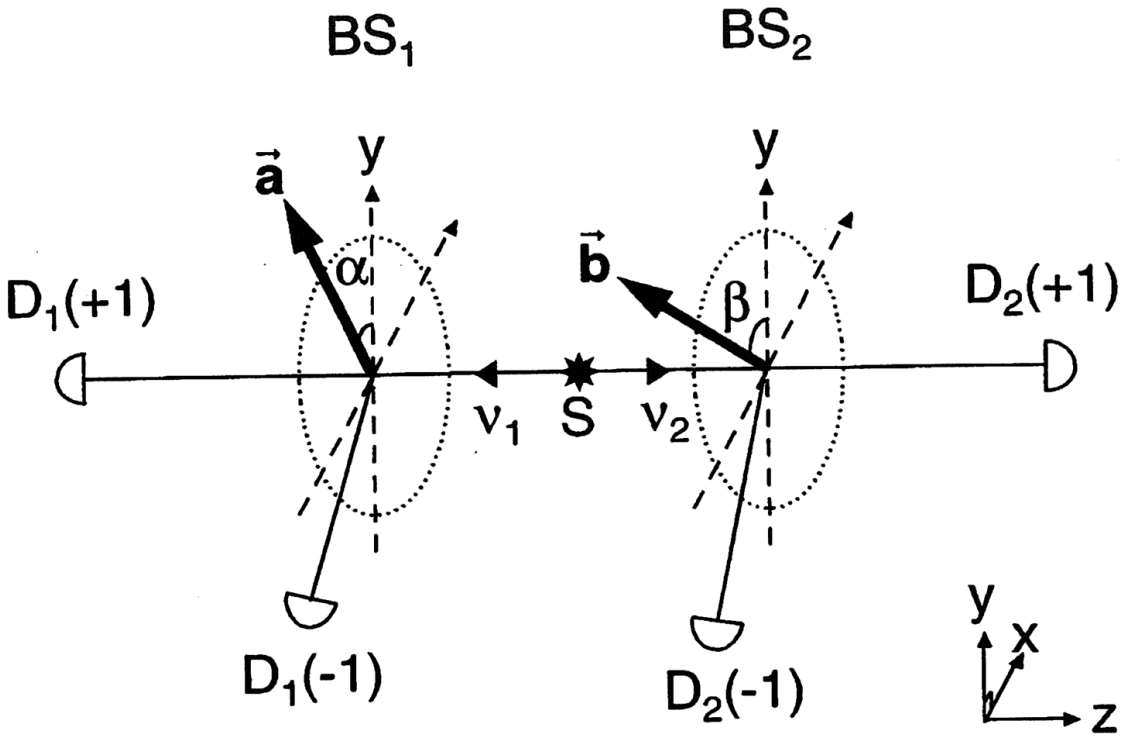

Figure 3
Apparatus for the optical version of the Bell inequality experiment using correlated photon pairs. The two photons ν_1 and ν_2 are incident on polarising beam-splitters BS_1 and BS_2 which do not have parallel axes. Single photon detectors $D_1(+1)$ and $D_2(+1)$ are set up to count the photons transmitted by the beam-splitters, while $D_1(-1)$ and $D_2(-1)$ count the reflected photons. $D_1(-1)$ and $D_2(-1)$ lie in the xy-plane of their beam-splitters, and are at right-angles to the axes $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ respectively.

If we were to try to analyse the Bell experiment using an LHV model, we might suppose for example that the source emits photon pairs in which both photons have the same polarisation angle θ, but that θ is random from pair to pair, as in §2. It is then easy to show that the probabilities expected are given by:

[6] In the EPRB experiment it is not necessary that $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ be parallel to $\vec{\mathbf{y}}$. The correlated two-photon sources used in the experiments are rotationally invariant, which implies that the whole apparatus is unaffected by a rotation about $\vec{\mathbf{z}}$. The important point is that $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ are parallel to each other, i.e. that $\alpha = \beta$.

$$
\begin{aligned}
\mathrm{P}_{++}(\alpha,\beta) &= \frac{90^\circ - \phi}{180^\circ} \\
\mathrm{P}_{+-}(\alpha,\beta) &= \frac{\phi}{180^\circ} \\
\mathrm{P}_{-+}(\alpha,\beta) &= \frac{\phi}{180^\circ} \\
\mathrm{P}_{--}(\alpha,\beta) &= \frac{90^\circ - \phi}{180^\circ}
\end{aligned}
\qquad (2)
$$

where $\phi = |\alpha - \beta|$ [16]. Note that these probabilities satisfy two necessary check rules:

1. The total of probability of getting a +1 or –1 for each photon is exactly ½ [7].
2. The correct results for the EPRB experiment are reproduced when $\alpha = \beta$ (i.e. $\phi = 0$), namely that $\mathrm{P}_{++}(\alpha,\beta) = \mathrm{P}_{--}(\alpha,\beta) = \frac{1}{2}$ and $\mathrm{P}_{+-}(\alpha,\beta) = \mathrm{P}_{-+}(\alpha,\beta) = 0$.

On the other hand, if we apply quantum-mechanical arguments, it can readily be shown that the expected probabilities are given by [16]:

$$
\begin{aligned}
\mathrm{P}_{++}(\alpha,\beta) &= \tfrac{1}{2}\cos^2\phi \\
\mathrm{P}_{+-}(\alpha,\beta) &= \tfrac{1}{2}\sin^2\phi \\
\mathrm{P}_{-+}(\alpha,\beta) &= \tfrac{1}{2}\sin^2\phi \\
\mathrm{P}_{--}(\alpha,\beta) &= \tfrac{1}{2}\cos^2\phi
\end{aligned}
\qquad (3)
$$

These also satisfy the two check rules listed above.

On comparing Eqs. 2 and 3, we see that we get the same probabilities for $\phi = 0°$, 45°, and 90°, but not at other angles. For example, at $\phi = 22.5°$, $\mathrm{P}_{++}(\alpha,\beta)$ is 0.375 according to the LHV model, and 0.427 in the quantum mechanical case. This simple example illustrates the possibility that the LHV and quantum mechanical approaches can give rise to different predictions for a particular experiment. One might object that the LHV model chosen to illustrate the point was a particularly simple one, and that it might be possible to construct a more complicated model which gives the same predictions as quantum mechanics. This would in fact be a fruitless task: not even Sherlock Holmes would succeed [17]. This is because Bell proved that for some measurable quantities, the LHV and quantum mechanical approaches will *always* give different predictions.

The actual Bell inequality refers to the results of the following gedanken experiment. Set the left-hand polariser at angle α, and the right-hand polariser at β. Measure a large number *N* of photon events. Some of the time we will obtain a +1 result on the left and a –1 result on the right. We count this number and call it $n[\alpha^+\beta^-]$ [8]. Now rotate the polariser on the right-hand side to a new angle γ while leaving the one on the left unchanged, and measure the same number *N* of photon events. We denote the number of times we obtain the result +1 on the left and –1 on the right $n[\alpha^+\gamma^-]$. Finally, rotate the polariser on the left until its angle is γ, and rotate the right-hand polariser to β. Record N events and call the number of times we observe +1 on the left and –1 on the right $n[\gamma^+\beta^-]$. Bell proved that the following inequality

---

[7] This can be seen by adding $\mathrm{P}_{++}(\alpha,\beta)$ to $\mathrm{P}_{+-}(\alpha,\beta)$ which covers both possibilities where +1 is obtained for ν_1. The answer is exactly ½. In the same way, the probability of getting –1 for ν_1 is $[\mathrm{P}_{-+}(\alpha,\beta) + \mathrm{P}_{--}(\alpha,\beta)]$, which is also exactly ½. The results for ν_2 follow similarly.

[8] This is equal to $N \times \mathrm{P}_{+-}(\alpha,\beta)$ in the notation of the previous paragraphs.

$$n[\alpha^+\beta^-] \le n[\alpha^+\gamma^-] + n[\gamma^+\beta^-] \qquad (4)$$

always holds in LHV models, but it can sometimes be violated in the quantum mechanical case. The article by P. Pliska in the present volume explains the reasoning behind the derivation of this inequality[18].

In practice Eq. 4 is difficult to test experimentally. This is because even the best single photon detectors are still inefficient. Ideally we want to have the case that either $D_i(+1)$ fires or $D_i(-1)$ fires. In fact, most of the time neither fire. For this reason, various authors have derived new generalised Bell inequalities which are easier to test experimentally. A particularly important contribution was made by Clauser, Horne, Shimony and Holt, who derived the following inequality [19]:

$$-2 \le S \le +2 \ , \qquad (5)$$

where

$$S = E(\alpha,\beta) - E(\alpha,\beta') + E(\alpha',\beta) + E(\alpha',\beta') \ ,$$

and

$$E(\alpha,\beta) = \mathrm{P}_{++}(\alpha,\beta) + \mathrm{P}_{--}(\alpha,\beta) - \mathrm{P}_{+-}(\alpha,\beta) - \mathrm{P}_{-+}(\alpha,\beta) \ .$$

The inequality of Eq. 5 holds for LHV theories, but can be violated quantum mechanically for certain choices of angles. For example, if $\alpha = 0°$, $\beta = 22.5°$, $\alpha' = 45°$, and $\beta' = 67.5°$, then quantum mechanics predicts a value of $S = 2\sqrt{2}$, which clearly violates Eq. 5 by a substantial margin.

Following the derivation of the Bell inequality, there has been much experimental interest shown in finding out whether violations of the Bell inequality actually exist in nature. D'Espagnat has reviewed the experimental results prior to the Aspect experiments [20], and in the following sub-sections, I shall describe the Aspect experiments and some of the more recent optical work in this field. The experimental evidence increasingly supports the quantum mechanical prediction that the inequalities are violated in some situations. It is worth stressing that the Bell inequality debate is often pitched as a contest between realistic theories and quantum mechanics. In fact, this is not the point. In Bell's own mind the key issue that is at stake is *locality*: *It is the requirement of locality, or more precisely that the result of a measurement on one system be unaffected by operations on a distant system with which it has interacted in the past, that creates the essential difficulty* [4]. Violations of Bell's inequality therefore do not prove that the Copenhagen interpretation is correct. Rather they show that there is significant difference between local and non-local theories, and only non-local theories correctly predict the violations. As summarised by Rae: *We have seen that non-locality is an essential feature of any model giving results in breach of Bell's theorem so that some form of 'action at a distance' is necessary .... Even if quantum theory were shown to be incorrect tomorrow, any new fundamental theory would also have to face the challenge of the violation of Bell's inequality and would have to predict the observed correlations between widely separated measurements* [21].

3.2 The Aspect experiments

Alain Aspect and his co-workers completed three optical experiments testing for violations of Bell's inequality between 1981 and 1982. These experiments studied the correlated pairs of photons emitted by atomic cascades in calcium. A simplified level scheme for the atomic transitions involved in the cascade is given in Fig. 4(a). The initial and final states for the cascade are both J=0 states, with even parity. The atoms are excited to the upper level by pumping them with a laser, and they relax by emitting two photons, ν_1 and ν_2, at 551.3nm and 422.7nm respectively. The fact that these two photons are at different wavelengths is

important for the experiment: the atoms emit photons pairs in all directions, but the difference in wavelength permits us to distinguish between ν_1 and ν_2. For example, by inserting a coloured filter that cuts out light at 422.7nm to the left of the source, we can ensure that only the photons at 551.3nm reach the detectors on the left-hand side. Similarly, a filter cutting out light at 551.3nm on the right ensures that only the 422.7nm photons reach the right-hand detectors. The fact that the initial and final levels are both J=0 states requires that the photon pairs carry no net angular momentum. In addition, the rotational invariance of J=0 states, and the fact that the initial and final levels are both of the same parity, requires that the photon pairs have the correlation properties required for the EPRB and Bell experiments [22].

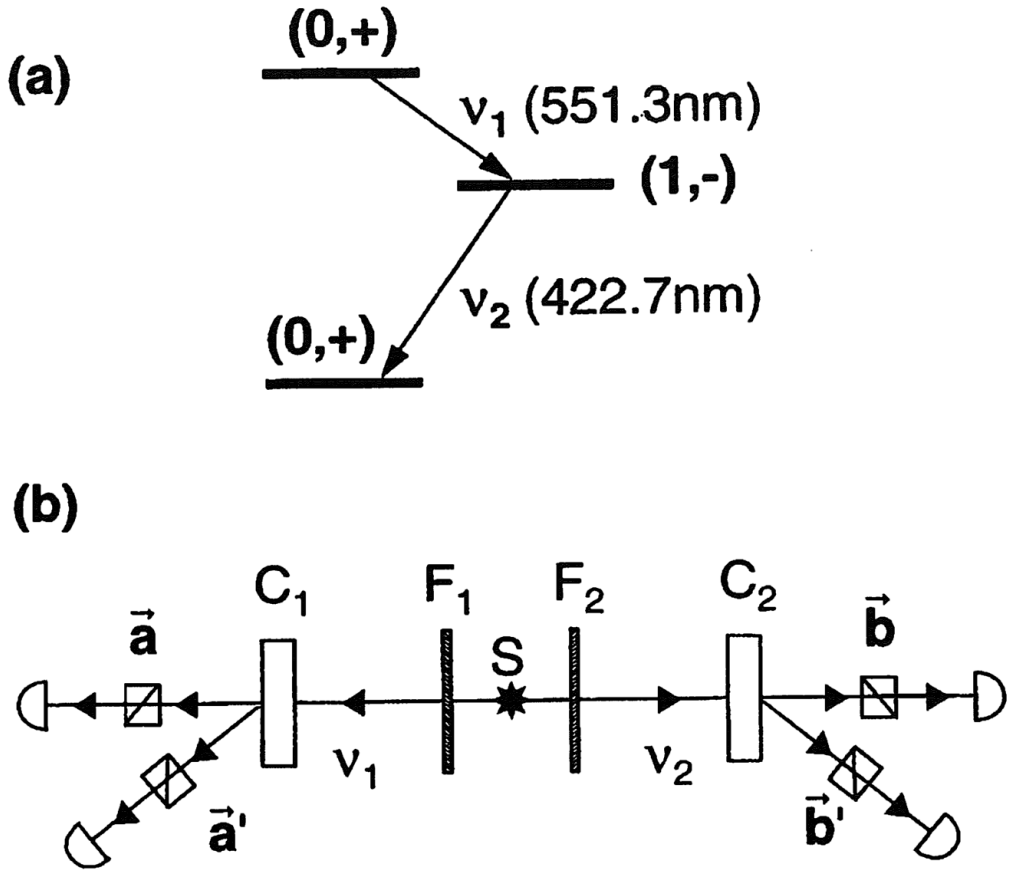

Figure 4
(a) Atomic level scheme in calcium used by Aspect *et al.* to generate correlated photon pairs. The states are labelled by J (the rotational quantum number), and the parity.
(b) Schematic diagram of the apparatus for the third Aspect experiment. *S* is the calcium source. F_1 and F_2 are colour filters that cut out photons at 422.7nm and 551.3nm respectively. C_1 and C_2 are high speed acousto-optic switches. The detectors are set up behind the polarizers with axes along $\vec{\mathbf{a}}$, $\vec{\mathbf{a}}'$, $\vec{\mathbf{b}}$ and $\vec{\mathbf{b}}'$.

The first two experiments approximated to the schematic apparatus shown in Fig. 3. The first experiment checked for violations of a generalised version of the Bell inequality [8]. This type of experiment compares count rates on the detectors $D_1(+1)$ and $D_2(+1)$ for different settings of the polariser angles α and β. The results were found to be in violation of the Bell inequality and in agreement with the quantum mechanical predictions. The second experiment measured the Clauser-Horne-Holt-Shimony inequality of Eq. 5, and found

violations of the inequality in agreement with the predictions of quantum theory [9]. The third experiment contained an additional important feature [10]. In the first two experiments, the polarizers were rotated mechanically between runs, which was a slow process compared to the speed of the photons. In order to establish the non-locality conclusion, it is necessary to change the angles of the polarizers in such a way that it is impossible for signals to pass from the left-hand side to the right-hand side and vice versa. This can be achieved if the polariser angles are changed in a time shorter than L/c, where L is the distance separating the polarizers, and c is the velocity of light. In Aspect's experiment, L/c was 40ns, which therefore requires switching of the polarizers in less than 40ns. The solution found by Aspect *et al.* is shown schematically in Fig. 4(b). Instead of trying to rotate the polarizers very fast, they switched the direction of the beam at rates up to 50MHz with acousto-optical modulators[9]. They arranged the apparatus so that the two switches C_1 and C_2 deflected the photons towards polarizers set at different angles. The results obtained were again in violation of the Bell inequalities, and in this way, they were able to confirm that the results for ν_1 and ν_2 are correlated faster than c.

3.3 More recent experiments

After the publication of the Aspect experiments, other workers have investigated different types of correlated photon sources. Most of the more recent work has concentrated on the correlated photons generated by techniques of nonlinear optics. These experiments rely on the properties of second-order nonlinear crystals that are capable of converting a single photon at angular frequency ω_0 into two photons at frequencies ω_1 and ω_2. Conservation of energy in the process requires that $\omega_0 = \omega_1 + \omega_2$, while conservation of momentum requires $\mathbf{k_0} = \mathbf{k_1} + \mathbf{k_2}$, where $\mathbf{k_i}$ is a vector of length $\omega_i n_i / c$ pointing in the direction of photon i, n_i being the refractive index of the crystal for that particular direction. A number of experiments of this kind have been performed. In 1990, Rarity and Tapster reported a violation of the Bell inequality be several standard deviations [23], while in 1993, Kiess *et al.* improved the accuracy of the violation to 22 standard deviations [24]. In the most recent work, Bell's inequality is observed to be violated when the photons are separated by over 4km [25].

One interesting development has been the work of Ou *et al*, who have managed to perform an EPR experiment with continuous variables [26]. This is significant because all the other experiments are of the EPRB type, where measurements on the particle can only yield one of two results (+1 or −1 in the notation used throughout this article). By contrast, the experiment of Ou *et al.* measured variables with a continuous spectrum, as was considered in the original EPR paper [2]. As explained in §2, EPR experiments in themselves cannot distinguish between the quantum and LHV theories, and as yet, there is no generalisation of the Bell inequalities for dynamical variables. However, it is noteworthy that the experiment originally proposed by Einstein, Podolsky and Rosen has now finally been performed.

3.4 Critique of the experiments

The far-reaching implications of the violation of Bell's inequality have prompted a number of authors to examine whether there might be a flaw in either the reasoning or the experimental

[9] These devices contain special crystals that can deflect light through an angle of 6° when an ultrasonic wave is induced by an electrical transducer. The efficiency of the deflection process is proportional to the voltage applied. Thus by modulating the applied voltage at 50MHz, it is possible to switch the direction of the beam at the modulation frequency.

method. In regard to the experimental methodology, the main issue is that the Bell experiments are gedanken experiments, while the actual experiments only approximate (often rather imperfectly) to the idealised case. It is in fact possible to construct complicated LHV models that explain the results of any individual experiment published to date. This has led Santos to argue that "no experimental test of a Bell inequality has shown a true refutation of LHV theories or local realism. Furthermore, no experiment, performed or planned, is able to show the contradiction between quantum mechanics and LHV theories. Only highly idealised (gedanken) experiments have shown that conflict. Consequently I can safely claim that the problem of whether a local realistic picture of the physical world is possible remains open" [27]. It should be pointed out, however, that the LHV model constructed to explain a particular experiment does not explain the others, while quantum mechanics agrees with all of them.

One major objection that is made against the experiments reported to date is that the detectors used are highly inefficient [28]. It is also argued that it is not justified to equate the counting rates of the detectors with the underlying probabilities required to evaluate the Bell inequality [29]. Moreover, it is pointed out that the cascade sources are inherently inefficient, because the photons are emitted in all directions, and only a small fraction of them are counted by the detectors. Kwiat *et al.* have recently described how a number of these experimental loopholes can be overcome [30]. Other objections are of a more fundamental nature. Barut has questioned whether it is really permissible to extrapolate between repeated events and individual events [31]. One might also object that the whole quantum mechanical argument is logically inconsistent because the measuring apparatus is treated classically [32].

This is by no means an exhaustive list of the possible objections made to the experiments. Bell summed up his own attitude to the more practical objections as follows: "Although there is an escape route there, it is hard for me to believe that quantum mechanics works so nicely for inefficient practical set-ups and is yet going to fail badly when sufficient refinements are made" [33]. In §4 I will discuss in more detail what the experiments actually establish.

3.5 Bell's theorem without inequalities

Several authors have explored ways of stating Bell's theorem in a more emphatic way than by means of an inequality. The first to do this were Greenberger, Horne and Zeilinger, who devised a three-particle gedanken experiment (the GHZ experiment) which gives completely different results for the LHV and quantum mechanical models [34]. Unfortunately, it is difficult to make a correlated three-particle source in the laboratory, and the GHZ experiment has not been performed at the time of writing. More recently, Lucien Hardy has devised an either/or version of the Bell's theorem which works with two correlated particles [35]. A non-technical description of this gedanken experiment is given in the chapter by P. Pliska [18]. In the Hardy experiment, quantum mechanics predicts that a particular result will occur with probability up to 9%, whereas the LHV theory predicts that this result can never occur. The experiment has recently been performed by two research groups, and the events forbidden by the LHV theory have in fact been observed with a probability in very good agreement with quantum mechanics [36]. The experimental confirmation of Hardy's analysis is a very beautiful demonstration of non-local interactions, although it is arguable whether it actually proves the point any more definitively than the Aspect and other experiments that establish the violation of the Bell inequality [37].

4. Discussion

The EPR paradox has been a point of argument among physicists for 60 years now, and the debate still rages. Bell's theorem adds a new dimension to the discussion. The locality issue was raised by the EPR experiments, and the Bell experiments should be seen as a test to check whether Einstein's solution to the EPR paradox is valid. Let us suppose that the experiments can be improved such that the practical loopholes discussed in §3.4 are sorted out. What, then, do the experimental confirmations of Bell's theorem tell us ?

D'Espagnat argued that the violation of Bell's inequality implies that the LHV approach is incorrect, and thus that some of the premises of the LHV theories must be wrong [20] [10]. He listed these premises as: realism, induction, and separability. In his words, realism is *the doctrine that regularities in observed phenomena are caused by some physical reality whose existence is independent of human observers*. Induction implies that *inductive inference is a valid mode of reasoning and can be applied freely, so that legitimate conclusions can be drawn from consistent observations*. Separability states that *no influence of any kind can propagate faster than the speed of light*. Few scientists are willing to renounce the validity of the inductive method, and so the issue is usually discussed in terms of the first and third premises, which together are called *local realism*. Many authors would agree with Bell that the weak link of local realism is the *local* part (i.e. separability in d'Espagnat's terminology). Referring to Bohm's pilot wave hypothesis [5], Bell stated that *a hidden variable interpretation of elementary quantum theory has been explicitly constructed. That particular interpretation has a grossly non-local structure. This is a characteristic, according to the result to be proved here, of any such theory which reproduces exactly the quantum mechanical predictions*[4].

In asserting that quantum mechanics was incomplete, Einstein's primary concern was to safeguard realism in response to some of the philosophical ideas that were entering physics with the Copenhagen interpretation. Since we only obtain deterministic knowledge of the properties of microscopic systems by measuring them, and the results obtained depend on the way the measurement is made, it is easy to argue that *real physical properties are possessed only by the combined system of microscopic object plus measuring apparatus*[39]. A natural progression of this line of thought is to argue that before the measurement is made, the particle's state is unreal, or indeterminate, implying that objective reality does not exist at the microscopic level. This approach is summed up succinctly by Bell: *Making a virtue of necessity, and influenced by positivistic and instrumentalist philosophies, many came to hold not only that it is difficult to find a coherent picture but that it is wrong to look for one - if not actually immoral then certainly unprofessional. Going further still, some asserted that atomic and subatomic particles do not have any definite properties in advance of observation .... Indeed even the particles are not really there* [40].

The EPR paradox was devised to counter this line of reasoning. The aim was to prove the logical inconsistency of the Copenhagen approach by showing that it led to the unacceptable notion of action at a distance. The definition of reality given in the EPR paper is as follows: *If, without in any way disturbing a system, we can predict with certainty (i.e. with probability equal to unity) the value of a physical quantity, then there exists an element of*

[10] This opinion is not universally accepted. For example, Brody has collected together a number of alternative derivations of the Bell inequality, and points out that only *some* of them are based on the assumption of local hidden variables [38]. Brody himself argues that the key issue is *joint-measurability*, which is the assumption that two or more physical quantities can be measured without mutual interference.

physical reality corresponding to this physical quantity [2]. The fact that the experiments now seem to show that nature does not subsribe to this definition of physical reality does not in any way imply that objective reality does not exist at the microscopic scale. The authors of the EPR paper acknowledge that their definition was only a *sufficient condition*, which therefore did not exhaust *all possible ways of recognizing a physical reality*.[11] Einstein was a physicist and not a philosopher, and so it is not surprising that he attempted to refute the philosophical intrusions into physics with physical arguments. The notion of local realism mixes a metaphysical notion (realism) with a physical one (locality). Metaphysical reality is here understood *in the radical sense of differing from nonreality* [41]. Thus it is not necessary to defend Einsteinian local realism (i.e. space-time causality) in order to remain a metaphysical realist. Dewdney *et al.* put it this way: *the fundamentally new feature of matter introduced in the quantum theory is a kind of wholeness in which the behaviour of an individual particle is irreducibly connected with its context (expressed through the wavefunction), evidenced in the two-particle case by the existence of non-local connections .... the elements of reality in quantum theory are essentially different to the elements of reality in Newtonian physics .... In our opinion, assuming reality has this fundamentally new feature of wholeness is preferable to assuming that it does not exist except when we are looking for it* [42]. There is much confusion in the literature, and frequently one finds scientists who presumably want to maintain metaphysical realism, and in fact only defend the limited notion of Einsteinian local realism. We should be careful to avoid confusing Einstein's common sense objections to the philosophical overtones of the Copenhagen approach with his (possibly flawed) physical argumentation.

It is worth closing this section by briefly considering whether the EPR correlations have any practical consequences. It might be thought that the apparently instantaneous correlations between separated measurements could provide a mechanism for the transfer of information faster than the speed of light. However, this is not in fact possible, because the sequence of events registered by either detector is completely random, and the correlations between the two sets of results are only apparent when they are compared. Since this comparison must be done by ordinary means of communication, one gains nothing from the EPR scheme [7,37]. The use of EPR correlations in quantum cryptographic-key-sharing and quantum teleportation schemes has been discussed in the literature [43,44]. However, in neither case is it possible to achieve faster-than-light communication, because the information cannot be decoded until a classical signal has been sent.

5. Conclusions

This article has meant to be a brief survey of the very active research field based around Bell's theorem. Bell's theorem explores whether quantum mechanics can be explained in terms of local hidden variables, and at the present time, the experimental evidence is weighing up very heavily against the LHV interpretation. Most authors (including Bell himself) interpret this evidence as implying that non-local correlations exist in nature. The notion of non-locality is a very strange concept and appears to break all our prior assumptions about "no action at a distance". However, as Mermin puts it: *if there is spooky action at a distance, then, like other spooks, it is absolutely useless except for its effect, benign or otherwise, on our state of mind* [7]. Einstein, Podolsky and Rosen's attempt to prove the incompleteness of the Copenhagen approach might appear to have back-fired on them. In fact, the debate about Einstein's more fundamental objections to the philosophical extrapolations of the Copenhagen approach continue as before. The difference is that any new discussion must include reference

[11] Private communication by Lucien Hardy.

to Bell's theorem, and explain the beautiful experimental work that has been performed as a result of Bell's insight.

Appendix: Polarised light

The classical description of light is based on electromagnetic waves, consisting of oscillating electric and magnetic fields. The electric and magnetic fields are perpendicular to each other, and both are perpendicular to the direction of the light wave. Thus if one takes the z-axis as the direction of propagation of the light, then the electric and magnetic fields must be oscillating at right angles to each other in the (x,y) plane. The *polarisation* of the light refers to the direction of the electric field. Direct light from the sun is *unpolarised*, which means that there is no preferred direction for the electric field. By contrast, *linearly polarised* light consists of electromagnetic waves in which the direction of the electric field is fixed with respect to a given axis. The polarisation is specified by the angle θ between the electric field and the reference axis. If the reference axis is the y-axis, then two common situations are *vertically* and *horizontally* polarised light, which consists of light polarised along the y-axis ($\theta = 0°$) or the x-axis ($\theta = 90°$) respectively. The polarisation angle θ can be determined with polaroid sheets. These are filters which have the property of absorbing electric fields perpendicular to the internal axis of the sheet. The fraction of the incident light which passes through the polaroid is given by $\cos^2\theta$, where the reference axis is now the known internal axis of the polaroid. Polaroids are commonly used in sunglasses, and are used to cut out sunlight reflected from roads or the sea. The light becomes horizontally polarised by the reflection and the sunglasses have polaroids with vertical axes to absorb the reflected light.

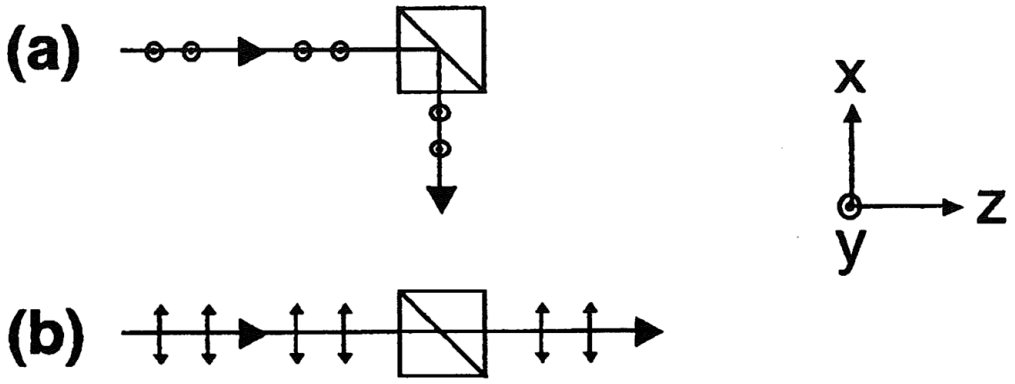

Figure 5
Effect of a polarising beam-splitter on linearly polarised light propagating in the +z direction. (a) Vertically polarised light. (b) Horizontally polarised light.

A polarising beam splitter operates slightly differently from a polaroid in that it is able to separate the two polarisation components of the light (see Fig.5). The axis of the beam splitter is defined by axes of the crystal from which it is made, and the beam splitter has the property that it transmits horizontally polarised light, but reflects vertically polarised light. If θ is the angle between the polarisation and the beam-splitter axis, then the fraction of the light transmitted is $\sin^2\theta$, and the fraction reflected is $\cos^2\theta$. After the beam-splitter, the transmitted light is horizontally polarised, and the reflected light is vertically polarised. If unpolarised light is incident on the beam splitter, 50% of the light will be transmitted and 50% will be reflected. This follows because the unpolarised light has a random polarisation (i.e. random θ), and the average values of $\sin^2\theta$ and $\cos^2\theta$ are both ½.

References

[1] J.S. Bell, *Speakable and unspeakable in quantum mechanics*, Cambridge University Press, Cambridge (1987), p. 89.
[2] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[3] D. Bohm, *Quantum Theory*, Prentice-Hall, Englewood Cliff, New Jersey (1951).
[4] J.S. Bell, *op. cit.*, chapter 2; originally published in Physics **1**, 195 (1964).
[5] D. Bohm, Phys. Rev. **85**, 166 and 180 (1952). A simplified explanation of this theory has recently been given by D.Z. Albert in Scientific American, **270,** no. 5 (May), 32 (1994).
[6] A. Suarez, this volume, chapter 10.
[7] N.D. Mermin, Physics Today, **38**, no. 4 (April), 38 (1985).
[8] A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. **47**, 460 (1981).
[9] A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. **49**, 91 (1982).
[10] A. Aspect, J. Dalibard, and G. Roger, Phys. Rev. Lett. **49**, 1804 (1982).
[11] The arguments followed here are a simplified discussion of those presented by A.I.M. Rae, *Quantum physics: illusion or reality?*, Cambridge University Press (1986).
[12] P.E. Hodgson, The Month **17**, 216 (1984).
[13] See T.A. Brody, *The Philosophy Behind Physics*, ed. L. de la Peña and P.E. Hodgson, Springer-Verlag, Berlin (1993), chapters 13 and 14.
[14] A.I.M. Rae, *op. cit.*, p.28.
[15] A. Einstein, Letter to M. Born, 3 March 1947, published in *The Born-Einstein Letters*, Ed. M. Born, Macmillan, London (1971) p.158.
[16] A derivation for the case of photons with perpendicular polarisations is given in A.I.M. Rae, *op. cit.*, chapter 3. J.S. Bell gives a similar proof for Stern-Gerlach experiments in ref.[1], chapter 16.
[17] C. Jack, Physics World, **8**, no. 4 (April), 39 (1995).
[18] P. Pliska, this volume, chapter VIII.
[19] J.F. Clauser, M.A. Horne, A. Shimony, and R.A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[20] B. d'Espagnat, Scientific American, **241**, no. 5 (November), 128 (1979).
[21] A.I.M. Rae, *op. cit.*, pp.52, 47.
[22] A. Aspect, Doctor of Physical Sciences Thesis, University of Paris-sud, Centre d'Orsay, 1983, pp 75-77.
[23] J.G. Rarity and P.R. Tapster, Phys. Rev. Lett., **64**, 2495 (1990).
[24] T.E. Kiess, Y.H. Shih, A.V. Sergienko, and C.O. Alley Phys. Rev. Lett., **71**, 3893 (1993).
[25] P.R. Tapster, J.G. Rarity, and P.C.M. Owens, Phys. Rev. Lett., **73**, 1923 (1994).
[26] Z. Y. Ou, S.F. Pereira, H.J. Kimble, and K.C. Peng, Phys. Rev. Lett., **68**, 3663 (1992).
[27] E. Santos, Phys. Rev. A, **46**, 3646 (1992).
[28] M. Ferrero, T.W. Marshall, and E. Santos, Am. J. Phys. **58**, 683 (1990).
[29] E. Santos, Phys. Rev. Lett. **66**, 1388 and 2702 (1991).
[30] P.G. Kwiat, P.H. Eberhard, A.M. Steinberg, and R.Y. Chiao, Phys. Rev. A **49**, 3208 (1994).
[31] A.O. Barut, Foundations of Physics, **22**, 137 (1992).
[32] J.S. Bell, Physics World, **3**, no. 8 (August), 33 (1990).
[33] J.S. Bell, *op. cit.* p.154.
[34] D.M. Greenberger, M.A. Horne, and A. Zeilinger, Physics Today, **46**, no. 8 (August), 22 (1993). The experiment is also discussed in ref. [18].
[35] L. Hardy, Phys. Rev. Lett., **68**, 2981 (1992); **71**, 1665 (1993).
[36] J.R. Torgerson, D. Branning, C.H. Monken, and L. Mandel, Phys. Lett. A, **204**, 323 (1995); D. Boschi, D. De Martini, and G. Di Giuseppe, (unpublished).
[37] N.D. Mermin, Am. J. Phys. **62**, 880 (1994).

[38] T.A. Brody, *op. cit.*, chapters 16-18.
[39] A.I.M. Rae, *op. cit.*, p.48.
[40] J.S. Bell, *op. cit.*, p.142. The statements were made in the context of discussing Stern-Gerlach experiments, which measure the deflection of spin-½ particles in a non-uniform magnetic field. It is found that the particles are deflected either up or down. This is equivalent to the optical case, where the photons are either transmitted or reflected.
[41] S.L. Jaki, Essay "Determinism and Reality", in *Great Ideas Today 1990*, Chicago Encyclopaedia Britannica 1990, p.277.
[42] C. Dewdney, P.R. Holland, A. Kyprianidis, and J.P. Vigier, Nature **336**, 536 (1988).
[43] A.K. Ekert, J.G. Rarity, P.R. Tapster, and G. Massimo Palma, Phys. Rev. Lett. **69**, 1293 (1992).
[44] C.H. Bennett, Physics Today, **48**, no. 10 (October), 24 (1995).

Chapter X

Nonlocal phenomena:

physical explanation and philosophical implications

Antoine Suarez
Institute for Interdisciplinary Studies, P.O. Box 304, 8044 Zurich, Switzerland, E-mail: suarez@leman.ch

Abstract

The nonlocality of quantum mechanics is so important and counterintuitive that it is important to perform as many experiments as possible to rule out definitively any other alternative explanations. In this paper we propose a new gedanken-experiment with photons in which the timing of the photon impacts at beamsplitters is considered, an issue which has not been explored in previous nonlocality experiments. The experiment is analyzed with the conventional quantum mechanical description, and also according to a new nonlocal realistic model which incorporates superrelativistic features. The superrelativistic model explains the phenomena by means of faster-than-light influences without needing a universal order of succession (quantum aether), and without superluminal signaling either. For multi-particle experiments, in which the particles undergo successive impacts at beam-splitters or polarizers, it predicts correlations which depend on the time-ordering of the arrival of the particles at the different choice devices (polarizers, beam-splitters). This prediction clearly conflicts with quantum mechanics. Experimental work in progress to carry out the gedanken-experiment is described, and the implications of the results are discussed, depending on whether they uphold quantum mechanics or the supperrelativistic view. In both cases, we expect that the results will tell us something new about the nature of space-time. Finally some philosophical implications are discussed.

1. Introduction

This century has given us two theories that have been most successful in calculating and predicting phenomena: relativity and quantum mechanics. The attempt to unify them is a characteristic of physics research today.

A famous conclusion in special relativity is that nothing in nature should happen faster-than-light. Theories which accept this assumption are often labeled local-realistic. After 1964, the local-realistic worldview has been considerably disturbed by developments related to the controversy on the Einstein-Podolsky-Rosen (EPR) assumptions [1]. Concerning the correlations implied by two-particle and multi-particle superposition (often referred to as „entanglement“), work by John Bell [2], by Daniel M. Greenberger, Michel A. Horne, and Anton Zeilinger [3], and by Lucien Hardy [4] and Thomas F. Jordan [5] pointed out that local-

1 A. Einstein, B. Podolsky, and N. Rosen, *Physical Review*, **47**, 777-780 (1935).
2 J.S. Bell, *Physics* **1**, 195-200 (1964); *Speakable and unspeakable in quantum mechanics*, Cambridge: University Press, 1987.
3 D.M. Greenberger, M. A. Horne, A. Zeilinger, *Physics today*, August, p. 24 (1993).
4 L. Hardy, *Physical Review Letters*, **71**, 1665 (1993).
5 T. F. Jordan, *Physical Review A*, **50**, 62-66 (1994).

realistic theories cannot account for the quantum mechanical description of physical reality. This implies that faster-than-light influences in phenomena have to be admitted, even if they cannot be used by human observers for practical purposes (impossibility of „superluminal signaling"). This feature is now mostly referred to as nonlocality.

First order correlations related to one-particle superposition imply the impossibility of attributing to a particle only one determined trajectory. Single-particle interference fringes already show that quantum mechanical entities are not localized: a "particle" may be in different places at one time. However, one-particle nonlocality does not involve faster-than-light influences at all, and is a subluminal nonlocality. Recent arguments on „single-photon nonlocality" either use an ambiguous definition of single-particle experiments [6] or refer to effects which can be explained without needing superluminal influences [7]. Throughout this article the term *nonlocality* is used with the meaning of *superluminal nonlocality*, unless stated otherwise.

In spite of the loopholes in the experiments [8], the existence of superluminal nonlocality is today largely admitted: most physicists will not be surprised, if a future "loophole free" Bell experiment [9] definitely demonstrates the violation of the locality criteria (Bell's inequalities or others). We would like to notice that the rejection of faster-than-light influences in nature (the special relativity postulate), does not follow at all from observation. Strictly speaking, the negative results of Michelson-Morley and related experiments only imply that human observers cannot communicate faster-than-light (we will come back to this point later). This view is perfectly coherent with quantum nonlocality.

However, the heart of relativity is the principle that there is no absolute space-time, no "aether". Simultaneity depends on the observer's state of movement, and the order of succession of two space-like separated events may change if one changes the inertial frame. This principle seems to be at odds with the quantum mechanical description of the "measurement process". Consider for instance the orthodox quantum mechanical description of the perfect EPR correlations in two-particle experiments with entangled polarized photons: The spin operator related to a measuring apparatus with two parallel oriented polarizing beam splitters has two eigenvectors $|+1,+1\rangle$ and $|-1,-1\rangle$, representing two orthogonal quantum eigenstates; the measurement causes the entangled state to jump into either the state $|+1,+1\rangle$ or the state $|-1,-1\rangle$ instantaneously, where the first state means that both photons are detected in the detectors monitoring the transmitted output ports, and the second one that both photons are detected in the detectors monitoring the reflected output ports[10]. Consequently, the measurement produces events which are simultaneously strictly correlated in space-like separated regions. But in which inertial frame are these correlated events simultaneous? Quantum mechanics does not answer this question.

---

6 L. Hardy, The EPR argument and nonlocality without inequalities for a single photon, in D.M. Greenberger and A. Zeilinger (Eds.), *Fundamental Problems in Quantum Theory*, New York: New York Academy of Sciences, 1995, p. 600-615.

7 D. Home, and G.S. Agarwal, *Physics Letters A*, **209**, 1-5 (1995).

8 A. Aspect, P. Grangier and G. Roger, *Physical Review Letters*, **49**, 91-94, (1982). A. Aspect, J. Dallibard and G. Roger, *Physical Review Letters*, **49**, 1804-1807 (1982). J.G. Rarity and P. R. Tapster, *Physical Review Letters*, **64**, 2495-2498 (1990). E. Santos, *Physical Review Letters* **66**, 1388-1390 (1991), and *Physical Review Letters* **68**, 2702-2703 (1992).

9 P.G. Kwiat, P. H. Eberhard, A. M. Steinberg, and R. Y. Chiao, *Physical Review A*, **49**, 3209-3220 (1994).

10 For the definition of horizontal/vertical polarization in relation with transmission/reflection from a polarizing beam-splitter see M. Fox, *Optical tests of Bell's theorem*, this volume, chapter IX, Appendix.

Moreover, because each measurement of polarization may lie outside the other's light cone (e.g. the two measurements may be space-like separated events), the measurement which is considered as the cause of the "jump" in a certain inertial frame, is no longer the cause in another inertial frame. For one observer the value measured at side 1 depends on which value has been measured at side 2, and for another observer the value measured at side 2 depends on which value has been measured at side 1. Different observers are led to contradictory descriptions of the same reality. That is the reason why there still seems to be no consistent relativistic interpretation of the "quantum jump" (also referred to as "reduction of the wave packet" or "wavefunction collapse") associated with the measurement process, or why the notion of collapse appears to have no meaning in a relativistic context [11].

Interpretations of quantum mechanics which account for superluminal non-locality and in which the cause-effect links do not depend on the observer's state of motion, have been developed. The first causal and explicitly nonlocal interpretation was formulated in 1952 by David Bohm [12], and has subsequently been further elaborated by several authors [13]. Bohm's theory relies on the idea of a quantum potential acting upon the particles, and is referred to as the *causal model*. To avoid causal paradoxes, one is led to assume an absolute space-time or quantum aether, or more accurately, to a universal order of succession [14]. Consequently, at the fundamental level, the theory cannot be considered to be relativistic or Lorentz-invariant. In this approach, relativity or Lorentz-invariance comes out as a statistical effect, not as an absolute one. Until now, the causal model does not give predictions conflicting with conventional quantum mechanics.

More recently, the wish for realistic and causal interpretations has generated proposals that conflict not only with Einstein's locality postulate, but also with the predictions of conventional quantum theory. The so called *dynamical reduction theories* intend to construct *a completely satisfactory relativistic theory of state vector reduction*, in which the state vector represents reality, and the conflict with fundamental Lorentz invariance is avoided. These alternative realistic theories make predictions for two-slit neutron interference experiments which deviate from the quantum mechanical predictions [15]. Curiously a conflict concerning single particle superposition appears, without being possible to remove completely the causal ambiguities concerning entanglement. *Eberhard's* proposal explains the EPR correlations through causal effects propagating forward in time in a privileged inertial frame with velocity V much larger than the velocity of light c. In this model *"superluminal signaling"* is possible. Although Eberhard's theory conflicts with Einstein's view, it remains local, since the influences follow well-defined trajectories in space-time. No violation of the predictions of quantum theory is expected, if the time intervals between measurements are larger than the time necessary to travel between the locations of these measurements at the velocity $V>c$ in the privileged rest frame. If the parameter V (the presumed new constant of nature defining the upper velocity limit) is not too large, by performing experiments with large distances d

---

11 A. Peres, Relativistic Quantum Measurement, in D.M. Greenberger and A. Zeilinger (Eds.*), Fundamental Problems in Quantum Theory*, New York: New York Academy of Sciences, 1995, p. 445.

12 D. Bohm, *Physical Review*, **85**, 166-193 (1952)

13 See the contributions in: A. van der Merwe, F. Selleri and G. Tarozzi (eds.) *Microphysical Reality and Quantum Formalism*, Volume 2, Part: Is a causal interpretation of quantum theory possible? Dordrecht: Kluwer, 1988. Further: F. Selleri, *Quantum Paradoxes and Physical Reality*, Dordrecht: Kluwer, 1990, p. 123 and 240. D. Bohm and B.J. Hilley, *The Undivided Universe*, New York: Routledge, 1993.

14 D. Bohm and B.J. Hilley, *The Undivided Universe*, New York: Routledge, 1993, p. 290-295. B. Hiley in: P.C.W. Davies & J.R. Brown, *The ghost in the atom*, Cambridge: University Press, 1986, p. 142.

15 G.C. Ghirardi, and P. Pearle, *International Centre for theoretical Physics,* Trieste, Preprint IC/91/22-23 (1991).

and time intervals smaller than *d/V* between measurements, violations of quantum theory may be expected[16].

There is also a Lorentz-invariant interpretation of quantum mechanics, Costa de Beauregard's "retrocausation", which violates the causality principle[17]. "Retrocausation" means the possibility that our present decisions influence the past propagating backwards within the light cone of present events. Thereby one accepts that an effect can exist before its cause, or, equivalently, that time is a stronger concept than causality.

From the point of view of the state of movement of the polarizers or beam-splitters, it can be said that the quantum formalism does not depend at all on the inertial frames of the polarizers or beam-splitters. Thus, in two-particle experiments with entangled polarized photons, the quantum mechanical correlation coefficient is assumed to be given by a Lorentz-invariant expression of the type $E(\theta) = \cos 2\theta$ [18], where θ is the angle between the axes of the polarizers. Consequently, for $\theta = 0°$, one should always get perfectly correlated results (either both photons are transmitted, or both are reflected), and for $\theta = 90°$, perfectly anti-correlated results (one photon is transmitted, and the other is reflected). This would be valid also in experiments with fast-moving polarizing beam-splitters. Bohm's causal model, and the other two alternative theories previously referred to, also predict frame-independent correlations.

In this article we try to get a better understanding of nonlocal phenomena by means of a new gedanken experiment. In Section 2 we stress that quantum nonlocality teaches us something about the nature of space-time we do not yet fully understand. In order to get deeper insight it could be profitable to do experiments in which the relevant space-time parameters are varied to the maximal extent possible. In Section 3 we present a new gedanken experiment in which the particles undergo successive impacts at beam-splitters or polarizers, thereby allowing three different time sequences of impacts at the beam-splitters. For such arrangements quantum mechanics predicts nonlocal correlations which are independent of the time ordering in which the particles arrive at the different beam-splitters or polarizers. On the basis of this experiment we show in Section 4 that nonlocal phenomena can be explained while respecting the causality principle, by means of faster-than-light influences, but without needing a quantum aether, and without superluminal signaling either. For this reason the proposed explanation is also referred to as a *superrelativistic* theory. This explanation has the following characteristics:

1) it distinguishes between *choices* (in devices like polarizers or beam splitters) and *detection*, resolving thereby the confusion involved in the concept of "measurement";
2) it introduces as many referential frames as choice devices there are in the setup;
3) assumes values existing before "measurement", and *conditional probabilities*;
4) establishes that the correlation functions depend nonlocally on the state of movement of the choice devices (polarizers, beam-splitters);
5) admits not only one, but *many entanglement rules*.

In Section 5 we present the following theorem of the superrelativistic theory: the nonlocal correlations in two-particle experiments with one of the photons undergoing two

---

16 P. Eberhard, A realistic model for quantum theory with a locality property, in: W. Schommers (ed.), *Quantum theory and pictures of reality*, New York: Springer, 1989, pp.169-215.

17 O. Costa de Beauregard, *Physical Review Letters*, **50**, 867-869 (1983); *Foundations of Physics Letters*, **5**, 489-491 (1992).

18 A. Aspect, P. Grangier and G. Roger, *Physical Review Letters*, **49**, 91-94, (1982).

successive impacts at choice devices, depend on the time ordering in which the particles arrive at the three choice devices. In Section 6 we discuss real experiments in preparation which will allow us to decide between quantum mechanics and the alternative proposal. Independent of the outcome, these experiments will give more insight in the nature of space-time. In Section 7, finally, we highlight that the new proposal integrates and harmonizes the main features of quantum mechanics and relativity within a nonlocal causal model. Thereafter we work out some philosophical implications concerning causality and nonlocality. The theory will be discussed without going into all the subtleties [19].

2. Quantum nonlocality teaches us something about the nature of space-time which is not fully understood at the present moment

A main postulate of conventional quantum mechanics is that only after a specific measurement has been made can we attribute a definite physical property to a quantum system. There are no pre-existing values prior to the measurement[20]. Therefore, the appearance of perfect nonlocal correlations necessarily expresses a link existing between *real* measured values. Consequently one is led to assume that the measurement at one of the regions produces either a value +1 or −1 after taking into account the value that has *actually* been measured in the other region. This means that there is an order of succession, and thus the perfect correlations should disappear in the case of simultaneous measurements. However, according to the superposition principle, the appearance of perfect correlations does not depend at all on the times at which the values are measured, in any inertial frame whatsoever. The superposition principle looks, therefore, to be at odds with the postulate of "no values prior to the measurement".

This argument points out that quantum nonlocality teaches us something about the nature of space-time that we do not yet fully understand. In order to test the general validity of the superposition principle, the ideal experiment would undoubtedly be one in which the observer in region 1 measures the value *before* the observer in region 2 performs his measurement, and vice versa. This would require fast-moving measuring devices, which is impossible to arrange by means of the techniques available today. Nevertheless it is possible to devise experiments which allow for a greater variation in the relevant space-time parameters than in the experiments which have been performed so far. We propose one such experiment in the next section.

3. Experiments with one particle impacting successively at two beam-splitters: quantum mechanics predicts independence of the time ordering.

Quantum mechanics highlights the relationship between superposition and the impossibility to obtain path information (indistinguishability). If one cannot distinguish (even in principle) between different paths from source to detector, the amplitudes for these alternative paths add coherently, and interference fringes appear. If it is possible in principle to distinguish, the interference patterns vanish. The dependence of superposition on

---

[19] A more detailed presentation is given in: A. Suarez, *Non-local phenomena: a superrelativistic theory with many superposition principles theory*, Center for quantum Philosophy, Preprint CQP-960107, 1996

[20] See for instance the contributions of J. Wheeler and R. Peierls in: P.C.W. Davies & J.R. Brown, *The ghost in the atom*, Cambridge: University Press, 1986.

indistinguishability is postulated equally for one-particle and for multi-particle superposition[21].

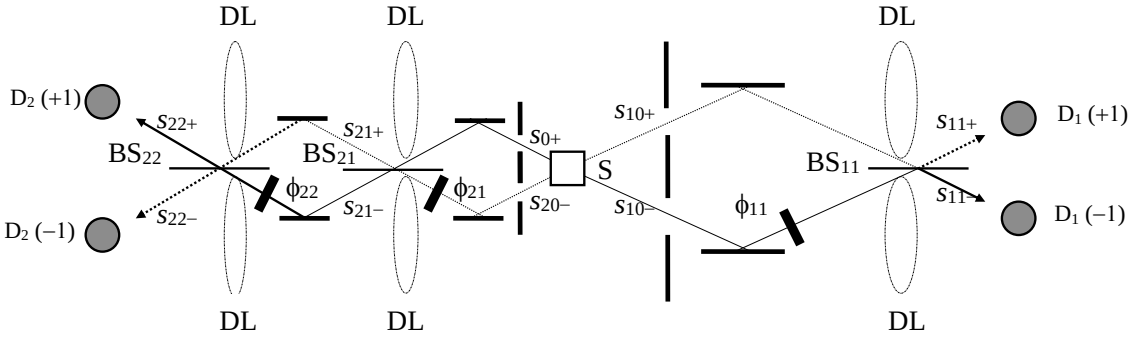

Figure 1:
Twofold double-slit experiment with the left photon impinging successively on two beam splitters. S: source; BS_{11} , BS_{21} and BS_{22}: beam-splitters; D_1 and D_2: detectors; ϕ_{11}, ϕ_{21}, and ϕ_{22}: phase shifters. By changing the length of the different delay lines DL, different time orderings for the impacts at the beam splitters can be arranged. The setup is supposed to be completely symmetric for the different trajectories.

Consider the gedanken-experiment represented in FIG. 1. Two photons emitted back-to-back, can travel by alternative pairs of paths from the source S to either one of the right-hand detectors $D_1(+1)$, $D_1(-1)$ and either one of the left-hand detectors $D_2(+1)$, $D_2(-1)$. Before being detected, photon 1 impacts at beam-splitterBS_{11}, and photon 2 impacts successively at beam-splitters BS_{21} and BS_{22}. By varying the lengths of the photon paths by means of delay lines, one can arrange three different time orderings:

1) the impact on BS_{22} occurs before the impact BS_{11};
2) the impact on BS_{11} occurs before the impact on BS_{21};
3) the impact on BS_{21} occurs before the impact on BS_{11}, and the impact on BS_{11} occurs
before the impact on BS_{22}.

The conventional application of the quantum mechanical superposition principle considers all three time orderings as being equivalent: if it is impossible to obtain path information, the superposition rule applies. The relative time ordering of the impacts at the *choice devices* does not influence the distribution of outcomes; in this respect only indistinguishability matters.

The quantum mechanical description of the experiment follows from the equations:

---

[21] J.D. Franson, *Physical Review Letters* **62**, 2205-2208 (1989). P.G. Kwiat, P. H. Eberhard, A. M. Steinberg, and R. Y. Chiao, *Physical Review A*, **49**, 3209-3220 (1994). D.M. Greenberger, M. A. Horne, A. Zeilinger, *Physics today*, August, p. 24 (1993).

$$
\begin{aligned}
&|\psi\rangle = \frac{1}{\sqrt{2}}|s_{10+}\rangle|s_{20-}\rangle + \frac{1}{\sqrt{2}}|s_{10-}\rangle|s_{20+}\rangle \\
&|s_{10+}\rangle = \frac{i}{\sqrt{2}}|s_{11+}\rangle + \frac{1}{\sqrt{2}}|s_{11-}\rangle \qquad |s_{10-}\rangle = e^{i\phi_{11}}\frac{1}{\sqrt{2}}|s_{11+}\rangle + e^{i\phi_{11}}\frac{i}{\sqrt{2}}|s_{11-}\rangle \\
&|s_{20+}\rangle = \frac{i}{\sqrt{2}}|s_{21+}\rangle + \frac{1}{\sqrt{2}}|s_{21-}\rangle \qquad |s_{20-}\rangle = e^{i\phi_{21}}\frac{1}{\sqrt{2}}|s_{21+}\rangle + e^{i\phi_{21}}\frac{i}{\sqrt{2}}|s_{21-}\rangle \\
&|s_{21+}\rangle = \frac{i}{\sqrt{2}}|s_{22+}\rangle + \frac{1}{\sqrt{2}}|s_{22-}\rangle \qquad |s_{21-}\rangle = e^{i\phi_{22}}\frac{1}{\sqrt{2}}|s_{22+}\rangle + e^{i\phi_{22}}\frac{i}{\sqrt{2}}|s_{22-}\rangle
\end{aligned}
\tag{1}
$$

where the factor $\frac{1}{\sqrt{2}}$ accounts for the assumption that each beam-splitter is 50-50, the complex factors $e^{i\phi_{jk}}$ account for the effects of the different phase shifters ϕ_{11}, ϕ_{21} and ϕ_{22}. Further it is assumed that reflections at a beam-splitter cause a 90° phase shift (which accounts for the factor *i*).

Summing in (1) the amplitudes for the different alternative paths from source to detector, and squaring the moduli, one obtains the probabilities of the different possible detection outcomes:

$$
\begin{aligned}
P(+1,+1) &= \frac{1}{16}\left|e^{i\phi_{21}} + e^{i\phi_{21}}e^{i\phi_{22}} + e^{i\phi_{11}} - e^{i\phi_{11}}e^{i\phi_{22}}\right|^2 \\
&= \frac{1}{4} - \frac{1}{8}\cos(\phi_{21} - \phi_{11} - \phi_{22}) + \frac{1}{8}\cos(\phi_{21} - \phi_{11} + \phi_{22})
\end{aligned}
$$

$$
\begin{aligned}
P(-1,-1) &= \frac{1}{16}\left|e^{i\phi_{21}} - e^{i\phi_{21}}e^{i\phi_{22}} - e^{i\phi_{11}} - e^{i\phi_{11}}e^{i\phi_{22}}\right|^2 \\
&= \frac{1}{4} - \frac{1}{8}\cos(\phi_{21} - \phi_{11} - \phi_{22}) + \frac{1}{8}\cos(\phi_{21} - \phi_{11} + \phi_{22})
\end{aligned}
$$

$$
\begin{aligned}
P(+1,-1) &= \frac{1}{16}\left|e^{i\phi_{21}} - e^{i\phi_{21}}e^{i\phi_{22}} + e^{i\phi_{11}} + e^{i\phi_{11}}e^{i\phi_{22}}\right|^2 \\
&= \frac{1}{4} + \frac{1}{8}\cos(\phi_{21} - \phi_{11} - \phi_{22}) - \frac{1}{8}\cos(\phi_{21} - \phi_{11} + \phi_{22})
\end{aligned}
$$

$$
\begin{aligned}
P(-1,+1) &= \frac{1}{16}\left|e^{i\phi_{21}} + e^{i\phi_{21}}e^{i\phi_{22}} - e^{i\phi_{11}} + e^{i\phi_{11}}e^{i\phi_{22}}\right|^2 \\
&= \frac{1}{4} + \frac{1}{8}\cos(\phi_{21} - \phi_{11} - \phi_{22}) - \frac{1}{8}\cos(\phi_{21} - \phi_{11} + \phi_{22})
\end{aligned}
$$

and the correlation coefficient $E^{QM}(\phi_{11}, \phi_{21}, \phi_{22})$:

$$
\begin{aligned}
E^{QM}(\phi_{11}, \phi_{21}, \phi_{22}) &= P(+1,+1) + P(-1,-1) - P(+1,-1) - P(-1,+1) \\
&= \frac{1}{2}\left(\cos(\phi_{11} - \phi_{21} - \phi_{22}) - \cos(\phi_{11} - \phi_{21} + \phi_{22})\right)
\end{aligned}
\tag{2}
$$

For $\phi_{22} = 90°$, and $\theta_{21} = \phi_{21} - 90°$:

$$E^{QM}(\phi_{11}, \theta_{21}) = -\cos(\phi_{11} - \theta_{21})$$

which is an equation of the same type as that leading to violations of Bell's inequalities in nonlocality experiments with two choices. According to the conventional superposition principle, equation (2) holds for all three time orderings, i.e. the correlation coefficient depends nonlocally on the parameters of the phase shifters, but not on the time sequence of the impacts at the beam-splitters.

By means of the successive impacts of one of the photons, it becomes possible to arrange more complicated sequences of the times at which the photons arrive at the beam splitters. This substantially increases the variability of the relevant space-time parameters. Moreover, as we will see in the coming sections, it is possible to elaborate an explanation for this new gedanken experiment which makes predictions that contradict the conventional superposition principle. In light of these possibilities, and of the unusual features of the quantum mechanical postulates pointed out in the preceding section, it seems reasonable to explore this new experimental direction, even if one's personal understanding is that quantum mechanics is so strong and consistent that it is unlikely that it will break down in any of these experiments.

4. Principles of an alternative nonlocal, causal theory, assuming many superposition principles

We present below the principles of an alternative nonlocal theory, which yields correlations depending on the time ordering of the impacts in experiments with particles impacting successively at several beam-splitters or polarizers. Accordingly, the theory is testable against quantum mechanics through real experiments. The theory is also referred to as *superrelativistic*, because it assumes superluminal nonlocality (faster-than-light influences), but does not require any absolute space-time or quantum aether.

4.1. Distinguishing between choice devices and detectors

A quantum mechanical "measuring apparatus" includes two main elements: a device such as a beam-splitter or a polarizer, that allows us to define several alternative outcomes, and the detectors monitoring the output ports of the beam-splitter or polarizer. Nevertheless, the quantum mechanical description does not distinguish between the event happening in the beam-splitter or polarizer and the detection event. Both appear somewhat confused within the concept of "measurement". So, for instance, the ensemble of a polarizing beam splitter and the corresponding two detectors is referred to as "polarimeter"[22], or a "detector" is considered to consist of "two phase shifters, a beam-splitter, and two particle counters"[23].

Consider a generalization of the experiment in FIG. 1, with both photons undergoing a number of successive impacts at beam-splitters. We say that a photon *chooses* the output port +[−] in the beam-splitter BS_{ik} ($i \in \{1,2\}, k \in \{1,2, ...n\}$) or leaves this beam-splitter through the channel +[−], if, with the two detectors $D_i(+1)$, $D_i(-1)$ in place, the photon would be detected at $D_i(+1)$ [$D_i(-1)$]. This does not mean that the particle leaves the splitter through one port,

---

22 A. Aspect, P. Grangier and G. Roger, *Physical Review Letters*, **49**, 91-94, (1982).

23 B. Yurke and D. Stoler, *Physical Review Letters* **68**, 1251 (1992).

and through the other port nothing is going. The *choice* consists in that the observable properties of the particle (the properties that lead the detector to click) go through one determined output port, and accordingly a click would be observed if there would be a detector monitoring this port. Unobservable information related to the particle travels through the other port, and no click will be observed, even if there is a detector monitoring this port. This is essentially the same view as that of Bohm's causal model[24].

Consider the experiment of FIG. 1, and suppose a photon impinging on BS_{21} makes the *choice* to produce a click on path segment s_{21+}. If the detectors are in place after BS_{21} the outcome would yield the value +1, and no click would be observed in the detector monitoring s_{21-}. If the detectors are removed after BS_{21}, then the observable part of the particle will reach BS_{22} by path segment s_{21+}, but an unobservable part of the particle will travel through s_{21-} and perceive the influence of the phase plate ϕ_{22}. According to this description we clearly distinguish between choice and detection, and refer to devices as polarizers or beam-splitters as *choice devices*[25].

4.2. Detection reveals the choice the particle made, and destroys the possibility for further choices

The experimental data support the view that for both single-particle and multi-particle superposition, the outcome distribution does not depend on the distances at which the detectors are placed with respect to the beam splitters (this is also the conventional quantum mechanical view). From this independence follows (in the light of the distinction between choice and detection) that the outcome is determined at the time the photon leaves the splitter. As far as the photon is not detected, it is, however, always possible for the physicist to let the photon pass to a further interferometer and to oblige it to change the outcome distribution. Detection reveals the *choice* the photon made at a *choice device*, and disables the photon to new choices. According to this view, the "reduction" which makes it impossible for a physicist to restore coherently the whole information the particle carries, occurs at the instant at which an observable or unobservable part of this information meets a detector. And the "jump" determining the value the measurement yields, occurs at the impact of a particle on the last *choice device* preceding the detectors. So, at the time of choosing, nothing irreversible happens, and at the time an irreversible detection happens, the photon makes no choice.

4.3. Indistinguishability

If the choice in BS_{ik} makes it impossible to know (through detection after this choice) to which input sub-ensemble of each $Bs_{ik'}$, $k' \leq k$, a particle belongs, the choice in BS_{ik} is referred to as choice leading to indistinguishability or uncertainty, and labeled u_{ik}. If the

---

24 D. Bohm and B.J. Hilley, *The Undivided Universe*, New York: Routledge, 1993, p. 290-295.

25 The expression 'a photon chooses' deserves some explanation. We assume the view that when a photon impacts at a beam-splitter, some free will (choice requires free will) makes a choice, for instance the choice to make a detector monitoring the transmitted path click, and to let undisturbed a detector monitoring the reflected path. This does not mean that we attribute free will to the 'photon', but rather that terms like 'photon' or 'particle' refer to contents or intentions of some mind (evidently not a human one). In a particular experiment, a detection reveals one intention this mind had, and destroys the possibility to know the other intentions of this mind regarding the alternative (undone) experiments. We use expressions like 'the photon makes a choice' to avoid complicated formulations. Notice that this view can be considered realistic in the sense, *that the regularities in the observed phenomena, and the intersubjective agreement, require a general explanation not exclusively based on the structure of the human mind*. (B. d'Espagnat, Nonseparability and the tentative descriptions of reality, in: W. Schommers (ed.), *Quantum theory and pictures of reality*, New York: Springer, 1989, p. 159).

choice in BS_{ik} allows one to distinguish the input sub-ensembles of each $Bs_{ik'}$, $k' \leq k$, it is labeled d_{ik}.

Let us emphasize that there may be situations in which it is in principle possible [impossible] to know to which input sub-ensemble of choice device BS_{21} the particle belongs by detecting it when it leaves this choice device. It is, however, impossible [possible] to acquire this knowledge (namely to which input sub-ensemble of choice device BS_{21} the particle belongs) by detecting it after it leaves a subsequent choice device BS_{22}. The discussion of such cases is omitted in this article.

4.4. Before and non-before choices

Every time one of the two particles impinges on one of the choice devices, we consider whether in the inertial frame of this choice device, the other photon has made a choice or not. We denote by $\left(T_{jh} - T_{ik}\right)_{ik}$ the difference between the time at which particle j makes its h choice, and the time at which particle i makes its k choice, with both times T_{jh} and T_{ik} measured in the reference frame of BS_{ik}.

Suppose a two-particle experiment in which the particles are detected after impacts at choice devices BS_{1l} and BS_{2m}. We consider the choices in devices BS_{ik} and BS_{jh} ($i,j \in \{1,2\}$, $i \neq j$, $l \geq k$ or $m \geq k$, respectively; $m \geq h$ or $l \geq h$, respectively), and introduce the following definitions:

> ***Definition 1****:* The choice of particle i in choice device BS_{ik} is a *before* event b_{ik}, if:
> 1) it is a u_{ik} choice (this requires that there is at least one BS_{jh} in which the choice of particle j is a u_{jh} one), and
> 2) for all u_{jh} it holds that $\left(T_{jl} - T_{ik}\right)_{ik} > 0$ (i.e. at the arrival time of particle i in BS_{ik}, in the inertial frame of BS_{ik}, the *other* particle did not yet made any choice generating uncertainty).
>
> ***Definition 2****:* The choice of particle i in choice device BS_{ik} is a *non-before* event $a_{ik[jh]}$, if
> 1) it is a u_{ik} choice,
> 2) the choice in BS_{jh} is a u_{jh} one,
> 3) $\left(T_{jh} - T_{ik}\right)_{ik} \leq 0$ (e.g. in the inertial frame of BS_{ik}, at the arrival time of particle i in this choice device, the *other* particle has met choice device BS_{jh}),
> 4) $\left(T_{jh+1} - T_{ik}\right)_{ik} > 0$.

Notations like $(a_{11[21]}, b_{12}, a_{13[22]}, a_{21[12]}, b_{22})_{+-}$ indicate the type of each particle's choices, and the values the detections yield. $(a_{11[21]}, b_{12}, a_{13[22]})_{-}$ refers to the times particle 1 is detected at $D_{13}(-1)$, independently on where particle 2 is detected. If it does not matter in a theorem whether the choice is *before* or *non-before*, we refer to it as x_{ij}.

Expressions like $P\left((a_{11[21]}, b_{12}, a_{13[22]}, a_{21[12]}, b_{22})_{+-}\right)$, $P\left((a_{11[21]}, b_{12}, a_{13[22]})_{-}\right)$, denote the probabilities to obtain the indicated detection values (e.g. photon 1 detected in

$D_{13}(+1)$ and photon 2 detected in $D_{22}(-1)$, or photon 1 detected in $D_{13}(-1)$, respectively) in an experiment in which the particles make the indicated choices. For non-before events, the subscripts in brackets may be omitted if no ambiguity results.

The notations p_{+-}, p_{-+} refer to the outcomes the two sub-ensembles initially prepared by the physicist would produce if the particles were detected before entering the first choice device. The corresponding probabilities are given by $P(p_{+-})$, $P(p_{-+})$. Notice that the two prepared sub-ensembles are such that detection (before choice) of only one of the particles suffices to establish to which of the sub-ensembles it belongs. Accordingly, the view that beam-splitter BS_{21} in FIG.1 is part of an extended, complicated source, cannot be shared.

The preceding definitions and considerations apply straightforwardly to experiments in which photon 1 and photon 2 originate from independent sources[26]. It is important to highlight, that the property of an impact at a choice device to be *before* or *non-before* is an absolute one, i.e. it does not depend of the observer's state of motion.

4.5. Two-particle experiments with time-like separated choice events imply that the particle choosing later takes into account the choices the other particle made before

Bell experiments with time-like separated choices have already been done[27], demonstrating the same correlations as when the choices are space-like separated. Consider the experiment proposed in FIG. 2, in which the choice photon 2 makes in BS_2 lies *time-like separated after* the choice photon 1 makes in BS_1. It is clear that at the time photon 1 makes its choice, it cannot account for choices in BS_2 because such choices do not exist at all, from any observer's point of view. In this case expressions like 'the later choice', and 'the former choice' make the same sense in every inertial frame. Therefore, the correlations appear because the choice photon 2 makes, depends somewhat on the choice photon 1 has made. Moreover, as it is always possible to obtain path information about photon 1 by detecting photon 2 before it reaches beam-splitter BS_2, photon 1 has to choose as it would choose in the absence of nonlocal influences.

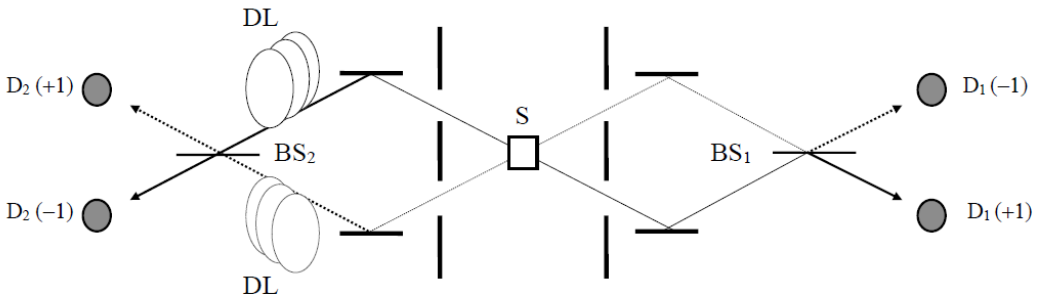

Figure 2:
Schematic diagram of a twofold double-slit experiment with the impact at beam-splitter BS_2 time-like separated from the impact on BS_1 that occurred previously.

---

[26] The conditions for entangling photons from independent sources are discussed by M. Zukowski, A. Zeilinger, and H. Weinfurter, in: D.M. Greenberger and A. Zeilinger (Eds.), *Fundamental Problems in Quantum Theory*, New York: New York Academy of Sciences, 1995, p. 91-102.

[27] P. R. Tapster, J.G. Rarity and P.C.M. Owens, *Physical Review Letters*, **73**, 1923-1926 (1994).

4.6. In two-particle experiments with space-like separated choices, the correlations arise because non-before choices take into account the outcomes of before choices

The preceding explanation for time-like separated events is now somewhat extended to space-like separated choices. In this case as well, every time a photon impacts on a choice device it takes account of whether the other photon has already impacted or not. In which inertial frame does the photon look at what happens in the other choice device? We assume that the inertial frame which matters for one photon's choice is that of the choice device at which this photon impacts. Suppose in FIG. 2 that the choice of photon 2 is a *before* event and the choice of photon 1 is a *non-before* event. We now assume that the correlations appear because photon 2 cannot take account of the choice photon 1 makes in BS_1, and photon 1 takes account of the choice photon 2 makes in BS_2. However, the choice photon 1 makes in a *non-before* impact does not take into account the choice photon 1 itself would have made if the impact would have been a *before* one.

4.7. Experiments with two before choices: the correlation functions depend nonlocally on the state of movement and the position of the choice devices

We now discuss what happens, if the outcomes on both sides result from before choices. In this case we assume that each of the photons cannot take account of the choice of the other photon. Consequently the choices produce the same outcome distribution which would originate if only local information were to matter. For a two particle experiment, with each particle making only one choice, one is led to the following relations:

$$\begin{aligned} P\left((d_{11},d_{21})_{++}\right) &= P\left((b_{11},b_{21})_{++}\right) \\ P\left((d_{11},d_{21})_{--}\right) &= P\left((b_{11},b_{21})_{--}\right) \\ P\left((d_{11},d_{21})_{+-}\right) &= P\left((b_{11},b_{21})_{+-}\right) \\ P\left((d_{11},d_{21})_{-+}\right) &= P\left((b_{11},b_{21})_{-+}\right) \end{aligned} \tag{3}$$

where the expressions $P\left((d_{11},d_{21})_{\pm\pm}\right)$, denote the probabilities of obtaining each of the four different possible outcomes (+1, +1), (+1, −1), (−1, +1), (−1, −1), when both choices occur under distinguishability conditions. The expression $P\left((b_{11},b_{21})_{\pm\pm}\right)$ denotes the probabilities of obtaining each of the four different possible outcomes, when both choices are *before* events. Similarly, for experiments with subsequent choices, relations such as the following hold:

$$\begin{aligned} P\left((d_{11},d_{21},b_{22})_{++}\right) &= P\left((b_{11},b_{21},b_{22})_{++}\right) \\ P\left((d_{11},b_{12},d_{21},b_{22})_{+-}\right) &= P\left((b_{11},b_{12},b_{21},b_{22})_{+-}\right) \end{aligned} \tag{4}$$

However:

$$\begin{aligned} P\left((d_{11},d_{21},d_{22})_{++}\right) &\neq P\left((b_{11},b_{21},b_{22})_{++}\right) \\ P\left((d_{11},d_{12},d_{21},d_{22})_{+-}\right) &\neq P\left((b_{11},b_{12},b_{21},b_{22})_{+-}\right) \end{aligned}$$

The argument in the preceding section 4.6, and the equations (3) and (4) lead to the following inequalities:

$$\begin{aligned} P\left((b_{11},a_{21})_{++}\right) &\neq P\left((b_{11},b_{21})_{++}\right) \\ P\left((a_{11},b_{21},b_{22})_{++}\right) &\neq P\left((b_{11},b_{21},b_{22})_{++}\right) \end{aligned} \tag{5}$$

and to similar inequalities for the other outcome values. Consequently, we are led to assume that the correlation functions depend nonlocally on the state of movement and the position of the choice devices.

4.8. Experiments with two non-before choices: Conditional probabilities

Now the question arises, what happens, if the outcomes on both sides result from non-before choices? We assume that photon 1 makes its choice in BS_{11} taking into account the choice photon 2 would have made in BS_{22} if the impact at this beam-splitter would have been a *before* event. This choice of photon 1, however, is independent of the choice photon 2 makes in the actual *non-before* impact. Similarly photon 2 makes its choice in BS_{22} depending on which choice photon 1 would have made in BS_{11} if the impact at this beam-splitter would have generated a *before* event, but independently of the choice photon 1 makes in the actual *non-before* impact. Once again the choice photon 1 [2] makes in a *non-before* impact does not take into account the choice photon 1 [2] itself would have made if the impact would have been a *before* one. Notice also that the choice photon 1 [2] makes does not depend on the choice photon 2[1] *really* makes, i.e. in this case there is no relation of dependence between *real* measured values. Accordingly, (contrary to quantum mechanics) no contradiction results.

The joint probabilities for two *non-before* impacts result from the key equation:

$$\begin{aligned} P\left((a_{11},a_{21})_{++}\right) = {} & P\left((b_{11},b_{21})_{++}\right)P\left((a_{11})_{+}\middle|(b_{21})_{+}\right)P\left((a_{21})_{+}\middle|(b_{11})_{+}\right) \\ & + P\left((b_{11},b_{21})_{--}\right)P\left((a_{11})_{+}\middle|(b_{21})_{-}\right)P\left((a_{21})_{+}\middle|(b_{21})_{-}\right) \\ & + P\left((b_{11},b_{21})_{+-}\right)P\left((a_{11})_{+}\middle|(b_{21})_{-}\right)P\left((a_{21})_{+}\middle|(b_{11})_{+}\right) \\ & + P\left((b_{11},b_{21})_{-+}\right)P\left((a_{11})_{+}\middle|(b_{21})_{+}\right)P\left((a_{21})_{+}\middle|(b_{11})_{-}\right) \end{aligned} \tag{6}$$

The term $P\left((a_{11})_{+}\middle|(b_{21})_{+}\right)$ gives the probability that a photon pair that would have produced the outcome $(b_{11},b_{21})_{\pm+}$ if both choices were *before* events, produces the outcome $(a_{11},b_{21})_{++}$ if photon 1 makes a *non-before* choice and photon 2 a *before* one. Similar $P\left((a_{21})_{+}\middle|(b_{11})_{+}\right)$ is the probability that a photon pair that would have produced the outcome $(b_{11},b_{21})_{+\pm}$ if both choices were *before* events, produces the outcome $(b_{11},a_{21})_{++}$ if photon 1 makes a *before* choice and photon 2 a *non-before* one; and so on.

Consequently, the theory must contain a theorem indicating how conditional probabilities like $P\left((a_{11})_{+}\middle|(b_{21})_{+}\right)$ derive from measurable statistical distributions like $P\left((a_{11},b_{21})_{++}\right)$. This theorem is given below as *Theorem 5.1.*

4.9. For the class of experiments with a given numbers of choices, a matrix of complex functions corresponds to each outcome

We associate to each of the prepared sub-ensembles and each of the four input-output possibilities at every choice device a complex function. This function depends on setup parameters (such as phases or orientations) that the experimentalist can choose arbitrarily. For the experiment represented in FIG. 1 these functions are labeled C_{+-} and C_{-+} for the prepared sub-ensembles: C_{+11+} , C_{+11-} , C_{-11+} , C_{-11-} for BS_{11}; C_{+21+} , C_{+21-} , C_{-21+} , C_{-21-} for BS_{21}; and C_{+22+} , C_{+22-} , C_{-22+} , C_{-22-} for BS_{22}; where for the moment we do not specify the variable domain of these functions. There is a correspondence between each possible outcome in an experiment with an arbitrary number of choices n, and a $2^{n-1} \times 2^{n-1}$ matrix the elements of which are products of $C_{\pm ij \pm}$.

The outcome (+1, +1) in any experiment with two choices (i.e. the outcomes $(d_{11}, d_{21})_{++}$, $(b_{11}, b_{21})_{++}$, $(a_{11}, b_{21})_{++}$, $(b_{11}, a_{21})_{++}$) corresponds to the 2×2 matrix:

$$\begin{pmatrix} C_{+-}C_{+11+}C_{-21+}\left(C_{+-}C_{+11+}C_{-21+}\right)^* & C_{+-}C_{+11+}C_{-21+}\left(C_{-+}C_{-11+}C_{+21+}\right)^* \\ C_{-+}C_{-11+}C_{+21+}\left(C_{+-}C_{+11+}C_{-21+}\right)^* & C_{-+}C_{-11+}C_{+21+}\left(C_{-+}C_{-11+}C_{+21+}\right)^* \end{pmatrix} \tag{7}$$

where each element in the matrix is a term in the expansion of the squared modulus:

$$\left|C_{+-}C_{+11+}C_{-21+} + C_{-+}C_{-11+}C_{+21+}\right|^2 \tag{8}$$

In the same way, in an experiment with two choices there is correspondence between the outcome (+1, −1) and the 2×2 matrix:

$$\begin{pmatrix} C_{+-}C_{+11+}C_{-21-}\left(C_{+-}C_{+11+}C_{-21-}\right)^* & C_{+-}C_{+11+}C_{-21-}\left(C_{-+}C_{-11+}C_{+21-}\right)^* \\ C_{-+}C_{-11+}C_{+21-}\left(C_{+-}C_{+11+}C_{-21-}\right)^* & C_{-+}C_{-11+}C_{+21-}\left(C_{-+}C_{-11+}C_{+21-}\right)^* \end{pmatrix} \tag{9}$$

where each element is a term of the sum resulting from expanding:

$$\left|C_{+-}C_{+11+}C_{-21-} + C_{-+}C_{-11+}C_{+21-}\right|^2 \tag{10}$$

Each outcome with values (+1, +1) in experiments in which photon 1 makes only one impact at a *choice device*, and photon 2 makes two such impacts, corresponds to the matrix which has as elements the single terms in the expansion of the squared modulus:

$$\left|C_{+-}C_{+11+}C_{-21+}C_{+22+} + C_{+-}C_{+11+}C_{-21-}C_{-22+} + C_{-+}C_{-11+}C_{+21+}C_{+22+} + C_{-+}C_{-11+}C_{+21-}C_{-22+}\right|^2 \tag{11}$$

Similarly one can build $2^{n-1} \times 2^{n-1}$ matrices for each of the four possible outcomes in any 2-particle experiment with an arbitrary number of choices *n*. They are referred to as $\underset{\pm\pm}{C_{k,l}}$ *matrices,* where the upper subscript indicates the number of choices each photon undergoes, and the lower subscript the outcome. The matrix elements are labeled c_{ij}.

4.10. Many superposition principles or entanglement rules

We introduce now a technique to generate different sums of matrix elements, each sum yielding a real value. Since the method proceeds by summing over different diagonals of elements or submatrices, in $\underset{\pm\pm}{C_{k,l}}$ or in submatrices of it, we refer it to as *diagonal* calculus.

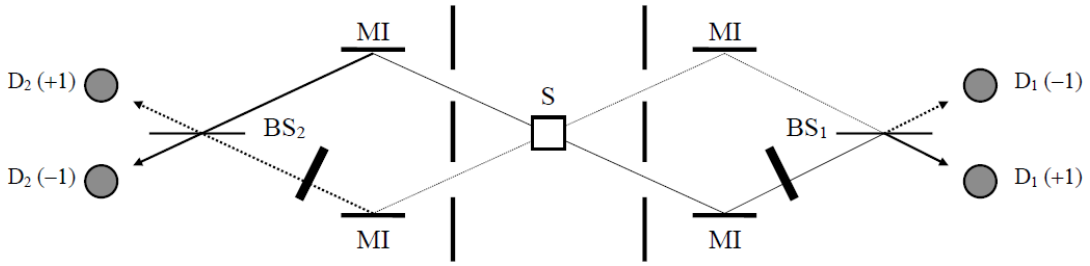

Figure 3:
Schematic diagram of a twofold double-slit experiment.

From the $\underset{++}{C_{1,1}}$ matrix, e.g. matrix (7), corresponding to the outcome (+1,+1) in a 2-choice experiment such as the standard Bell experiment in FIG. 3, arise the following two sums:

- the sum of all elements in the matrix:

$$\sum_{i=1,\,j=1}^{i=2,\,j=2} c_{ij} \tag{12}$$

which obviously is equal $\left|C_{+-}C_{+11+}C_{-21+} + C_{-+}C_{-11+}C_{+21+}\right|^2$, and

- the sum of all elements on the diagonal:

$$\sum_{i=1}^{i=2} c_{ii} \tag{13}$$

which represents the same value as $\left|C_{+-}C_{+11+}C_{-21+}\right|^2 + \left|C_{-+}C_{-11+}C_{+21+}\right|^2$.

We now associate (12) and (13) to different outcome probabilities in experiments, depending on the type of choice the photons make:

$$P\left((a_{11}, b_{21})_{++}\right) = P\left((b_{11}, a_{21})_{++}\right) = \sum_{i=1,\,j=1}^{i=2,\,j=2} c_{ij} \tag{14}$$

$$P\left((d_{11},d_{21})_{++}\right)=P\left((b_{11},b_{21})_{++}\right)=\sum_{i=1}^{i=2}c_{ii} \tag{15}$$

From the matrix $\underset{+-}{C_{1,1}}$, e.g. the matrix (8) corresponding to the outcome (+1, −1) in the experiment of FIG. 3, arise the probabilities:

$$P\left((a_{11},b_{21})_{+-}\right)=P\left((b_{11},a_{21})_{+-}\right)=\sum_{i=1,j=1}^{i=2,j=2}c_{ij}=\left|C_{+-}C_{+11+}C_{-21-}+C_{-+}C_{-11+}C_{+21-}\right|^2 \tag{16}$$

$$P\left((d_{11},d_{21})_{+-}\right)=P\left((b_{11},b_{21})_{+-}\right)=\sum_{i=1}^{i=2}c_{ii}=\left|C_{+-}C_{+11+}C_{-21-}\right|^2+\left|C_{-+}C_{-11+}C_{+21-}\right|^2 \tag{17}$$

With increasing numbers of choices, the *diagonal* calculus yields an increasing number of sums, and each of them can be associated to outcome probabilities. So for instance, from the matrix $\underset{++}{C_{1,2}}$ corresponding to the detection values (+1,+1) in the 3-choice experiment proposed in FIG. 1, the following probabilities arise:

$$\begin{aligned}&P\left((a_{11[22]},b_{21},b_{22})_{++}\right)=P\left((b_{11},a_{21[11]},a_{22[11]})_{++}\right)=\sum_{i=1,j=1}^{i=4,j=4}c_{ij}\\&=\left|C_{+-}C_{+11+}C_{-21+}C_{+22+}+C_{+-}C_{+11+}C_{-21-}C_{-22+}+C_{-+}C_{-11+}C_{+21+}C_{+22+}+C_{-+}C_{-11+}C_{+21-}C_{-22+}\right|^2\end{aligned} \tag{18}$$

$$\begin{aligned}&P\left((b_{11},b_{21},b_{22})_{++}\right)=P\left((d_{11},d_{21},u_{22})_{++}\right)=\sum_{i=1,j=1}^{i=2,j=2}c_{ij}+\sum_{i=3,j=3}^{i=4,j=4}c_{ij}\\&=\left|C_{+-}C_{+11+}C_{-21+}C_{+22+}+C_{+-}C_{+11+}C_{-21-}C_{-22+}\right|^2\\&+\left|C_{-+}C_{-11+}C_{+21+}C_{+22+}+C_{-+}C_{-11+}C_{+21-}C_{-22+}\right|^2\end{aligned} \tag{19}$$

$$\begin{aligned}&P\left((a_{11[21]},b_{21},b_{22})_{++}\right)=P\left((b_{11},a_{21[11]},b_{22})_{++}\right)\\&=\sum_{i=1,j=1}^{i=2,j=2}c_{ij}+\sum_{i=3,j=3}^{i=4,j=4}c_{ij}+c_{13}+c_{24}+c_{31}+c_{42}\\&=P\left((b_{11},b_{21},b_{22})_{++}\right)\\&+C_{+}C_{+11+}C_{+21+}C_{+22+}\left(C_{-}C_{-11+}C_{-21+}C_{+22+}\right)^{*}+\left(C_{+}C_{+11+}C_{+21+}C_{+22+}\right)^{*}C_{-}C_{-11+}C_{-21+}C_{+22+}\\&+C_{+}C_{+11+}C_{+21-}C_{-22+}\left(C_{-}C_{-11+}C_{-21-}C_{-22+}\right)^{*}+\left(C_{+}C_{+11+}C_{+21-}C_{-22+}\right)^{*}C_{-}C_{-11+}C_{-21-}C_{-22+}\end{aligned} \tag{20}$$

$$P\left((b_{11},b_{21},a_{22[11]})_{++}\right)=\sum_{i=1,j=1}^{i=2,j=2}c_{ij}+\sum_{i=3,j=3}^{i=4,j=4}c_{ij}+c_{14}+c_{23}+c_{32}+c_{41}=$$

$$=P\left((b_{11},b_{21},b_{22})_{++}\right)+$$

$$+C_{+}C_{+11+}C_{+21+}C_{+22+}\left(C_{-}C_{-11+}C_{-21-}C_{-22+}\right)^{*}+\left(C_{+}C_{+11+}C_{+21+}C_{+22+}\right)^{*}C_{-}C_{-11+}C_{-21-}C_{-22+}+ \quad (21)$$

$$+C_{+}C_{+11+}C_{+21-}C_{-22+}\left(C_{-}C_{-11+}C_{-21+}C_{+22+}\right)^{*}+\left(C_{+}C_{+11+}C_{+21-}C_{-22+}\right)^{*}C_{-}C_{-11+}C_{-21+}C_{+22+}$$

Similarly one obtains the formula corresponding to the other three outcomes (+1,−1), (−1, +1), (−1,−1).

As we will see later the different functions $C_{\pm ik\pm}$ can be identified to the quantum mechanical path amplitudes. Considering this fact, it appears that quantum mechanics allows only one rule to generate two-particle correlations, namely the sum of all elements in, for instance, the matrix $C_{\underset{++}{1,2}}$, as indicated in (18). On the contrary, the *superrelativistic* theory exploits other possibilities of summing over the elements of $C_{\underset{++}{1,2}}$ to define entanglement rules. This theory, therefore, is capable of accounting for phenomena which do not have any mathematical counterpart in quantum theory.

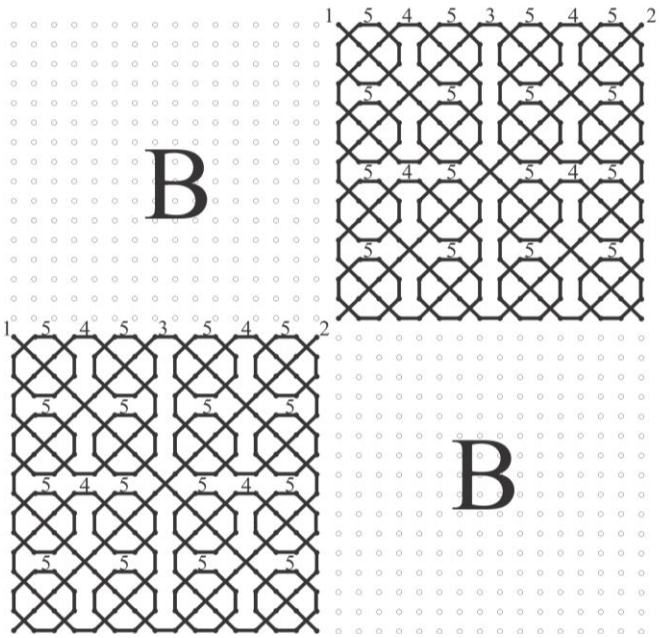

Figure 4:
Key to calculate the predictions of the *superrelativistic* theory through *diagonal* calculus for two-particles experiments with 6 impacts.

As an example, in FIG. 4 the key to apply the *diagonal* calculus in experiments with 6 choices is given. For experiments in which photon 1 undergoes only one impact, this impact is a *before* one, and photon 2 undergoes 5 choices, the rules to obtain the outcomes distribution are the following:

- All impacts of photon 2 are *before* events. Then one obtains the outcome distribution by adding the elements of the two submatrices labeled B.

- The first impact of photon 2 is a *non-before* event and all the other impacts are *before* events. Then one adds the elements of the two submatrices labeled B, and the elements on the two diagonals labeled 1.
- The second impact of photon 2, is a *non-before* event and all the other impacts at the beam-splitters are *before* events. Then one adds the elements of the two submatrices labeled B, and all the elements on the two diagonals labeled 2.
- The third impact of photon 2, is a *non-before* event and all the other impacts at the beam-splitters are *before* events. Then one adds all the elements of the two submatrices labeled B, and all the elements on the two octagons labeled 3.
- The four impact of photon 2, is a *non-before* event and all the other impacts at the beam-splitters are *before* events. Then one adds all the elements of the two submatrices labeled B, and all the elements on the 8 octagons labeled 4.
- The fifth impact of photon 2, is a *non-before* event and all the other impacts at the beam-splitters are *before* events. Then one adds all the elements of the two submatrices labeled B, and all the elements on the 32 octagons labeled 5.

The cases in which photon 2 undergoes several *non-before* impacts are derived from the preceding rules. So for instance:

- The third and the fifth impact of photon 2 are *non-before* events, and all the other three impacts are *before* events. Then one takes the elements of the two submatrices labeled B, plus all the elements on the two octagons labeled 3, and plus all the elements on the 32 octagons labeled 5.
- The five impacts of photon 2 are *non-before* events. Then one adds all the elements of the matrix.

We omit here the discussion of the rules for the experiments in which photon 1 undergoes two, three, or four choices, and photon 2 undergoes four, three, or two choices. As for the 3-choice experiments described by equation (18), the case in which photon 1 undergoes five choices and photon 2 undergoes one *non-before* choice is equivalent to the case in which photon 1 makes only one impact and all the five impacts of photon 2 are *non-before* events, i.e. one adds all the elements of the matrix.

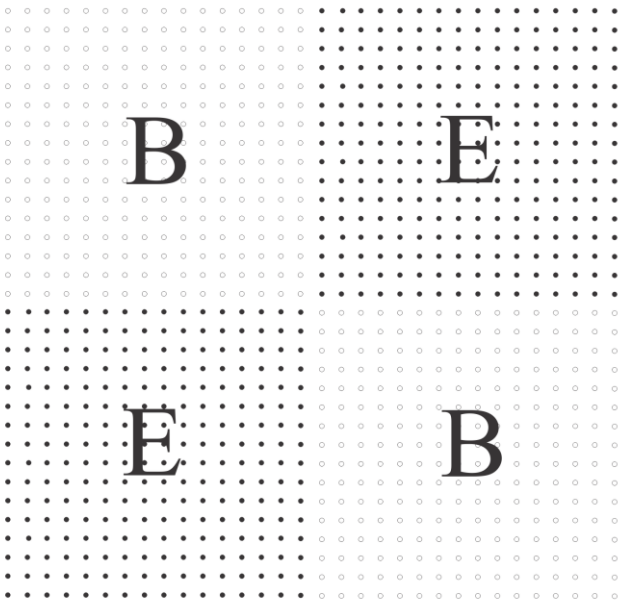

Figure 5:
Key to calculate the quantum mechanical predictions for two-particles experiments with 6 impacts.

FIG. 5 corresponds to the conventional quantum mechanical description, with only one entanglement rule: If it is possible to distinguish the input sub-ensemble, then the output distribution results from adding the elements of the two submatrices labeled B; if it is not possible to distinguish the input sub-ensemble, then one adds all the elements of the matrix: the submatrices labeled E account for the quantum mechanical entanglement.

4.11. Impossibility of communication without signaling

For 2-choice experiments, the following equalities are assumed:

$$\begin{aligned} P\left((a_{11})_+\right) &= P\left((b_{11})_+\right) \qquad P\left((a_{11})_-\right) = P\left((b_{11})_-\right) \\ P\left((a_{211})_+\right) &= P\left((b_{21})_+\right) \qquad P\left((a_{21})_-\right) = P\left((b_{21})_-\right) \end{aligned} \tag{22}$$

For 3-choice experiments, the following equalities are assumed:

$$\begin{aligned} P\left((a_{11[22]})_+\right) &= P\left((a_{11[21]})_+\right) = P\left((b_{11})_+\right) \\ P\left((a_{11[22]})_-\right) &= P\left((a_{11[21]})_-\right) = P\left((b_{11})_-\right) \\ P\left((a_{21},a_{22})_+\right) &= P\left((b_{21},a_{22})_+\right) = P\left((a_{21},b_{22})_+\right) = P\left((b_{21},b_{22})_+\right) \\ P\left((a_{21},a_{22})_-\right) &= P\left((b_{21},a_{22})_-\right) = P\left((a_{21},b_{22})_-\right) = P\left((b_{21},b_{22})_-\right) \end{aligned} \tag{23}$$

The physical meaning of (22) and (23) is the following: an experimentalist at place A cannot produce observable order (a message) at place B, if between A and B there is no observable connection. Accordingly, communication between (time-like or space-like) separated human observers requires energy propagating *in space-time* from one observer to the other. Indirectly the principle leads also to the impossibility of using nonlocal phenomena for superluminal signaling. Notice, however, that the principle works also in situations with time-like separated choices, like FIG. 2. In such a case, interference fringes at the level of the single detection (first-order correlations) would not imply any superluminal signaling. However, since there is no observable connection or signaling between any of the detectors D_1 and BS_2, first order interference fringes would imply subluminal signalless communication (i.e. the possibility of using energyless or unobservable connections for generating observable order), and this would mean the failure of the second law of thermodynamics.

The impossibility of superluminal signaling resulting from observations like those of Michelson-Morley implies the dependence of simultaneity on the inertial frame, and therefore the impossibility of an absolute time. This impossibility relates evidently to our definition of *before* and *non-before* choices. The motivation of equations (22) and (23) is primarily not the concern of limiting the speed of signaling, but rather to found the *second law* of thermodynamics within the theory.

4.12. Rules for conditional probabilities

For 2-choice experiments it holds that:

$$P\left((a_{11})_+\middle|(b_{21})_+\right) = P\left((a_{11})_-\middle|(b_{21})_-\right)$$
$$P\left((a_{11})_+\middle|(b_{21})_-\right) = P\left((a_{11})_-\middle|(b_{21})_+\right) \tag{24}$$

$$P\left((a_{21})_+\middle|(b_{11})_+\right) = P\left((a_{21})_-\middle|(b_{11})_-\right)$$
$$P\left((a_{21})_+\middle|(b_{11})_-\right) = P\left((a_{21})_-\middle|(b_{11})_+\right) \tag{25}$$

where the expression $P\left((a_{11})_+\middle|(b_{21})_+\right)$ means the probability that a particle pair which would have produced the outcome $(b_{11},b_{21})_{\pm+}$ in the case of two *before* choices produces the outcome $(a_{11},b_{21})_{++}$ in the case of a *non-before* choice in BS_{11} and a *before* choice in BS_{21}; and so on.

For 3-choice experiments:

$$P\left((a_{11[2i]})_+\middle|(b_{2i})_+\right) = P\left((a_{11[2i]})_-\middle|(b_{2i})_-\right)$$
$$P\left((a_{11[2i]})_+\middle|(b_{2i})_-\right) = P\left((a_{11[2i]})_-\middle|(b_{2i})_+\right) \tag{26}$$

$$P\left((a_{21},a_{22})_+\middle|(b_{11})_+\right) = P\left((a_{21},a_{22})_-\middle|(b_{11})_-\right)$$
$$P\left((a_{21},a_{22})_+\middle|(b_{11})_-\right) = P\left((a_{21},a_{22})_-\middle|(b_{11})_+\right) \tag{27}$$

$$P\left((b_{21},a_{22})_+\middle|(b_{11})_+\right) = P\left((b_{21},a_{22})_-\middle|(b_{11})_-\right)$$
$$P\left((b_{21},a_{22})_+\middle|(b_{11})_-\right) = P\left((b_{21},a_{22})_-\middle|(b_{11})_+\right) \tag{28}$$

Moreover:

$$P\left((b_{21},b_{22})_+\middle|(b_{21})_+\right) = P\left((b_{21},b_{22})_+\middle|(b_{21})_-\right) = P\left((b_{22})_+\right)$$
$$P\left((b_{21},b_{22})_-\middle|(b_{21})_+\right) = P\left((b_{21},b_{22})_-\middle|(b_{21})_-\right) = P\left((b_{22})_-\right) \tag{29}$$

where the expressions for conditional probabilities in (29) mean the probabilities that a particle which would have produced the indicated outcome if it had been detected after a *before* impact at BS_{21} produces the indicated detection value after a *before* impact at BS_{22}.

4.13. Summary

The proposed explanation of nonlocal phenomena is based on the clear distinction between *source*, *choice* and *detection*, and on the concepts of *before* and *non-before* choices. The main principles of the theory are three:

a) assumption of values before measurement and conditional probabilities.
b) assumption of many entanglement rules;
c) impossibility of signalless communication.

The proposed description is clearly causal and superluminally nonlocal, but it does not lead to the assumption of absolute time. In this sense it can be considered a *superrelativistic causal model*.

Although it is absolutely contrary to the spirit of orthodox quantum mechanics, the real existence of values before measurement and the conditional probabilities are basically the ingredients which make it possible to circumvent the causal paradoxes. So far only two-particle experiments and two-output-ports choice devices have been considered. The general description for *n-particle* systems and choice devices with a number of output ports greater than 2, will be presented in a future article.

5. Theorems

The preceding principles impose a number of conditions to the functions $C_{\pm ik\pm}$. It can be shown that the quantum mechanical path amplitudes in the equations (1) fulfill them and can be identified with the different functions $C_{\pm ik\pm}$ for experiments like the represented in FIG. 1. Moreover one is led to the following two main theorems (for proofs see ref. 19).

Theorem 5.1. For 2-choice experiments one can derive:

$$\begin{aligned} P\left((a_{11},b_{21})_+\right) &= P\left((a_{11})_+\middle|(b_{21})_+\right) = P\left((a_{11})_-\middle|(b_{21})_-\right) \\ P\left((a_{11},b_{21})_-\right) &= P\left((a_{11})_+\middle|(b_{21})_-\right) = P\left((a_{11})_-\middle|(b_{21})_+\right) \end{aligned} \tag{30}$$

$$\begin{aligned} P\left((b_{11},a_{21})_+\right) &= P\left((a_{21})_+\middle|(b_{11})_+\right) = P\left((a_{21})_-\middle|(b_{11})_-\right) \\ P\left((b_{11},a_{21})_-\right) &= P\left((a_{21})_+\middle|(b_{11})_-\right) = P\left((a_{21})_-\middle|(b_{11})_+\right) \end{aligned} \tag{31}$$

where:

$$P\left((a_{11},b_{21})_+\right) = P\left((a_{11},b_{21})_{++}\right) + P\left((a_{11},b_{21})_{--}\right)$$
$$P\left((a_{11},b_{21})_-\right) = P\left((a_{11},b_{21})_{+-}\right) + P\left((a_{11},b_{21})_{-+}\right)$$

and so on.

This Theorem establishes the rules to derive conditional probabilities from measurable quantities like $P\left((a_{11},b_{21})_{++}\right), P\left((a_{11},b_{21})_{--}\right)$, etc. Similar relations hold for 3-choice experiments, for instance:

$$\begin{aligned} P\left((a_{11[22]},b_{21},b_{22})_+\right) &= P\left((a_{11[22]})_+\middle|(b_{22})_+\right) = P\left((a_{11[22]})_-\middle|(b_{22})_-\right) \\ P\left((a_{11[22]},b_{21},b_{22})_-\right) &= P\left((a_{11[22]})_+\middle|(b_{22})_-\right) = P\left((a_{11[22]})_-\middle|(b_{22})_+\right) \end{aligned} \tag{32}$$

Theorem 5.2.

$$E(a_{11[21]},b_{21},a_{22}) = E(b_{11},b_{21})E(a_{11},b_{21})E(b_{11},b_{21},a_{22}) \tag{33}$$

where:

$$E(a_{11[21]},b_{21},a_{22}) = P\left((a_{11[21]},b_{21},a_{22})_{++}\right) + P\left((a_{11[21]},b_{21},a_{22})_{--}\right) - P\left((a_{11[21]},b_{21},a_{22})_{+-}\right) - P\left((a_{11[21]},b_{21},a_{22})_{-+}\right)$$

$$E(b_{11},b_{21}) = P\left((b_{11},b_{21})_{++}\right) + P\left((b_{11},b_{21})_{--}\right) - P\left((b_{11},b_{21})_{+-}\right) - P\left((b_{11},b_{21})_{-+}\right)$$

etc.

6. Real experiments in preparation

Evidently the proposed *superrelativistic*, nonlocal theory can in principle be tested versus conventional quantum mechanics by means of experiments with fast-moving choice devices. Such experiments are impractical within the realm of today's available laboratory techniques. Nevertheless, *Theorem 5.2.* opens a road for tests with choice devices at rest in the laboratory frame, if one of the particles undergoes two successive impacts.

Consider the double slit experiment of FIG.1. The functions $C_{+-}, C_{-+}, C_{\pm ik\pm}$ for this experiment are:

$$\begin{array}{llll} C_{+-} = \frac{1}{\sqrt{2}} & C_{-+} = \frac{1}{\sqrt{2}} & & \\ C_{+11+} = \frac{i}{\sqrt{2}} & C_{+11-} = \frac{1}{\sqrt{2}} & C_{-11+} = e^{i\phi_{11}}\frac{1}{\sqrt{2}} & C_{-11-} = e^{i\phi_{11}}\frac{i}{\sqrt{2}} \\ C_{+21+} = \frac{i}{\sqrt{2}} & C_{+21-} = \frac{1}{\sqrt{2}} & C_{-21+} = e^{i\phi_{21}}\frac{1}{\sqrt{2}} & C_{-21-} = e^{i\phi_{21}}\frac{i}{\sqrt{2}} \\ C_{+22+} = \frac{i}{\sqrt{2}} & C_{+22-} = \frac{1}{\sqrt{2}} & C_{-22+} = e^{i\phi_{22}}\frac{1}{\sqrt{2}} & C_{-22-} = e^{i\phi_{22}}\frac{i}{\sqrt{2}} \end{array} \qquad (34)$$

The coefficient E^{QM} predicted by quantum mechanics is given by equation (2). The Appendix gives the probabilities needed to calculate (33):

$$E(b_{11},b_{21}) = 0$$

$$E(a_{11},b_{21}) = \cos(\phi_{11}-\phi_{21})$$

$$E(b_{11},b_{21},a_{22}) = \frac{1}{2}\left(\cos(\phi_{11}-\phi_{21}-\phi_{22}) - \cos(\phi_{11}-\phi_{21}+\phi_{22})\right)$$

Therefore:

$$E(a_{11[21]},b_{21},a_{22}) = E(b_{11},b_{21})E(a_{11},b_{21})E(b_{11},b_{21},a_{22}) = 0 \qquad (35)$$

That means that the superrelativistic, causal theory conflicts with quantum mechanics for all values of $\phi_{11},\phi_{21},\phi_{22}$, which result in $E^{QM} \neq 0$, if the arrival of photon 1 in BS_{11} occurs (in

the laboratory frame) after the arrival of photon 2 in BS_{21}, and before the arrival of photon 2 in BS_{22}.

In particular, for $\phi_{11} = 45°$, $\phi_{21} = -45°$ and $\phi_{22} = 90°$:

$$E^{QM} = 1, \qquad \text{and} \qquad E(a_{11[21]}, b_{21}, a_{22}) = 0 .$$

A real experiment in preparation is aiming to test these predicted values through measuring the experimental quantity:

$$E = \frac{R_{++} + R_{--} - R_{+-} - R_{-+}}{R_{++} + R_{--} + R_{+-} + R_{-+}} \tag{36}$$

where $R_{\pm\pm}$ are the four measured coincidence counts in the detectors $D_1(\pm 1)$, $D_2(\pm 1)$ under the given time ordering of the impacts on the splitters and parameters of the phase shifters.

The realization of experiments allowing us to investigate different time sequences is interesting independently of the possibility they would offer of testing the proposed superrelativistic theory. The nonlocality of quantum mechanics is so important and counterintuitive that as many experiments as possible should be performed to rule out definitively any other alternative explanations. By ruling out the superrelativistic theory, the proposed experiments will offer us a useful demonstration of the general validity of the conventional superposition principle, and of the strangeness of quantum mechanics[28]. If the results prove to be, unexpectedly, in conflict with quantum mechanics, the road to a new description of physical reality would be opened.

7. Philosophical implications

No one has characterized the scientific attitude regarding correlated events more eloquently than John Bell: *The scientific attitude is that correlations cry out for explanations*[29]. Correlated events reveal causal links. Physics consists essentially in discovering new correlations and the connections behind them.

The conviction that physical causality necessarily relies on influences propagating in space and time is undoubtedly deeply rooted. It played a key role in the worldview of classical physics and of philosophers like Laplace, Kant and Comte. In particular Kant postulated that in the chain of causes responsible for a phenomenon each single cause is itself observable [30]. Kantian philosophy assumes determinism as the correct sharpening of the causality considerations [31]. According to this understanding of causality, each physical effect can be explained exclusively by causes working within space-time, or observable elements of reality propagating in space-time[32]. Consequently, the reality behind phenomena reduces to

---

28 A. Zeilinger, private communications, 20.2. and 25.6.1996.

29 J.S. Bell, *Speakable and unspeakable in quantum mechanics, Cambridge University Press, 1987, p. 152; see also J.S. Bell, Indeterminism and nonlocality,* this volume, chapter VII.

30 I. Kant, *Kritik der reinen Vernunft,* Hamburg: Feliy Meiner, 1956, p. 466*, 580, 589, 600 (B 483, B 637, B 649, B 664).

31 P. Bernays, Causality, Determinism and Probability, in: W. Yourgrau and A. van der Merwe, *Perspectives in Quantum Theory*, Cambridge: The MIT Press, 1971, p. 261.

32 I. Kant, *Träume eines Geistersehers,* erläutert durch Träume der Metaphysik, Textedition of the Königliche Akademie der Wissenschaften, Band II, Berlin:Walter de Greyter, 1968, p. 370-372; and *Vor*

information completely accessible by empirical means, and scientific knowledge should in principle enable man to a complete control of the natural causes. All that nature does, should be possible to man..

After Einstein's theory of special relativity, this deterministic belief was consolidated into the successful postulate of relativistic causality. Nevertheless, relativistic causality is not a necessary consequence of observations. In effect, the experiments that inspired Einstein's special relativity only demonstrated that energy cannot propagate from one place to another faster than light. Accordingly, any signaling or message transmission requiring energy propagation cannot occur at speeds greater than *c*. Therefore faster-than-light signaling between human beings is impossible. On the contrary, experiments like that of Michelson-Morley do not forbid in any way faster-than-light influences occurring without observable connections. The recent experimental evidence speaks in favor of superluminal nonlocality, but occurring through unobservable transfer of information. Accordingly, no contradiction arises. But evidently, one has to give up special relativity, i.e. the postulate that *c* is the upper limit for all causal influences, and one has to accept that man cannot use nature's superluminal influences for practical purposes.

Bohm's causal model and the theory proposed in this article demonstrate that a causal view accounting for superluminal nonlocality is definitely possible. In such a view the effects do not exist before the causes (*causality principle*). Bohm's causal model reproduces the quantum mechanical predictions, but it comes back to the concept of absolute space-time, remaining essentially nonrelativistic. In our opinion this model conflicts with the Michelson-Morley results and similar observations. The theory proposed in this article includes the essential elements of relativity, e.g. the impossibility of a unique universal time ordering, and includes quantum mechanics as a partial description for single-particle and certain classes of multi-particle phenomena. However, it differs from relativity because of the assumption of superluminal influences, and from quantum mechanics because of the assumption of more than one entanglement rule, and the correlated exclusion of a frame-independent "wavefunction collapse".

Moreover, the superrelativistic causal theory has the striking feature of sharing two properties conventionally considered as contradictory. On the one hand the theory is as *deterministic* as Einstein had desired, for all possible outcome values are already determined when the photon leaves the source, with the exception of the choice between the *before* value and the *non-before* one, which is determined at the moment the particle impinges at the choice device. On the other hand, the superrelativistic nonlocal theory goes, in a sense, even beyond Bohr's *complementarity* in the "decisive" role attributed to "measurements"[33]. In EPR experiments with entangled two-particle systems in which "measurements" of the same physical quantity (i.e. corresponding to "the same operator") are performed on both particles, knowledge of the outcome value for one of the particles does not in general allow us to predict the outcome value for the other particle. This prediction is only possible if one of the "measurements" involves a *before* choice, and the other "measurement" a *non-before* one.

In summary, nonlocality research, and in particular the *superrelativistic* causal theory, stresses that behind the phenomena there are unobservable causes. Superluminal nonlocality

---

dem ersten Grunde des Unterschiedes der Gegenden im Raume, Textedition of the Königliche Akademie der Wissenschaften, Band II, Berlin:Walter de Greyter, 1968, p. 383.

[33] N. Bohr, *Physical Review*, **48**, 696-702 (1935).

rules out the belief that physical causality necessarily relies on influences propagating in space and time, and in particular Kant's causality postulate. On the one hand, through the Michelson-Morley and related experiments we have been led to the conclusion that man cannot communicate faster than light. On the other hand, the recent experimental evidence about nonlocal effects seem to tell us that there are faster-than-light influences in nature. As long as we accept these two classes of experimental results, we have to accept *that nature always will do things that man is not able to do.*

APPENDIX: Matrices $C_{\underset{\pm\pm}{1,1}}, C_{\underset{\pm\pm}{1,2}}$, and related probabilities

The quantities (34) yield the $C_{\underset{\pm\pm}{1,1}}$ matrices:

$$C_{\underset{++}{1,1}} = C_{\underset{--}{1,1}} = \begin{pmatrix} \frac{1}{8} & \frac{1}{8}e^{-i(\phi_{11}-\phi_{21})} \\ \frac{1}{8}e^{i(\phi_{11}-\phi_{21})} & \frac{1}{8} \end{pmatrix} \qquad C_{\underset{+-}{1,1}} = C_{\underset{-+}{1,1}} = \begin{pmatrix} \frac{1}{8} & \frac{-1}{8}e^{-i(\phi_{11}-\phi_{21})} \\ \frac{-1}{8}e^{i(\phi_{11}-\phi_{21})} & \frac{1}{8} \end{pmatrix}$$

and the probabilities:

$$P\big((b_{11},b_{21}\big)_{++} = P\big((b_{11},b_{21}\big)_{--} = P\big((b_{11},b_{21}\big)_{+-} = P\big((b_{11},b_{21}\big)_{-+} = \frac{1}{4}$$

$$P\big((a_{11},b_{21}\big)_{++} = P\big((a_{11},b_{21}\big)_{--} = \frac{1}{4} + \frac{1}{8}\big(e^{i(\phi_{11}-\phi_{21})} + e^{-i(\phi_{11}-\phi_{21})}\big) = \frac{1}{4}\big(1+\cos(\phi_{11}-\phi_{21})\big)$$

$$P\big((a_{11},b_{21}\big)_{+-} = P\big((a_{11},b_{21}\big)_{-+} = \frac{1}{4} - \frac{1}{8}\big(e^{i(\phi_{11}-\phi_{21})} + e^{-i(\phi_{11}-\phi_{21})}\big) = \frac{1}{4}\big(1-\cos(\phi_{11}-\phi_{21})\big)$$

Similarly, one obtains from (34) the four $C_{\underset{\pm\pm}{1,2}}$ matrices. As an example we give $C_{\underset{++}{1,2}}$:

$$C_{\underset{++}{1,2}} = \begin{pmatrix} \frac{1}{16} & \frac{1}{16}e^{-i\phi_{22}} & \frac{1}{16}e^{-i(\phi_{11}-\phi_{21})} & \frac{-1}{16}e^{-i(\phi_{11}-\phi_{21}+\phi_{22})} \\ \frac{1}{16}e^{i\phi_{22}} & \frac{1}{16} & \frac{1}{16}e^{-i(\phi_{11}-\phi_{21}-\phi_{22})} & \frac{-1}{16}e^{-i(\phi_{11}-\phi_{21})} \\ \frac{1}{16}e^{i(\phi_{11}-\phi_{21})} & \frac{1}{16}e^{i(\phi_{11}-\phi_{21}-\phi_{22})} & \frac{1}{16} & \frac{-1}{16}e^{-i\phi_{22}} \\ \frac{-1}{16}e^{i(\phi_{11}-\phi_{21}+\phi_{22})} & \frac{-1}{16}e^{i(\phi_{11}-\phi_{21})} & \frac{-1}{16}e^{i\phi_{22}} & \frac{1}{16} \end{pmatrix}$$

From these matrices result the probabilities:

$$P\big((b_{11},b_{21},a_{22})_{++}\big) = P\big((b_{11},b_{21},a_{22})_{--}\big) = \frac{1}{4} + \frac{1}{8}\cos(\phi_{11}-\phi_{21}-\phi_{22}) - \frac{1}{8}\cos(\phi_{11}-\phi_{21}+\phi_{22})$$

$$P\big((b_{11},b_{21},a_{22})_{+-}\big) = P\big((b_{11},b_{21},a_{22})_{-+}\big) = \frac{1}{4} - \frac{1}{8}\cos(\phi_{11}-\phi_{21}-\phi_{22}) + \frac{1}{8}\cos(\phi_{11}-\phi_{21}+\phi_{22})$$

Chapter XI

Quantum Theory: A Pointer to an Independent Reality
A Discussion of Bernard d'Espagnat's "Veiled Reality"

Juleon M. Schins
Dept. of Applied Physics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

There is no quantum world. There is only an abstract quantum-mechanical description[1].

Introduction

Quite soon after the first predictions of quantum mechanics the founding fathers became aware of two major promises contained in this new theory: first, it seemed to have an exceptional predictive power in many physical and chemical disciplines; second, it would launch, like no other physical theory before, an equally exceptional philosophical discussion concerning the status of reality. The fundamental issue in this philosophical debate is whether things *are*, on quantum level, the same way as they *are* on a classical level.

Two major currents of thought have crystallized during these first decades of quantum mechanics. In *conventional realism*, the quantum and classical ways of being are considered as essentially equivalent, while in *radical idealism* only classical events carry some kind of existence, quantum states being 'undefined'. Among the many possible intermediary positions a very interesting one is represented by Bernard d'Espagnat's *veiled reality* [2]. The present contribution aims at presenting the main ingredients of veiled reality, as well as discussing the implications of veiled reality for the question of the existence of God.

I. Criticizing Radical Idealism

D'Espagnat presents his own world view, veiled reality, after having severely criticized both idealism and realism. Against radical idealism (which he also calls *radical phenomenism*), he mentions four major arguments.

The *first* one considers the failure of Kantianism to answer obvious questions: *Kantianism and neo-Kantianism quite naturally lead to the view that, since the physical laws are, when all is said and done, regulated by the a priori modes of our sensibility and understanding, the laws in question must be constructed in terms of 'visualizable' concepts only. Some science historians, such as A. Miller [3], have stressed the fact that trusting Kantianism or neo-Kantianism too much on this point may well have hindered some of the physicists who, during the first quarter of this century, took up the task of building up a theory of the atomic processes* [4].The concept of visualizability is very attractive in physics mainly because it makes

---

[1] N. BOHR quoted by A. Petersen, in: *The Philosophy of Quantum Mechanics*, M. Jammer (Ed.) John Wiley (1974).
[2] B. D'ESPAGNAT, *Veiled Reality*, Addison-Wesley P.C., Reading, Massachusetts (1995).
[3] A.I. MILLER, in: *Sixty-Two Years of Uncertainty*, A.I. Miller (Ed.), Plenum (NATO series), NY (1990).
[4] Ref. 2, p 314.

one feel comfortable. It confirms the traditional view that human mind is the highest form of intelligence. Especially in the times of Kant, it was generally believed in the scientific community that nature was nearly completely understood, 'dominated' by human mind. These scientists represented a receptive public for Kant's doctrine of categories (these are the ten basic ways in which human mind organizes sense perceptions) [5], since they fully account for man's dominion of nature - albeit in a quite revolutionary way. However, twentieth century science has dramatically changed the perspective of the previous century.

The two theories of relativity were quite a strong blow inflicted on Kantianism. The only way in which man can have some natural affinity with the concepts of special relativity, is that man's motion be comparable to the speed of light. This is not at all the case, and seems quite improbable in the future as well. An identical argument holds for general relativity and acceleration. And what about quantum theory? It meant another blow, even stronger than the one inflicted by relativity.

In spite of the fact that the quantum-mechanical formalism (the Schrödinger equation) determines the evolution of a wavefunction, and says nothing at all about particles, there exists, to date, not a single school of interpretation of quantum mechanics that can do without the particle concept. Yet, experiments all show a clear tendency: the better one looks, the less 'particle' one sees, and the more 'wave'. In quantum mechanics total visualizability seems to be incompatible with the basic ingredients of that theory. In the field of high-energy physics the same tendency can be appreciated: new discoveries always introduce new and rather poorly visualizable concepts (what about the flavor of an elementary particle, for example).

D'Espagnat comments: *But then, that is, if Kant's accounting for the set of all the concepts we use is not valid, a question arises as to where these concepts 'come from'. Why do we use such and such concepts instead of others? Clearly, the (obvious) answer 'It is because they work' is not totally satisfactory, for it immediately calls for the question already encountered above: 'Why do they work, while others do not?', which Kant, with his scheme, did not even have to consider. To sweep this last question away just 'by decree' ('no question beginning with a 'why' shall be considered') is somewhat of a poor answer. A slightly better one would be to point out that these new concepts are borrowed from mathematics, which is also a form, if not of our sensibility, at least of our understanding. But then, why precisely just these - curved spaces, tensors, Heisenberg-picture projectors and so on - instead of so many other ones, that swarm in textbooks on mathematics? When all is said and done, the idea that, even though the form of our scientific descriptions admittedly owes much to the mind structures, still it does not owe everything to it, seems to force itself upon us* [6].

This citation immediately introduces d'Espagnat's *second* argument against radical idealism, which considers the experimental refutation of physical theories that are mathematically consistent: *We also noted the well-known but still quite impressive argument based on the remark that we sometimes build up quite beautifully rational physical theories that experiment falsifies. Experiment cannot falsify the rules of the game of chess - nor those of any other game - because these rules are just created by us. In this case, therefore, there is nothing 'external' that could say 'no'. But in physics it sometimes (and even quite often!) happens that something does say no. How could this 'something' still be 'us'? It seems that the degree of intellectual contorsion necessary for answering such a question in any positive way exceeds what is acceptable*[7]. This argument has some aspects in common with the first

[5] I. KANT, *Kritik der reinen Vernunft*, B 102-169.
[6] Ref 2, p 334-335.
[7] Ref. 2, p 335-336.

argument. Most importantly, both of them point to the existence of an independent, external world, that constitutes the measure for our knowledge. This 'world', or intellect, is likely to be essentially superior to human mind, as long as it may keep saying 'no'.

The *third* argument d'Espagnat presents against radical idealism discusses the logical priority of existence over knowledge. He argues that it is impossible to conceive knowledge without accepting that the known object must exist, while the reverse does not hold true. This argument is quite natural for a philosopher with a realistic world view, but it does not indicate a contradiction *within* the radical idealistic system. Rather, it constitutes the essence of radical idealism.

The *fourth* argument is concerned with intersubjective agreement. In the words of d'Espagnat: *If (...) as the idealist claims, the statement that the teapot 'really exists' has no meaning beyond that, for Alice, of describing the way she mentally organizes her sensations (and same, of course for Bob), then this must also be the case concerning the assertion that the teapot 'is really there' (or not) at such and such times, and the fact that Alice and Bob agree that they always had the same sensations in this respect becomes puzzling: a kind of constantly renewed miracle, in fact. The realist is therefore very much entitled to press the idealist on this point, and ask what explanation the idealist has to offer that would be as simple as the one just stated. Surprisingly enough, few, if any, idealist philosophers seem to have worried about this problem and the corresponding possible attack on their views. In the relevant literature, practically only intersubjective agreement concerning noncontingent facts (mathematics) is discussed* [8].

As emphasized by d'Espagnat himself [9], the above mentioned arguments are not new. Worse, in a strictly philosophical sense, the four arguments are known not to be conclusive either: idealism has withstood so much criticism that it seems an internally consistent doctrine. But a physicist demands from whatever theory or scientific doctrine, not only that it be internally consistent, but also that it explain as many phenomena as possible, and in the easiest fashion, by explaining the observed complexity as resulting from a relatively simple set of basic rules. As far as the problem of intersubjectivity is concerned, and more generally, epistemology, idealism strongly gives the impression to complicate things rather than to simplify them. However, the interesting aspect of these arguments is given by the fact that they are formulated not by a professional philosopher, but by a physicist who has contributed abundantly to the field of quantum mechanics.

II. Criticizing Conventional Realism

In his criticism of the 'conventional realistic' world view, d'Espagnat mentions equally four inconsistencies.

The *first* one considers the inability of realism to cope with what d'Espagnat denominates 'weak objectivity': *conventional realism assumes strong objectivity, which is at variance with the standard - by far the most efficient - formulation of quantum mechanics* [10]. The definition of weak and strong objectivity is not at all easy, and d'Espagnat dedicates lengthy pages to that discussion. Roughly spoken, strongly objective statements inform *us directly of attributes of the things under study* [11], and *a statement is but weakly objective if, while being true for everybody, still it basically refers to what human beings actually do, or can do, or observe. In*

[8] Ref. 2, p 16-17.
[9] Cfr. Ref. 2, p 314.
[10] Ref. 2, p 354.
[11] Ref. 2, p 22.

parallel, we may form the idea of calling 'weakly objective' the concepts defined by specifically referring to some human procedure [12]. Using these definitions, d'Espagnat's first argument against realism can be paraphrased as follows: according to the standard interpretation of quantum mechanics, any system lacks well-defined properties as long as it is not measured. It is the measuring process that pins down its properties, *and very specifically with respect to the conditions imposed by the measuring device.* For example: before measurement, a spin ½ particle has no definite spin. At the moment of measurement, a wavefunction collapse occurs along a specific spatial direction, determined by the measuring apparatus. If the apparatus is oriented vertically, the particle will 'reveal' its spin along only that axis. Since realism only considers bodies with objective properties, it is incompatible with the standard interpretation of quantum mechanics.

Whereas the first argument against realism is formulated in the context of the standard formulation of quantum mechanics, the *second* one considers some more unconventional approaches. In the words of d'Espagnat: *Of them, admittedly, it can be said that they are genuinely ontological. They are the pilot-wave theory, some other hidden-variable theories (KHDF, GRW, CSL and others) and the spontaneous-reduction theories. These theories are extremely interesting, mainly because of the fact that they do restore strong objectivity (this is why they are 'ontological'), which is not a small achievement. On the other hand, this feature, which is their strong point, turns out to be at the same time a weak point because of the fact that in them there is an equivalent to nondivisibility by thought, namely nonlocality (Bell's theorems) and this feature is much more disturbing in their case than in the case of conventional quantum mechanics. The reason stems from the fact that (...) neither nondivisibility by thought nor nonlocality allows for superluminal signaling. When we have to do with a nonlocal, weakly objective theory this makes it possible to - if not completely forget about such things - at least speak, in Shimony's and Redhead's words, of a 'peaceful coexistence' between it and relativity theory (also conceived of as weakly objective only). Within the realm of a theory the strong objectivity of which is stressed, this kind of escape is of course no more admissible since, for consistency, also relativity must then be conceived of as a strongly objective theory: so that it is not only the 'signals' that should travel no faster than light but also any 'influence' whatsoever, which is hardly or not at all reconcilable with nonlocality (...). Hence what makes the interest and the value of these ontological theories is also, ironically, the source of their main weakness (...)*[13].

This rather lengthy argument deserves some explanation. The unconventional interpretations of quantum mechanics are characterized as 'ontological'. 'Ontological' refers to the being of the object. In an ontological interpretation of quantum mechanics, an object has well-defined properties at any time, independent of measurement. Measurements concerning two-particle correlation have pointed out that there is a superluminal kind of causality which, however, cannot be exploited for superluminal signaling: the reason is that this superluminal kind of causality can only be ascertained after comparison of the measurement data of the two distant particles by subluminal means. From the moment that one particle is detected, the outcome of the other is established, but it is not possible to program the outcome of the distant measurement by changing the interaction with the nearby particle.

According to d'Espagnat, these features do not bring quantum mechanics and relativity into contradiction when both are considered from a weakly objective point of view. They both do not allow for superluminal signaling, and superluminal causality can be conceived as somehow induced by the observer. (Since the argument is somewhat hazy, specialists

[12] Ref. 2, p 324.
[13] Ref. 2, p 320-321.

introduce the diplomatic term 'peaceful coexistence'.) In the context of strong objectivity, however, superluminal effects have their objective counterpart in reality, and the only way to restore compatibility with relativity is to reformulate relativity completely, and it is by no means clear that such a task will be possible.

The *third* argument concerns the conspicuous absence of an ontological interpretation of conventional measurement theory. Apart from the standard interpretation of quantum mechanics (which is weakly objective) and the ontological models (which are strongly objective) there exists a whole class of measurement theories that are built up within the conventional quantum framework. According to d'Espagnat, these theories are not ontologically interpretable, although they were formulated with that goal: *On this issue the review (...) of important theories of measurement and classical appearances built up within the conventional quantum framework is quite instructive. It is all the more so because most of the authors of these theories - reluctant as they, apparently, are to go into philosophical considerations - seem to have genuinely believed they were constructing theories capable, like old classical physics, of being understood as describing physical reality as it really is. And yet, by scrutinizing the premises of these theories we found out that, at the present stage in their development, if no new ingredient is added, not a single one of these descriptions - be it the environment-based theory, the formalism of operations and effects, the Griffiths theory, the Omnes theory, the Gell-Mann and Hartle theory (...) - is ontologically interpretable* [14].

The *fourth* argument concerns the inability of realism to cope with intersubjectivity since the advent of quantum mechanics: *Imagine for example that Alice and Bob both perform, one immediately after the other, a measurement of the position of one and the same electron, each using his own instrument. And assume further that before the first measurement the electron wave function is not a 'delta function', which is by far the most general case. Then the rules of quantum mechanics unambiguously predict intersubjective agreement: when Alice and Bob later compare their notebooks they will discover that they both saw the electron at the same place. Before the first measurement, however, as the quantum mechanical formalism tells us, the electron was at no definite place whatsoever, and therefore, in particular, it was not really at that place. The, allegedly obvious, realistic 'teapot-like' explanation of intersubjective agreement is, in this case, simply false* [15].

In this argument, one has the impression that d'Espagnat confuses intersubjectivity with the impossibility of measurement without disturbing the system. What d'Espagnat says about the behavior of the electron faced with the two different experimental devices is certainly true. However, in the context of intersubjectivity one needs to explain that there is a ground that two scientist should agree on a single fact (the electron position as measured by the first apparatus), *not that two different facts* (the electron position measured with different devices) *should be related in some special manner.* As a matter of fact, exactly the same problem, envisaged by d'Espagnat, occurs in classical, strongly objective physics as well: it is impossible to measure the position (or whatever strongly objective quality) of a particle without changing that quality. Measurement without physical interaction just boils down to plain prophecy!

It is interesting to note, by the way, that throughout these four arguments against realism, the word 'ontological' can be harmlessly replaced by 'deterministic'. More generally one has the impression that for d'Espagnat, realism is inevitable burdened with determinism. This association is quite common among philosophers of science, and for an obvious reason: in the

[14] Ref. 2, p 355.
[15] Ref. 2, p 19.

history of natural sciences, ontological realism has always coexisted with determinism. But what may be true for the domain of natural sciences is not at all the case for philosophy! Many pre-Cartesian philosophers were concerned with reconciling realism and human liberty, a task certainly unthinkable in a deterministic framework. However, when understanding realism in its reduced meaning of determinism, one can fully agree with d'Espagnat that 'conventional realism' is incompatible with a consistent interpretation of quantum mechanics.

III. The New Synthesis: Veiled Reality

As a consequence of the inconveniences in both the realistic and idealistic outlooks, d'Espagnat opts for *veiled reality*. The program of veiled reality is built on two main postulates.

The first postulate concerns the existence of two levels of reality, empirical and 'veiled': *(...) by definition, empirical reality is a set of phenomena, that is, that its description is one, not of how things really are, but of how they appear to the collectivity of mankind* [16]. Veiled reality is a distinct order of reality, essentially independent of man, and which can be known by man in a merely general or merely allegorical way [17]. This double layer of reality ensures, on one hand, that our theories cannot be purely ontological, since the veiled part of reality would never allow for such a description, and on the other hand, that phenomena cannot be the sole product of human mind, and that some 'external agent', veiled reality, is responsible for intersubjectivity. This double structure of reality is strongly reminiscent of Plato's 'allegory of the cave' [18]. Plato describes the condition of prisoners in a cave who can see nothing but shades on a cave wall, shed by hidden objects at firelight. The prisoners cannot see these objects directly. Plato assumes that the prisoners do not even think of these objects as existing: to them, it is the shades only which exist. The analogy is evident: Plato's hidden objects play the role of veiled reality, and his shades the role of empirical reality. Quite different are the purposes that led Plato and d'Espagnat to postulate the double level of reality: while the latter seeks to eliminate some inconsistencies in the interpretation of quantum mechanics, the former was concerned with educating the future ruling class of Greek society.

The second postulate of 'veiled reality' contemplates the existence of two types of causality, 'empirical' and 'extended': *Basic Postulate. Any observed regularity (statistical or otherwise) must have a cause (or a set of causes: the notion of 'oneness' is not stressed here), which (i) may be or not be located in time and (ii) may be or not be discoverable by men. (...) Point (i) implies in particular that the general causes of an event are not necessarily located in its past. According to it, the assertion that efficient causality is a scientific notion whereas final causality is not, is meaningful only within empirical reality. Point (ii) in fact extends the range of the notion of cause far beyond the one of the mere notion of explanation (...): reduction to the known and grouping under some general law. Here it is considered that it is not meaningless to speak of causes even if these are of such a nature that they conceivably can never be discovered, and the 'extended causes' we have in mind are considered as being prior to laws. After all, we should keep in mind the truism that the number of brain connections is (huge but) finite and that therefore it is by no means a logical truth that mind can discover the whole of what exists and has observable effects* [19].

Without asserting so explicitly, d'Espagnat at this point fully endorses the view that the information contained in nature is essentially superior to human mind. Let us resume the facts

---

[16] Ref. 2, p 371.
[17] Cfr. ref. 2, p 355.
[18] PLATO, *The Republic*, Book VII.
[19] Ref. 2, p 415-416.

that led d'Espagnat to postulate the existence of an undiscoverable cause (the second part of his postulate). First, the two-particle experiments in quantum mechanics show a specific correlation which is fully accounted for by the quantum-mechanical formalism, but the control of which is in plain contradiction with (i) relativity (insofar as superluminal signaling is implied) and (ii) the completeness of quantum mechanics (according to which this correlation is inherently uncontrollable, due to the lack of 'definedness' of the system before the instant of detection). Suarez explained the uncontrollable correlations in terms of unobservable causes[20], quite in agreement with veiled reality. The second fact which led d'Espagnat to postulate the existence of an undiscoverable cause, is that the uncontrollable quantum correlations manifest themselves on the level of the phenomena in a strongly objective way, i.e., as if the correlations were an intrinsic property of the system. In a single sentence: here we have an event that nobody can predict, not now nor in the far future, but everybody must agree on its existence.

'Veiled reality' represents an important step towards a consistent interpretation of quantum theory. The major advance of d'Espagnat consists in his recognizing the need of an independent reality, and in his stressing that regularity on the level of phenomena always urges for an explanation in terms of causality. Independent reality and causality have been so heavily attacked since the very birth of quantum mechanics, that d'Espagnat may rightfully be considered a pioneer in the field of philosophy of quantum mechanics. And to my opinion, his work provides a strong indication of the existence of God. To see this, let us concentrate on d'Espagnat's second remark to his own definition of extended causality: *It is true, however, that such a notion of extended causes, which theologians might identify to that of 'primary causes', should not be handled without care. A well-known objection to it is that it apparently leads to an infinite regression. 'Why is it that the Earth does not indefinitely fall down in empty space?' a disciple asked his guru in a famous tale. 'Because it rests on the back of an elephant', the latter replied. 'And why does not the elephant itself indefinitely fall down in empty space?' the insatiable disciple asked. 'Because', the tale goes on, 'it rests on the shell of a giant tortoise', and so on. The fable appropriately reminds us that the radical idealists and the Kantians are justified in questioning naive forms of the quest for causes. It can, however, not be considered as carrying any decisive objection. After all, Galileo was right when he asserted that preservation by an object of the rectilinear uniform motion it has requires no cause. Similarly, it is quite natural to counter the objection embodied in the fable by considering that the general structures of independent reality, which, we claim, are the causes of the observed objective regularities, do not themselves require causes* [21]. In the view of d'Espagnat, quantum mechanics has taught us the existence of a new kind of causality which is located outside space-time, and may be undiscoverable by man. This is the kind of causality responsible for the two-particle correlations in quantum mechanics. It is the kind of causality which connects independent, veiled reality, to empirical reality, the reality of our everyday impressions. It is a kind of causality which is not a mere tautology - like al causes are in a deterministic world view -, but it has a kind of 'existential flavor'. This causality obviously begs for the question of infinite regression. But the same d'Espagnat who earlier in his book heavily attacks the idealists for not receiving 'why-questions' now seems in urgent need to declare the question concerning infinite regression non-receivable! Nowhere stronger than at this point in 'veiled reality' one is faced with the question of the existence of God.

IV. Conclusion

[20] A. SUAREZ, in *'Naturwissenschaft und Weltbild'*, ed. H.-C. Reichel & E. Prat de la Riba, Verlag Hölder-Pichler-Tempsky Vienna (1992).

[21] Ref. 2, p 417.

D'Espagnat has convincingly shown that both conventional realism and radical idealism are unable to give a consistent interpretation of quantum mechanics. Conventional realism fails mainly due to its deterministic character. Idealism fails because it presupposes that the causes of all phenomena can be reduced to the human mind. According to d'Espagnat, this reduction makes it impossible to understand how typical quantum-mechanical effects, like many-particle correlations, can be inherently uncontrollable. The idealistic failure can be overcome by introducing the concepts of veiled reality and extended causality, which in d'Espagnat's very definitions imply the existence of a world that is radically independent from human mind.

A world view that allows for a rational explanation of quantum-mechanical effects has to accept a form of causality which originates from a 'veiled reality', underlying the world of phenomena. This veiled reality, as d'Espagnat conceives it, is the ultimate ground for the intersubjectivity of human observation.

D'Espagnat's introduction of 'veiled reality' finishes with the idealistic tradition of concentrating 'existence' in the constructs of human mind, at the expense of the underlying reality, which since Kant has gradually disappeared from the philosophical horizon. And his introduction of 'extended causality' finishes with the Kantian tradition of reducing causality to a mere chain of deterministic tautologies.[22]

The world view sketched by d'Espagnat is rather unusual. Starting from the empirical facts of intersubjectivity and quantum correlations, d'Espagnat is led to postulating the existence of 'veiled reality', a layer of reality that one can know but partially and indirectly. This veiled reality exists independently from human mind, and is found to be essentially superior to it, in the sense that it is intrinsically unpredictable, that it cannot be known by aprioristic arguments. If science is ever to give us an ingredient for a philosophical proof of the existence of God, what better suited fact could it offer but the existence of an independent, external world, that can impossibly originate from the human mind?

[22] I. KANT, *Kritik der reinen Vernunft*, B 232-257.

Chapter XII

Scientism and Scientific Knowledge of Things and God

Jacques Laeuffer, Paris, France

Introduction

What is scientism? It may be defined as a philosophical interpretation of modern science. In other words, scientism is not science but a philosophical view of science. It is based on the assumption that one day, in the future, science will be able to understand everything. One day human beings will enjoy absolute and total knowledge of the world. And this knowledge will imply an unlimited power and would finally lead to complete happiness. As a consequence, the idea of God is of no further help for man, because hope for a better future will have become obsolete.

How can we get a knowledge of things? The purpose of this paper is to examine this question, and show that it has exercised a great influence on modern thinking in the area of physics as well as on the question of the existence of God. In the second section, we will discuss how man is able to know things, and what he can say about God. The next section deals with the question of how scientism appeared in the 19th century. In the fourth section we shall examine some of the roots of scientism and try to analyse why it was so successful. After a discussion of the relationship between our present and future knowledge of the world, we end with some conclusions.

2. The nature of human knowledge and the question of the existence of God

When we begin to open our eyes, we discover the world. The things we see were already there before we came onto the scene. The earth did not wait for man to arrive in order to start existing. Things had the properties they have now, before we were able to understand them. For example, the earth did not wait for Copernicus to orbit around the sun. The Big-Bang occurred probably about thirteen billion years ago, and man appeared later. We are told all of this by science. Man discovers the world: “discover” means “removing the cover” from things. In other words, before our dis-covery, the thing was present under the “cover” of our ignorance. *There exists a physical reality outside ourselves, independent of our thoughts and our imperfect means of knowing it. If this was not the case, the unity of human knowledge, and the agreement of all men on the observation of facts, would be incomprehensible*[1].

Experience demonstrates clearly that our knowledge of particular things is not absolute and perfect. Every day we discover new aspects, new properties in things that we have already examined many times before. In addition, every answer leads us to new questions. One may say that man lives, in a certain sense, in a cage. Science can enlarge the size of the cage, but it does not open it.

Human beings are able to discover properties in things because they are gifted with intelligence. Things are not intelligent in the same sense as man, but nevertheless contain traces of intelligence. The equations of Maxwell or Schrödinger, for example, reveal a very

[1] Louis de BROGLIE : "Les idées qui me guident dans mes recherches", Albin-Michel, Paris, 1966, in *Un itinéraire scientifique*, Editions La Découverte, Paris, 1987, p. 149.

intelligent organization of things. But light and matter did not wait for these famous scientists to follow the above-mentioned equations. Therefore the order expressed by these equations does not originate from Maxwell and Schrödinger. What is their origin? Light and matter are things without intelligence. The intelligent thing or being from whom these laws originate, as well as many other forms of organization, has been called "God" by many people of all times.

One is inclined to object that this is not a real explanation. Postulating God is not an explanation in the sense that we know any details about the nature of the specific laws that apply for light and matter. These subjects are studied extensively and with great success in science. However, it is an explanation in the sense that all the above-mentioned scientific research does not give a real answer to the question: what is the cause of this organization? Introducing a new concept - God - is a way of recognizing that we have made a new discovery: that there exists a principle or being which is the cause of the organization of things.

This argument, already developed in the Greek philosophy by Plato[2] and Aristotle[3], was presented in the 13th century by Thomas Aquinas, in the so-called 5th way for the demonstration of the existence of God. *The fifth way is taken from the governance of the world. We see that things which lack intelligence, such as natural bodies, act for an end, and this is evident from their acting always, or nearly always, in the same way, so as to obtain the best result. Hence it is plain that not fortuitously, but designedly, do they achieve their end. Now whatever lacks intelligence cannot move towards an end, unless it be directed by some being endowed with knowledge and intelligence; as the arrow is shot to its mark by the archer. Therefore some intelligent being exists by whom all natural things are directed to their end; and this being we call God.*[4].

In this text, God is not considered as the right hypothesis from which deductions may afterwards be made. Here the existence of an intelligent being, which is called God, is the conclusion of a demonstration based on physical observation. A similar idea can be found in the Jewish tradition: *«The heavens are telling the glory of God; and the firmament proclaims his handiwork*[5]*»*.

Like Aristotle, Thomas Aquinas makes a distinction between "natural things" and "artificial things". Natural things can be found in Nature, whereas artificial things are made by human beings as a product of art or workmanship. The existence of natural things is sufficient for the demonstration of Aquinas given above. But the existence and the effectiveness of artificial things can be considered as a new argument. In fact, human technology is based upon the knowledge of the laws of nature, and artificial things are nothing but the result of combining or processing of natural things. As a consequence, artificial things work because the natural things they are made of, follow the laws of nature. For instance, copper is a natural element that is able to conduct electricity. An electrical device, that is an artificial thing, works properly, because the connections made out of copper conduct electricity like the natural element copper. The effectiveness of technology is the proof that natural things act *always or nearly always, in the same way, so as to obtain the best results*[6]. There is no technology without physics, and "physics" is the Greek word for the Latin term

[2] Plato, Timaeus 90 a-e.
[3] Aristotle, De philosophia, Fr. 10 R.
[4] THOMAS AQUINAS, *Summa Theologica* I, quaest.2, art.3.
[5] THE BIBLE, *Psalm 18.*
[6] THOMAS AQUINAS, op. cit., I, quaest.2, art.3.

"nature". Therefore, one may say that technology and its effectiveness reinforce the argumentation of the fifth way of Aquinas.

A lot of people, however, are convinced that belief in God is incompatible with a scientific attitude. The origin of this conviction can be found in the scientistic way of thinking. In the following section, a brief historical view of scientism is given.

3. The scientism of Laplace and Comte

Laplace (1749-1827) was a scientist at the time of the French revolution. Basically, he was a mathematician; in addition he applied the methodology used in mathematics to the area of physics. In this way he tried to treat science as a hypothetical-deductive system. As a member of the new Academy of Sciences, he presented the "scientific system" of that academy to Napoleon Bonaparte. Napoleon asked: "Where is the place for God in your system?" His answer was: "Majesty, we did not need that hypothesis."

In a well known text, Laplace defines his view of science: *We have to consider the actual state of the universe as the effect of its previous state, and as the cause of the state that will follow. Imagine an intelligent being which, at a certain time, would know all the forces at work in nature, and the respective location of all the beings that compose it. If it were powerful enough to analyse these data, it could embrace together in the same formula the movements of the largest bodies of the universe and those of the lightest atom. Nothing would be uncertain for this intelligent being and both the future and the past would be present to its eyes. The human mind offers, with the perfection it gave to astronomy, a weak impression of this intelligence. Its discoveries in the fields of mechanics and geometry, together with that of universal gravitation, now enable it to understand in a single analytical expression the present and future states of the system of the world.*[7]

This "mechanistic" point of view is a reduction of the material reality to the mathematical description we have of it. One could remark that this view reveals a confusion between the physical reality and the imperfect knowledge we have of it. The perfection of knowledge, i.e. true knowledge, consists primarily in the conformity of the intellect with reality, not in its mathematical simplicity. Furthermore, knowledge is not only about quantity or location, but also about qualities, and qualities are not reflected in mathematical equations.

If Laplace could be considered as a precursor of Scientism, Auguste Comte (1798-1857) should be seen as the real founder. From 1830 to 1842 he published his six volumes of the *Course on Positive Philosophy*. His *Positivistic Catechism Or An Overview Of The Universal Religion, In Eleven Systematic Conversations Between A Woman And A Priest Of Humanity* was published in 1856. Not a title a scientist nowadays would use. Comte states that humanity has now reached a new age, the age of science. He calls his theory "positivism", and it will later become to be known as "scientism". The title itself of the above-mentioned book shows that he intends to introduce science as a new religion.

Comte considers mathematics to be the starting point for human thinking. *Mathematical Science has to be the true beginning of any scientific and rational education, either general or specialised*[8]. Who would admit the priority of mathematics in the areas of literature, philosophy, history, etc.? Even in physics, the first experiences and knowledge are about qualitative properties: the child experiences that he may burn his fingers on a dark-red

[7] LAPLACE, *Essai philosophique sur les probabilités*, Gauthiers-Villars, Paris, 1921, p. 3.
[8] Auguste COMTE, *op.cit.*, Lesson 2, p. 64.

glowing oven plate. This experience is the basis of scientific knowledge, a knowledge of the causes. Comte comments in another place: *«Mathematical analysis is ... the true rational basis for the whole system of our positive knowledge. It constitutes the first and the most perfect of the fundamental sciences. The ideas, with which it deals, are the most universal, the most abstract and the simplest ones we can realistically conceive*[9]. One should remark once again that the first perfection of human knowledge is not simplicity, but its conformity to reality.
On this basis, Comte promises happiness to humanity: *We are building up directly the system of general ideas that will make this philosophy indefinitely prevailing in the human species. And the revolutionary crisis that torments the civilised peoples will essentially be terminated*[10]. He thinks that this new age will appear in the near future: *This general evolution of the human spirit is almost accomplished today. All that remains, as I already explained, is to complete the positive philosophy by including social phenomena, and then to resume it in one set of homogeneous doctrine. When this two-fold task has sufficiently progressed, the triumph of positive philosophy will occur spontaneously, re-establishing order in society.*[11]

Auguste Comte considered that science will increasingly appear as a substitute for religion. We shall no longer need God because science will provide us with absolute knowledge. Scientism assumes that man will take total control of the world. In 1900, people used to say: "Within ten years, laboratories will have to shut down, because everything will have already been discovered."

4. Some origins and consequences of scientism's success

Soon, to announce morning,
the sun will arise on her golden path,
soon shall superstition disappear,
soon the wise man will conquer.
Ah, gracious peace, descend,
return again to the hearts of men;
then the earth will be a heavenly kingdom,
and mortals like the gods[12]

Such was the dream at the end of the 18th century, when Mozart composed "The Magic Flute". This is poetry, but it reflects very well man's aspiration to achieve absolute happiness on this earth. This was the aim of the "Societies of thought" of the "Age of Enlightenment". The librettist wrote his "Finale" in the future tense. The dream is based on what "the wise man" knows. A new age must begin, symbolised by the morning and the sun. Such an aspiration cries out for justification by science. But up to now all approaches are neither very convincing nor accepted by the scientific community. It is remarkable, however, that whenever a "man of science" proposes such a justification, he seems to be successful, at least in the popular literature.

The consequences of scientism were recently analysed by Pierre-Gilles de Gennes, Nobel-Prize winner for Physics and director of the Ecole Nationale Supérieure de Physique et de Chimie, in Paris. He comments on the classification of sciences made by Auguste Comte: *I would now like to speak about a typical prejudice of the French culture, a prejudice inherited*

[9] Auguste COMTE, *op.cit.*, Lesson 3, p. 76.
[10] Auguste COMTE, *op.cit.*, Lesson 1, p. 39.
[11] Auguste COMTE, *op.cit.*, Lesson 1, p. 39.
[12] W. A. MOZART, *The Magic Flute*, Act two, Scene. 7, No. 21 Finale.

from Auguste Comte's positivism. This 19th century philosopher built up his glory by establishing a classification of sciences. At the top of the hierarchy, mathematics; at the bottom, chemistry which, according to him, "hardly deserves to be considered a science "(!), in the middle, astronomy and physics. This classification excluded geography and mineralogy, sciences that he considered concrete and descriptive, thus keeping only theoretical, abstract and general sciences.[13]

In his opinion, Auguste Comte is one of the main obstacles to the development of science: *Scientific education, the disciplines, and even the scientists themselves, still suffer from "the Auguste Comte prejudice". This prejudice may be the root for the subsequent contempt for manual work*[14].

As an illustrative example of the consequences of scientism for the way of thinking of students, Pierre-Gilles de Gennes describes a typical experience from his university teaching. *Some students came to study solid state physics for a "Diplôme d'Etudes Approfondies" at Orsay. They arrived convinced that they could know everything on the basis of calculations. By the end of the year, I gave them a problem.... They took an hour thinking at a corner of the black board. The solution was rather simple at this level of studies.... But these students I met at that time remained mute in front of the black board. One of them, at the end, told me (I shall never forget this phrase): "But Sir, which Hamiltonian do I have to diagonalize?" He was trying to hang on to theoretical ideas, that had nothing to do with this practical problem. This kind of answer explains, for a great part, the weakness of French industrial research.*[15]

5. Present and future knowledge

Human knowledge originates from the things we study. In science, we form our ideas of reality by observation. One can illustrate the way science deals with reality by means of a picture taken from Plato. In his famous passage of the shadows in a cave, Plato describes the difference between human knowledge and eternal Ideas. The ideas with a small "i" are those found in the human mind. The Ideas with a capital "I" are different: Regarding these, Plato explains that they are eternal; they existed before material things, and material things are made in imitation of eternal Ideas. These Ideas exist in God's mind, not in a human mind. Human ideas are images of things, and things are images of eternal Ideas. Therefore, human ideas are images, but only imperfect images of eternal Ideas. Between eternal Ideas and human ideas, the difference is not a matter of position along a continuous line, it is a difference of nature. Eternal Ideas are the causes of things, and things are the causes of human ideas. As eternal Ideas are seen by man only through the "shadows" of Plato's cave, human beings will never reach the level of God.[16]

The history of Physics shows man's impressive progress in dominating nature. But this raises a further question. Is there a limit beyond which man's true but imperfect knowledge of nature cannot go? Nobody can really say, because with respect to a phenomenon he does not understand, nobody would claim that he will never understand it. Nobody can say that a certain limit of knowledge will never be surpassed; this can be seen from the history of physics. When an obstacle seemed to be unsurpassable, it was often the consequence of a desire for absolute knowledge, which is fixed in an axiomatic system. When this system was changed, science progressed.

---

[13] Pierre-Gilles de GENNES, *Les objets fragiles*, Plon, Paris, 1994, pp. 231-232.
[14] Pierre-Gilles de GENNES, *op.cit.*, p. 232.
[15] Pierre-Gilles de GENNES, *op.cit.*, pp. 232-233.
[16] cf PLATO, *The Republic*, BookVII.

6. Conclusion

Scientism has been in the root of a lot of misunderstandings in the area of science as well as in the field of religion. Considering our ideas about things as being absolute will often lead to disappointments. For instance, we know from experience that equations and relationships will not always work to highest precision. Therefore we have to continue to do research, and this will result in the discovery of new influences and new equations.

At the same time, the scientific knowledge we do have is true. A first proof of this can be seen in the fact that many different researchers are able to reach the same conclusions. Technology provides a second proof: we can predict the behaviour of something that we have produced. This is valid in the area of scientific experiment, as well as in the area of industrial applications. Man's technological capability proves that his knowledge of nature is true. From time to time, however, technology fails, reminding us that our knowledge is not absolute.

What will we know in the future? *The researcher is not a prophet, but an explorer, often hesitating and tired*[17]. Nothing compels us to think that we could at one day reach absolute knowledge, as all past attempts to do so have failed. In any case, the uncertainty of our grade of knowledge in future will not change anything with regard to the actual question of the existence of God. The reason is that an actual present state cannot be caused or influenced by a potential future. The fact that we are not the cause of what we have discovered up to now shows that the organization of the actual present state of nature does not come from us. One may therefore state that science is compatible with the existence of God. Science, since it is both true and limited, points to the existence of an organized state of things and the existence of an organizer, and this is what people in general call God.

[17] Pierre-Gilles de GENNES, *op.cit.*, p. 14.

Chapter XIII

Physics and the Mind of God

The Templeton Prize Address*

Paul Davies
Department of Physics and Mathematical Physics, University of Adelaide, Australia
E-mail: pdavies@physics.adelaide.edu.au

1. Introduction

It is both an honour and a pleasure for me to deliver my acceptance address for the 1995 Templeton Prize to such a distinguished audience in this world-famous Abbey, just a few metres from the remains of Isaac Newton. Along with Einstein and Darwin, Newton is one of the few scientists known to almost every member of the population. He is one of the great heroes of my own discipline, physics, even if his career as a civil servant left a lot to be desired.

It was Newton, Galileo and their contemporaries who created science as we know it, three centuries ago. Today we take the scientific method of enquiry so much for granted that few people stop to think how astonishing it is that science works.

I was fascinated to learn that one of the judges for the Prize was Baroness Thatcher. This is in fact the second time she has been involved in giving me a prize. The first occasion was in 1962, at the Speech Day of Woodhouse Grammar School in North Finchley, when she presented me with a copy of Norton's Star Atlas for doing well in my O level exams. I doubt if her Ladyship recalls the encounter, but I can trace my own decision to become a scientist to more or less that event.

Like all school pupils, I learned science as a set of procedures that would reveal how nature works, but I never questioned why we were able to do this thing called science so successfully. It was only after a long career of research and scholarship that I began to appreciate just how deep scientific knowledge is, and how incredibly privileged we human beings are to be able to unlock the secrets of nature in such a powerful way.

Of course, science didn't spring ready-made into the minds of Newton and his colleagues. They were strongly influenced by two longstanding traditions that pervaded European thought. The first was Greek philosophy. Most ancient cultures were aware that the universe is not completely chaotic and capricious: there is a definite order in nature. The Greeks believed that this order could be understood, at least in part, by the application of human reasoning. They maintained that physical existence was not absurd, but rational and logical, and therefore in principle intelligible to us. They discovered that some physical processes had a hidden mathematical basis, and they sought to build a model of reality based on arithmetical and geometrical principles.

---

* *The following is the text of the address delivered in Westminster Abbey on May 3, 1995 by Professor Paul Davies on the occasion of his receiving the Templeton Prize for Progress in Religion.*

The second great tradition was the Judaic world view, according to which the universe was created by God at some definite moment in the past and ordered according to a fixed set of laws. The Jews taught that the universe unfolds in a unidirectional sequence - what we now call linear time - according to a definite historical process: creation, evolution and dissolution. This notion of linear time - in which the story of the universe has a beginning, a middle and an end - stands in marked contrast to the concept of cosmic cyclicity, the pervading mythology of almost all ancient cultures. Cyclic time - the myth of the eternal return - springs from mankind's close association with the cycles and rhythms of nature, and remains a key component in the belief systems of many cultures today. It also lurks just beneath the surface of the Western mind, erupting occasionally to infuse our art, our folklore and our literature.

A world freely created by God, and ordered in a particular, felicitous way at the origin of a linear time, constitutes a powerful set of beliefs, and was taken up by both Christianity and Islam. An essential element of this belief system is that the universe does not have to be as it is: it could have been otherwise. Einstein once said that the thing which most interested him is whether God had any choice in his creation. According to the Judaeo-Islamic-Christian tradition, the answer is a resounding yes.

Although not conventionally religious, Einstein often spoke of God, and expressed a sentiment shared, I believe, by many scientists, including professed atheists. It is a sentiment best described as a reverence for nature and a deep fascination for the natural order of the cosmos. If the universe did not have to be as it is, of necessity - if, to paraphrase Einstein, God did have a choice - then the fact that nature is so fruitful, that the universe is so full of richness, diversity and novelty, is profoundly significant.

2. The origin of laws in nature

Some scientists have tried to argue that if only we knew enough about the laws of physics, if we were to discover a final theory that united all the fundamental forces and particles of nature into a single mathematical scheme, then we would find that this superlaw, or theory of everything, would describe the only logically consistent world. In other words, the nature of the physical world would be entirely a consequence of logical and mathematical necessity. There would be no choice about it. I think this is demonstrably wrong. There is not a shred of evidence that the universe is logically necessary. Indeed, as a theoretical physicist I find it rather easy to imagine alternative universes that are logically consistent, and therefore equal contenders for reality.

It was from the intellectual ferment brought about by the merging of Greek philosophy and Judaeo-Islamic-Christian thought, that modern science emerged, with its unidirectional linear time, its insistence on nature's rationality, and its emphasis on mathematical principles. All the early scientists such as Newton were religious in one way or another. They saw their science as a means of uncovering traces of God's handiwork in the universe. What we now call the laws of physics they regarded as God's abstract creation: thoughts, so to speak, in the mind of God. So in doing science, they supposed, one might be able to glimpse the mind of God. What an exhilarating and audacious claim!

In the ensuing three hundred years, the theological dimension of science has faded. People take it for granted that the physical world is both ordered and intelligible. The underlying order in nature - the laws of physics - are simply accepted as given, as brute facts. Nobody asks where they come from; at least they don't in polite company. However, even the most atheistic scientist accepts as an act of faith the existence of a lawlike order in nature that

is at least in part comprehensible to us. So science can proceed only if the scientist adopts an essentially theological world view.

It has become fashionable in some circles to argue that science is ultimately a sham, that we scientists read order into nature, not out of nature, and that the laws of physics are our laws, not nature's. I believe this is arrant nonsense. You'd be hard-pressed to convince a physicist that Newton's inverse square law of gravitation is a purely cultural concoction. The laws of physics, I submit, really exist in the world out there, and the job of the scientist is to uncover them, not invent them. True, at any given time, the laws you find in the textbooks are tentative and approximate, but they mirror, albeit imperfectly, a really-existing order in the physical world. Of course, many scientists don't recognize that in accepting the reality of an order in nature - the existence of laws "out there" - they are adopting a theological world view. Ironically, one of the staunchest defenders of the reality of the laws of physics is the American physicist Steven Weinberg, a sort of apologetic atheist who, though able to wax lyrical about the mathematical elegance of nature, nevertheless felt compelled to pen the notorious words: *The more the universe seems comprehensible, the more it also seems pointless*.

Let us accept, then, that nature really is ordered in a mathematical way - that *the book of nature*, to quote Galileo, *is written in mathematical language*. Even so, it is easy to imagine an ordered universe which nevertheless remains utterly beyond human comprehension, due to its complexity and subtlety. For me, the magic of science is that we can understand at least part of nature - perhaps in principle all of it - using the scientific method of enquiry. How utterly astonishing that we human beings can do this! Why should the rules on which the universe runs be accessible to humans?

The mystery is all the greater when one takes into account the cryptic character of the laws of nature. When Newton saw the apple fall, he saw a falling apple. He didn't see a set of differential equations that link the motion of the apple to the motion of the moon. The mathematical laws that underlie physical phenomena are not apparent to us through direct observation; they have to be painstakingly extracted from nature using arcane procedures of laboratory experiment and mathematical theory. The laws of nature are hidden from us, and are revealed only after much labour. The late Heinz Pagels - another atheistic physicist - described this by saying that the laws of nature are written in a sort of cosmic code, and that the job of the scientist is to crack the code and reveal the message - nature's message, God's message, take your choice, but not *our* message. The extraordinary thing is that human beings have evolved such a fantastic code-breaking talent. This is the wonder and the magnificence of science; we can use it to decode nature and discover the secret laws that make the universe tick.

3. Time and eternity in the physical universe

Many people want to find God in the creation of the universe, in the big bang that started it all off. They imagine a superbeing who deliberates for all eternity, then presses a metaphysical button and produces a huge explosion. I believe this image is entirely misconceived. Einstein showed us that space and time are part of the physical universe, not a pre-existing arena in which the universe happens. Cosmologists are convinced that the big bang was the coming-into-being, not just of matter and energy, but of space and time as well. Time itself began with the big bang. If this sounds baffling, it is by no means new. Already in the fifth century St. Augustine proclaimed that *the world was made with time, not in time*. According to James Hartle and Stephen Hawking, this coming-into-being of the universe need

not be a supernatural process, but could occur entirely naturally, in accordance with the laws of quantum physics, which permit the occurrence of genuinely spontaneous events.

The origin of the universe, however, is hardly the end of the story. The evidence suggests that in its primordial phase the universe was in a highly simple, almost featureless state: perhaps a uniform soup of subatomic particles, or even just expanding empty space. All the richness and diversity of matter and energy we observe today has emerged since the beginning in a long and complicated sequence of self-organizing physical processes. What an incredible thing these laws of physics are! Not only do they permit a universe to originate spontaneously; they encourage it to self-organize and self-complexify to the point where conscious beings emerge, and can look back on the great cosmic drama and reflect on what it all means.

Now you may think I have written God entirely out of the picture. Who needs a God when the laws of physics can do such a splendid job? But we are bound to return to that burning question: Where do the laws of physics come from? And why those laws rather than some other set? Most especially: Why a set of laws that drives the searing, featureless gases coughed out of the big bang, towards life and consciousness and intelligence and cultural activities such as religion, art, mathematics and science?

If there is a meaning or purpose to existence, as I believe there is, we are wrong to dwell too much on the originating event. The big bang is sometimes referred to as "the creation", but in truth nature has never ceased to be creative. This ongoing creativity, which manifests itself in the spontaneous emergence of novelty and complexity, and organization of physical systems, is permitted through, or guided by, the underlying mathematical laws that scientists are so busy discovering.

Now the laws of which I speak have the status of timeless eternal truths, in contrast to the physical states of the universe which change with time, and bring forth the genuinely new. So we here confront in physics a re-emergence of the oldest of all philosophical and theological debates: the paradoxical conjunction of the eternal and the temporal. Early Christian thinkers wrestled with the problem of time: is God within the stream of time, or outside of it? How can a truly timeless God relate in any way to temporal beings such as ourselves? But how can a God who relates to a changing universe be considered eternal and unchangingly perfect?

Well, physics has its own variations on this theme. In our century, Einstein showed us that time is not simply "there" as a universal and absolute backdrop to existence, it is intimately interwoven with space and matter. As I have mentioned, time is revealed to be an integral part of the physical universe; indeed, it can be warped by motion and gravitation. Clearly something that can be changed in this manner is not absolute, but a contingent part of the physical world.

In my own field of research - called quantum gravity - a lot of attention has been devoted to understanding how time itself could have come into existence in the big bang. We know that matter can be created by quantum processes. There is now a general acceptance among physicists and cosmologists that spacetime can also originate in a quantum process. According to the latest thinking, time might not be a primitive concept at all, but something that has "congealed" from the fuzzy quantum ferment of the big bang, a relic, so to speak, of a particular state that froze out of the fiery cosmic birth.

If it is the case that time is a contingent property of the physical world rather than a necessary consequence of existence, then any attempt to trace the ultimate purpose or design of nature to a temporal Being or Principle seems doomed to failure. While I do not wish to claim that physics has solved the riddle of time - far from it - I do believe that our advancing scientific understanding of time has illuminated the ancient theological debate in important ways. I cite this topic as just one example of the lively dialogue that is continuing between science and theology.

A lot of people are hostile to science because it demystifies nature. They prefer the mystery. They would rather live in ignorance of the way the world works and our place within it. For me, the beauty of science is precisely the demystification, because it reveals just how truly wonderful the physical universe really is. It is impossible to be a scientist working at the frontier without being awed by the elegance, ingenuity and harmony of the lawlike order in nature. In my attempts to popularize science, I'm driven by the desire to share my own sense of excitement and awe with the wider community; I want to tell people the good news. The fact that we are able to do science, that we can comprehend the hidden laws of nature, I regard as a gift of immense significance. Science, properly conducted, is a wonderfully enriching and humanizing enterprise. I cannot believe that using this gift called science - using it wisely, of course - is wrong. It is good that we should know.

4. The position of God with respect to the universe

So where is God in this story? Not especially in the big bang that starts the universe off, nor meddling fitfully in the physical processes that generate life and consciousness. I would rather that nature can take care of itself. The idea of a God who is just another force or agency at work in nature, moving atoms here and there in competition with physical forces, is profoundly uninspiring. To me, the true miracle of nature is to be found in the ingenious and unswerving lawfulness of the cosmos, a lawfulness that permits complex order to emerge from chaos, life to emerge from inanimate matter, and consciousness to emerge from life, without the need for the occasional supernatural prod; a lawfulness that produces beings who not only ask great questions of existence, but who, through science and other methods of enquiry, are even beginning to find answers.

You might be tempted to suppose that any old rag-bag of laws would produce a complex universe of some sort, with attendant inhabitants convinced of their own specialness. Not so. It turns out that randomly-selected laws lead almost inevitably either to unrelieved chaos or boring and uneventful simplicity. Our own universe is poised exquisitely between these unpalatable alternatives, offering a potent mix of freedom and discipline, a sort of restrained creativity. The laws do not tie down physical systems so rigidly that they can accomplish little, nor are they a recipe for cosmic anarchy. Instead, they encourage matter and energy to develop along pathways of evolution that lead to novel variety, what Freeman Dyson has called the principle of maximum diversity that in some sense we live in the most interesting possible universe.

Scientists have recently identified a regime dubbed "the edge of chaos", a description that certainly characterises living organisms, where innovation and novelty combine with coherence and cooperation. The edge of chaos seems to imply the sort of lawful freedom I have just described. Mathematical studies suggest that to engineer such a state of affairs requires laws of a very special form. If we could twiddle a knob and change the existing laws, even very slightly, the chances are that the universe as we know it would fall apart, descending into chaos. Certainly the existence of life as we know it, and even of less elaborate

systems such as stable stars, would be threatened by just the tiniest change in the strengths of the fundamental forces, for example. The laws that characterize our actual universe, as opposed to an infinite number of alternative possible universes, seem almost contrived - fine-tuned some commentators have claimed - so that life and consciousness may emerge. To quote Dyson again: it is almost as if *the universe knew we were coming.* I can't prove to you that that is design, but whatever it is it is certainly very clever!

Now some of my colleagues embrace the same scientific facts as I, but deny any deeper significance. They shrug aside the breathtaking ingenuity of the laws of physics, the extraordinary felicity of nature, and the surprising intelligibility of the physical world, accepting these things as a package of marvels that just happens to be. But I cannot do this. To me, the contrived nature of physical existence is just too fantastic for me to take on board as simply "given". It points forcefully to a deeper underlying meaning to existence. Some call it purpose, some design. These loaded words, which derive from human categories, capture only imperfectly what it is that the universe is about. But, that it is about something, I have absolutely no doubt.

5. The place of man in the cosmic scheme

Where do we human beings fit into this great cosmic scheme? Can we gaze out into the cosmos, as did our remote ancestors, and declare: "God made all this for us!" Well, I think not. Are we then but an accident of nature, the freakish outcome of blind and purposeless forces, an incidental by-product of a mindless, mechanistic universe? I reject that too. The emergence of life and consciousness, I maintain, are written into the laws of the universe in a very basic way. True, the actual physical form and general mental make-up of homo sapiens contains many accidental features of no particular significance. If the universe were re-run a second time, there would be no solar system, no Earth and no people. But the emergence of life and consciousness somewhere and somewhen in the cosmos is, I believe, assured by the underlying laws of nature. The origin of life and consciousness were not interventionist miracles, but nor were they stupendously improbable accidents. They were, I believe, part of the natural outworking of the laws of nature, and as such our existence as conscious enquiring beings springs ultimately from the bedrock of physical existence - those ingenious, felicitous laws. That is the sense in which I have written in my book *The Mind of God*: *We are truly meant to be here*. I mean "we" in the sense of conscious beings, not homo sapiens specifically. Thus although we are not at the centre of the universe, human existence does have a powerful wider significance. Whatever the universe as a whole may be about, the scientific evidence suggests that we, in some limited yet ultimately still profound way, are an integral part of its purpose.

How can we test these ideas scientifically? One of the great challenges to science is to understand the nature of consciousness in general and human consciousness in particular. We still haven't a clue how mind and matter are related, nor what process led to the emergence of mind from matter in the first place. This is an area of research that is attracting considerable attention at present, and for my part I intend to pursue my own research in this field. I expect that when we do come to understand how consciousness fits into the physical universe, my contention that mind is an emergent and in principle predictable product of the laws of the universe will be borne out.

Secondly, if I am right that the universe is fundamentally creative in a pervasive and continuing manner, and that the laws of nature encourage matter and energy to self-organize and self-complexify to the point that life and consciousness emerge naturally, then there will

be a universal trend or directionality towards the emergence of greater complexity and diversity. We might then expect life and consciousness to exist throughout the universe. That is why I attach such importance to the search for extraterrestrial organisms, be they bacteria on Mars, or advanced technological communities on the other side of the galaxy. The search may prove hopeless - the distances and numbers are certainly daunting - but it is a glorious quest. If we are alone in the universe, if the Earth is the only life-bearing planet among countless trillions, then the choice is stark. Either we are the product of a unique supernatural event in a universe of profligate overprovision, or else an accident of mind-numbing improbability and irrelevance. On the other hand, if life and mind are universal phenomena, if they are written into nature at its deepest level, then the case for an ultimate purpose to existence would be compelling.

6. Concluding remarks

Finally, let me turn to the theme of the Templeton Prize itself: progress in religion. It is often pointed out that people are increasingly turning away from the established religions. However, it remains as true as ever that ordinary men and women yearn for some sort of deeper meaning to their lives. Our secular age has led many people to feel demoralised and disillusioned, alienated from nature, regarding their existence as a pointless charade in an indifferent, even hostile, universe, a meaningless three-score years and ten on a remote planet wandering amid the vastness of an uncaring cosmos. Many of our social ills can be traced to the bleak world view that three hundred years of mechanistic thought have imposed on us, a world view in which human beings are presented as irrelevant observers of nature rather than an integral part of the natural order. Some may indeed recoil from this philosophy and find comfort in ancient wisdom and revered texts that place mankind at the pinnacle of creation and the centre of the universe. Others choose to put their faith in so-called New Age mysticism, or resort to bizarre religious cults.

I would like to suggest an alternative. We have to find a framework of ideas that provides ordinary people with some broader context to their lives than just the daily round, a framework that links them to each other, to nature and to the wider universe in a meaningful way, that yields a common set of principles around which peoples of all cultures can make ethical decisions, yet remains honest in the face of scientific knowledge; indeed, that celebrates that knowledge alongside other human insights and inspirations. The scientific enterprise as I have presented it to you today may not return human beings to the centre of the universe, it may reject the notion of miracles other than the miracle of nature itself, but it doesn't make human beings irrelevant either. A universe in which the emergence of life and consciousness is seen, not as a freak set of events, but fundamental to its lawlike workings, is a universe that can truly be called our home.

I believe that mainstream science, if we are brave enough to embrace it, offers the most reliable path to knowledge about the physical world. I am certainly not saying that scientists are infallible, nor am I suggesting that science should be turned into a latter-day religion. But I do think that if religion is to make real progress, as Sir John Templeton so passionately advocates, it cannot ignore the scientific culture; nor should it be afraid to do so, for as I have argued, science reveals just what a marvel the universe is.

Sir John Templeton recognizes that if religion is to progress it must confront modern scientific thought. Many religious leaders also accept this, and over the years I have enjoyed fruitful discussions on science and religion with theologians of varying persuasions - behind closed doors. What has most impressed me about my encounters with these theologians has

been their open-mindedness and willingness to accept the conclusions of modern science. While the interpretation of the scientific account of the world may be contentious, there is considerable consensus on the scientific facts themselves. Basic notions like the big bang theory, the origin of life and consciousness by natural physical processes, and Darwinian evolution, seem to cause these theologians little difficulty.

Yet among the general population there is a widespread belief that science and theology are for ever at loggerheads, that every scientific discovery pushes God further and further out of the picture. It is clear that many religious people still cling to an image of a God-of-the-gaps, a cosmic magician invoked to explain all those mysteries about nature that currently have the scientists stumped. It is a dangerous position, for as science advances, so the God-of-the-gaps retreats, perhaps to be pushed off the edge of space and time altogether, and into redundancy.

The position I have presented to you today is radically different. It is one that regards the universe, not as the plaything of a capricious Deity, but as a coherent, rational, elegant and harmonious expression of a deep and purposeful meaning. I believe the time has now come for those theologians who share this vision to join me and my scientific colleagues to take the message to the people!

Chapter XIV

The question of the existence of God in the book of Stephen Hawking: A brief history of time

Alfred Driessen
Department of Applied Physics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

Abstract

The continuing interest in the book of S. Hawking "A Brief History of Time" makes a philosophical evaluation of the content highly desirable. As will be shown, the genre of this work can be identified as a speciality in philosophy, namely the proof of the existence of God. In this study an attempt is given to unveil the philosophical concepts and steps that lead to the final conclusions, without discussing in detail the remarkable review of modern physical theories. In order to clarify these concepts, the classical Aristotelian-Thomistic proof of the existence of God is presented and compared with Hawking's approach. For his argumentation he uses a concept of causality, which in contrast to the classical philosophy neglects completely an ontological dependence and is reduced to only temporal aspects. On the basis of this temporal causality and modern physical theories and speculations, Hawking arrives at his conclusions about a very restricted role of a possible creator. It is shown, that neither from the philosophical nor the scientific view his conclusions about the existence of God are strictly convincing, a position Hawking himself seems to be aware of.

1. Introduction

In 1988 Stephen Hawking, a mathematician and physicist, published a book[1] for the broader public, which soon after appearance became a best-seller. It was translated in more than 20 languages and, in parallel, a series of extended interviews were asked and given for important newspapers and magazines in many countries. Meanwhile more than 7 years have passed, and several studies have dealt with the physical[2] and philosophical[3],[4],[5] aspects treated in this book. In the following, a outline of the book and an analysis of the philosophical elements is given in the light of the metaphysics of Aristotle and Aquinas. The focus hereby is laid on the aspects relevant to the question of the existence of God.

In a first section a summary of the book will be presented based mainly on quotations of the book. The selections of the quotations of course are already a kind of comment, but in addition to this, explicit remarks are given which help to arrive to the conclusions of the present paper. In the following section the philosophical genre of this book will be identified: a speciality in the philosophical field, namely the proof of the existence of God. For comparison the view of the Aristotelian-Thomistic philosophical tradition will shortly be presented. Finally, in the last section, a discussion will be given, where the mutual relevance of the ideas of Hawking and the Aristotelian-Thomistic philosophy will be studied.

2. Summary of the book

The ***introduction*** by Carl Sagan already gives an important key for the understanding of the book. He writes[6]: *This is also a book about God.....The word God fills these pages....Hawking is attempting ....to understand the mind of God....the conclusion of the effort...: a universe with no edge in space, no beginning or end in time, and nothing for a Creator to do.* One clearly should have in mind, that besides presenting a popularisation of modern physical pictures about the universe, Hawking is entering the field of philosophy and eventually theology.

Hawking starts his book with a chapter called: ***Our picture of the universe***. He gives a short historical description of the different pictures of the universe. About the beginning of the universe he says[7]: *One argument for such a beginning was the feeling that it was necessary to have "First Cause" to explain the existence of the universe.* He adds then immediately an explanation: *Within the universe, you always explained one event as being caused by some earlier event.* It is remarkable, that in this description of cause the time-aspect is

---

[1] S. Hawking, *A brief history of time, from the big bang to black holes*, Bantam Books, New York, 1988.
[2] M. Sachs, *On Hawking's "A Brief History of Time" and the Present State of Physics*, Brit. J. Phil. Sci. 44(3) (1993), pp 543-547.
[3] W.L. Craig, *'What Place, then, for a creator?': Hawking on God and Creation*, Brit. J. Phil. Sci. (1990), pp 473-491.
[4] R.J. Deltete, *Hawking on God and Creation,* Zygon 28(4) (1993) pp 485-506.
[5] A. Driessen, *The question of the existence of God in the book of Stephen Hawking "A brief history of time"*, Acta Philosphica, **4**, (1995), pp. 83-93.
[6] ref. 1, p. X.
[7] ref. 1, p. 7.

essential, i.e. he neglects ontological causes, which are essential in the classical philosophy, and especially in metaphysics. On the basis of his definition of cause as working only from out the past he comes some two pages later to the first important conclusion about the role of a creator in a universe with a big bang[8]: *An expanding universe does not preclude a creator, but it does place limits on when he might have carried out his job.*

Interesting are his ideas about the fields of science, philosophy (metaphysics) and religion[9]: *Some people feel that science should be concerned with only the first part* (the laws that tell us how the universe changes with time)*; they regard the question of the initial situation as a matter for metaphysics or religion.* For him metaphysics and religion seem to be quite close to each other, and distant to science. Hawking is ending the first chapter with some remarks about a complete unified theory and concludes[10]: *And our goal is nothing less than a complete description of the universe we live in.* This remark gives rise to an important question: is a physicist able even with a perfect developed theory to give a complete description of the universe? What to say about the role of biology, medicine, sociology or even philosophy? Are they all included in physics?

In chapter 2 about ***Space and Time*** a history of science is given from the Greek up to the work of Penrose and Hawking which showed, that *Einstein's general theory of relativity implied that the universe must have a beginning and, possibly, an end.*[11] After speaking in ***The Expanding Universe*** about the understanding of the universe from general relativity up to the state of knowledge in 1970, he points out the necessity of quantum mechanics for a next step in a deeper understanding. In chapter 4 ***The Uncertainty Principle*** he explains some basic principles of quantum mechanics, especially the uncertainty principle, which he shows to be essential to avoid that *classical general relativity, by predicting points of infinite density, predicts its downfall.*[12] He remarks, that with the uncertainty principle a non-deterministic law in physics has been found. This has consequences also for the role of God, as *scientific determinism ...infringed God's freedom to intervene in the world.*[13]

In ***Elementary Particles and the Forces of Nature***, chapter 5, he describes, starting from the Greek atomists the way to an overall theory of the four basic forces: gravitational, electromagnetic, weak nuclear and strong nuclear force. Up to now, there is only a partial result, the grand unified theory (GUT), including electromagnetic, weak nuclear forces and strong nuclear forces. Hawking comments[14]: *This title is rather an exaggeration: the resultant theories are not at all that grand, nor are they fully unified, as they do not include gravity.*

Black Holes and ***Black Holes ain't so Black*** is treated in chapter 6 and 7. He first gives a historical overview, including the work of Penrose and himself, and shows, how general relativity gives rise to singularities, where the concept of space and time are seriously altered. A singularity, a concept taken from mathematical theories, denominates a special point or region in a function, where one has to divide by zero and where the function consequently is undefined.

---

[8]ref. 1, p. 9.
[9]ref. 1, p.11.
[10]ref. 1, p. 13.
[11]ref. 1, p. 34.
[12]ref. 1, p. 61
[13]ref. 1, p. 53.
[14]ref. 1, p. 74.

The functions used in the theory of general relativity can mathematically be considered as having a singularity, when they are applied to black holes. Later Hawking will speak about a second similar singularity, when he treats the big bang, the among physicists in general accepted starting point of the universe. For Hawking the concept of singularity is central in his reasoning. For within a singularity the known mathematical description of the physical reality breaks down, i.e. there is neither a deterministic nor a statistical description of the events of those regions. In this chapter Hawking is able to demonstrate quite convincingly, that the singularity in the center of a black hole can be circumvented, when one combines general relativity with the uncertainty principle. This seems to be the first combination of the two great theories of modern physics, general relativity and quantum mechanics, resulting in an unexpected and at a first sight paradoxical conclusion: black holes are not so black, i.e. they may emit energy or matter in form of radiation. Hawking considers this result as a *glimpse* of what a fully unified theory would bring in future. It is important to note that with this new approach, Hawking manages to get rid of the first class of singularities that are connected to black holes.

In the following chapter ***The Origin and Fate of the Universe*** Hawking tackles the problem of the second class of singularities, the big bang and eventually the big crunch. Unlike black holes, which are thought to be superabundant in the universe, the two species of the second class are unique; the big bang is considered as the starting point of the universe including time and all physical laws, the big crunch then is the final collapse with the end of time and the end of all known physical laws. After explaining in short the physical ideas connected to the big bang and big crunch, Hawking considers the philosophical implications of the big bang singularity: *space-time would have a boundary - a beginning at the big bang.*[15] He then makes a statement about the laws of sciences in accordance with his restricted concept of causality, i.e. only temporary causality: *These laws may have originally been decreed by God, but it appears that he has since left the universe to evolve according to them and does not now intervene in it.*[16] As one can see, only in the beginning, at the big bang singularity, a decisive role for God is possible.

In the next pages the anthropic principle[17] is introduced and different models of the development of the universe are presented. Hawking speculates about these models based on the general theory of relativity and quantum mechanics and ends with what he calls a *proposal* [18]: space and time could be finite without boundary or singularity[19], at least if one introduces the concept of imaginary time. Within his logic of the reduced concept of causality this proposal has *profound implications for the role of God in the affairs of the universe.*[20] These implications, which are the central point of his book, have already been presented in the introduction by Sagan, and is worthwhile to quote once again: *So long as the universe had a beginning, we could suppose it had a creator. But if the universe is really completely self-contained, having no*

---

[15]ref 1, p. 122.
[16]ref 1, p122.
[17]The anthropic principle has been introduced by Hawking and B. Carter, and can be summarized: *we see the universe the way it is because we exist.*
[18]ref 1, p 136.
[19] There is a certain similarity with the Gödel universe, where the past and the future is a loop, see K. Gödel, Collected works, volume II, Publications 1938-1974, edited by S. Feferman et al., Oxford University Press, New York, 1990, pp. 189-216.
See also: G.C. Chaitin, *Number and Randomness, algorithmic information theory - new results on the foundations of mathematics*, this volume, chapter II
[20]ref. 1 p.140.

boundary or edge, it would have neither beginning nor end: it would simple be. What place, then, for a creator?[21] Some pages earlier Hawking already used an expression for the universe, *it would just BE*[22], which resembles the name of God in the Bible[23]: Jahwe (I am who is). It should also be noted that Hawking in the development of his *proposal* is quite conscious of the speculative character of his argumentation: all statements, like the one just given, are expressed in terms of *would, could, if, may,* etc.

In ***The Arrow of Time*** Hawking considers the direction time passes, from past via the present to the future, this direction he calls the *arrow of time*. He considers three types of arrows: the thermodynamic arrow, related to entropy, i.e. the amount of disorder in a system, the psychological arrow, which relates to the human memory, as we only remember the past, and the cosmological arrow, which is the direction of time in which the universe is expanding. In the light of the "no boundary proposal" of the universe and the anthropic principle he shows the relation between the different arrows. His argumentation needs further philosophical study, because it is not clear whether the analogy between a computer memory and the human brain is strong enough to draw conclusions regarding the psychological arrow.

The Unification of Physics is the last chapter before the conclusion. Already the great aim of physics has been mentioned, the unification of the four basic forces in one single theory. But even with a complete unified theory, there are two reasons, why a physicist can't predict events in general: there is the uncertainty principle, where there *is nothing we can do to get around that*[24]. And there is another more practical inherent difficulty to solve exactly the equations given by the theory. It is, e.g. not possible to solve exactly the motion of three bodies in Newton's theory of gravity. Being conscious of these fundamental restrictions Hawking nevertheless puts an aim quite ambitious for a physicist: *our goal is a complete understanding of the events around us, and of our own existence*[25]

The last chapter ***Conclusion*** summarises the way Hawking had led through the exciting area of modern physics. But now he draws conclusions, which are presented like different pieces of a mosaic, and which in general go far beyond physics into the realm of philosophy and eventually theology. About the situation before the theories of gravity and quantum mechanics are united, he writes: *At the big bang and other singularities, all the laws would have broken down, so God would still have had complete freedom to choose what happened and how the universe began*[26]. According to Hawking, however, with the new still not available unified theory and the no boundary *proposal* the situation would have changed largely: *If the no boundary proposal is correct, he (*God*) had no freedom at all to choose initial conditions*[27].

After having made these statements, all in a conditional form, Hawking brings new pieces of thoughts into his mosaic of fundamental ideas regarding the universe, worthwhile to be quoted: *Even if there is only one possible unified theory, it is just a set of rules and equations. What is it*

[21]ref. 1 p. 140 f.
[22]ref. 1 p.136
[23]Ex. 3,15
[24]ref. 1 p. 168.
[25]ref. 1 p 169.
[26]ref. 1. p 173.
[27]ref. 1, p.174.

that breathes fire into the equations and makes a universe for them to describe?[28] With this almost lyric sentence Hawking expresses what in the Aristotelian-Thomistic metaphysics one could describe in terms of *causa formalis* and *causa efficiens.* The *causa formalis* is necessary, but not sufficient to cause the total effect. Besides this the *causa efficiens* is needed, who gives a set of ideas and 'formulas' an implementation in reality.

In the very same page Hawking invites the philosophers, the people who in contrast to scientists ask *why* instead of *what* the universe is, to *keep up with the advance of scientific theories*[29]. He hopes that after the discovery of a complete theory a new area will come: *Then we shall all, philosophers, scientists, and just ordinary people, be able to take part in the discussion of the question of why it is that we and the universe exist. If we find the answer to that, it would be the ultimate triumph of human reason - for then we would know the mind of God.*[30]

3. The central question: does God exist?

From the foregoing it may be clear, that although the book is written by a physicist, it deals with a philosophical subject, a specialised theme of metaphysics, namely the proof of the existence of God. Obviously the genre of the book is not affected by the positive or negative answer to the central question: "Does God exist?" Contemplating the two-three thousand years of history of philosophy from the ancient Greeks up to now, one observes a continuous interest in this central question. All the tools available to philosophers and scientists as, e.g. logic, metaphysics, history of philosophy and science itself, have been applied to clarify as much as possible the different aspects. Hawking as a scientist gives an important contribution to the scientific part of the question; regarding the philosophical aspects, he uses only a reduced selection of the knowledge until now obtained. The most comprehensive discussion of the proofs of the existence of God is probably given in the work of Aquinas, who resumed the different demonstrations in the famous five via's[31]. It is not the place here, to discuss in detail his argumentation, but a summary of the first way[32], which Aquinas called the *first and most obvious way*, may give a proof of the strength of the philosophical argumentation.

---

[28]ref.1, p. 174.
[29]ref. 1, p. 174.
[30]ref. 1, p 175.
[31]Thomas Aquinas, *Summa Theologiae*,I, q.2, a.3.
[32]For the interested reader an English translation of the first way is given below (from St Thomas Aquinas, *Summa Theologiae*, Latin text and English translation, Blackfriars 1964, Eyre&Spottiswoode, London). In contrast to our quotations in the text, which follow closely the Latin of Aquinas, this translation uses more the concepts of today's English. The main difference is the translation of *moveri*, being moved, which is translated as being in process of change.
The first and most obvious way is based on change (ex parte motus). Some things in the world are certainly in process of change: this we plainly see. Now anything in process of change is being changed by something else. This is so because it is characteristic to things in process of change that they do not have the perfection towards which they move, though able to have it; whereas it is characteristic of something causing change to have that perfection already. For to cause change is to bring into being what was previously only able to be, and this can only be done by something that already is: thus fire, which is actually hot, causes wood, which is able to be hot, to become actually hot, and in this way causes changes in the wood. Now the same thing cannot at the same time be both actually x and potentially x, though it can be actually x and potentially y: the actually hot cannot at the same time be potentially hot, though it can be potentially cold. Consequently, a thing in process of change cannot itself cause that same change; it cannot change itself. Of necessity therefore anything in process of change is being changed by something

3.1. The first way of Thomas Aquinas

In the first way Thomas uses ideas that already can be found with Plato[33], Aristotle[34] and Averoes. He starts from the common experience, that *it is sure, that in this world some things move*. Then he puts his first thesis: *all what moves, is moved by some other*. The proof of it is shortly given by an analysis of movement as being brought from being *in potentia* to being *in actu*, i.e. brought from being potentially in a certain state to being actually in that state. He comes to the conclusion: *It is therefore impossible, that something in the same aspect and in the same way brings into movement as well is moved or moves itself.* The next step in his argumentation is the thesis: *If the mover himself is moving, then he also has to be moved by some other*, which in fact is a logical extension of the first thesis, and shows that there is a cascade of movers which in turn are moved by other movers. Aquinas now states, that there can be no infinite chain of movers and moved, as otherwise there would be no first mover, and consequently nothing, which could start the movement. His conclusion therefore is, that there must be a first mover, which is not moved by anything. And he ends his proof with: *and this is what everybody understands by God.*

About this first via some remarks should be given. Speaking about moving Aquinas considers all kind of changes, like getting hot, changing of color, change of position etc. In his second way, a similar proof is given, but then one should read instead of moved: caused by. It is of extreme importance to note that in the via's moved or caused by is always moved or caused by *per se*, i.e. if the mover or cause stops to move or cause, the effect also stops. With other words, the mover or the cause is acting in the present time. That means that also the cascade of movers and moved or causes and caused is completely in the present. The following example of a cascade or hierarchy of movers, which in a shortened way Aquinas already has mentioned in the explanation of the first way, may be a good illustration. It is the case of man, who is moving a ball along a certain trajectory, let say a circle: The ball is moved by a stick. The stick is moved by a hand. The hand is moved by a set of muscles. The muscles are moved by neural commands. The neural commands are moved by the brains. The brains are moved by the will, etc. The exact identification of the different levels in this cascade of movers may be a point of discussion, but one sees clearly that all movers are acting simultaneously and are acting per se, i.e. if one of the movers fails, there is no effect, in this case the ball would not follow the original trajectory.

The proof of Aquinas is quite subtle and looses its strength if one introduces even minor changes in the different steps. In the foregoing example, one could consider a ball shot by a soccer player, once the direct contact between shoe and ball is broken, the ball follows a trajectory that could be the intended one, but also could be drastically changed or even stopped by other movers or causes, like wind or a keeper's hand. In the case of movers as presented in this last example Aquinas would never conclude that there must be necessarily a finite cascade or a first mover.

---

else. Moreover, this something else, if in process of change, is itself being changed by yet another thing; and this last by another. Now we must stop somewhere, otherwise there will be no first cause of the change, and, as a result, no subsequent causes. For it is only when acted upon by the first cause that the intermediate causes will produce the change: if the hand does not move the stick, the stick will not move anything else. Hence one is bound to arrive at some first cause of change not itself being changed by anything, and this is what everybody understands by God.

[33]Plato, *Phaedrus*

[34]Aristotle, *Physica* VIII

Aquinas ends his proof with: *and this is what everybody understands by God*. One has to realise, that all of his reasoning is still in the field of philosophy and not theology. Only by starting from the daily experience of the movement of material things and logical thinking, he arrives at the necessity of something, which is the first mover or, in the second via, the first cause. Having obtained this result, it seems that he looks around to see, where to find this first mover. And the results of this exploration: the first mover is just that, what people in general understand by God. The first mover, a pure philosophical concept, can be identified with God, the central theme in theology. Notice however that Aquinas is not evoking Revelation to define God, but only the general understanding of people. In the Introduction [35] an argument is given which may support the identification of the first mover with that people understand by God.

3.2. Hawking and the classical proof of the existence of God

It is useful, to compare the different steps, Hawking is making in his attempt to clarify the question of the existence of God, with the classical proof of the Aristotelian-Thomistic philosophy, as has been presented just before in a summarised form. Hawking starts by using a changed concept of causality. We already quoted his explanation of the meaning of being caused, which for him is exclusively causality in time: *Within the universe, you always explained one event as being caused by some earlier event...*[36]. The exclusive use of this kind of temporal causality[37], Aquinas explicitly excludes for his proof[38]. It may be clear that with the reduced concept of causality as used by Hawking, the original proof is strongly weakened.

Applying the temporal concept of causality, Hawking expects an intervention of a possible creator or God only in the beginning of the universe, as already has been shown by the quotations in section 2. As long as there is a beginning, which he identifies with the big bang singularity, there would be a role for a creator. If, however, the physical necessity of a beginning has been eliminated, the crucial question comes: *What place then, for a creator?*[39]. Hawking therefore comes in his main line of reasoning with the temporal concept of causality to the conclusion, that there is no logical need to assume the existence of a creator, as long as one only considers the universe of the physicists, which, as we have seen before, includes in his view the material world inclusive the human life. Nevertheless, he himself is convinced, that something is missing in his reasoning. Not only the question *what*, but also the question *why* should be asked: *Why does the universe go to all the bother of existing?* This question has not been answered yet, as *up to now, most scientists have been too occupied with the development of new theories that describe* ***what*** *the universe is to ask the question* ***why***[40].

4. Discussion

[35] A. Driessen and A. Suarez, *Introduction*, this volume.
[36] ref 1, p. 7.
[37] This corresponds also to the Kantian view of causality (see A. Driessen and A. Suarez *Introduction*, this volume, and A. Suarez, *Nonlocal phenomena*, this volume, chapter X).
[38] L. Elders, *De Metafysica van St. Thomas van Aquino in historisch perspectief, II: Filosofische godsleer,* Uitgeverij Tabor, Brugge, 1987, p.150, see also Thomas Aquinas, *In liber II Physicorum*, lectio 6, n 195.
[39] ref. 1, p. 141.
[40] ref. 1 p. 174.

After having gone through the book of Hawking and presented the proof of the existence of God in the Aristotelian-Thomistic philosophy, one may want to look for the mutual implications, at least if there are any. Scientists, like philosophers, have their own working field, and the methods in science are quite different from those in philosophy. There is however an overlap: in the object, as scientists are dealing with the material reality as being material and philosophers with the same reality, the material and beyond that also with the immaterial reality. And of course, always, the person, who is doing science, in other aspects is thinking as a philosopher, and often also the reverse is valid.

One may therefore say, there is an interaction between science and philosophy, and even between science and theology. Hawking himself gives an example, when introducing the Heisenberg uncertainty principle and discussing determinism: *The doctrine of scientific determinism was strongly resisted by many people, who felt that it infringed God's freedom to intervene in the world, but it remained the standard assumption of science until the early years of the century.*[41] If that theory of total determinism in the physical world would have been proven to be true, then God's intervention in the material world would be bound to deterministic laws, and regarding human freedom, one could only consider at most pure internal decisions, which would not affect any physical reality.

If one now considers the main line of argumentation of Hawking, one is at first confronted with his restricted concept of temporal causality, which we have shown is contrary to the one used in classical philosophy. Nevertheless, even if one accepts this concept, his "proof" of non-necessity of a creator is not supported by physical evidence, but of ideas with a highly speculative character. He starts with theories, like the of relativity and quantum mechanics, which are shown to be valid by thousands of experimental verifications and which are accepted by practically all physicist. When discussing big bang, black holes, etc., there the scientific evidence is much weaker, and the ideas have a more hypothetical nature. Introducing, however, imaginary time and the no boundary proposal, Hawking himself is conscious of the speculative nature of his reasoning. One should be aware, if the scientist Hawking calls his ideas a proposal and admits that is far from being proven, then a philosopher (say Hawking or any other) may not use this argumentation as a decisive proof for the existence or non-existence of a creator. And if one reads the remarks of Hawking in his last chapter (see quotation, ref. 23), he seems to be aware of it.

There is one very interesting question left. The title of the book *A brief history of time* promises worthwhile and perhaps new ideas about time. A widely discussed question in philosophy is, whether the universe is eternal, and - this is not the same question - whether the universe is created. Science was not able to give an answer. With the introduction of the big bang hypothesis, based on the work of Penrose and Hawking, many considered this as the proof, that there was a beginning and therefore a creation. With the no-boundary proposal Hawking has not proven, that the universe is eternal, simple being. What he has shown, is that for a scientist at the top of the knowledge about the universe, the older standard big bang hypothesis is not necessarily true, and that the idea of a universe without beginning can not be rejected on purely scientific reasons. It is therefore still a matter of discussion. Coming back to Aquinas, one finds the problem of creation of the universe in time or creation from eternity[42]. His conclusion is, that it is

[41]ref. 1 p. 53.
[42]for a discussion, see, e.g. L.J. Elders, *De natuurfilosofie van Sint Thomas van Aquino*, Uitgeverij Tabor, Brugge, 1990, p. 138 ff.

possible to demonstrate the ontological dependence of the universe from God, but not the beginning in time. Only additional information, as is given in theology by revelation, could give an answer[43]. For Aquinas evidently the answer to this question is not relevant for the demonstration of his 5 via's. This has an enormous impact on the philosophical value of the input of science as has been delivered by Hawking. The main line of his reasoning does not affect the philosophical proof of the existence or non-existence of a creator, at least in the philosophy of Aquinas. What then is the value? Not a small one, one may say, namely bringing people to think and stimulate them to ask why.

[43] In the Jewish-Christian tradition this information is found in Gen. 1.1: In the beginning.....

Chapter XV

Final remarks: Becoming aware of our fundamental limits in knowing and doing, implications for the question of the existence of God.

Alfred Driessen
Department of Applied Physics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
E-mail: A.Driessen@tn.utwente.nl

and

Antoine Suarez
Institute for Interdisciplinary Studies, P.O. Box 304, 8044 Zurich, Switzerland, E-mail: suarez@leman.ch

Having arrived at the end of this book the reader may not be completely satisfied. He has read about mathematics, about physics and about the way others think about science. But, so what? Why has this wide-ranging material been brought together? Are there firm conclusions regarding what the title promises: the existence of God? A comparison already mentioned in the Introduction could help answer these questions. Consider the parts of a puzzle, representing in their totality a masterpiece of art. The parts on their own are only small, colored pieces of cardboard. In this book we tried to deliver some important pieces of that puzzle, indicated by the set of questions. The reader should decide whether the pieces result in more than a random distribution of colored pieces. Being optimistic the editors and probably others, will see the vague but certainly discernible contours of a reality beyond the things we can touch, see, hear, smell or taste. In any case, however, we are aware that we have not reached completeness in solving our puzzle; some pieces are still missing that further reflection should work out.

The first piece of the puzzle is supplied by mathematics. Turing and Gödel have proved that in any formal (arithmetic) system one can formulate true statements, which are undecidable within the formal system in question. As a consequence, each formal arithmetic system lacks completeness. This means that man never can explore the full richness of mathematics, not because of limitations in our time or ability, but because of fundamental limits always present in any non-trivial formal system. The undecidability and lack of completeness in formal systems also have consequences for the origin of mathematical truth. The access of man to mathematical truth is fundamentally incomplete. Mathematical truth, therefore, cannot be an exclusive construction of the human mind.

The second piece of the puzzle, somewhat related to the foregoing, arises from the experience of mathematicians of down the ages. There is information, knowledge of abstract relations, which can be ‘discovered’ by man, but which is not a product of the human mind. In a certain sense the information discovered in mathematics at a certain moment has always existed and will exist forever. It is reasonable to relate this information to a reality, because information is something, it is not nothing. The question then arises: who, or which principle, supports this information? The human mind, with its limited access to mathematical information, surely cannot be considered the only candidate.

Coming now to physics a new set of pieces of the puzzle can be found. Since Laplace, and even before his time, the ideal of physics has been to find or derive a set of equations, which allow a complete description of the reality accessible by the methods of physics. In other words, physicists try to find a unique correspondence between physical phenomena and representations in a formal system. This formal system would permit, at least in principle, the calculation of all physical events. It is obvious that enormous progress has been made in physics that has resulted in a quantitative description of many physical effects. Nevertheless, the equations of physics are at best a well-developed formal system. And regarding these systems Gödel and Turing have proved that these can never be complete. There will therefore be physical events that cannot be adequately described by the formal system in question. It is therefore an illusion to hope that physical reality can be perfectly matched with a formal system, and therefore that physics can describe physical reality completely. Physical reality always will be more than a set of equations.

The foregoing was an argument about an inherent shortcoming of any physical theory based on a mathematical description of reality. And the "weak point" was in mathematics rather than in physics itself. One could argue that physics could somehow circumvent this problem. The questions now arise whether there are fundamental limits to physics itself and whether physics can be considered to give a complete description of the phenomena. The Einstein-Podolsky-Rosen paradox one could consider as the next part of the puzzle we are trying to solve. The EPR paradox - a two-particle gedanken experiment - and the consequent work by Bell, brings the physicist face to face with nonlocal correlations. And nonlocality contradicts all traditional physical theories. But things are more subtle. When we give a theoretical description of EPR-like experiments using the relevant physical theory (in this case Quantum Mechanics) we come up with results that are in complete agreement with experiment. We should remark, however, that these theoretical predictions are presented as probabilities that can be verified only by a large number of events. This probabilistic description therefore seems to be a correct approach. Only if one considers the single event, has one to assume unobservable causes which result in nonlocality.

The apparent contradiction could be solved if one assumes a fundamental incompleteness in physical theories. In some cases phenomena - or observable causes- which obey physical laws, are not the only actors that take part in the realization of the event. Other, unobservable actors that are not in contradiction with the statistical nature of our physical theories seem to affect the single physical effect. Consider for instance a quantum experiment in which the two output ports of a beam splitter are monitored by detectors D1 and D2. The fact that one of two detectors clicks may be explained by observable causes alone. But the particular alternative that D1 clicks and D2 remains at rest, or conversely, cannot be explained by observable causes alone. In general it can be stated that phenomena cannot be explained by a temporal chain of causes consisting of other phenomena.

Also single particle events in nature, like a decay of a radioactive atom, seem to demonstrate the lack of completeness in physical theories like Quantum Mechanics. We can predict the decay rate with high precision, but we can say nothing about a single atom, when it will decay. It is even less known why it decays in that certain moment. The philosophically not so satisfying answer is that we may not ask these questions or that it happens by chance. In the case of the EPR experiment (an essential two-particle experiment) one has to give up the "chance" explanation. There is experimental evidence of a correlation between the two particles even in the case of a single event, which can not be the result of local hidden variables (Bell). Could it not be possible that also in single particle events the "chance" theory

should be given up with the consequence of accepting the incompleteness of our physical theories?

With the pieces of the puzzle presented above, are we now in a position to come to some conclusions about the total picture? The *first* could be that man, with his formal approach in science, precisely due to this approach misses part of the information contained in physical reality. The *second* is that man will never have complete control over nature because in technology he uses exclusively physical observable causality. It seems that observable causes produce necessarily the expected effect but are not sufficient causes for a particular event.

Is this a question of our present ignorance? Will a future generation of mathematicians and scientists circumvent these difficulties? Or are we discovering traces of a powerful actor or acting principle who, as the support of information, has intelligence and is causing events in reality according to the physical laws? Certainly, we will not return to former primitive times when people appealed to a supernatural power in order to explain why the sun appears to go round the earth. But neither can we claim that we do not need such a being any more, simply because with modern science we now know considerably more about the reason why. And if we are obliged to give up the position that science will enable man to master all mysteries, the road to avoid absurdity may be the road that leads to a transcendent being. This is our *third* tentative conclusion.

But does this not mean to return to the *God-of-the-gaps, a cosmic magician invoked to explain all those mysteries about nature that currently have the scientists stumped,* as P. Davies[1] expresses? Is this not a dangerous position, as well? Probably not, because we have evidence that the gaps in our knowledge and in our ability to determine events are structural. It is not a question of knowing or doing more or less. In a certain sense nature can be considered to be a miracle, as Davies himself states, because natural phenomena have unobservable causes beyond the reach of any human power. There always will be unsolved mathematical problems, and we have to live with nonlocality.

Finally, we would like to comment on an apparently paradoxical situation. Modern science which accepts man's fundamental limits in knowing and doing, comes out to be more efficient than the science which believed that man will some day be able to know all. Now, when we have given up the postulate of absolute predictability, we predict and control better as ever before. We become capable of doing more and more marvelous things just because we have accepted that we will never be able to do everything.

Looking back now at the present book, we surely have not given a complete answer to all of the problems related to the title. Some results have been presented and conclusions have been given. One should bear in mind, however, that science is not the only access to reality. The rich world of human feelings and thoughts as expressed, for example, in literature, art, humanities and also in conversations in daily life, provides alternative routes to reality in all its dimensions. Could it be that the intelligent and powerful actor, whose outline, after much effort, seems becoming visible to the scientist, is identical with what people call God?

[1] P. Davies, *Physics and the mind of God*, this book, chapter XIII.

Mathematical Undecidability, Quantum Nonlocality and the Question of the Existence of God

Curriculum Vitae of Authors

F.T. Arecchi

After his Ph.D. in Electrical Engineering at the Politecnico di Milano in 1957, F.T. Arecchi worked on laser physics at Stanford (Cal.), CISE Milan, IBM Ruschlikon (CH) and the MIT Cambridge (Mass.). He held a Chair of Physics at the University of Pavia (1970-1977) and thereafter in Florence. Starting in 1975 he is President of the National Institute of Optics in Florence. He has worked and published extensively in the field of laser physics. His present interest is chaos in lasers and optical systems, pattern formation and competition in nonequilibrium systems and nonlinear dynamics. In addition he works on Epistemology.

J.S. Bell

John S. Bell, was born in Belfast in 1928, and died unexpectedly in Geneva on 1 October 1990. He received a B.Sc. in Experimental Physics in 1948 and one in Mathematical Physics a year later from Queens University (Belfast). From 1949 to 1960 he worked at the Atomic Energy Research Establishment (AERE) in Malvern and Harwell (Great Britain). In 1956 he received his Ph.D. at the University of Birmingham. Since 1960 he had been a pivotal figure in the Theory Division at CERN, Geneva. In 1964 he established the “Bell’s” inequalities that allow to distinguish experimentally between local realistic theories and quantum mechanics. The collection of his scientific work in *Speakable and unspeakable in quantum mechanics* (Cambridge University Press, 1987) belongs to the most important publications on the foundations of quantum mechanics. In the last years before his death he actively collaborated with the Center for Quantum Philosophy, promoting seminars and conferences on the philosophical consequences of nonlocality.

His contributions were recognized by the Dirac Medal of Physics, bestowed by the Institute of Physics (UK), and the Heineman Prize of the American Physical Society.

F. Cacace

Filippo Cacace received the M.S. degree in electronic engineering in 1988 and the Ph.D. degree in computer science in 1991, both from the Politecnico di Milano. He is currently working in the Dipartimento di Informazione e Sistemi of the Universita' di Napoli 'Federico II' , where his research interests include deductive and object-oriented database systems, parallel algorithms, and logic programming. He is the author of several articles published in international journals and conference proceedings, and he is the co-author of the book *Advanced Relational Programming* (Kluwer, 1996).

G.J. Chaitin

Gregory Chaitin is at the IBM Watson Research Center in New York. In the mid 1960s, when he was a teenager, he created algorithmic information theory, which combines, among other elements, Shannon's information theory and Turing's theory of computability. In the three decades since then he has been the principal architect of the theory. Among his contributions are the definition of a random sequence via algorithmic incompressibility, and his information-theoretic approach to Gödel's incompleteness theorem. His work on Hilbert's 10th problem has shown that in a sense there is randomness in arithmetic, in other words, that God not only plays dice in quantum mechanics and nonlinear dynamics, but even in elementary number theory. He is the author of four books: "Algorithmic Information Theory" published by Cambridge University Press; "Information, Randomness & Incompleteness" and

"Information-Theoretic Incompleteness," both published by World Scientific; and "The Limits of Mathematics," published electronically.

P. Davies

Paul Davies is Professor of Natural Philosophy at the University of Adelaide. He obtained a Ph.D. from the University of London and has held academic appointments at the universities of London, Cambridge and Newcastle-upon-Tyne. He emigrated to Australia in 1990. His research interests are in the field of black holes, cosmology and quantum gravity, and he has published over a hundred specialists papers, as well as several textbooks.

Prof. Davies has achieved an international reputation for his ability to explain the significance of advanced scientific ideas in simple language. He is the author of some twenty books, including *God and the New Physics.* He has also written and presented a number of TV and radio programmes. In 1995 he won the Templeton Prize, the world's largest award for intellectual endeavour, for his work on the deeper meaning of science.

A. Driessen

After his studies at the universities of Cologne, Bonn and Amsterdam he obtained his degree in Physics (1972) and his Ph.D. (1982) with an experimental study on Quantum Solids, both at the University of Amsterdam. After a 5-year period as postdoc in the Physics Department of the Free University of Amsterdam, he became associate professor at the Lightwave Devices Group at the Department of Physics of the University of Twente. In this function he works and publishes mainly in the field of integrated optics for telecommunication and optical nonlinear materials and devices. In addition he is largely interested in philosophical questions with an emphasis on the philosophy of science. He is Senior Member of the IEEE Laser and Electro-Optic Society (LEOS) and Member of the Board of the Institute for Interdisciplinary Studies (Zurich).

A.M. Fox

Dr. Mark Fox received the B.A. and D. Phil. degrees in physics from the University of Oxford in 1982 and 1987 respectively. His doctoral thesis was on the nonlinear optical spectroscopy of semiconductors. In 1986 he became a Junior Research Fellow of Christ Church College, Oxford. He held this post until 1988, when he became a postdoctoral member of the technical staff in the Photonic Switching Device Research Department at AT&T Bell Laboratories, Holmdel, New Jersey. In 1990 he returned to his present position as a Royal Society University Research Fellow in the Department of Physics at Oxford. His current research work is on the optical and electronic properties of low dimensional semiconductor structures, quantum optics, and ultrafast nonlinear optics.

J. Laeuffer

Jacques Laeuffer (1958) received a degree in electrical engineering at the Ecole Nationale Supérieure d'Ingénieurs Electriciens de Grenoble in 1980. He has been working for sixteen years in the area of power electronics, electrical energy conversion and controls. As a research engineer in different companies, he developed inverters and servomechanisms for radar, security power supplies for computers, and x-ray generators for medical imaging. He holds eleven patents. As an "Expert-Engineer" for G.E. Medical Systems, Paris, he is responsible for new developments, relationships with universities, and training courses.

P. Pliska

Born 1962 in Prague (Czechoslovakia). Primary and high schools in Prague, New York, and Zurich (Switzerland). Studies in physics at the Swiss Federal Institute of Technology *(ETH)* in Zurich. 1988-1995 assistant researcher and teaching assistant in the Optics

Laboratory at the Institute of Quantum Electronics of the Swiss Federal Institute of Technology *(ETH)* in Zurich; research on integrated optics, sensors, microtechnology, and related topics. 1994 dissertation (PhD) on „Integrated-Optical Nanomechanical Components". Since 1995 patent attorney in private business.

H.C. Reichel

Hans-Christian Reichel (1945) studied Mathematics and physics at the University of Vienna. After his Ph.D. (1969), Habilitation (1976) and guest professorship (1977-78) in Kassel (D), he is since 1979 professor at the University of Vienna. In this years he visited also Bremen and Oxford for a longer period. He publishes extensively on topology, didactic and philosophy of mathematics. He is author of several schoolbooks on mathematics and editor of two international journals on mathematics.

J.M. Schins

Born 1964 in Sorengo (Switzerland). Primary and High School in Varese (Italy). Graduation 1986 in Physics at the University of Amsterdam (The Netherlands). Dissertation 1992 prepared at FOM-Institute for Atomic and Molecular Physics on photodissociation and metal-induced dissociation of hydrogen molecules. Until 1995 post-doctoral position at the Commisariat à l'Energie Atomique, Saclay, and Ecole Nationale Supérieure de Techniques Avancées, Palaiseau (France) on electronic free-free transitions in intense laser fields using the Auger effect in rare gas atoms. Since 1995 assistant professor at University of Twente (The Netherlands) in applied optics and biophysics.

A. Suarez

Antoine Suarez is at The Institute for Interdisciplinary Studies (Leman Foundation) in Zurich and Geneva. He received his Ph.D (1975) in Natural Sciences from the Swiss Federal Institute of Technology (ETH) in Zurich. His research interests are in the fields of epistemology, theoretical biology and physical nonlocality. He introduced (1992) the principle of *unobservable causality* to explain quantum nonlocality, and established (1993) the biological theorem that the celullar differentiation in the embryo depends on embryonic information alone. He promoted (1989) the Center for Quantum Philosophy in which John Bell actively collaborated. At present he works on an alternative nonlocal theory which assumes many superposition principles, and on experiments to test this theory against conventional quantum mechanics.