Consciousness

Carlo Roselli

1. ***Theories related to the mind-body problem***

Dualistic theories are based on the hypothesis that mind and body are two fundamental kinds of entities or principles. Today there are essentially three main forms of dualism: a) - the dualism of substances (or strong dualism), on the basis of which the body and the mind are different substances, the first material, the second immaterial and characterized by equally immaterial states, in other words a substance that thinks; b) – the dualism of properties, which admits the existence of only one kind of substance which nevertheless possesses both physical and mental fundamental properties, the latter not reducible to physics, as if to say that there are physical entities which also have non-physical properties; c) – the dualism of predicates,[1] according to which the predicates referring to the mental world are not reducible to the predicates of the physical world and are considered indispensable for the description of a causally closed world.

Monism[2] is opposed to dualistic theories, which is articulated in three distinct conceptions: 1), idealist monism (mentalism, or simply idealism), which assumes the existence of only mental reality, where what we call "matter" is nothing else that a manifestation of it; 2), double aspect monism, according to which material and mental phenomena are only two aspects, or attributes, of the same reality; 3) materialist monism (or materialism), which today represents the most widespread philosophical position and which maintains that all phenomena of reality, including mental phenomena, can be traced back to the exclusive behavior of a single substance of a material kind, so that the mental world is held to be a product of material processes. From materialism came physicalism, according to which mental states are only physical states of the brain.

For over fifty years, the dualism of substances has been abandoned by most scholars, who have preferred to engage in the search for a possible material description of mental phenomena.

The reason lies in the fact that strong dualism implies a problem that cannot be tackled on a scientific level. In fact, it is completely incomprehensible how two ontologically distinct substances can influence each other.

Thus the study of the mind and its external manifestations dependent on neurophysiology, or the study of the physical or neural processes that take place in the brain, has been affirming. Its goal is to found the "science of behavior".

2. ***Behaviorism and cognitivism***

Behaviorism initially developed in the United States in the 1940s, and then spread to Europe as well. His investigation is aimed at understanding how neural processes are related to specific behaviors of living beings equipped with a highly evolved perceptual and intellectual apparatus and, therefore, above all of human beings, who are also equipped with a sophisticated language.

In the course of investigations of an experimental nature, the "cognitivist" hypothesis came to impose itself, on the basis of which in humans mental processes responsible for rather complex functions should be triggered, such as for example memory, associations of ideas, the

---

[1] This form of dualism is also called intensional conceptual dualism. Although the predicates referring to the mental and those referring to the physical are linguistically and phenomenologically distinct, it cannot be excluded that there is a correspondence between the extensions of the psychological and physical predicates.

[2] We recall that Spinoza, in his Acosmico, supports the point of view of monism, affirming that mind and body are two aspects of a single necessary, self-sufficient and ontologically indefinable substance. This substance is identified by him in the concept of Deus sive natura, whose attributes are extension and thought, and from it all the phenomena of the universe descend following a deterministic and teleological path. We also recall how Einstein approaches Spinoza's idea of God, sharing his attributes of extension and thought, but distancing himself from his pantheistic (or immanentist) conception. Einstein's God is in fact interpreted as an entity that transcends cosmic reality and yet pervades it everywhere.

ability to access own inner states and their reportability, i.e. the possibility of making them available through verbal reports. All these functions seem to be related to something much more complex than pure and simple behaviors.

Cognitivism thus becomes the new scientific approach to the study of the relationships between the complicated neural activity of the brain and the various cognitive functions. Through an increasingly elaborate experimental activity, many of these functions manage to find a scientific explanation thanks to the use of a constantly evolving technology and the use of particular non-invasive methods, such as Magnetoencephalography, Magnetic Resonance Mapping (MRI, from the French Imagerie par Résonance Magnétique) and Positon Emission Tomography (PET, from the English Positons Emission Tomography).

The acquired results lead scholars to convince themselves that in a more or less distant future the entire variety of neurophysiological and cognitive functions will be explained. However, there is one aspect of mental activity that cognitivism, for a long period of time, has completely kept out of its investigative program: consciousness, that is, first-person experience.

The first proposals aimed at laying the foundations for a science of consciousness date back to just over a couple of decades ago. However, these are extremely ambitious challenges in which several philosophers of mind and neuroscientists are engaged, generally attested on underlying premises in conflict with each other, and only slightly different in some, but for now none of them is able to boast a promising theoretical setting. This is due to the presence of a difficulty, consisting in the search for a correlation between the functional mechanisms triggered by the neural activity of the human brain and the conscious experience, i.e. the phenomenon which allows the owner of that brain to notice specific effects , to suffer them from his exclusive point of view.

The American philosopher Thomas Nagel (1937), in his reflections on the phenomenon of conscious experience, argues that

> it is a widespread phenomenon. It can be observed at different levels of animal life, even if we cannot be sure of it with regard to the simplest organisms [...]. There is no doubt that it exists in innumerable forms totally unimaginable to us, on other planets and in solar systems throughout the universe. But regardless of the variety of forms it might take, the very fact that an organism has conscious experience means, in essence, that it feels like something to be that organism.[3]

Furthermore, Nagel specifies that

> basically an organism experiences conscious mental states if and only if one feels something to be that organism – if the organism feels something to be what it is. We could call all this the subjective character of experience: this remains elusive for the typical reductionist analyzes of the mental developed in recent years [...]. The subjective character of the experience cannot in fact be analyzed in terms of any explanatory system of functional or intentional states, since these states could also be attributed to a robot or an automaton that behaved like men, even without having any subjective experience.[4]

3. ***The nature of consciousness***

To date, there is no theory of consciousness and not even a widely shared outline to pave the way towards a plausible understanding of it, so much so that many scholars of the mental world, some of whom are referred to as "mysterians",[5] declare themselves convinced that an answer to why and how cosmic reality contemplating the phenomenon of conscious experience is absolutely inaccessible to the intellect.

---

[3] T. Nagel, *Che cosa si prova ad essere un pipistrello?* Ed. Castelvecchio, Roma, 2013, cit pp. 7-8

[4] Idem, pp. 8-9 (v. anche sezsone 5 di questo capitolo).

[5] Mysterians are philosophers of mind who preach the amazement of experiencing one's own existence and who regard consciousness as an irreducible phenomenon.

David J. Chalmers (1966), one of the most authoritative philosophers of the mind and author of various essays, has carried out and is still carrying out a careful reflection on the most disparate theories proposed up to now to try to answer the inescapable questions emerging in the field of neurophysiology and neuroscience: what is consciousness? How can consciousness and all phenomena associated with it be immediately and vividly experienced first hand? Could consciousness arise from neural processes taking place in the brain?[6]

Chalmers' first concern consists in isolating the set of "easy problems" connected to the idea of consciousness from the so-called "hard problem". The former are susceptible to a relatively simple scientific explanation, as they can be objectively studied by neuroscience and classified as the performance of neural or computational mechanisms. All in all, these are explanations of specific functions that, according to Chalmers, seem to have to do exclusively with the physical world. On the other hand, the truly difficult problem consists in asking ourselves, as Nagel did before him, why the physical processes taking place in the brain are accompanied by conscious experience, i.e. can they light up, to use my personal expression, with their own light.

I do not deny - says Chalmers - that consciousness originates from the brain. We know, for example, that the subjective experience of vision is closely related to processes in the visual cortex. However, it is the link itself that perplexes us. Surprisingly, subjective experience appears to emerge from a physical process. But we have no idea how or why this happens.[7]

When asked if neuroscience will be able to open a window towards the understanding of consciousness, Chalmers replies that this could happen on one condition only: that what the philosopher Joseph Levine (1952) defined in 1983 as the "explanatory gap" ", that is, that a bridge can be built between physical processes and subjective experience. However, on the basis of his scrupulous arguments, Chalmers realizes that all the experimental methods practiced today by neuroscience and cognitive science, as well as all the most modern theoretical proposals in this direction, ranging from the different forms of dualism (with the exception of dualism of substances) to identity theory (also known as 'physicalism'), functionalism (see section 5) and eliminativist materialism,[8] to one degree or another fail in their task. None of them is in fact able to explain from which property or physical law conscious experience originates.

To hope to fill the explanatory gap will require a new theory based on significant discoveries that could come from the field of neurophysiology or from the study of algorithmic processes or, much more reasonably, from the discovery of something new at the level of quantum physics. Furthermore, the possible solution of some of the well-known mysteries of the latter would be of great help.

The new theory, in the hypothesis that it is prepared, will hardly be subjected to experimental tests and will therefore be based on conjectural procedures. In any case, to hope for its convincing elaboration, it will certainly be necessary to rethink the scientific method adopted so far for the formulation of physical theories.

Therefore, as I believe, one will essentially have to rely on the fecundity of the imagination and count on the possibility of intuiting new fundamental properties, or what Chalmers calls the "additional ingredient". He, believing that conscious experience must depend on physical processes but that this dependence cannot be derived only from physical laws, states that

---

[6] David J. Chalmers, *La mente cosciente* (Milano, Mc Graw-Hill, 1999 – originalmente *The conscious mind*, Oxford-New York, Oxford University Press, 1996).

[7] Chalmers, *L'enigma dell'esperienza conscia*, traduzione italiana del lavoro *The Puzzle of Conscious Mind*, Department of Philosophy, University of Arizona, Tucson, AZ 85721, chalmers@arizona.edu.

[8] Eliminativist materialism or, more simply, eliminativism holds that the mind should be studied like any other physical phenomenon. Its supporters assume that the mind is the product of two essential aspects, behavior and the brain; with this they exclude any sort of derivation of the mind from metaphysics.

The new basic principles postulated by a non-reductionist theory give us the extra ingredient we need to build an explanatory bridge. Of course, by viewing experience as fundamental, there is a sense in which this approach does not tell us why there is experience in the first place. But the same goes for any fundamental theory. Nothing in physics tells us why matter exists in the first place, but we don't consider that a fact that gets in the way of theories of matter. Some elements of the world must be considered fundamental in any scientific theory.[9]

Chalmers, starting from a philosophical position which I will mention later (in section 9), believes that certain criteria adopted to build physical theories, in particular those based on principles of simplicity and elegance and such as to suggest the existence of fundamental laws, can be equally valid for a theory of consciousness. The principles that unify physical and experiential processes will have to be conceived as explanatory fundamental, that is, not derivable from physical laws.

But a theory of experience, although it can benefit from reliable indirect sources, such as the verbal reports of subjects studied with the modern investigation methods of neurophysiopsychology and cognitive science, lacks objective data. It follows that any theory of experience, like all non-empirical theories, will always have a speculative character.

Sharing Chalmers' optimism, I do not exclude at all that, in the future, a well-formulated theory could be so convincing as to be generally shared. However, I would like to point out that any explanation of consciousness should be incorporated into a reformulation of quantum theory or, better still, into a theory of complete unification of the fields. Only on these assumptions could we point to the description of a real Theory of Everything, therefore also inclusive of the conscious experience.

But in this regard I would add a further observation: the possible Theory of Everything could not be such if it limited itself to describing the intrinsic and relational properties of the fundamental processes, since it should also be able to explain, contrary to what Chalmers maintains, what is the real ontological foundation of this procedural reality, the justification of its existence and the modalities of its operation (I will return later to argue on the reasonableness of the supplementary ingredient requested by Chalmers for a theory of consciousness).

4. *Intelligence, conscious thought and their relation to the physical world*

Each of us, through complicated networks of processes occurring inside our organism, which in turn interact with networks of processes taking place outside it, lives (or suffers) firsthand an incessant variety of experiences: sensory impressions produced by stimuli from the surrounding world, internal emotional states, perception of the self, associations of mental images, reasoning, desires, intentions, actions, reactions, projects, and so on.

Since each human being goes through a set of conscious experiences, it is reasonable to infer that, before the appearance of life in our universe, a complex variety of evolutionary and self-organizing processes still took place (albeit devoid, as is generally assumed, of minimal experiential value). Hypothetically wanting to adhere to materialism and to the idea that exclusively physical mechanisms were at work in that cosmic era (in accordance with our current theories) not associated with any kind of proto-experiential properties, it would be natural to conclude that some form of consciousness would unpredictably appear only at a given level of evolutionary complexity of biophysical processes occurring somewhere in our universe, at least on planet Earth. Based on this hypothesis, called "emergentism", certain properties, such as intelligence and awareness, would have arisen suddenly and incomprehensibly during the evolutionary history of very complex organisms.

[9] Given the partial nature of the theories we have in every cognitive field, only the knowledge of the ultimate elements (their ontological foundation, the reason for their existence and their intrinsic and relational properties) could unify them in a complete and self-consistent Theory.

Before delving into the subject concerning the phenomena of mental reality and, in particular, the phenomenon of conscious experience, I would like to recall the different philosophical positions of realists in the face of the problem of the relationship between the mental world and the physical world. Although these two realities seem profoundly different from each other, the possibility of understanding their common nature with the means of science is not excluded.

First of all, I will observe that the vast majority of realists do not hesitate to establish a distinction between two classes of entities or individuals (as I call them in some contexts) with fundamentally different properties: class M of individuals defined as "material" (or "inert") and class V of individuals defined as "living".[10] Individuals of class M have physical properties and the property of interacting with each other, while those of class V have, in addition to the aforementioned properties, also the ability to reproduce, curiosity (such as, for example, that of exploring the surrounding environment, obtaining food and preparing means of defense against dangers), attention and, at a given level of complexity, the ability to organize and plan one's own future and, of all this, being able to be fully aware.

This distinction makes the relationship between V and M clearly problematic, i.e. between the world of individuals capable of making observations and the world of individuals who are considered exclusively objects of observation, in other words, between a reality characterized by the experiential phenomenon and a reality in which this phenomenon seems to be absent.

As already pointed out at the beginning of the chapter, realists who assume the existence of two different realities as irreducible to each other are called "dualists" in the strong sense of the word. They are mostly pragmatic and disinterested in metaphysical questions, while some of them, as well as some supporters of monism (see Spinoza and Einstein in note 50) tend to embrace a mystical vision of cosmic reality.

Then there is a host of realist scientists called "physicists", who oppose Cartesian dualism with a physicalist monism. The latter propose the reduction of the mental world to the ontological supremacy of the physical world as their objective, seeking answers to questions concerning the characteristics of the mind, in particular the phenomenon of consciousness, in the field of physics only, calling themselves out, so to speak, from the narrow scope of biology. However, there are biologists who are against the idea of being able to reduce the mental world to the physical world but who, like the physicalists, are convinced that not all individuals, such as a jellyfish, a virus, a macromolecule or an electron, can to be endowed with protomental properties or a glimmer of consciousness.

Furthermore, there are physicalists who, starting from profoundly different philosophical positions, take into serious consideration the phenomenon of ~~*e*~~*mergentism.*[11] But as will be seen later, no reductionist proposal would seem able to lead to a causal closure of the physical world. On the other hand, according to some influential philosophers, it seems that this goal can be pursued only by anti-reductionist proposals. The latter provide interesting insights and the acquisition of some experimental results for the explanation of specific cognitive functions, but, like all the theories disclosed so far, they leave the question of how and why the experience takes place unanswered.

I now deem it appropriate to deal with the physicalists and, in particular, with those scientists convinced that all the phenomena (both physical and mental) of cosmic reality can be framed, in principle, in a single physical theory, in general based exclusively on ontology of quantum fields, and therefore also confident of being able to unify these fields and the gravitational field.

---

[10] "A living system is: a), a self-organizing system away from thermodynamic equilibrium such that; b), its processes are governed by a program stored symbolically and which, c), is able to reproduce itself, including the program"; this definition is given by the physicist Lee Smolin in his essay *The life of the cosmos*, Giulio Einaudi editore, Turin 1998, p. 197.

[11] The philosophers of mind who are in favor of emergentism argue that the mind and related phenomena are emergent properties of the brain..

However, to hope to achieve their goal, such scientists will have to find a way to reject or modify some fundamental ideas of quantum theory and/or general relativity, in order to obtain a unified description of the microcosm and the macrocosm, as well as to incorporate and expand and deepening it, Darwin's theory of evolution.

The physicalist project, from whatever philosophical position it intends to be tackled for an explanation of how to fill the explanatory gap that separates the physical from the mental, therefore appears to be such an ambitious undertaking that its realization should rely on the extraordinary resources of the human mind.

Such a project, if for example it wanted to be based on the physics of material particles, should be able to explain all physical systems, starting with those conventionally defined as "elementary" (leptons and quarks) and going up through a variety of composite systems (such as protons, atoms and molecules) towards those with a high organizational complexity regulated by feedback mechanisms and characterized by a relative autonomy, by sensitivity, perception, curiosity, discernment, learning, understanding, intelligence, intellectual reflection, to finally reach those processes responsible for conscious experience, which is the only phenomenon that is absolutely indubitable and pertinent at least to human beings.

In tackling the thorny problem of mind-body causation, I will consider some particular philosophical positions of physicalism, all confident in the possibility of a scientific description of the manifestations of the mental world.

5. *Functionalism, or Strong Artificial Intelligence Science*

Artificial intelligence (AI) is a branch of science pertaining to computer scientists, engineers, philosophers of mind and neuroscientists who study the mechanisms responsible for human behavior and cognitive faculties, and its aim is to reproduce them in machines controlled by suitably programmed computers.[12]

The science of AI operates in two fundamentally different realms, one called “strong” and the other “weak”. The former[13] represents a rather ambitious point of view of physicalism and is also called "functionalism". It argues that the proper characteristics of the mental world, such as intelligence, understanding and awareness, are the result emerging from complicated and appropriate calculations and that, in principle, they can be not only completely imitated by a machine controlled by a suitable program based on computation, but also to give this machine the same characteristics relevant to the human brain and, therefore, put it in a position to have thoughts and self-awareness.

The basic idea of strong AI can be found in the writings of the empiricist Thomas Hobbes (1588-167), who argued in Leviathan that “reasoning is nothing more than calculating the consequences of uniquely defined names”.

Then there are some students of strong AI who go much further, as they are confident that in the not too distant future intelligent computers will be able to far surpass all the abilities of human beings, acquire the ability to reproduce and evolve unpredictably, perhaps to the point of possessing complete knowledge of our universe.

It is not difficult to imagine that such a scenario would raise some disturbing questions. What role would humans play? Could they be a kind of inferior being wholly incapable of communicating with the superintelligent systems they themselves may have created? Is the human race doomed to extinction?

---

[12] Computer programs are defined in purely formal (or syntactic) terms and their limitations, as John Searle observes, are due to their lack of semantics.

[13] Strong AI is opposed to weak AI, on the basis of which it is possible, in principle, to construct a machine controlled by a sophisticated computer capable of simulating external manifestations of sensitivity, intelligence and awareness, such as those typical of a human brain, but that such a simulation does not actually possess these characteristics.

One of the best-known avant-gardists of functionalism is Marvin Minsky (1927-2016), a tireless student of mental processes and their applications in computerized machines, who theorized The Society Of Mind,[14] in which he argues that the mind would be the cooperative result of a set of brain "agencies" each responsible for carrying out a particular cognitive function. The agencies would communicate with each other through connections, so as to form an overall hierarchical structured system.

Minsk also introduced the notion of "frame",[15] a structure that collects all the information which, when associated with each other, would contribute to forming a given concept. The frames are described by him as interconnected according to a very articulated scheme, in such a way that "each component of each frame is connected to the frame that describes its structure."

An extremization of functionalism consists in considering the universe as an immense calculator and in hypothesizing that appropriate subcalculations that can be extrapolated from it would be able to produce the phenomenon of awareness.

As already mentioned, there are points of view in stark contrast to that of strong AI. One of them is based on the weak domain and states that conscious experience is due to a physical action of the brain and that a computer is certainly capable of simulating any physical action, including therefore the behavior of human beings, but that any simulation, however sophisticated, suggestive and even such as to mislead a human subject, cannot achieve awareness of what it is doing.

A second interesting point of view consists in maintaining that an appropriate physical action of the brain is able to arouse awareness and that this is a kind of non-computable property, so that no machine based on computation, however complex and equipped with a memory unlimited, could never equal all the characteristics of human beings. Finally, a third point of view hastily dismisses the problem of awareness, because it is considered absolutely foreign to the domain of science.

In summary, the physics we have at our disposal asks whether every physical action can, in principle, be simulated or not by a computer. It can be answered, as shown in Gödel's incompleteness theorem, that any deductive system is incomplete, i.e. that within the realm of mathematical logic there is a limitation to the explanatory possibilities based on computation.

An extremization of functionalism consists in considering the universe as an immense calculator and in hypothesizing that appropriate subcalculations that can be extrapolated from it would be able to produce the phenomenon of awareness.

As already mentioned, there are points of view in stark contrast to that of strong AI. One of them is based on the weak domain and states that conscious experience is due to a physical action of the brain and that a computer is certainly capable of simulating any physical action, including therefore the behavior of human beings, but that any simulation, however sophisticated, suggestive and even such as to mislead a human subject, cannot achieve awareness of what it is doing.

A second interesting point of view consists in maintaining that an appropriate physical action of the brain is able to arouse awareness and that this is a kind of non-computable property, so that no machine based on computation, however complex and equipped with a memory unlimited, could never equal all the characteristics of human beings. Finally, a third point of view hastily dismisses the problem of awareness, because it is considered absolutely foreign to the domain of science.

In summary, the physics we have at our disposal asks whether every physical action can, in principle, be simulated or not by a computer. It can be answered, as shown in Gödel's incompleteness theorem, that any deductive system is incomplete, i.e. that within the realm of mathematical logic there is a limitation to the explanatory possibilities based on computation.

[14] Marvin Minsky, *La società Della Mente*, Adelphi, Biblioteca Scientifica, 10 (1989).

[15] Marvin Minsky, in 1974, in the article *A Framework for Representing Knowledge*: frames allow knowledge to be divided into sub-structures that represent "stereotyped situations". Frames are the primary data structure of frame languages used in artificial intelligence.

Similarly, in physics, Alan M. Turung (1912-1954), the scientist who started the computer age, demonstrated that a robot controlled by any computer is not able to provide an answer to all questions formulated on the basis of a correct symbolic formalism, as there are some that involve the so-called "halting problem". In other words, there are questions to which the computer reacts by going on and on without giving an answer.

Therefore, given that the brain is a physical structure, one wonders how it is possible for conscious experience to take place in it by resorting to non-computational processes. One also wonders whether these processes responsible for awareness are really physical processes, or processes with particular characteristics that today's science does not currently have the faintest idea of how to conceive and describe.

For the most part, scientists tend to approach the problem of consciousness following a reductionist agenda. A relevant example of this trend is Minski's functionalism which I have just mentioned.

Then there are some reductionist scholars who hope to identify the brain circuits responsible for conscious experience or, more properly, the so-called "neuronal correlates of consciousness". Among the scientists who have set out to search for these neuronal correlates, the most authoritative is Francis Crick (1916-2004), discoverer in 1962, together with James Watson (1928), of DNA. He is convinced that he has found them and has come to the conclusion that experience is reflected in a set of processes that unfold in bundles of neurons; above all, he is convinced that only the in-depth study of the mechanisms of interaction between single neurons can provide incontrovertible experimental data to be used for the construction of a theory of consciousness. Furthermore, in 1990, Crick and his disciple Christof Koch (1956) elaborated a theory of consciousness based on neurobiology, more precisely on the oscillations (between 35 and 75 cycles per second) of the neuronal discharges of the cerebral cortex. They assume that

> oscillations are the basis of consciousness. Partly because the oscillations appear to correlate with awareness in a number of different ways – within the visual and olfactory systems, for example – and because they suggest a mechanism by which binding of information contents can be achieved. Binding is the process by which separately represented pieces of information about a single entity are joined together for use by subsequent processes, such as when information about the color and shape of a perceived object is integrated by separate visual pathways [. ..]. Crick and Koch hypothesize that binding can be achieved with synchronized oscillations of neuronal clusters representing relevant content. When two pieces of information are tied together, the relevant neuronal groups will oscillate with the same frequency and with the same phase.[16]

However this proposal, properly defined as "neuro-reductionist" and with which the two scientists would like to suggest that the aforementioned oscillations constitute the neuronal correlates of conscious experience, does not solve the difficult problem at all.

Crick and Koch, although they have the merit of having carried out a large amount of scientific work for the explanation of specific brain functions (in particular the visual function), do not provide an answer to the question "why the synchronized oscillations of neuronal discharges give rise to the subjective experience?"

6. *Roger Penrose's physicalist approach to the problem of consciousness*

Penrose demonstrated an extraordinary commitment in wanting to give a turn to the mind-body problem (or states of consciousness-brain), introducing interesting ideas and rigorous arguments in sharp contrast with those of functionalism. His arguments are available in three well-known popular essays: The Emperor's New Mind (ENM), Shadows of the Mind (SM), and The Great, the Small

[16] Francis Crick e Christof Koch, La scienza della coscienza, da Francis Crick ai TED Talk, 2016..

and the Human Mind (GSM), all three of which have come to attention of the scientific community and still the subject of heated discussions.

I will briefly outline Penrose's central thesis, but without going into detail, and then relate some important observations made in GSM by his fellow philosopher and physicist Abner Shimony, as well as Penrose's reply.

I myself, with all due respect to these scientists, will take the liberty of highlighting the essential points of my disagreement with their ideas and I will do so in the conclusions of this chapter which will lead me, in other writings, to propose a personal vision of reality through the introduction of a new geometric-mechanical concept.

Penrose, an openly Platonist[17] mathematician, but also a scientist with a profound interdisciplinary preparation, has been engaged for almost thirty years (with a group of his students from the University of Cambridge) in the search for a complete physical theory, the so-called "quantum gravity", which it should also be able to describe the functioning of specific areas of the brain related to the phenomenon of consciousness.

Penrose's thesis is articulated along three basic directives. First, he assumes that the mysterious manifestations of the brain are not based solely on computation and that, especially, the logic of abstract mathematics, the understanding of which is accessible to the human mind, could not be programmed into a digital computer, a no matter how sophisticated and broad his memory.

To give strength to his belief, Penrose brings as an example, in addition to the game of chess, the argument of the "Chinese room" by John Searle, designed to counter the theses in support of strong AI:

> […] As a matter of fact, I don't understand a word of Chinese. […] But let's imagine that I am locked in a room with some boxes full of Chinese symbols, and that I have a rules manual, in fact a computer program, which allows me to answer questions formulated in Chinese. I receive symbols which, unbeknownst to me, are questions; I look in the manual what I am expected to do; I take symbols from the boxes, manipulate them according to the program's rules, and send out the requested symbols, which are interpreted as responses. We can assume that I pass the Turing test for understanding Chinese, but still, I don't understand a word of Chinese. And if, while implementing the appropriate computer program, I don't understand Chinese, then no other computer understands it just by implementing the program, because no computer has anything that I don't have. [...] Let's imagine that in the same room I'm also asked questions in English, which I answer. Outwardly, my responses to English and Chinese questions look equally good. For both I pass the Turing test. But seen from the inside, the difference is huge. […] In English I understand what words mean, in Chinese I don't understand anything. For Chinese I am just a computer[18].

Penrose then uses a variant of the well-known Gödel theorem in a form almost similar to the argument proposed by Turing[19], once again with the aim of stating that

> mathematical understanding is not reducible to computation, but is something quite different that depends on our ability to be aware of things.[20]

---

[17] Penrose contrasted the hierarchical scheme of the three worlds theorized by Popper (the physical world, the mental world which is a product of the physical world and, lastly, the world of culture which, in turn, is a product of the mental world) with his own particular scheme in which he relates the physical world and the mental world (the latter, in the same way as Popper, conceived as a product of the former). These two worlds, together with the Platonic world of disembodied and eternal forms inclusive of mathematical "objects", constitute a triad of interdependent realities. See, Penrose, The small, the large and the human mind, Cambrige University Press, 1997, pp. 93-98.

[18] .. John R. Searle, *La mente*, 2005, Milano, Raffaello Cortina Editore, pp. 81-82 .

[19] The human brain, while being able to collide with problems that it would not be able to solve, at least at the current stage of its development, gives the impression of being governed by a structure of operational rules more powerful than that of the machines conceived up to now. The mathematician's mind, to be able to understand Gödel's theorem, cannot be isomorphic to any Turing machine, since if it were, it could not calculate and, at the same time, understand what it actually understands.

[20] GSM, p. 116.

Secondly, Penrose states that in order to understand the relationship between the body and the mind, a substantial modification of the formalism of quantum mechanics will be required, because it is based on two principles of dynamic evolution that are radically different and contradictory to each other: on the one hand, unitary, **U**, of the wave function governed by Schördinger's equation and, on the other, its reduction, **R**. The passage from **U** to **R**, i.e. the passage from the potentialities, implicit in the principle of superposition of states, to their actualization through a well-defined result, still remains incomprehensible. In short, we are dealing with the problem of measurement in quantum mechanics, to which some reference was made in the introduction.

Thirdly and lastly, Penrose proposes to replace **R** with a procedure he calls "Objective Reduction", briefly **OR** (from the English Objective Reduction), a form of the decoherence[21] phenomenon which is not based on computation and which requires quantum gravity (theory he sought with a certain optimism).

The **OR** phenomenon would be produced at a given instant through the pairs of microtubule bundles, each of which is present in a neuron and, more precisely, inside the button in synaptic contact with the dendritic spine[22]

[…] microtubules are small tubes made up of proteins called *tubulin*s. These [...] appear to have (at least) two different states, or conformations, and are able to switch from one configuration to another [...]. According to Hameroff, microtubules would behave like cellular automatons and complicated signals could be sent through them. Think of two different conformations of each tube represented by the 0 and 1 of a digital computer. A single microtubule could then behave like a computer, and this must be taken into account when analyzing what neurons do. Each neuron doesn't just act like a switch; instead it involves a lot of microtubules, each of which is capable of doing rather complicated things.[23]

Penrose argues that the two bundles of microtubules within each neuron are sufficiently isolated from the surrounding structures to be able to justify a coherence of quantum states in large areas of the brain of intense scope, enough to explain the mind in global terms. So, it goes on with the belief that

It might be that quantum mechanics is important in understanding these processes. […] It may well be that, within the tubes,there is somekind of large-scale coherent quantum activity, somewhat like a superconductor. Significant mass movements would be involved only when this activity begins to get coupled to the (Hameroff type) tubulin conformations , where now the "cellular automaton" behaviour would itself be subject to quantum superposition. [...] there would have to some type of coherent quantum oscillation taking place within the tubes which would need to extend to very large areas of the brain. […] It seems to me that […] any physical process responsible for consciousness would have to be something with an essentially global character. Quantum coherence certainly fits the bill in this respect. For such large-scale coherence to be possible, we need a high degree of isolation, as might be supplied by the microtubule walls. However, we also need more, when the tubulin conformations begin to get involved. But something more is needed when microtubule conformations come into play. This further degree of insulation from the environment might be supplied by ordered water just outside the microtubules. Ordered water (which is known to exist in living cells) would be likely also to be an important ingredient in any quantum coherent oscillations taking place inside the tubes.[24]

In his (anti-reductionist) thesis, Penrose resorts to two concepts: that of potentiality and that of plot. The first is introduced to argue that in a network of neurons of given size each neuron, being in a superposition of states, performs a so-called "quantum computation" (each part of the

---

[21] Decoherence, or desynchronization, Zurek, W.H.: *Decoherence and the transition from quantum to classical*, Physics Today, 44, 36-44 (1991).

[22] The basic idea of Penrose's thesis comes from the observation that even single-celled organisms exhibit purposeful behavior, despite the fact that there is no trace of neurons in them. These organisms are able to react to light, to circumvent obstacles in exploring the surrounding environment and possess a primitive form of memory. This happens, for example, in the paramecium, whose cilia responsible for movement are microtubular structures that produce synchronized oscillations. Microtubular structures are also found in the tail of sperms. But what is more interesting is that they are found in neurons. Therefore, it can reasonably be assumed that the fundamental element of cellular "intelligence" is to be found in the microtubule.

[23] Idem, pp. 130-131

[24] Idem, pp. 131-133

superposition performs its calculation independently of the calculation performed by the other or by the other parts of the overlap), and the second (which is a holistic concept and which Penrose prefers to define as "coherence") is called upon by him to explain the execution of the aforementioned calculations. And since the superposition of states of each neuron is propagated to the two bundles of the microtubule it contains, there will be a coherence extended to the entire network of neurons of that given size.

Finally, to justify the actualization of the states of consciousness, Penrose brings in the **OR** (an operation which, according to Gödel's theorem, would explain the non-computational aspects of mental activity), capable of transforming the global coherence of large areas of the brain (that is, the interweaving of quantum superpositions relating to a large number of microtubules) in a well-defined state of consciousness or, in other words, in a conscious experience that will be lived by the owner of that brain.

Later, Penrose also published together with Stuart Hameroff (1947) an article in Physics of Life Reviews, in which he restated his theory on the basis of empirical evidence.[25]

As announced, the fourth chapter of GSM is dedicated to a speech by Abner Shimony, who indicates some passages of Penrose's thesis with which he agrees, such as the belief that many ideas of quantum mechanics will be required for understanding of mental phenomena, as well as the need to modify their formal procedures.

However, Shimony identifies other passages that he does not share at all, above all Penrose's adherence (common to all the theories of physicalists and biologists) to emergentism, on the basis of which the manifestations pertinent to the mind suddenly supervene at a given level of organized complexity of the physical world.

In all honesty, Penrose declares that he does not know how to draw a line of demarcation between the consciousness of the human being and that of certain living beings. However, he suggests that such a quality is not possessed by a wide variety of individuals from a given level of unspecified biophysical complexity onwards.

The intervention that I consider most interesting from Shimony's writing is a sort of reproach addressed to Penrose for never having taken into consideration or even mentioned in his essays the organic realism of Alfred N. Whitehead [26] (1861-1947), on the basis of which the manifestations of the mental world, which are evident in complex organized structures such as human beings and, probably, in many other evolved species, derive from an ultimate level of entities called "current occasions" or "space-time quanta", each endowed with a protomental quality defined as “subjective immediacy and appetition”.

Although there are obvious analogies with the monads of Leibniz, unlike these, the quanta conceived by Whitehead are not absolute and permanent entities, but fundamental processes that express themselves in a monotonous and iterative way. As the spatio-temporal and protomental quanta organize themselves into more complex structures, subjectivity grows in intensity until it becomes significantly intense in humans.

Shimony proposes a modernization of Whitehead's panpsychist conception to the attention of Penrose with the aim of suggesting to his thesis a possible remedy to the problem of the non-derivability of the mental world from the physical world, which remains somewhat mysterious, just as the phenomena described by quantum mechanics, in particular, the passage from potentialities to their actualization, not to mention the future theory of quantum gravity (which Penrose sketches in a little detail and about which he expresses the conviction that it will be a non-computable theory); finally, the mysterious quantum phenomenon of non-local effects should be included.[27]

---

[25] Quantum oscillations within the tubes would have to be coupled in some way to the action of the microtubules, namely the cellular automaton activity that Hameroff talks about, but now his idea has to be combined with quantum mechanics.

[26] Norton Whithehead, *Processo e realtà*, Bompiani, Milano, (1965), p.

[27] LSM, Chapter 4.

How Shimony intends to modernize Whitehead's panpsychist conception is not at all clear. It seems that he simply wants to suggest to Penrose to rethink his physicalist conception by considering the idea of a protomind at the quantum level. However, Shimony does not provide him with any significant cue to be able to translate Whitehead's mentalist vision into a physicalist key, as he remains firmly of the idea that the manifestations pertaining to the mental cannot find a place in a physicalist theory.

Furthermore, I find it unsatisfactory that Shimony bases his proposal on the same principles of **QM** used by Penrose: indeterminacy, acausality, probability, potentiality and entanglement. However, Shimony is keen to point out that, in his attempt to modernize Whitehead's ideas, Penrose

he does not use quantum theory as a surrogate for the ontological state of the mental, but as a purely intellectual tool to explain the immense range of manifestations of the mental in the world, from the most intrinsic to the highest level expressions.[28]

In fact, he observes that if quantum mechanics were accepted as a complete theory, i.e. not susceptible to substantial changes, then all the concepts listed above should be considered objective aspects of the theory.

But let's get to the core of his reasoning. Instead of considering the leap from consciousness to non-consciousness (leap that also occurs in a highly evolved structure such as the brain) as a change from a mental condition to a physical condition, Shimony proposes to consider it

as a change of state from definiteness to indeterminacy and vice versa. In the particular case of a system as simple as an electron, one can only imagine a transition from the total indeterminacy of experience to a minimal glimmer".[29]

If then - continues Shimony - we take into consideration, instead of a single electron, a set of atomic systems each endowed with feeble mental aspects and we bring into the scene the concept of intertwining, the latter

can conceivably generate a large whole that contemplates all the possibilities between non-consciousness and high-level consciousness [30]

Penrose's theorization, observes Shimony, is incomplete because it lacks "an idea of the mental as something ontologically fundamental in the universe"[31] and, therefore, it is unsatisfactory. Although set out in a non-detailed way or, as the author himself defines it, "rudimentary", a modernization of Whitehead's conception is, in my opinion, full of interesting starting points for the development of a scientific theory (I will return to this topic in sections 8 and 9).

In his reply to Shimony, Penrose expresses his gratitude for the precious suggestions and declares that, while not ignoring Whitehead's works, he has never studied them sufficiently to be able to grapple with his ideas. However, he finds Shimony's attempt to modernize Whitehead's philosophy of the organism quite stimulating and confesses that he is also willing to believe that the mental can be ontologically fundamental in cosmic reality.

Penrose then reiterates the importance of the interweaving by stating that, if the phenomenon of interweaving did not take place in highly complex organizational structures (such as the human brain) capable of processing information, we would not know what other footholds to turn to to explain how all those mental manifestations of which we have direct experience can emerge. Furthermore, he admits that his ideas aimed at trying to explain the functioning of the brain and the phenomenon of consciousness are still quite confused.

[28] GSM, cit. p.
[29] Idem, cit. p-
[30] Idem, cit. p.
[31] Idem, cit. p.

Finally, wondering if and to what extent quantum effects are essential to explain the occurrence of conscious thoughts, he states that in this regard it will be good to think long and hard to avoid venturing into too hasty conclusions.

7. *Some observations on Penrose's thesis*

I will now present some of my personal observations on the ideas proposed by Penrose, and then on physicalism in general. I state that I share two particular beliefs of yours: the first states that the functioning of the brain is not based exclusively on computation and the second that the physics we currently have at our disposal is not able to provide us with the means to understand the phenomenon of perception and, even less, of conscious experience. What I do not agree with is the way in which Penrose intends to adjust to modify quantum mechanics. In fact, it seems to me that he limits himself to challenging only the dynamic process of reduction of the wave function, R, because it is obscure and the source of the well-known paradoxes connected to the measurement problem, and to replace it with his OR procedure, which he arrives at with rather forced upon the physiology of the brain and with which he finally obtains no satisfactory result, since it provides no explanation of what conscious experience is.

Penrose introduces the OR for the simple fact that, as Hawking observes,

it is somehow necessary to him since Gödel's theorem implies that conscious experience is not computable,[32]

that is, it cannot be genuinely simulated by a computer (this being based on algorithmic processes).

What I find unconvincing is the fact that Penrose aspires to account for the deepest mystery of nature, namely the phenomenon of consciousness, while tacitly accepting en bloc the various mysteries implicit in quantum mechanics, some of which he considered inaccessible to science. human mind and, more precisely, what he calls "mysteries Z",[33] such as for example *quantum non locality*.

My point of view is quite different: consciousness is a phenomenon that can only be explained on the basis of a radical monism (which I will mention at the end of the chapter), but on the condition that the paradoxes and anomalies that quantum mechanics. The latter, although it constitutes an extraordinary theoretical apparatus capable of making very accurate predictions, does not give us any genuine understanding of physical reality.

Unlike Penrose and many other distinguished scientists, I believe that quantum mechanics can be radically reinterpreted on a basis of rational concepts and integrated with gravity. I also believe that it will also be possible to understand the profound reasons why both dynamic procedures, **U** and **R**, on which the theory is based, although contradictory to each other, are so well functional for all practical purposes.

At this point I have to ask why Penrose, instead of consuming himself in age-old and complex reasoning, has never bothered to devise strategies aimed at questioning some of the incomprehensible concepts of quantum mechanics, such as the intrinsic indeterminism of atomic systems, superposition of states, non-epistemic probabilities, non-locality, and so on, nor to take steps to question the validity of some of the principles of classical logic.

Perhaps there are no more scholars willing to follow the example of Einstein, who worked tirelessly to devise conceptual experiments aimed at demonstrating the objective existence of atomic systems and their properties.

If Einstein's belief in the rationality of cosmic reality, i.e. based on the concepts of three-dimensional space, time and the law of cause, was officially discredited by the experiments performed in 1982 by Alan Aspect (1947), whose results violate Bell's inequality and be in agreement with those foreseen by quantum mechanics; however, it cannot be excluded that in the

[32] Idem, cit. p.171

[33] Roger Penrose, SM, cit. pp. 296.

future Einstein's conviction may be re-evaluated in the light of new ideas and new experimental results (see note 89).

The conclusions drawn from those experiments still seem questionable, but there is only a small minority of theoretical physicists willing to oppose them, such as Franco Selleri (1936-2013);[34] however, this minority remains unheard of by the dogmatism that dominates most of the scientific community. Furthermore, admitted and not granted the correctness of those results, the eventual legitimacy of the non-locality might not represent a necessarily unfathomable mystery. In fact, this phenomenon has been interpreted on the basis of rather convincing and generally shared results, but perhaps misleading and, in my view, attributable to a rational explanation.

8. *More general observations on physicalism*

It is not difficult to see that anti-reductionist physicalism, if it really intends to be a comprehensive theorization of both physical and mental entities that are non-physical, still seems to recur in a form of dualism.

According to this point of view, mental entities, although considered compatible with physical laws, are not deducible from them, but unpredictably occur at a given level of organizational complexity which, moreover, is not in any case specified in an unequivocal way. It can therefore be objected that the definition of "physicalism" conceived in this way, i.e. based on the mysterious phenomenon of the supervening of non-physical entities on physical entities, being unable to overcome dualism, proves to be completely inappropriate.

The one based on the theory of identity, according to which the mental is identified with the physical, would then seem preferable to this kind of physicalism. But even then, whether one takes the mental as something non-physical, or one takes it as something other than the physical (i.e., having some familiarity with the physical, while not transcending the physical), dualism cannot be avoided. Furthermore, it should be noted that all anti-reductionist theories based on the assumption that mental states are physical states would become, de facto nisi de iure, reductionist theories.

These considerations therefore seem to imply that our conception of the world is destined not to be able to free itself from some form of dualism in all those cases in which one wishes to assume a philosophical point of view based on any form of anti-reductionist monism.

I deem it appropriate to recall that the philosophy of mind and the science concerned with solving the problem of experience were severely criticized by the psychologist William James (1842-1910) even before the ideas and theses described above were developed. Indeed, the three main points of view adopted today by scholars of the mind in favor of monism are still open to criticism in the manner of James, not only for the vagueness and ambiguity of the various proposals, but above all for their inadequacy to take a significant step towards a full overcoming of dualism and to give a satisfactory answer to why and how physical processes are accompanied by conscious experience.

We do not know what the world itself is (an sich), we do not know whether it is actually material or of another nature, since, as James observes in his speech given in Rome in 1905 at the V International Congress of Experimental Psychology,

> the world is only an object of experience; and the indispensable condition for this is that it is referred to witnesses, that it is known by a subject […]. Object and subject are the two legs without which it seems that philosophy cannot take a step forward.

One wonders then whether it is possible to develop new and appropriate concepts in support of a coherent form of monism, i.e. such as to avoid the insurmountable obstacles encountered by the tortuous reasoning of philosophers, probably prisoners of a rationality based on logical categories and principles misleading.

[34] Franco Selleri, Fondamenti della fisica moderna, Jaca Book, 1996.

On the front of anti-reductionist conceptions regarding the mind-brain relationship, all converging in some form of dualism, James flaunts his pragmatism, arguing that the facts that constitute reality and the facts of the experiences we consciously make of them, when these facts they are produced, they are completely coincident with each other, that is, they are an inseparable whole.

Consciousness and matter do not correspond to two essences of a different nature, but they are both the same experience. What is called "physical" and what is called "mental", i.e. experienced object and subject of experience, are nothing more than simple conceptual creations, they are finally distinctions of a practical order. Those who strive to make ontological distinctions are accused by James as victims of a perverse way of thinking. James calls "experience" the stuff which he holds to be common to all the variety of existing things, whether one calls them "material objects," "physical processes," "dreams," or "thoughts."

However, there is a fourth anti-dualistic point of view which seems far more interesting than the others and which today enjoys widespread acceptance: the "neutral monism" proposed by Bertrand Russell (1872-1970) in 1927, according to which reality of the world is made up exclusively of events that do not underlie any substance. Therefore, neither a material substance nor a mental substance is associated with events. What is usually distinguished into material and mental derives from the same entity called by Russell "neutral substance".[35]

Although Russell does not go beyond the assumption of such a substance, that is, he does not try to propose an ontology of natural events, I wanted to mention this particular form of monism of his, not only because it was a source of inspiration for the work carried out recently by some scholars, in particular by Chalmers, with the aim of being able to build an explanatory bridge between mental activity and brain activity, but also because I personally consider it, among all the known philosophical ideas, the most reasonable starting point for the possible formulation of a Theory of Consciousness.

9. ***Hints at some "double aspect theories"***

In his thesis called "naturalistic dualism", Chalmers argues that phenomenal consciousness is constituted by intrinsic properties of the fundamental physical entities,[36] and that therefore it is one of the basic elements of nature. Its dualistic interpretation consists in attributing protophenomenal (or protoexperiential) and physical properties to the fundamental units of nature. Physics originates from the relationships between these units, while consciousness originates from their protophenomenal properties. On the basis of these assumptions, the philosopher and scholar of cognitive science Andrea Lavazza observes that the conception of Chalmers

it is compatible with the causal closure of the physical world and with the framework of physical theory, capable of overcoming the difficulties of mental causation [...]. There are fundamental neutral (protophenomenal) properties that constitute both the physical and the phenomenal domain. A panprotopsychism follows: everything is endowed with at least a minimal capacity for conscious experience (even a stone or a thermostat).[37]

——Thus the idea that Shimony suggested to Penrose seems to be making a comeback, that of taking into serious consideration Whitehead's philosophy of the organism (based on panpsychism) to modernize it in a physicalist key. Here, however, the question arises of how to conceive the minimal unit of consciousness in nature (provided that the stream of consciousness is quantized like other physical quantities).

[35] Bertrand Russell, *The Analysis of Matter*, new edition published in 2007 by Spokesman, Russell House, Nottingham, England.

[36] In filosofia della mente , la teoria del doppio aspetto, chiamata anche "monismo a doppio aspetto", sostiene che il mentale e il fisico siano due aspetti della stessa sostanza che non può essere compresa,

[37] Andrea Lavazza, *L'uomo a due dimensioni*, Ed. Bruno Mondadori, 2008, cit. p.60.

The philosopher Galen Strawson (1952), in a simpler and more direct way, supports the validity of a panpsychist conception, as

necessary (and unwanted) to get out of the impasse of those who firmly believe in physicalism (renamed 'realistic monism') and cannot deny the existence of conscious experience.[38]

The formulation of a theory of consciousness still appears to be a goal that is difficult to achieve. All the theories proposed so far on the problem of consciousness have a speculative character. Excluding those who deny the phenomenon and those who consider it a mystery inaccessible to the intellect, there remain those who, while taking the problem seriously, do not arrive at any convincing formulation, above all because they are all based exclusively on philosophical language (which, in my opinion judgement, remains unproductive without the support of an innovative language for describing the physical world and of a mathematical language that is not reduced to pure abstraction, such as that of topology, where concepts are combined with images of the physical space).

Some optimistic scholars trust in the future progress of cognitive sciences, but above all they seem not to ignore these last observations, as they do not exclude the possibility of finding significant connections between topics concerning apparently heterogeneous fields of knowledge, through which it would perhaps be possible to discover the logical scheme that connects all the pieces of cosmic reality in a unified way.

Among the optimists we meet Chalmers. He agrees with other philosophers that the problem of the explanatory gap between physical processes and experience does not necessarily imply that it is a metaphysical, and therefore non-interpretative, gap. In fact, it cannot be excluded that the experience is of a physical nature, that it is in short a phenomenon that can be described in some way in mechanical terms of three-dimensional space.

In any case - says Chalmers - this position still admits an explanatory hiatus between physical processes and experience. The principles that unite the physical and the experiential element will not be derivable from the laws of physics, these principles must therefore be considered explanatory fundamental [...]. We are already in a position to understand certain central facts of the relationship between physical processes and experience and of the regularities which connect them. Once the reductionist explanation is set aside, we can reconsider those facts to make them the elements of a non-reductionist theory of consciousness and the constraints on the fundamental laws that constitute an ultimate theory.[39]

With this premise, Chalmers undertakes to propose his thesis on the phenomenon of consciousness, starting by taking into consideration two principles: *the principle of structural coherence* and *the principle of organizational invariance*. [40]

Chalmers is well aware that the above principles are of a high standard and that they are not at the right level for

construct the fundamental laws of a theory of consciousness. However, they act as strong constraints. What is still missing are the underlying principles that fit these constraints and can ultimately explain them.[41].

Chalmers' thesis is based on Shannon's information theory and the double aspect principle.

Where there is information - explains the author - there are information states rooted in an information space..[42]

[38] Idem, cit. p. 61
[39] Idem, cit. p. 227.
[40] Idem, cit. p. 229-237
[41] Idem, cit. p. 235
[42] Idem, cit. p. 235

Information would therefore have two fundamental aspects, one physical and one experiential.

It is not the case here to go into the merits of Chalmers' arguments essentially for two reasons. Firstly, his reasonings are exclusively conjectural and, secondly, his initial premise of rejecting the reductionist hypothesis is, from my point of view, misleading and responsible for the impossibility of formulating a founded theory of consciousness, rather than on the idea of the double aspect of information states, on a purely monist position, i.e. on a single set of elementary processes, or units of space-time, having the property of being physically experiential. Even better, each basic unit of space-time should be conceived of as immaterial, having only one aspect that is both physical and experiential.

10. ***Further remarks on Penrose's thesis***

Anti-reductionist physicalists find themselves in a situation of great uncertainty when they propose to establish a boundary between individuals endowed with conscience and those without it. Penrose, for example (like other physicalists, biologists and cognitive science scholars), assumes that certain evolved individuals are endowed with a consciousness not very dissimilar from the human one and that therefore they too are subjects of conscious perceptions. But the idea seems unacceptable to him that this quality called consciousness (or awareness), in descending the evolutionary ladder of biophysical systems, can fade, attenuating its intensity from step to step, i.e. passing from higher order consciousness to a primary consciousness, for example that of a dog, and then to one of a lower order, for example that of a caterpillar, and then again towards and beyond the boundary of chemical and physical systems, without the need for it to be completely suppressed.

Obviously, at any less complex level of organization, rather than using the term consciousness, I would find it more appropriate to agree with less strong terms, such as self-sensitivity, self-management, self-action or, better yet, a very general expression such as , for example, experience lived from an exclusive point of view.[43] Sto insomma parlando di processi fisici che godono di una relativa autonomia organizzativa e che sono distinguibili in una gerarchia di livelli di complessità numericamente limitati, tutti caratterizzati dall'autoriferimento, a partire da un livello base di processi elementari.

By this I mean to argue that even the ultimate level (for example, a simple loop associated with the spin of atomic systems that I will describe in the next chapter) can be a physical process with even a faint familiarity with what we call "experience".

In this sense, there would be different levels of experience, each reducible to the one below, up to the fundamental level which should be characterized by a minimum degree of experience. The latter should consist of mechanical units of space, each capable of suffering the effects of its acting on itself and its interaction with other units (I am speaking out against radical atomism and in favor of the existence of immaterial elementary physical processes).

What could convince us that the phenomenon of conscious experience diminishes as one descends from the level of organizational complexity inherent in human beings to levels of lesser complexity, and then suddenly shuts down at some level of biological, chemical or physical reality, such as that of a mosquito, a bacterium, a macromolecule or an electron?

One of the questions that Penrose does not ask in his thesis and which seems to me of considerable importance is the following: what do all the varieties of individuals ranging from atomic systems to human beings have in common? Consider, for example, the following array of individuals of varying organizational complexity: a man, a chimpanzee, a canary, a fly, a gene, and an electron. Thinking carefully, we can ask ourselves: why not consider them all individuals (or individual physical systems), each with the generic property of self-acting, as well as that of being able to interact in some way with other individuals? It seems reasonable to me to think that this power to self-act and self-regulate, and, in a certain sense, to self-manage the information deriving

[43] The reference to Leibniz's monadological conception is clear.

.

from interactions with the surrounding environment, grows in intensity as the self-organizational complexity of individuals increases.

However, I would not like to be misunderstood: the philosophical position to which I feel close has nothing to do with the ontological hegemony of the mental, and I will propose one of my personal thesis with a completely different approach from that of Penrose, not only as regards the fundamental concepts and principles of quantum mechanics, but also a possible reformulation of the theory based on the three-dimensional space characterized by an unstoppable activity, as well as the description of a physical law which I would define as of a topological-mechanical order and which would allow us to explain, with plausible analogies viewable, the existence and behavior of proto-experiential elementary entities, as well as their methods of aggregation and disintegration, according to the specific conditions around them.

In addressing the problem concerning the phenomenon of conscious experience, I will therefore assume the existence of its precursor at the quantum level, adopting (as already expressly stated) a philosophical point of view that I will call "radical monism". The idea will not be new, but no one has so far proposed a satisfactory description (based on computer simulations) of the experiential phenomenon associated with fundamental physical processes.

At present, in Western culture there is no theory capable of making particular characteristics of the mental world intelligible, such as attention, intentionality, language, understanding and awareness, nor a generally acceptable definition of these terms (considering then that the human being does not seem to be the only individual in which these characteristics are present).

Penrose is well aware of this state of affairs, but he states that, at least from the mathematical point of view, it is possible to establish the relationships between these three terms: intelligence requires understanding and this, in turn, requires awareness . In other words, Penrose argues that talking about intelligence in the absence of awareness makes no sense. All of this helps him counter strong AI philosophical positions, but leaves the mystery of conscious experience intact.

11. ***A triad of mysteries***

The set of phenomena associated with consciousness, according to the point of view of the physicalists, if it were considered a true and proper mystery of nature, in the sense of not being deducible from the laws known to us nor compatible with them, would remain a subject that can be relegated to the field of metaphysics.

The current anti-reductionist proposals therefore seem to reflect a pessimistic attitude, as they all, in one way or another, come to the conclusion that the intellect cannot have access to the understanding of conscious experience.

However, there are scientists, like Chalmers, who are not willing to give up. And if they really intend to commit themselves to formulating a theory capable of answering the question "what is consciousness?", they must bear in mind that such a question cannot be addressed without being joined to two other equally profound questions. The first refers to this chapter and that is "why is there something instead of nothing?". In short, given that something actually exists, one cannot limit oneself to affirming, in the manner of Leibniz, that this something exists by virtue of a principle of sufficient reason. Rather, the explanation of why something exists will be required to be based on a sense of logical necessity.

The second question is "which properties do atomic processes possess to be able to give rise to self-organization and, with it, to the immense variety of more or less unstable entities which make up our universe and which imply the phenomenon of experience ?" The formulation of satisfactory answers to this kind of question would be equivalent to possessing valid prerequisites for the construction of a Theory of Everything. I have already expressed myself in other writings on the search for the Unitary Physical Theory, or Theory of Everything (see online my article "A new image of reality", p. 10, "*The search for a final theory*") . Such a Theory, in order to be able to declare its completeness, will require a definition of the fundamental principle, i.e. the principle

capable of justifying the existence of cosmic reality in an understandable way, possibly through a reasoning that reflects a sense of logical necessity.

The Theory will then require an innovative physical-mathematical support to explain how the individual fundamental processes of nature can take place and how these can interact with each other to give rise to self-organizational structures. In a certain sense, the Theory will have to make it understandable how all phenomena can descend from it, through an evolutionary path (of the Darwinian type): apparently material structures, all the forces that act on them, biology, sensible perception , intelligence and conscious experience.

In short, a vision of the world will be needed that takes into account all the issues addressed up to now starting from the two fundamental categories of thought: nothingness and being.

6. ***Self-referentiality and self-organizing principle.***

The principle of objectivity of nature represents the fundamental postulate that physicists committed to overcoming the unresolved questions of science oppose the obscure principle of complementarity, with which Bohr asserts the impossibility of reformulating quantum theory on a rational basis, that is, inclusive of concepts of three-dimensional space, of time and of causal law. An understandable description of the characteristics objectively possessed by quantum particles cannot be excluded. But the confidence that it can be accessed sooner or later, as long as the QM is not reformulated in intelligible and logically shareable terms, will leave physicists who are in line with Bohr's ideas completely indifferent.

In one of my papers "The role of paradox in scientific discourse" (available online) I mention the possibility of a Theory of Everything being formulated, but also my belief that this goal cannot be achieved through a description based on pure mathematical abstraction .

Finally, assuming that cosmic reality is actually that uncreated Totality, variously unstable and everywhere quivering with incorporeal activity that I have called "**Nih-Entity**" , I wonder if and how one can imagine its way of operating, that is, if it is possible to provide a hint of an answer to the question (3).

To aspire to that much, we will first need firm points of reference that can provide us with a guiding principle. Assuming objective realism, I will highlight the only three certainties available today:

(i) - there is the phenomenon of sensitive perceptions (at least ours) associated with a self-referential thinking apparatus; therefore, there are very complex systems structured in hierarchies of organizational levels regulated by feedback mechanisms from top to bottom and vice versa;

(ii) - there is an absolute limit to sensitive perception; from this it does not necessarily follow that the quantum world is intrinsically indeterminate, but rather that it is impossible to directly explore its characteristics in order to understand its nature and authentic functioning mechanisms;

(iii) - Gödel's two incompleteness theorems imply that the theory of scientific knowledge cannot be exhausted or ascertain the validity of its hypotheses through tools based on number theory.

And here is our guiding principle: self-referentiality, which plays an essential role in the context of these three certainties. But let's see why and to what extent.

Point (i) establishes that there are self-referential processes, at least those concerning the sensitive and intellectual apparatus of the human being, which nevertheless get lost in a tangle of conceptual contradictions when one tries to frame them in a theory that describes their nature, the unity and the law that governs them.

Point (ii) refers to the Heisenberg uncertainty principle, according to which it is not possible to extract from a quantum system all the information necessary to know its initial conditions; this limitation allows to make exclusively probabilistic forecasts. Moreover, the bizarre notions adopted for the description of quantum theory, such as wave-particle, superposition of states and non-localities, seem inaccessible to our intellect, just as inaccessible seems to be the undivided mind-brain unity.

It is also important to consider the power that, based on my personal realistic conception of the quantum world, the actors of the microcosm have, not only to self-act, but also to interact with each other and organize themselves in composite and complex physical systems.

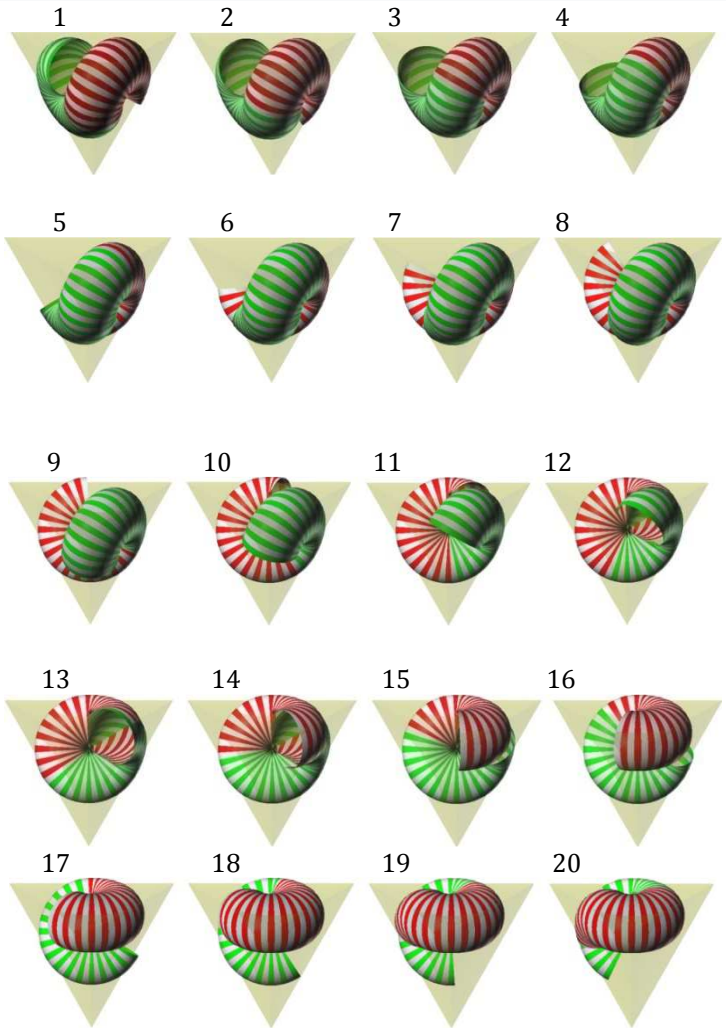
Such a power, as already partially explained in the book Geometrodynamics and Architecture of the Void, should be associated with a topological, dynamic and self-referential property capable of combining the continuous with the discrete, as suggested by figure 1.

This property of physical systems would be amplified, through a series of transitions, towards increasingly complex levels of self-organization starting from a fundamental level.

Finally, there is point (iii), which establishes a limit to mathematical knowledge and which is imposed on us by a theorem built on a self-referential basis, that is, on a proposition concerning arithmetic and which is formally true (correct), but which asserts its own indemonstrability.

However, it is clear that in this case, as already explained in section 7 of my article "Gödel's Shadows between the Mathematical and Physical World" (see note 63), self-referentiality does not play an autonomous role, as it is opportunely introduced as an expedient of creative thinking in order to make mathematical logic a system endowed with self-analytic faculties that it does not actually possess.

All in all, we can reasonably hypothesize that self-referentiality is a property of fundamental physical processes and of their particular levels of self-organization (at the top of which would appear that of human beings), and that the way in which these processes take place is associated with specific properties searchable in some field of geometry, such as topology or node theory (the latter, however, will not be taken into consideration in my proposals).



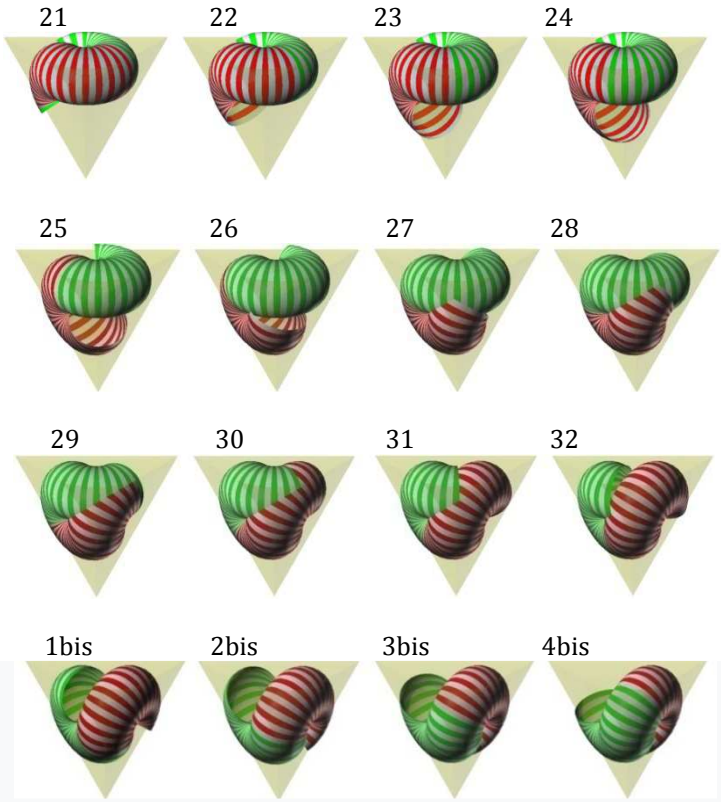


Fig. 1: a toroidal hemi-loop derived from the intersection of the tetrahedron with the sphere (extrapolated from a computer animation of my own design) and describing, in a kinematic way, the 32-phase period of an experiential flow composed of four elementary quanta , each with its own spatial orientation; as you can see, the self-penetration can somehow recall the Klein bottle (but without involving, in this case, a derivation from the Moebius strip).

For a clearer understanding of this figure, I suggest entering <www.carloroselli.com> in the 'Works' section and starting the video 'Periodic Self-Action Loop'.