Mind as Hardware and Matter as Software

Jan-Markus Schwindt[1]

[1] *Institut für Theoretische Physik, Philosophenweg 16, 69120 Heidelberg, Germany*
E-mail: Schwindt@thphys.uni-heidelberg.de

Abstract

We present an argument against physicalism in two steps: 1) Physics reduces the world to a mathematical structure; 2) The notion of "structure" only makes sense when carried by something and interpreted by something else. Physicalism does not allow such a carrier and interpreter at a fundamental level, hence it must be wrong. An extended notion of Mind is presented as the fundamental "hardware" which is necessary by the argument. In particular, qualia correspond to the "monitor component" of mind.
Some ideas are presented on how to extend this mind-matter relation to a more elaborate picture: 1) A system of two complementary reductionisms (one physical, the other mental) may hint toward a deeper reality in which mind and the physical world are closely entangled. 2) A division of mind into a conscious "monitor" and an unconscious "processor" is suggested using the analogy of dreams. Finally, the problem of Solipsism and the existence of "minds other than my own" is discussed.

1 Introduction

1.1 Schrödinger's Problem

"The world" (or what I may call such) is a construct of my mind. It is an extrapolation and generalization of my perceptions, generated by my thinking and intuition. I have learned to associate the dirty white crescent that I sometimes see in the night sky with the word "moon". I got used to the idea that the moon is an "object" that is "really out there". My intuition tells me to believe that the moon is still there when I do not see it. In fact, I do not see the moon at the moment, but I trust my memory and believe that I have seen it many times. I have never been to Australia, but I have good friends who used to live there for a while (at least I believe that they are my friends), and I trust them, so I believe that Australia "exists" and that the stories they tell me are "true". Even if no friends of mine had ever been there, I would still believe in Australia, since I trust the books and the TV reports. I even believe in the "existence" of the top quark particle, although only very few people have "seen" it in a very indirect way. I have observed

many regularities and causal connections in the "world", and in the most cases, I am not suspicious about them, I just take them for granted. So, what I call the world relies on an enormous amount of trust in my conclusions and beliefs. [1]

I have learned to associate myself with something that is almost always present in my sensations and that I use to call my body. Although I have never seen or felt it, I believe that my body contains an organ called "brain". I have learned that my thinking is somehow associated with processes taking place in that brain, but I know that a lot of work was needed for mankind to realize this (more precisely: to realize that thinking is in general associated with brain processes). It is widely accepted to go even further: Not only is thinking associated with brain processes, so it is claimed, but rather thinking *is* a part of the brain processes; and not only thinking, but also the entire consciousness. This is where things are getting problematic. All my direct experiences and sensations seem to be part of something which is only derived from these experiences and sensations: the brain.

Erwin Schrödinger was one of those who realized the strangeness of this remarkable impression. He wrote [1]: "The thing that bewilders us is the curious double role that the conscious mind acquires. On the one hand it is the stage, and the only stage on which this whole world-process takes place, or the vessel or container that contains it all and outside which there is nothing. On the other hand we gather the impression, maybe the deceptive impression, that within this world-bustle the conscious mind is tied up with certain very particular organs (brains), which [...] serve after all only to maintain the lives of their owners, and it is only to this that they owe their having been elaborated in the process of speciation by natural selection."

Schrödinger compares the situation with an artist who places a picture of himself as an inconsiderable minor character in one of his paintings. This seems to him the best allegory for the confusing double role of mind. On the one hand it is the artist who created everything; in the completed work, however, it is only an unimportant decoration which could have been left out without changing the total impression substantially.

1.2 Physicalism

The widely accepted hypothesis that mind can be reduced to physical phenomena relies on the idea that mind is a kind of "software" running on a "hardware" given by the brain. While the association of something material with the word "hardware" still sounds natural for a biologist, it turned out to be inappropriate with respect to the results of the last hundred years of physics. The loss of substance in matter was discussed by many physicists in consequence of quantum theory. Arthur Eddington, for instance, noted: "We have chased the solid substance from the continuous liquid to the atom, from the atom to the electron, and there we have lost it" [2]. James Jeans wrote: "The universe begins to look more like a great thought than a great machine" [3]. Such a loss of substance

[1] It is sometimes argued that our conclusions (concerning the laws of nature) are very likely correct in most cases, since natural selection forced us to develop correct thinking. Otherwise we would not have survived. This argument is, however, intrinsically wrong, since our theory of natural selection is itself a result of our conclusions.

moves matter from the "hardware" to the "software" side. In search for a fundamental theory, many researchers come to the conclusion that the physical world must in essence be some kind of mathematical structure, such as the "wave function of the universe", or a formal system. The lack of a "substance" in such a construction is often argued away by the very vague assumption that mathematical entities have an existence and a life on their own, and that therefore our existence can be reduced to their existence. A typical statement regarding such a structure is given in a paper by H. Nicolai et al [4]: "The 'wave function of the universe' [...] is supposed to contain the complete information about the universe 'from beginning to end'. A good way to visualise Ψ is to think of it as a film reel; 'time' and the illusion that 'something happens' emerge only when the film is played." Such a statement leads me to the following questions: 1) What is the "material" of which the film reel is made? 2) What is the "projector" that "plays" the film? 3) How can a part of the structure contained on it (namely the part which is identified with my mind) miraculously beam itself out of the film, sit in front of the screen and watch the part in which it appears? One could answer that one should not take the analogy too literally. But one aim of the present paper is to show that questions of this type are really necessary and cannot be argued away by the limited scope of analogies.

1.3 Qualia

"Consciousness is the way information feels when being processed" asserts Max Tegmark, the self-proclaimed "fundamentalist" in a recent paper on math, mind and matter [5]. But why should information "feel" anything when being processed? What does that mean? A similar question appears in the famous article "What is it like to be a bat?" [6], where it was asked why and how the information processing in the brain of a person/animal is accompanied by something "what it is like to be" that person/animal. I think the ongoing qualia debate which followed the article can essentially be boiled down to the following question: Is consciousness identical to the information it contains (or processes)? That which is hypothetically missed when one reduces consciousness to its information content is called qualia. I would like to ask a similar (but rhetoric) question: Is the monitor of my PC identical to the information contained in the picture displayed on it? I will make use of an analogy in which consciousness corresponds to the monitor, qualia corresponds to the picture displayed by it, and the information that can be extracted from the qualia may be isomorphic (in some sense which remains to be determined) to some processes in the brain.

A physicalist's objection to this analogy could be that there is a big difference between consciousness and a monitor: There is much more to say about a monitor than about its picture. I can talk about its cabinet, for instance, or about the chemical structure of its display. In the case of consciousness, there are no such things to talk about apart from its content. So, in contrast to the monitor, one could say that consciousness is isomorphic to the information it contains, and, as the fundamentalist of [5] asserts, "if two entities are isomorphic, then there is no meaningful sense in which they are not one and the same." The problem of such an argument is that it relies on a misconception of information. It assumes that information is something that is objective, that can stand alone. But the

crucial point is that **information is only information with respect to someone who decodes or interprets or observes it.** For information to appear, there are two entities necessary: a carrier and an interpreter. Information arises when the interpreter (decoder, observer) interprets (decodes, observes) the carrier. (I will explain this in detail in sections 2 and 3.) So, if nobody else looks into my brain, who is the interpreter of my brain processes? Who makes the brain processes contain specific information by looking at them?

Subjectively, I can clearly say that there is not a mathematical structure but a very concrete picture of what I am seeing displayed in my consciousness, which is much more than the information that can be abstracted from it, such as a monitor displays in a very concrete form the information it receives through the cable [2]. I cannot prove this to you, and you can take my statement as a pure physical output of a physical entity, but you can check your own consciousness to realize if it contains such a picture, too. In this way, although qualia are always subjective, at least their "existence" (not of a particular quale, which is quale to only one subject, but of qualia in general) becomes intersubjective.
Here I intend to present a point of view in which consciousness is part of a generalized form of "mind" which is a "hardware" that runs the physical world as a kind of "program" on it.

2 Matter as Software

The loss of substance in matter has often been viewed as a consequence of the strange properties of quantum mechanics. In fact, I think that it is not so much a consequence of quantum mechanics but of the scientific methodology in general. The loss of substance would have also occurred if physics had remained perfectly classical (if a purely classical world should be possible). It is just that quantum mechanics (and even more the relativistic underlying quantum field theories) brilliantly demonstrates how far this can lead us away from our solid everyday experience of the world.

In quantum mechanics, the "solid" atom has to be replaced by "wave functions" or abstract states in Hilbert spaces whose precise interpretation is still a matter of debate. But it is still a fact that the particle unambiguously "exists" (at least in one particle quantum mechanics, in a simple interpretation). The wave function just reflects the fact that it does not have a well-defined position. The situation gets worse in quantum field theory on Minkowski space. Now there can be, for instance, an overlap between a state in which a particle is somewhere and a state in which there is no particle at all. So, even the existence of a particle becomes a matter of tendencies and probabilities. Next, in quantum field theory on curved spacetime, there are situations in which a given state is seen by one observer as a vacuum (i.e. no particles), and by another observer as a thermal mixture of particles. This implies that the probabilty for the presence of a particle is now

[2] An example that is used in several places for experience beyond information content is our perception of time: "Our experience of time is that of a moving present, one that is very different from a tick mark somewhere halfway a time axis in a frozen four-dimensional spacetime" [7]

also a matter of the observer's perspective. Finally, from a yet to be discovered theory of quantum gravity, one might expect even worse effects.

All the physical entities of our everyday experience are completely deconstructed while we proceed to more and more fundamental levels of physics. There is nothing left which fits our usual ideas about the term "hardware" or "substance". What remains are only mathematical constructions.

I will now argue that even without quantum mechanics the physical world is necessarily a mathematical structure. The argument is similar to the one used by Max Tegmark [8, 9, 5], with slightly different connotations. For a detailed explanation of the term "mathematical structure", I refer the interested reader to Tegmark's work.

For the argument I will discuss a situation involving the famous Alice and Bob. Assume that Alice and Bob want to construct a scientific theory. Such a theory has to describe something objectively. A property can only be objective if it can be communicated. The content of Alice's consciousness (call it qualia or not) is subjective and cannot be transferred to Bob as such. But Alice can extract certain aspects of her perceptions and bring them into a symbolic form (words, numbers, etc.) that can be communicated to Bob [3]. Bob receives the message from Alice and understands it, which requires that they have a common code, for example a common language. If Bob is then able to extract the same symbolic pattern from his own perceptions, they would call the pattern an objective fact.

Example: Alice and Bob are standing in a garden. From her perceptions Alice extracts the picture of a red flower. She says: "The flower in front of us is red." Bob understands her because he also speaks English and has learned what a flower is and what the color red looks like. From his own perceptions he can also extract the picture of the flower and thereby he verifies Alice's statement. Hence they conclude that the sentence "The flower in front of us is red" is an objectively true statement. This is only a simplified description of how objectivity comes about. Things become much more complicated when Alice and Bob make "equivalent" experiments at different places and times, or if they try to construct objective statements about "all" objects of some kind. Or if more people take part in the communication who do not speak the same language.

The crucial point is that what is objective and therefore communicable is always a pattern of symbols.

To construct a scientific theory, Alice and Bob need to find many objective statements and relate them to each other. These relations are necessarily also patterns of symbols. The theory that arises is thus a structure of patterns of symbols. Such a structure is always equivalent to a mathematical structure.

The "physical world" is what remains when one isolates the objective patterns from Alice and Bob's perceptions and their communication process. The physical world *consists* of these patterns. It is therefore a mathematical structure.

If the universe consisted of classical billard balls instead of quantum particles, these billard balls can be described by some mathematical equations. And since there is nothing more to say objectively about these balls apart from the sets of numbers assigned to them

[3] Accentuation and facial expression are hereby not considered as scientific communication. A smile is in some sense a statement but not a scientific statement.

by the mathematical description, one can say that the balls *are* nothing else but these numbers.

A non-mathematical hardware (or substance), can be assigned only on an intermediate level, to things that have not been analyzed yet. When I write a sentence on a piece of paper with my pen, I can consider the ink and the paper as hardware, as substances which carry the information contained in the sentence. But when I take a closer look at the ink and the paper, I find that both are patterns of molecules, the molecules are patterns of atoms and the atoms consist of wavefunctions. The substance is lost.

To summarize the points made so far: *My world* is an interpretation of my first-person experiences. The *physical world* is a particular kind of interpretation of the first-person experiences experienced by different people who communicate with each other. The physical world is, in the first place, a description, not a "real thing" to be described. The only "real thing" we have at hand is first-person experience itself. The physical world can be considered as a model for the structures which can be "extracted" from this experience. As such a model, it can necessarily be reduced to a mathematical structure. This is true completely without referring to QM and its interpretation.

We can still ask whether the physical world, despite being only a description of "real things" in the first place, can also be a "real thing" itself. There are basically two possibilities:
1) One could say that only our observations are real, and the mathematical objects (which constitute the physical world) are just tools in order to describe these observations economically, without any deeper meaning or reality. This is the Positivist's point of view. In this case there are no objective physical entities left at all. In some sense, the physical world has no "existence".
2) Or one believes that the mathematical constructions are something that is "really out there", that exists independently of us. This is the Platonic point of view. In this case the physical world *consists* of the mathematical constructions.

In my point of view, it is not meaningful to speak of "physical entities" which are more than mathematical structures. Such entities would be more than what could objectively be described about them. But why should we call them "physical", then? The "physical world" was designed to describe what is objective in our experience. An electron, for example, is just a term for some specific physical information. The word "electron" refers to certain aspects of certain first-person experiences experienced by certain physicists. It appears very implausible to me that it refers at the same time to some mysterious electron *an sich* (electron as such) in a Kantian sense. Therefore, I do not add to the list above a third possibility which involves such blown up "physical entities".

In fact, I *do* argue that there are "entities" which are more than mathematical structures. These entities are the first-person experiences themselves, which constitute consciousness. I would not call these entities "physical" but "mind-like".

Given the two possibilities listed above, what conclusions do they allow for Physicalism? Physicalism asserts that mind consists of or can be reduced to physical entities. In

the Positivist's point of view, there are no physical entities at all, so Physicalism cannot be true.

The Platonic point of view asserts that mathematical entities are substance enough. The structures exist independently from the process that led to their discovery. A mind consisting of physical, i.e. mathematical, entities seems therefore possible.

If one agrees with the Platonic point of view, then there is also a possibility to solve questions like "Why is our universe the way it is? How did it come into being? Why is it described by this mathematical theory and not by another one?" Tegmark [8, 9] suggested an answer based on what he calls the "Democratic Principle" and the "Level IV multiverse" of mathematical structures. The Democratic Principle asserts that **all** mathematical structures have an existence on their own. The set of all these structures is the Level IV multiverse. Some of the structures have so-called "self-aware substructures" (SAS) which play the role of observers in their specific "universe". Our universe is suggested to be just a typical example of these SAS-containing structures.

Nevertheless, something is going terribly wrong. When we find a structure, this is always a part of some "communication" process, as was demonstrated above with Alice and Bob. The information is at first transmitted from the flower to Alice's eyes, carried by light. Then it is transformed into Alice's perceptions. When carried by the perceptions, it is interpreted by the analytic part of Alice's mind and thereby transformed into words. The words carry the information to Bob who interprets it by transforming the words into his imagination of a red flower.

A structure appears only in a process, where information is *carried* by something and *interpreted* by something else, and thereby brought into a new form.

This is the problem of a Platonic fundamental theory. There is no hardware left which could carry the structure defined by the theory. When we think of mathematical structures, we do so always in terms of words or other symbols. **Even mathematical structures need at least symbols to carry them, and at least one mind to interpret these symbols, and something different cannot even be imagined.** So where do the symbols that carry the structure that is our universe come from and what do they look like (what is the film reel made of)? A universe made of symbols is absurd, but it is a consequence of the isolation of the structure from the perception and communication process from which it arose. This isolation was illegal. A world consisting of pure objects, i.e. pure structure, is not well-defined.

Another problem is the Platonic use of the word "existence". I think the use of this word is not well-defined here. I will come back to that point in section 6.

3 Mind as Hardware

Any step by step evolution process inherent to some mathematical structure can be understood in terms of information processing. Since the physical world is a mathematical structure, Physicalism requires that consciousness (which, from a physical point of view, evolves in physical time) is something that can be described in this way. The processing takes place in the brain, and the relevant information is represented by firing and non-

firing neurons/synapses. Consciousness is (for a physicalist) a mathematical substructure of the physical world. It is *identical* to the information processing which defines the substructure.

In his "Story of a brain" [10], Arnold Zuboff takes the idea that conscious experience can be understood in this way *ad absurdum*. Although it might not have been his intention, the core of his argument is essentially the missing external interpreter of the brain states. In the brain, states are represented by firing and not firing neurons and synapses. It is usually assumed that the "firing" or "non-firing" is the only information relevant for the conscious experience, i.e. any furhter physical information about the neurons/synapses can be ignored. Even if this were not true, the argument below could be easily extended, since consciousness would (for a physicalist) *still* be information processing; one would only have to incorporate *more* information.

If an experience is given by a state/process in the above sense, it is not necessary that it is represented by neurons and synapses. An arbitrary representation, made of arbitrary components of arbitrary size and location would lead to the same experience, as long as (roughly speaking) the relation between the Ones (representing a firing neuron/synapse) and Zeroes (representing non-firing neurons/synapses) remains the same. This statement is called Functionalism. There could be very strange realizations [11], but how strange they may be, Functionalism tells that the system has the same conscious experience that a brain process of the same structure would imply.

In his story, Zuboff takes these "strange realizations" to the extreme, although the components are still neurons. These neurons are in several steps dislocated from their original brain and excited by hand, but in such a way that the pattern of excitation induces a specific experience. The people who perform the manipulation believe that the experience is felt by the total set of neurons which originate from the brain of a particular person, and the set is still identified with that person. In a next step, some of the neurons are substituted by other neurons, and finally one arrives at the conclusion that each neuron can be taken to be just one of the neurons in some other person's brain. So the strange realization of a process in one brain is given here by a set of neurons distributed over many other brains. One of the figures in the story even concludes that it is sufficient to excite one neuron of each type a single time, generating thereby all possible experiences. Brains would therefore be unnecessary, a single excitation simulates all imaginable processes of any brain at once.

Dennett and Hofstadter [12] reject Zuboff's argument in the following way: If Zuboff were right, they say, one could in a similar way argue that all books are unnecessary. It would be sufficient to print the alphabet a single time, and all possible orderings of the letters and therefore the content of any book would be already present in that alphabet. With the arguments of the previous sections it is easy and also illuminating to see why Dennett and Hofstadter are wrong: The books are read by someone who is outside these books. The reader receives the information contained in the books in a specific way, he has learned to follow a specific order in reading, e.g. to start at the upper left of the first page, read line by line until he gets to the lower right, then he turns to the next page and so on. So the words and sentences of the book have to be ordered in that specific way. Just writing the alphabet is not sufficient.

Such a necessity of following specific orders does not exist for a conscious experience, because Zuboff started from the assumption (which he wants to refute) that the functional state of the brain (the ones and zeroes given by the neurons and synapses) is itself the conscious experience, without an external interpreter.

I agree completely with Zuboff's argument. I did not repeat it here in its entirety, instead I would like to give a slightly different perspective on it, making the connection with our discussion in section 2 more explicit. Imagine that I want to generate a binary code for texts written in upper case letters (26 letters and a space character). The most obvious way is to take blocks of five bits, using for instance 00001 for A, 00010 for B, 00011 for C etc. Any other ordering would do as well. I could also be less efficient and take 10 for A, 110 for B, 1110 for C, 11110 for D, etc. Instead, I could also use codes not for single letters but for combinations, e.g. 10 for BERT IS EVIL, 110 for BLNGZUFHJ, ... This would be extremely inefficient, but it is possible. Decoding means, in this case, to regain the letters from the ones and zeroes. What series of letters a certain sequence of ones and zeroes stands for depends on the code, and without knowledge about the code I cannot even say whether a specific sequence stands for a short or a long text.

Now let's return to the strange realizations of conscious experiences. What does functionalism mean? If I assume that an experience is given by the functional structure of a specific brain process, I need a method to check whether another process is also a realization of the same structure. One has to transform or "decode" the other process into the original brain process to check whether they are isomorphic. In Zuboff's first steps, the "decoding" happens just by identifying the neurons and ignoring their spatial distances and orientation. The big question is, what kinds of transformations (decoders) are allowed. The crucial point in Zuboff's argument is that **if one allows any kind of "decoders", then any process is isomorphic to the original brain process** (such as any sequence of 1's and 0's could stand for any sequence of upper case letters).

Take again the fundamentalist's statement: "If two entities are isomorphic then there is no meaningful sense in which they are not one and the same." That sentence is only clear and true as long as one ignores the complications in the question how to *extract* the structure of an entity, and how to *compare* it with other structures. If we should one day meet some extraterrestric intelligent species who uses a completely different system of communication, it might be extremely difficult to judge whether something they "say" is "isomorphic" to something we say. This includes mathematics. Their use of mathematics may be so different from ours that we simply cannot see what corresponds to what. We need somehow to extract the information and transform (decode) it into our language. This process may be full of traps and may include steps which are a matter of taste or belief (belief in that something they mean is identical to something we would mean). The over-interpretation of the measures of the Egyptian pyramids by some occultists may serve as a warning of how dangerous such a translation is.

Again: A structure is only a structure with respect to someone who interprets/decodes it in a particular way. If there are no limits on what interpretations are allowed, anything can be isomorphic to anything. (Of course, SO(3) is never isomorphic to SO(4), but the symbols I write down to define SO(3) might by some maniac be interpreted as symbols defining SO(4)).

A part of my brain processes is "decoded" by my muscles, i.e. transformed into movements of my body, facial expression or speech, which can be seen/heard and interpreted by other people. The problem with my conscious experience is that it is visible only to myself. This distinguishes it from my brain. My brain can in principle be detected by other people, but even if they had full access to it, they could only see/measure the activity of neurons and synapses. When I see an apple, my brain contains only the information of the picture of an apple (and even that only when one interprets the structure of the brain activity in a specific way), but not the picture of an apple, such as my TV cable transmits the information of the pictures, but only the monitor *shows* these pictures.

It obviously requires some decoding process to get from the brain process to my conscious experience, from the pattern to the qualia, and these qualia are obviously (to me) more than the structures which can be extracted from them. It makes no sense to locate the decoding process inside the brain, since the brain is rather what has to be decoded, decoded by something "external". I may define my consciousness as precisely the device which transforms neural information into qualia. Consciousness defined in that way does not think, it only transforms the neural information of thinking into "what it is like to think" for me. I will regard consciousness as the passive part of mind (namely the monitor component), a relation which will be made more precise in the section 5.

This picture may be misleading in suggesting that the brain process is there at first, and qualia follow from it after the decoding process. Our discussion of the emergence of structure from perception and communication (Alice's and Bob's flower), in contrast, suggests something different: It might be a better picture to say that qualia are more real, more fundamental than brain processes, which are "only" structures and therefore only derived from qualia. Here we have arrived again at Schrödinger's problem, the double role of consciousness, but in a slightly milder form: Does the content of consciousness emerge from the brain, or vice versa? (See also ref. [13]). It is milder than the original problem since we ask only about the dependence of the *content* of consciousness, not of the dependence of consciousness itself.

In both ways, consciousness itself is a non-physical "hardware device", needed anyway by the argument of section 2. It is even a fundamental hardware: It cannot be further reduced. Other people cannot see it and therefore not analyze it, and everything I can see, feel or think is only a part of what is shown on its "surface" (the picture shown by the monitor). In contrast, a physical hardware is "hard" only at intermediate steps of analysis, before it loses its substance in the next step, as I demonstrated in section 2.

Before I go on, I would like to link the discussion above to some aspects of the qualia discussion going on in philosophy: One can describe typical qualia thought experiments in terms of the "monitor picture". A "zombie" (someone who has no qualia at all) would correspond to someone whose monitor is turned off, so that the processes in his computer are not visible to him. "Qualia inversion" is more interesting. It would correspond, for instance, to a monitor in which red and green pixels are exchanged, i.e. the picture contains the same information as it would on a usual monitor, but the code is different. If the person in front of the monitor has no other access to these colors, he would not

realize the difference.

4 Two Complementary Reductionisms

Descartes has shown that the first thing we can be sure of is our own mind. While we can imagine that the world around us is an illusion, we cannot deny that there is something *having* that illusion. It is quite natural to imagine a type of reductionism that tries to reduce everything to this most obvious and directly experienced "entity": mind. Or, if one wishes to emphasize the "fluid" aspect: mental processes. I will call such a reductionism a Type I Reductionism, or, a bit provocatively, a *Natural Reductionism*.

Its most impressive realizations are given by philosophies from East Asia, inherent in Buddhism, Hinduism or Taoism. In many of these philosophies, the whole world appears as a process happening on Mind, a kind of illusion (called *maya* in Sanskrit).

An example for the kind of reasoning used in these systems of thought is the Buddhist philosophy of *Pratitya-samutpada* (dependent origination), a beautiful description of the process of coming-into-being: The starting point for everything is ignorance. It is the first of the so-called "twelve links of conditioned existence": With ignorance as condition, mental formations arise. With mental formations as condition, consciousness arises. With consciousness as condition, name and form arise. With name and form as condition, sense gates arise. With sense gates as condition, contact arises. With contact as condition, feeling arises. With feeling as condition, craving arises. With craving as condition, clinging arises. With clinging as condition, becoming arises. With becoming as condition, birth arises. With birth as condition, aging and dying arise.

It is described how an Ego appears at the "surface" of Mind, how and why it suffers, and the conditions for the suffering to be removed.

Many of the systems assert that, at a deeper level, all the different minds of different people turn out to be aspects of one and the same Mind (they are reduced to it); the distinction between the object and subject aspects of Mind disappears on a fundamental level, so that there remains really only one "being", with all structures removed. Furthermore, there can be no "fundamental theory" in such a type I reductionism, since theories, language and concepts in general are part of the processes appearing at the surface or at intermediate layers of Mind and have to be removed when going to a deeper level. All theories have therefore rather pragmatic, intermediate character, being useful in order to give someone a positive attitude towards going on with his practice.

On the "experimental" side, techniques are given to verify the statements summarized above. In order to experience the deeper levels of mind, it is necessary (similar to experiments in physics) to remove disturbing "noise effects", and that is precisely what is done in meditation, when the eyes are closed, silence surrounds us and thoughts are coming to rest.

For some Western type I philosophies, Ducasse [14] has used the term **Hypophenomenalism**: "The living body is a hypophenomenon of the soul or mind or of some constituent of it - an effect or product or dependent of it, instead of the converse of this as epiphenomenalism asserts." Ducasse gives two examples: Plotin and Schopenhauer.

For Plotin, the universe arises from the ineffable One, God, by a series of emanations. The soul is not in the body, but the body is in, and dependent upon, the soul, which both precedes and survives it, and whose forces give form and organization to the matter of which the body is composed. For Schopenhauer ("The world as will and idea"), matter is just an objectification of will; our body in particular is an objectification of the will-to-live.

Husserl's Phenomenology may be also called a type I reductionism. It advocates knowledge based on the process of going back to direct and pure experience with all further concepts and prejudices mentally turned off.

In a type II or *Inverted Reductionism*, the order of reduction is reversed. From that point of view, our mental processes are the most complex of all things that are to be explained, and have to be reduced in a sequence of more and more fundamental theories: experience to psychology, psychology to physiology, physiology to cell biology, cell biology to molecular biology, molecular biology to chemistry, chemistry to quantum mechanics, quantum mechanics to quantum field theory, quantum field theory to a hypothetical yet to be discovered more fundamental theory, and finally, facing the full consequences, the hypothetical yet to be discovered more fundamental theory to the hypothetical Level IV Multiverse of mathematical structures. I.e. everything is reduced to what is *furthest away* from our direct experience.

Today, type II reductionism clearly dominates the scientific and other intellectual work all over the world. It is the present framework for science which held an impressive triumph procession over the last 500 years and grew to a system of extraordinary completeness and consistency. Nevertheless, it is not a *total* completeness and it can never be, as was demonstrated in the previous sections. It is a system of objects. The necessary subject ingredient is missing. One should not conclude from the successes of type II that type I reductionism is less true.

The two types both have their domain of application. Both generate a picture of the world and our role in it. Roughly speaking, type II deals primarily with the physical aspects of the world, type I with the mental ones. Both incorporate the idea that things can be described on different levels, where higher levels can in principle be reduced to deeper levels. It is just that the levels are in opposite orders. Both are *causal*, i.e. processes have consequences which influence later processes (or their likelihoods) in a causal way. In a type I framework, the Sanskrit name for the causality principle is *karma*. Both types are equally useless for predicting the financial markets. Both provide techniques to reach certain goals: Type II leads to technologies for dealing with the world "outside" by manipulating it physically (providing physical devices for communication over long distances, for instance); type I created techniques to deal with oneself, techniques for relaxation, meditation, equanimity, happiness, altered states of consciousness. Thereby, both types can give us an immense increase of power, in type II a more outward-directed power, in type I a more inward-directed one.

The two types of reductionism are complementary to each other. I believe that they are opposite points of view on reality which are somehow simultaneously true.

5 Processing and Observing Unit

When I am dreaming, I see my dream world through the eyes of the dream person to which my consciousness is associated - in most cases a person who is similar to the person I am when I am awake, or maybe to the person I used to be years ago. Sometimes I may have a different name in the dream, or even be a completely different person. But I always observe the dream world from a specific perspective. The rest of the dream world appears to be outside myself, different from me and full of surprises. The scenery of the dream may have regularities and structures that my dream person discovers with curiosity, but it is outside my conscious control. Other people are wandering around and I associate with them their own will, and in fact they may surprise me with actions I did not expect. But when I wake up, I realize that the whole scenery was set up by an unconscious part of my own mind, and that all the different dream persons with their seemingly separate wills originate from the same unconscious dream processing mind of mine.

Similarly, the processing of a computer takes place in the processor. I do not see it, I do not have direct access to its internal procedures. But I have the monitor. The monitor gives me a certain perspective on what is going on inside, it provides me with some of the results of the computations taking place in the unseen processor.

I would suggest a similar model for the "real" world. There is no dualism between mind and matter. There is only mind, and matter is just a part of the information "running" on it. But there is a duality *inside* mind: an unconscious processing unit is opposed to one or many (that question will be addressed in the next section) conscious observing units which "display" a part of the information being processed.

The unconscious processing unit is not "my" unconscious (I have no aims towards Solipsism). Similarly, the unconscious processing unit of a dream is not the unconscious of the dream figure, but of the dreamer. In the "real world", "I" am just one of the figures in the big real world dream. The real world appears very complex and stable to me, so I cannot seriously consider it as being generated by something belonging to my person.

There are several possible attitudes one can have concerning the duality discussed above:
1) One could accept it as fundamental. Such an attitude is presented by the *Samkhya* philosophy, where the conscious but passive *purusha* is opposed to the active but unconscious *prakriti*.
2) One could consider the duality as non-fundamental, as something that makes sense on an intermediate level, but has to be overcome finally when going to a deeper level. This attitude is, for example, represented by the *Advaita-Vedanta* philosophy, in which there is only one fundamental entity: *Brahman*.
I think the question whether one of these interpretations is more correct than the other is still unanswered, but personally I find the second, fundamentally monistic possibility more appealing.

One can speculate how far the analogy between the "real life" and a dream can be taken. What about lucid dreams, for example? A lucid dream is a dream in which you realize that you are dreaming and thereby get control over your dream, a good starting

point for many interesting explorations. An interesting question is how far an analogy of a lucid dream can be realized in the "real world" (a "Lucid Life", so to speak). In order to get the appropriate attitude towards this, it could be useful to start with "Considering life as a dream" experiments such as described in Piet Hut's "Life as a Lab" [15].

6 One Mind - Many Minds

In a dream there is only one mind responsible for the entire dream world. In the real world, however, there seem to be many minds, but only one world. Schrödinger, who regards mind in some sense as the creator of the world (as cited in the Introduction of this paper), calls this the "arithmetic paradox" [1]. As a solution he suggests the philosophy of the Vedanta, which asserts that there is really only one being. (I mentioned this as the second possibility in the previous section). All the seemingly different minds are in fact only one mind, and this oneness is not only theoretical but something that can be experienced, either spontaneously or by "training" (meditation). As a first hint towards oneness, Schrödinger writes, one may take the fact that consciousness is always experienced in singular, never in plural. It is not even imaginable to experience more than one consciousness at the same time.

The oneness of the Vedanta must not be mixed up with Solipsism. Solipsism asserts that *my* mind is the only one, and the other minds do not *exist*. It says that the entire world is something that *exists* only in *my* mind. For justification, Solipsists might refer to Ockham's razor: Conceding existence to other minds independent from my own mind is an unnecessary assumption that is not needed for a complete picture. I can directly experience my mind (or consciousness, or qualia), but not the mind of other people. But being more careful, one gets to an even more drastic conclusion. I can only experience my mind *now*. Access to the past is given only by my memory now, to the future only by my anticipation now. So by solipsistic argumentation, I should conclude that only my mind exists, and that it exists only now. My mind did not exist five minutes ago, and it will not exist five minutes from now. **Mind is always One, and it is always Now**. The difference between the transformation from me now to me in the past or future and the transformation from me to someone else is that the first transformation is *continuous*, whereas the second one is *discrete*. This makes it intuitively easier to identify myself now with "myself" in the past rather than with another person now.

I think the solution to the solipsistic problem is that "existence" is not a meaningful notion. "Existence" only makes sense when followed by a specification like the term *"within"*. A square root of -1 exists within the complex numbers, but not within the real numbers. When we talk about something that exists, we usually mean that it exists within the physical world. As I explained, I do not think that mind exists within the physical world, I rather think that the physical world exists within mind. So what could it mean that mind, or *a* mind, exists? Other minds do not exist within my mind, but what could it mean that they do or do not exist *per se*, without any specification "within" what they exist? Or what could it mean that the real world exists or not, without the addition "within my mind" or so. The notion of existence alone simply makes no sense.

Therefore one can not meaningfully ask if other minds exist or not.

In the Vedanta philosophy, the reality of individual minds is only a very relative one. The seemingly different minds are merely different reflections of or perspectives on the one and only being. A similar statement holds for the "same" mind at different times.

7 Models and beyond Models

How can we combine the results and reflections from above into a complete picture, a "model"? Here I will suggest a sequence of models, starting with the type II reductionism of today's science as the A-model. The A-model is enormously successful but incomplete, as was shown in sections 2 and 3.

A minimal extension of the A-model would be to leave it unchanged but add one processing and many observing hardware units. I call this the B-model. The processing unit runs the physical world as its software, the hypothetical fundamental theory of physics being its machine code. The observing units are the consciousnesses. They display a part of the information processed by the processing unit, namely information associated with the brains of conscious beings [4]. So far the B-model is reminiscent of the situation in the movie "The Matrix", where humans live in a world presented to them by a big computer program. Note that consciousness is here at the same time the display, that which watches the display, and the process of watching. The difference between the B-model and "The Matrix" is that consciousness is completely passive in the B-model while there is some free acting possible in "The Matrix".

The B-model is an improvement with respect to the A-model, but not yet very satisfying. As far as consciousness (and not the processing unit) is concerned, it is similar to Epiphenomenalism and shares similar difficulties. If consciousness is non-physical and completely passive, how can my physical body know something about it? How can my body write a paper on it? A solution could be that the processing unit also runs a kind of "BIOS program", a software component that knows something about the hardware. Nevertheless, I regard such a construction as a bit artificial.

Furthermore, when we think of consciousness literally as something that just "decodes" information, i.e. brings it from one form into another, we run into the same problems as with the fundamental theory in section 2. The statement that "consciousness is at the same time the display, that which watches the display, and the process of watching" does not really make sense, such as it didn't make sense for the wave function of the universe.

The problem with the B-model, I think, is that the whole world-process is still accounted for by the type II reductionism, while the mental "hardware" is implemented only at the endpoints. A further improvement, a C-model, would shift the whole world picture further to the middle between type II and type I, with subject and object more tightly entangled. The emphasis would be a bit less on structures and a bit more on sensation, feeling and will, or on "Thinking beyond pattern", as the philosopher and psychologist Eugene Gendlin calls it [16].

[4] For the Many Worlds Interpretation of Quantum Mechanics, the observing units display only the information corresponding to one of the many-world tracks of the brain.

The construction of a C-model is far from trivial, in particular from the viewpoint of the present state of scientific language and paradigms. But it is in my opinion the next important task for science. The theory of objects (the A-model) is in a state where it is essentially complete on the intermediate layers where it is consistent (most parts of physics, chemistry, biology) [5]. Its hardest remaining problems lie in the boundary regions where it *must remain* incomplete by the arguments of sections 2 and 3: 1. Fundamental Physics (including the interpretation of Quantum Mechanics) and 2. Psychology.

In the Outlook I will say a bit more about what kind of work the possible construction of a C-model will require.

Finally, it may be that all models are misleading when one approaches fundamental aspects of reality. Mind is that which is beyond structure ("beyond pattern"), and so are our qualia. But a philosophical system or any other kind of model is itself a structure and so I have doubts if the connection between mind and matter, between subject and object can be described by any model or any language. (This is in agreement with type I reductionism.)

8 Conclusions and Outlook

8.1 Conclusions

Altogether, I have basically presented six ideas:

1. Structure (or information) is always part of a process in which something carries the structure, and something or someone else interprets it by "decoding" it. It is objective in the sense that it is intersubjective. It is intersubjective in the sense that different subjects decode it in the same way.

2. Physicalism reduces the world to a mathematical structure which does not allow for any carrier or external interperter. Physicalism is therefore inconsistent.

3. The decoding of brain processes into qualia requires a non-physical entity. The only physical decoding of brain processes is that which presents an "output" to the environment via a response of the body. Consciousness is therefore a non-physical fundamental hardware.

4. There are two complementary reductionisms: one reduces everything to fundamental physical entities (which are essentially mathematical structures); the other one reduces everything to immediate experience from which the world is created as a kind of mental activity.

5. Mind consists of a conscious observing and an unconscious processing unit, like in a dream.

[5] By "essentially complete" I mean that no changes in the basic pictures are expected. Only the details and exploitation issues are still under study.

6. "Existence" is, when standing alone, not a meaningful notion, so the question if other minds than my own "exist" is not well-defined.

I agree with Arthur Eddington [17] who once compared mind with water, and matter (or the physical world) with waves. The ocean consists of water, not of waves. The waves are just the form in which this water appears to us at the surface. Similarly, reality consists of mind, not of matter. The entire physical world is just the form in which mind appears to us at the surface.

8.2 Outlook

Our desired task, the construction of a C-model, requires that we take the role of the subject much more seriously than in present science, which is a science of objects. In contrast to P. Hut, who makes a similar suggestion e.g. in ref. [5], I do not think we are starting almost from zero.

There is already an immense amount of knowledge present for the "science of the subject". One of the problems is that we are not very familiar with it, because it is not part of our education. Another problem is that the knowledge is spread over many different disciplines and approaches who did not and do not regularly communicate with each other. They have very different paradigmatic backgrounds, and a systematic, unified picture is not available:

- We can probably learn very much from the mystic and spiritual traditions around the world and their associated philosophies.
- The history of philosophy is full of ideas towards our desired direction, for example Schopenhauer's Hypophenomenalism or Husserl's Phenomenology.
- Some psychologists offside the mainstream have also interesting approaches. The work of C. G. Jung is probably the most impressive and prominent among them.
- One should also mention Gendlin's philosophy and psychology of "Thinking beyond pattern" and "Felt Sense". In my opinion, being "beyond pattern" is precisely what carries mind over matter.
- Also interesting in our context is Sheldrake's hypothesis of formative causation [18], which asserts that the laws of nature are not so fixed, i.e. they are habits rather than laws.
- The controversial consequences of Quantum Mechanics led to many discussions on the role of the subject, for example between Heisenberg and Bohr.
- Even more interesting is the philosophical work of Eddington and Schrödinger which I have cited here in several places. They show that Quantum Mechanics is not really needed to see the importance of the subject in a complete picture of the world. They go further than any other physicists I know in their conclusions that mind is the essence (or vessel) of everything.

All these puzzle pieces have to be critically reviewed with the right mixture of scepticism and openness. One should let them reflect each other and see if a unified picture is possible.

An absolutely necessary requirement is the development of a more elaborate language regarding mental issues. Words like "Mind", "Consciousness", "Soul" are used very vaguely, with nebulous connotations depending on who uses the word. Singular is often used where plural is meant. Many aspects of our "inner life" cannot be expressed sufficiently at all. The experiences described in the mystic and spiritual sources have to be translated somehow to our "English-speaking minds". I think such a more elaborate language will automatically emerge when more research is done in these directions.

Since we are dealing with subjective experience, the most important things to do are on the practical side. To get into the right attitude, one could start with playful experiments which aim to relax our mental habits concerning subject and object, such as suggested in P. Hut's "Life as a Lab". Meditation practise is certainly very helpful, and also experimenting with Lucid Dreams, or with Gendlin's Focusing method. As soon as some experience in these directions has been made, words concerning our inner life will get new meanings and associations. They will make *more sense* to us in a very literal way. Then we will be able to communicate these issues on a higher level; and to understand what others have already written.

For me the guiding question through this labyrinth of Mind, Matter and Meaning is Schrödinger's problem: the strange double role of consciousness.

References

[1] E. Schrödinger, "Mind and Matter" (Cambridge Univ. Press, 1958)

[2] A. Eddington, "The Nature of the Physical World" (MacMillan, 1928)

[3] J. Jeans, "The Mysterious Universe" (Cambridge 1930)

[4] H. Nicolai et al, "Loop Quantum Gravity: An Outside View", Class. Quant. Grav. **22**, R193 (2005)

[5] P. Hut, M. Alford and M. Tegmark, "On Math, Matter and Mind", http://www.arXiv.org/physics/0510188 (2005).

[6] T. Nagel, "What is it like to be a bat?", Philosoph. Rev. **LXXXIII 4**, 435 (1974)

[7] P. Hut and B.v. Fraassen, "Elements of Reality", J. of Consc. Stud. **4**, 167 (1997)

[8] M. Tegmark, "Is 'the theory of everything' merely the ultimate ensemble theory?", Annals of Physics **270**, 1 (1998)

[9] M. Tegmark, "Parallel Universes", in *Science and Ultimate Reality*, J.D. Barrow, P.C.W. Ellis, C.L. Harper (eds.) (Cambridge Univ. Press, 2004)

[10] A. Zuboff, "Story of a Brain", in [12].

[11] N. Block, "Troubles with Functionalism", in *Perception and Cognition: Minnesota Studies in the Philosophy of Science, Vol IX*, C.W. Savage (ed.) (Minnesota University Press, 1978)

[12] D. Dennett and D. Hofstadter, "The Mind's I: Fantasies and Reflections on Self and Soul" (Basic Books, 1981)

[13] P. Hut and R. Shepard, "Turning 'The Hard Problem' Upside Down and Sideways", J. of Consc. Stud. **3**, 313 (1996)

[14] C. J. Ducasse, "A Critical Examination of the Belief in a Life After Death", (Charles C Thomas Pub Ltd, 1961)

[15] P. Hut, "Life as a Lab" (2004), http://lab.kira.org/lab.

[16] E.T. Gendlin, "Thinking beyond Patterns", http://www.focusing.org/tbp.html

[17] A. Eddington, "New Pathways in Science", (Cambridge Univ. Press, 1935)

[18] R. Sheldrake, "A new Science of Life" (Blond & Briggs Limited, 1981)