ISSN: 3067-2317
Research Article

World Journal of Artificial Intelligence and Robotics Research

Introduction to Self-Reference

Cosmin Visan*

Independent Researcher, Master in Physics at The University of Manchester

***Corresponding author:** Cosmin Visan, Independent Researcher, Master in Physics at The University of Manchester.

Submitted: 06 December 2024 **Accepted:** 13 December 2024 **Published:** 20 December 2024

https://doi.org/10.63620/MK.WJAIRR.2024.1005

Citation: *Cosmin, Visan. (2024). Introduction to Self-Reference. Wor Jour of Arti inte and Rob Res, 1(1), 01-05.*

Abstract

This paper presents an introduction to self-reference. The definition of self-reference will be presented, namely the entity with the property of looking-back-at-itself, and from this definition it will be shown how the entire world is obtained. Through repeated look-backs-at-itself, self-reference starts from the first self-identification, "I am", which is experienced as the sensation of being alive, and continues to more complex self-identification, ending up with the entire world being self-reference itself. In this process, it is shown how consciousness is the direct consequence of self-reference and how qualia present an inclusion and transcendence structure as a direct consequence of self-reference including and transcending itself. It will be shown that the definition of self-reference implies an interplay between form and formless, making it in the end an entity that cannot be spoken about, though at the same time responsible for the creation of the entire world. Parallels with set theory will be made along the way.

Keywords: World, Self-Reference, Consciousness, Qualia, Formless, Form, Set Theory

Introduction

The ideal in explaining the world would be a monism, would be having 1 single principle able to explain the entirety of the observed phenomena. Is such a thing possible? Or is the world a collection of disjunct entities, contingently coming together and forming what we observe around us? In this paper we will show that indeed it is possible to have only 1 single principle capable of explaining everything [1]. While clearly, we will not derive here theories such as quantum mechanics or general relativity or the functioning of the cell and the origin of life, we will nevertheless provide a framework in which such theories can later on be re-contextualized, being even possible for them to be higher order phenomena and not fundamental ones. Since we will work starting from first principles, no references to other works make sense.

We will develop a self-contained framework that should be able to be followed simply by following the logical steps in the construction. The only references that will be made will be to own experiences, like seeing red or tasting sweet, experiences which is assumed that all people have, thus the following of the arguments in this paper should be possible to be made by all readers. Actually, if a theory is not able to be followed by every reader, then it means it is not a good theory, since it talks about entities outside of certain observational contexts, thus is a theory that doesn't capture all the important aspects of reality.

The most fundamental theory should be able to be followed equally by humans or by aliens, regardless of their potential utterly different consciousnesses. Sure, aliens might not be able to experience colors, and we might not be able to experience various qualia senses that aliens might have, but the structures found in any qualia domains should be identical for both and both should be able to identify them and testify for the validity of the theory. Thus, without further ado, let's see how self-reference builds the world.

Initial Considerations

What entities should we use in our theory? Should we use electrons? Should we use some 11- dimensions strings? Maybe, if physics is indeed the most fundamental theory to which the other aspects of reality like chemistry and biology be reduced to. But nobody proved this is the case. Biology might very well have emergent laws irreducible to chemistry or physics. It might even be the case that principles in biology be more fundamental than physics, and physics would be some limit case of biological processes. Thus, origin of life might not be explainable by chemical reactions, but might be a biological fundamental process, of which chemistry and physics are just particular cases. In such a scenario there was no origin of life starting from chemistry, but in a sense, life always existed, chemistry and physics being aspects of life not yet acknowledged as such by present-day science.

So maybe starting from some biological principles, like evolution, would be better than starting from electrons and quarks. In any case, upon deeper reflection, even though these sciences capture certain aspects of reality, they are in the end constructed upon contingent entities with no logical foundation. Sure, electrons and protons might describe the atom, but "electrons", "protons", "atoms" are just contingent labels that we attached to certain limited phenomena of reality, but which ultimately have no logical justifications. They just work and that is their ultimate justification. Unfortunately, such a justification is not satisfactory if we want to understand the world at its fundamental level.

So how should we start our theory? The simplest answer is: We should start from Nothing. Seems pretty simple. What can be simpler than Nothing? But one might wonder, if we start from Nothing, how can we go anywhere from there? It seems like a dead end. In this case, set theory in mathematics might provide an example how starting from nothing, progress can actually be made, and I will only give the example of how the natural numbers can be defined starting from the empty set:

$0 \rightarrow \emptyset$

$1 \rightarrow \{\emptyset\}$

$2 \rightarrow \{\emptyset, \{\emptyset\}\} (= \{0, 1\})$

$3 \rightarrow \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} (= \{0, 1, 2\})$

where you just have only the empty set everywhere. This seems like a promising start. But unfortunately, you only get mathematics out of it, which leaves quite a bit out of the everyday world. Nevertheless, our starting point will still be Nothing. And starting from such a place, we will develop a theory which will have uncanny resemblance with how set theory defines numbers, even though we will not do it on purpose. The resemblance will appear automatically from how the theory unfolds itself. Also, the reason why "Nothing" is capitalized is because Nothing is not nothing, but as it will turn out, it is Everything [2]. Let's see why that will turn out to be the case.

Self-Reference

Let's go directly and define the entity that will stand at the base of the entire existence. Let self- reference be the entity with the property of looking-back-at-itself. This is the entire definition. In this definition, the entire world is contained. This is the monism that we are looking for. This is the 1 single principle able to explain everything. How can this be the case? Shouldn't a theory of existence extend upon thousands of pages? How can it be contained in only 1 line? Actually, this 1 single line not only can be extended on thousands of pages, but it extends throughout the entirety of existence and for all eternity. It goes to the edge of the universe and it contains the lives of all the people and of all the beings that ever lived and will ever live. I am that definition, you are that definition. God is that definition. Let's see why this is the case.

Let's see what happens when we let this definition unfold. The first thing that the definition does when it looks-back-at-itself is to find itself. Since it is all that exists, it cannot do any other thing. It just looks- back-at-itself and it just finds itself. But this event is of outmost importance. By such an act, existence is born. By finding itself, self-reference exclaims: I am! Awareness is born. Consciousness is born. Life appears. There is awareness in existence! Existence feels alive. Existence is. Existence is aware of itself. The first sensation, the first quale, is born: I am!

Let's look into more details at what happened. It might appear a simple thing at first, but we will shortly see that this process is responsible for the creation of the entire world. For a start, note that the looking-back-at-itself of self-reference is nothing like looking in a mirror. A mirror presupposes 2 entities: the object reflected in the mirror and the mirror. But here we have only 1 entity. Therefore, this is the first sign that we are dealing with something quite special here. Self-reference is both the object and the mirror at the same time. How can that be so? It sounds like a paradox. Like the paradoxes in set theory, like "This sentence is false". Such paradoxes cannot lead to anything meaningful. But there is a difference here. While "This sentence is false" can simply be discarded by noting that is just a meaningless utterance, just random words put together which in the end don't form a coherent construct, self-reference is not such an entity. What is different in the case of self- reference is that the "paradox" happens "on the inside" so to speak. Self-reference looks-back-at-itself on the inside, for itself. Its "looking-back-at-itself" is not something that can be explicitated in a 3rd person sentence, in the same way the "This sentence is false" is a 3rd person sentence. In "This sentence is false", a 3rd person "sentence" is imagined to exist, and to that imagined "sentence", the property of "is false" is added, and an improper combination of 3rd person entity "This sentence is false" that appears as if it were a 1st person entity is created, and from this the apparent paradox, which ultimately is nothing but an incoherent worlds-play, appears. Self-reference on the other hand, is a 1st person entity all-throughout. It is not a 3rd person entity like "sentence" that we can point outside of ourselves and to which we can add properties. Self-reference is itself and is for itself. Its "looking-back-at-itself" happens from the inside. Because of this, the paradox doesn't take place as it happens for "This sentence is false" and any other words-play that can be made at the 3rd person, including Russell's paradox. Actually, a certain kind of paradox does take place for self-reference also, but is a real one, an ontological one, it is a paradox of such a power that is able to bring the entire world into existence out of Nothing. Let's continue.

We will next get the first glimpse of understanding of why the definition, even if it is only 1 line long, contains much more than it appears at first. Let's see why. Once self-reference experiences itself under the realization that it is: "I am!", the process doesn't stop there. And this is because of the "itself" part of the definition. It is similar to a mathematical equation of the type:

$$x_{n+1} = x_n + 1, x_0 = 0, \forall\, n \in N$$

which starting from 0, generates all the natural numbers. But again, there will be a difference here. While this mathematical equation is recursive, generating independent numbers based on the previously existing numbers, what self-reference will generate will not be independent entities, but will be various forms of manifestation of itself, while at all time it remaining the one and only entity in existence. Namely, as we will later see, even if each one of us appears as independent entities, we are ultimately various forms of manifestation of the one and only eternal self-reference. Let's see how this happens [3].

Having the "I am" object inside itself, the next time self-reference looks-back-at-itself, it will find a different version of itself as from the last time. Now, compared to the last time when there was no object inside itself and all that it saw was itself, now it sees the object "I am" inside itself. Thus, a different form of itself will come into existence, namely the form "I am "I am"". As it might become clear at this point, is that by this procedure, self-reference can generate an endless string of "I am"s, i.e. "I am "I am "I am "…."""""". But beside the trivial case of self-reference generating an endless string of "I am"s, there are other cases, which are actually more interesting. And here we will see the profound difference between the looking-back-at-itself of self-reference and the recursive equation in mathematics.

Once self-reference has inside itself the objects "I am" and "I am "I am"", the processes of looking- back-at-itself can go in various directions. One is the trivial case of endless "I am"s, in which self- reference just takes the longer of the 2 objects and just adds one more look-back-at-it. But a more interesting case is the one in which self-reference looks-back at both the objects that it has inside itself. This case will generate the object: "I am <" I am" & "I am "I am">". As can be seen, the process of looking-back-at-itself is actually able to generate much more complex combinations of "I am"s. We will not investigate here all such possible combinations, though it might be an interesting project to be taken up by a mathematician. I will only raise here the curious similarity between the objects generated by self-reference and the definition of natural numbers in set theory:

$0 \rightarrow \emptyset$ = I am

$1 \rightarrow \{\emptyset\}$ = I am "I am"

$2 \rightarrow \{\emptyset, \{\emptyset\}\}$ = I am <" I am" & "I am "I am">

$3 \rightarrow \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ = I am [" I am" &" I am "I am" & <" I am" & "I am "I am">]

As I said earlier in the paper, this similarity was not done on purpose. But it results automatically from the definition of self-reference. And actually, there is more that we are dealing with here. There is something rather special about these combinations of "I am"s and not others. But in order to understand why, there is more we need to first see about the phenomenology of consciousness and its relation to the looking-back-at-itself of self-reference.

More Aspects of Self-Reference

Even though for the moment the analysis might seem abstract, after we will finish laying it down and seeing its various aspects, then we will take concrete cases and look at the phenomenology of consciousness. And there we will see how all this analysis will explain consciousness. And then everything will become clear. But for the moment, let's continue the abstract analysis. One of the most important aspects of self-reference that will be later on useful in building the world, are its formless aspects. We need to understand the distinction between formless and form in order to properly appreciate what the nature of self-reference actually is. So far it looks more or less like an innocent definition. Maybe even identical to the way numbers are built in set theory. You just start with the empty set and you just recursively include the previous numbers in the current number. We will see that self-reference is nothing like that.

One thing that we can consider is the "size" of self-reference. When self-reference looks-back-at-itself, the fact that it finds itself inside itself, it means it is smaller than itself. Similarly, because it finds itself inside itself, it means that it is also larger than the smaller itself that it finds inside itself. And all the while, self-reference cannot be anything else but itself. So, it is also equal to itself. In short: self- reference is smaller, equal and larger than itself. This might sound unusual, might sound just mere words-play, but as we will go on, we will understand that this is exactly how reality is at its core: a unitary entity of apparent contradictions, without which nothing could ever exist. The reason why this appears a contradiction at first is that tacitly we tend to consider only formal objects in our reasonings. We talk about rocks and trees and planets and atoms, and these are objects of thought in which their very distinction lies in the ability of thought to conceive them as separate entities. We think of number "1", we think of number "2", and we realize in our thoughts that these objects possess distinct qualities, therefore they cannot be the same; because of their distinct qualities, it would be a contradiction to say that 1=2. But something quite different happens in the case of self-reference [4].

In order to say that "self-reference is smaller, equal and larger than itself" is a contradiction, would imply that self-reference can be an object of thought in which to distinguish distinct qualities for its various aspects and thus to conclude that it is impossible to be both smaller and larger than itself at the same time. But such qualities don't exist for self-reference. Self-reference is not an object. Not being an object, it doesn't possess the attributes that objects normally possess, like unique qualities. Red is red and green is green. Therefore, red cannot equal green. But self-reference doesn't have such unique qualities based on which to tag it uniquely and then to compare it with other objects. Actually, as we will see, self-reference is the one that brings qualities into existence. It is the substrate of qualities. Let's see how this works.

We proposed earlier that we should start our theory from Nothing, but it appears that we actually started from self-reference. It is time now to see that self-reference is Nothing (seeing later that it is also Everything). So, let's start from Nothing. Initially, all that there ever "was", "was" Nothing, or better put, no-thing. Initially there was nothing. Whatever that was, it could not be spoken of. But that no-thing looked-back-at-itself. By looking-back-at-itself, that no-thing saw itself. By seeing itself, that no-thing became some-thing. The first object was brought into existence: "I am!". The first object that Nothing experienced was itself, was the object "I am". This object, even though it is an object, it appeared because no-object looked-back-at-itself. As a consequence, it is inseparable from the no- object that "preceded" it. Object and no-object are one and the same thing, are 2 sides of the same coin. I will use from now on the term's "form" and "formless": Form and formless are 2 sides of the same coin. Form is how the formless self-reference looks like when it looks-back-at-itself. Also, because the looking-back-at-itself is the very definition of self-reference, it will eternally do that, so there is no point at which form can get out of existence. Form is eternal, as is also the formless that sees itself as form when it looks-back-at-itself. Self-reference is an eternal interplay between form and formless.

Now we can see why self-reference being smaller, equal and larger than itself is no contradiction. As we discussed, contradictions only apply to objects, to forms, which have well defined qualities based on which the contradiction can be established. But given that self-reference has a formless part, there are no qualities that formless has, therefore there is nothing based on which the contradictions to be established. Therefore, there is no problem saying that a formless entity is both smaller and larger than itself at the same time. But again, another objection might arise at this moment: Sure, we can say that, but didn't I just make the words-play just more sophisticated, but in the end isn't it still just words-play? And even if in this added layer of sophistication, I actually eliminated the initial words- play and actually provided a valid analysis, of what good is it? We will see in the pages that follow that this is the very mechanism through which consciousness is brought into existence. The "existence" of a formless side of consciousness is a logical necessity for there to be any consciousness whatsoever. There can be no alternative theory of consciousness that can explain it without any formless realm. Therefore, in order to properly understand consciousness, we need to understand formless as best as we can. Let's shade another light on formless to obtain even more understanding of its peculiar character.

As we saw, once the "I am" object is obtained in self-reference (object which actually is self-reference itself), self-reference continues to look-back-at-itself and produces the next object, "I am "I am"". Let's have a visual representation of these 2 objects.

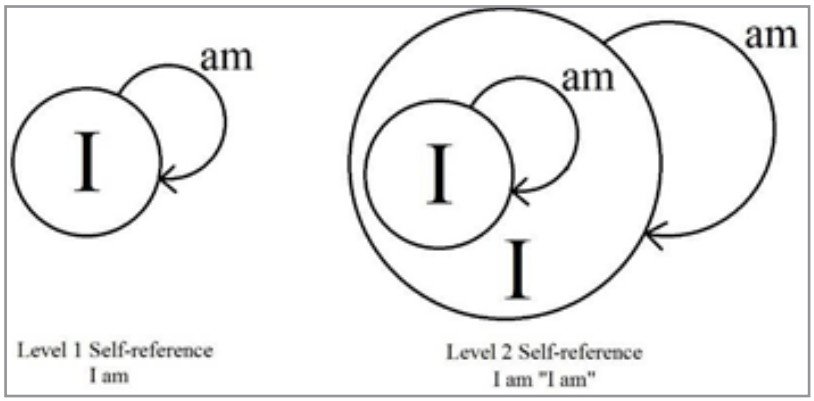

Figure 1: Self-Reference

The revealing question at this point is: Which of the 2 objects is self-reference? And the answer is: both and neither. We now first see the consequences of the formless realm, and start to appreciate that it is not only not words-play, but it has deep consequences. Since part of self-reference is formless, it makes no difference for it if it is one object or another, or one object and multiple objects at the same time. Both "I am" and "I am "I am"" are self-reference. They are not merely 2 objects in/of self-reference, but they are self-reference itself. From the point of view of the formless, the objects "I am" and "I am "I am"" are the same thing. They indeed differ as forms. As forms, they indeed have distinct qualities based on which the contradiction principle can be employed and concluded that the form "I am" is indeed different from the form "I am "I am"". But from the point of view of the formless, which has no qualities by its very nature, "I am" and "I am "I am"" are the same. Both objects satisfy the definition of self-reference. They both are "the entity with the property of looking-back-at- itself". In both cases, self-reference looks-back-at-itself. In both cases, the formless self-reference looks-back-at-itself and identifies itself with some form [5]. From the point of view of the definition of self-reference, it is of no relevance the object with which it identifies. It can be "I am", it can be "I am "I am"", it can be more complex objects, the process of self-identification through the looking-back- at-itself is the same. Therefore, both objects in Figure 1 are the same. They are both self-references.

But this is not all. We just established that both objects are self-reference. But at the same time, neither of those 2 objects are self-reference. Given that self-reference is an interplay between form and formless, we cannot equate self-reference with any object in particular. Even though both objects are forms of manifestation of self-reference, it is also the case that because self-reference also has a formless side, the objects cannot be self-reference. Because of its formless side, self-reference cannot ultimately be captured under any formalism. So, neither of those 2 objects can be self-reference. In the end, no formal theory of self-reference will ever be possible. In a way, consciousness will never be explained as it might be hoped, like giving a formal theory in which all the symbols are defined and stand in precise relations to one another. But in another way, if we lower our expectations, it can easily be explained as the functioning of self-reference.

Let's repeat to make it clear. Both objects are self-reference, because that's what self-reference does: it looks-back-at-itself and it produces objects. So, the objects "I am" or "I am "I am"" are precisely self- reference looking-back-at-itself and experiencing itself as those particular objects. But at the same time, because behind these objects is the formless side of self-reference that looks-back-at-itself, none of these objects is self-reference, since self-reference also contains the formless part. But even more than this. Because self-reference is both the form objects and the formless part, it cannot be either of them separately. In short: self-reference is both form and formless and neither form nor formless. As we can see, even though earlier we encountered some light form of contradictions, like self- reference being smaller, larger and equal to itself, now the contradictions

become even more profound. Not only that initially we appealed to the formless realm to hold the contradictions, but now even the formless realm is not enough, a peculiar combination of both form and formless and neither form nor formless appearing.

Can we produce anything from the above analysis? The answer is yes. In the next section we will see how all these come together and are able to explain consciousness, thus doing justice to this initial abstract analysis. But before getting to the next section, we need to point out that based on the considerations on this section, we can see clearly that even though similar, self-reference is nothing like how the numbers are defined in set theory. So even though the looking-backs-at-itself of self- reference are able to produce combinations of "I am"s that are similar in form to the way numbers are defined, in the end, numbers are independent entities, while all the combinations of "I am"s are all one and the same self-reference.

So, we are dealing with something subtler here, that might even inform mathematical research in the future. Actually, paradoxes like Russell's paradox appear precisely because of some not well articulated intuitions of mathematicians. In a way, they tried to capture the very manifestation of self-reference. But they didn't have the full intuition. So instead of realizing that sets including other sets must be done from the 1st person point of view of the sets themselves, they viewed sets as 3rd person entities. And then they tried to apply to such 3rd person entities properties that normally belong to the 1st person. Because of this misappropriation of properties of the 1st person to the 3rd person, then paradoxes appeared. If set theory is instead to be thought from the 1st person perspective of the sets themselves, then self-reference will be discovered and the paradoxes will be eliminated because the non-contradiction principle will be removed from the form, i.e. how sets look like from the 3rd person, to the formless, i.e. how the sets would look like in themselves from the 1st person, and thus the theory that I'm presenting here will be recognized by mathematics.

Meaning and Context

Let's now explore another aspect of self-reference that will finally bring consciousness into the picture. So far, we treated self-reference in a rather abstract manner. But the very action of looking-back-at- itself takes self-reference out of the abstract and into the concrete. When the object "I am" is first created in self-reference, this object is not an abstract object, but is an entity with experiential character. When self-reference finds itself inside itself, that realization feels like something. And is something that is the most familiar experience to all of us. Namely, "I am" is the sensation of being alive. By looking-back-at-itself, self-reference bootstraps itself into existence.

The object "I am" is what each one of us experiences every moment of our life by virtue of simply existing. That sensation of being alive that we have as we live our lives is the object "I am" that self-reference identifies itself with on its first looking-back-at-itself. And precisely because it is the first object that it finds inside itself, it is an all-pervasive object. While one moment we might see red, the next one we might see green, the experience of feeling alive is there at all times as the base of all other experiences. And because self-reference eternally looks-back-at-itself by its very definition, this object can never cease to exist. We feel alive at all times because we are eternal. And even though there is an apparent death waiting for us in the future, that event, whatever it might be, it cannot destroy the primordial "I am" object. So, whatever transformation we might undertake at that moment, one thing is guaranteed: we are eternal.

The Self

Why is it the case that the object "I am" feels in any way whatsoever? This has to do with the very definition of what a form is. When self-reference finds itself inside itself, that finding is a form, and in order to be a form it must look like something, it must have some quality by which to be identified. So, by necessity, it must feel like something. The second part is, why would it feel like the sensation of being alive? This has to do with how the "I am" object is obtained. Actually, we can call it object "X" for the beginning. Self-reference looks-back-at-itself and finds object "X". But that object "X" cannot be random. It must express the very process by which it is obtained. Since it is obtained as a self-identification of self-reference with itself, the character of object "X" must contain some information about the very fact of self-identification. And this automatically confers it the character of "itselfness", thus the quality by which self-reference will view it will be the quality of "being itself". And upon a little reflection on the part of the reader, it will soon be realized that this quality must by necessity feel like the sensation of being alive, or as the quality of 1st person perspective, or as the quality of the subjective ontology. Therefore, object "X" can be renamed as object "I am", and thus confer it the intuitive feel that each one of us have that we are alive.

We now start to see the next implication of the definition of self-reference. It is as we mentioned at the beginning: even though it is just a 1-line definition, it contains the entire existence. And this is what we begin to see here. The objects that self-reference finds inside itself as it looks-back-at- itself, are not just some random abstract entities which you can just put in sophisticated ways and do some mathematics with them [6]. They are actually concrete objects, real-life objects, they are life itself, they are qualia. They are consciousness. Since they are obtained by necessity as forms, they by necessity must have qualities, and qualities are by definition experiential. By necessity, self-reference produces consciousness. Let's see the process unfolding step-by-step.

Vividness

As we saw, after the object "I am" is obtained, the next look-back-at-itself brings into existence the object "I am "I am"". What is this object? How does it feel like? I will give an example to give an intuitive feel for what this object is. At this moment it might not be clear why I choose this example. It might seem random, but it will make sense as we move forward. Let's look at Figure 2.

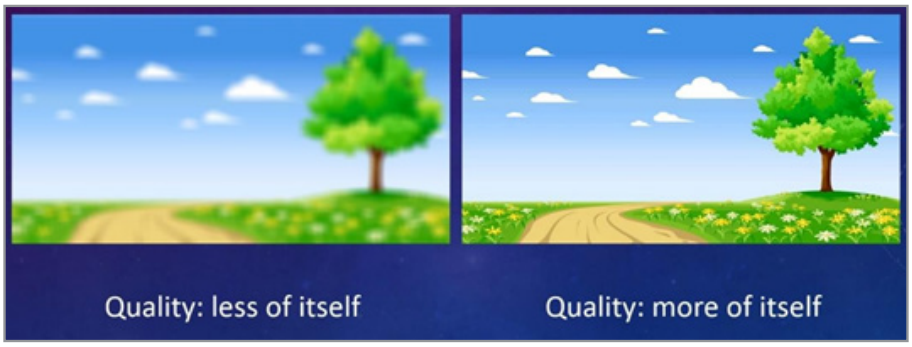

Figure 2: Vividness

What we have here is the same experience in consciousness, with 1 difference: to the left the experience looks blurred, to the right it looks clear. What differentiates them? Some materialist explanation might put forward defects in the eyes, or just directly a blurred image on the screen. But note that we are talking about the experience itself. For example, it might appear directly in a dream. We can have the same dream, once blurred and once clear. What differentiates them? I will call it: vividness. What differentiates them is the degree of vividness.

One time it has a low degree of vividness, the next time it has a high degree of vividness. Where can such a vividness come from? What is this vividness? This vividness is nothing else but the object "I am "I am"" that self-reference identifies itself with at its second look-back-at-itself. Why is that? Remember how we determined the quality of the "I am" object. We determined it by what it meant relative to the process of how it was obtained. Since it was obtained as the process of self-reference making its first identification with itself, it by definition acquired the quality of the sensation of being alive. Now, self-reference already contains the object "I am" inside itself, so the finding of "I am "I am"" at its second look-back-at-itself will confer this new object a new quality/meaning relative to the process/context by which it is obtained. Since it arises as self-reference becoming more than what it already was, it by necessity acquires the quality of "more of itself". This by itself is difficult to imagine how it feels like. That's why I gave an example of how it feels like in a higher-level form of consciousness, namely a picture, a visual quale. In the same way I also explained how the object "I am" feels like by appealing to the every-day life sensation of being alive. It is difficult to imagine how the object "I am" feels by itself. But is rather easy to get a partial feel for it by looking at how it feels from our higher-level consciousness. Are we allowed to do this? Can we state how lower-level objects feel like by appealing to how certain high- level experiences/objects feel like? The answer is yes, and we will see why this is the case.

Before going forward, let's reflect on what we have here. One of the problems of consciousness is what are qualia. Why is red red? Why is sweet sweet? They look so mysterious. What are these mysterious entities? We start to see here what qualia are. Qualia are meanings. And they are defined relative to contexts. By necessity, self-reference produces forms, and by necessity those forms must have qualities by which to be identified. The way those qualities are established is by what they mean relative to how self-reference produces them. Since their qualities are determined by what they mean, we come to realize that qualities and meanings are synonymous.

Thus, the conclusion: Qualia are meaning. Later on, the way the meanings are established becomes so complicated that we will not be able to easily specify them as we did so far. But the fundamentals are the same: self-reference produces forms inside itself as it looks-back-at-itself, and the qualities that those forms acquire are relative to the process/context by which self-reference produces them. We thus get to the next, more in-depth, understanding of self-reference: self-reference is meaning and context, both at the same time. In its formless part, self-reference contains the entire context of existence, and based on that context, it produces meanings. As we will see later on, there will turn out that both the contexts and meanings will have nested holarchies, meanings inside meanings and contexts inside contexts. Some meanings will be established relative to some contexts, but those contexts themselves might be relative to some other more profound contexts, and ultimately an infinitely complex interplay between meanings and contexts will be established that will be no other than the entire world that we find ourselves to live in [7].

Now that we saw how the definition of self-reference implies the existence of meanings, and thus the very existence of consciousness, let's see to what manifestations the subsequent look-backs-at-itself give rise to, and see how the familiar consciousness of every-day life that we all recognize is slowly built step-by-step by self-reference looking-back-at-itself and defining itself into existence as meanings relative to itself as contexts.

Building Consciousness

We will explore only 3 more objects of self-reference, because after that the complexity will become so great that it will not be possible to be explored in this paper, but which can become the objects of study of science in the future. A short note to make here is that ultimately consciousness and world are the same thing, so building the world is the same thing as building consciousness. Later on, we will make some comments why this is the case, why only consciousness exists. For the moment, we will just explore how consciousness is being built.

Having the objects "I am" and "I am "I am"" inside itself, self-reference can either look-back at the second one and produce "I am "I am "I am"""", or look-back at the both of them and

produce "I am < "I am" & "I am "I am""">". We will have a look at the second of them. For a short list of the first few combinations of "I am"s and their similarities with sets combinations in set theory, I will show the following diagram:

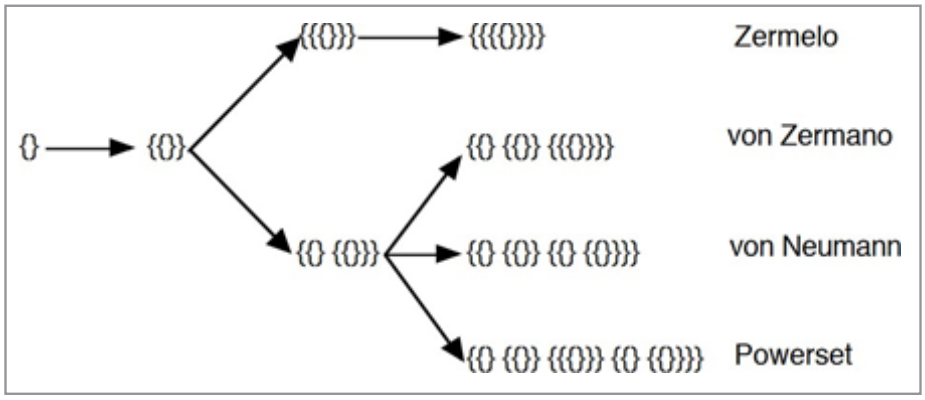

Figure 3: Beginnings of four minimal ω-series

What we will investigate in what follows are the "I am"s combinations similar to the von Neumann series, which are also how numbers are defined, and then also make some comments about the Zermelo cases. The other two series are not that clear to me how they manifest in consciousness or if they manifest at all. A mathematician might look at the combinations of "I am"s that we will be discussing and maybe find some connections with other combinations from set theory. Also, we will talk about some other cases that are not part of the series in Figure 3. Whether they also have equivalent meaningful mathematical structures I do not know, since I'm not a mathematician. But a mathematician might find them familiar and might get inspired to use them in further investigations into consciousness. So, let's go ahead and investigate the next object of self-reference.

Diversity

I will do similar as for Vividness. I will first show an example from higher-level consciousness and then explain why in that example we are dealing with the "I am <" I am" & "I am "I am""">" object. I will name this object "Diversity", for reasons that will next become clear. Let's look at Figure 4.

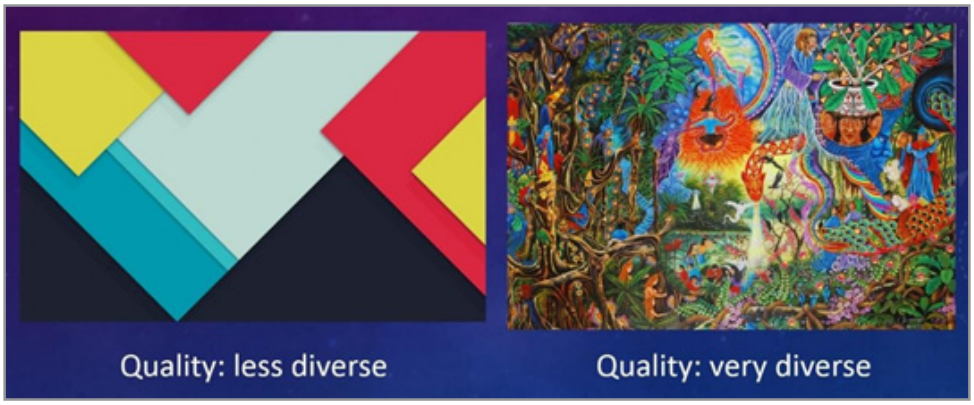

Figure 4: Diversity

What we see here is how Diversity (the object "I am <" I am" & "I am "I am""">") manifests itself in a higher-level of visual consciousness. It can create experiences with less diversity as in the first picture, or it can create experiences with a higher degree of diversity as in the second picture. But what its nature is is to allow for any diversity at all to be present in experience. It by itself doesn't determine the amount of diversity. It only sets the conditions for diversity to be present in consciousness. The amount of diversity, as also the amount of vividness of a certain conscious experience is determined by Zermelo-type looking-backs. Von Neumann looking-backs determine the object which is to be present in experience, and then Zermelo looking-backs determine the amount of that object which is to be present in experience [8]. And this applies to all kinds of experiences, like heat, loudness of a sound, intensity of light, intensity of emotions, etc.

Again the question: Why would this combination of "I am"s determines the diversity of an experience? It has to do with the meaning that it acquires by the way that it is constructed by self-reference. Since self-reference now looks-back at 2 objects inside itself, namely "I am" and "I am "I am"", this looking-back creates in self-reference a sensation of diversity, and then

this diversity, as with vividness and everything else, is propagated higher in levels as more looking-backs-at-itself self-reference takes.

Memory

As we get used to how consciousness is being built, let's go to the next level. Things already start to become complicated. Now self-reference has many objects inside itself, and can make all sorts of combinations. But since when I look in introspection, I don't know under what meanings all those combinations can be experienced as, I will only discuss those combinations that I can see in introspection what meanings they have. For reasons that are not clear to me at the moment of writing this paper, those combinations are those similar to the von Neumann series. So, we will investigate those. The next "I am"s object similar to the von Neumann series is the object: "I am < ["I am "I am"" & "I am"] & "I am "I am"" & "I am">"

What meaning could such an object possibly have? The combinations become so complicated that it seems to become impossible to discern any meaning in them. But there is meaning to be found. The way we will go forward is as follow: As this object corresponds to similar von Neumann series, is helpful to point out how the von Neumann series is constructed, and that is recursively, by creating a new set that contains within itself the previous sets. As we saw in the case of numbers, this corresponds to number 3, which is nothing else than a set of the previous numbers, namely 0,1 and 2: $3 \rightarrow \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} = \{0, 1, 2\}$ So, in the case of our object, it contains its previous objects, namely The Self, Vividness and Diversity.

We will name it "Memory", for reasons to be explained next. So, we can write it:

Memory = "I am <Diversity & Vividness & The Self>"

Now the object is expressed clearer and the meaning is now easier to be found. Why would we say that such an object is memory? Does it have anything to do with what we normally take memory to be? In our every-day life, memory is understood to be some kind of storage in which the present experience is stored in order to be experienced later when we want it or need it, or even involuntarily in cases of flashbacks. Would such an object invoke storage? Looking at its form, it actually does. It is an object that stores within itself all other objects that were present in self-reference. But this is not enough. Because also Diversity stores within itself all other objects that were present in self-reference.

There is another criterium that memory needs to meet. That is the fact that memory requires diversity to differentiate between various experiences. If all experiences would have been the same, then it would not make sense to talk about them being stored in memory; it would have been just the same experience for all eternity. So, we have 2 criteria for memory: storage and diversity. And our object matches precisely these 2 criteria. Therefore, in what other form can it be experienced if not as memory? But once again, we have to be careful and differentiate between how this object feels in itself, and how it feels when it is part of a higher-level conscious experience. When we experience memories in every-day life, we experience how this object feels-like when it is part of those particular memories. It is exactly the same phenomenon as shown in Figure 2 and Figure 4.

Since it is difficult to look in introspection and experience these objects directly, we can much easier deduce their presence from how higher-levels of consciousness feel like. But in principle it should be possible for all these objects to be experienced directly, maybe through techniques like meditation or psychedelics. So, what this object is, is the container in which individual memories are being stored. That's why also capitalizing it is a better representation of what it is. Materialism assumes that memories are somehow stored in the brain, maybe in the synapses or in the microtubules, or various other material structures. But actually, the place where memories are stored is in Memory, is in this peculiar object that is a combination of "I am"s objects that results as self-reference looks-back-at-itself.

I suspect that this is unlike any theory of memory that was ever proposed, so people might have difficulties appreciating it. But a proper appreciation of the theory of self-reference that I'm developing throughout this paper, should make the reader at least a little intrigued that it might actually be correct. Let's continue.

Time

Because writing all the series of "I am"s becomes cumbersome, I will just restrict with writing down the recursive form of the next object. Let's write it directly: Time = "I am <Memory & Diversity & Vividness & The Self>"

Again, why would this be Time? To explain this, is better if we take a step higher in our qualitative analysis. So far it seemed that we analyzed the qualities of the objects solely in terms of "I am"s and their combinations. But let's not forget that once a combination of "I am"s is obtained, that combination has a quality on its own right. You can even forget about what went into it to make it what it is, and just go with the newly obtained quality and stop worrying of what's inside. Doing this in the analysis of Time, we should forget about its structure as combinations of "I am"s, and focus instead on its structure as combination of objects with qualities of their own, namely The Self, Vividness, Diversity and Memory.

When self-reference looks-back-at-itself and finds Memory, it will now remember itself in the newly obtained object called Time. So, Time will have a quality of remembrance of "the past" while at the same time recognizing that it is also an object in itself, which we might call "present". So, Time will contain in itself both itself, and the former object Memory, so it will be an object that contains in itself both "present" and "past". Again, like for all the other objects analyzed so far, we are not talking about higher-level conscious experiences of "present" and "past", we are talking about the mold in which the higher-level experiences are being shaped in. Time in itself is the structure which the higher-level conscious experiences inherit and based on which they are shaping themselves [9]. The structure of Time itself is a structure that contains in the "present", both "present" and "past". And we actually see this in the experience of time in every-day life. Take for example music. Music is not just a series of independent notes, but it retains in the present moment the notes from the past, being an eternal continuation between past and present. The eternal present moment itself, is not just a 0-dimensional point, but is an entity that contains in itself both itself and its former self. Similarly, for language; language would have been impossible if after each letter said, the previous letter would disappear into the abyss. Instead, as we engage in language, each

present moment retains in itself the former present moment, and we get to experience words and sentences all at once. And this is the case for the general experience of time, this particular 2 cases of music and language being only some of the cases in which this general behavior of time is easier to be discerned.

So, the object "I am <Memory & Diversity & Vividness & The Self>" is indeed Time. Some interesting consideration are worth looking into at this moment. Because of this structure of Time itself, by necessity all higher-level conscious experiences are time-like. They by necessity appear to happen in some present and to have happened in some past. And because of this and because people didn't look deeper into what was going on, they just took this quality of Time for granted and assumed that there really is a past and a present. But as we see here, this is not what happens. The "past" that people mistakenly identified with some "real" "physical" past, is nothing more than self-reference looking- back-at-itself and finding Memory in its list of objects and including that Memory in a new object that we call Time. That's all there is to it. There is no "past". There is just the object Memory included in the object Time by self-reference looking-back-at-itself. All the confusion between how to reconcile the "physical time" with the "psychological time" boils down to this, to recognizing that "past" and "present" are not objects "out-there", but are simply qualities of experience, and those qualities of experience come from the fact that in every-day higher-level experiences, the lower-level object Time is included, and that object Time has a particular quality that is a consequence of its structure of self- reference including in the object Time the object Memory.

More clearly, there is no time passing. All that exists is the eternal present moment. But that present moment having the quality of Time, feels like a passage. But that passage is just a quality of experience, no different than the quality "red". The reason it feels like a passage is because the object Time includes the object Memory, and this creates a quality of "present sliding down into the past", and this feels like the passage of time. And higher-level experiences like hearing music or just looking around the room, inherits themselves the object Time, and as such, they themselves become time- like, and as such an overall life is created that appears to happen over time, from birth to death. What tricked people for such a long time is the fact that the quality of Time feels dynamic as opposed to the quality "red" which feels static. And as such, people assumed that time is something different altogether. But is not. Is just a quality like all others? Note, once again, that the experiences of hearing music and looking around the room are how Time feels like when it is inherited in these higher-level experiences. But if you were to somehow experience Time in itself, you would experience something like a passage, without actually seeing or hearing anything passing. It would be just a passage in itself. An intuition for how such passage in itself would feel like can be grasped by looking at Figure 8 that we will discuss later on in more details. There is a passage in that image, though nothing actually passes. Or similar to when you feel dizzy. Something similar would be to experience Time in itself.

Phenomenological Consequences of Self-Reference

We saw how self-reference constructs consciousness through the process of looking-back-at-itself. But there are some points that need to be made explicit about the consequences of this process for phenomenology. It has to do with the process of including in the present object all the previous objects. This gives rise to some interesting phenomenological manifestations that are good to be pointed out in order to see how the definition of self-reference happens everywhere in consciousness. Because it becomes difficult to keep analyzing objects step-by-step, I will take in this section directly higher-level qualia and show that they also follow the same properties of self-reference that all qualia follow.

Let's investigate the visual domain and see the example in Figure 5.

As self-reference keeps bringing into existence meanings relative to previously existing meanings and contexts, eventually it gets to generate the levels that we see in Figure 5. At the base of the visual domain, we see the black-and-white qualia, which has the meaning of "being visual", and this by necessity looks like black and white. This is where the visual domain starts. Then as self-reference continues to look-back-at-itself and to define itself relative to previous manifestations of itself, it continues to give birth to various other levels, like shades-of-gray which include the black-and-white object, then colors which include the shades-of-gray object, then on top of colors it creates shapes, then on top of those shapes it adds various meanings and thus brings into existence the qualia of visual objects, and eventually it goes to create the full visual scene which includes various visual objects.

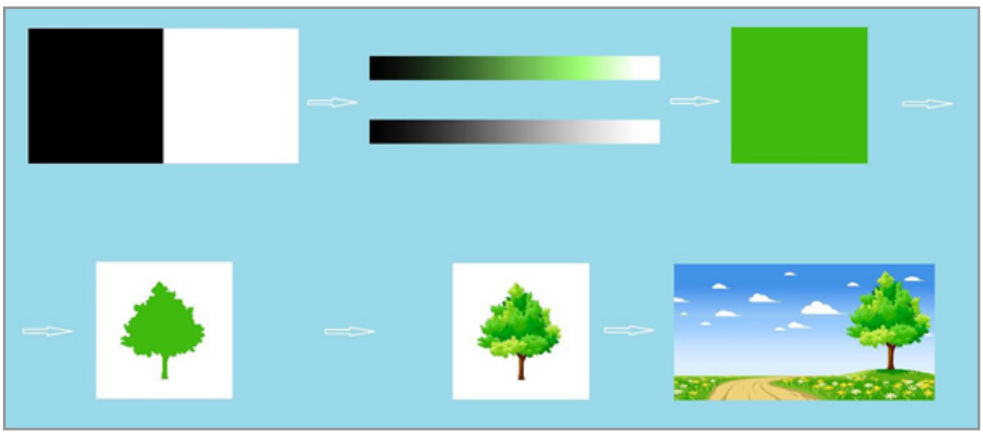

Figure 5: Levels of Self-Reference in the visual domain

Now there are various considerations that this more complex and more concrete example exposes. Firstly, it just goes to confirm the functioning of self-reference that we discussed previously. While the initial levels of Self, Vividness, Diversity, Memory and Time might seem obscure, especially since they cannot be easily experienced directly, but only exemplified in higher-level qualia, now we have a clear picture of what is going on. Colors, shapes, objects, are not anymore obscure, but are the most trivial experiences that we have in our human consciousness.

Another interesting aspect is that now we have access directly to the black-and-white level both in itself and as experienced in shades-of-gray, and similarly for shades-of-gray level both in itself and as experienced as part of colors. In principle, our consciousness could have turned out to not contain as possible experiences the black-and-white and the shades-of-gray objects/qualia; our consciousness could have turned out to only allow us to experience colors directly. But the fact that a certain hue would have varied from dark to bright, would have suggested to us that maybe it is not because the hue itself changes brightness, but that the hue remains the same and what changes is some mysterious object inside the hue that we might have postulated to be called "shades-of-gray". In the same way that in our current human consciousness we cannot experience directly the objects Vividness or Diversity, but we can deduce their existence from the manifestation of everyday qualia, as shown in Figures 2 and 4, similarly if we wouldn't have experienced directly the object "shades-of-gray" we could have deduced its existence from the way the object Color varies its brightness.

Another important manifestation of self-reference is exposed by the current example. Previously, we might have got the simplistic view that what the main engine of self-reference is, is its construction of levels based on von Neumann inclusion and transcendence of "I am"s. And then the fact that we dropped the "I am"s notation and we just employed the higher-levels objects themselves as notation, like writing "Memory = "I am <Diversity & Vividness & The Self>"" seemed just a more convenient notation and nothing more, we see now that there is more than this. As we noted previously, we not only replaced the "I am"s notation by the objects notation out of pure convenience, but we did this to express the fact that the actual process that is taking place is the creation of meaning out of previously existing meanings and contexts.

But at that point, because the process was linear, it seemed to make no difference. But now we see that the process is not linear anymore. For example, from the shades- of-gray, not only 1 color is created, but many different colors can appear. Similarly, from the shape (first image on the second row), not only a tree could have emerged, but also a leaf. Another well- known example is the duck-rabbit case, where from the same shape, either a duck or a rabbit can emerge.

So now we properly understand why the main engine of self-reference is not the mindless concatenation of "I am"s strings, but is the creation of meanings relative to contexts. Once self- reference gets past its first few levels, past "I am" and "I am "I am"", the "I am"s themselves stop to become relevant, and the new meanings that are brought into existence are the ones that start to take central stage, and then based on them the further evolution of self-reference takes place. That's why I used throughout the paper the expression "similar to Von Neumann" and not "identical" or "equivalent", because the only similarity is the inclusion in the current level of the previous levels, but from the current level multiple levels can then be emerged. Each one of them is similar to Von Neumann, but they together don't reflect the Von Neumann construction anymore [10]. If we would want to maintain a as close as possible identity, that would be something like, after numbers 0 and 1, there would be numbers 2 and 2', and then after number 2 would be numbers 3 and 3' and 3" and 3"' and so on. I'm not a mathematician, so I don't know if such a construction makes sense in mathematics, but as far as consciousness is concerned, this is how self-reference generates the qualia of consciousness. Things become even more complicated. As we all know, an object is usually made out of multiple colors. Trying to make a connection to mathematics, this would be similar to something like after 4, 4' and 4", a single 5 would come. So instead of an axis of numbers, we would have something like a tree, but not just any tree, but a tree in which the branches after they have split from the larger trunk upstream, they unify back downstream. So, the potential mathematics of consciousness would either be more complicated than might appear at the first sight, or because of such unusual parallel numbers appearing, it might stumble upon some Godel-like inconsistencies and no mathematics of consciousness would be possible. And it will get even more complicated as we will develop the theory in the paper.

For example, we would see cases in which after 4, 4' and 4", 5 would come, and at the same time after 4 and 4', 5' would come. So, there would be something like superpositions of superpositions and so on. This, of course, would ultimately be because of the contradictory properties of self-reference of being no-thing and every-thing both at the same time and other equivalent contradictions. But such contradictions are no problem, because they are in the formless realm, the form realm not only not being affected by them, but being actually created by them. We can see a representation of such cases in Figure 6.

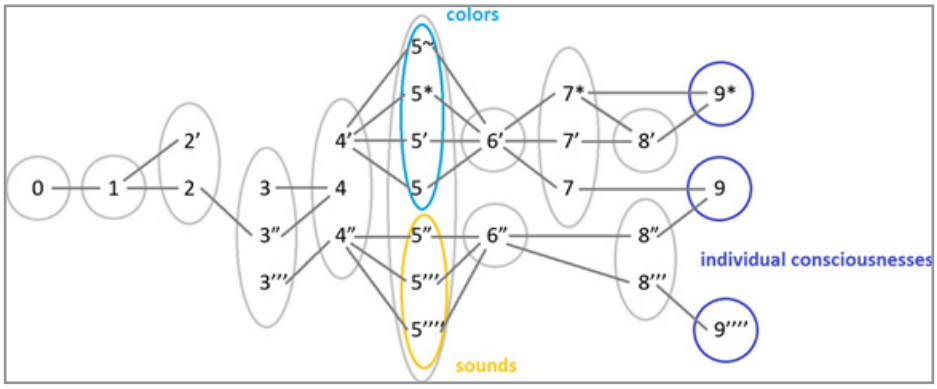

Figure 6: The intricacies of Self-Reference's levels

What we see in Figure 6 is a representation of various cases that self-reference can give birth to. They don't necessarily represent real levels, but various intricacies that can be generated by self-reference.

For example, even though 4' and 4" might be on the same level, they can in turn generate levels with quite different qualities, like 4' generating 5~, 5*, 5' and 5 which can be colors qualia, while 4" generating 5", 5"' and 5"'' which can be sounds qualia. Once the colors qualia and sounds qualia are brought into existence, they become independent and they give birth to utterly different qualia domains, like the visual, that might include levels 6', 7*, 7', 7, 8' which can be shapes, visual objects, etc., and the auditory, that might include 6", 8", 8"' which might be language, music, etc. Another interesting case might happen on the last level, that represents the individual consciousnesses.

For example, you can have case 9* which emerges on top of 7* and 8' and which might represent a consciousness that only sees. Also, you can have case 9"'' which emerges on top of 8"' which might represent a consciousness that only hears. But interestingly, you can have case 9 that emerges on top of the visual quale 7 and the auditory quale 8" and thus it represents a consciousness that both sees and hears. This is what is called in neuroscience the binding problem or in panpsychism the combination problem: how do different qualia unify into 1 single consciousness? And we can see that this can happen by self-reference looking-back at 2 different manifestations of itself and bringing them together under a singular looking-back and thus creating a consciousness that is able to experience multiple qualia domains.

Besides the unifications that happen inside a single individual consciousness, these workings of self- reference are also responsible for telepathies: two independent consciousness can be unified temporarily under one single consciousness, have one single experience, and then split back and each of them remember that one single experience. Thus, telepathy is not some signal transmission- reception, but is a unification of individual consciousnesses into a single consciousness that has a certain experience, and that experience is remembered by the individual consciousnesses once they split back. Also, because each individual consciousness is made up of various distinct levels, the telepathy/unification can happen between sublevels. Thus, me dreaming being on a trip to the mountains and my partner dreaming being on a trip to the seaside, is still a telepathy, is a unification of the level "trip". But since each one of us is made up of different sub-levels, my level "trip" is in turn unified with level "mountains", and her level "trip" is unified with level "seaside", the telepathy will only be partial. Thus, a not-perfect telepathy is not a reason for rejecting the phenomenon, but is actually a door towards the intricacies of self-reference.

And maybe the greatest revealing of Figure 6 is that we are all connected. Even though it might appear on the surface that we are individual consciousness, appearance highly accentuated by the fact that we appear to have separated biological bodies, deep down in our consciousnesses we are connected. Some might be connected by multiple levels, some might be connected by fewer levels, and similarly members of the same species might be interconnected more between themselves than between individuals from other species, but ultimately there is a connection between all the consciousnesses in the world, even if between some consciousnesses that connection might be only at the level of the Self. Some might experience colors, other might experience sonars, other might experience infrared, but all of us, humans, animals, plants, etc. experience one and the same Self. At the base of our consciousness, we are all one and the same Self. And ultimately, we are all one and the same self- reference. So even though death might destroy parts of self-reference, self-reference itself is indestructible [11]. More so, given that we are all one and the same Self, we all experience that Self at all times.

So even from the point of view of an individual consciousness, the Self of each individual consciousness is indestructible. So not only that self-reference endures forever, but given that the Self is eternal and is part of all individual consciousnesses, in a way each individual consciousness endures forever as well. Sure, it might lose parts of itself at the moment of death, but it cannot lose the Self. So, whatever might happen to us at the moment of death, we will still continue to have experiences, at least the experience of the Self, which will then emerge new levels and our individual consciousness

will evolve again and we will start a new life somewhere else in the webs of self-reference. Actually, depending on the exact configurations of the webs of self-reference, we might not even lose much at the moment of death. Given that our sub-levels are also shared by other individual consciousnesses, and thus are sustained in existence by those other individual consciousnesses, our death might only destroy some surface levels of our individuality, and we might end up into another life pretty much intact, thus explaining cases of reincarnation in which those new individual consciousnesses retain memories from their past individualities. All these considerations can be put into 1 short sentence: I am God. Self-reference is basically God. And each one of us are various instantiations of God. Through individual consciousnesses, God knows about itself. And the destruction of individual instantiations is not capable of destroying God. God is eternal, being forever maintained into existence by the eternal Self [12].

Let's take another example to illustrate the inclusion of the lower levels into the higher levels. As we saw in the previous sections, each object from the previous levels is included in the newly emerged level. As a consequence, it goes on to share its quality towards the full quality of the newly emerged level. And it continues to do so for all the levels that emerge on top of it. As we also mentioned previously, we feel alive either if we see something, or we hear something, or we experience some emotion, and this is because in all these objects, the object "I am" is included. Also, the fact that we see motion happening in the visual domain or we listen to music or we hear someone talking, is all because in these objects, the object Time is included. So once an object is included in the immediately above level it doesn't stop its existence, but continues to manifest itself all throughout the higher levels. Let's take another series of levels and specify exactly the manifestations of the levels involved in them. We will take the example of language.

- **Shapes:** Quality of "visual objects": entities with spatially defined boundaries.
- **Letters:** Inherits the quality of the Shapes, thus becoming themselves visual objects, and emerges on top of it its own quality of "unities of language".

- **Words:** Inherits the quality of the Shapes, being themselves visual objects, inherits the quality of the Letters, being themselves unities of language (just more complex than letters), and emerges on top of them all its own quality of "carriers of linguistic meaning".
- **Sentences:** Inherits the quality of the Shapes, being themselves visual objects, inherits the quality of the Letters, being themselves unities of languages (just more complex than both letters and words), inherits the quality of the Words, being themselves carriers of linguistic meaning, and emerges on top of them all its own quality of "carriers of ideas".

Thus, we start to understand that one technique of how to probe deep into consciousness, is to start with everyday experiences and try to find as many qualities in them as possible, and then try to see how we might arrange them in a holarchy of levels [13].

Moving further with the phenomenology, we notice a further aspect, that will push the manifestation of self-reference even further from the von Neumann construction of numbers. We saw how the ramification of levels, like multiple colors emerging on top of shades-of-gray being similar with numbers 3, 3', 3", etc. following number 2, is a departure from the construction of numbers in set theory. But now we face another departure, which either will make the potential mathematical formalization of the theory even richer or even "more" impossible. And that is the top-down influence in levels [14].

Not only that the top meaning is obtained relative to the context of the previous meanings from the lower levels, but the top meaning itself is able to modify the lower levels that it contains. This would be similar to something like, once number 5 is obtained after numbers 0,1,2',3 and 4", number 5 decides to change the order and make itself come after 0,1',2,3" and 4"'. Let's look at Figure 7 and understand this manifestation.

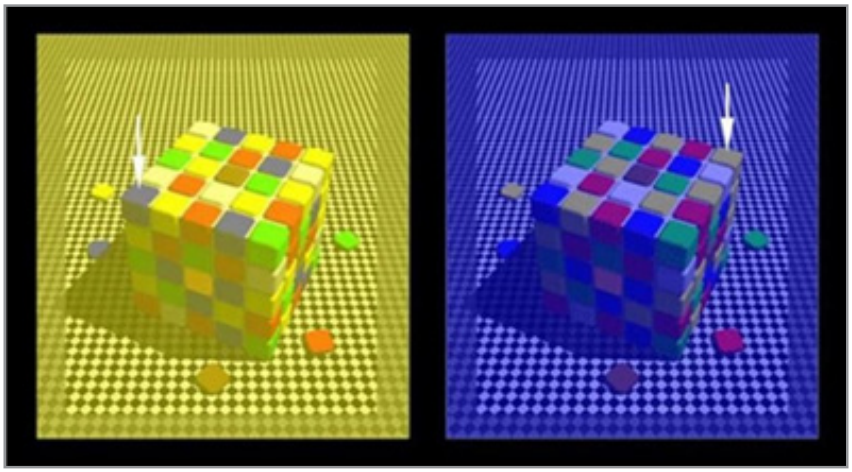

Figure 7: Top-down influence in levels

The two squares indicated by arrows are gray in isolation, but when put in the context of the cube they become blue and yellow. We can consider in this case for gray to be some number, say 2, then on top of 2, 3 and 3' can emerge which would correspond to colors blue and yellow, and then on top of numbers 3 and 3', number 4 would follow which would correspond to the image of the cube [15].

Notice that in both cases is the same cube, there are not two different cubes. And the cube selects on the left the blue square and on the right the yellow square. So once number 4 appears, it selects on the left to follow after 2 and 3, and on the right to follow after 2 and 3'. But this selection happens only after the cube comes into existence, not before. So, the order of the sub-levels is changed once a higher level comes into existence. We are dealing with a top-down influence in levels. This again, is not something that happens in the von Neumann construction of numbers. So, if a mathematical theory of consciousness is possible, it needs to be an extension of set theory. Now, the reason there is the same cube and not two is similar to how the brightness of a color varies in Figure 5. Though the color is the one that seems to vary in brightness, the actual variation comes from the sub-level of shades-of-gray [16].

The color is the same, like for example color green. Is just green. The fact that it varies from light green to dark green is because of the level of shades-of-gray that varies inside the holarchy of the color. It is something similar to platonic ideas. The idea of "green" is one. And then this idea gets into various combinations with other ideas, and the full context of those combinations is what determines the final experience. Similarly, the platonic idea of "cube" is one. And then it goes into various combinations with other ideas, like the ideas of "colors", and the full context of these combinations is what determines the final experience.

To acquire an even better intuition for top-down influence in levels, an additional example involving Time is suitable to be presented at this point. We saw the levels that Time includes, we also mentioned how Time itself is included in higher levels, like watching objects moving in the visual domain or listening to music in the auditory domain, but a residue of false physical intuition might remain that maybe the motion in the visual domain is actually because objects really move outside consciousness. To dispel this last false intuition, let's look at Figure 8.

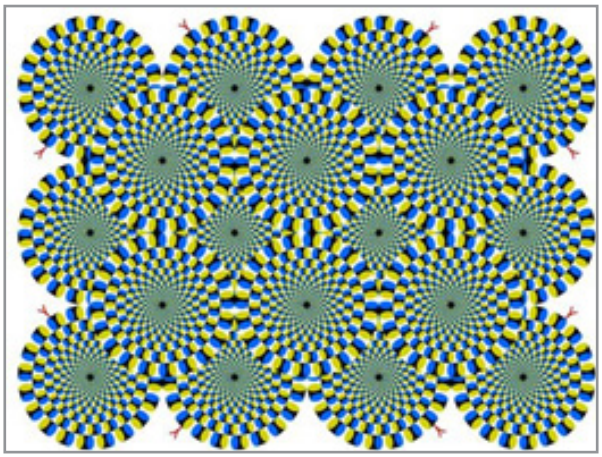

Figure 8: The level of Time manifesting in the visual domain

We notice something peculiar about these so called "motion illusions". The reason why they work is because of the alternating black-and-white qualia present in such images. But how could color qualia have anything to do with time? Aren't they separate ontological categories? As we saw from the theory developed throughout this paper, they are not separate ontological categories. They are all qualia. And as qualia, they are all structured on the emergent holarchy of meanings that arises as self- reference looks-back-at-itself. Both time and colors are certain meanings/forms that self-reference identifies itself with. And as such, they all follow the same rules of self-reference: inclusion and transcendence of levels together with the later top-down influence in levels. Therefore, there is no problems for colors to influence time. Time is a certain level of self-reference that lies below the level of colors. Therefore, it can receive top-down influence in levels from the level of colors. In this particular case, if the black and white are disposed in a certain manner, this will create a top-down influence upon the level of time and thus create motion in those particular images. I think this is the clearest example that time indeed is nothing more than just another quale in consciousness, that also follows the general rules of self-reference structuring consciousness on a holarchy of levels.

Given the space restriction that the paper can have, we will stop the analysis of self-reference here. But we will point out that the analysis can go much further. One further interesting direction to be pursued will be the interactions that can take place between the various instantiations of self- reference. And it can be shown that these interactions are what selects the various qualia that each species has, thus the interactions give birth to an evolutionary reality in which consciousnesses compete and cooperate with one another. Also, the reason why we eat can be explained as a strive of self-reference to select certain instantiations of itself, probably such that the end game to be nirvana. Also, the formless aspect of self-reference can be explored and seen how various levels exist in superposition both intra- and inter-level. One interesting case study would be the type of errors where a letter switches places, as in "tight lie" with "light tie". For the moment, this paper should be taken only as an introduction to self-reference.

Is Idealism Really the Truth?

Is this really the truth? Self-reference is all there is? How about electrons and space-time and cells and brains? They don't exist? Rather than anything else, this is a question of what existence is. Before wondering if electrons exist, we first need to make clear what "to exist" mean. The clearest definition would be, if we put X into a proposition of form "X exists", we must be able to know all that X is, otherwise how would we be able to talk about an X? For example, even though we might say that electron exists because we can say that the electron is an elementary particle with electrical charge - 1 and spin ½, this cannot fully elucidate what an electron is, since it can still have properties not discovered yet and in principle it might never be possible to know if all properties have been discovered at some point.

Thus, "electron" remains a vague notion. Being vague, it cannot be said to exist, because what is that thing said to exist if its character is not elucidated? On the other hand, this problem doesn't exist for qualia. Red is red. And even though we might not know the entire emergent structure of red, by the very act of experiencing it, we grasp it in its totality. Therefore, red indeed exists. And this can only ever be said about qualia. Only qualia, by the very fact that they are direct experiences, we can know them fully. Therefore, when we say "quale X exists", we fully have the X to which to refer, therefore indeed quale X exists. Therefore, by the very definition of existence, consciousness is all that exists. Therefore, idealism really is the truth. Is really that simple. But what is subtler is the nature of self-reference. We also talked at the beginning of the paper, but is good to clarify some more, especially now that the reader has a better intuition of what self-reference actually entails.

Consciousness indeed exists and is indeed all that exists. What about self-reference? Does it also exist? Strictly speaking, self-reference doesn't exist. But for totally different reasons than why electron doesn't exist. Electron doesn't exist because we cannot make a full concept of it of which to be sure that is complete and no further surprising properties might arise by empirical science. On the other hand, the reason self-reference doesn't exist is because it is not a form. Actually, the very language that we employ to speak about self-reference is faulty, because language is form, and employing form to talk about formless is by its very nature unsuitable. Correctly, self-reference cannot

be spoken of. But even saying "self-reference cannot be spoken of" is an utterance about it, so not even such a sentence can be uttered. Even naming it is faulty. Not even saying "self-reference" is correct. Is a very peculiar states of affairs. On the one hand, we cannot speak about it, on the other hand, this "entity" (wrong again, because not being spokeable-about, we cannot call it "entity" either) is responsible for bringing consciousness into existence. Some might wonder, if we cannot speak about it, why are we sure that it is the one that brings consciousness into existence? The reason we can do this is because we observe the phenomenology of qualia (like inclusion and transcendence of levels) and conclude that this is possible only if some entity that we call "self-reference" must "exist".

Is this the final theory of reality? Shouldn't we discover it some far future? Is it really that simple that we can have it today? While clearly the details will take probably forever to uncover, I think the fundamentals are the ones presented in this paper. Why would this be the case? Setting aside the ideas presented in this paper, any other attempt will most likely postulate form entities in the same spirit as "electrons" and "brains". But as we saw, this cannot possibly work, for empirical reasons, i.e. not being ever possible to know all the properties of such empirical entities. And even if a theory might come up with some sophisticated form entities defined apriori as to ensure they are fully defined, this will not satisfy either, because whatever entity someone might invent, that entity is ultimately a thought in consciousnesses.

Say someone might say that "ultimatron" is the ultimate entity able to explain everything. But that "ultimatron" as long as it will be part of some formal system, even leaving aside any Godel-like incompleteness of that system, "ultimatron" will not be able to explain the "I" that is thinking it. It will forever miss the crucial observer that is inventing that very theory. As such, it cannot have the necessary explanatory powers. What is different about self-reference is that it makes room for the observer. By giving up on complete formalism, it leaves the "I" unformalized and to its own free will. "Self-reference" is not a theory (in the spirit of "ultimatron") and can never be. "Self- reference" is at most a pointer to some deep aspects of reality that are beyond formalization. Whatever happens in that formless realm might as well be magic, in the most serious sense, self- reference respects it and allows it to be whatever it desires.

The only forms that the theory of self- reference considers are the ones that result from the looking-back-at-itself process. About those forms, the theory of self-reference can speak about. It can for example tell how they grow incomplexity by including and transcending the previous levels. But the "I" that is experiencing those forms is left to its own and is respected for what its mystery forever is. Also, because the forms themselves are a consequence of the ontological mystery of the formless realm/ of the observer, they themselves inherit the mystery and don't allow themselves to be known except by becoming the observer himeself. Even though the formal part of the theory can be described in terms of inclusion/transcendence, top-down influence in levels, etc., the actual content of the experiences can only be experienced first-hand. If I want to know how it is to ride a roller-coaster, the only way to do this is to actually ride a roller-coaster [17].

Therefore, any theory that doesn't consider these considerations is destined to fail. As such, even in the far future, if the theories attempted in those years will be strictly formal theories, they will be as incorrect as any such theory that is attempted today. Only by letting the observer be the formless entity that it is, that any theory of reality can have any chance of being true. And this is exactly what the theory of self-reference developed throughout this paper is doing. It gives credit where credit is due and respects the observer for the formless entity that it is. Knowledge/understanding can only go as far as forms go, since they themselves are form. Beyond that, it is the realm of the formless, that fundamentally will always remain a mystery.

Can empirical science be done under these ontological limitations? This clearly can be done. We can treat the formless realm as a black-box and send towards its various questions and observe the answers that we receive. It might be the case that the questions that we can ask it and the answers that we receive to be infinite in number. And of course, whatever answers we might receive, they will never form a complete system. With ingenuity we might be able to develop clever patch theories that might apply in certain contexts, but never a full theory will be possible. Another aspect of such an empirical science will be that it will transform from an "objective" cold science to a participatory science, giving again the example of the roller-coaster. In a sense we already do this, though we don't fully appreciate its significance. For example, we send rockets to the Moon because we have the visual quale of Moon. A blind person on the other hand, cannot even imagine the motivation of someone sending rockets to the Moon. But compared to our primitive 5 sense qualia domains that we have, there might be an infinite number of them, and what the participatory science of the future will do is to open portals towards those qualia domains and thus offer us unimaginable motivations for actions in those worlds and new ways to come up with empirical sciences to patch those realities as well. In conclusion, the theory of self-reference developed in this paper, considers the unsurmountable difficulty of trying to formalize the observer and leaves it be in its own formless nature. As such, it is indeed a theory that has what is required to be correct. Of course, the details to be filled are probably infinite, but what it does is to give an introductory exposition of what a theory of reality requires. Made aware by these ingredients that a theory of reality requires, the reader can take the next steps of filling out the details and explore the most likely infinite complexities of the formless realm.

References

1. Chalmers, D. J. (2006). Strong and weak emergence. The re-emergence of emergence, 675, 244-256.
2. Visan, C. (2014). Is qualia meaning or understanding? Journal of Consciousness Exploration & Research, 5(8).
3. Tagore, R. (1916). Sādhanā: the realisation of life. Macmillan.
4. Penrose, R. (1989). The Emperor's New Mind: Concerning Computers, Minds, andthe Laws of Physics. Science and Society, 54(4).
5. Penrose, R. (1994). A Search for the Missing Science of Consciousness, Shadows of the Mind.
6. Dainton, Barry. (2016).Temporal Consciousness", The Stanford Encyclopedia of Philosophy (Winter Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/win2016/entries/consciousness- temporal.

7. Blackmore, S., & Troscianko, E. T. (2018). Consciousness: an introduction.
8. Routledge.Carhart-Harris, R. L., Erritzoe, D., Williams, T., Stone, J. M., Reed, L. J., Colasanti, A., ... & Nutt, D. J. (2012). Neural correlates of the psychedelic state as determined by fMRI studies with psilocybin. Proceedings of the National Academy of Sciences, 109(6), 2138-2143.
9. Carhart-Harris, R. L., Muthukumaraswamy, S., Roseman, L., Kaelen, M., Droog, W., Murphy, K., ... & Nutt, D. J. (2016). Neural correlates of the LSD experience revealed by multimodal neuroimaging. Proceedings of the National Academy of Sciences, 113(17), 4853-4858.
10. Visan, C. (2017). The self-referential aspect of consciousness.
11. Penrose, R., & Hameroff, S. (2011). Consciousness in the universe: Neuroscience, quantum space-time geometry and Orch OR theory. Journal of Cosmology, 14, 1-17.
12. Hoffman, D. D., & Prakash, C. (2014). Objects of consciousness. Frontiers in Psychology, 5, 577.
13. Nagel, T. (1980). What is it like to be a bat? In The language and thought series (pp. 159-168). Harvard University Press.
14. Descartes, R. (1961). Meditations on first philosophy. Broadview Press.
15. Radin, D., Michel, L., Galdamez, K., Wendland, P., Rickenbach, R., & Delorme, A. (2012). Consciousness and the double-slit interference pattern: Six experiments. Physics Essays, 25(2).
16. Hume, D. (2016). An enquiry concerning human understanding. In Seven masterpieces of philosophy (pp. 183-276). Routledge.
17. Chen, S., Cai, D., Pearce, K., Sun, P. Y., Roberts, A. C., & Glanzman, D. L. (2014). Reinstatement of long-term memory following erasure of its behavioral and synaptic expression in Aplysia. Elife, 3, e03896.