Asking physics about physicalism, zombies, and consciousness

Ovidiu Cristinel Stoica
Dept. of Theoretical Physics, NIPNE—HH, Bucharest, Romania.
Email: cristi.stoica@theory.nipne.ro, holotronix@gmail.com
(Dated: November 9, 2023)

If the mind of a sentient being would be reducible to its structure, any system with identical structure should be equally sentient. Based on physics, I prove that this thesis has two unexpected consequences:

1) There would be an inflation of minds, living in apparently different worlds.

2) The content of these minds would be independent of the properties of the external world. That is, minds would be unable to know anything about the world.

Since this contradicts empirical observations, structure alone is insufficient for sentient experience.

This excludes the purely physicalist approaches to physics and consciousness. For physics to be as we know it, all physical properties have to be grounded in something sentiential.

Keywords: Physicalism; observers; philosophy of mind; consciousness; philosophical zombies; no-go theorem

I. INTRODUCTION

This article is about claims that consciousness reduces to the structure of physical substance and its dynamics, without invoking other features of the substance.

A purely physicalist explanation should not appeal to anything that escapes the possibility of empirical observations or measurements that can be independently reproduced and publicly verified. It should ignore the nature of the material constituents, their *ontology*. Ontology belongs to *Metaphysics*, so it makes sense to be considered outside of the domain of Physics, by definition. Physical substance should play exclusively the role of realizing the relations and the dynamics.

And indeed measurements and observations inform us only about the relations, and not about the nature of the relata. They tell us about structures and dynamics, and not about the nature of the substances having those structures and following the dynamics. In the theoretical formulations, ontology may be invoked to ground our intuition into something concrete, but the equations of Physics are blind to it.

Therefore, if Physicalism is true, everything should be captured in the way the systems are structured and how they behave. Consciousness shouldn't be any different.

This position can be formulated in the following way:

Thesis: Physicalism. Any two systems with the same structure and dynamics are equal in every relevant aspect, including their level of consciousness.

This should be true whatever "sentient" means. According to Physicalism, even if the substances making two systems are of different nature, as long as their nature doesn't affect the structure and the dynamics, the differences are irrelevant. Matter or the substance that makes up everything is "phenomenally inert", *i.e.* it doesn't have additional powers needed only to give rise to consciousness. Its structure and dynamics are sufficient to account for anything that exists, this should apply to consciousness as well.

The idea of philosophical zombies was proposed to illustrate the supposed possibility of qualitative differences between sentient and insentient beings, and as a test of materialism (Kirk and Squires 1974, Kirk 2023).

Definition 1. A *philosophical zombie* is a hypothetical insentient entity identical in structure and dynamics to a sentient being.

Since they are identical in structure and dynamics, there is no publicly verifiable difference between a "truly" conscious entity and a philosophical zombie. Therefore, proving the simultaneous existence of both insentient zombies and sentient beings identical in structure and dynamics would refute Physicalism (see Stoljar (2023), §5.1, Kirk and Squires (1974)). It would show that there is more to consciousness than the structure and dynamics of the physical substrate. In particular, it would refute proposals that mind is reducible to computations (Colombo and Piccinini 2023, Rescorla 2020), *functionalism* (Levin 2018), *illusionism* (Dennett 2016, Frankish 2016), and even *identity theories* (Smart 2022) that comply with Physicalism.

In this article I prove that there are insentient systems identical in structure and dynamics with sentient beings. This is a mathematical proof based on standard Physics. It builds on (Stoica 2021, 2023a,c), but it is self-contained. The proof is simple but rigorous. It doesn't appeal to ineffable qualities, subjective experience, qualia, or the *hard problem of consciousness* (Chalmers 1995). I will show that, if consciousness would be reducible to structure and dynamics, we wouldn't even be able to know simple facts about the external world. I realize that this claim is so unbelievable, that at this point many would simply quit reading this article. But since it's a mathematical proof, it can be verified, and the reader can try to find a fatal mistake.

Section §II explains why Physicalism, as defined here, seems to be justified in Physics.

Section §III gives a simple and self-contained mathematical description of physical systems in terms of structure and dynamics.

Section §IV uses this framework to define zombies.

Section §V shows that the relation between the theoretical description of properties and their physical meaning is highly ambiguous.

Section §VI explains how including the observers in the theory partially resolves this ambiguity. But this leads to an inflation of observer-like structures. Are some of them zombies?

Section §VII contains the main result: even when the theory includes the observers, the ambiguity from Section §V leads to the impossibility for the observer to know the world. The fact that we can know the world proves that there must be a sentiential ontological difference between us and most observer-like structures identical to us, so they must be zombies. We are more than the structure. The ontology matters precisely because of its sentiential powers. Physics as we know it can be restored only if it accepts sentience as (part of) its ontology.

Section §VIII discusses several physicalist positions about the world and about consciousness excluded by the result. Only the proposals that accept consciousness as fundamental pass the no-go theorem.

II. PHYSICALISM AND ONTOLOGY

In this Section I will give concrete examples showing that the relational or structural aspects of ontology are the only ones that matter in Physics. This seems to justify Physicalism as the default position.

Example 1 (Relativity of space)**.** Is space absolute or relative (Hoefer et al. 2023)? The relative position and velocity of an object depend on the reference frame. The thesis of *absolute space* states that there are absolute positions and absolute velocities. This requires an absolute distinction between reference frames at rest and in motion. However, even if this distinction were true, no known experiment could detect absolute space or an absolute reference frame at rest. The known laws of Physics don't discriminate among inertial reference frames.

For this reason, physicists adopt the thesis of *relative space*, that there is no preferred reference frame, and no absolute space. Even if space is absolute, this escapes Physics, it's a metaphysical distinction. Whatever the nature of space is, its only role in Physics is to embody relative positions and velocities. Physics is blind to other ontological aspects of space that may make it absolute.

Understanding this led, in conjunction with other postulates, to both Special and General Relativity. □

Example 2 (Relativity of gauge)**.** Classical electromagnetism can be understood best in terms of *gauge symmetry*, which refers to the symmetry of an "internal" space present at any point of spacetime. A gauge is a choice of a basis in the internal space, dependent of the point in spacetime. Even if there is an absolute gauge, Physics is blind to it. This allowed the understanding of the electromagnetic, electroweak, and strong interactions as *gauge theories*. □

The entire edifice of Physics is based on relativity and gauge symmetry. Their independence of ontology seems to justify Physicalism. Let's make the central point explicit:

Rule 1 (No appeal to ontology)**.** Physicalism should not appeal to the intrinsic nature of the physical substance. The only role of matter or substance comes through its structure and how the structure changes due to the dynamical laws.

Experiments that can be publicly and independently reproduced can only give us relations. They tell us nothing about the nature of the relata. All that we can measure, we express as a ratio between the measured quantity and a standard unit. Measuring a position results in a ratio between a distance and a unit length. Measuring the mass of a system results in a ratio between mass and a standard mass unit. In general, all observables are obtained as such ratios between the measured values and the values of similar properties of more familiar objects. When we examine the structure or the composition of an object, we obtain structural relations. When we register events or count particles resulting from an interaction, or molecules that result from a chemical reaction, all of these are relations.

This was remarked in various forms by thinkers like Poincaré (Poincaré 2022) and Russell (Russell 1927), resulting in the thesis called *epistemic structural realism* (Ladyman 2020).

Theoretical models can't go beyond relations either. Any logically consistent theory admits a faithful mathematical model, as we know from *model theory* (Chang and Keisler 1990, Hodges 1997). But mathematical models are mathematical structures, which are sets endowed with *relations* – subsets of Cartesian products of sets (Grätzer 2008).

Dynamical laws are also relational, being about the relations of various parameters or properties across time. Therefore, both experiments and theoretical models deal only with relations.

This makes many scientists expect consciousness to be explained as well without appealing to anything beyond structure and dynamics, for example to the "phenomenal powers" of the physical substance. This is the main reason why it is often considered unscientific to suggest that consciousness may not be reducible to structure and dynamics. But in Section §VII we will see that structure and dynamics alone are insufficient even to allow the knowledge of the value of a single physical property, and ontology will have to be invoked.

III. STRUCTURE AND DYNAMICS

Let us clarify what structure and dynamics are, according to Physics. There are structures in other sciences too, like Chemistry, Biology, Neuroscience *etc.* But any such structure is also a physical structure that follows the dynamical laws of Physics. So I'll explain structures and dynamics in Physics, as simple as possible for our needs.

All possible states of the system form a set $\mathcal{S}$, called *state space*.

At any time, a physical system is in a particular state. The *dynamical law* specifies the state $s(t)$ of the system at the time t in function of its state $s(t_0)$ at a time t_0

$$s(t) = \mathcal{U}_{t-t_0}\big(s(t_0)\big). \tag{1}$$

The change in time of the state of the system is described by a function of time $s : \mathbb{R} \to \mathcal{S}$, called the *history* of the system.

Different states from $\mathcal{S}$ are distinguished by their *properties*. The value of a property consists of one or more real numbers associated with the state. For example, the position of a point-particle in space is a property of that particle, and its value consists of three real numbers, x, y, and z. The coordinates x, y, and z are properties too. This allows us to consider only properties whose values are numbers, without loss of generality.

If the system contains more particles, the position of one of the particles is a property of that particle, but also a property of the entire system. For example, in Classical Physics, the configuration of a system of particles is fully characterized by the positions of the particles.

In general, a property is represented by an *observable*, a real-valued function of the state of the system,

$$A : \mathcal{S}_A \to \mathbb{R}, \tag{2}$$

where $\mathcal{S}_A \subseteq \mathcal{S}$ is a set of states on which the observable A is defined. $\mathcal{S}_A$ may be different for different observables.

For classical systems, every observable A always has a definite value for any possible state of the system, so $\mathcal{S}_A = \mathcal{S}$.

But in the case of quantum systems $\mathcal{S}_A \neq \mathcal{S}$, and for $A \neq A'$, in general, $\mathcal{S}_A \neq \mathcal{S}_{A'}$. For example, if an electron has a definite momentum, it doesn't have a definite position, and *vice versa*. Therefore, in quantum theory, an observable A has definite values only for states in a subset $\mathcal{S}_A$ of the state space $\mathcal{S}$, and not for other states.

The dynamics of a quantum system is described by the Schrödinger equation. The states are represented by unit vectors in a very high-dimensional complex vector space, and any observable A by a linear operator $\widehat{\mathbf{A}}$. Then, $\mathcal{S}_A$ consists of those unit vectors $|\psi\rangle$ that are transformed by the linear operator $\widehat{\mathbf{A}}$ into a vector of the form $\lambda|\psi\rangle$, where λ is a real number representing the value of the property A for the state represented by $|\psi\rangle$. The solutions of the Schrödinger equation satisfy equation (1), where s is a unit vector and $\mathcal{U}_{t-t_0}$ a *unitary transformation* (a rotation in the complex vector space $\mathcal{S}$). The operators $\mathcal{U}_{t-t_0}$ are called *evolution operators*, and they are all of the form

$$\mathcal{U}_{t-t_0} = e^{-i/\hbar \widehat{\mathbf{H}} t}, \tag{3}$$

where $\hbar$ is the *reduced Planck constant*, and $\widehat{\mathbf{H}}$ is called the *Hamiltonian operator*. The Hamiltonian operator $\widehat{\mathbf{H}}$ is the same for all t for closed systems, in particular for the entire universe, so it fully encodes the dynamical law.

Remark 1. In Quantum Physics there is an additional law that seems to contradict the dynamical law (1): whenever we measure an observable A of a quantum system, we find a definite value. Even if, according to the dynamical law, we would expect the system to be in a state for which the observable A is not defined! This suggests that, if the state of the system is not a state from $\mathcal{S}_A$, measurements make it somehow (appear to) jump in a state from $\mathcal{S}_A$. This jump is called *projection* or *wavefunction collapse*. The system can jump in one of more possible states from $\mathcal{S}_A$, according to a probabilistic rule (the *Born rule*). This tension between the law (1) and measurements leads to the *measurement problem*. We will not be concerned with this here, but we notice that a history of a quantum system may break the law (1) once in a while, when observations take place. This will not affect our discussion.

Let's put all of these together.

Definition 2. A *physical system* is characterized by

1. A *state space* $\mathcal{S}$ of possible states of the system.
2. A *dynamical law* given by the maps $\mathcal{U}_{\Delta t} : \mathcal{S} \to \mathcal{S}$ as in equation (1), for all time intervals $\Delta t \geq 0$.
3. A set $\mathcal{A}$ of *observables*, which are real functions defined on subsets of $\mathcal{S}$, as in equation (2).

A physical theory is a description of a physical system by specifying the elements from Definition 2. We can formulate a theory in other ways, but both Quantum and Classical Physics admit formulations as in Definition 2, including General Relativity (Arnowitt et al. 2008).

In Classical Physics, states can be distinguished from one another because some of their observables have different values. For example, two systems of point-particles can be distinguished by the positions of their particles. For classical systems, there is always a subset $\mathcal{A}_0 \subset \mathcal{A}$ of observables whose values can uniquely identify each state. For example, a classical system of n point-particles can be parametrized by the coordinates of the n particles,

$$\boldsymbol{q} := \big(\underbrace{x_1, y_1, z_1}_{\text{particle 1}}, \underbrace{x_2, y_2, z_2}_{\text{particle 2}}, \dots, \underbrace{x_n, y_n, z_n}_{\text{particle } n}\big) \tag{4}$$

together with the components of their momenta,

$$\boldsymbol{p} := \big(\underbrace{p_{x1}, p_{y1}, p_{z1}}_{\text{particle 1}}, \underbrace{p_{x2}, p_{y2}, p_{z2}}_{\text{particle 2}}, \dots, \underbrace{p_{xn}, p_{yn}, p_{zn}}_{\text{particle } n}\big). \tag{5}$$

Then, $(\boldsymbol{q}, \boldsymbol{p})$ is a parametrization of the state space $\mathcal{S}$. Similar parametrizations exist for systems containing classical fields.

In Quantum Physics, an observable A can distinguish two states only if both of them are from $\mathcal{S}_A$. There is no set of observables that distinguishes any two states. However, there is a set of observables $\mathcal{A}_0$ whose values can parametrize a basis of the state space, so that they can distinguish any two basis vectors. This is called a *complete set of commuting observables* (Dirac 1958). The basis consists of those vectors common to all subsets $\mathcal{S}_A$, for all $A \in \mathcal{A}_0$. Then, any state vector has a unique expression in this basis, which is what we call *wavefunction*.

For example, a quantum system of n scalar particles can be parametrized by positions, so that each state s is represented by a wavefunction of the form

$$\psi_s\big(\underbrace{x_1, y_1, z_1}_{\text{particle 1}}, \underbrace{x_2, y_2, z_2}_{\text{particle 2}}, \ldots, \underbrace{x_n, y_n, z_n}_{\text{particle } n}\big). \tag{6}$$

Similar parametrizations are possible for all types of quantum particles and quantum fields (Hatfield 2018).

Definition 3. In both the classical and the quantum cases, we call the set of observables $\mathcal{A}_0$ a *parametrization* of $\mathcal{S}$. We call *parameter space* the set of all possible combinations of values of the observables from $\mathcal{A}_0$.

In Classical Physics, the points of a parameter space are in one-to-one correspondence with the possible states. In Quantum Physics, they are in one-to-one correspondence with a basis of the vector space representing the possible states.

Observation 1. The physical structure of a system is characterized by the values of the parameters that identify the state, if we know what each parameter represents.

In the quantum case, the basis and the wavefunction expressing the state vector in that basis encodes everything that is to be known physically about the system.

Explanation. This is not obvious at first sight, but it should be familiar to physicists. For a classical system of n particles, we can read in the values of the positions how the particles are arranged in space, and from their momenta how they move. We can also read, from the dynamical law, which particles are charged, which attract or repel each other, and everything there is to know about the system.

In Quantum Physics, an atom can be described by its wavefunction. The wavefunction encodes the orbitals. More atoms, separated or parts of molecules, are described by a wavefunction on a higher dimensional parameter space, in a vector space whose basis is parametrized by the observables that correspond to more particles. The interactions are encoded in the dynamical law. Everything about a physical system can be read from the wavefunction, *if we know what each parameter means.* □

IV. PHYSICALISM AND ZOMBIES

Before characterizing Physicalism in terms of the physical formalism from Section §III, let us revisit some of the definitions of Physicalism.

Quote 1. A way to state Physicalism is (Stoljar 2023):

> Physicalism is true at a possible world w iff any world which is a physical duplicate of w is a duplicate of w *simpliciter*.

"Simpliciter" means without exceptions, qualifications, or specific conditions.

Quote 2. Also, according to Goff (2017), page 31:

> Let us call physical facts that can be captured in the mathematico-nomic vocabulary of physics "pure physical facts," and physicalism in conjunction with the view that fundamental reality wholly consists of such facts "pure physicalism."

With the formalism from Section §III, two systems are physically duplicate or equivalent iff their elements are in a one-to-one correspondence. More precisely,

Definition 4. A *morphism* between two physical systems (or even possible worlds) is a correspondence between their elements, *i.e.* a map between the state spaces of the two systems, $\alpha : \mathcal{S} \to \mathcal{S}'$, so that

1. If the first system is in the state s, the second system is in the state $\alpha(s)$.
2. For any state $s \in \mathcal{S}$ and any time interval Δt,
$$\mathcal{U}'_{\Delta t}\big(\alpha(s)\big) = \alpha\big(\mathcal{U}_{\Delta t}(s)\big). \tag{7}$$
3. For any state $s \in \mathcal{S}$ there is a correspondence between the observables of s and those of $\alpha(s)$, so that any observable A of s is related to the corresponding observable A' of $\alpha(s)$ by
$$A'\big(\alpha(s)\big) = A(s). \tag{8}$$

If α is one-to-one, this morphism is called *isomorphism*.

Two isomorphic systems or worlds are equivalent.

Equation (7) expresses the correspondence between the dynamical laws in the two possible worlds, but also between the histories. For the quantum case we can consider histories that include collapses as in Remark 1.

Remark 2. Definition 4 makes Physicalism as stated in Quote 1 equivalent with the Physicalism Thesis. Since the mathematico-nomic vocabulary from Section §III is known to provide a complete description of physical reality, Physicalism as stated in Quote 2 is also equivalent with Physicalism. This characterization of Physicalism is also consistent with other accounts of Physicalism as the thesis that everything can be described by physical quantities see Hempel (1969), Vicente (2011), and Jalloh (2023). □

Now let's introduce zombies.

Question 1. *Is it possible for two isomorphic systems to exist, one sentient and the other one insentient?*

An affirmative answer would prove that consciousness is irreducible to structure and dynamics, refuting Physicalism.

V. RELATIVITY OF STRUCTURE

Two theories are equivalent if they can be formulated as isomorphic descriptions of the same physical system. Any theory is, of course, equivalent to itself, by taking α to be the identity function $\alpha(s) = s$. Isomorphisms between a theory and itself are called *automorphisms*.

But a theory can be equivalent to itself in many ways simultaneously. Infinitely many "permutations" $\alpha : \mathcal{S} \to \mathcal{S}$ give different valid descriptions of the same system.

Definition 5 (Structural symmetries)**.** A *structural symmetry transformation* is an automorphism $\alpha : \mathcal{S} \to \mathcal{S}$.

Structural symmetry transformations form a *group*.

In Classical Physics, the structural symmetry transformations are *canonical transformations*. They are *generalized coordinate transformations* from a set of generalized coordinates and momenta $(\boldsymbol{q}, \boldsymbol{p})$ to another one $(\boldsymbol{q}', \boldsymbol{p}')$. Each new coordinate is a function of the old coordinates, $\boldsymbol{q}' = \boldsymbol{q}'(\boldsymbol{q}, \boldsymbol{p})$ and $\boldsymbol{p}' = \boldsymbol{p}'(\boldsymbol{q}, \boldsymbol{p})$.

In Quantum Physics, the structural symmetry transformations are *unitary transformations*, complex rotations that change the basis of the high-dimensional complex vector space used to represent the states.

Structural symmetry transformations change the parametrization that labels the states from $\mathcal{S}$ as in equations (4-6). Since a parametrization consists of the possible values of the observables from $\mathcal{A}_0$, $\mathcal{A}_0$ itself is transformed, usually in a different set of observables $\mathcal{A}_0'$.

For example, the wavefunction expressed in terms of positions in equation (6) can also be expressed in terms of momenta,

$$\psi_s'\big(\underbrace{p_{x1}, p_{y1}, p_{z1}}_{\text{particle 1}}, \underbrace{p_{x2}, p_{y2}, p_{z2}}_{\text{particle 2}}, \dots, \underbrace{p_{xn}, p_{yn}, p_{zn}}_{\text{particle } n}\big). \quad (9)$$

The symmetry transformation that takes the position representation (6) and gives as a result the momentum representation (9) is the *Fourier transform*. But there are infinitely many different parametrizations. The wavefunction looks different in different parametrizations.

When physicists make such transformations, they keep track of what physical property each observable represents, by notation and by names for the properties. For example, they do this by calling (6) *position* representation and (9) *momentum* representation.

Structural symmetry transformations are used to find representations that help us better understand the behavior of the system and solve problems more easily.

Translations and rotations of space are a particular case of structural symmetry transformations. They transform the positions $\boldsymbol{q}$ in a reference frame into the positions $\boldsymbol{q}'$ in another reference frame. But there are infinitely many more structural symmetry transformations, in both Classical and Quantum Physics.

The following result shows that the symmetry group of the structure is extremely large.

Recall from Definition 3 that the possible values of the observables from $\mathcal{A}$ form a parameter space $\mathcal{C}$ which, according to Observation 1, characterizes the structure of each state. A structural symmetry transformation α results in a different set of observables $\mathcal{A}'$, and therefore in a different parameter space $\mathcal{C}'$.

Proposition 1. Let $s \in \mathcal{S}$ be a state and $\mathcal{C}$ a parameter space. For any other state $s' \in \mathcal{S}$ with equal number of degrees of freedom as s, there is another parameter space $\mathcal{C}'$ so that the structure of s' on $\mathcal{C}'$ is identical with the structure of s on $\mathcal{C}$.

Proof. In Classical Physics, the parameter space of a system with n degrees of freedom is a $2n$-dimensional *symplectic manifold* $\mathcal{C}$. For any two points of the manifold, there is a canonical transformation α that maps the first point into the second point (Boothby 1969, Theorem A, page 98). This transformation can be used to map s to s', and the set of observables $\mathcal{A}_0$ that gives the parametrization of $\mathcal{C}$ into another set of observables $\mathcal{A}_0'$. Any observable A from $\mathcal{A}_0$ is mapped into an observable $A' = A \circ \alpha^{-1}$, therefore $A'(s') = A(s)$. They form a set $\mathcal{A}_0'$ which gives a new parameter space $\mathcal{C}'$. Since $A'(s') = A(s)$, the observables of s' from the set of observables $\mathcal{A}_0'$ have the same values as the observables of s from the set $\mathcal{A}_0$. Therefore, due to Observation 1, the structure of s' expressed in the parameter space $\mathcal{C}'$ is identical with the structure of s expressed in the parameter space $\mathcal{C}$.

In Quantum Physics, for any two unit vectors in a Hilbert space there is a unitary transformation $\widehat{\mathbf{S}}$ that maps the first vector into the second vector. The complete set of commuting observables $\mathcal{A}_0$ determine and parametrize a basis of the Hilbert space. The transformation $\widehat{\mathbf{S}}$ transforms any observable $\widehat{\mathbf{A}}$ into another observable $\widehat{\mathbf{S}}\widehat{\mathbf{A}}\widehat{\mathbf{S}}^{-1}$. Therefore, it transforms $\mathcal{A}_0$ into another complete set of commuting observables $\mathcal{A}_0'$, which determine and parametrize another basis. The components of the first state vector in the first basis coincide with the components of the second state vector in the second basis, so the wavefunction of s on $\mathcal{C}$ coincides with the wavefunction of s' on $\mathcal{C}'$. □

Remark 3. The transformations from Proposition 1 make the structure of $s(t)$ on $\mathcal{C}$ and that of $s'(t)$ on $\mathcal{C}'$ be identical, but in general these structures evolve differently, so they are different at another time $t' \neq t$. Such a transformation is a structural symmetry transformation (Definition 5) only if it preserves not only the structure, but also the dynamical law (1). To be a structural symmetry transformation, a transformation has to commute with the Hamiltonian operator $\widehat{\mathbf{H}}$ (Stoica 2021). In (Stoica 2023c), it was shown that distinct histories $s(t)$ and $s'(t)$, corresponding to distinct outcomes of quantum measurements, can be related by a structural symmetry transformation $\widehat{\mathbf{S}}$. This provides unlimited physically concrete examples of structural symmetries. □

An important implication of Proposition 1 is

Observation 2. The structure and the dynamics alone are insufficient to endow observables with physical meaning, *i.e.* to tell which physical property it represents. Nothing in the structure and the dynamics tells that $\boldsymbol{q}$ represents positions and $\boldsymbol{q}'$ doesn't. Then how do we know this? The physical correspondent of each observable is introduced by comparing the theory with the reality it describes. But this meaning is not part of the relational structure of the system, it may only seem so because we name and label the observables and we keep track of these names and labels when describing the system and solving the equations.

Therefore, Proposition 1 implies the following:

Principle 1 (Relativity of structure)**.** Assuming Physicalism, the symmetry group of any system should be its group of structural symmetry transformations.

But since in reality the symmetry group of any physical system consists of the spacetime and gauge symmetries, this leads to a question:

Question 2. *What reduces the large structural symmetry, associating a physical property to each observable A?*

Provisional answers. Proposition 1 excludes the possibility of the emergence of a unique correspondence between observables and physical meanings. Also see (Stoica 2021, 2023c).

But why not simply postulate this correspondence? Postulating it would require the ontology to do more than simply supporting the structure and following the dynamics, violating Rule 1 and therefore Physicalism. Also each transformation maps observables to observables, so whatever meaning we give to an observable, the transformation would reassign it to another observable.

Can we then add more structure to reduce the symmetry? We can try, but the added structure would not be observable, so it would be physically irrelevant.

The most sensible answer seems to be that the relation between observables and their physical meaning is established by experiments. For example, we measure the positions of other objects relative to our own position. We build measurement devices that translate the values of other observables into positions on the dial of the measuring apparatus, or in numbers displayed by the apparatus. All measurements of the observables are translated into data that we can perceive. But for this explanation to work, we have to include the observers in the physical description of the world. We will see that this opens a can of worms. □

VI. THE OVERLOOKED ROLE OF THE OBSERVER

Since Physicalism assumes that the world is causally closed, we need a complete description that includes the observers as subsystems of the world. The essential role played by the observer in Physics is not appreciated enough. Obviously, its role in the quantum measurement problem was extensively discussed, but this is a different problem. Our provisional answer to Question 2 shows that even in Classical Physics, where there is no measurement problem as in Quantum Physics, observers play an essential but ignored role.

The answer suggested to Question 2 is that the relation between observables and their physical meaning comes from the experiments. But experiments can find only relations between various observables of the observed system and the instruments used to perform the experiment (Section §II). And they are ultimately translated into observables familiar to the human observers performing the experiment, who give them a physical meaning.

But what endows with physical meaning the properties of the observer? We just take them as reference. For example, we take the position or the velocity of a system that we observe relatively to our own position or velocity.

But there is much more than this. The observers not only bring in their own positions, they also "smuggle" into the description of the world the very notion of position. The observers experience their own observables as physical properties.

Principle 2. The observers give physical meaning to the observables of the physical systems by anchoring them in the parameter space in which they appear as observers.

Principle 2 as such doesn't introduce sentience. It only states the role of the observers in anchoring the relations and structures in the physical reality relatively to their own structure. From this, it seems that there are two options, depending on the answer to the following question:

Question 3. *Is the structure of the observer sufficient to give physical meaning to the observables of the system?*

Our discussion will make use of the following notion:

Definition 6. An *observer-like structure* is a system isomorphic with an observer.

Proposition 1 shows that any structure possible in a parameter space it is possible in any other parameter space. This applies to observer-like structures as well. If the structure of the observer were sufficient to give physical meaning to a observable A of the system in the parameter space $\mathcal{C}$, an observer-like structure from another parameter space $\mathcal{C}'$ would give the same meaning to the observable $A' = A \circ \alpha^{-1}$, and to A would assign a different physical meaning. By giving different physical meaning to the observables of a state, they would perceive the same state as having different structures, and therefore, as different physical worlds.

On the other hand, a negative answer to Question 3 would imply that some observer-like structures can give physical meaning to observables, and others can't do this, so they are zombies, at least from this point of view.

Let's ask Physics to answer Questions 1, 2 and 3.

VII. ASKING PHYSICS

In this Section I prove the main result:

Theorem 1. *Any sentient physical being has insentient structural duplicates in other parameter spaces.*

Proof. An observer O is a subsystem of the world. Let E be the observer's external world (or *environment*). The observer O can know the values of some properties of the environment. This knowledge is encoded in the structure of her brain.

Let s, s' be states of the world, including the observer and her environment so that, in the parameter space $\mathcal{C}$, a property $\mathscr{A}$ of the environment of O, represented by the observable A, has the value a in s, respectively a' in s'. We choose the two states so that the observer O has the same structure in both of them, and this structure encodes the knowledge that the value of A is a. Both s and s' are possible states, even though the environment in s' is not as the observer thinks it is. According to the Lemmas from (Stoica 2023c), it is possible to choose s, s' so that they are related by a structural symmetry transformation α. In (Stoica 2023c) this was shown for quantum systems like our world.

From Proposition 1, there is a parameter space $\mathcal{C}'$ on which the state s has the same structure as s' on $\mathcal{C}$. They also evolve according to the same law. On $\mathcal{C}'$, s looks like s' on $\mathcal{C}'$, and since the observer-like structure O' has the same structure as O, it also encodes the information that the value of A is a. But on $\mathcal{C}'$ the structure of the state s is so that the property $\mathscr{A}$ is represented by the observable $A' = A \circ \alpha^{-1}$, whose value is a', and not a, as encoded in the observer's structure on $\mathcal{C}'$.

Since the observer only knows its own structure, it doesn't know whether it is O from $\mathcal{C}$ or O' from $\mathcal{C}'$. Therefore, the observer-like structure doesn't know whether the value of the property $\mathscr{A}$ really is a, as encoded in the structure of her brain, or any other possible value a'.

If observers were reducible to their structure, all such observer-like structures would be observers. An observer wouldn't know which of these observer-like structures she is. Therefore, she wouldn't know if the environment's property A has the value a or any other value a'. And this applies to all observable properties of the environment. Therefore, the probability that an observer has correct knowledge about her environment would be zero.

But we have correct knowledge about our environment. Therefore, we find ourselves every time in a parameter space with the right properties. This is either perpetual pure luck with vanishing chances to happen, or we simply can't be one of the observer-like structures from the other parameter spaces. This can happen only if we are sentient but those other observer-like structures are insentient. ☐

Theorem 1 shows that if all observer-like structures were sentient, our knowledge about the environment would be just like a random guess, contradicting our observations. Here are some of its implications:

Implication 1 (Irreducibility of sentience)**.** The answer to Question 1 is affirmative, duplicate zombies of sentient beings exist on other parameter spaces.

Remark 4. This is probably unexpected. When one says that there are no two identical structures that differ by their sentience, one usually either assumes that the two structures are from different possible worlds, or, if they are in the same world, they are as well on the same parameter space. But in our case the identical structures are in the same physical state, but on different parameter spaces, with respect to which the world appears to have different structures.

Implication 2 (Ontology)**.** Since structure and dynamics are insufficient to guarantee the sentience of an observer-like structure, the ontology of the preferred parameter space should provide the sentiential powers that structure and dynamics can't provide.

Definition 7. We call *ontic parameter space* a parameter space in which observer-like structures are sentient.

Implication 3 (Uniqueness)**.** The ontic parameter space is unique up to unobservable physical symmetries (space or spacetime isometries and gauge symmetries).

Proof. Any property as in the proof of Theorem 1 is observable, so it is invariant to gauge symmetries. Otherwise there would exist observables that can be measured but cannot be known, leading to a contradiction. The same applies to changes of reference frames in space. ☐

Implication 3 answers Question 2.

Implication 4 (Observers and physical meaning)**.** The answer to Question 3 is negative: the structure of the observers alone cannot give physical meaning to the observables. This is possible only in conjunction with the ontic parameter space.

Therefore, Theorem 1 shows that all physical properties are grounded in sentient experience.

Implication 5 (Sentience)**.** By associating sentient experience not only with the structure and dynamics, but also with ontology, Theorem 1 supports the following double understanding of sentience:

1. as the additional ingredient making an observer-like structure be an observer (intrinsic role),
2. as what gives physical meaning to the observables (extrinsic role).

Implication 6 (Phenomenology)**.** All observer-like structures encode information that can be interpreted as representational or aboutness or *access consciousness*, but only those from the ontic parameter space can experience it. Therefore, sentience can't be reduced to access consciousness.

Remark 5. The existence of memories that correspond to the past events, the so-called *epistemic arrow of time*, is related to the Second Law of Thermodynamics, which requires that the universe's initial state had a very special structure (Boltzmann 1964, Albert 2000, Stoica 2022a). This may make one think that we can avoid the conclusion of Theorem 1 by invoking the epistemic arrow of time. But since on another parameter space the structure of the initial state can appear like the structure of any other state, the initial conditions on another parameter space don't guarantee the validity of memories. This is why most observer-like structures from other parameter spaces would be like Boltzmann brains, they wouldn't know the properties of the environment on their parameter spaces.

VIII. PHYSICALISM AND CONSCIOUSNESS

Theorem 1 implies that

Implication 7. Physicalism is invalid.

Maybe some readers accepted Physicalism as defined at the beginning of this article, and now they disagree that it expresses their own physicalist views. It may even appear to the reader that the Physicalism Thesis is a strawman which doesn't capture the real Physicalism as they see it. But if Physicalism as they see it requires an objectively preferred parameter space, this would mean to break Rule 1.

Some people who regard themselves as physicalists may be calling physical everything that exists. This would be a very vague notion of Physicalism, to the extent that it would be contentless. Such notion of Physicalism would accommodate everything, so it would lack both explanatory and predictive power (Hempel 1980). The point of Physicalism was, like that of materialism, to reject precisely certain ideas about consciousness, including attributing sentiential powers to ontology.

Theorem 1 is relevant, because it excludes several important theses about the world and consciousness:

Refuted Hypothesis 1*. Ontic structural realism*, the thesis that structure alone is the complete ontology (Ladyman 2020).

Refuted Hypothesis 2*. Epistemic structural realism*, the thesis that only the structure, and not the ontology, is accessible to our knowledge. We have seen that without ontology there would be no knowledge.

Refuted Hypothesis 3*.* Tegmark's *Mathematical Universe Hypothesis*, the proposal that mathematical existence (*i.e.* logical consistency) equates physical existence (Tegmark 2014). Mathematical objects are pure structures, and therefore they are insufficient to describe the world.

Refuted Hypothesis 4*.* The thesis that only the state vector and the (spectrum of the) Hamiltonian are fundamental, and everything else can be derived from them, see for example (Carroll 2021, Carroll and Singh 2019) and references from (Stoica 2021). In (Stoica 2021) it was already shown in full generality that these are insufficient to recover even space or the decomposition of the world into subsystems. Numerous counterexamples of this thesis were given in (Stoica 2022b). A much more physically intuitive refutation was given in (Stoica 2023c). But now we see that even adding more structure is still insufficient.

Refuted Hypothesis 5*.* Other purely relational approaches like *Relational Quantum Mechanics* (Rovelli 1996) and various quantum-first proposals discussed in (Stoica 2021).

But the most important are the implications for the Philosophy of Mind. The following proposals qualify as Physicalism, so they are refuted:

Refuted Hypothesis 6*.* Versions of *behaviorism* in which consciousness reduces to behavior (Graham 2023).

Refuted Hypothesis 7*.* The *computational theory of mind* (Rescorla 2020, Colombo and Piccinini 2023). It was already shown (without appealing to other parameter spaces) that the mind can't be reduced to a computation as understood in Computer Science (Stoica 2023b).

But sometimes by "computation" it is understood that the structure of the machine implementing the computation has to be isomorphic to that of the data processed by the computation (Chalmers 1994, Piccinini 2015). This contradicts Computer Science, in particular Turing universality (Stoica 2023b). And now we have seen that even if we take structure into account, the mind can't be reduced to such a restricted notion of computation.

Refuted Hypothesis 8*. Functionalism*, the proposal that consciousness reduces to functionality (Levin 2018), if functionality is seen as structure and dynamics alone.

Refuted Hypothesis 9*. Representationalism*, the proposal that consciousness reduces to representations (Lycan 2019). Representations are just morphisms between the structure of the brain and that of the represented systems, so they are structures. Also see Implication 6.

Refuted Hypothesis 10*. Identity theory* (Smart 2022) based on structure and dynamics alone. Theorem 1 applies down to the finest structural details of the atoms and elementary particles, because these structures are possible in any parameter space.

Refuted Hypothesis 11*. Integrated Information Theory* proposes that the degree of consciousness is measured by a numerical value Φ, exclusively determined by the structural characteristics of the system (Tononi et al. 2016). Since all isomorphic observer-like structures have the same Φ, Theorem 1 refutes the claim that Φ is a measure of consciousness. Note that both Tononi and Koch consider IIT to be a rigorous theory of panpsychism, but it can't be if it relies on structure alone.

Refuted Hypothesis 12. *Illusionism* (Dennett 2016, Frankish 2016) and *eliminativism* (Ramsey 2022), the attempts to explain away phenomenal consciousness by reducing it to one of the proposals based on structure and dynamics alone, for example to computation or representation. Also see Implication 6.

The results from this article make no proposal to explain consciousness. They don't provide a theory, a description, or an explanation of consciousness. They don't show what kind of structure conscious beings should have. To explore this, Hypotheses 1-12 are useful and constructive, because they focus on structure, dynamics, functionalism, representations *etc*. Any of these may advance our understanding of the mind. But, since Theorem 1 refutes Physicalism by revealing a connection between ontology and consciousness, it refutes the reductionist claims made by all Hypotheses 1-12.

Theorem 1 supports the following

Thesis: Sentiential Ontology. Physical systems are characterized by their structure, dynamics, *and sentience*, which grounds consciousness and the physical meaning of all observables.

Sentiential Ontology doesn't require nonphysical entities, neither modifications of the structures and the dynamics, so it doesn't violate causal closure.

Let us see what theories of mind are compatible with Sentiential Ontology.

Compatible Hypothesis 1. *Dualism*, with matter and mind as fundamental but distinct kinds of things (Robinson 2023) that interact or mirror each other's properties.

For dualism to be compatible with Sentiential Ontology, the mental stuff have to duplicate the structure, the dynamics, and the properties of matter, as seen in Implication 3, so it has to be maximally redundant.

Monistic proposals avoid this redundancy. Here are, very broadly, the remaining options.

Compatible Hypothesis 2. *Panpsychism*, "the view that mentality is fundamental and ubiquitous in the natural world" (Goff et al. 2022). Compatibility with Implication 3 requires that sentiential properties correspond to observable physical properties, and this means that the ubiquity of sentience is maximal, as in the monistic forms of panpsychism (James 1904, Mach 1914, Russell 1927, Eddington 1928, Stubenberg and Wishon 2023).

But if physical properties are sentiential, the most natural form of Sentiential Ontology is one in which the ontology is purely sentiential. After all, all that Physicalism lacked was sentience (Implication 5).

Compatible Hypothesis 3. Variants of *idealism* (Berkeley 1881, Guyer and Horstmann 2023) that are not antirealistic, in which the reality of properties is grounded in a sentiential substance (Indich 1995, Kastrup 2019, Stoica 2020). The sentiential substance can have its own structure and follow its own dynamics like any system as in Definition 2. As it happens, its structure and dynamics are those that we observe in our universe. But sentience also gives physical meaning to observables.

Sentiential Ontology is not anti-physicalist, if Physicalism is understood based on the fact that *physis* means "nature". Nature includes ontology, even if ontology has sentiential powers. A fair physicalist position in this sense, which is consistent with Sentiential Ontology is advocated by Strawson (2006), page 3:

> You're certainly not a realistic physicalist, you're not a real physicalist, if you deny the existence of the phenomenon whose existence is more certain than the existence of anything else: experience, 'consciousness', conscious experience, 'phenomenology', experiential 'what-it's-likeness', feeling, sensation, explicit conscious thought as we have it and know it at almost every waking moment.

This self-evident truth should have been sufficient, but its rejection became the scientific norm, because it doesn't show up in the publicly observable structures of the brain. Now we have seen that even if we deny our own sentient experience and its scientific relevance, even if we try to exclude it from Science, Physics itself tells us that structure is insufficient, and without sentience no property of the world would be known to us.

REFERENCES

Albert, D. (2000). *Time and chance.* Harvard University Press, Cambridge, Massachusetts; London, England.

Arnowitt, R., Deser, S., and Misner, C. W. (2008). Republication of: The dynamics of general relativity. *General Relativity and Gravitation*, 40(9):1997–2027.

Berkeley, G. (1881). *A treatise concerning the principles of human knowledge.* JB Lippincott & Company.

Boltzmann, L. (1964). *Lectures on gas theory.* University of California Press,, Berkeley.

Boothby, W. M. (1969). Transitivity of the automorphisms of certain geometric structures. *Trans Am Math Soc*, 137:93–100.

Carroll, S. (2021). Reality as a vector in Hilbert space. Technical Report CALT-TH-2021-010, Cal-Tech.

Carroll, S. and Singh, A. (2019). Mad-dog Everettianism: Quantum Mechanics at its most minimal. In Aguirre, A., Foster, B., and Merali, Z., editors, *What is Fundamental?*, pages 95–104. Springer.

Chalmers, D. J. (1994). On implementing a computation. *Minds and Machines*, 4(4):391–402.

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of consciousness studies*, 2(3):200–219.

Chang, C. and Keisler, H. (1990). *Model theory.* Elsevier.

Colombo, M. and Piccinini, G. (2023). *The Computational Theory of Mind.* Cambridge University Press, New York.

Dennett, D. (2016). Illusionism as the obvious default theory of consciousness. *J Conscious Stud*, 23(11-12):65–72.

Dirac, P. (1958). *The Principles of Quantum Mechanics*. Oxford University Press, Oxford.

Eddington, A. (1928). *The nature of the physical world*. Dent, London.

Frankish, K. (2016). Illusionism as a theory of consciousness. *J. Conscious. Stud.*, 23(11-12):11–39.

Goff, P. (2017). *Consciousness and fundamental reality*. Oxford University Press.

Goff, P., Seager, W., and Allen-Hermanson, S. (2022). Panpsychism. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2022 edition.

Graham, G. (2023). Behaviorism. In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2023 edition.

Grätzer, G. (2008). *Universal algebra*. Springer Science & Business Media.

Guyer, P. and Horstmann, R.-P. (2023). Idealism. In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2023 edition.

Hatfield, B. (2018). *Quantum field theory of point particles and strings*. CRC Press.

Hempel, C. (1969). Reduction: Ontological and linguistic facets. In *Readings in Philosophy of Psychology*, pages 179–199. St. Martin's Press, New York.

Hempel, C. (1980). Comments on goodman's "ways of worldmaking". *Synthese*, pages 193–199.

Hodges, W. (1997). *A shorter model theory*. Cambridge University Press, 1 edition.

Hoefer, C., Huggett, N., and Read, J. (2023). Absolute and Relational Space and Motion: Classical Theories. In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2023 edition.

Indich, W. M. (1995). *Consciousness in Advaita Vedanta*. Motilal Banarsidass Publ.

Jalloh, M. (2023). The nature of the physical and the meaning of physicalism. *THEORIA*, 38(2):205–223.

James, W. (1904). A world of pure experience. *The Journal of Philosophy, Psychology and Scientific Methods*, 1(20):533–543.

Kastrup, B. (2019). *Analytic Idealism: A consciousness-only ontology*. Bernardo Kastrup.

Kirk, R. (2023). Zombies. In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2023 edition.

Kirk, R. and Squires, R. (1974). Zombies v. materialists. *Proc. Aristot. Soc., Supplementary Volumes*, 48:135–163.

Ladyman, J. (2020). Structural realism. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2020 edition.

Levin, J. (2018). Functionalism. In Zalta, E., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2018 edition.

Lycan, W. (2019). Representational Theories of Consciousness. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2019 edition.

Mach, E. (1914). *The analysis of sensations, and the relation of the physical to the psychical*. Open Court Publishing Company.

Piccinini, G. (2015). *Physical computation: A mechanistic account*. OUP, Oxford.

Poincaré, H. (2022). *Science and Hypothesis*. Bloomsbury Publishing.

Ramsey, W. (2022). Eliminative Materialism. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2022 edition.

Rescorla, M. (2020). The Computational Theory of Mind. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2020 edition.

Robinson, H. (2023). Dualism. In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2023 edition.

Rovelli, C. (1996). Relational quantum mechanics. *Int. J. Theor. Phys.*, 35(8):1637–1678.

Russell, B. (1927). *The analysis of matter*. London: Kegan Paul.

Smart, J. J. C. (2022). The Mind/Brain Identity Theory. In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2022 edition.

Stoica, O. C. (2020). The negative way to sentience. *Preprint on http://philsci-archive.pitt.edu/17036 - last accessed November 9, 2023*.

Stoica, O. C. (2021). 3d-space and the preferred basis cannot uniquely emerge from the quantum structure. *To appear in Advances in Theoretical and Mathematical Physics. Preprint arXiv:2102.08620*.

Stoica, O. C. (2022a). Does quantum mechanics require "conspiracy"? *Preprint arXiv:2209.13275*.

Stoica, O. C. (2022b). Versatility of translational quantum dynamics. *Preprint arXiv:2204.01426*.

Stoica, O. C. (2023a). Are observers reducible to structures? *Preprint arXiv:2307.06783*.

Stoica, O. C. (2023b). Does a computer think if no one is around to see it? *philsci-archive:22304*.

Stoica, O. C. (2023c). The prince and the pauper. A quantum paradox of hilbert-space fundamentalism. *Preprint arXiv:2310.15090*.

Stoljar, D. (2023). Physicalism. In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2023 edition.

Strawson, G. (2006). Realistic monism: Why physicalism entails panpsychism. In Freeman, A., editor, *Consciousness and its place in nature: Does physicalism entail panpsychism?*, pages 3–31, UK. Imprint Academic.

Stubenberg, L. and Wishon, D. (2023). Neutral Monism. In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2023 edition.

Tegmark, M. (2014). *Our Mathematical Universe: My Quest for the Ultimate Nature of Reality*. Knopf Doubleday Publishing Group.

Tononi, G., Boly, M., Massimini, M., and Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. *Nature Reviews Neuroscience*, 17(7):450–461.

Vicente, A. (2011). Current physics and 'the physical'. *Br J Philos Sci*.