

Thoughts as Selected Set-Theoretical Constructions, and topics in Philosophy of Mind by way of Mathematical Analogy

Stephen Gutwald

stgutwald@gmail.com

Abstract

A theory of mind is provided by assuming thoughts are mathematical objects (more specifically, constructible using set-theory). Problems from the philosophy of mind are probed using mathematical analogy, and the relation of minds to bodies is clarified using relations that are typical between mathematical structures.

Assumptions

The following positions in the Philosophy of Mind will be taken as postulates:

1. *Supervenience*. Physical indiscernibility implies mental indiscernibility. It is impossible for a physical system to realize a mental-state, and a particle-to-particle duplicate of that system realize a different mental-state (or none at all).
2. *Multi-realizability*. Mental-states are not identical with physical-states (i.e., brain-states). The relation between brain-states and mental-states is not a 1-1 correspondence. Mental-states have multiple physical realizations (as the experience of pain being realized across species with inequivalent physiologies).
3. *Causal Closure of the Physical Domain*. Physical causes are sufficient to explain physical events. The causes of a physical event are physical, and the causes of those causes are physical, and at no point does the chain of explanation escape the physical domain.

These positions are not a universal consensus among philosophers of mind. They are taken as axioms to stage a logical exploration.

Additionally, the physical domain (or substrate) through which mentality is realized, is conceived of as a Wittgensteinian world: As true states-of-affairs arranged in logical space (Wittgenstein, 2001). Here we only mean that physical-states are bits of matter (macroscopic and microscopic objects) configured into a definite arrangement. A physical-state is a way in which physical constituents bind within a structure. For example, in the human brain, the mutual positions of the neurons, their connections and excitations, the way in which all this “hang one in another like links in a chain” (2001, p. 9) is a physical-state (the state of the physical system). My purpose for evoking such a world is to transition from the more conventional ontology of *things*, to an ontology of *physical-states*, and by maintaining those states as points somewhere in logical space, obtain a set $P = \{m: m \text{ is a state}\}$ (as opposed to $\{o: o \text{ is a thing}\}$).

Notice that (1) is the logical requirement for a mathematical function from physical-states to mental-states, and (2) says this function is many-to-one (ie, not injective). Write this function as $e: P \rightarrow M$ from physical-states P to mental-states M . A “null experience,” written as 0, will be included in M to account for the images $e(p)$ of physical-states $p \in P$ without the smallest traces of consciousness.¹ Such an inclusion is required for the mathematical form. Observe that the writing of $e: P \rightarrow M$, though effortless in symbols, is heavy with suppositions. To start, it takes for granted that physical-states can be collected together and made the elements of a set. Such is natural to our world comprised of facts in logical space (here the set and the world are the same); but there is some cause to dismiss the project of cleanly demarcating physical-states and unambiguously collecting them together as naively optimistic. True, at times it is hard to know how to apply labels in nature and to decide where a phenomena begins and ends, but one would admit a great many cases where this rendering into elements proceeds well: My brain-state and your brain-state might be made separate elements $b_1, b_2 \in P$ without controversy, as could the state of a flowering plant $f \in M$, or the internal state of our sun $s \in P$ (all these, of course, at specific times). We assume this technique to be universalizable, which is perhaps not precisely true, but like any other scientific model, an idealization must be made, and ambiguities trimmed, to derive consequences which, while also ideal, model approximately. The writing of $e: P \rightarrow M$ also assumes that our right-hand-side M is worthy of objecthood. That the set M (of beliefs, sensations, emotions, and of mental species generally) is a set of objects, where critics might object that M is a bag of tricks and illusions (Dennett, 1993) and is better left unwritten. It should be recognized that if P is accepted as a mathematical set, and (1) is assumed, then mental-states necessarily gain set-hood. For example, if identity-theory were correct, then $e: P \rightarrow M$ is a bijection, and mental-states are alike to a relabeling of physical-states (the set of one implies the set of the other). We assume multiple-realizability (2), in which case the set of physical-states still grounds the set of mental-states, provided mental-states can be identified across physical realizations. Provided pairs of mental-states can be decided as either identical or non-identical, then define an equivalence-relation \sim on P which makes physical-states equivalent when they realize the same mental-state. Since P is a set-theoretically defined, *as is the equivalence relation*, the set of equivalence classes is also defined, and these classes label mental-states. Either the equivalence classes P/\sim are a proper subset of M (that is, there exist mental phenomena that are not physically realizable), or we might make the reasonable guess that $M = P/\sim$, which is equivalent to the function $e: P \rightarrow M$ being surjective.

The function also brackets differences among the varieties of mental phenomena. A belief is not the same as an emotion. While a single experience might be analyzed into a constellation of propositional attitudes and complexes of feeling. The fact that a physical-state p has a unique image $p \rightarrow e(p)$ implies that $e(p)$ is seen as point-like. Though this is not unlike the Wittgensteinian picture of physical-states: Complex arrangements of objects (as all the people on a subway car) is taken as an existing state-of-affairs and is made a single point in logical space. In that same way, a mind stocked with psychic-travelers becomes a point of its space. The function e maps physical-complexes to mental-complexes.

Mathematical Analogies

Consider a physical domain of classical particles moving through space. The dynamics of these particles is determined by laws with some likeness to the laws governing our Universe but cannot be assumed to be identical. As observers of this domain, we are given the task of conceptualizing a scientific model

¹ This formulation is also compatible with pan-psychism, which would say that: $e(p) \neq 0$ for all $p \in P$

whose predictions correspond to events as they happen. As data, we are provided the worldlines of each particle over a set interval of time. For simplicity, assume the interval of time begins at 0 and ends at 1 (that is, one unit), so that these worldlines are represented as continuous functions $p: [0,1] \rightarrow S$ from the unit-interval to a space S . Going forward, a *space* will always mean a path-connected topological space. It will also be assumed that there are enough physical paths to supply each construction (for example, that each homotopy class is represented by a physical path).

Suppose we assume a strict ontologically conservative stance for this research project. We accept, at least, that the particles themselves are real and that space is real, but nothing more. In all matters we refuse to multiply entities and eliminate external causes wherever proffered. Whenever a rival school claims spirit is at work among the particles, or the finger of some deity is gently guiding from behind, we are quick to reduce the issue to particles and space. Our eliminative approach is supportable: All events reduce to particle-locations, therefore knowing the positions of all particles at all times is a complete understanding without room for outside revelation.

Suppose events are smooth enough that differential calculus is applicable. Consider this construction: For a point $x \in S$ and a distinguished time $t_* \in [0,1]$ consider a subset of paths p where $p(t_*) = x$. Now on this subset of paths create an equivalence relation which equates $p \sim q$ when their derivatives coincide at the distinguished point $p'(t_*) = q'(t_*)$ (forgiving some abuse of notation). The equivalence classes are tangent vectors and together give the tangent space at x . Intuitively speaking, these are the possible directions through the point x .

Here is a challenge to our school: Are tangents real? Do they deserve the distinction of “exists”? It seems a silly question, since tangent vectors are exactly what the physicist applies routinely in their calculations (as related to velocity). That is no good, because *physical* to a physicist is often whatever works. A physicist is ontologically uncommitted and will borrow from distant spheres whatever aids their calculation. We are partisans of ontological minimalism. For us, particles and space exhaust everything that is, and how can you put our chosen substance together, laying space on top of space, or placing particles next to particles, to get a tangent?

Consider our school an analogue of reductive materialism. In classical mechanics, velocities are not thought of as determined by a base dataset of spatiotemporal positions. That is because the future positions of objects are not a given and are the task to be calculated. In our case, the particle worldlines are disclosed in full, and that makes tangents a supervenient property. They are determined by $p: [0,1] \rightarrow S$ considered pointwise while being unable to make alterations to the given path.² I propose a situation where the reduction to particles and space is sufficient to correctly predict *about* particles and space. Where the domain particles-space is causally closed and tangents are not needed. It could be that the particles are classified by a taxonomy of readily calculated functional-types. That each path adheres to a form (think of $t \rightarrow \sin(t)$ or $t \rightarrow \ln(t)$ from the real-valued case) and it is enough to determine which by tells in its movement. Such would be an example of a physical model without a need to appeal to tangents. Despite this, mathematically speaking, any particle at any time still has a

² The situation is alike to calculating the derivative of a differentiable function $f: \mathbb{R} \rightarrow \mathbb{R}$. The derivative $f'(t)$ is determined by the function pointwise and cannot alter any value of the function.

tangent.³ Further, it is possible that the particles appear to respond to their tangents, to be visibly motivated or influenced by them, despite their epiphenomenal status in our model.⁴

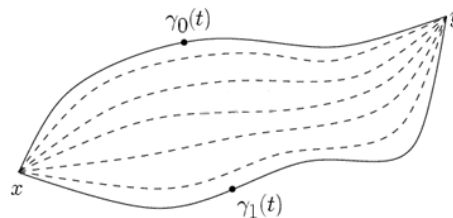
Does a strict commitment to an ontological base (particles-space) necessitate the transcendence of that base (in the form of tangents)?

Before this is affirmed as obvious, it should be appreciated that a similar mathematical process could incorporate mental phenomena. Above an equivalence relation was defined on physical-states when they realize the same mental-state, assuming all mental-states are realizable, mental phenomena might identified with equivalence classes of physical-states, so that:

$$M = P/\sim$$

That is not unlike the production of tangents, which calls paths through a distinguished point equivalent when they share a direction and uses the resulting equivalence classes. Both share a form: Whatever might be the ontological substrate (let us say U), first take the Cartesian product $U \times U$, define an equivalence relation $\sim \subset U \times U$, then take the equivalence classes of that relation U/\sim . If this sequence is permitted in one case, why not the second? The response might be that tangents are recognizably physical, and mentality is not, but does that recognition come from tangents being native to the substrate, or is it because physicists have a way of transgressing ontological boundaries without being conscious of it?

Another construction. Say that two paths (traced by particles) are equivalent when the first is continuously deformable into the second while preserving endpoints. The technical details are not very relevant, intuitively it should look something like:



Note. A homotopy between two paths γ_0 and γ_1 which preserves endpoints x and y from Wikimedia Commons.

For a distinguished point $x_* \in X$ consider the subset of paths $p: [0,1] \rightarrow S$ so that $p(0) = x_* = p(1)$. In other words, loops which start and end at x_* . Call the equivalence classes *homotopy classes*. The

³ Allowing set-theoretical closures, tangents exist and are assigned to particles independently of whether they appear in the physical model of our reductive school.

⁴ The influence in question would be limited to a supervenient influence (which will be explored shortly but is presently undefined). The case holds that $p: [0,1] \rightarrow S$ is known in full, so that every tangent is determined and is unable to change the value of p at any time t . Whatever this influence might be, it is difficult to fit into a system that is already completely determined. The difficulty is of course analogous to fitting minds into a physically deterministic system. For now, I would ask the reader to imagine tangents as influencing particles much as thoughts influence bodies. In the same way that human beings gravitate towards a happiness that cannot interrupt or redirect physical causal chains in the brain (by our third premise); imagine particles conforming to tangent-choosing behaviors, even when those tangents cannot change what is already determined.

product of two homotopy classes is defined as the homotopy class of the composition of loops (the first loop followed by the second). The inverse of a homotopy class is the homotopy class of the inverse loop (following it in reverse). The identity homotopy class is the homotopy class of the constant loop. This gives a mathematical group called the Fundamental Group. Now the same question: Does the Fundamental Group exist?

This construction has less physical appeal, but that is likely because all loops are deformable to a constant loop in the flat space we experience, making these groups degenerative. When space is punctured, teared, or wound into itself (as the domain of our investigation may be) these groups will have a say over dynamics. They will announce themselves to a physical investigation.

Still partisans of ontological minimalism: Do we allow these groups the dignity of objecthood? Tangents are easy to square with the substrate. They appear to travel with the object, attending to its movements, and are always present as physical parts are pushed about. If not themselves matter, simple enough to give them an honorary designation of “material.” But how to locate the elements of a fundamental group? To put it in perspective: The Fundamental Group of the circle is the integers (under addition). How much of the structure of space is carried to the integers? Not only is there nothing left of space in the integers – there is no place in which to make a place – but the integers are their own mathematical structure, and one of a completely different kind. Think of it this way: Having arrived at the structure of the integers under addition, consider its subgroups which all have the form $m\mathbb{Z} = \{mn: n \in \mathbb{Z}\}$. Order these through the subset relation and observe that $m\mathbb{Z} \subset k\mathbb{Z}$ exactly when k divides m . What are the largest subgroups that are not \mathbb{Z} itself? Those of the form $p\mathbb{Z}$ where p is a prime number. A mathematical bridge has been crossed from a world of angles, distances and volumes, to a world of algebra, divisibility and prime numbers. How have we teleported to this other side? A side so substantially different, and whose questions and meanings are so alien to what we began with.

Tangents show how the inner workings of a substrate might entail a transcendence of that substrate. Entities “over and above” mere particles and space considered in their baseness. The Fundamental Group (whose construction is not so dissimilar to that of tangents) shows how a transcendence might leap into radically different structures and meanings. These results are certainly evocative when brought to the Philosophy of Mind, and we will explore this suggestion further.

Derived Properties

What both the Tangent Space and the Fundamental Group have in common is their constructability using set theory. In the case of tangents:

1. Space S exists and time $[0,1]$ exists.
2. The set P_{x,t_*} , consisting of all paths $p: [0,1] \rightarrow S$ such that $p(t_*) = x$, exists.
3. The Cartesian product P_{x,t_*}^2 exists.
4. The equivalence relation $\sim \subset P_{x,t_*}^2$ exists.
5. The equivalence classes exist.

Conclusion: Tangents exist.

Going forward I will also call a set-theoretic construction a *derivation* or a *derived property*. The moral of the above examples is that derivations are capable of ontological transversals. Consider the production

of numbers: The integers are derivable from the natural numbers, the rational numbers are derivable from the integers, the real numbers are derivable from the rational numbers, the complex numbers are derivable from the real numbers.

I would like to contrast this with the comparatively limited constructive power of first-order logic.⁵ In first-order logic a signature of formal symbols is interpreted into a set U (a domain of interpretation) such that an n -ary relation R is interpreted as a subset $R \subset U^n$, an n -ary function f as $f: U^n \rightarrow U$, and constants as elements of U . Formulas are then recursively generated through logical combination. The resulting constructions tend to be ontologically conservative, with all the productions resulting in another U -element or more U -stuff. Take the Natural Numbers as a model of Peano Arithmetic. We are provided with a domain \mathbb{N} and a handful of linguistic tools to make sense of it. Much is possible with this limited beginning, and with some ingenuity it is possible to define divisibility, greatest common divisors, prime numbers, Diophantine equations, and so on. However, we cannot ascend our ladder to the complex numbers (and still beyond). These horizons are intuitively visible in the logical forms themselves. How could one compose functions $f(g(x, y), h(x, y))$ in a way that does not return an element of \mathbb{N} ? How could the existential quantifier be affixed to a relation $\exists x R(x, y)$ without a solution being somewhere in \mathbb{N} ? More rigorously, any n -ary formula ψ specifies a definable subset of tuples that satisfy it:

$$\{(a_1, \dots, a_n) \in N^n: N \models \psi(a_1, \dots, a_n)\}$$

so that these subsets represent the limits of expression. One is unable to use second-order powers to quantify over relations, make formulas themselves the objects, and so on. It is still more impossible to pair \mathbb{N} with another mathematical object and observe the functions or relations between them.

Returning to our universe of classical particles. Suppose that law-like behaviors emerge amid their jostling:

(Mutual exclusion) If two particles x and y have the same spacetime coordinate then $x = y$.

Or:

(Recurrence) If a particle has space-coordinate s at time t , then there exists a time $\tau > t$ such that the particle has space coordinate s at τ .

Whatever are the observed behaviors, let us assume they are capturable as a first-order theory. A type of physical model where a language L is interpreted into a universe of objects U such that certain laws (axioms representable as sentences) hold true.

If first-order descriptions are the working methodology of our eliminative school, we will be pleased that investigations into the logos of this universe never escapes our favored ontology. That there are no facts which do not combine our preferred types of objects. These descriptions will tend to picture configurations of U -objects (the formula $\psi(a_1, \dots, a_n)$ as the objects a_1, \dots, a_n arranged in a structure). Assuming a correspondence between descriptions and events, our eliminative school has the epistemological advantage of making correct predictions while simultaneously minimizing entities. We

⁵Set theory is itself formulated using first-order logic. By limited constructive power, I mean a linguistic structure interpreted into a *single set* as a domain of interpretation, and what is expressible through the resulting logical formulae.

eliminate anything outside U because – and this is true by assumption – such entities are not required for explanation. Yet the success of our eliminative approach hides a deficiency. Since we cannot incorporate law-like regularities among properties derived from the substrate. For example, we cannot incorporate the derivation of tangents (1-6), even when they influence the motion of particles, and are immediately implied by particle-space considerations. Assuming these particles are mathematical objects, and deferring to the way *in which mathematics understands mathematical objects*, transverse derivations are anticipated in a complete description.

Selection and Derived-Properties

Thoughts as analogs of tangents is suitable, since our lifelines are also paths of a sort, and as tangents travel along a path so our inner-lives move with our bodies. The mathematical derivation of tangents and thoughts might also share a resemblance: The former uses an equivalency between bodies traveling the same direction, while the later depends on some conjectured equivalence between bodies. Were behaviorism correct: Two bodies are in the same mental-state when they are behaviorally indistinguishable. Were machine-functionalism correct: Two bodies are in the same mental-state when both realize the same machine-state. Either would agree that $M = P/\sim$ where \sim is an equivalence relation on physical-states. Like tangents, thoughts are both supervenient and multi-realizable. Hence tangent-vectors and mental-states might not be so different, both traveling with the moving body, at each point of time assigning an object which expresses a class of equivalent objects in the substrate.

Are tangent vectors causally inert? In a way yes, since to know the path $p: [0,1] \rightarrow S$ is to know the life of a particle completely. Tangents have no power to modify substrate events (determined by the data of paths), and are superfluous to queries like “What is the terminal point of this particle?” or “Will these two particles cross paths?” Tangents cannot add anything to our physical theory, which was described as a first-order linguistic structure (modeling a theory) whose formulas combine objects from the ontological base.

However, imagine a struggle for existence on the surface of S favors particles with the ability to perform right-angles.⁶ Perhaps this angle-making facilitates exact navigation, minimizes expense, or confers some other advantage. Selection will leave those particles with a “causal power” expressed in tangents: At time $t + \epsilon$ the tangent $p'(t + \epsilon)$ is orthogonal to the tangent $p'(t)$, where $\epsilon > 0$ is a small real number. Grid-following particles dominate after sufficiently many generations. By one perspective, behaviors are following grid-like instructions, received from whichever ontological realm contains grids, and by another perspective, advantaged particles are selected from within the substrate.

Let us enrich the thought. For simplicity, let us assume our space S is the real numbers, so that the path of a particle has the form $p: [0,1] \rightarrow \mathbb{R}$. Further assume that p is analytic (ie, has all its derivatives at every time). Then p at a time τ is expressible as a power-series:

$$f(t) = \sum_{n=0}^{\infty} \frac{f^{(n)}(\tau)}{n!} (t - \tau)^n$$

Now instead of single directions (tangents) we have the characteristics $(f^{(0)}(\tau), f^{(1)}(\tau), \dots)$. In a struggle for existence, it might occur that integer coefficients secure a greater advantage. Or coefficients

⁶ Depending on how much smoothness is assumed, perfect-right angles might not be possible.

that are integers and where the 0-th coefficient is a prime number. Or where coefficients where even indices are zero. Under selection, we observe particles – automatons of a sort – struggle into *types* of coefficients. Under selection, certain combinations obtain, others are discarded; regardless, *every* combination is an object “over and above” the substrate. Analogized to minds, supposing a mental-state (as an equivalence class of physical-states) had a form like (a_0, a_1, \dots) we can imagine combinations of those a-terms, like a frequency, corresponding to “pain,” “love,” “surprise,” and so on.

There is one mental-state per physical-state (as the function $e: P \rightarrow M$ maintains), thus it is inconsistent to argue: An organism responded to physical-state p with mental-state m , when there was greater survival value responding m' . The mental-state $e(p)$ can only be what it is. Consider the physical-state p as the entire state of the organisms body, which includes the states of sub-systems (as the contractions of muscles, metabolic processes, communications of the nervous system, etc.). Given that mental-states are consolidated by brain processes, it is consistent for organisms to prosper by coordinating the states of organs and other processes together with appropriate brain-states (thereby, harmonious mental-states). For example, where the anatomy of the physical-state coordinates tissue damage together with brain-state mapping onto a painful mental-state. Natural selection is as *blind* here as it is elsewhere. It does not intentionally design a harmonious correspondence of physical-states and mental-states (as Leibniz’s “preestablished harmony”). Simply: Organisms whose physical-states coordinated tissue damage with a brain-state mapping onto “pain” enjoyed differential success. That pain responses are visible in so many species suggests that among the possibilities of experience produced by physical combinations, the experience of pain is something of a locus which is settled into for of its advantage. Whenever the physical structure of an organism is complexified to the point of accessing the phenomenology of pain, and coordinates appropriate events with its appearance, that configuration prospers.

A *Category* is a family of mathematical objects which together form something like a mathematical species together with structure-preserving transformations between those objects. For example, the category of *groups* with *group homomorphisms* as transformations (maps $f: A \rightarrow B$ such that for $a_1, a_2 \in A$, $f(a_1 \cdot a_2) = f(a_1) \cdot f(a_2)$). These transformations are composable (given group homomorphisms $f: A \rightarrow B$ and $g: B \rightarrow C$ the composition $g \circ f$ is a group homomorphism from A to C since $(g \circ f)(a_1 \cdot a_2) = g(f(a_1 a_2)) = g(f(a_1)f(a_2)) = g(f(a_1))g(f(a_2)) = (g \circ f)(a_1)(g \circ f)(a_2)$). Compositions are associative $h \circ (g \circ f) = (h \circ g) \circ f$, and each object has an identity transformation to itself (the map $id_A: A \rightarrow A$ sending a to a is a group homomorphism).

Imagine the mathematical objects of a category C as the units of selection. Intuitively speaking, two C -objects are isomorphic when they agree on C -properties. Put another way: C -objects are described by their C -properties. Consider this an analogue of physical objects being described by all the attendant physical facts. Given an object $A \in C$ let $Aut(A)$ denote the set of isomorphisms from A to itself (transformations $f: A \rightarrow A$ such that $f^{-1}: A \rightarrow A$ is a transformation). Observe that:

1. If $f, g \in Aut(A)$, then the composition $f \circ g$ is in $Aut(A)$.
2. If $f \in Aut(A)$, then $f^{-1} \in Aut(A)$.
3. The identity $id: A \rightarrow A$ is in $Aut(A)$, and $f \circ id = f = id \circ f$ for any $f \in Aut(A)$.

Therefore $Aut(A)$ is a mathematical group with composition as the group operation. There is a sense that the group $Aut(A)$ somehow *belongs* to A (though A need not have anything whatever to do with

groups). Certainly, it forms part of the description of A , or its classification, and this is because: Given an object $B \in C$, if B is C -isomorphic to A , then $Aut(B)$ is isomorphic to $Aut(A)$ as groups (this perspective has a historically important application to geometry in Klein's *Erlangen program*, which classifies geometric systems according to their symmetry groups).

Using C as an analogue of physical, and imagining generations of C -objects under selection, specific automorphism groups might emerge as dominant. Respective to C these groups are simultaneously inert and significant. Inert, because $Aut(A)$ supervenes on A , so that knowing A gives $Aut(A)$, and conversely $Aut(A)$ cannot add additional properties to A as a C -object. Significant, because $Aut(A)$ characterizes A , and the characteristics A obtains by having the derived-property $Aut(A) = G$ could be significant in C -terms.

The logical contents and insides of A as a C -object are complete and cannot be added to, and yet $Aut(A)$ might tell the most about A in the least space, and, under specific conditions, is pushed to the fore by selection. The mysteries of mind apply *mutatis mutandis* here: Where is $Aut(A)$? Where does it fit into A ? When we open the lid of A , and look at its logical works, there is no group – how can that be explained? The group $Aut(A)$ is determined by A , and not the reverse, therefore it appears to have no behavioral influence at all. And so on.

In a way, nothing strange is happening here: If derived-properties are constructed set-theoretically from the physical substrate, they are determined by physical-states and covary with their modifications; additionally, because derived-properties characterize their physical systems, they are physically detectable and selectable. A derived-property is always *about* its object. When a obtains the derived-property q it becomes of a type- q as the physical structure of the state conforms to that characterization. Given two physical states a and b , if there exists a derived property q such that $q(a)$ but not $q(b)$, then $a \neq b$, and the distance between a and b implied by q could entail measurable effects in the substrate. Selection, as blind towards philosophical disputes as it is anything else, chooses properties by *their measurable effects* without first consulting our opinions about what counts as "material."

This theory resembles Searle's (1992, 2002, 2004) *Biological Naturalism* which conceives of mental phenomena as higher-order properties of physical systems that are causally reducible to physical processes but not ontologically reducible. Indeed, Searle's favorite analogies, like temperature or solidity, have some resemblance to what is meant by a derived-property. The temperature of a system of particles is not reducible to particle-properties and one could not pin-point *the* particle that is -10 degrees Celsius. The same is true of solidity. Though, I would stress that Searle's explanations of ontological irreducibility are different than those proposed here (also less plausible). Searle borrows an intuition of natural science: Physical phenomena have different levels of explanation: A macro-explanation, involving macroscopic objects and their causal relations; and, a micro-explanation, involving the states of all the included particles. Hence bringing water to a boil is explained by putting the pot on the burner, turning oven-nobs, the convection of heat in its macroscopic aspect, etc. It is also explained "because the kinetic energy transmitted by the oxidation of hydrocarbons to the H₂O molecules has caused them to move so rapidly that the internal pressure of the molecule movements equals the external air pressure" (Searle, 1992, p. 87). It is true, and often-recited, that the quantum world and the world of human scale obey different rules and forms of explanation. However, it is uncertain that macroscopic perspectives are enough to transcend the ontological closure of the physical

domain. Physicalism holds that existing objects are bits of matter assembled into higher complexes, and it is not clear how putting together enough bits of matter and then looking at it from outside would be anything but another view of the same material stuff. Assuming classical particles, it is at least plausible to describe macro-properties as states-of-affairs in which particles “hang together like links in a chain.” A macro-property like *solidity* might have a very complicated first-order expression $\psi(a_1, \dots, a_n)$ describing how the particles a_1, \dots, a_n fit together in a lattice structure. According to Wittgenstein (2001), all a state-of-affair like this can do is display objects in a logical structure, and we have seen how combining objects using first-order logical powers is ontologically conservative.

Certainly, the higher-properties of Searle are admissible for selection. Temperature, as average kinetic energy, is selectable since it is possible for specific averages to enjoy differential success. It is simple enough represent the process of selection: A temperature characterizes the behavior of a physical system, when specific temperatures and temperature ranges prove successful, those *numbers* emerge as selected properties. We do not question any of this. But is it so past questioning? A radical skeptic might critique the existence of temperatures as determined by the kinetic energy of the particles, while themselves being unable to add or remove anything from a single particle, nor to add a cause or effect that was not there from the start. Opening the system, and inspecting the particles, a “temperature” is nowhere found. But we know by observation that these higher-properties evolutionarily entrench themselves as features of nature. It is enough to expand the ontology of supervenient properties from higher-properties to *derived-properties*. Then, as our primary mathematical analogies show:

1. At times, derived-properties are the implications and immediate logical growths of a substrate (as tangents are to a world of particles).
2. Derived properties supervene. They are entailed by the substrate, and while being unable to make alterations to that base, they simultaneously characterize (the Automorphism Group).
3. Derived properties are not ontologically reducible, and their native meanings are semantically incongruent with those of the substrate (the Fundamental Group).

Intentionality

How thoughts become about states-of-affairs in the world can be understood through the logic of derived-properties. A tangent vector is about the course of a particle. The Fundamental Group $\pi(S)$ of a space S , is an image of S , a representation of it, in algebraic terms. Derivations perform a type of measurement upon their object, or represent it, and for that reason possess “aboutness.”

Of course, that mentality is about physicality falls short of providing the details of mental representation (say, in the style of Kant, where sense data becomes structured through space, time, etc. into intelligible experiences and the “mental pictures” are exactly described). The premise that $a \rightarrow e(a)$, and that $e(a)$ is about a , is rather silent on what $e(a)$ will be. Still, the thesis of intentionality is broadly corroborated. Abstractly, consider a mapping from the unit interval $[0,1]$ (representative of time) into a family of mathematical objects A :

$$h: [0,1] \rightarrow A$$

Now suppose a derived-property q applies to all A so that $q: A \rightarrow B$. By universal applicability, I mean the way in which the property of *mass* applies to all physical objects (possibly having value zero). Or, where A is a mathematical category, how $Aut(\cdot)$ might be applied to any object $a \in A$ forming the

group $Aut(a)$. By composition, $q \circ h: [0,1] \rightarrow B$. The path of mathematical objects $h(t)$ through A evokes a parallel path of objects $(q \circ h)(t)$ through B . Where at each time $(q \circ h)(t)$ is about $h(t)$ – further, this directedness is inescapable and iron-chained, $(q \circ h)(t)$ can no more stop referring to $h(t)$ than might a color unbind from its object and go its own way.

Considering the life of a human body as function from time into physical-states $h: [0,1] \rightarrow P$ (which is only to say that the state of my body at any time is a physical-state), and by composing with $e: P \rightarrow M$, we obtain $e \circ h: [0,1] \rightarrow M$. The inner-life $(e \circ h)(t)$ runs parallel to the bodily-life $h(t)$, to which it forever refers, being powerless to redirect its directedness.

Mental Causation

Mental-states have the paradoxical quality of being powerless to add cases to a physically closed domain, while also having hyper-significance to the embodied subject (as phenomenological introspection attests). Psychological states are somehow both prominent and inert.

The logic of a mathematical function $p \rightarrow e(p)$ implies that the image $e(p)$ has no power *over itself*, let alone the power to reach backwards and manipulate p . If physical law dictates a causal chain physical-states p_1, p_2, \dots, p_n , mental-states can only *trail behind* $e(p_1), e(p_2), \dots, e(p_n)$ as a *shadow following its object*. At first glance, this appears to affirm the epiphenomenalism of Huxley (2002), whose famous words are worth repeating here:

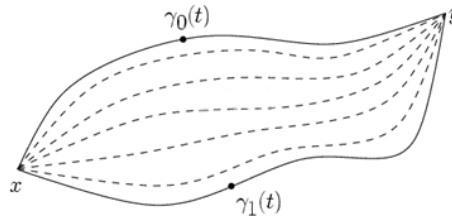
It may be assumed, then, that molecular changes in the brain are the causes of all the states of consciousness of brutes. Is there any evidence that these states of consciousness may, conversely, cause those molecular changes which give rise to molecular motion? I see no such evidence . . .

The consciousness of brutes would appear to be related to the mechanism of their body as a collateral product of its working and to be as completely without any power of modifying that working as the steam-whistle which accompanies the work of a locomotive engine is without influence upon its machinery (p. 29).

As for human mentality “It seems to me that in men, as in brutes, there is no proof that any state of consciousness is the cause of change in the motion of the matter of the organism” (Huxley, 2002, p. 30). Still, mental-states have at least *one power*, which is to characterize or in some way measure their physical-states, and through this they are able to impart physical transformations from their side. This can be seen by analogy to the temperature example. Many organisms (endotherms) regulate the temperature of their bodies by homeostatically responding to disturbances. Write temperature as a function $t: P \rightarrow \mathbb{R}$ from physical-states to real numbers. Suppose a genetic program includes the instruction: If bodily temperature falls beneath T then act to increase temperature, if bodily temperature increases above T then act to decrease temperature. The number T has made a causal entrance, since our organism will always behave in a way to keep $t(p)$ close to T . Hence the body orbits this higher-property – even while $t(p)$ is a kind of *shadow* of the state p . This paradox is resolved by natural selection acting on higher-properties: Temperatures are selectable because a temperature characterizes the physical state of a body. By mathematical example it was demonstrated that derived-properties characterize objects despite being ontologically and semantically separated from them. The function $e: P \rightarrow M$ is much like $t: P \rightarrow \mathbb{R}$, and there is no reason that joys and pains cannot regulate

animal behavior as does hot and cold, the only difference being that the meanings of experience are further from the physical substrate than are temperatures (the latter being immediately read from states-of-affairs).

In what follows, I wish to take an unconventional approach to ontology by imagining an ontological kind as a mathematical category. In the category of topological spaces, the objects are spaces,⁷ and the transformations are continuous functions. Say X, Y are topological spaces and $f: X \rightarrow Y$ is a continuous function. If $p: [0,1] \rightarrow X$ is a path in X , then the composition $f \circ p: [0,1] \rightarrow Y$ is a path in Y . In other words, a transformation between X and Y sends paths in X to paths in Y . Recall the construction of the Fundamental group, without dwelling on the exact details, one would agree that a homotopy of paths in X ,



passed through f , is a homotopy of paths in Y . Blessedly, everything works out so that the spatial transformation $f: X \rightarrow Y$ becomes an algebraic transformation $\pi(f): \pi(X) \rightarrow \pi(Y)$ (ie, a group-homomorphism). Such a mapping between mathematical categories is known as a *functor*. Briefly stated: A functor sends mathematical objects to mathematical objects and transformations to transformations.

Given physical-states $a, b \in P$, let us say that $c: a \rightarrow b$ when a causes b according to the physical causal relation c . Immediately these transformations are composable and associative, and we might also add identity transformations (as a mathematical formality). Therefore, the physical-states P form a mathematical category. When the physical-state a causes the physical-state b , by supervenience, $e(b)$ must follow $e(a)$, and it is felt vividly that the mental event $e(a)$ has *caused* $e(b)$ to appear in the sequel. That it was thirst, the phenomenological thirst, which led to the phenomenological experience of reaching for the glass, and so on. This suggests the physical-causal arrow $c: a \rightarrow b$ entails an arrow of mental-causation

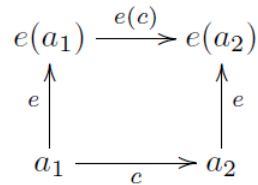
$$e(c): e(a) \rightarrow e(b).$$

Mathematically speaking, e has become a functor (before only a function), which not only sends physical-states to mental-states, but further sends physical causal relations to mental causal relations.

For the above to work, a lawlike connection between physical-states and mental-states must be assumed. Such a hypothesis has its dissenters. Notably, Davidson's (2001) *anomalous monism* denies the existence of such psychophysical laws. It is also necessary to assume the existence of mental causal relations, psychological laws that regulate the stream of consciousness, as the transformations $e(c): e(a) \rightarrow e(b)$. Psychical causal relations cannot be mapped onto mental causal relations if there *are no* mental causal relations. With these assumptions, a plausible account of psychophysical causality

⁷ As before spaces are path-connected to minimize complexity.

can be given. Behavioral programs that employ mental-states for their compassing are effective because mental-states characterize physical-states. A simple behavioral program might take the form: When in physical-state a_1 with $m_1 = e(a_1)$, initiate physical transformations of a_1 until reaching a_2 with mental-state $m_2 = e(a_2)$. For example: When thirsty, act to quench thirst. The program is effective because mental transformations run parallel with physical transformations (through the functor-relation). The transformation of $e(a_1) \rightarrow e(a_2)$ in M occurs in the diagram:



and the physical-state a_2 which fits in the diagram is the one which has accomplished the bodily task.

It can now be understood how mental-states direct bodies: Such occurs when mental-states, as derived properties which characterize their respective physical-states, are pushed to the foreground by selection. When the simple rule “avoid pains, repeat whatever gives pleasures,” performs better than tracking all the biochemical details of the body. Hence the quality of our being which centers the struggle for psychic satisfaction. All share the end of happiness and the satisfaction of drives (in their psychic aspect) while the material details pass as a backdrop. Something like a Hegelian sublation, where mind germinates in the hard clay of matter and shoots upwards until its flowering body is seized upon and the steam-whistle that could only follow locomotion foregrounds itself – perhaps mirroring the very words of Hegel (2001) “The instinctive movement – the inherent impulse in the life of the soul – to break through the rind of mere nature, sensuousness, and that which is alien to it, and attain to the light of consciousness, *i.e.*, to itself” (p. 73). All this transpires under the closure of a closed physical substrate.

On What There Is

Consider Quine’s (1948) *ontological commitment*, which takes the form “A theory is committed to those and only those entities to which the bound variables of the theory must be capable of referring in order that the affirmations made in the theory be true” (p. 33). Euclidean geometry needs points, lines, planes, etc. to make geometric theorems come true, but nothing more. Ptolemy requires celestial bodies, since without them, what moves with celestial motion? Ptolemy does not require sense-data or species of vegetation.

Plainly, many successful physical theories are *physically* affirming. Using our Wittgenstenian picture: A state-of-affairs of the form $\psi(a_1, \dots, a_n)$ occurs. What caused it to be so? Another state-of-affair of the form $\theta(b_1, \dots, b_k)$ (or some collection of these). What are these a and b-terms? They are what the bound variables of the theory must refer to: Physical entities. Complex arrangements of physical constituents, atoms and void, quantum fields, or whatever proves best. Which is to say a physical theory can be (and most often is) fully explanatory without referencing a beyond. So it was that when Lagrange was asked by Napoleon why his work on mechanics made no reference to God, he responded “I had no need for that hypothesis.” There was no need to refer to a deity, or any transcendent over and above the physical, to make the affirmations of the theory be true.

The human subject acts in a physically closed domain. Something “over and above” the physical (like a thought) is not required to explain the next movement of my body or whether I reach here or step there. It is enough to know myself microphysically, in terms of atoms, and charges, and so on. The brain moves through a series of states, each necessitated by the last, and mind is incapable of adding anything to the unfolding of that chain. No reference to a deity, spirit, subjectivity, inwardness, etc. can make anywhere I go come out *truer*. What is odd, the oddity through which mind seems to enter, is how an object or system can be simultaneously determined and insufficiently expressed. Natural numbers are sufficient for the theory of natural numbers since these by themselves make all the formulas of arithmetic true. But is the theory of numbers fully expressed by only referring to numbers? Not in the least. Transcendent objects leave impressions on the natural numbers in a thousand ways. The Riemann Zeta function comes first to mind, which is defined (for specific s) by the series:

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

and is then analytically continued to the complex plane. This function is “about” the natural numbers (for example, the probability that two randomly chosen numbers is relatively prime is given by $1/\zeta(2)$). Its expected (but still conjectural) behavior also determines the distribution of the prime numbers. Should I say “determines”? Of course, the prime numbers have the distribution they will always have the moment that $\{0, 1, 2, \dots\}$ is written down. The case is already the case, and does not require external objects like ζ to be so. In fact, ζ cannot change the position of a single number. Then why does the behavior of ζ promise to resolve deep puzzles within $\{0, 1, 2, \dots\}$? What type of power is that?

Whatever that power may be, modular forms appear to share it. Somehow lining up the patterns of numbers while being descended from another mathematical world. There is a quote ascribed to Martin Eichler, to the effect of “there are five elementary arithmetical operations: addition, subtraction, multiplication, division, and modular forms.” P-adic numbers (and p-adic analysis) have a similar power. But I will not tire the point, as even a basic survey of the most important connections would run for another hundred pages.

The natural numbers are a closed system, able to adjudicate its own facts by internal law, and whose native objects (numbers) are sufficient for those facts to obtain. No further object has the power to make any truth truer. But somehow, consistent with those closures, transcendent objects become so logically entangled as to assert themselves and demand incorporation into the theory. All these numbers have been frozen in place by the immutable laws of arithmetic, and still this petrified world gives itself over to an outer play of higher-objects, none of which can change a thing, and still their super-structural drama transpires.

Such a situation is rather puzzling under conventional analysis: How can a thing be powerless to change what is collected into its influence? How can the steam-whistle following after the train become an “about which” the train organizes? Puzzling, but not inconsistent, since mathematical objects are interconnected this way without contradiction. Any mathematical object is a position within a network of objects, all of which mutually interpose while descending from distinct ontological planes. Seen so, perhaps the puzzle is created by our conceptual biases and philosophical anticipations, our expectations of connection and relation, while the thing in-itself is perfectly consistent.

Monism or Dualism?

It was stated that $e: P \rightarrow M$ is a functor between mathematical categories. Typically, categories silo different species of mathematical object (the example of topological spaces vs groups is instructive here). One could interpret *separate categories* as *separate realms*. Following this, ought we accuse the separation of the physical and the mental into two mathematical categories of dualism? The matter may be one of interpretation.

The physicist has a liberal attitude towards physicality and is ready to include all the mathematical kinds that become tangled in the dynamics of the Universe as *physical*. Indeed, without hesitation, the physicist will use connections of the above form (functors) to bring new objects into the analysis. Mentality might be *just another one of those*. Under this interpretation, mentality is a branch of physicality as forces and fields are branches.

Still, the relation of the mental to the physical could also be defended as dualistic, at least where monism means an undivided mathematical kind. If physical monism is given Smart's (2002) meaning "There does seem to be, so far as science is concerned, nothing in the world but increasingly complex arrangements of physical constituents" (p. 61), then the monist picture is incomplete, as there is an exterior, or a dual side, forming part of reality. Under this interpretation, mentality is dualized by misfitting ontological criteria.

Property dualism is another viable interpretation. So many physical models incorporate distinctive mathematical kinds for the purpose of describing what is ultimately the motion of one physical substance, and the epistemological form of one substance interpreted under mathematical pluralism could be seen as paradigmatic. Under this interpretation, when physical processes are interrupted, when the body dies let us say, the soul perishes at once, since it was always a property of that body and has nowhere else to be.

But stranger dualistic possibilities remain open. As any mathematician knows, a functor between categories $F: X \rightarrow Y$ *does not* imply that the objects of Y *depend* on the objects of X . Appealing once again to the Fundamental Group: The Fundamental Group of the circle are the integers under addition, but the integers are not ontologically dependent upon the circle, since other shapes share that fundamental group, and secondly, other structures besides spaces are connected to the integers (one of these being the integers in themselves). There is a functor $\pi: Spaces \rightarrow Groups$ mapping multiple spaces to the integers, and further, there exist functors from other categories $F: C \rightarrow Groups$ connecting non-spatial structures to the integers. Such variable connections lends the mathematical object a certain independence. Analogized to mind: A primitive emotion is realizable across the animal kingdom (by the assumption of multi-realizability), it is also realizable on distant worlds, strange worlds, and possible worlds – further off still, in a soup of universes, where ours is but one floating bubble, and others drift bearing their own rules and their own physicality, within any of these the same feeling could appear. With all these connections reaching up to the same emotion, it begins to seem as though mind is something its own. My thought is present now, but it has appeared to another, and in the unlimited combinatorics of all that is, it may have occurred infinitely many times already, or perhaps it has *always been*. Under this interpretation, the soul is rescued from matter:

As the sun, who is the eye of the world,

Cannot be tainted by the defects in our eyes

Or by the objects it looks on,
So the one Self, dwelling in all, cannot
Be tainted by the evils of the world.
For this Self transcends all!
(The Upanishads, p. 88)

As with other grand questions, the traces of fact leave so many gaps as to make answers a matter of interpretation. As to why suffering exists or the purpose of humanity, there is a scattering of fact and the Universe's characteristic silence. Or as for the meaning in one's own life, there are many disparate fragments, but no clear answers. The mathematical connection between minds and bodies is clear enough, supposing what has been said is correct, what it *means* is a hermeneutical project.

References

- Davidson, D. (2001). *Essays on Actions and Events*. Oxford University Press.
<https://doi.org/10.1093/0199246270.001.0001>
- Dennett, D. C. (1991). *Consciousness explained*. Black Bay Books.
- Hegel, G. (2001). *The Philosophy of History* (J. Sibree, Trans.). Batoche Books.
- Huxley, T. (2002). On the Hypothesis That Animals Are Automata, and Its History. In D. Chalmers (Ed.), *Philosophy of Mind: Classical and Contemporary Readings* (1st ed.) (pp 24-30). Oxford University Press.
- Quine, W. V. O. (1948). On what there is. *The Review of Metaphysics*, 2(5), 21-38.
- Searle, J. (1992). *The Rediscovery of the Mind*. MIT Press.
- Searle, J. (2002). Why I am not a property dualist. *Journal of Consciousness Studies*, 9(12), 57-64.
- Searle, J. (2004). *Mind: A Brief Introduction*. Oxford University Press.
- Smart, J. J. C. (2002). Sensations and Brain Processes. In D. Chalmers (Ed.), *Philosophy of Mind: Classical and Contemporary Readings* (1st ed.) (pp 60-68). Oxford University Press.
- The Upanishads* (E. Easwaran, Trans.; 2nd ed.). (2007). Nilgiri Press.
- Wittgenstein, L. (2001). *Tractatus Logico-Philosophicus* (D. F. Pears & B. F. McGuinness, Trans.). Routledge.