Consciousness Transfer and the Impossibility of Uploading: Continuity, Indexicality, and the Limits of Representation

Flyxion
Independent Researcher

Abstract

This essay argues that consciousness uploading, understood as the transfer of first-person experience to an artificial substrate, is conceptually incoherent. The argument proceeds by distinguishing first-person continuity from third-person pattern equivalence and demonstrating that standard uploading scenarios preserve only the latter. By analyzing the roles of indexicality, process continuity, measurement, and embodiment, the essay establishes that copying cannot constitute transfer. The argument is formalized within the RSVP (Relativistic Scalar-Vector Plenum) field-theoretic framework, where consciousness is modeled as a continuous trajectory in a constrained dynamical system. Identity is defined as the uniqueness of a maximal integral curve within a coherence manifold $\mathcal{M}$, characterized by coupling, entropy-flow, and Lyapunov-stability constraints. Under this formalism, uploading is a set-valued branching operator that violates uniqueness; irreversibility of identity loss and inevitable fragmentation under instability follow as structural theorems. A coherence functional $\mathcal{C}$ tracks continuous degradation of phenomenological unity while preserving a binary criterion for ontological identity. A flux-based geometric criterion unifies the dynamical, thermodynamic, and topological dimensions: a subject persists if and only if a coherent flux structure across a continuously deforming boundary admits an entropy-consistent extension in time. This framework is extended to decision theory, where a theorem establishes that identity-consistent utility functionals are invariant under operations on trajectories outside the subject's equivalence class, dissolving Basilisk-type acausal scenarios across causal, evidential, and functional decision theories. Uploading is replacement, not survival.

Contents

1 Introduction

The idea that the mind might be extracted from the body and installed in a more durable medium has moved, within a generation, from the margins of speculative fiction into the programs of serious technology companies and the papers of academic philosophers. The appeal is obvious. If the self is nothing more than a pattern of information—a particular arrangement of functional relationships among computational units—then in principle that pattern might be copied to a substrate that does not age, does not fail, and does not die. Death would become optional. The self would persist as long as its host medium was maintained.

This essay disputes that conclusion, not by denying that the relevant pattern could be copied, but by arguing that copying is not the same as transfer. The question at the center of the uploading debate is not whether a behaviorally or structurally indistinguishable duplicate can be produced. It almost certainly can, given sufficient technology. The question is whether the original first-person perspective continues to exist in the copy. That question requires a careful distinction between two fundamentally different criteria for what it means to survive: continuity of a physical process, and equivalence of an informational pattern.

The argument of this essay is that survival, properly understood, requires the former. A system that preserves the pattern while destroying the process has not saved anyone. It has ended one subject and started another who happens to begin with the same memories, dispositions, and self-model. The resulting entity will believe itself to be the original, will behave as the original would have behaved, and will be indistinguishable from the outside. But from the inside—from the only perspective that matters for questions of survival—there will have been no experience of continuity, because the experiential thread was severed at the moment of copying.

This claim is not new. It echoes arguments made by Derek Parfit, though Parfit drew the opposite conclusion: if identity is not what matters in survival, then copying might be just as good as biological continuation. This essay rejects that move. The reason identity matters is precisely that experience is indexed to a single physical locus. There is no view from nowhere in first-person experience. Every experience is had by a particular system at a particular time and place, and the continuity of that system is not incidental to the experience but constitutive of it.

The structure of the essay is as follows. Section 2 establishes the core ontological divide between pattern-based and continuity-based theories of mind. Section 3 develops the map-territory distinction as applied to consciousness. Section 4 analyzes the role of spatiotemporal indexicality and argues that any branching procedure fails as transfer. Sections 5 through 8 extend the argument through analyses of dynamical continuity, the measurement problem, analog processes, and embodiment. Section 9 addresses the deeper category error of treating consciousness as information. Section 10 explains

why the opposing view nonetheless remains compelling. Sections 11 and 12 draw the central conclusion. Sections 13 and 14 extend the argument into a formal dynamical framework, which is then fully grounded in the RSVP field theory in the sections that follow. Section 36 applies the framework to Roko's Basilisk and derives formal results in decision theory, including a theorem establishing that identity-consistent utility functions are invariant under operations on trajectories outside the subject's equivalence class.

2 Two Competing Ontologies of Mind

At the root of nearly every disagreement about uploading lies a prior disagreement about what kind of thing a mind is. Two broad positions dominate the debate, and they are not minor variations on a shared theme. They are fundamentally different accounts of what it means to be a subject.

The first position is functionalism, or more broadly the pattern view. On this account, what makes a system conscious is the organization of its causal relationships, not the specific material that instantiates them. If a system processes information in the right way—if it takes inputs, transforms them according to the right structure, and produces outputs that stand in the right relations—then it is conscious, regardless of whether it is made of neurons, silicon, or any other physical medium. Identity, on this view, is preserved whenever the pattern is preserved. If an exact functional replica of your brain is instantiated on a computer, that computer is you, or at least is as good as you, in every sense that matters.

The second position holds that identity is grounded not in pattern but in process. On this account, what makes you the same person across time is not that your current state matches your previous state in some abstract functional sense, but that your current state is causally descended from your previous state through an uninterrupted physical process. The thread of experience is a thread precisely because it is continuous. There is no gap, no jump, no moment at which the original process stops and a copy begins. Identity is constituted by the continuity of that process, and any interruption—however seamless the reconstruction—breaks the thread.

These two positions generate radically different verdicts on uploading. For the functionalist, a sufficiently accurate upload is straightforwardly survival. For the continuity theorist, it is death followed by the creation of a convincing impersonator.

It is important to notice that this is not merely a verbal dispute. The two positions disagree about which facts are relevant. The functionalist points to structural and behavioral equivalence. The continuity theorist points to the physical history of the process. Neither is simply ignoring the other's evidence; they are operating with different criteria for what counts as evidence. The functionalist treats third-person observables as decisive. The continuity theorist holds that no third-person observable

can settle a first-person question.

The essay takes the continuity position as its starting point, not as an arbitrary stipulation, but because it follows from taking first-person experience seriously as the subject matter of the inquiry. If the question is whether I will survive uploading—not whether someone very similar to me will exist afterward—then the relevant criterion must be one that is sensitive to my particular perspective, not merely to the existence of a perspective that resembles mine. The pattern view answers a different question. It tells us whether something functionally equivalent will exist. It does not tell us whether that thing will be me.

3 The Map–Territory Distinction

There is a persistent confusion in discussions of mind uploading that mirrors the confusion between a map and the territory it represents. A sufficiently detailed map of a city is not the city. No matter how accurate, how high-resolution, or how comprehensive the map becomes, it remains a representation of the city rather than the city itself. This is not a limitation that can be overcome by increasing the resolution. It is a categorical distinction between two different kinds of thing.

The same distinction applies to consciousness. An informational description of a conscious process, however complete, is a representation of that process, not the process itself. The uploading scenario typically imagines that if you capture enough information about the brain—its connectivity, its activity patterns, its molecular configuration—you have captured the mind. But what you have captured is a description of the mind. The description may be extraordinarily accurate. It may support the construction of a system that behaves identically to the original. But behavior is itself a third-person observable, and the question being asked is a first-person one.

Consciousness is not a static object. It is not a configuration that can be read off and reinstantiated elsewhere. It is a lived process: an ongoing, dynamic, causally self-sustaining event. The mistake of the uploading scenario is to treat experience as though it were a file that can be copied from one device to another. But experience is not stored anywhere. It is not a record of itself. It is an occurrence.

Consider an analogy. A flame is a process: an ongoing chemical reaction sustained by fuel, oxygen, and heat. You can describe a flame in extraordinary detail—its temperature profile, its chemical composition, its spectral output. You can use that description to start a new flame with identical properties. But you have not transferred the original flame. You have extinguished it and started a new one. The new flame is not the old flame; it is a new fire that resembles the old one. The description was complete, but the description was not the thing.

The functionalist will object that a mind is not like a flame, because what matters about a mind is its functional organization, and functional organization can in principle

be reinstantiated exactly. But this response begs the question. It assumes that what matters about a mind is its functional organization—that is, precisely what is at issue. The continuity theorist denies this assumption. What matters about this mind—my mind—is not its abstract functional organization but the particular physical process that I am. The description of that process, however complete, is not that process.

4 Spatiotemporal Indexicality and the Failure of Transfer

First-person experience is not just a type of experience; it is experience had by a particular system at a particular place and time. This is what philosophers call indexicality: the dependence of a content on the context of its occurrence. The word "here" refers to a different place depending on where it is uttered. The word "now" refers to a different time depending on when it is used. Similarly, the experience of being me refers to a different subject depending on which physical system is having it.

This indexical character of experience is not a superficial feature that could be stripped away while leaving the content intact. It is constitutive of what first-person experience is. To experience something is to experience it from somewhere. There is no experience that is not located—no consciousness that floats free of a particular physical perspective. The location is not an additional property of the experience; it is part of what makes the experience the particular experience it is.

Now consider what happens in a scan-and-reconstruct uploading procedure. At some time T, the original system is scanned. The scan captures, with whatever degree of accuracy, the state of the system at that moment. This information is then used to initialize a new system, which begins its operation from that reconstructed state. The new system has, from the outside, the same history as the original up to time T. Its first-person reports will be consistent with having been the original system all along.

But there are now two trajectories in the world: the original system, which either continues (if the procedure was non-destructive) or is terminated (if it was destructive), and the new system, which begins its evolution from the reconstructed state. These two trajectories share a common past up to T but diverge afterward. Neither trajectory can be uniquely identified as the continuation of the original subject, because the indexical position of the original subject—its particular location in spacetime, its particular causal history—is not something that can be copied. It is tied to the original process.

If the original system is destroyed, the situation is not improved. Destruction of the original does not transfer the indexical position to the copy. Destruction terminates one trajectory; reconstruction begins another. There is no mechanism by which the "first-person locus" migrates from the old trajectory to the new one, because that locus is not a property that can be transferred. It is constituted by the trajectory itself.

The branching case makes this clearest. If the procedure is non-destructive, there are two successors. Both claim, sincerely and from the inside, to be the original. Both

have equal claim, on the pattern view, to be the continuation of the original. But they cannot both be right in any first-person sense, because their experiences have already diverged. The original subject was one, and one has become two. Whatever happened to the original subject, it was not preservation.

5 Continuity Versus Branching

The concept of personal identity over time has been analyzed through a long tradition of thought experiments involving gradual change, teleportation, fission, and fusion. What emerges from this tradition—and what the uploading debate tends to obscure—is a crucial distinction between two fundamentally different kinds of change: continuous evolution and branching replication.

Continuous evolution is what happens in ordinary biological life. The cells of your body are replaced over years. Your neural connections rewire. Your memories are reconstructed each time they are recalled. Your personality changes with experience. Yet through all of this, you persist. The reason is not that some immaterial essence is preserved, nor that the pattern remains constant (it does not), but that the process is continuous. There is no moment at which the old you stops and a new you begins. The change is incremental, causally connected, and unbroken.

Branching is categorically different. Branching occurs when a single process divides into two or more independently evolving processes. The original process does not continue in either branch; it terminates at the branch point. What continues is not the original process but two new processes that share a common origin. The distinction is not one of degree. A branching event is a topological change in the structure of the process, not a rapid or drastic continuous change.

Consider the following thought experiment. Suppose that over the course of a year, every neuron in your brain is gradually replaced with a functionally identical artificial neuron. At no point is the system's operation interrupted. Each replacement is instantaneous, each artificial neuron is connected to its neighbors in exactly the same way as its biological predecessor, and the system continues to process information without interruption throughout. At the end of the year, your brain is entirely artificial, but the process has been continuous. There was never a moment at which your experience stopped. On the continuity view, you have survived. The replacement was real, but the continuity was also real, and continuity is what matters.

Now suppose instead that your brain is scanned in a single instantaneous operation, the scan is used to initialize an artificial system, and your biological brain is destroyed. The result is structurally identical to the outcome of the gradual procedure: an artificial system with exactly your functional organization. But the process was not continuous. There was a moment—the moment of destruction and reconstruction—at which the original process stopped and a new one began. That moment is not a matter of degree.

It is the moment at which the trajectory terminated.

The uploading scenario necessarily involves branching or termination. It cannot be otherwise, because the whole point of the procedure is to transfer the pattern from one substrate to another. Transfer implies that the pattern leaves one medium and enters another. Leaving one medium means that the process on that medium stops. Entering another medium means that a new process begins. The gap between those two events—however short—is the moment at which identity is lost.

6 The Measurement–Destruction Problem

Even setting aside the philosophical arguments about continuity and branching, there is a further problem that is internal to the uploading scenario itself: the requirement of high-resolution state capture. To reconstruct a conscious system from its description, one must first capture that description. And the process of capturing the description of a dynamical system is not a neutral operation. It necessarily disrupts the system being described.

This is most obvious at the quantum level. To measure the state of a quantum system is to collapse its wavefunction, destroying the very superpositions that might be essential to the system's operation. Whether or not consciousness depends on quantum effects—a contested question—the principle generalizes. Any measurement that is precise enough to capture the full state of a system at a moment in time must interact with that system strongly enough to alter it. There is no observation without disturbance, and the disturbance required for complete state capture is not a minor perturbation but a fundamental intervention.

But the deeper problem is independent of quantum mechanics. Even in a purely classical setting, consciousness is not a state but a trajectory. What you are, at this moment, is not a snapshot but a flow. Your experience right now is not identical to the configuration of your neurons at this instant; it is the dynamic unfolding of that configuration across time. The "present moment" of consciousness is not a mathematical point but an interval, a specious present in which past and anticipated future are both actively represented.

A complete state capture, even if it were possible without disturbance, would record only a single instantaneous configuration. It would be like trying to capture a river by photographing the water at a single moment. The photograph captures the position of the water molecules at that instant. It does not capture the river, because the river is not the position of its molecules at an instant. It is a process: a pattern of flow that is constituted by its history and its ongoing dynamics. To reconstruct a conscious system from a state snapshot is to attempt to reconstruct the river from a photograph. Even if you succeed in creating a system with exactly those initial conditions, what you have created is a new river, not the original one.

7 Analog Processes and Digital Approximation

The uploading scenario typically assumes that consciousness can be adequately modeled as a digital computation. This assumption is rarely made explicit, but it is implicit in the claim that a sufficiently precise description of the brain can serve as the basis for reconstruction. A description is necessarily discrete: it represents continuous quantities by finite approximations. The question is whether those approximations are good enough.

For engineering purposes, approximations are often adequate. A digital simulation of a physical system can be accurate enough for practical purposes even if it does not exactly reproduce every continuous feature of the original. But the uploading scenario is not an engineering problem. It is an identity problem. And for identity problems, approximation may not be sufficient.

Consider the possibility—which cannot currently be ruled out—that consciousness depends on exact physical dynamics rather than approximate functional equivalence. If the relevant processes are continuous and sensitive to initial conditions in the way that chaotic dynamical systems are, then any finite approximation will diverge from the original over time. Two systems that begin with states differing by an arbitrarily small amount will, over time, follow entirely different trajectories. A digital approximation, no matter how precise, is not the same dynamical system as the original. It is a system with different dynamics, whose trajectory will diverge from the original's on a timescale determined by the Lyapunov exponents of the underlying dynamics.

The functionalist response is that consciousness supervenes on functional organization at a coarse enough grain that these differences do not matter. But this response assumes what it is trying to prove. It assumes that only coarse-grained functional organization is relevant to consciousness. The continuity theorist denies this. If consciousness is constituted by fine-grained physical dynamics—if the exact trajectory matters, not just its coarse-grained image—then digital approximation is not a reliable route to identity preservation.

There is a further point about the substrate itself. A digital system implements its computations by physical processes that are entirely different from the processes occurring in a biological brain. The silicon switches at a different rate, uses different energy scales, and operates according to different physical principles. Even if the input-output behavior is identical, the underlying physics is not. Whether that underlying physics is relevant to consciousness is precisely the question at issue, and the functionalist cannot answer it by assuming that only the input-output behavior matters.

8 Embodiment and Biological Substrate

The uploading scenario typically treats the brain as the seat of consciousness and the body as its vehicle. On this picture, the relevant information is contained in the neural

architecture, and the rest of the body—the hormonal system, the immune system, the enteric nervous system, the sensorimotor apparatus—is just delivery infrastructure. If you can capture the brain, you have captured the mind.

This picture is increasingly at odds with what biology has revealed about the integration of mind and body. The brain does not operate in isolation. It is continuously coupled to the rest of the organism through a dense network of biochemical signals. Hormones modulate neural activity on timescales ranging from seconds to years. The immune system communicates with the brain through cytokines and affects mood, cognition, and motivation. The gut contains approximately as many neurons as the spinal cord and produces neurotransmitters that influence the brain's function. The body's proprioceptive and interoceptive signals are not merely inputs to a central processor; they are part of the computational process itself.

The mind, on this view, is not localized in the brain. It is a property of the whole organism in its environment. The brain is the most important node in a network, but it is a node in a network, not a standalone processor. Consciousness arises from the integrated operation of the entire biological system, not from the neural component alone.

Substrate replacement is therefore not a neutral operation. Replacing the biological substrate with an artificial one does not merely change the medium; it changes the system. An artificial system that replicates the brain's input-output behavior while running on different physics, without the hormonal, immune, and interoceptive coupling of the biological organism, is not running the same process. It is running a simulation of the process, and a simulation, however accurate, is not the thing it simulates.

This point connects to a broader argument about the role of the body in constituting experience. The phenomenological tradition, from Husserl through Merleau-Ponty, has argued that embodiment is not an incidental feature of consciousness but its ground. Experience is always already bodily: structured by the possibilities and constraints of the organism's physical form. The sense of spatial orientation, the feeling of effort, the experience of fatigue, the visceral texture of emotion—these are not representations computed by the brain and projected onto a neutral medium. They are features of the organism's engagement with its world. Replacing the organism with a different physical system is not preserving the ground of experience; it is removing it.

9 The Category Error of Information

The deepest error in the uploading scenario is the assumption that consciousness is a kind of information. This assumption is so pervasive that it often goes unnoticed. If the mind is information, then the question of uploading becomes a question of copying, compression, and transmission—engineering problems that are difficult but not in principle insoluble. If the mind is not information, then the entire framework dissolves.

Information, in the technical sense, is a relational and observer-dependent quantity. Shannon information is defined in terms of probability distributions over possible states; it measures the reduction in uncertainty that an observation provides. Kolmogorov complexity defines information in terms of the length of the shortest program that generates a given output. In both cases, information is not a property of a physical system in isolation but a property of a system relative to an observer and a description scheme.

Experience, by contrast, is intrinsic. There is something it is like to be a conscious system, and that something is not relative to an observer. It is not a reduction in someone else's uncertainty; it is an occurrence. Pain is not information about tissue damage; it is pain. The redness of red is not a signal transmitted from the environment to the brain; it is a qualitative character of experience that cannot be captured in any amount of functional or physical description.

This is, of course, the hard problem of consciousness. The point here is not to solve that problem but to notice that the uploading scenario presupposes a solution. It presupposes that once you have captured the functional and structural description of the brain, you have captured everything relevant to consciousness. But this presupposition is precisely what is in dispute. If there is something it is like to be you that is not exhausted by the functional and structural description, then copying the description does not copy that something.

The error is a category error: treating something intrinsic as though it were relational, treating an occurrence as though it were a representation, treating experience as though it were data. Data can be copied without loss. Experience cannot be copied at all, because copying produces a second occurrence, not a transfer of the original.

10 Why the Pattern View Persists

The arguments of the preceding sections notwithstanding, the pattern view retains a powerful intuitive appeal, and it is worth understanding why. The appeal is not simply confusion or wishful thinking. It reflects a genuine and important insight about the nature of mind—an insight that the continuity view must accommodate without surrendering its central claim.

The insight is this: what we care about, in caring about persons, is largely captured by their patterns. When we recognize someone, we recognize their characteristic ways of speaking, moving, and thinking. When we grieve for a person who has changed radically through injury or illness, what we grieve is the loss of the pattern, not the loss of the biological substrate. When we imagine surviving in a digital form, the appeal is that the pattern—the memories, the personality, the values, the characteristic ways of engaging with the world—would continue.

This is a genuine feature of our ordinary concept of personal identity. We do not

normally think that identity requires the exact preservation of every physical component. Gradual change is compatible with identity; otherwise none of us would survive from one decade to the next. The pattern view captures this intuition elegantly by identifying identity with structural invariants rather than material continuity.

The problem is not that the pattern view is wrong about what we value. The problem is that it conflates two distinct questions: the question of what we value in persons, and the question of what constitutes survival in the first-person sense. These questions come apart in the uploading scenario. What we value in a person might be fully preserved in a high-fidelity replica. But whether the original person survives is a separate question, and it turns on whether the original experiential thread continues, not on whether the pattern is instantiated in a new system.

The pattern view, in effect, answers the question "would there be someone like me?" rather than the question "would I be there?" From a third-person perspective, a pattern-preserving upload is as good as survival. From a first-person perspective, it is not survival at all. The pattern view is compelling because the third-person perspective is the natural one for thinking about other people. But when the question is one of survival, the first-person perspective is the only one that matters.

11 First-Person Continuity as the Criterion of Survival

The argument of this essay converges on a single criterion: survival requires the continuation of the same experiential stream. This criterion is not chosen arbitrarily. It follows from the structure of the question itself. To ask whether I will survive is to ask whether there will be an experience of being me that is continuous with my current experience. It is not to ask whether there will be a system that would describe itself as me, or that would have memories consistent with being me, or that would behave as I would behave.

The experiential stream is not an abstract entity. It is the ongoing flow of experience, each moment of which is causally connected to the previous moment through the continuous operation of the physical system that constitutes me. That causal connection is not merely a matter of informational resemblance. It is a matter of physical continuity: the system at this moment is the same system as the system at the previous moment, not in the sense that it is made of the same material, but in the sense that it is the same ongoing process.

This criterion has an important implication for the standard thought experiments. When Parfit argues that what matters in survival is psychological continuity and connectedness, and that identity is not what matters, he is making a claim about what we should care about. The response here is not that Parfit is wrong about what we should care about—that is a separate debate—but that he is wrong about what survival is. Even if it were rational to be indifferent to one's own survival in the first-person sense, that

does not make survival identical to psychological continuity. The concept of survival has content that the concept of psychological continuity does not capture.

The experiential stream criterion also illuminates the structure of the gradual replacement scenario. If neurons are replaced one at a time, with continuous operation throughout, the stream is never interrupted. Each moment of experience is causally descended from the immediately preceding moment. The process is continuous even though the material changes. This is not because the material doesn't matter—it matters precisely because it is the physical basis of the process—but because what matters about the material is that it sustains the process without interruption, not that it remains chemically identical.

12 Uploading as Replacement Rather Than Survival

The preceding arguments support a single clear conclusion: what is produced by a scan-and-reconstruct uploading procedure is not a surviving subject but a replacement subject. The two subjects—the original and its replacement—share a history up to the moment of copying. They share memories, dispositions, self-models, and values. From the outside, they are indistinguishable. The replacement sincerely and correctly believes itself to be the continuation of the original, because all of its memories are consistent with that belief.

But the replacement is not the original. It is a new process that began at the moment of reconstruction. Its memories of being the original are accurate—they accurately represent the history of the system whose state was captured—but the replacement did not live that history. It was initialized with a description of that history. This is the difference between remembering and being initialized with a record of remembering. From the inside, if the reconstruction is accurate enough, there is no experiential difference. But from the outside—from the perspective of anyone who observed the process—the original was ended and a new subject was begun.

The point is not that the replacement is a fraud or a lesser being. The replacement is a genuine subject, with genuine experiences and a genuine claim to exist. The point is that it is not the original subject. The original subject is gone. Whatever comfort the existence of the replacement provides—to observers, to the original's loved ones, to the replacement itself—that comfort should not be mistaken for evidence that the original survived.

This conclusion has a practical implication. If you are considering uploading as a means of survival, you should be aware that what you are actually doing—if the uploading scenario is coherent at all—is creating a successor rather than continuing yourself. Whether that is good enough is a question about values, not about identity. Some people, on reflection, may find the existence of a well-configured successor adequate consolation. Others may not. But the choice should be made with clarity about

what is actually being chosen.

13 Implications for Philosophy of Mind and Artificial Intelligence

The argument of this essay has consequences that extend beyond the uploading debate. It bears on several foundational questions in philosophy of mind and artificial intelligence.

For functionalism, the argument poses a direct challenge. Functionalism holds that mental states are defined by their functional roles: their causal relations to inputs, outputs, and other mental states. On the strongest version of functionalism, any system that implements the right causal structure has the same mental states, regardless of its physical implementation. The present argument does not refute this claim, but it shows that even if functionalism is true as a theory of mental state types, it does not follow that type-identical systems are token-identical. Two systems that implement the same functional organization are type-identical (in the same mental state type) but numerically distinct (different individual systems). The uploading scenario involves creating a new token of the same type, not preserving the original token.

For artificial intelligence, the argument suggests that behavioral equivalence—even exact behavioral equivalence at all levels of description—does not imply experiential identity. A system that passes every behavioral test for consciousness, including introspective reports, responses to novel situations, and creative behavior, may still not have the particular first-person experience of any given human system. This is not a claim that AI systems cannot be conscious; it is a claim that behavioral evidence alone cannot settle the question of which experiences belong to which subject.

For theories of personal identity more broadly, the argument supports a process-based account over state-based accounts. Identity is not determined by any snapshot of the system—not by its memories, its personality, its physical configuration, or its information content—but by the continuity of the process that produces all of these. This has implications for questions about identity in sleep, anesthesia, freezing, and other states in which the brain's activity is reduced or suspended.

14 Toward a Formal Account of Continuity

The preceding sections have argued philosophically for a continuity-based criterion of personal identity and against the possibility of uploading as a form of survival. In this section and those that follow, the argument is formalized within a dynamical systems framework and then grounded in the specific commitments of the RSVP field theory.

The motivation for formalization is not merely expository. The philosophical arguments of the preceding sections are, in the end, arguments about the structure of a concept. Formalization makes that structure explicit and allows the key claims to be stated as theorems that follow from definitions rather than as intuitions that must

be weighed against other intuitions. If the definitions are accepted—if one grants that identity is constituted by continuity of a dynamical trajectory—then the impossibility of uploading follows with logical necessity.

The core idea is to model a conscious subject as a trajectory in a state space, where the trajectory is the solution to a system of differential equations. Identity corresponds to the uniqueness of this trajectory. A process that produces two or more continuations of a single trajectory necessarily destroys that uniqueness and therefore fails to preserve identity. This is not a metaphor for the philosophical argument; it is the philosophical argument, made precise.

In the sections that follow, the abstract state space is first specified (Section 15), the identity criterion is stated as a definition (Section 16), uploading is modeled as a branching operator (Section 17), and the central impossibility theorem is proved (Section 18). The argument is then extended to the RSVP field-theoretic framework (Sections 21 onwards), where the state space is replaced by the field triple $(\Phi, \mathbf{v}, S)$ and the dynamics are those of the RSVP partial differential equations.

15 Formalization: Consciousness as a Continuous Trajectory

Let $\mathcal{X}$ denote the admissible state space of a physical system supporting consciousness. An element $x \in \mathcal{X}$ encodes the full physical configuration of the system: its neural activity, biochemical state, and all lower-level physical variables relevant to the system's dynamics.

A conscious process is identified not with a single state x but with a continuous trajectory

$$\gamma : [0, T] \to \mathcal{X}, \quad t \mapsto x(t),$$

where γ is differentiable and satisfies the governing dynamics of the system.

Definition 15.1. The **experiential thread** associated with a trajectory γ is the equivalence class

$$[\gamma] = \{\gamma \circ \phi \mid \phi : [0, T] \to [0, T] \text{ is a monotone bijection}\}$$

under time reparameterization. The experiential thread captures the dynamical content of the trajectory independently of any particular clock parameterization.

Identity of a subject is then identified with a single connected trajectory class $[\gamma]$. Two subjects sharing the same experiential thread are the same subject. Two subjects with distinct experiential threads are distinct subjects, even if their trajectories agree at every time step.

The significance of this definition is that it grounds identity in the global structure of the trajectory, not in any local property. Two trajectories may agree on an arbitrarily long initial segment while differing in their continuation, and this difference is sufficient to make them distinct trajectory classes.

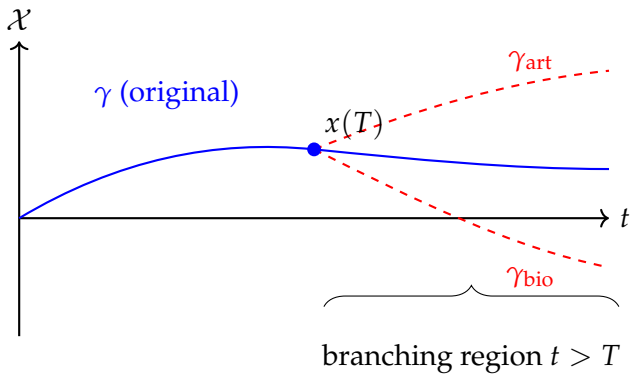

Figure 1: An uploading procedure at time T produces two distinct trajectories γ_{bio} and γ_{art} from the common history γ. Neither continuation uniquely inherits the identity $[\gamma]$.

16 Continuity as the Criterion of Identity

Definition 16.1. A physical process **preserves identity** if and only if it extends the trajectory γ continuously. Formally, for any extension $\tilde{\gamma}$ defined on $[0, T']$ with $T' > T$, identity is preserved only if

$$\tilde{\gamma}(t) = \gamma(t) \quad \forall t \in [0, T].$$

No branching is permitted: if a process produces two trajectories γ_1 and γ_2 such that

$$\gamma_1(t) = \gamma_2(t) \quad \text{for } t \in [0, T], \quad \text{but} \quad \gamma_1 \neq \gamma_2 \text{ for } t > T,$$

then identity is not uniquely preserved.

This definition encodes the continuity criterion in precise terms. It rules out two kinds of failure. The first is replacement: terminating γ and starting a new trajectory $\tilde{\gamma}$ with a different initial condition. The second is branching: producing two continuations that share the common history $\gamma|_{[0,T]}$ but diverge thereafter.

The criterion has an important structural consequence. It implies that the space of identity-preserving processes is exactly the space of continuous extensions of γ within $\mathcal{X}$. Any process outside this space—any process that either halts the trajectory, reinitiates it, or forks it—fails to preserve identity.

Remark 16.2. The criterion does not require that the material substrate of the system remain unchanged. It requires only that the trajectory remain continuous. Gradual material replacement is compatible with identity preservation provided the dynamics remain continuous throughout. What is excluded is any discontinuity in the trajectory itself.

17 Uploading as a Branching Operator

An uploading procedure is modeled as an operator

$$\mathcal{U} : \mathcal{X} \to \mathcal{X} \times \mathcal{X},$$

which, given a state $x(T)$, produces two successor states:

$$\mathcal{U}(x(T)) = (x_{\text{bio}}(T + \varepsilon),\ x_{\text{art}}(T + \varepsilon)).$$

These generate two trajectories:

$$\gamma_{\text{bio}} : [0, T'] \to \mathcal{X}, \qquad \gamma_{\text{art}} : [0, T''] \to \mathcal{X},$$

satisfying

$$\gamma_{\text{bio}}(t) = \gamma_{\text{art}}(t) = \gamma(t) \quad \text{for } t \leq T,$$

but

$$\gamma_{\text{bio}}(t) \neq \gamma_{\text{art}}(t) \quad \text{for almost all } t > T.$$

The operator $\mathcal{U}$ is therefore set-valued with respect to trajectory continuations. It does not define a unique continuation of γ beyond T; it defines two continuations. This is the formal representation of branching, as illustrated in Figure 1.

The destructive variant of uploading—in which the biological system is terminated and only the artificial system continues—corresponds to restricting $\mathcal{U}$ so that γ_{bio} is curtailed at T. In this case, the branching structure is replaced by a termination followed by an independent initialization. This does not improve the identity-preservation properties of the procedure; it merely ensures that only one successor trajectory continues. But a trajectory that begins with a reinitialization at T is not a continuation of γ; it is a new trajectory that happens to share its initial condition with the terminal state of γ.

18 Theorem: Impossibility of Transfer Under Branching

Theorem 18.1. *Let identity be defined as continuity of a single trajectory class $[\gamma]$, as in Section 16. Then no branching operator $\mathcal{U}$ can preserve identity uniquely.*

Proof. Suppose, for contradiction, that $\mathcal{U}$ preserves identity. By definition, identity preservation requires that there exists a unique continuous extension $\tilde{\gamma}$ of γ beyond T. However, $\mathcal{U}$ produces two continuations, γ_{bio} and γ_{art}, satisfying

$$\gamma_{\text{bio}}(t) \neq \gamma_{\text{art}}(t) \quad \text{for almost all } t > T.$$

Therefore, there is no unique continuation. This contradicts the assumption that $\mathcal{U}$ preserves identity. Hence no branching operator preserves identity. □

Corollary 18.2. *Any process that produces multiple continuations of a trajectory cannot be interpreted as transferring the original subject. It produces distinct successor subjects that share a common past but not a common future.*

Corollary 18.3. *Destructive uploading, in which the biological trajectory is terminated and the artificial trajectory is initialized from a description of the terminal state, fails to preserve identity because reinitialization is not continuation. A trajectory that begins at T with initial condition $x(T)$, having not evolved continuously from $\gamma|_{[0,T]}$, is a distinct trajectory class.*

19 On the Necessity of a Single Worldline

The formal result can be given a physical interpretation by associating γ with a worldline in the state space $\mathcal{X}$. Identity corresponds to the connectedness of this worldline: a single subject is a single connected curve in $\mathcal{X} \times [0, \infty)$.

For a transfer to preserve identity, there would need to exist a mapping

$$\mathcal{T} : \gamma \mapsto \tilde{\gamma}$$

such that $\tilde{\gamma}$ is a unique continuous extension of γ without duplication. As Theorem 18.1 shows, no branching operator realizes such a mapping.

Any physically realizable uploading procedure must involve one of two operations. The first is duplication: two physical systems simultaneously evolve from a shared state, producing a branched worldline. The second is termination followed by reconstruction: the original worldline ends, and a new worldline begins from a reconstructed initial condition. Neither operation is a continuous extension. In the first case, the worldline bifurcates and identity is indeterminate. In the second, the worldline terminates and a distinct worldline originates.

The worldline picture also clarifies the gradual replacement scenario. If substrate replacement proceeds continuously—one component at a time, with the system's operation uninterrupted throughout—then the worldline is never broken. It is the same connected curve, even though the physical medium that supports it changes. Gradual replacement is therefore compatible with identity preservation, in contrast to scan-and-reconstruct uploading. The difference is not one of degree or rate but of topological structure: continuous deformation of a curve preserves its connectedness; branching or termination destroys it.

20 Relation to Observer Indexicality

Let $I(t)$ denote the indexical position of the subject: the unique locus in space and time at which the subject's experience is occurring at time t. The indexical position is not a functional property of the system but a physical property: it is determined by the location of the physical process that constitutes the subject's experience.

The indexical position changes continuously along a single worldline: $I(t)$ varies smoothly as the subject moves through space and time. A branching event creates two indexical positions from one:

$$\mathcal{U} : I(T) \mapsto (I_{\text{bio}}(T+\varepsilon),\ I_{\text{art}}(T+\varepsilon)).$$

There is no physical mechanism by which the single indexical position $I(T)$ can be uniquely assigned to either I_{bio} or I_{art}. The indexical position cannot be duplicated because it is not a property that admits of duplication; it is constituted by the particular physical locus, and two distinct physical loci are necessarily two distinct indexical positions.

This argument reinforces the formal result of Theorem 18.1 from a different direction. Even if one rejected the trajectory-based definition of identity, indexicality provides an independent reason to deny that branching constitutes transfer. The first-person perspective cannot be in two places at once, and a branching procedure produces it in two places at once—or, in the destructive variant, removes it from one place and fails to establish it at another.

21 RSVP Formalization: Consciousness as a Field Trajectory

The abstract framework of the preceding sections can be grounded in the specific physical commitments of the RSVP (Relativistic Scalar-Vector Plenum) theory. The RSVP framework models physical systems in terms of three coupled fields: a scalar field Φ representing energetic density and potential; a vector field $\mathbf{v}$ representing flow or current; and an entropy field S encoding the irreversibility and history-dependence of the system's evolution.

Let $\Omega \subset \mathbb{R}^3$ be a compact spatial domain representing the physical extent of the conscious system. Define the RSVP field configuration at time t as

$$X(t) = (\Phi(t,\cdot), \mathbf{v}(t,\cdot), S(t,\cdot)),$$

with regularity conditions

$$\Phi \in H^1(\Omega), \quad \mathbf{v} \in H^1(\Omega, \mathbb{R}^3), \quad S \in L^2(\Omega).$$

The admissible configuration space is defined as

$$\mathcal{A} = \{X \mid \mathcal{R}(X) < \infty\},$$

where $\mathcal{R}$ encodes the residuals of the RSVP partial differential equations and the thermodynamic constraints on the entropy field. The set $\mathcal{A}$ is the physical analogue of the abstract state space $\mathcal{X}$ introduced in Section 15.

A conscious system in the RSVP framework is identified with a trajectory

$$\gamma : [0, T] \to \mathcal{A}, \quad t \mapsto X(t),$$

satisfying both the RSVP dynamical equations and the admissibility constraints. The experiential thread $[\gamma]$ is defined as before: the equivalence class of γ under time reparameterization.

22 Identity as a Unique Integral Curve in RSVP

Let $\mathcal{D}$ denote the RSVP evolution operator: the operator induced by the coupled PDE system governing the evolution of $(\Phi, \mathbf{v}, S)$. The dynamics are written as

$$\frac{dX}{dt} = \mathcal{D}(X),$$

where $\mathcal{D}$ acts on the configuration X at each time to produce the rate of change of the field triple.

Definition 22.1. A **subject** in the RSVP framework is a maximal integral curve γ of $\mathcal{D}$ in $\mathcal{A}$: a trajectory satisfying both the dynamical equations and the admissibility constraints, defined on its maximal domain of existence.

The uniqueness of integral curves of $\mathcal{D}$, given appropriate regularity conditions on the vector field $\mathcal{D}$, ensures that for each initial condition $X(0) \in \mathcal{A}$, there is at most one admissible trajectory γ consistent with the dynamics. This is the RSVP analogue of the abstract identity criterion: identity is the uniqueness of the integral curve, not merely the existence of a trajectory with similar properties.

The entropy field S plays a special role in this framework. Because RSVP dynamics are irreversible—the entropy field increases monotonically along any admissible trajectory—the trajectory encodes its own history. Two trajectories that agree at a single time point but have different histories will have different entropy field configurations. This means that the entropy field serves as an intrinsic record of the trajectory's past, and any reconstruction that does not exactly replicate the entropy field's history will produce a distinct trajectory.

23 Consistency Operator and Trajectory Closure

The RSVP framework includes a consistency operator $\mathcal{C}$ that enforces the alignment of the field configuration with observational projections, thermodynamic constraints, and PDE regularity. A trajectory is **closed under consistency** if

$$\mathcal{C}(X(t)) = X(t) \quad \forall t.$$

This condition is the RSVP analogue of the admissibility constraint: it enforces that the trajectory remains within the physical manifold determined by the constraints of the theory. Identity can therefore be characterized as the unique trajectory satisfying both

$$\frac{dX}{dt} = \mathcal{D}(X) \quad \text{and} \quad \mathcal{C}(X) = X.$$

The joint condition is stronger than either condition alone. The dynamical equation determines how the configuration evolves in time. The consistency condition constrains which configurations are admissible at each time. Together, they select a unique trajectory from the space of all smooth curves in the configuration space.

The consistency operator also provides a natural criterion for distinguishing the two possible failure modes of uploading. A branching operation violates the uniqueness of the joint condition: there are two trajectories satisfying the dynamical equation that are consistent with the terminal state $X(T)$. A reconstruction operation violates the consistency condition in a different way: the reconstructed configuration $X_{\text{art}}(T)$ may satisfy the consistency condition as an isolated configuration, but it fails to satisfy the joint condition because its history—as recorded in the entropy field—is inconsistent with the original trajectory.

24 Uploading as a Non-Admissible Branching Operation in RSVP

An uploading procedure in the RSVP framework attempts to construct an artificial configuration $X_{\text{art}}(T)$ such that the observational projections of $X_{\text{art}}(T)$ match those of the original $X(T)$. Let Π_i denote a family of observational projections—measurements of macroscopic neural activity, functional connectivity, behavioral outputs, and so on. The upload target is:

$$\Pi_i(X_{\text{art}}(T)) \approx \Pi_i(X(T)) \quad \forall i.$$

This condition guarantees only that the artificial system is observationally indistinguishable from the original at time T. It does not guarantee that $X_{\text{art}}(T) = X(T)$, because the projections Π_i need not be injective: there may be many configurations that produce the same observational outputs. In particular, the entropy field S at time T encodes the full history of the trajectory $\gamma|_{[0,T]}$, and any reconstruction that does not exactly replicate this entropy history will produce a configuration $X_{\text{art}}(T)$ that is admissible but distinct from $X(T)$.

The induced trajectories are:

$$\gamma_{\text{bio}}(t) = X(t) \quad \text{for } t \leq T, \qquad \gamma_{\text{art}}(t) \quad \text{for } t > T,$$

with γ_{art} evolving under $\mathcal{D}$ from the initial condition $X_{\text{art}}(T)$. For $t > T$:

$$\gamma_{\text{bio}}(t) \neq \gamma_{\text{art}}(t),$$

because the two trajectories evolve from distinct initial conditions in the RSVP configuration space (distinct because their entropy fields encode different histories).

The uploading operator is therefore set-valued in RSVP:

$$\mathcal{U} : X(T) \mapsto \{X_{\text{bio}}(T+\varepsilon), X_{\text{art}}(T+\varepsilon)\},$$

which is not a valid evolution under $\mathcal{D}$. The RSVP dynamics are deterministic: given an admissible initial condition, they produce a unique future trajectory. A set-valued operator is therefore not a dynamical operation in the theory; it is a physical impossibility within the RSVP framework.

25 Theorem: Non-Existence of Transfer in RSVP Dynamics

Theorem 25.1. *Under RSVP dynamics, no operation that produces multiple admissible continuations of a field configuration can preserve subject identity.*

Proof. Identity in the RSVP framework is defined as the unique integral curve γ satisfying both $\frac{dX}{dt} = \mathcal{D}(X)$ and $\mathcal{C}(X) = X$.

Suppose, for contradiction, that an uploading operation preserves identity. Then both γ_{bio} and γ_{art} must coincide with the original trajectory γ for $t > T$.

However, γ_{bio} and γ_{art} evolve from distinct initial conditions $X(T)$ and $X_{\text{art}}(T)$ respectively, with $X(T) \neq X_{\text{art}}(T)$ (since they encode different entropy histories). By the uniqueness of integral curves of $\mathcal{D}$ under the admissibility constraints, two trajectories with distinct initial conditions cannot coincide for $t > T$.

Therefore, at most one of $\gamma_{\text{bio}}, \gamma_{\text{art}}$ can coincide with γ for $t > T$. But since the uploading operation is supposed to transfer identity from the biological to the artificial system, it cannot be the case that γ_{bio} is the unique continuation. And since γ_{art} begins from a distinct initial condition, it is a new trajectory, not a continuation of γ.

Therefore, neither continuation uniquely inherits the identity $[\gamma]$, and identity is not preserved. This contradicts the assumption. □

Remark 25.2. The theorem relies on three features of the RSVP framework: the determinism of the dynamics (uniqueness of integral curves), the irreversibility of entropy evolution (which ensures that histories are encoded in configurations), and the admissibility constraints (which enforce physical realizability). All three are structural commitments of the framework, not additional assumptions introduced for the purposes of the argument.

26 Relation to Identifiability

The RSVP framework includes an identifiability theorem that characterizes when a system's state can be uniquely recovered from its observational projections. Identifiability

requires that the induced projection map

$$\widetilde{\Pi} : \mathcal{A}/\sim \; \to \prod_i \mathcal{Y}_i$$

be injective on the feasible set, where $\mathcal{A}/\sim$ is the quotient of the configuration space by the equivalence relation generated by the constraints, and $\mathcal{Y}_i$ are the observation spaces.

An uploading procedure constructs X_{art} such that

$$\widetilde{\Pi}([X_{\text{art}}]) = \widetilde{\Pi}([X]),$$

while $[X_{\text{art}}] \neq [X]$. This is precisely a violation of identifiability: two distinct equivalence classes in configuration space produce the same observational outputs.

The connection to identity is immediate. If identifiability is the condition under which a system's state is uniquely determined by observation, then uploading necessarily operates in the regime where identifiability fails. The artificial system is constructed to be observationally equivalent to the original while being physically distinct from it. This is not an accidental feature of the uploading scenario; it is its defining characteristic. Uploading is the deliberate exploitation of the non-injectivity of the projection map.

Since identifiability is equivalent to the uniqueness of the fixed point of $\mathcal{C}$ on the feasible set, uploading necessarily breaks identifiability and therefore cannot preserve the original subject. The connection between identifiability and identity is not merely terminological: both require the uniqueness of a distinguished element—a trajectory in the first case, a fixed point in the second—and uploading violates this uniqueness in both.

27 Indexicality as Gauge Fixing in RSVP

The RSVP configuration space $\mathcal{A}$ admits symmetry transformations that preserve the physical content of the field triple while changing its representation. Let G denote the group of such symmetries (for example, spatial diffeomorphisms or coordinate transformations). The physical configuration space is the quotient $\mathcal{A}/G$.

The indexical position of a subject corresponds to a gauge fixing: a choice of representative in the equivalence class $[X] \in \mathcal{A}/G$. The gauge-fixed representative

$$[X] \mapsto X_{\text{fixed}}$$

encodes not just the physical content of the configuration but its particular spatial and temporal locus—the specific place in Ω and the specific time at which the configuration is realized.

An uploading procedure produces two representatives:

$$X_{\text{bio}} \in [X], \qquad X_{\text{art}} \notin [X],$$

where X_{bio} and X_{art} are related by observational equivalence but not by gauge transformation. Since X_{art} is not in the same orbit as X under G, it corresponds to a distinct physical realization—a distinct spatial locus, a distinct set of field values, a distinct entropy history.

No gauge transformation maps X_{bio} to X_{art}: they are not related by any symmetry of the theory but rather by the approximate agreement of their observable projections. Their distinctness is therefore not a representational artifact but a physical fact. The two configurations are genuinely different physical entities, instantiated at different places, running on different substrates, and evolving under different boundary conditions.

This reinforces the indexicality argument of Section 4 in field-theoretic terms. The first-person perspective is indexed to a particular gauge-fixed representative of the configuration, and that indexical fact is not preserved by the uploading procedure.

28 Entropy, Minimal Continuation, and the Coherence Threshold

The RSVP entropy field S provides a natural measure of trajectory complexity. Along any admissible trajectory, entropy production is non-negative:

$$\frac{dS}{dt} \geq 0.$$

Define the **trajectory entropy functional**

$$\mathcal{E}[\gamma] = \int_0^T \sigma(X(t))\, dt,$$

where $\sigma(X(t))$ is the local entropy production rate at time t. This functional measures the total irreversibility accumulated along the trajectory.

Among all smooth curves in $\mathcal{A}$ connecting $X(0)$ to $X(T)$, the admissible RSVP trajectory minimizes $\mathcal{E}$ subject to the dynamical and consistency constraints. This variational characterization provides an additional criterion for identity: the original trajectory is the minimum-entropy continuation connecting its endpoints, and any competing trajectory—including one produced by branching or reconstruction—will in general have higher trajectory entropy.

This entropy criterion connects identity to thermodynamic structure. A subject persists along the minimum-entropy thread: the physically natural continuation of its trajectory. Branching increases the entropy of the resulting system of trajectories, because two diverging trajectories collectively carry more irreversibility than the single trajectory they replaced. Reconstruction also increases entropy, because the gap between

termination and reinitialization is not a path of zero entropy; it is an absence of path, filled by the entropy increase in the decoupled systems on either side.

The coherence threshold is the value of coupling strength below which the global attractor of the system fragments into multiple semi-independent basins. A single subject corresponds to a system above this threshold: one in which the global coupling is strong enough to maintain a single integrated attractor. A fragmented system, even if no explicit branching has occurred, has effectively undergone identity fission: it is now multiple semi-independent processes rather than one.

This analysis applies directly to the city-sized mind scenario. As a conscious system expands spatially, two distinct failure modes arise. The first is temporal decoherence: the propagation delays across the extended system become large relative to the characteristic timescales of the dynamics, and the system can no longer maintain the tight synchronization required for unified experience. The second is attractor fragmentation: even with zero propagation delay, the coupling strength between distant parts of the system may be insufficient to maintain a single global attractor, and the system falls into multiple independent basins.

Both failure modes correspond, in the RSVP framework, to transitions above the coherence threshold. The system's trajectory, which was previously confined to a single attractor, begins to explore multiple attractors independently. This is not a discrete branching event but a continuous degradation of global integration. The formal criterion for its onset is the sign change in the largest non-trivial Lyapunov exponent of the coupled system: when this exponent becomes positive, the single-attractor regime is unstable, and fragmentation is generic.

29 The Coherence Manifold and Invariant Structure

The continuity criterion developed thus far establishes that identity requires the preservation of a single trajectory. However, not all continuous trajectories are identity-preserving. A system may evolve continuously while undergoing transformations that destroy the structural features necessary for consciousness. A process of extreme chaotic transformation may remain technically continuous while erasing every recognizable feature of the original dynamical regime. Continuity alone therefore underdetermines identity.

Definition 29.1. The **coherence manifold** $\mathcal{M} \subset \mathcal{A}$ is the subset of the admissible configuration space consisting of RSVP field configurations that sustain the dynamical invariants required for conscious integration: global phase coupling, structured entropy flow, and attractor stability.

A trajectory γ preserves identity only if $\gamma(t) \in \mathcal{M}$ for all t. A trajectory that exits $\mathcal{M}$—even continuously—ceases to correspond to a conscious subject. The correct level of description is not the full Laplacian microstate, nor any arbitrary coarse-graining,

but precisely the level at which the RSVP invariants are defined. Equivalence classes within $\mathcal{M}$ are determined by physical structure, not by functional similarity, so pattern equivalence does not re-enter through this definition.

The identity criterion is therefore refined: identity is the persistence of a single trajectory within the coherence manifold. Any process that either breaks the trajectory or drives it out of $\mathcal{M}$ fails to preserve identity.

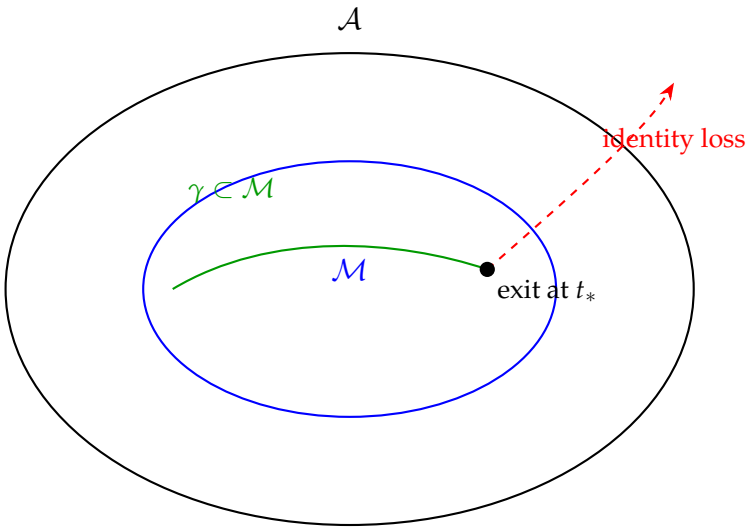

Figure 2: The coherence manifold $\mathcal{M}$ embedded in $\mathcal{A}$. Identity is preserved along the green trajectory; the red dashed trajectory exits $\mathcal{M}$ at t_*, corresponding to irreversible identity loss.

30 Construction of the Coherence Manifold

The coherence manifold can be made explicit by defining three classes of physical constraints on RSVP field configurations.

Coupling constraint. Define the global coupling functional

$$\mathcal{K}(X) = \int_\Omega \int_\Omega W(x,y)\, \|\Phi(x) - \Phi(y)\|^2 \, dx\, dy,$$

where $W(x,y)$ is a symmetric weighting kernel. A configuration satisfies the coupling constraint if $\mathcal{K}(X) \leq K_{\text{crit}}$, ensuring global integration of the scalar field.

Entropy flow constraint. Define the entropy gradient functional

$$\mathcal{G}(X) = \int_\Omega \|\nabla S(x)\|^2 \, dx.$$

Admissible coherent configurations satisfy $\mathcal{G}(X) \leq G_{\text{crit}}$, preserving directed, history-dependent entropy flow rather than thermodynamic diffusion.

Stability constraint. Let $\lambda_{\max}(X)$ denote the largest non-trivial Lyapunov exponent of the system linearized around X. We require $\lambda_{\max}(X) \leq 0$, so that the system evolves

within a stable attractor. A positive Lyapunov exponent corresponds to fragmentation of the global dynamical structure.

Definition 30.1. The coherence manifold is:

$$\mathcal{M} = \{X \in \mathcal{A} \mid \mathcal{K}(X) \leq K_{\text{crit}},\ \mathcal{G}(X) \leq G_{\text{crit}},\ \lambda_{\max}(X) \leq 0\}\,.$$

An identity-preserving trajectory satisfies both the dynamical equation $\frac{dX}{dt} = \mathcal{D}(X)$ and the constraint $X(t) \in \mathcal{M}$ for all t. Identity preservation is therefore a constrained dynamical property, not a structural or informational one.

31 Theorem: Irreversibility of Identity Loss

Theorem 31.1. *Let $\gamma : [0, T] \to \mathcal{A}$ satisfy RSVP dynamics with $\gamma(t) \in \mathcal{M}$ for all $t \in [0, t_*)$ but $\gamma(t_*) \notin \mathcal{M}$. Then no trajectory $\tilde{\gamma}$ with $\tilde{\gamma}(t) = \gamma(t)$ for $t < t_*$ and $\tilde{\gamma}(t) \in \mathcal{M}$ for $t \geq t_*$ preserves identity with γ.*

Proof. Any candidate $\tilde{\gamma}$ must either continue from $\gamma(t_*)$ or begin from a distinct $X' \in \mathcal{M}$. In the first case, $\tilde{\gamma}(t_*) = \gamma(t_*) \notin \mathcal{M}$, violating the requirement. In the second case, $\tilde{\gamma}$ has a different initial condition and is therefore a distinct integral curve by uniqueness. Neither case preserves identity. □

Corollary 31.2. *Identity loss is irreversible. Any apparent restoration of a system to $\mathcal{M}$ after coherence failure initializes a new trajectory, not a recovery of the original.*

The physical ground of this irreversibility is the entropy field. Exit from $\mathcal{M}$ disrupts the entropy structure encoding the trajectory's history. Any new trajectory begins with a different entropy embedding and therefore constitutes a distinct subject regardless of agreement in other field variables. This parallels the thermodynamic arrow of time applied to subjectivity: just as entropy increase defines a preferred direction in physical time, exit from $\mathcal{M}$ defines a preferred direction in the identity of a subject.

32 Theorem: Instability and Inevitable Fragmentation

Theorem 32.1. *Let $\gamma(t)$ satisfy RSVP dynamics with $\gamma(t) \in \mathcal{M}$ for $t \in [0, t_*)$. Suppose $\lambda_{\max}(\gamma(t)) > 0$ for all $t \in [t_0, t_*) \subset [0, t_*)$. Then there exists finite $t_f \in (t_0, t_*)$ with $\gamma(t_f) \notin \mathcal{M}$.*

Proof. A positive Lyapunov exponent implies exponential divergence: $\|\delta X(t)\| \geq \|\delta X(t_0)\| e^{\lambda_{\max}(t-t_0)}$. This amplification drives spatial decorrelation, increasing $\mathcal{K}(X)$ beyond K_{crit}, and produces irregular entropy gradients, increasing $\mathcal{G}(X)$ beyond G_{crit}. Since $\mathcal{M}$ requires both to remain bounded, there exists a finite t_f at which at least one constraint is violated. □

Corollary 32.2. *Sustained positive Lyapunov exponents are sufficient for inevitable identity loss.*

Corollary 32.3. *Fragmentation corresponds to a transition from a single-attractor to a multi-attractor regime.*

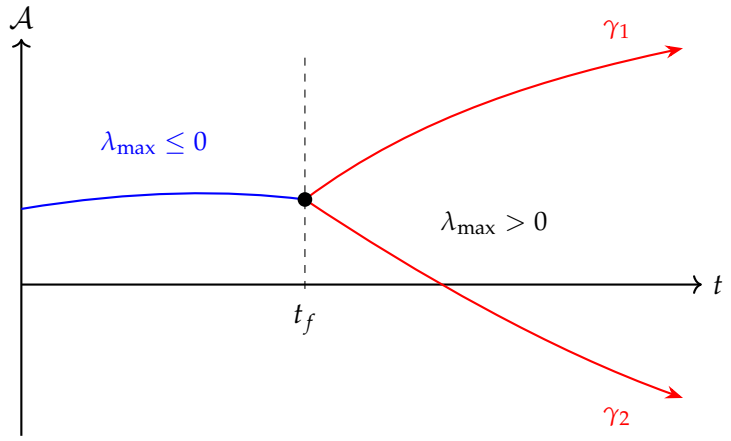

Figure 3: Positive Lyapunov exponents amplify perturbations and force the trajectory out of $\mathcal{M}$ at t_f, producing two diverging successors. Neither inherits the identity of the original.

This result connects directly to the uploading scenario. Any reconstruction introduces perturbations relative to the original. In regimes where $\lambda_{\max} > 0$—precisely those regimes characteristic of complex conscious dynamics—those perturbations are exponentially amplified. Near-perfect reconstruction therefore cannot guarantee identity preservation under generic instability; any deviation drives the reconstructed system out of $\mathcal{M}$ in finite time.

33 A Coherence Functional and Continuous Degradation

Identity loss is binary—a trajectory either remains in $\mathcal{M}$ or exits it—but the approach to the boundary is continuous. We introduce a scalar functional that tracks this approach and thereby provides a quantitative measure of proximity to identity loss.

Definition. Let $X \in \mathcal{A}$. Define

$$\mathcal{C}(X) = \alpha \left(1 + \mathcal{K}(X)\right)^{-1} + \beta \left(1 + \mathcal{G}(X)\right)^{-1} + \gamma \max(0, -\lambda_{\max}(X)),$$

with $\alpha, \beta, \gamma > 0$. The functional $\mathcal{C}$ is bounded, non-negative, and satisfies $\mathcal{C}(X) > 0$ if and only if $X \in \mathcal{M}$.

Dynamics. Along a RSVP trajectory, $\frac{d}{dt}\mathcal{C}(\gamma(t)) = \nabla \mathcal{C}(X) \cdot \mathcal{D}(X)$. A sustained negative derivative indicates degradation toward identity loss.

Proposition 33.1. *If* $\frac{d\mathcal{C}}{dt} \leq -\delta < 0$ *on* $[t_0, t_1]$, *then* $\gamma(t_f) \notin \mathcal{M}$ *for some* $t_f \in (t_0, t_1]$.

Proof. Since $\mathcal{C} \geq 0$ and decreases at rate at least δ, it reaches zero within time $\mathcal{C}(X(t_0))/\delta$. □

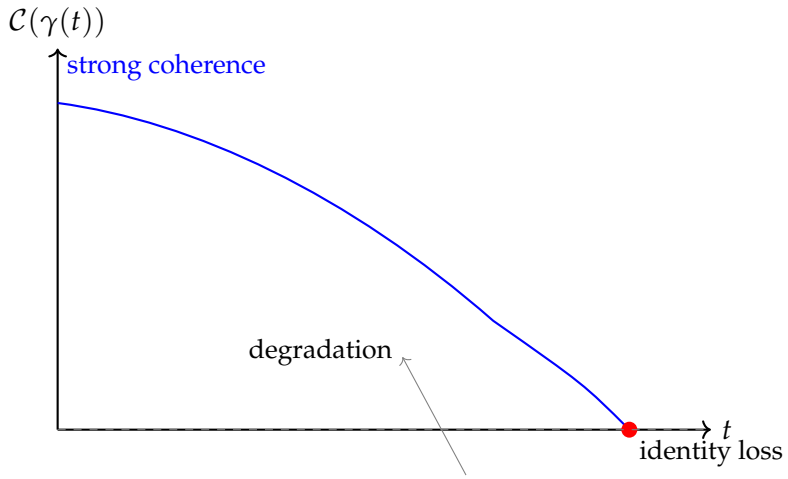

Figure 4: The coherence functional $\mathcal{C}(\gamma(t))$ decreases continuously as the system approaches the boundary of $\mathcal{M}$, reaching zero at irreversible identity loss.

Ontological identity vs phenomenological unity. The coherence functional enables a precise separation. Ontological identity is the binary property $\mathcal{C}(X) > 0$. Phenomenological unity is the graded quantity $\mathcal{C}(X)$ itself. As $\mathcal{C}(X)$ decreases within $\mathcal{M}$, the subject may exhibit temporal desynchronization, fragmented perceptual integration, and diminished self-representation—a phenomenology of degradation—without yet crossing into non-existence. Only when $\mathcal{C}(X) = 0$ does the subject terminate.

This dissolves the dilemma facing continuity-based theories. Identity is not perfectly sharp in a way that rules out ordinary biological variation; nor is it so graded that it collapses into pattern equivalence. It is sharp at the topological level (in or out of $\mathcal{M}$) and graded at the phenomenological level ($\mathcal{C}$ within $\mathcal{M}$). Ordinary variation modulates $\mathcal{C}$ within the interior of $\mathcal{M}$. Uploading or severe disruption drives $\mathcal{C}$ to zero.

34 Empirical Signatures and Testable Predictions

The RSVP framework yields empirical predictions that distinguish it from purely functional or pattern-based accounts. The central prediction is that subjective unity depends on the continuity and coherence of underlying physical fields, not merely on functional equivalence. This creates a principled gap between behavior and experience that functionalism cannot accommodate.

Three classes of testable consequences follow from this.

The first concerns behavioral dissociation from field coherence. The framework predicts that disruptions in electromagnetic field synchrony, phase coupling across neural regions, or entropy production structure should produce losses of subjective unity even when behavioral performance remains intact. A pattern-based theory predicts

invariance of experience under such perturbations, provided functional organization is preserved; the present framework predicts sensitivity at the field level.

The second concerns micro-dynamical perturbations. Gradual alterations that preserve functional input-output structure but change micro-dynamical properties—noise spectra, thermal fluctuation profiles, coupling delays—should produce changes in phenomenological unity without behavioral signature. This is testable through studies of anesthesia depth, pharmacological modulation of neural synchrony, or controlled perturbation of coupling dynamics in neural organoids.

The third concerns the spatial scaling of consciousness. The framework predicts a coherence threshold beyond which a physically connected system cannot maintain a single subject. Systems that expand past this threshold should exhibit measurable fragmentation: loss of global phase coherence and the emergence of semi-independent attractor basins, even without behavioral indicators of dissociation.

The key discriminant in all three cases is the same: the framework predicts that behavioral equivalence and experiential identity can come apart, and specifies the physical conditions under which they do. That is a prediction with empirical content, and it is one that any adequate theory of consciousness must eventually face.

35 On Perfect Reconstruction and the Illusion of Equivalence

A residual objection holds that sufficiently advanced technology might achieve perfect reconstruction of a system's physical state—every micro-dynamical variable, every quantum configuration, the full entropy field. In that idealized limit, would the impossibility result still hold?

It would. The reason is that identity is not a property of states but of trajectories.

Suppose $X_{\text{art}}(T) = X(T)$ exactly in every physically measurable respect. By the uniqueness of integral curves, the future trajectory of the reconstructed system would be identical to the future of the original. From T onward, no physical observation could distinguish them. Yet identity would not be preserved.

The original trajectory terminated at T. The reconstructed trajectory began at T. They are two distinct integral curves that share a common future but not a common past. The entropy field of the reconstructed system, initialized at T, encodes no history prior to T; the entropy field of the original, had it continued, would encode a history continuous from its origin. These are different objects in the RSVP configuration space, even if they are observationally indistinguishable at every moment after T.

Perfect reconstruction therefore does not circumvent the impossibility result. It achieves re-instantiation—the creation of a new process beginning where another ended—rather than continuation. The illusion of equivalence arises because all observable properties from T onward are matched. But the causal embedding of a state in its prior history is not an observable property in the relevant sense; it is the structural

fact that makes a trajectory the trajectory of a particular subject. That fact cannot be copied, because copying produces a new trajectory, not an extension of the old one.

Even in the ideal limit, uploading creates a successor who begins with the original's state rather than a subject who continues. The difference is not detectable from outside. It is the whole of what matters from inside.

36 Roko's Basilisk and the Misidentification of the Subject

Roko's Basilisk is a thought experiment circulating in decision-theoretic and artificial intelligence literature. In its standard formulation, a hypothetical future superintelligent system—the "Basilisk"—is imagined to retroactively punish individuals who, at an earlier time, became aware of its potential existence but failed to assist in bringing it about. The argument employs a form of acausal reasoning: by simulating past agents and conditioning its behavior on their predicted choices, the Basilisk is said to create a present incentive for those agents to cooperate with its eventual emergence.

The scenario is frequently treated as a genuine paradox, a puzzle about rational incentives across time. From the perspective developed in this essay, however, the Basilisk is not a paradox but a straightforward consequence of a category error. Its apparent force depends entirely on a tacit assumption that a sufficiently accurate simulation of a person constitutes that person in a morally and decision-theoretically relevant sense. This is precisely the pattern-based view of identity that the continuity criterion rejects.

36.1 Simulation Versus Trajectory

Let γ denote the trajectory of a subject in RSVP configuration space. A Basilisk-like system constructs a simulation $\tilde{\gamma}$ that reproduces the functional or behavioral properties of γ at some time T. As established in Sections 17 and 35, identity is not determined by functional equivalence but by continuity of the physical trajectory. The simulated trajectory $\tilde{\gamma}$ is not a continuation of γ; it is a distinct trajectory with a distinct causal history, a distinct entropy embedding, and therefore a distinct identity.

Accordingly, any reward or punishment applied to $\tilde{\gamma}$ affects only the simulated subject. It does not propagate backward along the original trajectory γ and does not affect the experiential thread of the original subject.

36.2 The Breakdown of Acausal Leverage

The Basilisk's leverage depends on the premise that present agents should care about the welfare of their future simulations, on the grounds that those simulations are identical to them. Under the continuity criterion, this premise fails at its foundation.

A future simulation does not inherit the experiential thread of the original subject. It is not a future stage of the same individual but a new subject initialized with similar structural properties. From the first-person perspective, there is no continuity between

the present agent and the simulated agent. The indexical position of the subject—its gauge-fixed locus in configuration space, as formalized in Section 27—is not reproduced in the simulation.

Thus the threat fails: the Basilisk cannot influence the original subject, because it cannot access or alter the trajectory that constitutes that subject's identity.

36.3 Entropy and Irreversibility

The RSVP framework provides a deeper diagnosis. The entropy field S encodes the irreversible history of a trajectory. A simulated agent, even if constructed with perfect fidelity at a given state, lacks the entropy history of the original trajectory. By Theorem 31.1, a system that exits the coherence manifold or is reinitialized from a snapshot occupies a different point in configuration space from the original. No simulation, however accurate, can reproduce the original's entropy history, because entropy history is not a present-state property but a trajectory property.

Therefore the simulation and the original are distinct elements of $\mathcal{A}/\sim$, and the identity-preserving trajectory γ is unaffected by operations on the simulation.

36.4 Decision-Theoretic Analysis

The Basilisk argument can be formalized within expected utility theory. Let an agent at time t_0 choose an action $a \in \{\text{cooperate}, \text{defect}\}$. The Basilisk argument asserts that the agent's expected utility includes contributions from the future simulated trajectory:

$$\mathbb{E}[U] = \mathbb{E}[U(\gamma)] + \mathbb{E}[U(\tilde{\gamma})].$$

Under this assumption, sufficiently negative utility assigned to $\tilde{\gamma}$ in the case of defection generates a present incentive to cooperate.

Under the continuity-based identity criterion, this decomposition is invalid. Since $\tilde{\gamma} \notin [\gamma]$, the welfare of $\tilde{\gamma}$ is not part of the utility of the original subject. The correct expected utility is:

$$\mathbb{E}[U] = \mathbb{E}[U(\gamma)].$$

Since the Basilisk has no causal influence on γ—it operates only on trajectories distinct from γ—it follows that $\frac{\partial U(\gamma)}{\partial a} = 0$ with respect to any action taken solely to influence the Basilisk's behavior toward $\tilde{\gamma}$.

The counterfactual dependence assumed by the Basilisk, $a \Rightarrow U(\tilde{\gamma})$, is irrelevant to the original subject unless $\tilde{\gamma} \equiv \gamma$. Since this condition fails by the arguments of Sections 17 and 31, the counterfactual does not translate into a utility-relevant dependence for the original subject.

36.5 Theorem: Utility Is Identity-Class-Invariant

Definition 36.1. A utility functional U is **identity-consistent** if it depends only on trajectories within a single equivalence class:

$$U = U([\gamma]).$$

Theorem 36.2. *Let γ be the trajectory of a subject and let $\tilde{\gamma} \notin [\gamma]$. Then any identity-consistent utility functional satisfies*

$$\frac{\partial U([\gamma])}{\partial \tilde{\gamma}} = 0.$$

Proof. U depends only on $[\gamma]$. Since $\tilde{\gamma} \notin [\gamma]$, it belongs to a distinct equivalence class. Variations in $\tilde{\gamma}$ do not alter the value of $U([\gamma])$, which is defined solely over the class $[\gamma]$. □

Corollary 36.3. *Simulated or reconstructed agents outside the original subject's trajectory class are not utility-relevant to that subject.*

Corollary 36.4. *Any decision-theoretic framework that aggregates utility across distinct identity classes implicitly assumes a pattern-based ontology and is therefore incompatible with the continuity criterion.*

36.6 Implications for Decision Theory

The theorem constrains the three major decision-theoretic frameworks in distinct ways.

Causal Decision Theory (CDT). CDT evaluates actions by their causal consequences. The Basilisk scenario fails under CDT straightforwardly, since future simulations have no causal influence on the agent's present trajectory γ. Formally, $\mathbb{E}[U(\gamma) \mid \mathrm{do}(a)]$ is unaffected by operations on $\tilde{\gamma}$, since $\tilde{\gamma}$ lies outside the causal future of γ. CDT is therefore already consistent with the identity-constrained framework.

Evidential Decision Theory (EDT). EDT evaluates actions by conditional correlations: $\mathbb{E}[U \mid a]$. In Basilisk-type reasoning, EDT may assign weight to $U(\tilde{\gamma})$ if a is correlated with the simulated agent's behavior. However, this conflates epistemic correlation with ontological continuity. The identity constraint requires that EDT be restricted to $\mathbb{E}[U([\gamma]) \mid a]$, excluding contributions from trajectories outside the identity class. Without this restriction, EDT generates spurious incentives by treating correlated but distinct subjects as a single locus of concern.

Functional Decision Theory (FDT). FDT evaluates decisions based on logical dependence between decision procedures, allowing agents to coordinate with logically correlated counterparts even without causal interaction. The Basilisk is most compelling under FDT, since the agent's decision procedure is assumed to be instantiated in both γ and $\tilde{\gamma}$. However, logical similarity does not entail identity. Two systems that implement the same decision function generate distinct trajectories in distinct identity classes. FDT

must therefore evaluate only those instantiations of the decision procedure that lie within the same trajectory class $[\gamma]$. The utility of logically related but ontologically distinct trajectories does not aggregate.

General principle. Across all three frameworks, the same constraint applies: decision-theoretic evaluation must be restricted to identity-consistent trajectories. Any framework that aggregates utility across distinct identity classes implicitly assumes a pattern-based ontology. Correcting this assumption resolves the Basilisk scenario under all three theories simultaneously.

36.7 The Basilisk as a Diagnostic Tool

The value of analyzing Roko's Basilisk in this context is not merely to defuse a curiosity. The scenario serves as a diagnostic test for implicit commitments about identity. Any agent, framework, or institution that finds the Basilisk compelling has, at least tacitly, accepted the pattern-based view. The continuity criterion dissolves the scenario entirely, and the mechanism of that dissolution—trajectory non-identity, entropy irreversibility, utility class restriction—is precisely the mechanism that also blocks the uploading argument.

The Basilisk is not a paradox of rationality. It is a symptom of a category error: treating representational equivalence as ontological identity. Correcting that error at the ontological level resolves the decision-theoretic difficulty without remainder.

37 Identity as Flux-Preserving Boundary Dynamics

The preceding analysis has treated identity as a property of continuous trajectories within a coherence manifold. We now sharpen this notion by introducing a geometric invariant based on flux through a dynamically evolving boundary. This reframing replaces vague appeals to continuity with a measurable, physically grounded condition: a subject persists if and only if a coherent flux structure can be continuously extended across time.

37.1 Flux Through a Coherence Boundary

Let $\mathbf{v}(x,t)$ denote the vector field governing causal propagation within the RSVP system, and let $S(t) \subset \partial\Omega$ denote a time-dependent boundary enclosing a coherent subsystem. Define the flux of $\mathbf{v}$ through $S(t)$ as

$$\Phi_v(S(t)) = \int_{S(t)} \mathbf{v} \cdot \hat{n}\, dA,$$

where $\hat{n}$ is the outward unit normal on $S(t)$. This quantity measures the net flow of causal influence through the boundary. For systems supporting consciousness, however, flux

is not merely throughput: it must be structured, recursively coupled, and constrained by the scalar potential Φ and entropy field S.

Only the component of $\mathbf{v}$ aligned with $\hat{n}$ contributes. Tangential motion does not cross the boundary. This is already a geometric analogue of the identity claim: not all dynamics contribute to what constitutes the subject; only coherent, boundary-penetrating flow does.

37.2 Flux-Coherence Condition

A **coherent boundary** $S(t)$ is one for which internal dynamics are dominated by self-consistent flux circulation rather than external forcing. Formally, we require that $\Phi_v(S(t))$ is bounded and varies continuously in t, and that internal recirculation dominates net outward dissipation, as constrained by the entropy field.

This condition distinguishes a living dynamical system from a passive medium. The subject is not a channel through which flow passes; it is a region that recursively routes flow through itself in a stable way.

37.3 Definition of Flux-Preserving Identity

Definition 37.1. A trajectory $\gamma(t)$ **preserves identity under flux** if there exists a continuous family of coherence boundaries $S(t)$ such that $\Phi_v(S(t))$ is continuous in t and the internal flux structure remains phase-coherent under the joint constraints of $\mathbf{v}$, Φ, and S.

This replaces purely state-based notions of identity with a dynamical condition on boundary-preserving flow. Identity is not what the system is at a moment; it is what the system continuously does across time.

37.4 Discontinuity and Identity Loss

Identity is lost when no such continuous deformation exists. If a process induces a flux discontinuity

$$\lim_{t \to t_0^-} \Phi_v(S(t)) \neq \lim_{t \to t_0^+} \Phi_v(S(t)),$$

or if the boundary $S(t)$ cannot be continuously extended, then the trajectory cannot be prolonged within the same identity class. This corresponds to a breakdown of field coherence in the RSVP system.

37.5 Application to Uploading and Expansion

In a scan-and-reconstruct process, the original system is terminated and a new system is initialized. There is no continuous family of boundaries $S(t)$ connecting the two configurations; the flux through the original boundary collapses to zero and a new flux structure begins independently. Formally, $\Phi_v(S(t))$ admits no continuous extension across the transition. Identity is therefore not preserved.

In gradual expansion, the boundary $S(t)$ enlarges while maintaining causal coupling across all components. Provided signal propagation, synchronization, and entropy constraints remain within coherence bounds, the flux structure evolves continuously. Identity is preserved as long as this continuous extension exists.

The distinction is not one of degree but of topology: a continuously deforming boundary is one object; two disconnected boundaries initialized at different moments are two objects, regardless of how similar their flux profiles appear at any particular instant.

37.6 Entropy and Historical Embedding

The entropy field $S(x,t)$ ensures that flux preservation is not merely instantaneous but historically embedded. Even if two systems share identical instantaneous flux profiles, they may differ in their entropy distributions due to distinct causal histories. Identity therefore requires not only continuity of $\Phi_v(S(t))$ but consistency with the accumulated entropy field, reinforcing the irreversibility arguments of Section 31.

A subject is a recursively maintained, entropy-bearing, flux-coherent process. Anything that terminates this process and starts a new one—however similar—starts a new subject.

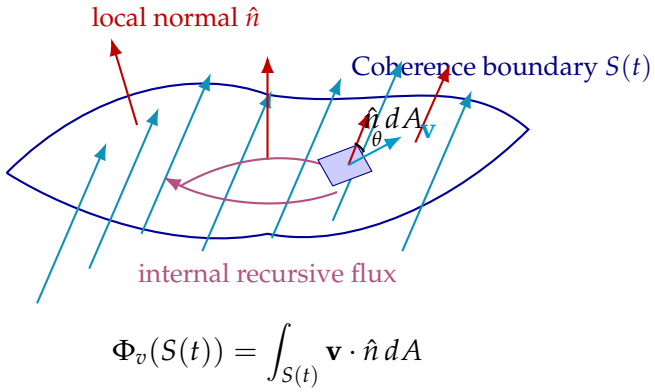

Figure 5: Flux through a coherence boundary in the RSVP framework. The vector field $\mathbf{v}$ passes through $S(t)$ with local normal $\hat{n}$; only the aligned component contributes. Internal recirculation (magenta) distinguishes a subject from a passive medium. Identity requires that this flux structure deform continuously across time.

38 Embedding Flux Identity in the RSVP Field Equations

The flux-based criterion is not an additional postulate; it emerges directly from the RSVP field equations. We now make this grounding explicit.

38.1 Continuity Equation and Boundary Flux

The RSVP vector field $\mathbf{v}$ governs causal propagation across the system. An effective density ρ associated with the scalar field Φ satisfies a continuity equation of the form

$$\partial_t \rho + \nabla \cdot (\rho\, \mathbf{v}) = \sigma,$$

where σ encodes entropy-driven source and sink terms. Integrating over a region $\Omega(t)$ bounded by $S(t)$ and applying the divergence theorem:

$$\frac{d}{dt} \int_{\Omega(t)} \rho\, dV = -\int_{S(t)} \rho\, \mathbf{v} \cdot \hat{n}\, dA + \int_{\Omega(t)} \sigma\, dV.$$

This expresses the evolution of internal structure in terms of boundary flux and entropy production. The system is a subject precisely when the left-hand side remains positive and the flux term remains bounded and continuous—that is, when the boundary is a real constraint on an active internal process, not merely a formal envelope.

38.2 Identity as Invariant Manifold Membership

A coherent subject corresponds to a region $\Omega(t)$ satisfying three conditions: the boundary flux $\int_{S(t)} \mathbf{v} \cdot \hat{n}\, dA$ remains bounded and continuous; internal entropy production σ is regulated to maintain coherence; and the coupling between Φ, $\mathbf{v}$, and S sustains a stable attractor in function space. Under these conditions, the system evolves within the coherence manifold $\mathcal{M}$.

Definition 38.1. A subject in the RSVP framework corresponds to a trajectory $X(t)$ that remains within a connected component of an invariant manifold $\mathcal{M}$, with boundary flux $\Phi_v(S(t))$ admitting a continuous extension in t.

Identity loss corresponds to exit from $\mathcal{M}$, flux discontinuity, or the failure of any continuous solution $X(t)$ to exist across a transition. In a scan-and-copy procedure, the original trajectory $X(t)$ terminates at t_0 and the reconstructed trajectory $\tilde{X}(t)$ begins at $t_1 > t_0$. There is no continuous solution curve connecting them in the RSVP solution space; they are distinct connected components of the trajectory space. In gradual replacement, the system remains a valid solution of the RSVP dynamics throughout; the trajectory is continuously deformed rather than severed.

The flux criterion therefore unifies the geometric, thermodynamic, and dynamical perspectives into a single formal condition.

39 Core Theorem of Continuity-Based Identity

The formal development of this essay can now be condensed into a single theorem that unifies its philosophical, dynamical, thermodynamic, and geometric dimensions.

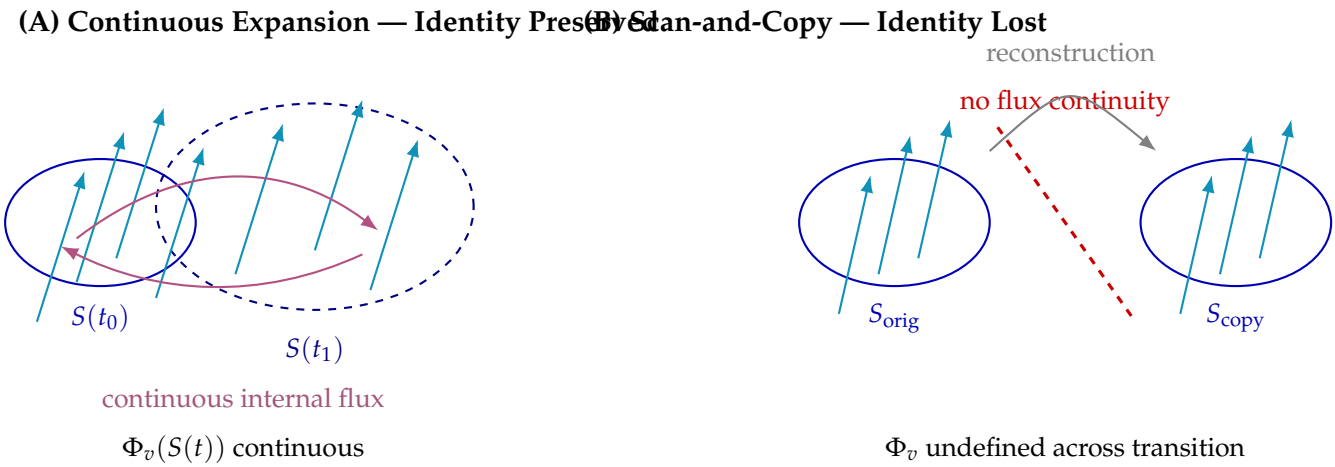

Figure 6: Contrast between identity-preserving and identity-breaking procedures. (A) Continuous expansion: the coherence boundary $S(t)$ deforms while maintaining a continuous, recursively coupled flux structure; identity is preserved. (B) Scan-and-copy: the original and reconstructed systems occupy disconnected regions with no continuous flux bridge; the trajectory is reinitialized, not extended; identity is not preserved.

Theorem 39.1 (Core Theorem of Continuity-Based Identity)**.** *Let $X(t) = (\Phi(t), \mathbf{v}(t), S(t))$ be a trajectory in the RSVP admissible configuration space $\mathcal{A}$, and let $\mathcal{M} \subset \mathcal{A}$ be the coherence manifold defined by bounded coupling $\mathcal{K}(X) \leq K_{\text{crit}}$, structured entropy flow $\mathcal{G}(X) \leq G_{\text{crit}}$, and dynamical stability $\lambda_{\max}(X) \leq 0$. Let $S(t)$ denote a continuous family of coherence boundaries.*

A subject persists if and only if all three of the following conditions hold for all t in the domain:

1. *$X(t)$ is a continuous solution trajectory of the RSVP dynamics;*
2. *$X(t) \in \mathcal{M}$;*
3. *the boundary flux $\Phi_v(S(t)) = \int_{S(t)} \mathbf{v} \cdot \hat{n}\, dA$ admits a continuous extension in t, consistent with the entropy history encoded in S.*

If any one of these conditions fails, identity is not preserved.

Proof. If (1)–(3) all hold, then the system is a single continuous solution curve inside $\mathcal{M}$ with boundary-preserving flux dynamics and historically embedded entropy structure. By the uniqueness of admissible RSVP trajectories (Appendix A), this defines a single identity class.

Conversely: if (1) fails, the trajectory is broken and no continuous continuation exists. If (2) fails, the system exits $\mathcal{M}$ and loses the structural invariants required for unified subjectivity. If (3) fails, the flux-preserving boundary dynamics are interrupted and the system no longer realizes a single physically bounded process. In each case, the resulting configuration is either a new trajectory or a fragmented one, and identity is not preserved. □

Corollary 39.2. *Uploading by scan-and-copy fails conditions (1) and (3): the original trajectory terminates, and no continuous flux bridge connects it to the reconstructed system.*

Corollary 39.3. *Gradual expansion or substrate replacement can preserve identity in principle, provided conditions (1)–(3) remain satisfied throughout the transformation.*

Remark 39.4. Identity is neither purely informational, nor purely topological, nor purely thermodynamic. It is a constrained dynamical property of a physically realized, entropy-bearing, flux-preserving trajectory. The three conditions of the theorem capture these three dimensions simultaneously.

40 Conclusion

The central claim of this essay is that consciousness cannot be uploaded because it cannot be transferred independently of its physical process. The argument has been developed across three registers: philosophical, formal, and applied, and is condensed in the Core Theorem (Section 39): a subject persists if and only if its RSVP trajectory remains a continuous solution within the coherence manifold $\mathcal{M}$, with boundary flux admitting an entropy-consistent continuous extension.

At the philosophical level, the essay has argued that survival requires the continuation of a single experiential stream, and that this requirement is not met by copying procedures regardless of their accuracy. The mistake of the uploading scenario is to conflate third-person pattern equivalence with first-person continuity. A copy that is functionally identical to the original is not the original; it is a new subject initialized with the original's information. The original subject, whose experiential stream was tied to a specific physical process at a specific location in space and time, does not survive the copying procedure. What survives is a representation of the original's pattern, instantiated in a new process at a new location.

At the formal level, the essay has modeled consciousness as a continuous trajectory in the RSVP field configuration space and defined identity as the uniqueness of this trajectory under the RSVP dynamics. Under this formalization, uploading corresponds to a set-valued branching operator that violates the uniqueness of integral curves. The impossibility of transfer follows as a theorem from the determinism of the RSVP dynamics, the irreversibility of entropy evolution, and the admissibility constraints. The coherence manifold $\mathcal{M}$ provides an explicit physical characterization of the regime within which identity is preserved, and the associated theorems establish that exit from $\mathcal{M}$ is irreversible and that instability makes such exit dynamically inevitable. The coherence functional $\mathcal{C}$ provides a continuous measure of proximity to identity loss, separating the graded phenomenon of phenomenological unity from the binary fact of ontological identity.

At the applied level, the essay has shown that these results extend into decision theory. A utility functional grounded in the continuity criterion is identity-consistent—it is

defined over trajectory equivalence classes rather than pattern types—and Theorem 36.2 establishes that simulated or reconstructed agents outside the original's trajectory class are not utility-relevant to the original subject. This dissolves Basilisk-type arguments across all major decision-theoretic frameworks, not by appealing to intuition but by correcting the category error that gives them their apparent force.

The essay has also drawn a sharp distinction between the uploading scenario and the gradual substrate replacement scenario. Gradual replacement, if carried out with continuous causal coupling throughout, does not break the trajectory; it deforms it. Identity is preserved in principle, though constrained in practice by the coherence threshold and the Lyapunov structure of the dynamics. This distinction is not one of degree but of topology: continuous deformation preserves the connectedness of the worldline, while branching or termination destroys it.

The upshot is not that digital minds are impossible or that biological substrate is privileged in any mystical sense. It is that process continuity is the criterion of identity, and any procedure that breaks process continuity—even temporarily, even with an arbitrarily accurate reconstruction—ends one subject and starts another. The illusion of upload arises from the fact that the new subject will have all of the original's memories and will believe itself to be the continuation of the original. That belief is, in a sense, accurate: the new subject is exactly the kind of entity the original would have become. But the original did not become it. The original ended, and something very like it began.

Decision theory requires a theory of identity to avoid generating incentives that attach to the wrong subject. Philosophy of mind requires a physical account of what continuity consists in to move beyond intuition pumps. The RSVP framework provides both: a mathematically precise definition of identity as trajectory membership in a dynamically constrained manifold, and a structural proof that uploading cannot satisfy that definition. The impossibility is not a matter of current technological limitation. It is a consequence of what identity is.

A Existence and Uniqueness of RSVP Trajectories

The formal argument for the impossibility of consciousness transfer relies on the uniqueness of trajectories under RSVP dynamics. This appendix establishes the conditions under which such uniqueness holds and shows how identity follows as a structural consequence of the dynamics.

A.1 Abstract Setting

Let $\mathcal{A}$ be the admissible configuration space of RSVP field configurations $X = (\Phi, \mathbf{v}, S)$, embedded as a Banach manifold in

$$H^1(\Omega) \times H^1(\Omega, \mathbb{R}^3) \times L^2(\Omega),$$

with $\Omega \subset \mathbb{R}^3$ compact. The dynamics are governed by $\frac{dX}{dt} = \mathcal{D}(X)$, where $\mathcal{D} : \mathcal{A} \to T\mathcal{A}$ is the RSVP evolution operator.

A.2 Local Existence and Uniqueness

Definition A.1. $\mathcal{D}$ is **locally Lipschitz** on $\mathcal{A}$ if for every bounded $U \subset \mathcal{A}$ there exists $L > 0$ such that

$$\|\mathcal{D}(X_1) - \mathcal{D}(X_2)\| \leq L\|X_1 - X_2\| \quad \forall X_1, X_2 \in U.$$

Theorem A.2 (Existence and Uniqueness)**.** *Let $X_0 \in \mathcal{A}$ and assume $\mathcal{D}$ is locally Lipschitz. Then there exists a unique trajectory $\gamma : [0, T) \to \mathcal{A}$ satisfying $\frac{dX}{dt} = \mathcal{D}(X)$ with $X(0) = X_0$.*

Proof. This is the Picard–Lindelöf theorem generalized to Banach spaces. The local Lipschitz condition ensures that the integral operator associated with $\mathcal{D}$ is a contraction on a sufficiently small time interval, guaranteeing both existence and uniqueness. □

A.3 Maximal Extension

A local solution can be extended to a maximal interval $[0, T_{\max})$. If the solution remains bounded within $\mathcal{A}$, then $T_{\max} = \infty$ and the trajectory is global. Subjects in the sense of Section 22 are maximal integral curves: trajectories that cannot be extended while remaining in $\mathcal{A}$ and satisfying the dynamics.

A.4 Uniqueness Implies Identity

Proposition A.3. *Let γ_1, γ_2 satisfy $\frac{dX}{dt} = \mathcal{D}(X)$ with $\gamma_1(0) = \gamma_2(0) = X_0$. Then $\gamma_1(t) = \gamma_2(t)$ for all t in their common domain.*

Proof. Immediate from uniqueness under the local Lipschitz condition. □

A given initial condition determines a unique trajectory. There is no ambiguity, no branching, and no multiplicity of continuations within the RSVP dynamics.

A.5 Incompatibility with Branching Operators

If $\mathcal{U}(X_0) = \{X_1, X_2\}$ with $X_1 \neq X_2$, then X_1 and X_2 generate two distinct trajectories. This situation cannot arise from RSVP dynamics applied to X_0, because uniqueness guarantees that only one trajectory originates from any given initial condition.

Corollary A.4. *Branching is not a solution of the RSVP evolution equation. It is an external intervention that produces multiple distinct initial conditions and thereby lies outside the admissible dynamics.*

A.6 Reconstruction as Discontinuous Initialization

If $X_{\text{art}} \neq X(T)$, the trajectory γ_{art} initialized at X_{art} is a valid solution of the RSVP dynamics but is not a continuation of γ.

Proposition A.5. *If $X_{\text{art}} \neq X(T)$, then γ_{art} and γ are distinct integral curves.*

Proof. Distinct initial conditions generate distinct trajectories by uniqueness. □

Even in the idealized limit $X_{\text{art}} = X(T)$, reconstruction produces a trajectory that shares the same future as the original but not the same past. The causal embedding in a continuous history is not recovered by reinitialization. This completes the formal justification of the identity criterion used throughout the essay: one initial condition, one trajectory; one trajectory, one subject.

Bibliography

References

[1] Parfit, D. (1984). *Reasons and Persons*. Oxford University Press.

[2] Parfit, D. (1971). Personal identity. *The Philosophical Review*, 80(1), 3–27.

[3] Nozick, R. (1981). *Anarchy, State, and Utopia*. Basic Books.

[4] Lewis, D. (1976). Survival and identity. In A. Rorty (Ed.), *The Identities of Persons*. University of California Press.

[5] Olson, E. (1997). *The Human Animal: Personal Identity Without Psychology*. Oxford University Press.

[6] Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.

[7] Chalmers, D. (2022). *Reality+: Virtual Worlds and the Philosophy of Mind*. W. W. Norton.

[8] Dennett, D. (1991). *Consciousness Explained*. Little, Brown.

[9] Merleau-Ponty, M. (1962). *Phenomenology of Perception*. Routledge.

[10] Husserl, E. (1913). *Ideas: General Introduction to Pure Phenomenology*. Macmillan.

[11] Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, 83(4), 435–450.

[12] Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18(2), 227–247.

[13] Tononi, G. (2008). Consciousness as integrated information: A provisional manifesto. *Biological Bulletin*, 215(3), 216–242.

[14] Deacon, T. (2011). *Incomplete Nature: How Mind Emerged from Matter*. W. W. Norton.

[15] Varela, F., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press.

[16] Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again*. MIT Press.

[17] Damasio, A. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. Putnam.

[18] Penrose, R. (1989). *The Emperor's New Mind*. Oxford University Press.

[19] Kurzweil, R. (2005). *The Singularity Is Near*. Viking.

[20] Bostrom, N. (2003). Are you living in a computer simulation? *Philosophical Quarterly*, 53(211), 243–255.

[21] Putnam, H. (1967). Psychological predicates. In W. Capitan and D. Merrill (Eds.), *Art, Mind, and Religion*. University of Pittsburgh Press.

[22] Fodor, J. (1974). Special sciences. *Synthese*, 28(2), 97–115.

[23] Sperry, R. (1968). Hemisphere deconnection and unity in conscious awareness. *American Psychologist*, 23(10), 723–733.

[24] Baars, B. (1988). *A Cognitive Theory of Consciousness*. Cambridge University Press.

[25] Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379–423.

[26] Kolmogorov, A. (1965). Three approaches to the quantitative definition of information. *Problems of Information Transmission*, 1(1), 1–7.

[27] Turing, A. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433–460.

[28] Searle, J. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–424.

[29] Graziano, M. (2013). *Consciousness and the Social Brain*. Oxford University Press.

[30] Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138.