A Framework for Consciousness: Autopoiesis, Causal Poiesis, Causal Collapse

From Life to Consciousness

Introduction

Where does consciousness come from? Despite numerous theories – from neuronal correlates to information processing to quantum physical hypotheses – this question remains unanswered if it is considered only within narrow disciplinary boundaries. This essay proposes an alternative approach: consciousness is not understood here as an emergent phenomenon or product of neuronal activity, but as the result of a multi-stage, systemic-thermodynamic process. The key concepts are: autopoiesis, causal poiesis and causal collapse. They mark three fundamental stages on which living systems emerge, center, and finally, under certain conditions, develop consciousness.
The aim of this essay is to systematically reconstruct this path from life to consciousness on the basis of the concepts mentioned above – not in the sense of a goal-oriented development, but as a consequence of certain structural and exchange relations in open systems.
This approach deliberately differs from established information-theoretical models such as Integrated Information Theory (IIT) or Global Neuronal Workspace Theory (GNWT). While these try to explain consciousness through functional information integration, the approach proposed here is based on structural-thermodynamic processes as they occur in the self-organization of living systems. Terms such as "causal core" or "causal collapse" are conceptual proposals that are based on concretely describable systemic properties – even if they have so far only been formalized and operationalized in initial theoretical work (cf. Stegemann, 2025a, 2025b). The present text is therefore not intended as a complete theory, but as a theoretical framework for the development of an interdisciplinary model of understanding.

1. Autopoiesis – The Framework of Autocatalysis

Living systems differ from non-living matter in their capacity for self-preservation through internal organization. This is captured by the term autopoiesis, which was introduced by Maturana and Varela. Autopoiesis describes a system that produces its own components and at the same time forms a boundary to its environment – typically a membrane at first.
The physical basis of this organization is autocatalysis: chemical reactions that reinforce themselves, are cyclically maintained and are protected from dissolution by spatial demarcation. An autopoietic system is in an open exchange with its environment, but organizes it internally in such a way that its structure is preserved.
This can be exemplified by simple protocells whose membrane has selective permeability. This makes it possible to absorb certain molecules and catalyze reactions

within the cell, which in turn ensure the structure of the membrane and the catalytic units themselves.
Autopoiesis is thus not simply a property, but a dynamic-stable configuration that generates local structural stability under conditions of energetic imbalances.

2. Causal Poiesis – The Framework of the Causal Core

With increasing complexity of biological systems, local centers with increased reaction density emerge within open structures. This development is referred to here as causal poiesis: a process in which a causal core emerges from energetic and functional asymmetries (Stegemann 2025a).
The causal nucleus is not only a functional configuration, but in many cases it can also be determined materially. Examples are the nucleus in eukaryotic cells or central neuronal structures in the brain. Such centers bundle causal interactions and enable coordinated system dynamics.
From a thermodynamic point of view, this centering results from non-equilibriums, where certain local configurations are more stable or reactive than others. The formation of a core means an increase in internal control capacity. It is a functional center in which systemically relevant processes converge, stabilize and mutually regulate.
The transition from autopoiesis to causal poiesis can be described mechanistically: local gradients and unstable flows form structural condensations that create internally coupled reaction spaces. This development is visible in biological evolution, e.g. in the formation of eukaryotic cell nuclei or central nervous systems.

3. Causal Collapse – The Framework of Integration

3.1 The Structural Threshold

Causal collapse refers to a transition in which segmented causal pathways transition into an integrated system state that correlates with the experience of consciousness (Stegemann 2025b,c).
In complex systems – such as the human brain – numerous sub-processes are active at the same time. Under certain structural and dynamic conditions, these processes are recursively coupled, resulting in a temporary integration.
It is a qualitative change in the system constellation: a coherent state is created in which previously separate functional areas are coupled within the system. This configuration results from the interplay of dynamic properties such as plasticity, depth of interconnection and energetic exchange.
One example is conscious perception: various sensory, motor, emotional and cognitive processes take place simultaneously, often independently. Only in certain moments does a coordinating connection occur, which is subjectively experienced as consciousness. This linkage is not synonymous with simultaneity, but with systemic coherence: a state in which recursive references produce relative integration.
The transition from causal poiesis to causal collapse can be described as a structural threshold event: When recursive interconnections, plastic coupling and energetic

dynamics exceed a certain degree of internal reconnection, a stable state of activity arises that can be experienced as a subjective presence. The concept of collapse refers to the simultaneity of integration and reduction – many paths condense into a state that can be experienced.

3.2 Physical Realization Through Oscillatory Coherence

What realizes this abstract structural threshold physically in neural systems is oscillatory coherence. Neural oscillations are not mere accompaniments of information processing, but the physical condition that enables recursive feedback to close within a critical time window.

A single nerve cell fires and returns to its resting state. This local event would be reversible and factorizable in isolation. Only when millions of such events are synchronized through oscillatory coupling does the temporal coherence emerge that fulfills the condition of collapse.

Oscillatory synchronization creates a time window within which distributed processes count as simultaneous. When cortical areas oscillate phase-synchronously in the gamma band (30-100 Hz), this means their activity maxima coincide within a window of approximately 10-30 milliseconds. Within this window, no causal order is distinguishable. What was temporally separated becomes temporally fused. This fusion is the physical realization of what is structurally described as the collapse of factorizability.

Three parameters of the oscillatory structure determine how the collapse manifests:

The spatial phase structure $\varphi(x)$ determines which cortical regions are included in the fusion. Regions with consistent phase relations form a functional unit whose parts can no longer vary independently. This phase structure is not static but reconfigures depending on attention, task, and state of consciousness. This reconfiguration is the physical basis for changes in what is present in the epistemic interior space.

The amplitude distribution $A(x)$ modulates the relative strength with which different regions contribute to the integrated unit. High amplitude in visual areas combined with low amplitude in auditory areas means the collapsed state is primarily constituted by visual content. This weighting is not subsequent but part of the collapse dynamics itself. It determines which sensory modalities and which aspects of the environment appear in phenomenal space.

The frequency ω determines the temporal resolution of the collapsed state. High-frequency gamma oscillations (50-100 Hz) create short integration windows of 10-20 milliseconds in which rapid sensory events are bound. Low-frequency theta oscillations (4-8 Hz) create longer integration windows of 125-250 milliseconds in which event sequences fuse into narrative units. The coexistence of multiple frequency bands creates a hierarchical temporal structure of consciousness in which momentary percepts are embedded in longer episodic contexts.

Crucially, these three dimensions of oscillatory structure (phase, amplitude, frequency) are not independently variable. They form a non-factorizable unity. A change in phase structure necessarily affects amplitude distribution because phase-synchronized regions mutually amplify their amplitudes. A change in dominant frequency restructures phase patterns because different frequencies generate different spatial coherence patterns.

This mutual dependence is not accidental but the physical manifestation of structural non-factorizability.
The transition from factorizable to collapsed state can be followed empirically. In states of low integration like deep sleep or anesthesia, EEG measurements show that phase coherence between distant cortical regions breaks down. Oscillations remain locally, but no longer synchronize globally. Thus, they lose their integrative function. The system state becomes again factorizable into local, independent processes. Correspondingly, the interior perspective disappears.
In the transition from sleep to waking, the reverse process occurs. Global phase coherence builds up within seconds. Different frequency bands begin to synchronize over large distances. The amplitude distribution reorganizes so that certain regions become dominant. This physical process is the collapse, which is structurally described as loss of factorizability and phenomenally experienced as awakening.
The mathematical formalization $M(x,t) = S(x,t) \cdot A(x) \cdot \cos(\omega t + \varphi(x))$ captures precisely this structure. The term $\cos(\omega t + \varphi(x))$ represents the oscillatory coherence that realizes temporal and spatial integration. The term $A(x)$ represents the weighting that determines which contents enter integration. The product of these terms with the input pattern $S(x,t)$ is not additive but multiplicative, which mathematically expresses non-factorizability.
The differentiation of the epistemic interior space thus arises directly from oscillatory structure. Different phase configurations generate different spatial content structures. Different frequencies generate different temporal structures. Different amplitude profiles generate different qualitative modalities. These differentiations are not subsequently added to the collapsed state, but are intrinsically contained in the oscillatory dynamics that realizes the collapse.

4. Philosophical and Empirical Implications

A current empirical reference emerges from the study by Luppi et al. (2023) published in 2023. There, functional MRI data and the Integrated Information Decomposition (ΦID) method were used to investigate how information integration in the brain changes under conditions of altered consciousness (e.g. due to anesthesia or impaired consciousness). Particularly striking is the reduction of synergistic information integration in the default mode network (DMN) in the event of loss of consciousness – as well as its restoration when returning to the waking state.
These findings can be interpreted as empirical support for the structural proposed relationships in the present model. In this study, the DMN assumes a gateway function for global integration – corresponding to the causal core described here in the context of causal poiesis. The alternation between integrated and disintegrated states corresponds to the transition to causal collapse – a state of recursive coupling identified in the model with consciousness emergence. The measures used, such as ΦR (revised integrated information) and PCI (Perturbational Complexity Index), represent dynamic correlates for the structural integration described here.
Oscillatory Evidence: The oscillatory framework developed in section 3.2 receives strong empirical support from multiple lines of research:
Melloni et al. (2007) demonstrated that gamma-band synchronization across cortical areas directly correlates with conscious perception in binocular rivalry. When subjects

consciously perceived one of two competing stimuli, gamma oscillations synchronized between frontal and parietal regions specifically for the perceived stimulus. This supports the model's claim that phase coherence $\varphi(x)$ determines which contents enter the collapsed state.

Mathewson et al. (2009) showed that the phase of alpha oscillations at stimulus presentation time predicts detection probability. Stimuli presented during the excitatory phase of the alpha cycle were detected more reliably than those presented during the inhibitory phase. This confirms that the temporal position within the oscillatory cycle ω determines whether an input enters the integration window.

Fries et al. (2001) found that attention selectively increases gamma amplitude and coherence in task-relevant areas. This validates the model's proposition that amplitude distribution $A(x)$ modulates which regions contribute to the integrated state, with attention reconfiguring this distribution.

Massimini et al. (2005) demonstrated that during deep sleep, cortical effective connectivity breaks down while local activity persists. The global synchronization collapses while local oscillations continue. This empirically confirms the transition from non-factorizable (conscious) to factorizable (unconscious) states predicted by the model.

This theory avoids reductionist interpretations by focusing not on individual causes or substrates, but on systemic configurations. It replaces vague assumptions of emergence with comprehensible transitions between different system states. Consciousness does not appear as a separate entity, but as the result of structural conditions within complex, open systems.

At the same time, current empirical findings offer possible starting points for testing this model. For example, studies on neuronal avalanche criticality (Beggs & Plenz, 2003) and chaotic edge dynamics (Shew & Plenz, 2013) show that conscious states typically arise in a dynamic region between order and chaos. This corresponds to the threshold structure of causal collapse described here.

In addition, the Perturbational Complexity Index (PCI) (Schartner et al., 2015; Casali et al., 2013) provides a measurable correlate for the extent of conscious integration. PCI quantifies the complexity of the brain response to external stimulation and has been shown to be a robust indicator of different levels of consciousness. The theory presented here suggests that high PCI values are related to the state of maximally integrated recursive coupling (causal collapse).

Concepts such as the synergistic workspace – which couples gateway regions such as the default mode network with executive networks – can also be interpreted as functional manifestations of a structural core, as described by the concept of causal poiesis.

5. Mathematical Sketch of the Model

The model presented here can be structured mathematically in parts, but without claiming to make subjective experience completely formally comprehensible. Rather, the aim is to correlate the transitions between the system states – autopoiesis, causal poiesis and causal collapse – with dynamic and structural variables.

5.1. Autopoiesis

Autopoietic systems can be described as open, nonlinear flow systems whose stability is based on feedback and boundary formation. Formally, system dynamics can be modeled by a set of coupled differential equations:

$$\frac{dC_i}{dt} = f_i(C_1, \dots, C_n) - \gamma_i C_i + I_i(t) \tag{1}$$

Here, the concentration of a system component, describes autocatalytic interactions, $C_i f_i \gamma_i$ stands for decay constants and for external inputs. The condition of structural coupling is shown by the fact that the system retains its organization despite exchange with the environment.$I_i(t)$

5.2. Causal poiesis

The transition to causal poiesis can be understood as the formation of an attractor center. A causal density function describes the local concentration of pathways:

$$\rho(x) = \int_\Omega K(x, y) \cdot w(y)\, dy \tag{2}$$

A causal nucleus is formed when the density forms a local maximum within a range:$\rho\Omega$

$$x_{\text{core}} = \arg\ \max \rho(x) \tag{3}$$

This corresponds to the emergence of a structural center of increased causal intensity.

5.3. Causal collapse

Causal collapse is described as a transition to a highly integrated recursive state. This can be modeled as a critical threshold process:

$R(t)$describes the density of recursive loops, $\lambda(t)$ the instantaneous coupling strength. If $\lambda(t) \cdot R(t) > \lambda_{\text{crit}}$, a stable state of integration is created:

$$\Psi_{\text{integrated}} \tag{4}$$

This state can be correlated in some way with high values of the Perturbational Complexity Index (PCI), even if the model is not based on measured values but on structural configurations.

5.4. The Oscillatory Equation of Causal Collapse

The physical realization of causal collapse by neuronal oscillations can be formalized as a phase-amplitude coupling equation:

$$M(x,t) = S(x,t) \cdot A(x) \cdot \cos(\omega t + \varphi(x)) \quad (5)$$

where is the integrated system state (the collapsed state), $M(x,t)S(x,t)$ the spatiotemporal input pattern from sensory processing, the spatial amplitude distribution that determines the regional contribution strength, the dominant oscillation frequency that determines the temporal resolution, and the spatial phase structure that determines which regions are functionally coupled.$A(x)\omega\varphi(x)$

This equation captures three critical properties of collapse:

Non-factorizability: cannot be decomposed M into independent contributions of S and the oscillatory term. Multiplication produces irreducible fusion.

State dependency: The same input pattern S produces different integrated states depending on $MA(x)$, and $\omega\varphi(x)$. This explains why identical sensory inputs can be experienced differently in different states of consciousness.

Temporal coherence: Only inputs that arrive within the coherent time window are fully integrated. Inputs outside of this window interfere destructively and are not integrated.$T_{\text{coh}} = 2\pi/\omega$

The threshold condition for collapse can be formulated as:

Collapse occurs when:

$$\int \int |A(x)| \cdot |\cos(\omega t + \varphi(x))| \, dx \, dt > \Phi_{\text{crit}} \quad (6)$$

where represents the critical integration threshold. This combines the abstract structural condition (non-factorizability) with measurable oscillatory parameters.Φ_{crit}

6. Discussion and Outlook

The present theory on the emergence of consciousness via the three stages of autopoiesis, causal poiesis and causal collapse combines structural self-organization with dynamic integration. Methodologically, it differs from classical emergence theories in its systemic-thermodynamic approach and gains plausibility through current empirical studies.

6.1 Empirical Connectivity

This connectivity becomes particularly clear in the light of the study by Luppi et al. (2023), which shows the connection between consciousness and synergistic information integration. The gateway regions (e.g. default mode network) and broadcaster regions (e.g. executive control network) identified there correspond structurally to the concept of the causal core introduced here in the context of causal poiesis. The restoration of global synergies in the transition from the unconscious to the conscious state is also in line with the dynamics of causal collapse.

Likewise, studies on the Perturbational Complexity Index (PCI) (e.g., Casali et al., 2013) confirm the dynamics described in the model: high PCI values stand for integrated, conscious states, while low values are associated with disintegration and unconsciousness. These threshold transitions can be interpreted as quantifiable manifestations of the collapse scheme.
The oscillatory framework adds a crucial layer of mechanistic specificity. It explains not only that integration correlates with consciousness, but how this integration is physically realized through phase-amplitude coupling. This bridges the gap between abstract structural descriptions and concrete neurophysiological mechanisms.

6.2 Open Challenges

Despite these similarities, several questions remain unanswered that should be the subject of future research:
Mathematical precision: The equation for the causal coupling value as the threshold value for the causal collapse requires a more precise connection with experimentally accessible parameters such as PCI or network density. The aim would be to calibrate the theoretical threshold with empirical complexity values.
Falsifiability: The model should formulate specific predictions, e.g., on the dynamics of conscious transitional states (e.g., when falling asleep, during absences, or psychedelic experience) that go beyond existing models. The oscillatory framework provides testable predictions: disruption of phase coherence at specific frequencies should produce specific alterations in consciousness quality, not merely threshold changes.
Step dynamics: The transitions between autopoiesis, causal poiesis and collapse could be interpreted as phase transitions and quantitatively modeled – for example by gradients in energy flows, network configuration or entropy.

6.3 Perspective: Testability via TMS-EEG

A promising next step would be to try to test the model assumptions experimentally within the framework of existing TMS-EEG protocols. The following hypotheses could be operationalized:
The causal nucleus (e.g. DMN) shows a particularly high degree of re-integration (high-frequency synergies) under TMS excitation during waking states.
Causal collapse is associated with a sudden increase in global complexity (PCI > 0.3), which occurs specifically in deliberate transitions.
When returning from unconsciousness, the gateway cores (DMN) reconfigure first before broadcaster networks (ECN) generate full integration – which harmonizes with the phase model proposed here.
Phase-locked TMS at different phases of the alpha cycle should produce differential effects on conscious perception, with maximal disruption at critical phase angles predicted by the model.

7. Conclusion

Autopoiesis, causal poiesis, and causal collapse describe three successive processes by which structurally centered, functionally integrated systems emerge from physical, open systems – up to the emergence of what we call consciousness. The strength of

this model lies in its interdisciplinary connectivity: it combines thermodynamics, systems theory, biology, neuroscience and philosophy of mind – without falling into metaphysical speculation or reductionist short-circuits.
The critical innovation lies in specifying how the abstract structural threshold of causal collapse is physically realized through oscillatory coherence. The phase-amplitude coupling equation $M(x,t) = S(x,t) \cdot A(x) \cdot \cos(\omega t + \varphi(x))$ provides a concrete bridge between structural non-factorizability and measurable neural dynamics. This transforms consciousness from an explanatory gap into an empirically tractable phenomenon.
The thesis is: With this perspective, consciousness as the result of coherent system states under conditions of nonlinear, open dynamics can be understood – within scientific explainability and without recourse to teleological or dualistic assumptions. The oscillatory framework makes this understanding not only conceptually coherent but empirically testable.

8. Literature

- Beggs, J. M., & Plenz, D. (2003). Neuronal avalanches in neocortical circuits. Journal of Neuroscience, 23(35), 11167–11177.
- Casali, A. G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K. R., ... & Massimini, M. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. Science Translational Medicine, 5(198), 198ra105.
- Fries, P., Reynolds, J. H., Rorie, A. E., & Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. Science, 291(5508), 1560-1563.
- Luppi, A. I., Mediano, P. A. M., Rosas, F. E., Holland, N., Fryer, T. D., O'Brien, J. T., ... & Stamatakis, E. A. (2023). A synergistic workspace for human consciousness revealed by Integrated Information Decomposition. eLife, 12, e88173.
- Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H., & Tononi, G. (2005). Breakdown of cortical effective connectivity during sleep. Science, 309(5744), 2228-2232.
- Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To see or not to see: prestimulus alpha phase predicts visual awareness. Journal of Neuroscience, 29(9), 2725-2732.
- Maturana, H. R., & Varela, F. J. (1980). Autopoiesis and cognition: The realization of the living. D. Reidel Publishing Company.
- Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W., & Rodriguez, E. (2007). Synchronization of neural activity across cortical areas correlates with conscious perception. Journal of Neuroscience, 27(11), 2858-2865.
- Schartner, M. et al. (2015). Complexity of multi-dimensional spontaneous EEG decreases during propofol induced general anaesthesia. PLoS One, 10(8), e0133532.
- Shew, W. L., & Plenz, D. (2013). The functional benefits of criticality in the cortex. The Neuroscientist, 19(1), 88–100.

- Stegemann, W. (2025a). The Emergence of Causal Cores: From Thermodynamic Foundations to Biological Functionality. https://doi.org/10.5281/zenodo.15558264
- Stegemann, W. (2025b). Consciousness – Integrated Model of Causal Collapse. Zenodo. https://doi.org/10.5281/zenodo.15350050
- Stegemann, W. (2025c). Consciousness - Expanding the Model of Causal Collapse. https://doi.org/10.5281/zenodo.15351381
- Tononi, G. (2004). An information integration theory of consciousness. BMC Neuroscience, 5(1), 42.