A Unified Information Theory of Subjective Cognition

Authors: Jian-Sheng Kang, Jian-Sheng Kang

Date: 2025-09-22T00:00:00+00:00

Abstract

Consciousness is considered the most challenging problem in the science of the mind, especially Chalmers' hard problem. From the standpoint of the systems science philosophy, many consciousness theories overlook the structure-function correlation principle, thereby reducing their explanatory power regarding consciousness. Therefore, we initially introduce a philosophy for systems science. Based on the structure-function correlation principle, we realize that the neck structure of the dendritic spine is a key for information encoding in the frequency domain. Consequently, the whole work is discussed in the frequency domain. At the level of sensory modalities, the intrinsic function of simple auditory neurons encodes information as a wave function. Subsequently, at the mesoscopic level, we successfully resolve the quantum mechanism for frequency adaptation, a representation of energy constraint. Particularly, we find that the accommodation mechanism and the multi-layered retina are specialized to employ the principles of Fourier optics and quantum optics. Specifically, the stratified structures, positions, and functions of retinal ganglion cells and bipolar cells are specialized for the Fock states in quantum optics. The entire brain can be considered as a neuronal quantal field, comprising billions of neurons and trillions of synapses, which function intrinsically as wave functions and perceptive units, respectively. Furthermore, we address both the "hard problem" and "easy problems" of consciousness by analyzing them in the frequency domain, thereby fully resolving the pseudo-metaphysical hard problem. In summary, it is thrilling to understand the mechanism of neuronal function and consciousness, and this Unified Information Theory (UIT) provides a bridge between microscopic and macroscopic levels. Additionally, building on the identification of time perception as an intrinsic manifestation of causal asymmetry and a representation of imaginary temporal dynamics, we derive a quantum geometric derivation of Newton's gravitational constant and resolve the cosmological constant problem through discrete spacetime structures, unified by Hodge duality mechanisms.

Full Text

Preamble

A Unified Information Theory of Subjective Cognition: Quantum Information Processing from Neuron to Consciousness

Jian-Sheng Kang*
Clinical Systems Biology Laboratories, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, China
*Correspondence should be addressed to J.-S. Kang, Clinical Systems Biology Laboratories, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, 450052, China. E-mail: kjs@zzu.edu.cn ORCID ID: 0000-0002-2603-9718.

Abstract: Consciousness represents the most challenging problem in the science of the mind, particularly Chalmers' hard problem. From a systems science philosophy standpoint, many consciousness theories overlook the structure-function correlation principle, thereby reducing their explanatory power. We therefore introduce a philosophy for systems science and, based on the structure-function correlation principle, identify the dendritic spine neck as a key structure for information encoding in the frequency domain. Consequently, our entire framework operates in the frequency domain. At the sensory modality level, the intrinsic function of simple auditory neurons encodes information as a wave function. At the mesoscopic level, we resolve the quantum mechanism for frequency adaptation as a representation of energy constraints, finding that accommodation mechanisms and the multi-layered retina specialize in employing principles of Fourier optics and quantum optics. Specifically, the stratified structures, positions, and functions of retinal ganglion cells and bipolar cells are specialized for Fock states in quantum optics. The entire brain can be considered a neuronal quantal field comprising billions of neurons and trillions of synapses that function intrinsically as wave functions and perceptive units, respectively. Furthermore, we address both the "hard problem" and "easy problems" of consciousness through frequency-domain analysis, thereby resolving the pseudo-metaphysical hard problem. In summary, this Unified Information Theory (UIT) provides a thrilling mechanistic understanding of neuronal function and consciousness, bridging microscopic and macroscopic levels. Building on the identification of time perception as an intrinsic manifestation of causal asymmetry and imaginary temporal dynamics, we derive a quantum geometric derivation of Newton's gravitational constant and resolve the cosmological constant problem through discrete spacetime structures unified by Hodge duality mechanisms. Moreover, neural systems encode information through frequency adaptation, aligning spike intervals with prime-rich regions to realize the Hilbert-Pólya conjecture, providing a novel pathway to explore deep connections between biology, geometry, and number theory while offering insights into proving the Riemann Hypothesis.

Keywords: Frequency adaptation; Frequency Domain; Quantum Information; Unified Information Theory (UIT); Hilbert-Pólya conjecture; Riemann Hypothesis; Consciousness; Systems Science Philosophy

Fundamental Conceptual Architecture with Argumentative Synopsis

UIT-01-Section 1: Frequency-Domain Neural Computation and Integrated Framework for Fisher-Shannon Dynamics

Abstract: This work presents a unified framework integrating systems science philosophy, information theory, and neurophysiology to decode neuronal computation. At its core is frequency-domain analysis of neural signals, where action potentials are modeled as Bernoulli trials, revealing that Fisher information (precision) and Shannon entropy (uncertainty) are intrinsically linked through the Tangled Factor (TF), a geometric curvature quantifying their mutual constraint. The Logit function serves as a canonical parameter for perceptual decisions, enabling linear feature integration in neural circuits while adhering to the Weber-Fechner Law's logarithmic scaling. Complementing this is Mental Knots (MK), a metric derived from non-commutative Lie brackets that captures cognitive tension via directional bias and information-entropy coupling (TF). These concepts are grounded in differential geometry, where Fisher information defines local curvature and Shannon entropy reflects global variability. Key findings include: (1) neuronal structures like dendritic spines act as finite-frequency filters; (2) synaptic integration operates via convolution of Poisson-distributed vesicle release with frequency responses; (3) energy-efficient coding emerges from the interplay of ATP metabolism and CK/PCr buffering. By bridging algebraic (Logit, MK) and geometric (TF) perspectives, this framework elucidates how perception balances computational efficiency with dynamic adaptability, offering a principled approach to understanding neural decision-making under uncertainty.

UIT-02-Section 2.1: Topological Protection in Auditory Neuronal Coding: Quantum Dynamics, Energy Constraints, and Information Encoding

Abstract: This study elucidates the role of topological protection in quantum neuronal mechanisms within the auditory system, focusing on how energy constraints and spatiotemporal dynamics govern information encoding. Auditory neurons decompose sound into frequency-specific signals via tonotopic organization, where simple cells encode stimuli through wave functions invariant under translational symmetry, reflecting characteristic frequencies and phase stability. Spike frequency adaptation and refractory periods arise from energy-efficient processes mediated by the creatine kinase/phosphocreatine (CK/PCr) system, with creatine molecules acting as electric dipoles influenced by oscillatory neuronal fields. Using quantum perturbation theory, we derive an analytic solution for time-dependent Hamiltonian evolution, demonstrating that topological protection emerges from full circular dipole trajectories that ensure robust phase coherence against thermal noise and decoherence. This mechanism links ATP consumption ($\sim 10^6$ molecules per action potential), creatine concentration (5-10 mM), and refractory periods, establishing a unified framework for metabolic

constraints in information encoding. The resulting quantized phase evolution mirrors topological conditions in condensed matter systems, while the statistical sparsity of aligned dipoles explains millisecond-scale firing intervals. By replacing abstract constructs with testable quantum dynamics, this theory provides a biophysically plausible explanation for cognitive limits like Miller's "magic number seven," rooted in energy-optimized neural computation. The interplay between topological protection, energy coupling, and spatiotemporal adaptation underscores a fundamental principle governing neuronal information processing under biological constraints.

UIT-03-Section 2.2-2.3: Hierarchical Visual Processing: Fourier, Radon, Gabor, and Their Inverse Transformations in Retinal to Cortical Pathways

Abstract: The visual system processes information through a hierarchical pathway involving the eye, lateral geniculate nucleus (LGN), and primary/secondary visual cortices. Light is focused on the retina, where photoreceptors convert it into electrical signals, with rod cells detecting single photons and cone cells encoding color and detail. The retina functions as a Fourier plane, capturing spatial frequency, phase, and amplitude information, which is further decomposed via Radon transforms in the LGN-visual cortex pathway. Retinal ganglion cells (RGCs) act as pixel units integrating gradients from bipolar cells, while LGN neurons perform Gabor-like transformations, combining Gaussian windows with Fourier analysis to encode visual features. Simple cells in V1 cortex respond selectively to specific orientations and spatial frequencies, their activity reflecting a Fourier-sliced hologram of the visual scene. Complex cells in extrastriate areas, such as V4 and V5, integrate this information through inverse Radon transforms, reconstructing detailed representations for tasks like motion detection and color perception. These processes involve quantum-inspired principles, including coherent states and Fock states, with inhibition modulating temporal frequency tuning. Experiments on primate visual cortex demonstrate that simple neurons exhibit phase-specific firing patterns, while complex neurons integrate multi-dimensional features through reciprocal projections, enabling conscious perception of objects and scenes.

UIT-04-Section 2.4: Structure-Function Correlation in Sensory Systems: Biophysical and Computational Principles

Abstract: This section synthesizes structural-functional principles, quantum mechanics, and meta-metrics governing sensory perception. Auditory/visual systems decompose stimuli into frequency domains via Fourier-like transformations (cochlea/lens), with tonotopic/retinotopic maps encoding spatial information through inverse Radon transforms. Quantum optics and Bayesian inference reveal probabilistic neuronal coding, exemplified by retinal Fock states and gamma-distributed priors in the ITC enabling efficient visual recognition. Meta-metrics like confidence gain (CG), mastery level (ML), and temporal precision quantify cognitive dynamics, bridging neural computation with subjective experience. Time perception emerges from causal asymmetry, with imaginary

time encoding probabilistic coherence. Sensory cortices integrate spatiotemporal meta-metrics, highlighting how neural plasticity and Bayesian updates shape perception and cognition.

UIT-05-Section 2.4.5: Time Perception and the Fundamental Cosmic Architecture: Quantum Geometry, Hodge Duality, and the Analytic Derivation of Newton's G

Abstract: Building on the identification of time perception as an intrinsic manifestation of causal asymmetry and imaginary temporal dynamics, this work presents a unified framework linking quantum geometry, spacetime emergence, and fundamental constants through an 11-dimensional supersymmetric structure. By extending Hamiltonian mechanics into hyperkähler manifolds via imaginary time, we show how temporal dynamics arise from symmetry breaking, with effective spacetime dimensions defined as 11 ($2 \times 4 + 3$). The cosmological constant problem is addressed by connecting supersymmetry breaking to time asymmetry, where vacuum energy suppression emerges naturally from geometric constraints. Newton's gravitational constant G_0 is derived analytically through Hodge duality between photon and graviton forms, dimensional flux dilution, and quantum corrections involving the Barbero-Immirzi parameter $\gamma \approx -0.241873$, a value consistent with loop quantum gravity bounds. The framework reveals gravity as an emergent phenomenon from discrete geometric structures, while dark energy scales are tied to vacuum fluctuations in an 11D photon-graviton duality field. Notably, the derived upper energy bound for vacuum fluctuations ($E_{ub} \approx 7.95$ meV) closely matches cosmic microwave background photon energies, suggesting a quantum-gravitational origin for dark energy. The analysis also implies that real time and classical spacetime emerge from an underlying discrete, imaginary-time framework, with implications for quantum propulsion concepts leveraging geometric asymmetries. This approach bridges electromagnetism, quantum gravity, and cosmology, offering a first-principles explanation for interaction strengths while redefining fundamental physical constants through topological and symmetry-based principles.

UIT-06-Section 3: Hierarchical Modular Neural Mechanisms in the Emergence of Self-Consciousness: Integrating Fault-Tolerance, Quantum Dynamics, and Metabolic Efficiency

Abstract: This work explores the emergence of self-consciousness through hierarchical and modular neural mechanisms, emphasizing the integration of sensory modalities and their submodalities into a fault-tolerant system. Self-consciousness arises from unified internal representations generated by active synapses across distributed brain regions, where complex cells process multidimensional information via wavefunction dynamics and structure-function correlations. Clinical disorders such as achromatopsia, akinetopsia, and spatial disorientation highlight the role of specific cortical areas in maintaining perceptual and cognitive functions, revealing fault-tolerant designs that sustain functionality despite lesions. The "hard problem" of consciousness—explaining subjective experience—is addressed through quantum wave mechanics and energy-efficient

information processing, demonstrating that both hard and easy problems share foundational principles in complex cell dynamics. By quantifying information emergence via neural bandwidth, spike frequency adaptation, and energy constraints, this study bridges biophysical mechanisms with computational models, showing how adaptive processes optimize information capacity while minimizing metabolic costs.

Alzheimer's disease is framed as a systemic energy crisis in the self-awareness network, where hypometabolism in key regions like the retrosplenial cortex and prefrontal cortex disrupts spatial integration and self-referential processing, underscoring the critical role of metabolic efficiency in sustaining consciousness. These findings unify clinical observations, neurobiological principles, and computational frameworks to advance understanding of how neural systems generate conscious experience through dynamic, energy-constrained interactions.

UIT-07-Section 4: Frequency Adaptation and Prime-Based Information Encoding in Neural Systems

Abstract: This study explores how frequency adaptation in neurons might encode prime-based information by drawing parallels between neuronal dynamics and prime number distribution. Using geometric and algebraic frameworks inspired by Lie group theory and number theory, we model neuron refractory periods and spike timing through logarithmic transformations, revealing a hierarchical structure underlying neural timing mechanisms. This suggests that neurons dynamically adapt their firing frequencies to leverage "prime-rich" intervals for efficient information processing, such as the "magic seven" phenomenon observed in working memory. The work further bridges neuroscience and number theory by demonstrating that this scaling law corresponds to the Selberg trace formula for modular surfaces. The spectral characteristics of neural activity hint at a mathematical pathway linking physical system spectra to the non-trivial zeros of the Riemann zeta function, thereby converging toward the Hilbert-Pólya conjecture. This framework offers a novel biological pathway to prove the Riemann Hypothesis through frequency-adapted prime encoding.

Collectively, these findings demonstrate how biological neural systems inherently leverage prime number distributions to achieve optimal information representation. This interdisciplinary approach provides transformative insights for both neuroscience and pure mathematics, suggesting that neurons may use prime-based patterns to enhance computational efficiency and robustness in complex environments.

Preface: A Philosophy of Systems Science

The mind-body relationship constitutes not only a scientific problem but also a profound philosophical theme. Surveys reveal that 73.3% of Americans and 43.2% of Europeans believe in an afterlife, reflecting a dualistic view that the

mind can survive death (Riekki et al., 2013), while the majority of college students similarly endorse mind-brain separateness (Demertzi et al., 2009). In *Introduction to Systems Philosophy* (1972), Ervin Laszlo rejected dualism and pluralism in favor of holism and system monism. However, the prevalence of dualistic views suggests fundamental problems in scientific explanations of the mind-body relationship, necessitating a brief discussion of relevant philosophical history and the introduction of a systems science philosophy.

In Western thought, dualism is closely linked to René Descartes' meditations (1641), which posited that consciousness and self-awareness are nonphysical. In East Asia, Zhu Xi's Neo-Confucian Li-Qi dualism has been influential since China's Southern Song Dynasty. "Li" represents abstract, unchanging universal patterns governing the cosmos, while "Qi" constitutes the tangible, mutable substance of the material world (Fung and Bodde, 1942, 1948). This dualism originates from the systematic worldview of the *I Ching* (Book of Changes), whose "Xici" appendix states: "There is in the Changes the Great Primal Beginning (Taiji). This generates the two primary forces. The two primary forces generate the four images. The four images generate the eight trigrams" (Wilhelm and Baynes, 1967). The Great Primal Beginning (Taiji) represents a monistic concept.

In modern systems theory, Austrian biologist Karl Ludwig von Bertalanffy (1901-1972) pioneered general systems theory in the West, while Chinese scientist Qian Xuesen (1911-2009) distinguished himself in engineering cybernetics in the East. In a 1982 report, Qian advocated integrating general systems theory, information theory, and cybernetics within the broader field of systems science, with systems theory at its center (Qian, 1984). Subsequently, Wei Hongsen and Zeng Guoping expanded this concept in *System Theory - Systems Science Philosophy* (1995), promoting systematicism based on general system theory, information theory, cybernetics, dissipative structure theory, synergetics, hypercycle theory, catastrophe theory, chaos theory, and fractal theory. They condensed systematicism into eight key concepts: integrity, hierarchy, openness, teleonomy/purposefulness, catastrophe, stability, self-organization, and similarity, along with five fundamental laws: structure-function correlation, feedback in information, competition and cooperation, order emerging from fluctuations, and optimization of evolutionary processes (Wei and Zeng, 1995).

Embracing systematicism is essential for addressing today's intricate challenges, requiring an interdisciplinary approach that weaves together insights from multiple disciplines, particularly when examining theoretical constructs and fundamental understandings of neuron and brain functions and consciousness. The following analyses will primarily rely on the laws of structure-function correlation and optimization of evolutionary processes, along with the principles of integrity, hierarchy, openness, teleonomy, and stability. These discussions prioritize logical coherence and clarification of fundamental concepts over rigorous mathematical and physical details; readers are encouraged to consult reference materials for deeper comprehension of technical terminology.

1 Matter (Structure), Energy and Information

From a physical perspective, the brain can be characterized as an open system that exhibits integrity through coordinated interactions of its hierarchical components, openness through exchange of matter, energy, and information with its environment, and stability through maintenance of functional states despite external disturbances. Matter, energy, and information constitute three indispensable elements for maintaining the brain's integrity and stability over time. Time remains a fundamental component, though it often goes unexpressed explicitly when discussions are conducted primarily in the rate/frequency domain, which is analogous to a projected or reciprocal space in mathematical and physical contexts. The frequency domain representation offers a more concise and lucid description compared to the time domain, as it circumvents the need to engage with an abundance of dynamic parameters and intricate temporal details that have been intensively documented, such as the biophysics of spike timing in single neurons (Koch, 1998) and spike rates as Bayesian posterior probabilities (Rieke et al., 1999).

The brain is an evolutionarily optimized structure that emerged from matter and is subject to constraining pressures of energy on evolutionary processes. Its hierarchical structures and functions have evolved for processing information through storage, encoding, decoding, transmission, and quantitation. Concurrently, the precise definition of information, including its emergent properties, holds the key to unlocking the intricacies of neuronal and brain mechanisms. Subsequent analyses will dissect the links between the brain's hierarchical structures and its information-processing functions.

1.1 Information Storage in Binary Form

The brain constitutes approximately 2% of human body mass yet accounts for roughly 20% of the body's energy consumption (Raichle and Gusnard, 2002). Notably, the brain's energy efficiency significantly surpasses that of contemporary computing and artificial intelligence systems by approximately 9×10^8 times (Stiefel and Coggan, 2023). This efficiency likely results from evolutionary optimization due to energy constraints, indicating that the all-or-none action potential represents an efficient form for storing and processing information at any given moment.

This concept is supported by Norbert Wiener's conclusion in Chapter V of the 1965 second edition of *Cybernetics*, where he established the minimum cost required to reliably store one bit of information in a binary system (Wiener, 1965). These constraints apply not only to information stored in action potentials at any given moment but also likely pertain to long-term information storage, such as via the stability of synaptic structures and circuits for recurrence of action potential patterns and potential epigenetic modifications (Santoni et al., 2024).

In addition to its all-or-none nature, the action potential has a short refractory period and shows adaptation to stimulation; furthermore, neurons encode in-

formation intensity as the frequency of impulses (Adrian and Zotterman, 1926; Koch, 1998; Kandel et al., 2021). The action potential initiates at the distal end of the axon initial segment (AIS) [FIGURE:1], where a high concentration of low-threshold sodium channel Nav1.6 is found; concurrently, the proximal region of the AIS is enriched with high-threshold sodium channel Nav1.2, which is crucial for backpropagation of the action potential (Hu et al., 2009). Upon arrival at the presynaptic terminal, voltage-gated calcium channels open, allowing calcium influx detected by the calcium sensor synaptotagmin (Xu et al., 2007), which activates the SNARE complex to facilitate neurotransmitter release from presynaptic vesicles (Südhof and Rothman, 2009).

1.2 The Release Probability of Presynaptic Vesicles

Synaptic vesicle pools, containing from nearly a hundred to approximately a million vesicles, can be categorized into the readily releasable pool (RRP) docked to the presynaptic active zone, recycling pool, and reserve pool (Rizzoli and Betz, 2005) [FIGURE:1]. In contrast, the average number of synaptic vesicles per active zone is around two hundred (Rizzoli and Betz, 2005). RRPs of rat hippocampal varicosities contain $\sim 5-10$ vesicles (Schikorski and Stevens, 2001). Similarly, RRPs per active zone have $\sim 2-9$ vesicles in the rat Calyx of Held (~ 600 active zones) (Sätzler et al., 2002), a giant presynaptic terminal in the auditory system (Guo et al., 2015) with a total RRP size around $\sim 3000-5000$ vesicles.

contrasts time-domain and frequency-domain viewpoints. In Chapter 4 of *Biophysics of Computation: Information Processing in Single Neurons* (Koch, 1998), synaptic input is modeled in the time domain from the dendritic tree perspective. The probability of quantal release is typically represented using a binomial distribution, while synaptic weight is averaged over a Poisson-distributed presynaptic spike pattern (Koch, 1998). Rate theory regards spike rates as Bayesian posterior probabilities (Rieke et al., 1999). Currently, activation functions in artificial neural networks, such as the Rectified Linear Unit (ReLU), emulate neurotransmitter release at the synapse to introduce non-linearity.

In the frequency domain, we initially focus on a single active zone within the presynaptic terminal. The RRP, as its name suggests, contains vesicles ready for immediate release upon action potential arrival. As noted above, RRP size is usually less than 5% of the total number of vesicles per active zone. Considering that synaptic transmission is stochastic and release probability can be quite low (Koch, 1998), the Poisson distribution appropriately models vesicle release probability. More precisely, the non-homogeneous Poisson process is crucial for presynaptic modeling, especially when release probability varies over time or with synaptic activity, such as through modulation by presynaptic inhibition (Naumann and Sprekeler, 2020).

Mathematically, the random arrival of action potentials at presynaptic terminals can be modeled using Bernoulli trials with a time-varying firing rate $\lambda(t)$. The

probability of firing within a time interval Δt is $\lambda(t) \cdot \Delta t$. Given that k vesicles are released per action potential, the probability $p(k; \lambda(t)\Delta t)$ follows a non-homogeneous Poisson process (Table 1.2.1), generalized from the Poisson process with independent but nonstationary increments (Ross, 1996; Lu and Zhang, 2012):

$$p(k; \lambda(t)\Delta t) = \frac{(\lambda(t)\Delta t)^k e^{-\lambda(t)\Delta t}}{k!} \tag{1.2.1}$$

When the product $\lambda(t) \cdot \Delta t$ is time-invariant, Equation 1.2.1 becomes a standard Poisson distribution. Intriguingly, for simplicity without loss of generality, the standard Poisson distribution can discuss RRP size per presynaptic active zone [FIGURE:1]. As demonstrated in Figure 1.2.1, if Δt is brief (e.g., 1 ms) and $\lambda(t)\Delta t \leq 1$, the probability of $k \geq 5$ vesicles being released concurrently during interval Δt approximates zero, implying that 5-10 vesicles in the RRP per active zone may encompass full probabilities for short intervals.

1.3 The Function of Dendritic Spine in the Frequency Domain

We now examine the postsynaptic terminal model in the frequency domain . As shown in [FIGURE:1], the spine neck has an average diameter of 0.2 ± 0.06 μm for layer 2/3 pyramidal cells from mouse visual cortex (Arellano et al., 2007). In the time domain, spine necks function as diffusion barriers for large proteins, small ions, and molecules (Adrian et al., 2014; Tønnesen et al., 2014). Molecular diffusion is acutely sensitive to variations in spine neck width and inversely proportional to its cross-sectional area (Tønnesen et al., 2014), implying that spine neck structure dictates diffusion rates. Consequently, the spine neck's role as a diffusion barrier can be accurately modeled using a point spread function (PSF), with the one-dimensional $\mathrm{sinc}(x)$ function, defined as $\sin(x)/x$ with $\mathrm{sinc}(x = 0) = 1$, serving as an appropriate and simple example. The sinc function is fundamental in probability and information theory (Woodward and Davies, 1952).

An interesting property of the sinc function is that integrals of $\mathrm{sinc}(x)$ and $\mathrm{sinc}^2(x)$ over all real numbers are equal (Baillie et al., 2008), simplifying extension from one to two dimensions. The sinc function possesses remarkable properties in summation, multiplication, and integration (Kac, 1959; Baillie et al., 2008; Ortiz Gracia and Oosterlee, 2016). As illustrated in Figure 1.3.1, the Fourier transform of the sinc function is a finite rectangle window in the frequency domain, signifying that dendritic spine structure effectively represents a finite frequency response.

1.4 Neuron as a Summator of Synchronous Frequencies and a Bernoulli Binary Generator

Having discussed presynaptic and postsynaptic structures, functions, and models, we can now address synaptic integration, which is fundamentally a convolu-

tion process between the presynaptic input signal and the postsynaptic impulse response. In the frequency domain, the overall dendritic output results from superposition of individual outputs, each obtained by convolving presynaptic input with various finite frequency responses of spines on the dendritic tree. Convolving the non-homogeneous Poisson distribution parameterized by $\lambda(t) \cdot \Delta t$ (Equation 1.2.1) with finite frequency responses of spines eliminates the time interval Δt due to the scaling property of Fourier transformation [FIGURE:1]. Additionally, given the all-or-none nature of action potentials, amplitude scaling is redundant. As the dendritic tree receives presynaptic spikes conforming to a Poisson distribution (Koch, 1998), and considering vesicle release probability characterized by the non-homogeneous Poisson process with rate parameter $\lambda(t)$ (Equation 1.2.1), the resulting postsynaptic activity can be described as:

$$B(t) = \sum_{i=1}^{N} \lambda_i(t) k_i e^{-\lambda_i(t)} \tag{1.4.1}$$

where $B(t)$ represents the summation of individual spine outputs in response to presynaptic inputs; N denotes the substantial count of postsynaptic spines, exemplified by over 17,000 spines on mouse layer III pyramidal neurons (Ballesteros-Yáñez et al., 2006) and up to 97,853 spines on human Purkinje cell dendrites (Masoli et al., 2024).

Equation 1.4.1 can also be derived by modeling neuronal function as a busy period $B(t)$ in an M/G/1 queueing system using its AIS and dendritic spines, where spikes arrive following a Poisson process with rate $\lambda(t)$. Here, M stands for "Markovian" arrivals (memoryless Poisson process), G represents "General" service time distribution (non-exponential patterns), and 1 indicates a single server (AIS) (Ross, 1996, section 2.3.1). Alternatively, as discussed in Section 2.2.2.1, Equation 1.4.1 can be formulated by treating synapses as coherent states, since the n-th number state in a coherent state follows a Poisson distribution in quantum optics (Gerry and Knight, 2005).

Moreover, as signaling molecules such as Na^+ from excitatory synapses and Cl^- from inhibitory synapses diffuse along ion gradients through the dendritic tree to the distal AIS to trigger or inhibit action potentials, the dendritic tree structure can enhance synchronous frequencies while attenuating non-synchronous frequencies, effectively transforming a conditional compound Poisson process into more discrete, all-or-none behavior resembling a Bernoulli process. Intriguingly, given that simple cells receive one simple type of input (Hubel and Wiesel, 1962), Equation 1.4.1 implies that simple cells function as linear time-invariant (LTI) units in the frequency domain [FIGURE:1]. Equation 1.4.1 also yields Bernoulli binary outputs of simple cell activations with an average firing rate/frequency of $\lambda(t)$. For clearer understanding of $\lambda(t)$, refer to Equation 2.1.3 or coherent states described in Section 2.2.2.1. Hence, a neuron can be considered to function as both a summator of synchronous frequencies and a Bernoulli binary generator.

1.5 The Bernoulli Process and Ergodic Properties

The Bernoulli process models random phenomena where each trial outcome is either success or failure. Named after Swiss mathematician Jacob Bernoulli, it describes neuronal function in the frequency domain [FIGURE:1], offering both benefits and drawbacks. The independence of Bernoulli trials, guaranteed throughout the action potential refractory period [FIGURE:1], enables quick adaptability and decision-making based solely on current information without considering prior outcomes (Markov property). However, ignoring long-term dependencies is also disadvantageous since the Bernoulli process is memoryless.

The discrete nature of Bernoulli processes can conflict with continuous sensory data streams, a tension resolved through the concept of ergodicity. A Bernoulli process can be ergodic if it satisfies certain conditions, such as simple cells with time-invariant frequencies in the frequency domain discussed above. Ergodicity is essential in statistical mechanics, probability theory, functional analysis, and dynamical systems theory (Ashley, 2015). In statistical mechanics, ergodicity ensures equivalence between time averages and ensemble averages. Ergodic systems are characterized by stationary or invariant measures, signifying that the probability distribution of system states is time-invariant, as in Bernoulli shifts and Markov shifts (Quas, 2011). Additionally, a non-homogeneous Poisson process driven by an almost periodic intensity function possesses ergodic properties (Rolski, 1990), providing a natural rationale for vesicle release probability in presynaptic terminals adopting a non-homogeneous Poisson distribution (Equation 1.2.1 and [FIGURE:1]).

1.6 The Modular Group and Congruency Subgroups

Bernoulli trial outcomes (failure or success) can be expressed as $\{0, 1\}$, forming an additive cyclic group $\mathbb{Z}_2$ under addition modulo 2, or as sign changes $\{-1, 1\}$, forming a multiplicative cyclic group isomorphic to $\mathbb{Z}_2$. The success outcome (action potential) can always be mapped to 1.

Before discussing group representations, we must introduce certain concepts for clarity, particularly given potential confusion from texts emphasizing non-trivial groups while disregarding trivial ones. A group is a closed set equipped with a binary operation adhering to associativity, identity element existence, and inverse element existence (Simon, 1996). A cyclic group is generated by a single element. The digit 1 and identity matrices satisfy definitions of multiplicative and cyclic groups with the simplest structures—trivial yet fundamentally important for later discussions.

In mathematics, the modular group is represented by the projective special linear group $\mathrm{PSL}(2, \mathbb{Z})$, consisting of 2×2 matrices with integer coefficients and determinant 1 (Diamond and Shurman, 2006). To streamline notation in the frequency domain (a projective space), we use the special linear group $\mathrm{SL}(2, \mathbb{Z})$ to denote the modular group without ambiguity:

$$\mathrm{SL}(2,\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{Z}, ad - bc = 1 \right\} \tag{1.6.1}$$

The modular group is generated via fractional linear transformation:

$$A(\tau) = \frac{a\tau + b}{c\tau + d} \tag{1.6.2}$$

with two matrix generators: the translation matrix T (Equation 1.6.3) and inversion matrix S (Equation 1.6.4). The fractional linear transformation maps the modular group from the complex upper half-plane onto itself ($\mathrm{Im}(\tau) > 0$) [FIGURE:2], where τ is a complex number with positive imaginary part (Serre, 1973; Apostol, 2012). Note that matrices A and $-A$ are identical in the modular group, making it a quotient group $\mathrm{SL}(2,\mathbb{Z})/\{I, -I\}$ in projective space, so $S^2 = I$ in Equations 1.6.6.

Beyond regular expression $\{0, 1\}$ and sign expression $\{-1, 1\}$, the $\mathbb{Z}_2$ group can be expressed as a matrix representation using 2×2 matrices. The Pauli matrices form a set of three 2×2 complex matrices $(\sigma_x, \sigma_y, \sigma_z)$:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.6.7-1.6.9}$$

The 2×2 matrix representation of $\mathbb{Z}_2$ includes the identity matrix I (Equation 1.6.5) and either Pauli σ_x (Equation 1.6.7) or σ_z (Equation 1.6.9), which generates the multiplicative cyclic group $\mathbb{Z}_2$. The action potential can be mapped to the 2×2 identity matrix I, while Bernoulli failure maps to σ_x or σ_z. As discussed above, the action potential expressed as the 2×2 identity matrix I alone forms a trivial group. Moreover, the identity matrix I belongs to every principal congruence subgroup of $\mathrm{SL}(2,\mathbb{Z})$, denoted as the gamma congruence subgroup $\Gamma(n)$:

$$\Gamma(n) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{n} \right\} \tag{1.6.10}$$

All elements of $\Gamma(n)$ are matrices congruent to the identity matrix modulo n, where diagonal elements a and $d \equiv 1 \pmod{n}$ and off-diagonal elements b and $c \equiv 0 \pmod{n}$. The principal congruence subgroup $\Gamma(2)$ is the smallest non-trivial congruence subgroup of $\mathrm{SL}(2,\mathbb{Z})$, and the action potential expressed as the 2×2 identity matrix I as a trivial group encapsulates the simplest action form within $\Gamma(2)$ of $\mathrm{SL}(2,\mathbb{Z})$.

Beyond principal congruence subgroups, the notable congruence subgroups $\Gamma_1(n)$ and $\Gamma_0(n)$ hold significance for later classifications:

$$\Gamma_1(n) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \pmod{n} \right\} \tag{1.6.11}$$

$$\Gamma_0(n) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \pmod{n} \right\} \tag{1.6.12}$$

Every congruence subgroup Γ has finite index in $\mathrm{SL}(2,\mathbb{Z})$ (Diamond and Shurman, 2006) and follows a strict inclusion hierarchy: $I \subset \Gamma(2) \subset \Gamma(n) \subset \Gamma_1(n) \subset \Gamma_0(n) \subset \mathrm{SL}(2,\mathbb{Z})$.

1.7 The Definition of Information and the Minimum Uncertainty Principle

In group theory, the 2×2 identity matrix representation of the action potential serves as both a fundamental element and the smallest congruence subgroup of $\mathrm{SL}(2,\mathbb{Z})$. Correspondingly, in neuroscience, the action potential represents the smallest known unit of information carried by neurons. Given that additivity is an essential feature of information, modeling the action potential as a Bernoulli trial provides a solid foundation for quantifying information in neurons and the brain.

1.7.1 Shannon Information and Fisher Information Addressing Shannon's inquiry within communication systems—"How does one measure amount of information?"—requires a precise definition. Shannon's solution for determining "capacity of a communication channel" was information entropy, or Shannon entropy (S), measured in bits (Shannon and Weaver, 1998):

$$S = -\sum_i p_i \log_2(p_i) \tag{1.7.1}$$

For an action potential firing with probability p as a Bernoulli trial outcome with failure rate $1-p$, its Shannon information (S_B) is:

$$S_B = p \log_2\left(\frac{1}{p}\right) + (1-p)\log_2\left(\frac{1}{1-p}\right) \tag{1.7.2}$$

In physics, entropy associates with randomness, disorder, or uncertainty, while information links to accuracy, precision, invariance, or stability. In statistics, precision of a random variable x is defined as inverse variance. The lower variance bound is the Cramér-Rao bound (CRLB) (Cramér, 1946), defined as the inverse of Fisher's information (Rao, 1992). Fisher's information is an additive "intrinsic accuracy" defined by Fisher (1922). For probability mass function

$f(x,p)$ of discrete random variable x with unbiased estimator p, Fisher information $F(p)$ is:

$$F(p) = -E\left[\left(\frac{\partial}{\partial p}\ln f(x,p)\right)^2\right] \tag{1.7.3}$$

where the partial derivative of the natural logarithm of $f(x,p)$ is called the score, and Fisher's information $F(p)$ is the expected value (E) of the score variance.

1.7.2 Neuronal Information and the Minimum Uncertainty Principle

If x is one sample from a Bernoulli trial, the Fisher information carried by x is:

$$F(p) = -E\left[\frac{\partial^2}{\partial p^2}\ln\left(p^x(1-p)^{1-x}\right)\right] = \frac{1}{p(1-p)} \tag{1.7.4}$$

where p is success probability. Therefore, the variance $p(1-p)$ of a Bernoulli trial is the inverse of Fisher's information , meeting the CRLB lower bound and indicating minimal variance or uncertainty. Consequently, Fisher information in a Bernoulli trial equals its precision, signifying that intrinsic accuracy is synonymous with precision.

The concept of Fisher information's intrinsic invariance is more intuitively understood through information geometry (Nielsen, 2022). Similar to how curvature in differential geometry depicts manifold curvature, Fisher information's second-order property (Equation 1.7.3) captures local curvature of the conditional probability space, related to relative entropy or Kullback-Leibler divergence (Gourieroux and Monfort, 1995). In differential geometry, curvature is recognized as an intrinsic invariant, a concept established by Gauss's Theorem Egregium (Gu and Yau, 2008). In summary, neuronal information in Bernoulli processes can be defined and quantified as Fisher information, adhering to the principle of minimum uncertainty.

1.7.3 Relationship Between Fisher Information and Shannon Entropy

1.7.3.1 Isomorphic Relationship

Through appropriate mathematical transformations (such as taking the logarithm of Equation 1.7.4), Fisher information and Shannon entropy (Equation 1.7.2) exhibit an isomorphic relationship for Bernoulli trials. This isomorphism reflects shared behavior across probability space, as the Shannon entropy vector can be expressed as a linear action of the logarithmic Fisher information vector via multiplication by a diagonal probability matrix:

$$\begin{pmatrix} p\log_2\left(\frac{1}{p}\right) \\ (1-p)\log_2\left(\frac{1}{1-p}\right)\end{pmatrix} = \begin{pmatrix} p & 0 \\ 0 & 1-p\end{pmatrix}\begin{pmatrix} \log_2\left(\frac{1}{p}\right) \\ \log_2\left(\frac{1}{1-p}\right)\end{pmatrix} \tag{1.7.5}$$

However, this isomorphism does not imply direct equivalence or interchangeability. Shannon entropy measures overall system uncertainty, while Fisher information focuses on parameter estimation precision. Despite structural similarities, they serve distinct statistical purposes.

1.7.3.2 Dialectical Relationship

The extremal behavior of Shannon entropy and logarithmic Fisher information exhibits opposite trends. For a Bernoulli distribution with parameter p, Shannon entropy achieves its maximum value $\log_2 2$ at $p = 0.5$ (maximum uncertainty) and approaches 0 as $p \to 0$ or 1. Fisher information and its logarithm attain minimum values of 4 and $2 \log_2 2$, respectively, at $p = 0.5$ (least precision), while approaching infinity as $p \to 0$ or 1 (highest precision). Thus, logarithmic Fisher information and Shannon entropy exhibit opposite yet complementary behavior at extreme values, embodying a dialectical unity of opposites that underscores their deep interconnection and complementary roles in describing information systems.

1.7.3.3 Information Geometry and Group Representations

Information geometry studies the geometric structure of statistical models, treating probability distributions as points on a manifold (Nielsen, 2022). The Fisher information metric defines the Riemannian metric on this manifold, providing a way to measure distances between distributions and understand how parameter changes affect distribution shape and information content. In group theory and Lie groups, generators are elements used to construct all other group elements via the exponential map (Greub, 1975; Hall, 2003). In information geometry, this concept broadens to represent transformations or flows on the statistical manifold, where generators correspond to infinitesimal parameter changes causing distribution shifts (Petersen, 1998). The logarithmic isomorphism between Fisher information and Shannon entropy not only highlights structural similarities but also uncovers deeper theoretical connections rooted in information geometry and group theory. Viewing entropy as a representation of transformation generators on the statistical manifold enriches our understanding of the interplay between global uncertainty and local parameter sensitivity.

1.7.3.4 Dialectical Unification Relations

1.7.3.4.1 One-Parameter Transformation Groups and Fisher Geometry in Statistical Manifolds

In Lie theory, a one-parameter subgroup of Lie group G is a smooth homomorphism from the additive group of real numbers $(\mathbb{R}, +)$ to G, written as $\gamma(t) = \exp(tX)$, where X is an element of G's Lie algebra and $t \in \mathbb{R}$ is a real parameter. The exponential map connects the infinitesimal generator X (an element of the tangent space at identity) to group elements through left-invariant vector field flow (Hall, 2003). In information geometry, we can consider a one-parameter transformation group acting on probability space with infinitesimal generator $X = p(1-p)\partial/\partial p$. This vector field encodes directional fluctuations analogous to thermal energy $k_B T$ (with k_B as Boltzmann constant and T as temperature in Kelvin), corresponding to Bernoulli process variance. Fisher

information emerges as the metric dual of X, quantifying local curvature of the statistical manifold, analogous to $1/(k_BT)$ in thermodynamics. The group action on probability space $p \in (0,1)$ is given by $\Phi_t(p) = pe^t/(1-p+pe^t)$, satisfying both group composition properties ($\Phi_s(\Phi_t(p)) = \Phi_{s+t}(p)$) and consistency with infinitesimal generator verification ($\partial\Phi_t/\partial t|_{t=0} = p(1-p)$). This formulation bridges stochastic fluctuations in binary systems to geometric structures, mirroring thermodynamic analogies between statistical variability and thermal noise.

To achieve transparent insight into the relationship between Fisher information and Shannon entropy, we can rewrite Equation 1.7.4 as it relates to one-parameter subgroup form generated by associated Lie algebra elements:

$$F(p) = \frac{1}{p(1-p)} = \exp\left[\ln\left(\frac{1}{p}\right)\right] \cdot \exp\left[\ln\left(\frac{1}{1-p}\right)\right] \tag{1.7.6}$$

In information theory (Alajaji and Chen, 2018), the terms $\ln(1/p)$ and $\ln(1/(1-p))$ are defined as self-information of Bernoulli outcomes. Equation 1.7.6 demonstrates that Fisher information is the product of reciprocal probabilities of spiking and silence, and the product of exponentials of these self-information terms. Crucially, this Bernoulli process identity holds as an algebraic equality, not merely a group-theoretic analogy.

1.7.3.4.2 Shannon Entropy: A Weighted Average of Self-Information Quantifying System-Wide Uncertainty

If the probability matrix in Equation 1.7.5 is expressed as a probability vector, Shannon entropy (S_B) of a Bernoulli trial can be expressed as the dot product of the probability vector and the self-information vector of Fisher information:

$$S_B = (p, 1-p) \cdot \frac{\left(\ln\left(\frac{1}{p}\right), \ln\left(\frac{1}{1-p}\right)\right)}{\ln(2)} = p\log_2\left(\frac{1}{p}\right) + (1-p)\log_2\left(\frac{1}{1-p}\right) \tag{1.7.7}$$

In neuronal information modeled as a Bernoulli process, Shannon entropy serves as a weighted average of self-information, capturing overall uncertainty in neuronal firing patterns. Shannon entropy quantifies average surprise, while Fisher information quantifies parameter estimation precision. This perspective highlights the dialectical unity between Shannon entropy and Fisher information concerning information uncertainty and encoding efficiency. When neural activity is described by independent and identically distributed Bernoulli trials, each trial represents a discrete time point where the neuron may fire a spike ("success") or remain silent ("failure"). Shannon entropy quantifies outcome uncertainty by calculating self-information for all possible spiking events, weighted by their probabilities. The principle that less probable events carry more information because they provide greater "surprise" is central to this understanding. Conversely, Fisher information focuses on the precision of estimating neuronal

firing probability from observational data, measuring the responsiveness of observed spike trains to minute changes in underlying firing probability. While Fisher information serves as a metric for parameter estimate precision, Shannon entropy provides a gauge for total uncertainty conveyed by these estimates.

1.7.3.4.3 Fisher-Shannon Tangling in Bernoulli Processes: Unifying Information and Entropy in Tangled Factor

1.7.3.4.3.1 Unifying Information and Entropy in Tangled Factor

For Bernoulli parameters p (success probability) and q (failure probability) treated formally as independent variables for gradient decomposition (with constraint $p + q = 1$ retained), gradients can be taken with respect to either parameter separately. The formal sum of partial derivatives of Shannon entropy with respect to p and q is $\nabla S_B = A\partial/\partial p + B\partial/\partial q$, where $A = \log_2(1/p) + 1$ and $B = \log_2(1/q) + 1$.

For the Fisher geometry statistical manifold defined by the one-parameter transformation group (Section 1.7.3.4.1), we define self-information vector fields $\ln(1/p)\partial/\partial p$ and $\ln(1/(1-p))\partial/\partial p$ as Lie algebra elements $dA = (A-1)\ln 2\partial/\partial p$ and $dB = (B-1)\ln 2\partial/\partial p$, naturally derived from entropy gradient components through the Fisher information metric. Under the linear constraint $p + q = 1$, which induces $\partial/\partial q = -\partial/\partial p$, these vector fields do not commute, leading to a non-vanishing bracket $[dA, dB]$ (Equation 1.7.8) that quantifies covariance between precision (Fisher information) and uncertainty (entropy). This bracket satisfies anti-commutativity and the Jacobi identity, defining properties of Lie algebras that emerge directly from constraint-induced non-commutativity of dA and dB, reflecting how geometric structures encode functional relationships in information systems. This mathematical rigor aligns with the systems science principle that structure determines function.

$$[dA, dB]f = \left[\ln\left(\frac{1}{p}\right)\frac{\partial}{\partial p}, \ln\left(\frac{1}{1-p}\right)\frac{\partial}{\partial p}\right] f = F(p)S_B \ln 2 \frac{\partial f}{\partial p} \tag{1.7.8}$$

This bracket quantifies a fundamental constraint: firing precision $F(p)$ and uncertainty S_B are tangled such that their product scales the mismatch between spiking and silent states. This coupling defines the Tangled Factor (TF):

$$TF = F(p)S_B \ln 2 = \frac{1}{p(1-p)}\left(p\log_2\frac{1}{p} + (1-p)\log_2\frac{1}{1-p}\right)\ln 2 \tag{1.7.9}$$

TF arises directly from the non-vanishing Lie bracket $[dA, dB]$ (Equation 1.7.8), where the coupling of $F(p)$ (precision) and S_B (uncertainty) is quantified by their product.

1.7.3.4.3.2 Tangled Factor as a Local Geometric Curvature

Figures

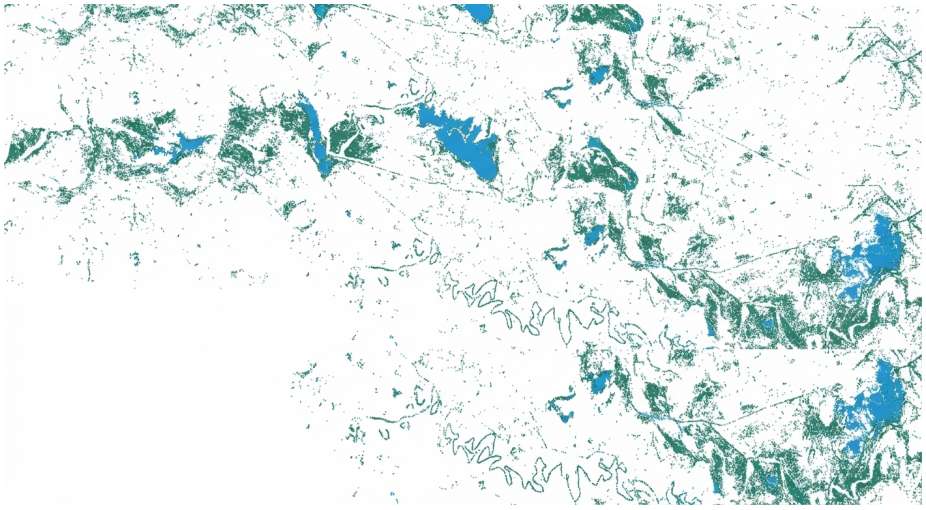

Figure 1: Figure 24