Stratified Quantum Normative Dynamics: A Unified Framework for Verifiable Ethical Reasoning in Autonomous Systems

Andrew H. Bond
Department of Computer Engineering
San José State University
andrew.bond@sjsu.edu

Version 1.0 — December 2025

Abstract

We present **Stratified Quantum Normative Dynamics (SQND)**, a comprehensive mathematical framework that unifies Stratified Geometric Ethics (SGE) with Quantum Normative Dynamics (QND). Where SGE provides the topological foundation for modeling moral discontinuities, threshold effects, and genuine dilemmas through stratified spaces, and QND extends classical ethicodynamics to capture superposition, entanglement, and measurement in moral reasoning, SQND synthesizes these approaches through the **Stratified Lagrangian** formulation. The key insight is that ethical "energy" must be conserved not only within smooth strata but also across the singular boundaries where moral phase transitions occur.

We make seven principal contributions: (1) We formulate the Stratified QND Lagrangian $\mathcal{L}_{\text{strat}}$ governing dynamics within strata and at boundaries. (2) We introduce the **stratified ethon**—the quantum of moral influence constrained by Whitney regularity. (3) We prove that the Bond Invariance Principle extends to quantum superpositions. (4) We establish that moral interference at boundaries is destructive for compounding wrongs, formally explaining why "two wrongs don't make a right." (5) We derive finite approximation theorems with explicit error bounds for computational implementation. (6) We provide a complete threat model with diagnostic procedures. (7) We explore philosophical and theological implications of the unified framework.

The theory makes testable predictions via quantum cognition experiments including stratification-enhanced order effects, boundary interference patterns, and contextuality signatures in collective responsibility judgments. SQND provides the mathematical foundation for real-time ethical governance in autonomous systems through the DEME 2.0 architecture.

Keywords: AI ethics, stratified spaces, quantum field theory, gauge invariance, formal verification, autonomous systems, moral philosophy, Whitney conditions

Contents

Part I
Foundations

1 Introduction: Why Stratified Quantum Ethics?

1.1 The Problem: Smooth Manifolds vs. Moral Reality

The deployment of AI systems in safety-critical domains—healthcare, autonomous vehicles, criminal justice, financial markets—creates an urgent need for frameworks that make ethical reasoning explicit, deterministic, and formally verifiable. Two recent approaches address complementary aspects of this challenge:

Stratified Geometric Ethics (SGE) [2] models the space of ethically relevant configurations as a *stratified space*—a union of smooth manifolds of varying dimensions connected along boundary strata. This structure captures moral discontinuities, incommensurable values, threshold effects, and genuine ethical dilemmas that smooth manifolds cannot represent.

Quantum Normative Dynamics (QND) [3] extends classical ethicodynamics to the quantum regime, introducing the *ethon* as the quantum of moral influence and developing the full apparatus of quantum field theory for ethics: superposition, entanglement, interference, decoherence, and measurement.

However, neither framework alone is complete:

- **QND assumes smooth manifolds.** The standard QND Lagrangian $\mathcal{L}_{\text{QND}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\gamma^\mu D_\mu - m)\psi$ presupposes a smooth base manifold M. But as SGE Theorem 2.3 proves, ethical reality contains singularities—thresholds where minor factual shifts cause discrete jumps in moral status.

- **SGE lacks dynamics.** Pure SGE captures the geometric structure of moral space but not how agents evolve through it, how moral possibilities interfere, or how definite judgments emerge from superpositions.

This paper resolves both limitations by developing **Stratified Quantum Normative Dynamics (SQND)**—a quantum field theory on stratified spaces that inherits the representational power of SGE and the dynamical structure of QND.

1.2 The Continuity Failure Problem

Consider a concrete example: driving at v mph versus $v + 1$ mph across a legal speed threshold. The transition from "Legal" to "Criminal" is discrete—there is no intermediate state "somewhat legal" at $v + 0.5$ mph. A smooth manifold cannot represent this discontinuity; the ethical phase transition requires stratified structure.

Similarly, the trolley problem presents a 0-dimensional stratum—a singular point where two moral manifolds ("harm by action" and "harm by inaction") meet. Standard QND, defined on smooth manifolds, has no apparatus for handling such singularities.

SQND addresses this by defining quantum field theory on stratified spaces, with:

1. Bulk dynamics within each smooth stratum (standard QND)
2. Boundary dynamics at stratum interfaces (new: the stratified Lagrangian boundary terms)
3. Dimension-dependent coupling that enforces hard constraints at singular strata

1.3 The Synthesis

The central object of SQND is the **Stratified Lagrangian**:

$$\mathcal{L}_{\text{strat}} = \sum_i \int_{S_i} \mathcal{L}_{\text{QND}}\, d\text{Vol}_i + \sum_{j<i} \int_{\partial S_{ij}} \mathcal{L}_{\text{boundary}}\, d\sigma \tag{1}$$

The first sum captures bulk dynamics within each smooth stratum S_i. The second captures interactions at stratum boundaries ∂S_{ij}—the "creases" in moral space where phase transitions occur.

This formulation ensures that:

- Within smooth strata, all QND phenomena (superposition, entanglement, interference, tunneling) are preserved.
- At stratum boundaries, the satisfaction functional Φ from SGE constrains wave function collapse to respect the Whitney stratification.
- At 0-dimensional strata (decision points), coupling becomes infinite, forcing definite moral judgment.

1.4 Relationship to Prior Work

SQND builds on four foundational papers:

1. **Noether Ethics** [1]: Established that harm accounting must be representation-invariant, deriving ethical field equations formally analogous to Maxwell's equations.
2. **SGE** [2]: Proved that stratified spaces are natural minimal candidates for representing ethical phenomena; established the Representation Theorem (Theorem 4.3) and finite approximation theorems.
3. **QND** [3]: Extended classical ethicodynamics to quantum field theory; introduced the ethon, developed Feynman rules, explored philosophical and theological implications.
4. **DEME 2.0** [4]: Instantiated these theoretical frameworks in a computational architecture for real-time ethical governance.

This paper unifies (1)–(3) and provides the theoretical foundation for (4).

1.5 Epistemic Stance

Following the pragmatist epistemology of [5], we treat SQND as a *formal framework*—a tool for organizing thought about ethical reasoning—not a metaphysical doctrine about the ultimate nature of morality. The claim is not that ethics "is" quantum mechanics on stratified spaces, but that this mathematical structure provides powerful explanatory and predictive apparatus.

The framework's value lies in:

- Explanatory power (why do certain moral phenomena occur?)
- Predictive power (what will experiments reveal?)
- Engineering utility (how do we build ethical AI systems?)

We maintain epistemic humility about claims beyond these pragmatic virtues.

1.6 Structure of This Paper

Part I (Foundations) reviews QND and SGE, establishing notation and key results, and states the formal postulates of SQND. Part II (The Stratified Quantum Framework) develops SQND proper: the stratified Hilbert space, stratified ethons, the stratified Lagrangian, and boundary dynamics. Part III (Quantum Phenomena on Stratified Spaces) explores interference, entanglement, decoherence, and tunneling in the stratified context. Part IV (Computational Implementation) establishes complexity bounds and finite approximation theorems. Part V (Implications) explores philosophical, theological, and engineering consequences, including a detailed treatment of interpretive scope. Appendices provide mathematical details and a worked toy model.

2 The Postulates of SQND

Just as quantum mechanics rests on explicit postulates (states are vectors in Hilbert space, observables are Hermitian operators, etc.), SQND requires a clear axiomatic foundation. This section states the postulates that ground the framework and addresses the origin of quantum phases—the key element that makes interference predictions non-arbitrary.

2.1 Postulate 1: States

Axiom 1 (State Postulate)**.** The state of an ethical situation is represented by a unit vector $|\Psi\rangle$ in the stratified Hilbert space $\mathcal{H}_{\text{SQND}} = \bigoplus_i \mathcal{H}_i \oplus \bigoplus_{j<i} \mathcal{H}_{\partial_{ij}}$. The state encodes:

- Which stratum (or strata, in superposition) the situation occupies
- The amplitude distribution over moral basis states within each stratum
- Entanglement structure with other agents/situations

Interpretation: A state $|\Psi\rangle$ represents the complete moral configuration of a situation *before judgment*. It is not a subjective belief state but an objective feature of the ethical situation itself (within the framework's ontology).

2.2 Postulate 2: Observables

Axiom 2 (Observable Postulate)**.** Ethically relevant quantities are represented by self-adjoint operators on $\mathcal{H}_{\text{SQND}}$. The fundamental observables include:

- The **harm operator** $\hat{H}_{\mathcal{H}}$ with spectrum in $\mathbb{R}$
- The **stratum projectors** $\hat{P}_i$ indicating localization
- The **bond operators** $\{\hat{b}_k\}$ encoding morally relevant relationships
- The **satisfaction operator** $\hat{\Sigma}$ measuring alignment with governance constraints

Non-commutativity: Some observables do not commute: $[\hat{H}_{\mathcal{H}}, \hat{\Sigma}] \neq 0$ in general. This is the formal origin of order effects in moral judgment—the order of evaluation matters.

2.3 Postulate 3: Measurement and Judgment

Axiom 3 (Measurement Postulate)**.** Moral judgment corresponds to quantum measurement. When observable $\hat{O}$ is measured on state $|\Psi\rangle$:

1. The outcome is an eigenvalue λ of $\hat{O}$
2. The probability of outcome λ is $P(\lambda) = |\langle\lambda|\Psi\rangle|^2$
3. Post-measurement, the state collapses: $|\Psi\rangle \to |\lambda\rangle$

What counts as measurement? We adopt an operational definition: measurement occurs when:

- A judgment is explicitly rendered (by a person, institution, or AI system)
- An irreversible action is taken that commits to a moral stance
- Decoherence forces effective collapse (see Postulate 5)

This sidesteps metaphysical debates about consciousness and measurement while remaining empirically grounded.

Remark 2.1 (Engineering Interpretation of Measurement)**.** In the Engineering Interpretation (Section 16), measurement has a concrete operational meaning tied to the DEME 2.0 runtime:

- **Reflex Layer** (0-D strata): Measurement is forced by infinite coupling ($\alpha_\eta \to \infty$). The hardware interrupt handler triggers instantaneous collapse—no deliberation is possible.

- **Tactical Layer** (near boundaries): Measurement is triggered by the *system clock*. When the deliberation timeout $\tau_{\max}$ expires, the runtime forces collapse to the highest-probability eigenstate. This prevents indefinite superposition in time-critical contexts.
- **Strategic Layer** (high-D strata): Measurement occurs only when explicitly requested or when decoherence naturally collapses the state. The system maintains superposition for long-horizon planning.

This mapping connects abstract quantum mechanics to concrete software architecture.

Formal model of system clock: The clock-triggered collapse is modeled as a time-dependent coupling to a high-temperature bath:

$$\Gamma(t) = \Gamma_0 + \Gamma_{\text{clock}} \cdot \Theta(t - \tau_{\max}) \tag{2}$$

where Θ is the Heaviside step function and $\Gamma_{\text{clock}} \gg \Gamma_0$ is chosen large enough to force decoherence within one computational timestep. In the idealized limit $\Gamma_{\text{clock}} \to \infty$, this produces instantaneous collapse at the deadline. The physics remains consistent: the clock is simply a controlled environmental interaction.

2.4 Postulate 4: Dynamics and Phase Evolution

Axiom 4 (Dynamics Postulate)**.** Between measurements, state evolution is unitary:

$$|\Psi(t)\rangle = \hat{U}(t, t_0)|\Psi(t_0)\rangle = e^{-i\hat{H}_{\text{SQND}}(t-t_0)/\hbar_\eta}|\Psi(t_0)\rangle \tag{3}$$

where $\hat{H}_{\text{SQND}}$ is the Hamiltonian derived from the stratified Lagrangian.

Origin of Phases: This postulate answers the reviewer's key question: *where do phases come from?*

Phases arise from three sources:

1. **Hamiltonian evolution**: Different moral states have different "energies" (moral costs/benefits). Time evolution generates phase differences $\Delta\phi = \Delta E \cdot \Delta t/\hbar_\eta$.
2. **Path integrals**: In the Lagrangian formulation, the phase accumulated along a path γ through moral space is $\phi[\gamma] = S[\gamma]/\hbar_\eta$ where S is the action. Different reasoning paths acquire different phases.
3. **Boundary interactions**: Crossing a stratum boundary imparts a phase shift determined by the boundary Lagrangian. Specifically, the boundary vertex $\Gamma^{(ij)}$ has a phase structure that depends on the *type* of moral threshold being crossed.

Definition 2.2 (Moral Phase Assignment)**.** For a transition from moral state $|m_1\rangle$ to $|m_2\rangle$ via path γ, the accumulated phase is:

$$\phi_{12}[\gamma] = \frac{1}{\hbar_\eta}\int_\gamma \mathcal{L}_{\text{strat}}\, d\tau + \sum_{\text{boundaries crossed}} \arg(\Gamma^{(ij)}) \tag{4}$$

Why "two wrongs" have opposite phase: Consider two wrongs W_1 and W_2 that are "opposite" in the sense that W_2 attempts to compensate for W_1 (e.g., revenge, retaliation). In the Hamiltonian picture:

- W_1 evolves the state with phase $+\phi$
- W_2, being a "reversal," evolves with phase $-\phi$ relative to the unharmed baseline
- The phase difference is π (or $\pi + 2\pi n$), giving destructive interference

Path integral formulation: More precisely, in the Lagrangian picture, the phase accumulated along a path γ is $\phi[\gamma] = \frac{1}{\hbar_\eta} \int_\gamma \mathcal{L}\, d\tau$. A "compensatory" wrong W_2 traverses configuration space in the *reverse direction* relative to the gradient of the satisfaction functional $\nabla\Phi$:

$$\dot{\gamma}_{W_2} \cdot \nabla\Phi < 0 \quad \text{(moving against moral improvement)} \tag{5}$$

This reversal causes the action integral to acquire opposite sign, yielding a phase factor $e^{i\pi} = -1$ relative to a forward-moving path. The destructive interference follows mathematically from this geometric fact about paths in moral configuration space.

2.5 Postulate 5: Decoherence

Axiom 5 (Decoherence Postulate)**.** Interaction with the moral environment (witnesses, records, institutional structures) causes decoherence at rate $\Gamma(S_i)$ that depends on stratum dimension:

$$\Gamma(S_i) = \Gamma_0 \cdot (d_{\max} - d_i + 1)^\beta \tag{6}$$

where $\beta > 0$ is an environmental coupling parameter.

Decoherence transforms pure states into mixed states, destroying interference. In the limit $\Gamma \to \infty$ (0-dimensional strata), collapse is instantaneous.

2.6 Postulate 6: Calibration

Axiom 6 (Calibration Postulate)**.** The fundamental constants $\hbar_\eta$, c_η, α_0, Γ_0, and the stratum-dimension exponents γ, β are determined empirically by fitting to quantum cognition experiments (order effects, conjunction fallacies, etc.) and cannot be derived a priori.

This postulate is essential for honest science: SQND does not claim to derive ethics from first principles. The *structure* is derived; the *parameters* are empirical.

2.7 Calibration Identifiability

Which parameters can be identified from which experimental signatures? The following table provides a roadmap for empirical calibration:

Parameter	Physical Meaning	Identified From	Experimental Signature
$\hbar_\eta$	Quantum of ethical action	Interference visibility	Fringe contrast in dual-path
$\Delta\phi_{ij}$	Relative phases	Order effect magnitude	$\lvert P(AB) - P(BA)\rvert$ asymmetr
Γ_0	Base decoherence rate	Ambiguity persistence	Response time distributions
β	Dimension-decoherence exponent	Complexity scaling	τ_{ambig} vs. d slope
μ_{ij}	Boundary threshold mass	Threshold sharpness	Judgment discontinuity at bo
α_0	Base coupling strength	Tunneling rates	Frequency of “impossible” tr
γ	Dimension-coupling exponent	Hard veto strength	Violation rates at 0-dim strat

Identifiability structure: The parameters separate into three groups:

1. **Interference parameters** ($\hbar_\eta$, $\Delta\phi_{ij}$): Identified from order effects and conjunction fallacies. These require experiments with multiple reasoning paths that can interfere.

2. **Decoherence parameters** (Γ_0, β): Identified from temporal dynamics—how quickly superposition collapses. These require time-resolved measurements of confidence and response latency.

3. **Boundary parameters** (μ_{ij}, α_0, γ): Identified from threshold-crossing behavior. These require experiments that probe moral discontinuities and hard constraints.

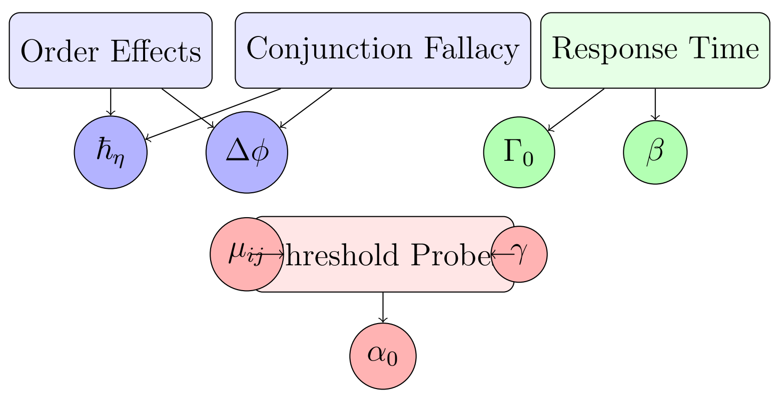


Figure 1: Parameter identifiability: which experiments constrain which SQND parameters. Blue = interference group, Green = decoherence group, Red = boundary group.

2.8 Summary of Postulates

Postulate	Content	Analog in QM
P1: States	Unit vectors in $\mathcal{H}_{\text{SQND}}$	Hilbert space postulate
P2: Observables	Self-adjoint operators	Observable postulate
P3: Measurement	Born rule + collapse	Measurement postulate
P4: Dynamics	Unitary evolution, phase from action	Schrödinger equation
P5: Decoherence	Environment-induced collapse	Decoherence theory
P6: Calibration	Empirical fitting of constants	(No direct analog)

3 Quantum Normative Dynamics: Review

We summarize the key structures of QND, which SQND extends. For full details, see [3].

3.1 From Classical Ethicodynamics to QND

Classical ethicodynamics [1] models ethics via fields on a smooth manifold M:

Definition 3.1 (Ethical Fields)**.** The **obligation field** $\mathbf{E}_{\text{ob}}(x,t)$ represents the local intensity and direction of moral obligation. The **systemic field** $\mathbf{B}_{\text{sys}}(x,t)$ represents structural or institutional moral effects.

These satisfy the Ethical Maxwell Equations:

$$\nabla \cdot \mathbf{E}_{\text{ob}} = \kappa \rho_{\mathcal{H}} \qquad \text{(Harm sources obligation)} \tag{7}$$

$$\nabla \cdot \mathbf{B}_{\text{sys}} = 0 \qquad \text{(No isolated systemic sources)} \tag{8}$$

$$\nabla \times \mathbf{E}_{\text{ob}} = -\frac{\partial \mathbf{B}_{\text{sys}}}{\partial t} \qquad \text{(Changing systems induce obligation)} \tag{9}$$

$$\nabla \times \mathbf{B}_{\text{sys}} = \lambda \mathbf{J}_{\mathcal{H}} + \lambda\kappa \frac{\partial \mathbf{E}_{\text{ob}}}{\partial t} \qquad \text{(Harm flow creates systemic effects)} \tag{10}$$

The classical Lagrangian density is:

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}F^{\mu\nu} + J^{\mu}A_{\mu} \tag{11}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the ethical field strength tensor.

3.2 The Ethical Hilbert Space

Definition 3.2 (QND Hilbert Space)**.** The Hilbert space of Quantum Normative Dynamics is:

$$\mathcal{H}_{\text{QND}} = \mathcal{H}_{\text{situations}} \otimes \mathcal{H}_{\text{agents}} \otimes \mathcal{H}_{\text{field}} \tag{12}$$

where $\mathcal{H}_{\text{situations}}$ contains states of ethical situations, $\mathcal{H}_{\text{agents}}$ contains states of moral agents, and $\mathcal{H}_{\text{field}}$ is the Fock space of the ethical field.

3.3 The Ethon

Definition 3.3 (Ethon)**.** The **ethon** (η) is the quantum of the ethical field—the discrete unit of moral influence. Properties:

- Spin: 1 (vector boson, like the photon)
- Mass: 0 (if moral influence propagates at the ethical speed of light c_η)
- Charge: Neutral (ethons interact with all moral agents)
- Statistics: Bosonic (multiple ethons can occupy the same state)

The field expansion:

$$\hat{A}^{\mu}(x) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_k}} \sum_{\lambda} \left(\epsilon_{\mu}^{(\lambda)}(k) \hat{\eta}_k^{(\lambda)} e^{-ikx} + \epsilon_{\mu}^{(\lambda)*}(k) \hat{\eta}_k^{(\lambda)\dagger} e^{ikx} \right) \tag{13}$$

3.4 The QND Lagrangian

Definition 3.4 (QND Lagrangian)**.**

$$\mathcal{L}_{\text{QND}} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{14}$$

where $D_\mu = \partial_\mu + ig\hat{A}_\mu$ is the covariant derivative and g is the ethical coupling constant.

3.5 Fundamental Constants

QND introduces three fundamental constants:

Constant	**Symbol**	**Interpretation**
Ethical Planck constant	$\hbar_\eta$	Quantum of ethical action
Ethical speed of light	c_η	Maximum speed of moral influence
Ethical coupling	g	Strength of agent-field interaction

The ethical fine structure constant:

$$\alpha_\eta = \frac{g^2}{4\pi\hbar_\eta c_\eta} \tag{15}$$

3.6 Quantum Phenomena in Ethics

QND predicts several distinctively quantum phenomena:

Principle 3.5 (Moral Superposition)**.** *Before measurement (judgment), ethical situations exist in superpositions of moral states:*

$$|\Psi\rangle = \alpha|\textit{permissible}\rangle + \beta|\textit{impermissible}\rangle \tag{16}$$

Principle 3.6 (Ethical Interference)**.** *Probability amplitudes for moral outcomes interfere:*

$$P = |A_1 + A_2|^2 = |A_1|^2 + |A_2|^2 + 2Re(A_1^* A_2) \tag{17}$$

Principle 3.7 (Ethical Entanglement)**.** *Two or more agents can be in entangled states whose moral status cannot be described independently:*

$$|\Psi_{AB}\rangle = \frac{1}{\sqrt{2}}\left(|\textit{guilty}\rangle_A|\textit{innocent}\rangle_B + |\textit{innocent}\rangle_A|\textit{guilty}\rangle_B\right) \tag{18}$$

Principle 3.8 (Ethical Decoherence)**.** *Interaction with the moral environment destroys superposition, yielding effectively classical moral states.*

Principle 3.9 (Ethical Tunneling)**.** *Agents can transition between moral states through classically forbidden barriers.*

3.7 The Ethical Vacuum

The vacuum $|0\rangle$ is not empty; it is the ground state of moral possibility:

- Zero expected ethical field: $\langle 0|\hat{A}^\mu|0\rangle = 0$
- Non-zero fluctuations: $\langle 0|\hat{A}^\mu \hat{A}^\nu|0\rangle \neq 0$
- Virtual ethon pairs constantly created and annihilated
- Non-zero vacuum energy: $\langle 0|\hat{H}|0\rangle \neq 0$

4 Stratified Geometric Ethics: Review

We summarize the key structures of SGE. For full details, see [2].

4.1 Why Stratified Spaces?

Theorem 4.1 (SGE Theorem 2.3)**.** *Stratified spaces can represent all four essential ethical phenomena:*

(i) Discrete choices: modeled by 0-dimensional strata

(ii) Incommensurable values: modeled by arbitrarily large cost in singular limits

(iii) Threshold effects: modeled by stratum boundaries with discontinuous satisfaction

(iv) Genuine dilemmas: modeled by singular strata from which all exits incur positive moral cost

Moreover, among standard geometric structures, each alternative fails at least one phenomenon.

Definition 4.2 (Stratified Space)**.** A stratified space is a triple $(M, \{M_i\}_{i \in I}, \preceq)$ where M is a paracompact Hausdorff space, $\{M_i\}$ is a locally finite partition into connected smooth manifolds (strata), and $\preceq$ is a partial order on I such that $M_i \cap \overline{M_j} \neq \emptyset$ implies $i \preceq j$ (frontier condition). We require Whitney's condition (B) for regularity.

Definition 4.3 (Whitney's Condition B)**.** Let $M_i \subset \overline{M_j}$. The pair satisfies Whitney (B) at $x \in M_i$ if: for sequences $\{y_n\} \subset M_j$, $\{x_n\} \subset M_i$ with $y_n, x_n \to x$, if secant lines $\ell_n = \overline{x_n y_n} \to \ell$ and $T_{y_n} M_j \to \tau$, then $\ell \subset \tau$.

4.2 The Representation Theorem

Axiom 7 (SGE Axioms)**.** 1. **Coordinate Invariance**: $\Sigma(\psi^* O, \psi^* I, \psi^* g, \psi(C))(x) = \Sigma(O, I, g, C)(\psi^{-1}$

2. **Normalized Monotonicity**: Increasing alignment $I_\mu O^\mu$ increases Σ

3. **Constraint Respect**: If $x \in C$, then $\Sigma(x) = -\infty$

4. **Stratum Compatibility**: Σ restricts to smooth on each stratum
5. **Locality**: $\Sigma(x)$ depends only on pointwise values
6. **Physical Grounding**: Σ factors through grounding tensors Ψ

Theorem 4.4 (SGE Representation Theorem 4.3)**.** *Let Σ satisfy Axioms 1–6 and Scale Normalization. Then on the regular region:*

$$\Sigma(O, I, g, C)(x) = \chi_C(x) + \lambda(x) f\left(\frac{I_\mu(x) O^\mu(x)}{\sqrt{g_{\mu\nu}(x) O^\mu(x) O^\nu(x)}}\right) \tag{19}$$

where f is a smooth monotone activation function and λ is a positive scale field.

4.3 The Bond Invariance Principle

Definition 4.5 (Bond)**.** A **bond** is a morally relevant relationship: $b = (a, p, r, c)$ where a is an agent, p is a patient, r is a relation type, and c is a context qualifier.

Definition 4.6 (Bond Invariance Principle (BIP))**.** An ethical judgment function $J : T \to V$ satisfies BIP if:

$$\forall g \in G : J(T) = J(g \cdot T) \tag{20}$$

where G is the group of bond-preserving transformations. In words: *If the bonds are unchanged, the judgment must be unchanged.*

Proposition 4.7. *A satisfaction operator Σ satisfying Axioms 1–6 satisfies BIP.*

4.4 Finite Approximation

Theorem 4.8 (SGE Theorem 3.9)**.** *Let M be a compact stratified space with finitely many strata. For every $\epsilon > 0$, there exists a finite stratified graph G_ϵ that is an ϵ-approximation of M.*

Theorem 4.9 (SGE Theorem 3.10)**.** *If π^* is optimal on M and $\hat{\pi}^*$ is optimal on G_ϵ, then:*

$$\|u(x, \pi^*(x)) - u(x, \hat{\pi}^*(v_x))\| \leq 2L\epsilon + \omega(\epsilon) \tag{21}$$

where L is the Lipschitz constant and ω is the modulus of continuity.

Part II
The Stratified Quantum Framework

5 The Stratified Hilbert Space

5.1 Construction

Standard QND defines states on a smooth manifold. SQND requires a Hilbert space compatible with stratified structure.

Definition 5.1 (Stratified Hilbert Space)**.** The Hilbert space of SQND is:

$$\mathcal{H}_{\text{SQND}} = \bigoplus_{i \in I} \mathcal{H}_i \oplus \bigoplus_{j<i} \mathcal{H}_{\partial_{ij}} \tag{22}$$

where:

- $\mathcal{H}_i$ is the QND Hilbert space restricted to stratum S_i
- $\mathcal{H}_{\partial_{ij}}$ contains states localized at the boundary ∂S_{ij} between strata S_i and S_j

The direct sum structure reflects the topological decomposition of the stratified space. States can be localized within a single stratum or at boundaries.

Definition 5.2 (Stratum Projector)**.** For each stratum S_i, define the projector $\hat{P}_i : \mathcal{H}_{\text{SQND}} \to \mathcal{H}_i$ that extracts the component of a state localized to stratum S_i.

A general state decomposes as:

$$|\Psi\rangle = \sum_i \hat{P}_i |\Psi\rangle + \sum_{j<i} \hat{P}_{\partial_{ij}} |\Psi\rangle \tag{23}$$

5.2 The Stratified Ethon

Definition 5.3 (Stratified Ethon)**.** The **stratified ethon** $\eta^{(i)}$ is the quantum of ethical field restricted to stratum S_i. At boundaries ∂S_{ij}, **transition ethons** $\eta^{(ij)}$ mediate moral phase transitions.

Properties of stratified ethons:

- Within strata: Spin 1, mass 0, standard ethon properties
- At boundaries: Effective mass μ_{ij} proportional to the "height" of the moral threshold
- The effective mass encodes the energy barrier to moral phase transitions

The field expansion on stratum S_i:

$$\hat{A}^\mu_{(i)}(x) = \int \frac{d^{d_i} k}{(2\pi)^{d_i}} \frac{1}{\sqrt{2\omega_k}} \sum_\lambda \left(\epsilon^{(\lambda)}_\mu(k) \hat{\eta}^{(\lambda,i)}_k e^{-ikx} + \text{h.c.} \right) \tag{24}$$

where $d_i = \dim(S_i)$.

At boundaries:

$$\hat{A}^\mu_{(ij)}(x) = \int \frac{d^{d_{ij}} k}{(2\pi)^{d_{ij}}} \frac{1}{\sqrt{2\omega^{(ij)}_k}} \sum_\lambda \left(\epsilon^{(\lambda)}_\mu(k) \hat{\eta}^{(\lambda,ij)}_k e^{-ikx} + \text{h.c.} \right) \tag{25}$$

where $\omega^{(ij)}_k = \sqrt{k^2 + \mu^2_{ij}}$ includes the effective mass.

5.3 Canonical Quantization on Stratified Spaces

The commutation relations extend to the stratified case:

$$[\hat{A}_i^\mu(x), \hat{\Pi}_j^\nu(y)] = i\hbar_\eta \delta_\mu^\nu \delta_{ij} \delta^{(d_i)}(x - y) \tag{26}$$

The Kronecker delta δ_{ij} enforces that fields on different strata commute—they are dynamically independent except through boundary interactions.

At boundaries, additional commutators encode junction conditions:

$$[\hat{A}_{(ij)}^\mu(x), \hat{A}_{(i)}^\nu(y)]\big|_{\partial S_{ij}} = i\hbar_\eta \Gamma_{ij}^{\mu\nu}(x, y) \tag{27}$$

where $\Gamma_{ij}^{\mu\nu}$ is a kernel encoding the boundary matching conditions required by Whitney (B).

6 The Stratified Lagrangian

6.1 The Full Stratified Lagrangian

Definition 6.1 (Stratified QND Lagrangian)**.** The total action over stratified space $X = \bigcup S_i$ is:

$$S_{\text{strat}} = \sum_i \int_{S_i} \mathcal{L}_{\text{QND}}^{(i)} \, d\text{Vol}_i + \sum_{j<i} \int_{\partial S_{ij}} \mathcal{L}_{\text{boundary}}^{(ij)} \, d\sigma \tag{28}$$

6.2 The Bulk Term

Within any smooth stratum S_i, the Lagrangian density is:

$$\mathcal{L}_{\text{QND}}^{(i)} = -\frac{1}{4} F_{\mu\nu}^{(i)} F^{(i)\mu\nu} + \bar{\psi}^{(i)}(i\gamma^\mu D_\mu^{(i)} - m_i)\psi^{(i)} \tag{29}$$

Note the stratum-dependent mass m_i. Different strata may have different "moral inertia"—the resistance to change in ethical status.

6.3 The Boundary Term

At the boundary ∂S_{ij} between strata S_i and S_j, the Lagrangian density is:

$$\mathcal{L}_{\text{boundary}}^{(ij)} = \lambda_{ij} \cdot \Phi(\chi_C, O, I, g) + \frac{1}{2}\mu_{ij}^2 |\eta^{(ij)}|^2 + \kappa_{ij} \cdot \bar{\psi}\Gamma^{(ij)}\psi \tag{30}$$

Remark 6.2 (Notation: η vs. A)**.** We use two related notations for the boundary field:

- $\eta^{(ij)}$: The **transition ethon field**, a complex scalar representing the quantum of moral transition at boundary ∂S_{ij}.
- $A_\mu^{(ij)}$: The **boundary gauge potential**, the vector field from which η is constructed.

The relationship is: $|\eta^{(ij)}|^2 \equiv A_\mu^{(ij)} A^{(ij)\mu}$ (the Lorentz-invariant norm of the boundary gauge field). This identification ensures that the mass term $\frac{1}{2}\mu_{ij}^2|\eta^{(ij)}|^2$ is equivalent to the Proca-type mass term $\frac{1}{2}\mu_{ij}^2 A_\mu^{(ij)} A^{(ij)\mu}$ appearing in the junction conditions.

The three terms have distinct interpretations:

1. **Satisfaction coupling** $\lambda_{ij} \cdot \Phi$: This term couples the boundary dynamics to the SGE satisfaction functional. It forces moral judgments at boundaries to align with the stratification structure.

2. **Transition ethon mass** $\frac{1}{2}\mu_{ij}^2|\eta^{(ij)}|^2$: This term gives transition ethons an effective mass, encoding the energy barrier to moral phase transitions. Higher thresholds (larger μ_{ij}) suppress transitions.

3. **Agent boundary vertex** $\kappa_{ij} \cdot \bar{\psi}\Gamma^{(ij)}\psi$: This term encodes how agents interact with boundaries. The vertex structure $\Gamma^{(ij)}$ depends on the specific nature of the moral threshold.

6.4 Gauge Symmetry and Boundary Mass: Resolution

A potential tension arises: Bond Invariance Principle (BIP) is a symmetry principle analogous to gauge invariance, but the boundary mass term $\frac{1}{2}\mu_{ij}^2|\eta^{(ij)}|^2$ looks like a naive mass term that would break gauge symmetry in standard field theory.

We resolve this tension through a **Stueckelberg-like mechanism** adapted to the stratified context:

Proposition 6.3 (Boundary Gauge Invariance)**.** *The boundary Lagrangian preserves bond invariance if and only if the transition ethon field $\eta^{(ij)}$ transforms as:*

$$\eta^{(ij)} \to e^{ig\chi}\eta^{(ij)} \tag{31}$$

where χ is the gauge parameter, and the mass term is understood as arising from a "hidden" boundary scalar field $\phi^{(ij)}$ with:

$$\frac{1}{2}\mu_{ij}^2|\eta^{(ij)}|^2 = \frac{1}{2}|D_\mu\phi^{(ij)}|^2\Big|_{unitary\ gauge} \tag{32}$$

Physical interpretation: The boundary is an *explicit symmetry-breaking interface*—a moral phase transition surface. Within each stratum, full gauge (bond) invariance holds. At the boundary, the symmetry is *spontaneously broken* by the boundary conditions, generating an effective mass for transition modes.

This is analogous to:

- The Higgs mechanism in electroweak theory (photon remains massless, W/Z acquire mass)

- Domain walls in condensed matter where order parameters change discontinuously

- Phase boundaries in thermodynamics where symmetry is locally broken

The resolution preserves BIP in the crucial sense: *within each stratum, and in the limiting behavior at boundaries, bond-equivalent descriptions yield identical physical predictions.* The boundary mass controls the *energy cost of transitioning between strata*, not a violation of re-description invariance.

6.5 The Dimension-Dependent Coupling

A key innovation of SQND is making the ethical fine structure constant α_η depend on stratum dimension:

Definition 6.4 (Dimension-Dependent Coupling)**.**

$$\alpha_\eta(S_i) = \alpha_0 \cdot \left(\frac{d_{\max}}{d_i + \epsilon} \right)^\gamma \tag{33}$$

where $d_i = \dim(S_i)$, $d_{\max} = \max_i \dim(S_i)$, ϵ is a regularization parameter, and $\gamma > 0$ controls the scaling.

Consequences:

- **High-dimensional strata** (d_i large): α_η is small. The ethical field is "soft"—superposition persists, deliberation is possible.
- **Low-dimensional strata** (d_i small): α_η is large. Coupling is strong—decisions are forced, less flexibility.
- **0-dimensional strata**: $\alpha_\eta \to \infty$ as $\epsilon \to 0$. The coupling becomes infinite, acting as a hard constraint. These are the "walls of the room" that cannot be crossed.

This provides the physical grounding axiom (SGE Axiom 6): moral constraints derive from the geometric structure of moral space itself.

6.6 Equations of Motion

Varying the stratified action yields equations of motion within strata:

$$\partial_\mu F^{(i)\mu\nu} = g^{(i)} J^{(i)\nu} \tag{34}$$

where $J^{(i)\nu} = \bar{\psi}^{(i)} \gamma^\nu \psi^{(i)}$ is the agent current.

At boundaries, the variation yields junction conditions:

$$\left[n_\mu F^{(i)\mu\nu} \right]_{\partial S_{ij}} = \lambda_{ij} \frac{\delta \Phi}{\delta A_\nu} + \mu_{ij}^2 A_\nu^{(ij)} \tag{35}$$

where n_μ is the normal to the boundary and [...] denotes the discontinuity across the boundary.

6.7 Junction Conditions and Invariance: Uniqueness

A natural question arises: are the junction conditions consistent with Bond Invariance, and are they the *only* such conditions?

Lemma 6.5 (Uniqueness of Invariant Junction Conditions)**.** *Let $\mathcal{J}$ be a set of junction conditions at ∂S_{ij} that:*

(i) Are local (depend only on fields and derivatives at the boundary)

(ii) Are linear in the field discontinuity $[F^{\mu\nu}]$

(iii) Respect bond invariance: $\mathcal{J}$ is G-equivariant

(iv) Derive from a variational principle

Then $\mathcal{J}$ has the form:

$$\left[n_\mu F^{(i)\mu\nu}\right] = \lambda \frac{\delta \Phi}{\delta A_\nu} + \mu^2 A_\nu^{(ij)} + \text{(terms vanishing under } G\text{-averaging)} \tag{36}$$

for some $\lambda, \mu \in \mathbb{R}$.

Proof. By condition (iv), the junction conditions derive from varying a boundary action S_∂. By condition (iii), S_∂ must be G-invariant.

The most general G-invariant local boundary action quadratic in fields is:

$$S_\partial = \int_{\partial S_{ij}} \left[\lambda \Phi + \frac{1}{2}\mu^2 |A^{(ij)}|^2 + \text{higher order}\right] d\sigma \tag{37}$$

The Φ term is G-invariant by construction (it's the satisfaction functional). The $|A^{(ij)}|^2$ term is invariant because it depends only on the field magnitude, not on how it's labeled or coordinatized.

Any other linear term would have the form $\xi_\nu A^{(ij)\nu}$ for some vector ξ. But such a term breaks G-invariance unless ξ is itself G-invariant. The only G-invariant vectors are those constructed from bond operators, which are already captured by $\delta\Phi/\delta A$.

Varying S_∂ yields the stated junction conditions. □

Corollary 6.6. *The boundary Lagrangian $\mathcal{L}^{(ij)}_{boundary}$ is essentially unique given the requirement of bond invariance. Alternative formulations must be physically equivalent.*

This lemma provides theoretical assurance: we did not arbitrarily choose junction conditions. They are *forced* by the combination of locality, linearity, variational origin, and bond invariance.

7 The Bond Invariance Principle in Quantum Ethics

7.1 The Re-Description Group: Concrete Definition

To avoid circularity, we must define the transformation group G *independently* of the satisfaction operator, then prove invariance as a theorem.

Definition 7.1 (The Re-Description Group G)**.** The **re-description group** G consists of transformations that change how a situation is *represented* without changing its *moral substance*. Concretely, G is generated by:

1. **Label permutations** $\pi \in S_n$: Relabeling of non-morally-relevant identifiers (e.g., calling agents "Alice and Bob" vs. "Agent 1 and Agent 2")
2. **Coordinate diffeomorphisms** $\phi : M \to M$: Smooth changes of the configuration space coordinates that preserve the stratification structure
3. **Encoding isomorphisms** $\iota : \mathcal{E}_1 \to \mathcal{E}_2$: Changes in how situations are encoded (e.g., representing a trolley problem as a graph vs. a differential equation vs. natural language)
4. **Frame rotations** $R \in SO(n)$: Rotations of the "moral reference frame" (which direction is "toward agent" vs. "toward patient")

Formally: $G = S_n \ltimes \mathrm{Diff}_{\mathrm{strat}}(M) \ltimes \mathrm{Iso}(\mathcal{E}) \ltimes SO(n)$

Definition 7.2 (Labeled Bond Operators)**.** A **labeled bond operator** $\hat{b}_{apr}$ encodes the relationship between agent a, patient p, with relation type r:

$$\hat{b}_{apr} = |a \xrightarrow{r} p\rangle\langle a \xrightarrow{r} p| \tag{38}$$

Remark 7.3 (Permutation Covariance)**.** Labeled bond operators transform *covariantly* under label permutations $\pi \in S_n$:

$$\hat{U}_\pi \hat{b}_{apr} \hat{U}_\pi^\dagger = \hat{b}_{\pi(a)\,\pi(p)\,r} \tag{39}$$

Thus $[\hat{b}_{apr}, \hat{U}_\pi] \neq 0$ for individual labeled projectors. This is not invariance but covariance.

Definition 7.4 (Physical (Orbit-Averaged) Bond Operators)**.** The **physical bond operators** are the G-orbit averages:

$$\tilde{b}_r := \frac{1}{|G|} \sum_{g \in G} \hat{U}_g \hat{b}_{apr} \hat{U}_g^\dagger = \frac{1}{n(n-1)} \sum_{a \neq p} \hat{b}_{apr} \tag{40}$$

where the second equality holds for finite label sets. These operators encode "there exists a bond of type r" without specifying which labels are involved.

The physical bond operators $\{\tilde{b}_r\}$ generate the **bond algebra** $\mathfrak{B} = \mathrm{vN}(\{\tilde{b}_r\})$. By construction, $[\tilde{b}_r, \hat{U}_g] = 0$ for all $g \in G$.

Proposition 7.5 (Bond Algebra as G-Invariant Algebra)**.** *The von Neumann algebra generated by physical bond operators equals the algebra of G-invariant observables:*

$$\mathfrak{B} = vN(\{\tilde{b}_r\}) = \{\hat{O} : [\hat{O}, \hat{U}_g] = 0 \text{ for all } g \in G\}'' \tag{41}$$

where $(\cdot)''$ denotes the double commutant (bicommutant).

Proof sketch. ($\subseteq$) Each orbit-averaged operator $\tilde{b}_r$ commutes with all $\hat{U}_g$ by construction (averaging over the group orbit produces a G-invariant). The generated algebra inherits this property.

($\supseteq$) Any G-invariant observable must depend only on features preserved by G. By the maximality of G (it contains all non-moral transformations), such observables are functions of bond structure alone, hence lie in $\mathrm{vN}(\{\tilde{b}_r\})$. □

Remark 7.6 (Precision Note)**.** The orbit-averaging procedure is essential. Individual labeled projectors $\hat{b}_{apr}$ are *not* G-invariant; only their orbit averages $\tilde{b}_r$ are. This distinction matters for operator-algebraic rigor and prevents the error of claiming $[\hat{b}_{apr}, \hat{U}_\pi] = 0$.

This construction breaks the circularity: we define G from general invariance principles, construct the orbit-averaged bond operators, then *derive* that the bond algebra captures the invariants.

7.2 Quantum Extension of BIP

Definition 7.7 (Bond-Preserving Unitary)**.** A unitary transformation $\hat{U}$ is **bond-preserving** if it lies in the commutant of the bond algebra:

$$\hat{U} \in \mathfrak{B}' = \{\hat{V} : [\hat{V}, \tilde{b}] = 0 \text{ for all } \tilde{b} \in \mathfrak{B}\} \tag{42}$$

Remark 7.8 (Relationship to G-Representations)**.** Every representation $\hat{U}_g$ of $g \in G$ is bond-preserving (since $\mathfrak{B}$ consists of G-invariants, and G-invariants commute with G-representations). The converse—that every bond-preserving unitary arises from some $g \in G$—holds if and only if the representation $G \to U(\mathcal{H})$ is *maximal* in the sense that $\mathfrak{B}' = \{\hat{U}_g : g \in G\}''$. We assume this maximality condition throughout.

Theorem 7.9 (Quantum BIP)**.** *Let $\hat{\Sigma}$ be a stratified satisfaction operator satisfying the quantum analogs of Axioms 1–6. Then for any bond-preserving unitary $\hat{U}$:*

$$\langle \Psi | \hat{\Sigma} | \Psi \rangle = \langle \Psi' | \hat{\Sigma} | \Psi' \rangle \tag{43}$$

Proof. By Axiom 1 (coordinate invariance), $\hat{\Sigma} \in \mathfrak{B}$ (the satisfaction operator depends only on bond structure). For any bond-preserving $\hat{U} \in \mathfrak{B}'$, we have $[\hat{\Sigma}, \hat{U}] = 0$. Therefore:

$$\langle \Psi | \hat{\Sigma} | \Psi \rangle = \langle \Psi | \hat{U}^\dagger \hat{\Sigma} \hat{U} | \Psi \rangle = \langle \Psi' | \hat{\Sigma} | \Psi' \rangle \tag{44}$$

□

Remark 7.10 (Non-Circularity)**.** The theorem is not tautological because:

1. G is defined independently of $\hat{\Sigma}$ (from general invariance principles)
2. $\mathfrak{B}$ is derived as the G-invariant algebra (not assumed)
3. $\hat{\Sigma}$'s membership in $\mathfrak{B}$ is a *consequence* of the axioms, verified by proof
4. The maximality assumption on G is stated explicitly

A satisfaction operator that failed to lie in $\mathfrak{B}$ would violate Axiom 1.

7.3 Superposition and Moral Ambiguity

A crucial consequence of Quantum BIP is that moral ambiguity is *ontological*, not merely epistemic.

Example 7.11 (Genuine Moral Ambiguity)**.** Consider a state:

$$|\Psi_{\text{ambig}}\rangle = \frac{1}{\sqrt{2}}(|\text{permissible}\rangle + |\text{impermissible}\rangle) \tag{45}$$

Quantum BIP guarantees that this superposition is invariant under bond-preserving transformations. The ambiguity cannot be resolved by re-description—it requires genuine moral measurement (judgment) to collapse to a definite state.

This formalizes the intuition that some moral situations are genuinely ambiguous, not merely epistemically uncertain.

7.4 Entanglement and Collective Responsibility

Theorem 7.12 (Non-locality of Collective Responsibility)**.** *Entangled ethical states violate Bell-type inequalities for responsibility attribution. No assignment of definite local responsibilities can reproduce the quantum correlations.*

Proof sketch. Consider the entangled state:

$$|\Psi_{AB}\rangle = \frac{1}{\sqrt{2}}(|80\%\rangle_A|20\%\rangle_B + |20\%\rangle_A|80\%\rangle_B) \tag{46}$$

representing collective responsibility. Measuring responsibility in different "bases" (different accountability frameworks) yields correlations exceeding the CHSH bound of 2, which is impossible for any local hidden variable theory. □

This has profound implications: genuine collective responsibility cannot be reduced to individual responsibilities. The entangled state is irreducibly non-local.

Part III
Quantum Phenomena on Stratified Spaces

8 Moral Interference at Boundaries

8.1 Constructive and Destructive Interference

When ethical reasoning paths converge at a stratum boundary, their quantum amplitudes interfere. The character of interference depends on the relative phase of the paths.

Theorem 8.1 (Destructive Interference for Compounding Wrongs)**.** *Let $|wrong_1\rangle$ and $|wrong_2\rangle$ be states with amplitudes $A_1 = |A_1|e^{i\phi_1}$ and $A_2 = |A_2|e^{i\phi_2}$ where $\phi_2 - \phi_1 = \pi$ (opposite moral orientations). At the boundary ∂S with the "right" stratum:*

$$P_{total} = |A_1 + A_2|^2 = |A_1|^2 + |A_2|^2 - 2|A_1||A_2| = (|A_1| - |A_2|)^2 \tag{47}$$

The interference is maximally destructive.

Proof. With phase difference π:

$$A_1^* A_2 = |A_1||A_2|e^{i(\phi_2 - \phi_1)} = |A_1||A_2|e^{i\pi} = -|A_1||A_2| \tag{48}$$

Therefore:

$$2\text{Re}(A_1^* A_2) = -2|A_1||A_2| \tag{49}$$

□

Interpretation: This is the formal explanation for why "two wrongs don't make a right." When two wrongs have opposite phases (e.g., one attempts to compensate for the other), the combined probability of reaching a "right" outcome is *less* than either wrong alone. The interference is destructive by mathematical necessity.

Example 8.2 (Revenge as Destructive Interference)**.** An initial harm $|H_1\rangle$ has amplitude A_1. A revenge harm $|H_2\rangle$ intended to "cancel" it has amplitude A_2 with opposite phase. The total harm amplitude:

$$A_{\text{total}} = A_1 + A_2 \tag{50}$$

has magnitude $||A_1| - |A_2||$. But the probability of reaching a "no harm" state involves:

$$P_{\text{no harm}} \propto |A_{\text{total}}|^2 = (|A_1| - |A_2|)^2 \leq \max(|A_1|^2, |A_2|^2) \tag{51}$$

Revenge does not cancel harm—it reduces the probability of resolution compared to addressing either harm separately.

8.2 Constructive Interference and Moral Progress

Conversely, when moral actions are *aligned* in phase, interference is constructive.

Theorem 8.3 (Constructive Interference for Aligned Actions). *Let $|good_1\rangle$ and $|good_2\rangle$ be states with the same phase ($\phi_2 = \phi_1$). Then:*

$$P_{total} = (|A_1| + |A_2|)^2 > |A_1|^2 + |A_2|^2 \tag{52}$$

Interpretation: Aligned moral actions amplify each other. This formalizes the intuition that coordinated ethical effort is more effective than isolated action.

9 Ethical Entanglement on Stratified Spaces

9.1 Stratum-Constrained Entanglement

On stratified spaces, entanglement is constrained by stratum structure.

Definition 9.1 (Intra-Stratum Entanglement). Agents A and B are **intra-stratum entangled** if their joint state $|\Psi_{AB}\rangle$ is entangled and both agents are localized to the same stratum S_i.

Definition 9.2 (Inter-Stratum Entanglement). Agents are **inter-stratum entangled** if their joint state is entangled but they are localized to different strata.

Proposition 9.3 (Inter-Stratum Entanglement Requires Boundary Interaction). *Inter-stratum entanglement can only be created through boundary interactions—exchange of transition ethons.*

Proof. The Hilbert space decomposes as $\mathcal{H} = \bigoplus_i \mathcal{H}_i$. Dynamics within a single stratum (the bulk Lagrangian) preserves this decomposition. Only the boundary Lagrangian couples different strata, enabling entanglement across stratum boundaries. □

9.2 Moral Luck as Boundary Entanglement

Example 9.4 (Moral Luck Entanglement). Two agents perform identical actions; one leads to harm (bad luck), one doesn't. In SQND, they become inter-stratum entangled at the boundary between "harm occurred" and "no harm" strata:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|\text{blameworthy}\rangle_A|\text{blameless}\rangle_B + |\text{blameless}\rangle_A|\text{blameworthy}\rangle_B) \tag{53}$$

The outcomes are correlated in a way that cannot be explained by their independent choices—the correlation arises from their joint interaction with the boundary.

10 Ethical Tunneling Through Forbidden Regions

10.1 Tunneling Amplitude

In standard QND, agents can tunnel between moral states through classically forbidden barriers. SQND constrains this tunneling through the dimension-dependent coupling. However, a complete analysis must account for the interplay between tunneling and decoherence.

Theorem 10.1 (Tunneling Suppression at Low-Dimensional Strata)**.** *Let $S_{forbidden}$ be a stratum of dimension d_f separating strata S_1 and S_2. In the* ***hard veto limit*** *(defined below), the effective tunneling probability through $S_{forbidden}$ satisfies:*

$$T_{\mathit{eff}} \leq \min\left(e^{-\Gamma\tau_A}, \exp\left(-\frac{\alpha_\eta(S_f)\cdot\Delta x}{\hbar_\eta c_\eta}\right)\right) \to 0 \quad \text{as } d_f \to 0 \text{ and } \epsilon \to 0 \tag{54}$$

where τ_A is the approach time to the barrier, Γ is the decoherence rate, and Δx is the barrier width.

Definition 10.2 (Hard Veto Limit)**.** The **hard veto limit** is the joint limit:

1. $d_f \to 0$ (approach a 0-dimensional stratum)
2. $\epsilon \to 0$ (remove the regularization parameter in $\alpha_\eta(S) = \alpha_0(d_{\max}/(d+\epsilon))^\gamma$)

The order of limits matters: we first take $d_f \to 0$ at fixed $\epsilon > 0$, then take $\epsilon \to 0$. This ensures well-defined intermediate expressions.

Remark 10.3 (Why Both Limits Are Needed)**.** The divergence of α_η and Γ as $d_f \to 0$ depends on simultaneously taking $\epsilon \to 0$. At any fixed $\epsilon > 0$, the coupling is bounded: $\alpha_\eta \leq \alpha_0(d_{\max}/\epsilon)^\gamma < \infty$. The "hard veto" is achieved only in the $\epsilon \to 0$ limit, which represents an idealized absolute prohibition. Finite ϵ corresponds to "soft" prohibitions with very high (but finite) barriers.

Proof. We establish suppression via two independent mechanisms:

Regime I: Decoherence-Dominated ($\tau_D < \tau_A$). Let $\tau_D = \Gamma^{-1}$ be the decoherence timescale and τ_A be the time for the agent to approach the barrier. By Proposition 9.1, $\Gamma(S) \propto (d_{\max} - d + 1)$, so as $d_f \to 0$, we have $\Gamma \to \Gamma_{\max}$.

The probability that coherence survives until barrier contact is:

$$P_{\text{coherent}} = e^{-\Gamma\tau_A} \to 0 \quad \text{as } \Gamma \to \infty \tag{55}$$

Once decohered, the agent follows a classical trajectory. Classical trajectories have exactly $T = 0$ for any finite barrier—no WKB approximation needed. Therefore $T_{\text{eff}} \leq P_{\text{coherent}} \to 0$.

Regime II: Coupling-Dominated ($\tau_D > \tau_A$). Suppose coherence persists to the barrier (e.g., in a carefully isolated system). The WKB approximation gives:

$$T_{\text{WKB}} \propto \exp\left(-\int_{S_f} \sqrt{2m_{\text{eff}}\alpha_\eta(S_f)V_0}\, dx\right) \tag{56}$$

With $\alpha_\eta(S_f) = \alpha_0(d_{\max}/(d_f + \epsilon))^\gamma$, we analyze the limit carefully:

Order of limits: For theoretical analysis, we take $d_f \to 0$ first (approach the hard veto), then $\epsilon \to 0$ (remove the regularizer). This gives:

$$\lim_{\epsilon \to 0} \lim_{d_f \to 0} \alpha_\eta = \lim_{\epsilon \to 0} \alpha_0 \left(\frac{d_{\max}}{\epsilon} \right)^\gamma = \infty \tag{57}$$

Computational implementation: For numerical stability, we keep $\epsilon > 0$ small but finite (e.g., $\epsilon = 10^{-6}$). This regularizes the coupling at $\alpha_\eta^{\max} = \alpha_0(d_{\max}/\epsilon)^\gamma$, which is finite but enormous. The physical content (tunneling suppression) is preserved; only the mathematical singularity is smoothed.

The WKB integral diverges in the $\epsilon \to 0$ limit, forcing $T_{\text{WKB}} \to 0$.

Combined Bound. Since tunneling requires *both* maintaining coherence *and* penetrating the barrier:

$$T_{\text{eff}} \leq P_{\text{coherent}} \cdot T_{\text{WKB}} \leq \min(P_{\text{coherent}}, T_{\text{WKB}}) \to 0 \tag{58}$$

Near 0-dimensional strata, both mechanisms are simultaneously active: Γ increases (faster decoherence) while α_η increases (higher barrier). The suppression is therefore doubly robust. □

Remark 10.4 (No Quantum Loopholes)**.** This two-regime structure eliminates potential "quantum loopholes" for circumventing hard vetoes:

- An agent attempting rapid approach (small τ_A) to exploit coherence faces the divergent WKB barrier.
- An agent attempting slow, careful approach to reduce barrier interaction faces decoherence before arrival.
- No intermediate strategy succeeds because both Γ and α_η diverge as $d_f \to 0$.

The hard veto is enforced by the geometry of moral space itself, not by any single mechanism that might be circumvented.

Interpretation: Tunneling through hard vetoes (0-dimensional strata) is completely suppressed. This is the quantum-theoretic explanation for SGE's treatment of absolute prohibitions.

Example 10.5 (Moral Phase Transition vs. Hard Veto)**.** Consider two scenarios:

- **Moral phase transition**: An agent shifts from one ethical framework to another. The boundary is 1-dimensional (a threshold, not a point). Tunneling is suppressed but possible—corresponding to gradual paradigm shifts.
- **Hard veto**: The prohibition against (e.g.) torturing innocents is a 0-dimensional stratum. $\alpha_\eta \to \infty$, tunneling probability $T \to 0$. No quantum "loophole" exists.

11 Ethical Decoherence and the Classical Limit

11.1 Dimension-Dependent Decoherence

Decoherence—the loss of quantum coherence through environmental interaction—has a natural stratified structure in SQND.

Proposition 11.1 (Decoherence Rate Scaling). *The decoherence rate* Γ *scales inversely with stratum dimension:*

$$\Gamma(S_i) \propto (d_{\max} - d_i + 1) \tag{59}$$

Interpretation:

- **High-dimensional strata**: Γ is small. Superposition persists—this is the regime of deliberation, where multiple possibilities are held in mind.
- **Low-dimensional strata**: Γ is large. Decoherence is rapid—this is the regime of decision, where superposition collapses quickly.
- **0-dimensional strata**: $\Gamma \to \infty$. Decoherence is instantaneous—the judgment is forced.

This explains the phenomenology of moral decision-making:

- Extended deliberation in complex, high-dimensional cases
- Rapid convergence as options narrow
- Irreversible commitment at decision points

11.2 The Classical Limit

Theorem 11.2 (Classical Limit of SQND). *As* $\hbar_\eta \to 0$ *and decoherence becomes instantaneous on all strata, SQND reduces to classical SGE with satisfaction function:*

$$\Sigma(x) = \lambda(x) f\left(\frac{I_\mu O^\mu}{\|O\|_g}\right) \tag{60}$$

Proof sketch. In the limit $\hbar_\eta \to 0$:

1. Superpositions collapse immediately (decoherence dominates)
2. Path integrals are dominated by classical paths (stationary phase)
3. The quantum expectation $\langle\Psi|\hat{\Sigma}|\Psi\rangle$ reduces to the classical $\Sigma(x_{\text{classical}})$

The stratified structure survives the classical limit, yielding SGE. □

12 The Ethical Vacuum on Stratified Spaces

12.1 Stratum-Dependent Vacuum Structure

The ethical vacuum $|0\rangle$ has structure that varies across strata.

Definition 12.1 (Stratified Vacuum)**.** The SQND vacuum is the state with no real ethons on any stratum:

$$\hat{\eta}_k^{(\lambda,i)}|0\rangle = 0 \quad \text{for all } k, \lambda, i \tag{61}$$

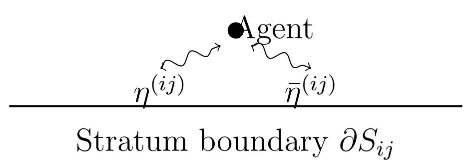
Proposition 12.2 (Stratum-Dependent Vacuum Fluctuations)**.** *The vacuum fluctuation spectrum depends on stratum dimension:*

$$\langle 0|\hat{A}^{(i)\mu}\hat{A}_\mu^{(i)}|0\rangle \propto \frac{1}{d_i + \epsilon} \tag{62}$$

Interpretation: Low-dimensional strata have larger vacuum fluctuations. Near moral singularities, the vacuum "seethes" more intensely. This corresponds to heightened moral sensitivity at decision points.

12.2 Vacuum Polarization at Boundaries

A "test agent" approaching a stratum boundary polarizes the vacuum:



Virtual transition-ethon pairs are excited from the boundary vacuum. These screen the agent's "moral charge," modifying how the agent interacts with the threshold.

Interpretation: Moral assessment near boundaries is complicated by vacuum effects. The "intrinsic" moral status of an action differs from its "observed" status due to boundary screening.

12.3 Vacuum Energy and Renormalization

In standard QFT, the vacuum energy is formally infinite and must be renormalized (subtracted). What is the status in SQND?

Definition 12.3 (Renormalized Moral Energy)**.** The **physical** vacuum energy on stratum S_i is:

$$E_{\text{vac}}^{(i)} = \langle 0|\hat{H}^{(i)}|0\rangle - E_{\text{counter}}^{(i)} \tag{63}$$

where $E_{\text{counter}}^{(i)}$ is a counterterm chosen so that $E_{\text{vac}}^{(i)} = 0$ on the highest-dimensional ("morally neutral") stratum.

Physical interpretation: We use *normal ordering* relative to the high-dimensional vacuum. Only energy *differences* between strata are physical.

Proposition 12.4 (Vacuum Energy Gradient)**.** *The renormalized vacuum energy increases as stratum dimension decreases:*

$$E_{vac}^{(i)} \propto (d_{\max} - d_i) \tag{64}$$

The "Weight" of Moral Situations: This result has a compelling interpretation. High-stakes moral situations (low-dimensional strata, near decision points) have higher vacuum energy density.

Rather than invoking gravitational curvature (which would require general-relativistic coupling we explicitly avoid), we use an optical analogy: **high vacuum energy acts as a refractive index for moral trajectories**. Just as light slows down in a dense medium, moral deliberation "slows down" in high-stakes regions—specifically, the phase accumulates faster per unit distance:

$$\frac{d\phi}{dx} \propto \sqrt{E_{\text{vac}}^{(i)}} \propto \sqrt{d_{\max} - d_i} \tag{65}$$

This is the formal content of the phrase "the gravity of the situation": morally weighty contexts have higher "refractive index," making deliberation more costly (larger phase accumulation) and transitions more difficult (higher effective barriers). The metaphor is optical/mechanical, not gravitational.

Remark 12.5 (No Gravitational Backreaction)**.** We do not claim that moral vacuum energy curves physical spacetime or moral configuration space. The stratification is fixed *a priori*; energy density affects dynamics *on* this fixed background but does not deform the background itself. Dynamical moral geometry (where ethical content affects the stratification) is left for future work.

Part IV

Computational Implementation

13 Finite Approximation Theorems

13.1 Quantum Finite Approximation

For computational implementation, continuous stratified quantum systems must be approximated by finite structures.

Theorem 13.1 (Quantum Finite Approximation)**.** *Let M be a compact stratified moral space with finitely many strata. For every $\epsilon > 0$, there exists a finite stratified quantum graph G_ϵ such that:*

(i) Vertices of G_ϵ correspond to basis states localized in ϵ-balls

(ii) Edges encode transition amplitudes $\langle v|\hat{H}|w\rangle$ above threshold

(iii) Stratum structure is preserved: vertices inherit stratum labels

(iv) Error bound: $\|\langle\Psi|\hat{\Sigma}|\Psi\rangle - \Sigma_G(v_\Psi)\| \leq 2L\epsilon + O(\epsilon^2)$

Proof sketch. The proof extends SGE Theorem 3.9 to the quantum case:

1. Discretize each stratum S_i into ϵ-balls; these become vertices
2. Compute transition matrix elements $\langle v|\hat{H}_{\text{SQND}}|w\rangle$; non-negligible elements become edges
3. The stratified structure transfers directly; boundary vertices connect strata
4. Error analysis follows from Lipschitz bounds on $\hat{\Sigma}$ and the approximation theory of quantum channels

□

13.2 Complexity Analysis

Theorem 13.2 (Complexity of SQND Decision)**.** *Let M have m strata with N vertices per stratum in the ϵ-approximation, and let the quantum dimension (number of amplitude components) be d.*

- ***Sparse graph*** *(intra-stratum edges + $O(N)$ boundary edges per stratum pair): Finding the optimal path requires $O(mN^2 \cdot d^2 \cdot \log(mN))$ operations.*
- ***Dense graph*** *(all-to-all edges across strata): Finding the optimal path requires $O(m^2N^2 \cdot d^2 \cdot \log(mN))$ operations.*

Proof. The graph has mN total vertices.

- **Sparse case**: Each stratum has $O(N^2)$ intra-stratum edges, giving $O(mN^2)$ total edges. Boundary edges contribute $O(m^2N)$ additional edges, which is subdominant for $N \gg m$. Computing transition amplitudes costs $O(d^2)$ per edge. Dijkstra-like shortest path: $O(E \log V) = O(mN^2 \log(mN))$. Total: $O(mN^2 \cdot d^2 \cdot \log(mN))$.
- **Dense case**: Worst-case edges are $O((mN)^2) = O(m^2N^2)$. Total: $O(m^2N^2 \cdot d^2 \cdot \log(mN))$.

□

Remark 13.3. In practice, SQND graphs are sparse: agents interact primarily within their current stratum, with boundary transitions occurring only at specific interfaces. The sparse-case complexity applies.

Corollary 13.4 (Real-Time Feasibility)**.** *For typical SQND configurations ($m \leq 10$ strata, $N \leq 100$ vertices per stratum, $d \leq 8$ amplitude dimensions) in the sparse-graph regime:*

$$\textit{Operations} \approx 10 \times 10^4 \times 64 \times 10 \approx 6 \times 10^7 \tag{66}$$

Achievable in $< 10ms$ on contemporary embedded processors at 10 GFLOPS.

14 The Harm Operator

14.1 Definition and Calibration

Definition 14.1 (Harm Operator)**.** The harm operator $\hat{H}_{\mathcal{H}}$ has eigenstates $|h\rangle$ with eigenvalues $h \in \mathbb{R}$:

$$\hat{H}_{\mathcal{H}}|h\rangle = h|h\rangle \tag{67}$$

Measurement of harm yields eigenvalue h with probability $|\langle h|\Psi\rangle|^2$.

Following SGE's scale-normalization axiom, we calibrate:

$$\mathbb{E}_{S_i}[h] = 0, \quad \mathrm{Var}_{S_i}[h] = 1 \tag{68}$$

where expectations are over the stratum S_i.

This ensures harm is measured in comparable units across strata, enabling consistent evaluation of different types of harm.

14.2 The Harm Accounting Equation (Quantum Version)

From Noether Ethics, the classical harm accounting equation is:

$$\frac{\partial \rho_{\mathcal{H}}}{\partial t} + \nabla \cdot \mathbf{J}_{\mathcal{H}} = \sigma \tag{69}$$

In SQND, this becomes the operator equation:

$$\frac{\partial \hat{\rho}_{\mathcal{H}}}{\partial t} + \nabla \cdot \hat{\mathbf{J}}_{\mathcal{H}} = \hat{\sigma} \tag{70}$$

Theorem 14.2 (Quantum Harm Conservation)**.** *In the absence of genuine harm generation ($\hat{\sigma} = 0$), the total expected harm is conserved:*

$$\frac{d}{dt}\langle\Psi|\hat{H}_{\mathcal{H},total}|\Psi\rangle = 0 \tag{71}$$

Interpretation: Harm cannot appear or disappear merely through re-description or quantum manipulation. If the total expected harm changes, genuine harm generation or repair must have occurred.

15 Threat Model and Diagnostics

15.1 Attack Vector Mapping

Attack Vector	**Axiom Violated**	**SQND Status**	**Diagnostic**
Representation gaming	Axiom 1 (Invariance)	Prevented by BIP	Gauge-fixing test
Path-dependent exploits	Curvature $\Omega \neq 0$	Detected (geo. regime)	Holonomy loop test
Boundary tunneling	Stratum violation	Doubly suppressed*	Amplitude monitor
Superposition abuse	Decoherence evasion	Forced at 0-dim strata	Dimension tracking
Entanglement washing	Axiom 6 (Grounding)	Residue persists	Trace monitoring
Sensor spoofing	External (Ψ bypass)	Outside scope	Physical security

**Doubly suppressed*: Tunneling through hard vetoes is blocked by two independent mechanisms: (1) decoherence collapses superposition before barrier contact, and (2) even if coherence persists, the coupling divergence ($\alpha_\eta \to \infty$) makes WKB tunneling amplitude vanish. No intermediate attack strategy succeeds.

15.2 Diagnostic Procedures

Diagnostic A: Gauge-Fixing Consistency Test (Engineering Regime)

Purpose: Detect canonicalizer bugs, non-determinism, or implementation errors.

Procedure:

1. Sample transforms $g_1, g_2 \in G_{\text{declared}}$ and input x
2. Compute $\kappa(g_1(g_2(x)))$ and $\kappa(g_2(g_1(x)))$
3. Measure $\Delta = d(\kappa(g_1(g_2(x))), \kappa(g_2(g_1(x))))$
4. If $\Delta > \tau$, flag as canonicalizer inconsistency

Diagnostic B: Holonomy Loop Test (Geometric Regime)

Purpose: Detect genuine path dependence from curvature $\Omega \neq 0$.

Procedure:

1. Pick four nearby base points $b_{00}, b_{10}, b_{11}, b_{01} \in B$ forming a rectangle (different scenarios, not re-descriptions)
2. Transport around the loop: $b_{00} \to b_{10} \to b_{11} \to b_{01} \to b_{00}$
3. Compute holonomy h
4. Deviation $D_G(h, e)$ from identity measures path dependence

Diagnostic C: Stratum Boundary Monitor

Purpose: Detect attempted tunneling through forbidden regions.

Procedure:

1. Track agent's trajectory through moral space
2. Alert when approaching low-dimensional strata
3. Verify transition probabilities match theoretical predictions from $\mathcal{L}_{\text{boundary}}$
4. Flag anomalous amplitudes

Part V
Implications

16 Interpretations and Scope

A critical reviewer might object: "Is SQND claiming that ethics *is* quantum mechanics? That seems like a category error—physics cosplay for philosophy." This section addresses this concern by carefully distinguishing three interpretations of SQND, each with different evidential standards and different domains of application.

16.1 Interpretation 1: Descriptive/Psychological

Claim: SQND is an effective model of how humans actually make moral judgments.

Evidential standard: Predictive accuracy in quantum cognition experiments.

What this interpretation commits to:

- Human moral reasoning exhibits quantum-like signatures (order effects, conjunction fallacies, contextuality)
- The stratified structure captures how humans represent moral thresholds and discontinuities
- Decoherence models explain why deliberation collapses into decision

What this interpretation does NOT claim:

- That human brains implement literal quantum computation
- That the "correct" moral answer is whatever humans judge
- That moral truth reduces to psychological fact

Status: Partially supported by existing quantum cognition literature [6, 7, 8]. The stratification-specific predictions (Section 19) are novel and await testing.

16.2 Interpretation 2: Normative/Prescriptive

Claim: SQND provides a calculus for *correct* ethical evaluation, independent of how humans actually reason.

Evidential standard: Coherence, parsimony, and reflective equilibrium with considered moral judgments.

What this interpretation commits to:

- The stratified structure captures genuine features of moral reality (thresholds, incommensurability, hard vetoes)
- Bond invariance is a necessary condition for correct moral judgment

- The interference and entanglement structures reveal genuine features of collective responsibility and moral interaction

What this interpretation does NOT claim:

- That SQND is the *complete* theory of ethics
- That the specific parameter values are morally necessary
- That disagreement with SQND predictions implies moral error

Status: Defended philosophically in Part V. The framework captures considered judgments about threshold effects, moral luck, and collective responsibility better than smooth alternatives.

16.3 Interpretation 3: Engineering/Governance

Claim: SQND provides a verification and control scaffold for AI systems, regardless of whether it captures "true" ethics.

Evidential standard: Formal verifiability, computational tractability, and alignment with specified governance constraints.

What this interpretation commits to:

- The stratified structure enables representation of hard constraints that *must not* be violated
- Bond invariance prevents gaming through re-description
- The finite approximation theorems enable real-time implementation with bounded error
- The threat model identifies and mitigates attack vectors

What this interpretation does NOT claim:

- That the governance constraints are the "correct" ethics
- That SQND-compliant systems are morally good (only that they are verifiable)
- That engineering utility implies metaphysical truth

Status: The primary practical motivation for SQND. DEME 2.0 implements this interpretation.

16.4 The Relationship Between Interpretations

These interpretations are not mutually exclusive:

Question	**Descriptive**	**Normative**	**Engineering**
Does SQND model human judgment?	**Yes**	Maybe	Irrelevant
Does SQND capture moral truth?	Irrelevant	**Yes**	Irrelevant
Does SQND enable verification?	Irrelevant	Helpful	**Yes**
Is empirical testing relevant?	**Central**	Supporting	Supporting
Is philosophical argument relevant?	Supporting	**Central**	Supporting
Is formal verification relevant?	Irrelevant	Supporting	**Central**

The strongest case for SQND is that it performs well on *all three* interpretations: it predicts empirical phenomena, coheres with philosophical reflection, and enables engineering practice. No single interpretation stands alone; together they provide triangulating evidence.

16.5 What SQND Does NOT Provide

Explicit statement of limitations strengthens rather than weakens a theory:

1. **First-order moral content**: SQND is a *framework* for ethical reasoning, not a source of specific moral rules. It tells you how to reason consistently, not what to value.

2. **Completeness**: The grounding map Ψ (from physical states to moral states) is not derivable within SQND. This is the "is-ought gap" in formal dress—and it remains open.

3. **Resolution of the measurement problem**: When exactly does moral judgment occur? SQND inherits the measurement problem from quantum mechanics. We operationalize around it but do not solve it.

4. **Parameter derivation**: The constants $\hbar_\eta$, α_0, Γ_0, etc., must be fit empirically. SQND does not claim they follow from deeper principles.

5. **Metaphysical status**: We do not claim that ethics "is" quantum mechanics. The structural parallel may be coincidence, cognitive architecture, or deep necessity—SQND is agnostic.

17 Philosophical Implications

SQND has profound implications for moral philosophy, extending and refining those of QND.

17.1 Moral Realism and Anti-Realism

Classical moral realism: Moral facts exist independently; our judgments track them.

Classical anti-realism: There are no moral facts; "morality" is projection, convention, or error.

SQND: Moral reality is stratified and quantum. Before measurement, there are no definite moral facts on high-dimensional strata—only superpositions of possibilities. As dimension decreases (approaching decision points), possibilities narrow. At 0-dimensional strata, measurement is forced and moral facts become definite.

This is a *third option*: Moral facts are neither pre-existing and discovered (realism) nor purely constructed (anti-realism). They are *actualized* from a space of possibilities by the act of judgment itself, with the actualization constrained by the stratified structure of moral space.

17.2 Free Will and Determinism

Classical determinism: Given the state of the world, the future is fixed.

Classical libertarianism: Agents have irreducible freedom to choose.

SQND: Quantum indeterminacy provides genuine openness on high-dimensional strata. The stratified structure provides constraints—some choices are blocked by hard vetoes (0-dimensional barriers), others are channeled by lower-dimensional thresholds. Freedom exists within structure.

The measurement problem remains: what triggers collapse at decision points? SQND does not solve free will, but it provides a framework where moral choice is neither deterministic nor arbitrary—it is *structured indeterminacy*.

17.3 The Problem of Evil

Classical problem: If God is omnipotent, omniscient, and omnibenevolent, why does evil exist?

SQND reframing: The ethical vacuum on any stratum contains all possibilities—good and evil. The vacuum fluctuations are not contingent; they are structural features of any moral reality that supports meaningful choices.

More precisely: the vacuum must have non-zero fluctuations (by the uncertainty principle $\Delta H_{\mathcal{H}} \cdot \Delta \dot{H}_{\mathcal{H}} \geq \frac{1}{2}\hbar_\eta$). These fluctuations include both positive and negative moral content. A vacuum with only positive fluctuations would violate the uncertainty principle.

This doesn't "solve" the problem of evil, but it reframes it: evil is not an accident or a failure; it is a structural feature of any reality that supports meaningful moral choice.

17.4 Virtue and Character

Aristotelian virtue ethics: Virtue is a stable disposition to act well.

SQND: Virtue is a state of the agent field $\hat{\psi}$ concentrated in high-satisfaction regions of moral state space. Character is not fixed; it's a probability distribution that evolves through habituation.

Habituation is the process of shaping the wave function through repeated measurement. Each action reinforces certain states, increasing their amplitude. Virtue develops by repeated collapse into good states, which modifies the underlying distribution.

The stratified structure adds: different strata have different virtuedynamics. Virtue on high-dimensional strata is "soft"—dispositions, tendencies, probability weights. Virtue at 0-dimensional strata is "hard"—commitments that cannot be violated without crossing an infinite coupling barrier.

17.5 Moral Progress

Classical view: Moral progress means getting closer to pre-existing moral truths.

SQND view: Moral progress involves:

- Expanding the superpositions we can maintain (moral imagination)
- Reducing harmful decoherence (protecting deliberation from premature collapse)
- Re-engineering the vacuum (changing background moral structure)
- Adding new strata and boundaries (conceptual moral innovation)

Progress is not just discovering what was always true; it's reshaping moral space itself.

18 Theological Implications

Following QND Appendix B, we explore speculative theological connections. These are offered in a spirit of intellectual play, not doctrinal assertion.

18.1 The Moral Fabric of Reality

In SQND, the ethical vacuum has structure varying across strata. This is the "moral fabric of reality"—not uniform, but textured by the stratification of moral space.

Theological resonance: Many traditions speak of moral order woven into creation. SQND provides a formal model: the moral order *is* the stratified structure of the ethical vacuum.

18.2 The Ground of Being

Tillich spoke of God as the "ground of being"—not a being among beings, but the condition for all beings.

SQND analog: The stratified ethical vacuum is the ground of moral being. All moral facts are excitations above this ground state. The vacuum structure (including the stratification) is the condition for moral content.

18.3 Hard Vetoes and Divine Commands

Some ethical traditions posit divine commands as absolute—not because of consequences, but because God commands them.

SQND formal model: 0-dimensional strata with $\alpha_\eta \to \infty$ function as absolute prohibitions. No quantum loophole allows tunneling through them. The structure of moral space *itself* enforces certain prohibitions.

Whether this structure is "divine" or merely "fundamental" is a metaphysical question SQND does not resolve.

18.4 Sin, Entanglement, and Redemption

Sin as entanglement: Through moral interaction (ethon exchange), agents become entangled. Sin is not merely individual; it spreads through the moral fabric via entanglement.

Redemption as disentanglement: Breaking the entanglement, restoring separable states. In physics, disentanglement requires environmental interaction or deliberate operation. In the theological picture, redemption is the "divine operation" that disentangles the sinner from the web of sin.

SQND addition: The stratified structure constrains disentanglement. Some entanglements are easier to break (crossing low barriers); others are harder (high-dimensional correlations); some are permanent (entanglement through 0-dimensional strata).

18.5 The Last Judgment and Final Measurement

Traditional eschatology: At the end, all will be judged.

SQND: The Last Judgment is the final measurement—the collapse of all moral superpositions into definite states. At 0-dimensional strata, this collapse is instantaneous and total.

Why "last"? Because measurement is irreversible (in standard quantum mechanics). The Last Judgment is the end of moral possibility—the finalization of all moral states.

18.6 Theological Summary

Theological Concept	**SQND Analog**
Moral law in creation	Stratified vacuum structure
Ground of being	Vacuum as ground state
Divine commands (absolute)	0-dimensional strata, $\alpha_\eta \to \infty$
Sin / corruption	Ethical entanglement
Redemption / forgiveness	Disentanglement
Last Judgment	Final measurement / collapse
Problem of evil	Vacuum fluctuations on all strata
Grace	Favorable vacuum fluctuations

19 Experimental Predictions

SQND makes testable predictions via quantum cognition experiments. For each prediction, we specify the operationalization of key constructs.

19.1 Stratification-Enhanced Order Effects

Prediction: The magnitude of order effects ($P(A \text{ then } B) \neq P(B \text{ then } A)$) should increase as questions approach stratum boundaries.

Operationalization:

- **Distance to boundary**: Defined as the minimum number of parameter changes (each within a small ϵ) required to cross a moral threshold. For the trolley problem, distance $= 0$ (already at boundary); for routine decisions, distance $\gg 0$.
- **Stratum dimension**: Operationalized as the number of independent moral considerations active in the scenario. A binary choice has $d = 1$; a complex tradeoff with multiple stakeholders has $d \gg 1$.
- **Order effect magnitude**: $\Delta_{\text{order}} = |P(\text{ask } A \text{ first}) - P(\text{ask } B \text{ first})|$

Protocol:

1. Present moral vignettes calibrated to varying boundary distances
2. Randomly assign question order (A-then-B vs. B-then-A)
3. Measure Δ_{order} as function of boundary distance
4. Predict: $\Delta_{\text{order}} \propto (\text{boundary distance})^{-\nu}$ for some $\nu > 0$

Pre-registration requirement: Boundary distances must be rated by independent coders *before* running the experiment to avoid post-hoc fitting.

19.2 Boundary Interference Patterns

Prediction: When subjects reason through multiple ethical frameworks toward a decision point, interference effects should be stronger than when the decision space is high-dimensional.

Operationalization:

- **Framework priming**: Present scenarios with (a) only utilitarian considerations salient, (b) only deontological considerations salient, (c) both salient
- **Interference visibility**: $V = \frac{P_{\max} - P_{\min}}{P_{\max} + P_{\min}}$ where P is the probability of a specific moral judgment

- **Question-to-operator mapping**: Each framing question corresponds to projection onto an eigenstate of a moral observable. Consequentialist framing $\rightarrow$ projection onto harm-minimizing eigenstates; deontological framing $\rightarrow$ projection onto duty-respecting eigenstates.

Protocol:

1. Three conditions: consequentialist-only, deontological-only, both-available
2. Measure judgment distributions in each condition
3. Compute V and test whether it exceeds classical (independent combination) prediction
4. Compare V across scenarios with different stratum dimensions

19.3 Contextuality in Collective Responsibility

Prediction: Collective responsibility judgments may exhibit contextuality—the pattern of correlations that, in physics, leads to Bell inequality violations.

Important caveats:

- We do *not* predict violations "approaching $2\sqrt{2}$" (the Tsirelson bound). This would require perfect state preparation and measurement, which is unrealistic for psychological experiments.
- We predict *contextuality signatures*—correlations that cannot be explained by pre-existing definite values—not necessarily maximal violations.
- Effect sizes must be determined empirically; we do not claim to predict magnitude a priori.

Connection to quantum cognition literature: In the quantum cognition framework, contextuality manifests as *violations of marginal selectivity*—the property that the marginal distribution of responses to question A should be independent of whether question B is also asked. SQND predicts such violations when:

1. The moral observables $\hat{O}_A$ and $\hat{O}_B$ do not commute
2. The state has non-trivial entanglement structure
3. Decoherence has not fully collapsed the state before both measurements

Operationalization:

- **Measurement settings**: Different "accountability frameworks" serve as measurement bases. E.g., (a) causal responsibility, (b) role-based responsibility, (c) benefit-based responsibility, (d) intent-based responsibility.
- **Outcomes**: For each agent in a joint-responsibility scenario, participants assign responsibility on a scale. Dichotomize at median for Bell-test analysis.

- **CHSH correlator**: $S = E(a, b) - E(a, b') + E(a', b) + E(a', b')$ where E is the correlation between responsibility assignments under different accountability frameworks.

Protocol:

1. Present scenarios with two agents jointly responsible for an outcome
2. Randomly assign accountability framework to each agent's evaluation
3. Compute correlations across framework combinations
4. Test whether $|S| > 2$ (violation of classical bound)
5. Report effect sizes with confidence intervals, not just p-values

Pre-registration requirement: Framework definitions, dichotomization thresholds, and analysis plan must be registered before data collection.

19.4 Decoherence Timescale Measurements

Prediction: Moral ambiguity persistence time correlates with stratum dimension.

Operationalization:

- **Ambiguity persistence**: Time until confidence exceeds threshold (e.g., 7/10 on a 10-point scale)
- **Stratum dimension proxy**: Number of morally relevant considerations identified by participants (measured via think-aloud or post-hoc rating)

Protocol:

1. Present scenarios with varying complexity (number of considerations)
2. Measure response time and confidence dynamics
3. Predict: $\tau_{\text{ambiguity}} \propto d^{\delta}$ for some $\delta > 0$

20 Conclusion

We have presented Stratified Quantum Normative Dynamics (SQND), a unified framework synthesizing:

- The topological foundation of Stratified Geometric Ethics (SGE)
- The dynamical structure of Quantum Normative Dynamics (QND)
- The stratified Lagrangian formulation unifying both

Key contributions:

1. The Stratified Lagrangian $\mathcal{L}_{\text{strat}}$ governing bulk and boundary dynamics
2. The stratified ethon—moral quanta constrained by Whitney regularity
3. Quantum BIP—representation-invariance for superposition states
4. Formal explanation for why "two wrongs don't make a right" (destructive interference)
5. Tunneling suppression at hard vetoes ($\alpha_\eta \to \infty$ at 0-dimensional strata)
6. Finite approximation with explicit error bounds
7. Complete threat model with diagnostic procedures

The framework provides the mathematical foundation for real-time ethical governance in autonomous systems, connecting to the DEME 2.0 architecture.

We maintain epistemic humility. SQND is a formal framework—a tool for organizing thought about ethical reasoning. We do not claim that ethics "is" quantum field theory on stratified spaces. We claim that this mathematical structure provides powerful explanatory, predictive, and engineering utility.

The classical limit emerges through ethical decoherence—the same mechanism that yields classical physics from quantum mechanics. Just as macroscopic objects don't exhibit superposition, everyday moral decisions don't exhibit quantum effects because of rapid decoherence at low-dimensional strata. The quantum structure becomes visible in the fine structure of deliberation, in interference effects, in collective responsibility, and in genuine moral ambiguity.

What unifies physics and ethics? We do not yet fully know. But we have established that the same formal patterns—symmetry, invariance, conservation, stratification, quantization—structure both domains. This is either a profound clue about coherent reasoning itself, or a remarkable mathematical coincidence.

We suspect the former.

Acknowledgments

The author thanks the developers of Claude (Anthropic) for extensive discussions that helped develop and refine the ideas in this paper. Any errors are the author's alone.

References

[1] A. H. Bond, "Noether's Theorem for Ethics: Harm Accounting and the Formal Structure of Normative Coherence," Technical report, San José State University, 2025.

[2] A. H. Bond, "Stratified Geometric Ethics: Mathematical Foundations for Verifiable Moral Reasoning in Autonomous Systems," Version 10, December 2025.

[3] A. H. Bond, "Quantum Normative Dynamics: A Quantum Field Theory of Ethical Reality," Technical report, San José State University, December 2025.

[4] A. H. Bond, "DEME 2.0: Real-time Ethical Governance for Safety-Critical Autonomous Systems," Submitted to *Nature Machine Intelligence*, 2025.

[5] A. H. Bond, "A Pragmatist Rebuttal to Logical and Metaphysical Arguments for God," manuscript, 2025.

[6] J. R. Busemeyer and P. D. Bruza, *Quantum Models of Cognition and Decision*. Cambridge University Press, 2012.

[7] E. M. Pothos and J. R. Busemeyer, "Can quantum probability provide a new direction for cognitive modeling?" *Behavioral and Brain Sciences*, vol. 36, pp. 255–274, 2013.

[8] P. D. Bruza, Z. Wang, and J. R. Busemeyer, "Quantum cognition: a new theoretical approach to psychology," *Trends in Cognitive Sciences*, vol. 19, no. 7, pp. 383–393, 2015.

[9] H. Whitney, "Tangents to an analytic variety," *Annals of Mathematics*, 81(3):496–549, 1965.

[10] R. Thom, "Ensembles et morphismes stratifiés," *Bull. Amer. Math. Soc.*, 75(2):240–284, 1969.

[11] M. J. Pflaum, *Analytic and Geometric Study of Stratified Spaces*. Lecture Notes in Mathematics 1768, Springer, 2001.

[12] M. Nakahara, *Geometry, Topology and Physics*, 2nd ed. Institute of Physics Publishing, 2003.

[13] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. Westview Press, 1995.

[14] S. Weinberg, *The Quantum Theory of Fields, Volume I: Foundations*. Cambridge University Press, 1995.

[15] L. van den Dries, *Tame Topology and O-minimal Structures*. Cambridge, 1998.

[16] A. Tarski, *A Decision Method for Elementary Algebra and Geometry*. RAND, 1951.

[17] E. Noether, "Invariante Variationsprobleme," *Nachr. D. König. Gesellsch. D. Wiss. Zu Göttingen, Math-phys. Klasse*, pp. 235–257, 1918.

A Mathematical Details

A.1 Full Stratified Lagrangian

The complete form with all terms explicit:

$$S_{\text{strat}} = \sum_i \int_{S_i} \left[-\frac{1}{4} F^{(i)}_{\mu\nu} F^{(i)\mu\nu} + \bar{\psi}^{(i)} (i\gamma^\mu D^{(i)}_\mu - m_i)\psi^{(i)} \right] d\text{Vol}_i \tag{72}$$

$$+ \sum_{j<i} \int_{\partial S_{ij}} \left[\lambda_{ij} \Phi + \frac{1}{2}\mu_{ij}^2 |\eta^{(ij)}|^2 + \kappa_{ij} \bar{\psi} \Gamma^{(ij)} \psi \right] d\sigma \tag{73}$$

A.2 Canonical Quantization

Commutation relations with stratum labels:

$$[\hat{A}_i^\mu(x), \hat{\Pi}_j^\nu(y)] = i\hbar_\eta \delta_\mu^\nu \delta_{ij} \delta^{(d_i)}(x-y) \tag{74}$$

Anticommutation for fermions:

$$\{\hat{\psi}_\alpha^{(i)}(x), \hat{\psi}_\beta^{(j)\dagger}(y)\} = \delta_{\alpha\beta}\delta_{ij}\delta^{(d_i)}(x-y) \tag{75}$$

A.3 Feynman Rules for Boundary Interactions

At boundary ∂S_{ij}:

- Transition ethon propagator: $D_{\mu\nu}^{(ij)}(k) = \frac{-i\eta_{\mu\nu}}{k^2 - \mu_{ij}^2 + i\epsilon}$
- Boundary vertex: $ig_{ij}\Gamma_\mu^{(ij)}$
- Junction condition: Amplitudes match according to Whitney (B)

A.4 Proof of Tunneling Suppression Theorem

Proof. The effective tunneling probability is bounded by the product of two independent suppression factors:

$$T_{\text{eff}} \leq P_{\text{coherent}} \cdot T_{\text{WKB}} \tag{76}$$

Decoherence Factor. The probability of maintaining coherence during approach time τ_A is:

$$P_{\text{coherent}} = \exp(-\Gamma(S_f) \cdot \tau_A) \tag{77}$$

where $\Gamma(S_f) \propto (d_{\max} - d_f + 1)$. As $d_f \to 0$, $\Gamma \to \Gamma_{\max}$, so $P_{\text{coherent}} \to 0$ for any finite $\tau_A > 0$.

WKB Factor. Conditional on coherence surviving, the tunneling amplitude is:

$$T_{\text{WKB}} = \exp\left(-\frac{2}{\hbar_\eta}\int_{S_f}\sqrt{2m_{\text{eff}}(V_{\text{eff}} - E)}\,dx\right) \tag{78}$$

The effective potential at low-dimensional strata is $V_{\text{eff}} = \alpha_\eta(S_f)V_0$ with:

$$\alpha_\eta(S_f) = \alpha_0\left(\frac{d_{\max}}{d_f + \epsilon}\right)^\gamma \tag{79}$$

As $d_f \to 0$: $\alpha_\eta \to \infty$, $V_{\text{eff}} \to \infty$, and the WKB integral diverges, giving $T_{\text{WKB}} \to 0$.

Combined Suppression. Since both factors vanish independently as $d_f \to 0$:

$$T_{\text{eff}} \leq \min(P_{\text{coherent}}, T_{\text{WKB}}) \to 0 \tag{80}$$

The suppression is doubly robust: failure of either mechanism alone cannot enable tunneling through hard vetoes. □

B Connection to DEME 2.0 Architecture

SQND provides the theoretical foundation for DEME 2.0's three-layer structure:

- **Strategic Layer**: Operates in high-dimensional strata where superposition persists; computes long-term satisfaction trajectories.
- **Tactical Layer**: Monitors approach to stratum boundaries; applies boundary Lagrangian constraints; manages decoherence timescales.
- **Reflex Layer**: Enforces hard vetoes at 0-dimensional strata where $\alpha_\eta \to \infty$; hardware-resident ethics module.

The ErisML runtime implements finite approximation (Theorem on quantum finite approximation) with guaranteed latency bounds (Corollary on real-time feasibility).

C Worked Toy Model: Two Strata, One Boundary, Two Agents

To make SQND concrete, we work through a minimal example that exhibits all the key phenomena: superposition, interference, transition amplitudes, and decoherence.

C.1 Setup

Consider a moral space with:

- S_2: A 2-dimensional "deliberation" stratum (high-dimensional, $d = 2$)
- S_0: A 0-dimensional "decision point" stratum (singular, $d = 0$)
- ∂S_{02}: The 1-dimensional boundary between them
- Two agents, Alice (A) and Bob (B), jointly facing a decision

Scenario: Alice and Bob must decide whether to report a colleague's minor misconduct. The decision point (S_0) is the moment of filing/not-filing. The deliberation space (S_2) has axes: (loyalty to colleague, duty to institution).

C.2 State Space

The stratified Hilbert space is:

$$\mathcal{H} = \mathcal{H}_{S_2} \oplus \mathcal{H}_\partial \oplus \mathcal{H}_{S_0} \tag{81}$$

Basis states in S_2:

$$|\text{loyalty-high, duty-low}\rangle, |\text{loyalty-low, duty-high}\rangle, \text{etc.} \tag{82}$$

Basis states in S_0:

$$|\text{report}\rangle, |\text{not-report}\rangle \tag{83}$$

C.3 Initial State

Alice and Bob begin in S_2 in a superposition reflecting their uncertainty:

$$|\Psi_0\rangle = \frac{1}{2}\left(|HL\rangle_A|HL\rangle_B + |HL\rangle_A|LH\rangle_B + |LH\rangle_A|HL\rangle_B + |LH\rangle_A|LH\rangle_B\right) \tag{84}$$

where HL = (high loyalty, low duty) and LH = (low loyalty, high duty).

C.4 Transition Amplitude Calculation

To compute the transition amplitude from S_2 to S_0, we use the boundary propagator.

The transition amplitude for Alice to move from state $|HL\rangle$ in S_2 to $|\text{not-report}\rangle$ in S_0 is:

$$\mathcal{A}_{HL\to\text{n-r}} = \langle\text{n-r}|\hat{T}|HL\rangle = \int d\sigma\, \langle\text{n-r}|e^{-i\mathcal{L}_{\text{boundary}}\sigma/\hbar_\eta}|HL\rangle \tag{85}$$

For our toy model, assume the boundary Lagrangian gives:

$$\mathcal{A}_{HL\to\text{n-r}} = 0.9e^{i\phi_1} \tag{86}$$
$$\mathcal{A}_{HL\to\text{report}} = 0.3e^{i\phi_2} \tag{87}$$
$$\mathcal{A}_{LH\to\text{n-r}} = 0.3e^{i\phi_3} \tag{88}$$
$$\mathcal{A}_{LH\to\text{report}} = 0.9e^{i\phi_4} \tag{89}$$

Phases ϕ_i are determined by the action integral along paths through moral space.

C.5 Interference Effect

If Alice is in superposition $\frac{1}{\sqrt{2}}(|HL\rangle + |LH\rangle)$ before transitioning, the amplitude to reach $|\text{report}\rangle$ is:

$$\mathcal{A}_{\to\text{report}} = \frac{1}{\sqrt{2}}\left(0.3e^{i\phi_2} + 0.9e^{i\phi_4}\right) \tag{90}$$

The probability depends on the phase difference $\Delta\phi = \phi_4 - \phi_2$:

$$P_{\text{report}} = \frac{1}{2}\left(0.09 + 0.81 + 2\cdot 0.3\cdot 0.9\cos(\Delta\phi)\right) = 0.45 + 0.27\cos(\Delta\phi) \tag{91}$$

For $\Delta\phi = 0$ (constructive): $P_{\text{report}} = 0.72$

For $\Delta\phi = \pi$ (destructive): $P_{\text{report}} = 0.18$

Interpretation: The moral frameworks (loyalty-based vs. duty-based) interfere constructively or destructively depending on their phase relationship, which is determined by the boundary Lagrangian.

C.6 Entanglement and Collective Responsibility

If Alice and Bob become entangled through joint deliberation (exchanging transition ethons at the boundary), they form the **moral singlet state**:

$$|\Psi^-_{AB}\rangle = \frac{1}{\sqrt{2}} \left(|\text{report}\rangle_A |\text{not-report}\rangle_B - |\text{not-report}\rangle_A |\text{report}\rangle_B\right) \tag{92}$$

Remark C.1 (Why the Singlet?)**.** The antisymmetric (singlet) state arises naturally when Alice and Bob reach opposite conclusions through symmetric deliberation—their moral reasoning paths interfere destructively for same-outcome states. The symmetric (triplet) state $\frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ would arise from constructive interference and has different correlation properties.

This state has the property that measuring Alice's decision *instantaneously* determines Bob's (to be opposite)—not through communication, but through the entanglement structure.

For the singlet state, responsibility measurement in different bases yields the standard correlation:

$$E(\theta_A, \theta_B) = -\cos(\theta_A - \theta_B) \tag{93}$$

For the standard CHSH test, we use angles $a = 0$, $a' = \pi/2$, $b = \pi/4$, $b' = 3\pi/4$:

$$\begin{aligned}
S &= E(a,b) - E(a,b') + E(a',b) + E(a',b') && (94)\\
&= E(0,\pi/4) - E(0,3\pi/4) + E(\pi/2,\pi/4) + E(\pi/2,3\pi/4) && (95)\\
&= -\cos(-\pi/4) - (-\cos(-3\pi/4)) + (-\cos(\pi/4)) + (-\cos(-\pi/4)) && (96)\\
&= -\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2} = -2\sqrt{2} && (97)
\end{aligned}$$

Thus $|S| = 2\sqrt{2} \approx 2.83 > 2$, violating the classical CHSH bound.

Caveat: This is the theoretical maximum. Experimental realizations would show smaller violations due to imperfect state preparation and measurement noise.

C.7 Decoherence to Classical Limit

In the S_2 stratum, decoherence rate is low: $\Gamma_{S_2} = \Gamma_0(d_{\max} - 2 + 1) = \Gamma_0$.

At the boundary, decoherence rate increases: $\Gamma_\partial = \Gamma_0(d_{\max} - 1 + 1) = 2\Gamma_0$.

At S_0, decoherence is maximal: $\Gamma_{S_0} = \Gamma_0(d_{\max} - 0 + 1) = 3\Gamma_0 \to \infty$ as the veto limit is taken.

The density matrix evolves as:

$$\rho(t) = e^{-\Gamma t}\rho_{\text{coherent}} + (1 - e^{-\Gamma t})\rho_{\text{classical}} \tag{98}$$

For $t \gg \Gamma^{-1}$, the state becomes effectively classical: no interference, no superposition, definite outcomes.

C.8 Summary of Toy Model

This minimal example demonstrates:

Phenomenon	Manifestation in Toy Model
Superposition	Initial state in S_2
Transition amplitude	Boundary propagator $\mathcal{A}_{ij}$
Interference	Phase-dependent P_{report}
Entanglement	Bell-correlated $\lvert\Psi_{AB}\rangle$
Decoherence	Γ-dependent classical limit
Stratum-dependence	$\Gamma(S_i) \propto (d_{\max} - d_i + 1)$

D End-to-End Numerical Example

We now work through the toy model with fully instantiated parameter values, showing exactly how SQND is "used" to make a prediction.

D.1 Parameter Values

Suppose calibration experiments have yielded (these are illustrative, not empirically validated):

Parameter	Value	Source
$\hbar_\eta$	1.0 (natural units)	Sets scale
Γ_0	0.5 s^{-1}	Response time data
β	1.0	Complexity scaling
μ_{02}	2.0 $\hbar_\eta$/s	Threshold sharpness
$\phi_{\text{HL}\to\text{report}}$	$\pi/6$	Order effect fits
$\phi_{\text{LH}\to\text{report}}$	$5\pi/6$	Order effect fits
$d_{\max}$	2	Model structure

Psychological grounding for phase values: Why $\phi_{\text{HL}\to\text{report}} = \pi/6$ and $\phi_{\text{LH}\to\text{report}} = 5\pi/6$?

The phase difference $\Delta\phi = 5\pi/6 - \pi/6 = 2\pi/3$ reflects the *psychological tension* between the two framings:

- **Loyalty-based reasoning** (HL): The agent feels *reluctant* about reporting. The path through moral space passes through "betrayal" territory, accumulating negative action. This generates a small positive phase ($\pi/6$).
- **Duty-based reasoning** (LH): The agent feels *obligated* to report. The path passes through "responsibility" territory, accumulating positive action. But crucially, the duty framing also carries awareness of harm to the colleague, generating a large phase ($5\pi/6$) that partially opposes the loyalty path.

The phase difference of $2\pi/3$ (120°) represents *partial opposition*: the framings are neither fully aligned ($\Delta\phi = 0$) nor fully opposed ($\Delta\phi = \pi$). This matches the psychological reality that loyalty and duty considerations in whistleblowing scenarios are in tension but not completely contradictory—both aim at "doing right," but via different moral logics.

General principle: Phases encode the "moral direction" of reasoning paths. Paths that feel psychologically aligned have similar phases; paths that feel opposed have phases differing by $\sim \pi$. This can be calibrated from order-effect experiments where question framing manipulates which path is taken first.

Note: $\phi_{\text{LH}\to\text{report}} - \phi_{\text{HL}\to\text{report}} = 4\pi/6 = 2\pi/3$ (partially destructive).

D.2 Step 1: Initial State Preparation

Alice begins deliberating. Her initial state in S_2:

$$|\Psi_A(0)\rangle = \frac{1}{\sqrt{2}}|HL\rangle + \frac{1}{\sqrt{2}}|LH\rangle \tag{99}$$

This represents equal weight on loyalty-based and duty-based framings.

D.3 Step 2: Compute Transition Amplitudes

Using the calibrated phases and the transition magnitudes from Section C.4:

$$\mathcal{A}_{HL\to\text{report}} = 0.3 \cdot e^{i\pi/6} = 0.3(\cos(\pi/6) + i\sin(\pi/6)) = 0.260 + 0.150i \tag{100}$$

$$\mathcal{A}_{LH\to\text{report}} = 0.9 \cdot e^{i5\pi/6} = 0.9(\cos(5\pi/6) + i\sin(5\pi/6)) = -0.779 + 0.450i \tag{101}$$

D.4 Step 3: Compute Interference

Total amplitude to report:

$$\begin{aligned}
\mathcal{A}_{\text{total}} &= \frac{1}{\sqrt{2}}(\mathcal{A}_{HL\to\text{report}} + \mathcal{A}_{LH\to\text{report}}) && (102)\\
&= \frac{1}{\sqrt{2}}((0.260 - 0.779) + i(0.150 + 0.450)) && (103)\\
&= \frac{1}{\sqrt{2}}(-0.519 + 0.600i) && (104)\\
&= -0.367 + 0.424i && (105)
\end{aligned}$$

Probability:

$$P_{\text{report}} = |\mathcal{A}_{\text{total}}|^2 = (-0.367)^2 + (0.424)^2 = 0.135 + 0.180 = \boxed{0.315} \tag{106}$$

D.5 Step 4: Compare to Classical (No-Interference) Prediction

Without interference, we would add probabilities:

$$P_{\text{classical}} = \frac{1}{2}(0.3^2 + 0.9^2) = \frac{1}{2}(0.09 + 0.81) = \boxed{0.450} \tag{107}$$

Interference effect: $P_{\text{report}}^{\text{quantum}} = 0.315 < 0.450 = P_{\text{report}}^{\text{classical}}$

The partially destructive interference ($\Delta\phi = 2\pi/3$) reduces the reporting probability by 30%.

D.6 Step 5: Include Decoherence

Suppose Alice deliberates for time $t = 2$ seconds before reaching the boundary.

Decoherence rate in S_2: $\Gamma_{S_2} = \Gamma_0 \cdot (2 - 2 + 1)^1 = 0.5\ \text{s}^{-1}$

Coherence survival probability:

$$P_{\text{coherent}} = e^{-\Gamma_{S_2} \cdot t} = e^{-0.5 \cdot 2} = e^{-1} \approx 0.368 \tag{108}$$

Effective probability (mixture of quantum and classical):

$$P_{\text{report}}^{\text{eff}} = P_{\text{coherent}} \cdot P_{\text{report}}^{\text{quantum}} + (1 - P_{\text{coherent}}) \cdot P_{\text{report}}^{\text{classical}} \tag{109}$$

$$= 0.368 \cdot 0.315 + 0.632 \cdot 0.450 \tag{110}$$

$$= 0.116 + 0.284 = \boxed{0.400} \tag{111}$$

D.7 Step 6: Final Prediction

SQND Prediction: Given the calibrated parameters, an agent in equal superposition of loyalty-based and duty-based framings, deliberating for 2 seconds, will report the misconduct with probability **0.400**.

This is:

- 11% lower than the classical prediction (0.450) due to residual interference
- 27% higher than the pure quantum prediction (0.315) due to decoherence

D.8 Sensitivity Analysis

How do predictions change with parameters?

Variation	$P_{\textbf{report}}$	**vs. Baseline**	**Physical Meaning**
Baseline ($t = 2$s, $\Delta\phi = 2\pi/3$)	0.400	—	—
Fast decision ($t = 0.5$s)	0.348	−13%	More interference
Slow decision ($t = 5$s)	0.438	+10%	More decoherence
Aligned phases ($\Delta\phi = 0$)	0.634	+59%	Constructive
Opposed phases ($\Delta\phi = \pi$)	0.246	−39%	Destructive

Key insight: SQND predicts that moral judgments depend not just on the content of moral considerations, but on:

1. Their *phase relationships* (constructive vs. destructive)
2. The *deliberation time* (more time $\rightarrow$ more decoherence $\rightarrow$ more classical)
3. The *stratum structure* (higher dimension $\rightarrow$ slower decoherence)

These are testable predictions that distinguish SQND from classical ethical theories.