Measurements and Erasures in a Universe Governed by Unitary Quantum Mechanics

Paul Cousin

Abstract

In unitary quantum mechanics, the time evolution of a finite-dimensional isolated system corresponds to an orthogonal transformation in a real Hilbert space. All geometric structures defined by constraints are therefore strictly conserved. This conservation has profound implications for the measurement and erasure of a qubit, as we demonstrate that extreme fine-tuning of the system's entire trajectory is required for either of these phenomena to occur. By modeling perfect and imperfect measurements, as well as rederiving Landauer's principle for isolated systems via dimensional constraints, we cast these physical processes in purely geometric terms. Ultimately, we argue that resolving this fine-tuning problem—distinct from the standard arrow of time—requires supplementing unitary mechanics with a past hypothesis that restricts the early universe to a highly specific initial state.

Contents

1. Introduction 1
2. Real space time evolution operator 2
3. Measurements 3
   3.1. Perfect measurements 3
   3.2. Imperfect measurements 5
   3.3. Decoherence 6
4. Erasure 6
   4.1. The thermal approach 6
   4.2. Landauer's principle for isolated systems 9
5. Conclusion 10
6. Acknowledgments 10
References 10

1. Introduction

Applying unitary quantum mechanics to an isolated system, like the universe as a whole, presents some difficulties. One of the prominent puzzles is the perceived asymmetry in time. Whereas our macroscopic experience is marked by an omnipresent arrow of time, unitary quantum mechanics is time-symmetric. The framework must therefore be supplemented with some extra postulate that breaks this symmetry. Postulates that restrict the initial state of the universe are collectively referred to as 'past hypotheses', a term coined by David Albert in [1]. They take several forms in the literature. In the quantum case, the entanglement past hypothesis [2] postulates that the universe started in a low entanglement state, providing a plausible explanation for the decoherent arrow of time. Carlo Rovelli proposed that the arrow of time might be purely perspectival [3], but how such a perspective-bearing spacetime volume arises in the first place remains unclear.

In this paper, we present an adjacent problem rooted in geometry conservation. We will first introduce a real space for quantum states and prove that the time evolution operator in this space is

an orthogonal matrix (Section 2). From this simple fact, we will see that both the measurement (Section 3) and the erasure (Section 4) of a qubit require a fine-tuning of the entire trajectory.

This general approach presents a problem that could be solved either by a past or a future hypothesis. It is therefore distinct from the arrow of time problem. However, because it is natural to think in terms of initial conditions and because this framing provides a clearer, more intuitive language for the proofs, we will primarily employ this perspective in Sections 3 and 4. We will return to the broader implications for the system's entire trajectory in Section 5.

2. Real space time evolution operator

In this section, we provide a simple proof that unitary time evolution preserves all structures imposed on a quantum state in its real-space representation. In group-theoretical terms, this corresponds to the natural embedding of the complex unitary group $U(2^N)$ into the real special orthogonal group $\mathrm{SO}(2^{N+1})$. While this is a well-established mathematical fact, we provide an explicit proof here to build the geometric intuition and establish the notation required for the subsequent sections.

Definition 1: The **real Hilbert space** of a quantum system is the real vector space formed by mapping the complex probability amplitudes of the original complex Hilbert space to their separate real and imaginary components.

Theorem 1: In the real Hilbert space representation, the unitary time-evolution operator corresponds to a special orthogonal transformation.

Proof: Let $\hat{U}(t) = e^{-i\hat{H}t}$ be the unitary time-evolution operator of an isolated system.

$$|\psi(t)\rangle = \hat{U}(t)|\psi(0)\rangle \tag{1}$$

We can separate this operator into its real and imaginary parts, giving

$$\hat{U}(t) = \hat{A} + i\hat{B} \tag{2}$$

where $\hat{A}$ and $\hat{B}$ are real-valued. Now, using unitarity, we see that

$$\mathbb{I} = \hat{U}^\dagger\hat{U} = \left(\hat{A} - i\hat{B}\right)^T\left(\hat{A} + i\hat{B}\right) = \hat{A}^T\hat{A} + \hat{B}^T\hat{B} + i\left(\hat{A}^T\hat{B} - \hat{B}^T\hat{A}\right). \tag{3}$$

Because $\mathbb{I}$ is real-valued, we obtain two relations.

$$\hat{A}^T\hat{A} + \hat{B}^T\hat{B} = \mathbb{I} \tag{4}$$

$$\hat{A}^T\hat{B} = \hat{B}^T\hat{A} \tag{5}$$

In the following, computational basis states will be written using either binary or decimal representations. This choice should always be clear from context. For example:

$$|1\rangle|0\rangle|1\rangle \equiv |101\rangle \equiv |5\rangle. \tag{6}$$

To construct the time-evolution operator in the real space representation of quantum states, it is easier to work in the "stacked" basis (Equation 7).

$$\begin{pmatrix} \Re\langle 0|\psi\rangle \\ \Re\langle 1|\psi\rangle \\ \vdots \\ \Re\langle 2^N-1|\psi\rangle \\ \Im\langle 0|\psi\rangle \\ \vdots \\ \Im\langle 2^N-2|\psi\rangle \\ \Im\langle 2^N-1|\psi\rangle \end{pmatrix} \in \mathbb{R}^{2^{N+1}} \tag{7}$$

Let $\hat{O}(t)$ be the time-evolution operator in this basis.

$$\hat{O} = \begin{pmatrix} \hat{A} & -\hat{B} \\ \hat{B} & \hat{A} \end{pmatrix} \tag{8}$$

We can show that $\hat{O}$ is orthogonal by using Equations 4 and 5 in the following block matrix multiplication.

$$\hat{O}^T\hat{O} = \begin{pmatrix} \hat{A}^T & \hat{B}^T \\ -\hat{B}^T & \hat{A}^T \end{pmatrix} \cdot \begin{pmatrix} \hat{A} & -\hat{B} \\ \hat{B} & \hat{A} \end{pmatrix} \tag{9}$$

$$= \begin{pmatrix} \hat{A}^T\hat{A} + \hat{B}^T\hat{B} & -\hat{A}^T\hat{B} + \hat{B}^T\hat{A} \\ -\hat{B}^T\hat{A} + \hat{A}^T\hat{B} & \hat{B}^T\hat{B} + \hat{A}^T\hat{A} \end{pmatrix} \tag{10}$$

$$= \begin{pmatrix} \mathbb{I}_{2^N} & 0 \\ 0 & \mathbb{I}_{2^N} \end{pmatrix} = \mathbb{I}_{2^{N+1}} \tag{11}$$

□

$\hat{O}$ therefore acts as a rotation in a high-dimensional real Hilbert space. Consequently, any manifold of quantum states will retain its exact geometric structure throughout the dynamics.

3. Measurements

In the unitary quantum mechanics and decoherence literature [4], the measurement of a qubit is often modeled by a simple entanglement process.

$$|\psi\rangle_{\text{system}}|\phi\rangle_{\text{observer}} = (\psi_0|0\rangle_{\text{s}} + \psi_1|1\rangle_{\text{s}})|\phi\rangle_{\text{o}} \quad \longrightarrow \quad \psi_0|00\rangle_{\text{so}} + \psi_1|11\rangle_{\text{so}} \tag{12}$$

Because the dynamics are reversible, this exact process cannot occur. Some "environment" degrees of freedom $|\chi\rangle_{\text{e}}$ are needed to absorb the pre-measurement state of the observer. The global state after a perfect measurement will therefore have the general form of Equation 13.

$$\psi_0|00\rangle_{\text{so}}|\chi_0\rangle_{\text{e}} + \psi_1|11\rangle_{\text{so}}|\chi_1\rangle_{\text{e}} \tag{13}$$

3.1. Perfect measurements

Equation 12 is an example of a perfect measurement (Definition 2). In this section, we will prove that having a perfect measurement performed in an isolated system requires some fine tuning of the initial conditions. This has already been proven in [5], but we will use a different, more general approach.

Definition 2: A **perfect measurement** occurs when the system and observer qubits are perfectly correlated, meaning all cross-terms representing mismatched states are exactly zero.

Theorem 2: In a N-qubit isolated system, the manifold of initial conditions compatible with a perfect measurement occurring at a later time has codimension 2^N in the space of possible states, which is approximately half its dimension for large systems. It is therefore of Lebesgue measure zero.

Proof: Let $|\psi\rangle$ be the state of a N-qubit system. To have a better picture of the distinguishable states of this system, and enable some geometric intuition, we will work in the real Hilbert space representation. Let V be the vector defined as

$$V \equiv \begin{pmatrix} \Re\langle 0|\psi\rangle \\ \Im\langle 0|\psi\rangle \\ \Re\langle 1|\psi\rangle \\ \Im\langle 1|\psi\rangle \\ \vdots \\ \Re\langle 2^N-1|\psi\rangle \\ \Im\langle 2^N-1|\psi\rangle \end{pmatrix} \in \mathbb{R}^{2^{N+1}}. \tag{14}$$

Quantum dynamics are independent of the global phase, or the norm of the state vector. This degeneracy can be removed by the two following constraints.

$$\sum_{i=1}^{2^{N+1}} v_i^2 = 1 \tag{15}$$

$$v_{2^{N+1}} = 0 \tag{16}$$

Equation 15 imposes normalization and Equation 16 selects the phase where $\langle 2^N-1|\psi\rangle$ is real. These two constraints define a manifold M of codimension 2 in $\mathbb{R}^{2^{N+1}}$.

$$\dim(M) = 2^{N+1} - 2 \tag{17}$$

Let the first qubit, represented by the most significant digit in the binary number representation, be the system and the second the observer. If we want $|\psi\rangle$ to be a post-measurement state as described by Equation 13, we have to impose $\langle 01\ldots|\psi\rangle = \langle 10\ldots|\psi\rangle = 0$.

$$v_i = 0 \quad \forall i \in [\![2^{N-1}+1, 3\cdot 2^{N-1}]\!] \tag{18}$$

Because the 2^N constraints defined in Equation 18 are independent from those of Equations 15 and 16, they define a submanifold $\widetilde{M}$ of codimension 2^N in M.

$$\dim\big(\widetilde{M}\big) = 2^{N+1} - 2 - 2^N = 2^N - 2 \tag{19}$$

Let $\hat{O}(t)$ be the real space time evolution operator. As proven in Section 2, it is orthogonal and therefore preserves the dimension of manifolds. The codimension of $\widetilde{M}$ in M and the ratio between their dimensions are also preserved.

$$\frac{\dim\big(\hat{O}(t)\widetilde{M}\big)}{\dim(M)} = \frac{\dim\big(\widetilde{M}(t)\big)}{\dim(M)} = \frac{\dim\big(\widetilde{M}\big)}{\dim(M)} = \frac{2^N-2}{2^{N+1}-2} \underset{N\to\infty}{\longrightarrow} \frac{1}{2} \tag{20}$$

If the measurement happens Δt after the initial conditions are set, then $\widetilde{M}(-\Delta t)$ is the set of initial states that will end up giving a perfect measurement. We have therefore proven the theorem.

□

3.2. Imperfect measurements

However, the dynamics of decoherence do not produce perfect measurements but exponentially decaying cross-terms.

$$\langle 01...|\psi(t)\rangle \quad \propto \quad \langle 10...|\psi(t)\rangle \quad \propto \quad e^{-t} \tag{21}$$

It is therefore useful to consider an ε-neighborhood around $\widetilde{M}$.

$$\widetilde{M}_\varepsilon \equiv \left\{ V \in M \;\middle|\; \sum_{i=2^{N-1}+1}^{3\cdot 2^{N-1}} v_i^2 = P_\mathrm{e} \leq \varepsilon^2 \right\} \tag{22}$$

P_e is the self-locating probability of error of the observer. This quantity defines an imperfect measurement.

Definition 3: An **imperfect measurement** is a measurement with a self-locating probability of error $P_\mathrm{e} \leq \varepsilon^2 < \frac{1}{2}$ in the final state.

Here, we will not work with the dimension of the manifold since allowing an ε-neighborhood will make its codimension null. We will consider its volume instead.

Theorem 3: In a N-qubit isolated system, the space occupied by the manifold of initial conditions compatible with an imperfect measurement occurring at a later time tends to zero as the system becomes large.

Proof: Let us split the coordinates v_i into two groups.
- $X \equiv (v_{2^{N-1}+1}, ..., v_{3\cdot 2^{N-1}}) \in \mathbb{R}^{2^N}$, the "failure" coordinates,
- $Y \equiv (v_1, ..., v_{2^{N-1}}, v_{3\cdot 2^{N-1}+1}, ..., v_{2^{N+1}-1}) \in \mathbb{R}^{2^N-1}$, the "success" coordinates.

These vectors must obey $\|X\|^2 + \|Y\|^2 = 1$ and $\|X\|^2 < \varepsilon^2$.

We can use a coordinate system enforcing the first condition. Let θ be such that $\|X\| = \sin\theta$ and $\|Y\| = \cos\theta$. In this setup, X lives on a (2^N-1)-sphere of radius $\sin\theta$ and Y on a (2^N-2)-sphere of radius $\cos\theta$. To obtain the ratio of volumes $V\left(\widetilde{M}_\varepsilon\right)/V(M)$, we do not need the exact areas of the two spheres, but only how they scale.

$$\frac{V\left(\widetilde{M}_\varepsilon\right)}{V(M)} = \frac{\int_0^{\arcsin\varepsilon} (\sin\theta)^{2^N-1}(\cos\theta)^{2^N-2}\,\mathrm{d}\theta}{\int_0^{\pi/2} (\sin\theta)^{2^N-1}(\cos\theta)^{2^N-2}\,\mathrm{d}\theta} \tag{23}$$

Changing variables for $u = \sin^2\theta$, we obtain

$$\frac{V\left(\widetilde{M}_\varepsilon\right)}{V(M)} = \frac{\int_0^{\varepsilon^2} u^{2^{N-1}-1}(1-u)^{2^{N-1}-\frac{3}{2}}\,\mathrm{d}u}{\int_0^1 u^{2^{N-1}-1}(1-u)^{2^{N-1}-\frac{3}{2}}\,\mathrm{d}u} = I_{\varepsilon^2}\left(2^{N-1}, 2^{N-1}-\frac{1}{2}\right) \tag{24}$$

where I_{ε^2} is the Regularized Incomplete Beta Function.

As before, we obtain

$$\frac{V\left(\hat{O}(t)\widetilde{M}_\varepsilon\right)}{V(M)} = \frac{V\left(\widetilde{M}_\varepsilon(t)\right)}{V(M)} = \frac{V\left(\widetilde{M}_\varepsilon\right)}{V(M)} \quad \underset{N\to\infty}{\longrightarrow} \quad 0. \tag{25}$$

$\widetilde{M}_\varepsilon(-\Delta t)$ being the set of initial states that will end up giving a measurement with a probability of error $P_\mathrm{e} \leq \varepsilon^2$, we have proven the theorem. □

This shows that allowing for imperfect measurements does not solve the fine-tuning problem. Given the size of the universe, it almost does not matter that a measurement be imperfect. As long as the probability of the cross-terms is smaller than a half, only a vanishing portion of the initial states would lead to the desired outcome.

3.3. Decoherence

A measurement requires a mechanism through which entanglement grows. This mechanism is typically called decoherence [6], the simplest model of which is a qubit becoming entangled with its environment.

$$(\psi_0|0\rangle_\mathrm{s} + \psi_1|1\rangle_\mathrm{s})|\chi\rangle_\mathrm{e} \quad\longrightarrow\quad \psi_0|0\rangle_\mathrm{s}|\chi_0\rangle_\mathrm{e} + \psi_1|1\rangle_\mathrm{s}|\chi_1\rangle_\mathrm{e} \tag{26}$$

Let $\frac{\psi_0}{\psi_1} \equiv \alpha + i\beta$. In our real-space representation, ensuring the system is a tensor product in its initial state requires imposing 2^N real constraints coupling the real and imaginary components:

$$v_i = \begin{cases} \alpha\, v_{i+2^N} - \beta\, v_{i+2^N+1} & \text{if } i \text{ is odd} \\ \beta\, v_{i+2^N-1} + \alpha\, v_{i+2^N} & \text{if } i \text{ is even} \end{cases} \qquad \forall i \in [\![1, 2^N]\!] \tag{27}$$

The codimension in M of the manifold defined by these 2^N constraints is $2^N - 2$ due to the freedom allowed by α and β. We therefore find a 2^N-dimensional submanifold for the initial state, whereas the final state seems completely generic in M. However, because the time evolution operator $\hat{O}(t)$ is an orthogonal matrix, the final state cannot be generic. It has to live on a very specific manifold of the same dimension and shape as the initial state.

4. Erasure

Let us now turn to the erasure of a qubit as described by Equation 28.

$$a_0|0\rangle_\mathrm{s} + a_1|1\rangle_\mathrm{s} \quad\longrightarrow\quad b_0|0\rangle_\mathrm{s} \tag{28}$$

Landauer's principle [7] states that the minimum energy required to erase one bit of information is $k_\mathrm{B}T\ln(2)$. This quantity strongly suggests working with a thermal state. We will derive this limit when using a N-qubit thermal bath to perform the erasure.

4.1. The thermal approach

Let $\hat{\rho}^i_\mathrm{bath}$ be the initial density matrix of a bath of non-interacting qubits with energy 0 for $|0\rangle$ and energy $E > 0$ for $|1\rangle$.

$$\hat{\rho}^i_\mathrm{bath} = \frac{1}{\mathcal{Z}} e^{-\beta \hat{H}_\mathrm{bath}} = \sum_{i_1,\ldots,i_N} \frac{1}{\mathcal{Z}} e^{-\beta E \sum_{k=1}^N i_k} |i_1, \ldots, i_N\rangle\langle i_1, \ldots, i_N| \qquad \text{with} \quad \beta = \frac{1}{k_\mathrm{B}T} \tag{29}$$

The partition function $\mathcal{Z}$ can be expressed as

$$\mathcal{Z} = \sum_{i_1,\ldots,i_N} e^{-\beta E \sum_{k=1}^N i_k} = \sum_{k=0}^{N} \binom{N}{k} e^{-\beta E k} = \left(1 + e^{-\beta E}\right)^N \tag{30}$$

and the initial expected energy E^i_bath as

$$E^i_\mathrm{bath} = \mathrm{Tr}\left(\hat{\rho}^i_\mathrm{bath}\hat{H}_\mathrm{bath}\right) = -\frac{\partial \ln \mathcal{Z}}{\partial \beta} = \frac{NE}{1 + e^{\beta E}}. \tag{31}$$

We obtain the expected limits.

$$\lim_{T\to 0} E^i_{\text{bath}} = 0 \tag{32}$$

$$\lim_{T\to \infty} E^i_{\text{bath}} = \frac{NE}{2} \tag{33}$$

We want to erase an unknown qubit with initial density matrix

$$\hat{\rho}^i_0 = \frac{1}{2}|\uparrow\rangle\langle\uparrow| + \frac{1}{2}|\downarrow\rangle\langle\downarrow|, \tag{34}$$

which is equivalent to setting $\langle\uparrow|\hat{H}_0|\uparrow\rangle = \langle\downarrow|\hat{H}_0|\downarrow\rangle$.

The full initial density matrix will be given by

$$\hat{\rho}^i_{\text{full}} = \hat{\rho}^i_0 \otimes \hat{\rho}^i_{\text{bath}}. \tag{35}$$

The Von Neumann entropy of the system is

$$S_{\text{full}} = -\operatorname{Tr}\left(\hat{\rho}^i_{\text{full}} \ln \hat{\rho}^i_{\text{full}}\right) = S^i_0 + S^i_{\text{bath}} \tag{36}$$

with

$$S^i_0 = -\operatorname{Tr}(\hat{\rho}^i_0 \ln \hat{\rho}^i_0) = \ln(2) \tag{37}$$

and

$$S^i_{\text{bath}} = -\operatorname{Tr}\left(\hat{\rho}^i_{\text{bath}} \ln \hat{\rho}^i_{\text{bath}}\right) \tag{38}$$

$$= -\operatorname{Tr}\left(\hat{\rho}^i_{\text{bath}} \ln \frac{e^{-\beta \hat{H}_{\text{bath}}}}{\mathcal{Z}}\right) \tag{39}$$

$$= \beta \operatorname{Tr}\left(\hat{\rho}^i_{\text{bath}} \hat{H}_{\text{bath}}\right) + \operatorname{Tr}\left(\hat{\rho}^i_{\text{bath}}\right) \ln \mathcal{Z} \tag{40}$$

$$= \beta E^i_{\text{bath}} + \ln \mathcal{Z}. \tag{41}$$

We can define a permutation matrix $\hat{U}$ that sorts states by ascending energies. Because the state of the first qubit has no impact on the total energy, every energy level is at least doubly-degenerate. Therefore, many different matrices exhibit this sorting behavior. Let $\hat{U}$ be any one of them.

$$\hat{U}\begin{pmatrix} \langle\uparrow 000...000|\Psi\rangle \\ \langle\uparrow 000...001|\Psi\rangle \\ \langle\uparrow 000...010|\Psi\rangle \\ \vdots \\ \langle\downarrow 111...101|\Psi\rangle \\ \langle\downarrow 111...110|\Psi\rangle \\ \langle\downarrow 111...111|\Psi\rangle \end{pmatrix} = \begin{pmatrix} \langle\uparrow 000...000|\Psi\rangle \\ \langle\downarrow 000...000|\Psi\rangle \\ \langle\uparrow 100...000|\Psi\rangle \\ \vdots \\ \langle\downarrow 011...111|\Psi\rangle \\ \langle\uparrow 111...111|\Psi\rangle \\ \langle\downarrow 111...111|\Psi\rangle \end{pmatrix} \tag{42}$$

We now want to show that the probability $p^f_\downarrow$ of measuring $|\downarrow\rangle$ after applying $\hat{U}$ goes to zero as N becomes large. Assuming N is even for simplicity, we have

$$p^f_\downarrow = \frac{1}{\mathcal{Z}} \sum_{k=N/2}^{N} \binom{N}{k} \frac{e^{-\beta E k}}{2^{\mathbb{1}[k=N/2]}} \leq \frac{1}{\mathcal{Z}} \sum_{k=N/2}^{N} \binom{N}{k} e^{-\beta E k} \tag{43}$$

with $\mathbb{1}$ the indicator function.

$$\mathbb{1}[k = N/2] = \begin{cases} 1 & \text{if } k = N/2 \\ 0 & \text{otherwise} \end{cases} \tag{44}$$

Inserting the expression for the partition function (Equation 30), we obtain

$$p_{\downarrow}^{f} \leq \frac{1}{\left(1+e^{-\beta E}\right)^{N}} \sum_{k=N/2}^{N} \binom{N}{k} e^{-\beta E k} = \sum_{k=N/2}^{N} \binom{N}{k} \left(\frac{e^{-\beta E}}{1+e^{-\beta E}}\right)^{k} \left(\frac{1}{1+e^{-\beta E}}\right)^{N-k} \tag{45}$$

where we recognize the binomial distribution. Defining the probability of success as $p = \frac{e^{-\beta E}}{1+e^{-\beta E}} < \frac{1}{2}$ and the probability of failure as $q = 1 - p$, we get

$$p_{\downarrow}^{f} \leq \sum_{k=N/2}^{N} \binom{N}{k} p^{k} q^{N-k} = q^{N} \sum_{k=N/2}^{N} \binom{N}{k} \left(\frac{p}{q}\right)^{k} \tag{46}$$

$$\leq q^{N} \left(\frac{p}{q}\right)^{\frac{N}{2}} \sum_{k=N/2}^{N} \binom{N}{k} \tag{47}$$

$$\leq (\sqrt{qp})^{N} \sum_{k=0}^{N} \binom{N}{k} = (2\sqrt{qp})^{N} \tag{48}$$

Now, because $(p+q)^2 = 1$ and $(p-q)^2 > 0$, we have

$$1 = (p+q)^2 = p^2 + q^2 + 2pq = (p-q)^2 + 2pq + 2pq > 4pq \quad \Rightarrow \quad 1 > 2\sqrt{pq} \tag{49}$$

and we have therefore confirmed that $\lim_{N\to\infty} p_{\downarrow}^{f} = 0$.

$$\hat{U}\hat{\rho}_{\text{full}}^{i}\hat{U}^{\dagger} = \hat{\rho}_{\text{full}}^{f} \quad \underset{N\to\infty}{\longrightarrow} \quad |\uparrow\rangle\langle\uparrow| \otimes \hat{\rho}_{\text{bath}}^{f} \tag{50}$$

In this limit, the Von Neumann entropy of the system can be expressed as

$$S_{\text{full}} = -\operatorname{Tr}\left(\hat{\rho}_{\text{full}}^{f} \ln \hat{\rho}_{\text{full}}^{f}\right) \quad \underset{N\to\infty}{\longrightarrow} \quad 0 + S_{\text{bath}}^{f} \tag{51}$$

Because $\hat{U}$ is unitary, S_{full} is constant. Therefore, in the limit,

$$\ln(2) + S_{\text{bath}}^{i} = 0 + S_{\text{bath}}^{f} \qquad \Rightarrow \qquad \Delta S_{\text{bath}} \equiv S_{\text{bath}}^{f} - S_{\text{bath}}^{i} = \ln(2) \tag{52}$$

We now want to compute the required change in the bath's energy ΔE_{bath}. To do so, we use the quantum relative entropy, also known as the Kullback-Leibler divergence.

The relative entropy between the exact final state of the bath $\hat{\rho}_{\text{bath}}^{f}$ and the initial thermal equilibrium state $\hat{\rho}_{\text{bath}}^{i} = \frac{e^{-\beta \hat{H}_{\text{bath}}}}{\mathcal{Z}}$ is strictly non-negative:

$$D\left(\hat{\rho}_{\text{bath}}^{f} \middle\| \hat{\rho}_{\text{bath}}^{i}\right) = \operatorname{Tr}\left(\hat{\rho}_{\text{bath}}^{f} \ln \hat{\rho}_{\text{bath}}^{f}\right) - \operatorname{Tr}\left(\hat{\rho}_{\text{bath}}^{f} \ln \hat{\rho}_{\text{bath}}^{i}\right) \geq 0. \tag{53}$$

Using the definition of the Von Neumann entropy $S(\hat{\rho}) = -\operatorname{Tr}(\hat{\rho}\ln\hat{\rho})$, the first term is simply $-S_{\text{bath}}^{f}$. We expand the second term using the definition of the thermal state:

$$-\operatorname{Tr}\left(\hat{\rho}_{\text{bath}}^{f} \ln\left(\frac{e^{-\beta \hat{H}_{\text{bath}}}}{\mathcal{Z}}\right)\right) = \beta \operatorname{Tr}\left(\hat{\rho}_{\text{bath}}^{f} \hat{H}_{\text{bath}}\right) + \operatorname{Tr}\left(\hat{\rho}_{\text{bath}}^{f}\right) \ln \mathcal{Z} \tag{54}$$

$$= \beta E_{\text{bath}}^{f} + \ln \mathcal{Z}. \tag{55}$$

From Equation 41, we know the initial thermal state's entropy is $S_{\text{bath}}^{i} = \beta E_{\text{bath}}^{i} + \ln \mathcal{Z}$. Expressing $\ln \mathcal{Z}$ in these terms and substituting it gives:

$$-\operatorname{Tr}\left(\hat{\rho}_{\text{bath}}^{f} \ln \hat{\rho}_{\text{bath}}^{i}\right) = \beta E_{\text{bath}}^{f} + S_{\text{bath}}^{i} - \beta E_{\text{bath}}^{i} = \beta \Delta E_{\text{bath}} + S_{\text{bath}}^{i}. \tag{56}$$

Inserting these terms back into the relative entropy inequality yields:

$$D\big(\hat{\rho}^f_{\text{bath}} \| \hat{\rho}^i_{\text{bath}}\big) = -S^f_{\text{bath}} + \beta \Delta E_{\text{bath}} + S^i_{\text{bath}} \geq 0. \tag{57}$$

Rearranging for the change in energy gives the exact bound:

$$\beta \Delta E_{\text{bath}} \geq S^f_{\text{bath}} - S^i_{\text{bath}} = \Delta S_{\text{bath}}. \tag{58}$$

Taking the limit as $N \to \infty$ and substituting the entropy change $\Delta S_{\text{bath}} = \ln(2)$ from Equation 52, yields Landauer's limit:

$$\Delta E_{\text{bath}} \geq \frac{1}{\beta} \ln(2) = k_{\text{B}} T \ln(2). \tag{59}$$

4.2. Landauer's principle for isolated systems

The issue with the derivation of Landauer's principle presented in the previous section is that it relies on density matrices. Classical probabilities make sense when talking about subsystems. However, if our goal is to model the universe as a whole, classical probabilities are only expected to arise as an emergent phenomenon: there is no one outside the universe to have uncertainty about its composition. Here, we will present a different approach that casts Landauer's principle in terms that are compatible with a full-universe model.

Consider an initial state analogous to the pre-measurement state defined in Section 3, but now the observer / experimenter wants to erase the first qubit. We will use a structure similar to one presented in [5] but for a different purpose.

Let $|\Psi_i\rangle$ be the initial state and $\big|\Psi_f\big\rangle$ the final state.

$$|\Psi_i\rangle = (a_0|0\rangle_{\text{s}} + a_1|1\rangle_{\text{s}})|\phi\rangle_{\text{o}}|\chi_i\rangle_{\text{e}} \quad \longrightarrow \quad \big|\Psi_f\big\rangle = (b_0|0\rangle_{\text{s}})|\phi\rangle_{\text{o}}\big|\chi_f\big\rangle_{\text{e}} \tag{60}$$

Expressing the possible final states $\big|\Psi_f\big\rangle$ in term of constraints gives the 2^N constraints of Equation 61.

$$v_i = 0 \quad \forall i \in [\![2^N + 1, 2^{N+1}]\!] \tag{61}$$

The post-erasure state therefore lives, as before, on a $(2^N - 2)$-dimensional manifold that has codimension 2^N in M. Now, since this dimension is preserved under the dynamics, the manifold of states that are compatible with the erasure will have the same dimension at all times.

$$\dim(M) - \dim\Big\{\hat{O}(t)V \;\Big|\; V \in M, \quad v_i = 0 \;\; \forall i \in [\![2^N + 1, 2^{N+1}]\!]\Big\} = 2^N \tag{62}$$

Going back to the language of quantum mechanics, we have

$$\begin{aligned} \dim\{|\Psi_i\rangle\} &= \dim\{a_0|0\rangle_{\text{s}} + a_1|1\rangle_{\text{s}}\} \cdot \dim\{|\phi\rangle_{\text{o}}\} \cdot \dim\Big\{\big|\chi_i\big\rangle_{\text{e}}\Big\} \\ \dim\Big\{\big|\Psi_f\big\rangle\Big\} &= \dim\{b_0|0\rangle_{\text{s}}\} \cdot \dim\{|\phi\rangle_{\text{o}}\} \cdot \dim\Big\{\big|\chi_f\big\rangle_{\text{e}}\Big\} \end{aligned} \tag{63}$$

with

$$\dim\{a_0|0\rangle_{\text{s}} + a_1|1\rangle_{\text{s}}\} = \dim\{\Re(a_0), \Im(a_0), \Re(a_1), \Im(a_1)\} = 4 \tag{64}$$

$$\dim\{b_0|0\rangle_{\text{s}}\} = \dim\{\Re(b_0), \Im(b_0)\} = 2 \tag{65}$$

where we remember that the normalization constraint is imposed externally, on the entire state space.

Now, equating the dimension of the space of possible initial and final states, we obtain a relation between the dimensions of the space of possible environment states.

$$\dim\{|\Psi_i\rangle\} = \dim\{|\Psi_f\rangle\} \qquad \Rightarrow \qquad 2 \cdot \dim\Big\{|\chi_i\rangle_{\mathrm{e}}\Big\} = \dim\Big\{|\chi_f\rangle_{\mathrm{e}}\Big\} \tag{66}$$

If we define the entropy of the environment as $S = \ln(\dim\{|\chi\rangle_{\mathrm{e}}\})$, in the spirit of Boltzmann's formula $S = k_{\mathrm{B}} \ln(W)$, we obtain

$$\begin{aligned} \Delta S &= \ln\Big(\dim\Big\{|\chi_f\rangle_{\mathrm{e}}\Big\}\Big) - \ln\Big(\dim\Big\{|\chi_i\rangle_{\mathrm{e}}\Big\}\Big) \\ &= \ln\Big(2 \cdot \dim\Big\{|\chi_i\rangle_{\mathrm{e}}\Big\}\Big) - \ln\Big(\dim\Big\{|\chi_i\rangle_{\mathrm{e}}\Big\}\Big) \\ &= \ln(2) \end{aligned} \tag{67}$$

This framework provides a derivation very close to Landauer's principle, but only in terms of constraints. We think this is the way to approach erasure in a full-universe model.

5. Conclusion

In this paper, we have demonstrated that the constraints imposed on a quantum state at any given time propagate throughout its entire trajectory due to the unitarity of the time-evolution operator. Consequently, the dimensionality of the allowed state space before and after a measurement—or an erasure—remains invariant. Because post-measurement and post-erasure states occupy a highly restricted manifold within the total state space, the initial conditions must be equally fine-tuned.

If unitary quantum mechanics is to provide a complete description of our universe, it must be supplemented by an additional postulate—such as a past hypothesis—to account for the macroscopic complexity we observe. The measurement and erasure of a single qubit represent merely minimalist constraints. In reality, we observe cosmological-scale decoherence leading to classicality, where macroscopic events generate a vast proliferation of physical records. Accounting for these complex dynamics implies that the universe's initial conditions must be constrained to an extraordinarily specific submanifold.

While the geometric arguments presented here are time-symmetric and agnostic regarding where along the trajectory constraints should be imposed, there are compelling empirical reasons to focus on initial conditions. Constraining the universe's origin—often conceptualized as a low-entropy state—provides a simpler and more natural boundary condition than enforcing extreme fine-tuning on its future. In our geometric framework, this suggests that the early universe naturally occupied an extraordinarily specific, low-dimensional submanifold.

6. Acknowledgments

This work was conducted during a three-month internship at the Max Planck Institute for Astrophysics, as part of my final year in the international master's program in Physics of Complex Systems at Université Paris Cité. I would like to express my sincere gratitude to the entire Information Field Theory group for providing such a welcoming environment. Special thanks go to Torsten Enßlin for our stimulating discussions, and to Vishal Johnson, whose collaboration profoundly shaped my experience. Our regular meetings were immensely valuable.

References

[1] D. Z. Albert, *Time and chance*. Harvard University Press, 2003.

[2] J. Al-Khalili and E. K. Chen, "The decoherent arrow of time and the entanglement past hypothesis," *Foundations of Physics*, vol. 54, no. 4, p. 49, 2024.

[3] C. Rovelli, "Is time's arrow perspectival," *The philosophy of cosmology*, pp. 285–296, 2017.

[4] W. H. Zurek, "Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse?," *Physical review D*, vol. 24, no. 6, p. 1516, 1981.

[5] V. Johnson, A. Singh, R. Leike, P. Frank, and T. Enßlin, "Correlation as a Resource in Unitary Quantum Measurements," *arXiv*, 2022, doi: 10.48550/arxiv.2212.03829.

[6] W. H. Zurek, "Decoherence, einselection, and the quantum origins of the classical," *Reviews of modern physics*, vol. 75, no. 3, p. 715, 2003.

[7] R. Landauer, "Irreversibility and heat generation in the computing process," *IBM journal of research and development*, vol. 5, no. 3, pp. 183–191, 1961.