is that when we affirmatively measure $|\psi\rangle$ to have that property, the $|\psi\rangle$ assumes the new state $\frac{P|\psi\rangle}{\sqrt{\langle\psi|P|\psi\rangle}}$. This state now *has* the property, since

$$P \cdot \frac{P|\psi\rangle}{\sqrt{\langle\psi|P|\psi\rangle}} = \frac{P|\psi\rangle}{\sqrt{\langle\psi|P|\psi\rangle}}$$

Said another way, if we measure a state to affirmatively have a property (whether or not it definitely had the property before), it subsequently *assumes* that property. This is different from classical mechanics: for example, classical mechanics stipulates that if we measure a particle to have position x then it definitely had position x before. In quantum mechanics, by contrast, measuring a particle to be in position x just tells us that the particle is in position x now, even though it might not 'definitively' have had that property before.

We notice another peculiarity of the fourth axiom, which is that the map

$$|\psi\rangle \longmapsto \frac{P|\psi\rangle}{\sqrt{\langle\psi|P|\psi\rangle}} \tag{11}$$

is not in general unitary (unless $P = \mathbb{1}$ in which case the map is the identity since $|\psi\rangle$ has unit norm). This would appear to violate the second axiom, which necessitates unitary dynamics. However, we were careful in the second axiom to specify that unitary dynamics happens for *closed* systems; in ordinary circumstances, the measurement apparatus is external to the system that it interrogates, and so the non-unitary of (11) is not in conflict with the second axiom. However, the fourth axiom tempts us to consider the following: if we described the detector (which itself is quantum-mechanical) as *part of* the closed system, then the total detector-system dynamics must be unitary; then can the fourth axiom somehow be derived from the other three? This question is both challenging and profound. Its core difficulty is that the first three axioms do not speak of probability whereas the fourth axioms does speak of probability; as such, the question posed would mandate that probability is *emergent* in quantum mechanics. There have been a vast number of attempts to weaken the fourth axiom or to in some sense 'derive' it from the other three (which often involves covertly bringing in a weakening of the fourth axiom anyway). For our purposes, we can think of the fourth axiom is *pragmatic*, in that it tells us what happens, *in practice*, when we measure a quantum system with an external measurement device.[7]

Having abstractly discussed the axioms, some examples are in order.

Example 9 (Dynamics and projective measurements for a single qubit). We work in the two-dimensional Hilbert space $\mathcal{H} \simeq \mathbb{C}^2$ with the *computational basis*

[7]Related to the previous footnote, we might wonder how we can test quantum mechanics as a theory if we require quantum theory to build the measurement apparatus needed for the tests themselves. As before, the answer is that we are testing the *consistency* of quantum mechanics, and its alignment with empirical reality. One cannot generally test quantum mechanics with detectors solely intelligible through Newtonian mechanics, i.e. you cannot solely use classical to test quantum (see [**Mah18**] for a quantum cryptographic wrinkle in this story). But it is fine to use quantum to test quantum, so long as it all works out empirically. And it very much does.

$\{|0\rangle, |1\rangle\}$ where $|0\rangle := \begin{bmatrix}1\\0\end{bmatrix}$ and $|1\rangle := \begin{bmatrix}0\\1\end{bmatrix}$. We introduce the **Pauli matrices**[8]

$$X = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} = |0\rangle\langle 1| + |1\rangle\langle 0|$$
$$Y = \begin{bmatrix}0 & -i\\ i & 0\end{bmatrix} = -i\,|0\rangle\langle 1| + i\,|1\rangle\langle 0|$$
$$Z = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} = |0\rangle\langle 0| - |1\rangle\langle 1|.$$

They are Hermitian, satisfy $X^2 = Y^2 = Z^2 = \mathbb{1}$, and obey

$$[\sigma_j, \sigma_k] = 2i\,\varepsilon_{jk\ell}\,\sigma_\ell, \qquad \{\sigma_j, \sigma_k\} = 2\,\delta_{jk}\,\mathbb{1},$$

where $(\sigma_1, \sigma_2, \sigma_3) = (X, Y, Z)$. Their eigenvalues are ± 1, with $Z|0\rangle = |0\rangle$ and $Z|1\rangle = -|1\rangle$.

Measuring "spin along z" corresponds to the compatible, complete pair of projectors

$$P_0 = |0\rangle\langle 0| = \frac{\mathbb{1} + Z}{2}, \qquad P_1 = |1\rangle\langle 1| = \frac{\mathbb{1} - Z}{2}.$$

Likewise, "spin along x" has eigenstates $|\pm\rangle := \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$ with projectors

$$P_+^{(x)} = |+\rangle\langle +| = \frac{\mathbb{1} + X}{2}, \qquad P_-^{(x)} = |-\rangle\langle -| = \frac{\mathbb{1} - X}{2}.$$

More generally, for any unit vector $\hat{n} = (n_x, n_y, n_z) \in \mathbb{R}^3$ we have

$$P_\pm^{(\hat{n})} = \frac{\mathbb{1} \pm \hat{n}\cdot\vec{\sigma}}{2}, \qquad \hat{n}\cdot\vec{\sigma} := n_x X + n_y Y + n_z Z,$$

which indeed satisfy the properties of projectors.

For dynamics, consider unitary rotations generated by the Pauli matrices. For any unit vector $\hat{n}$ and real angle θ, define

$$R_{\hat{n}}(\theta) := \exp\left(-i\,\frac{\theta}{2}\,\hat{n}\cdot\vec{\sigma}\right) = \cos\left(\frac{\theta}{2}\right)\mathbb{1} - i\,\sin\left(\frac{\theta}{2}\right)\hat{n}\cdot\vec{\sigma}.$$

Physically, $R_{\hat{n}}(\theta)$ is the time-t propagator of a closed qubit with Hamiltonian $H = \frac{\Omega}{2}\,\hat{n}\cdot\vec{\sigma}$ and $\theta = \Omega\, t$. That is, $R_{\hat{n}}(\theta)$ can be written as e^{-iHt} for the above choices of H and t.

Suppose we prepare the qubit in the $+1$ eigenstate of Z, namely $|\psi_0\rangle = |0\rangle$. If the system evolves under the Hamiltonian $H = \frac{\Omega}{2}Y$ for time t, the unitary $U(t) = R_y(\theta)$ acts with $\theta = \Omega\, t$. Acting on $|0\rangle$ and using $Y|0\rangle = i|1\rangle$, the evolved state is

$$|\psi_t\rangle = \cos\left(\tfrac{\theta}{2}\right)|0\rangle + \sin\left(\tfrac{\theta}{2}\right)|1\rangle.$$

Now consider measuring in the Z basis. The Born rule with projectors P_0, P_1 gives

$$p_Z(0\,|\,t) = \cos^2\left(\tfrac{\theta}{2}\right), \qquad p_Z(1\,|\,t) = \sin^2\left(\tfrac{\theta}{2}\right).$$

[8]The hardest one to remember is Y, in particular the placement of the minus sign in the matrix elements. High energy physicist Howard Georgi has a useful mnemonic: the 'minus i' is lighter so it floats all the way to the top. Now hopefully you will never forget where the minus sign goes.

If outcome 0 is observed, the state collapses to $|0\rangle$; if outcome 1 is observed, it collapses to $|1\rangle$.

If instead we measure in the X basis, the probabilities are

$$p_X(\pm \,|\, t) = \tfrac{1}{2}(1 \pm \langle \psi_t | X | \psi_t \rangle)\,.$$

Since $\langle \psi_t | X | \psi_t \rangle = \sin\theta$, we find

$$p_X(+ \,|\, t) = \tfrac{1+\sin\theta}{2}, \qquad p_X(- \,|\, t) = \tfrac{1-\sin\theta}{2}.$$

To connect with the Bloch sphere, define for any $|\psi\rangle$ the triple

$$\vec{r} = (\langle X \rangle, \langle Y \rangle, \langle Z \rangle) \in \mathbb{R}^3.$$

For the state $|\psi_t\rangle$, we obtain $\vec{r}(t) = (\sin\theta, 0, \cos\theta)$, a unit vector rotating about the y-axis. The Born rule in this language becomes

$$\Pr[\text{outcome } \pm \text{ along } \hat{n}] = \frac{1 \pm \hat{n}\cdot\vec{r}}{2}.$$

Example 10 (Quantum Zeno effect for a single qubit). We revisit the single-qubit system in $\mathcal{H} \simeq \mathbb{C}^2$ with computational basis $\{|0\rangle, |1\rangle\}$ and Pauli matrices X, Y, Z. As a baseline, take the closed-system Hamiltonian to point along y,

$$H = \frac{\Omega}{2}\, Y,$$

and prepare the qubit in $|\psi_0\rangle = |0\rangle$, the $+1$ eigenstate of Z. Under uninterrupted unitary evolution for time T, the state rotates as in Example 9:

$$|\psi_T\rangle = \exp\Bigl(-i\,\tfrac{\Omega T}{2}\,Y\Bigr)|0\rangle = \cos\Bigl(\tfrac{\Omega T}{2}\Bigr)|0\rangle + \sin\Bigl(\tfrac{\Omega T}{2}\Bigr)|1\rangle,$$

so a single projective measurement of Z at time T returns outcome 0 with probability $\cos^2(\Omega T/2)$ and outcome 1 with probability $\sin^2(\Omega T/2)$.

To exhibit the **quantum Zeno effect**, we now intersperse frequent, strong (projective) measurements in the Z basis during the evolution. Fix a total observation time T and an integer $N \geq 1$. Partition the interval into N equal steps of duration $\tau := T/N$, and at the end of each step perform the two-outcome measurement $\{P_0, P_1\}$ with

$$P_0 = \frac{\mathbb{1} + Z}{2} = |0\rangle\langle 0|, \qquad P_1 = \frac{\mathbb{1} - Z}{2} = |1\rangle\langle 1|.$$

Between checks, the system evolves unitarily for time τ. After a single step, the unmeasured state is

$$U(\tau)|0\rangle = \cos\Bigl(\tfrac{\Omega\tau}{2}\Bigr)|0\rangle + \sin\Bigl(\tfrac{\Omega\tau}{2}\Bigr)|1\rangle, \qquad U(\tau) = \exp\Bigl(-i\,\tfrac{\Omega\tau}{2}\,Y\Bigr).$$

Measuring Z now yields "still in $|0\rangle$" with probability

$$p_{\text{stay}}(\tau) = \cos^2\Bigl(\tfrac{\Omega\tau}{2}\Bigr),$$

and, conditioned on that outcome, the post-measurement state collapses back to $|0\rangle$. Because the state is reset to $|0\rangle$ after each successful check, the trials are identical and independent. Iterating N times, the survival probability (the chance to see outcome 0 at *every* check, hence to remain in $|0\rangle$ throughout) is therefore

$$p_{\text{Zeno}}(T, N) = \Bigl[\cos^2\Bigl(\tfrac{\Omega\tau}{2}\Bigr)\Bigr]^N = \Bigl[\cos^2\Bigl(\tfrac{\Omega T}{2N}\Bigr)\Bigr]^N.$$

A short-time expansion makes the scaling explicit. Since $\cos^2 x = 1 - x^2 + O(x^4)$, we have

$$\cos^2\Big(\tfrac{\Omega T}{2N}\Big) = 1 - \frac{\Omega^2 T^2}{4N^2} + O\Big(\frac{1}{N^4}\Big),$$

and thus

$$\log p_{\text{Zeno}}(T, N) = N \log\bigg(1 - \frac{\Omega^2 T^2}{4N^2} + O\Big(\frac{1}{N^4}\Big)\bigg) = -\frac{\Omega^2 T^2}{4N} + O\Big(\frac{1}{N^3}\Big).$$

Exponentiating yields the large-N behavior

$$p_{\text{Zeno}}(T, N) = 1 - \frac{\Omega^2 T^2}{4N} + O\Big(\frac{1}{N^2}\Big) \xrightarrow{N\to\infty} 1\,.$$

In words, for fixed total time T, increasing the measurement frequency ($N \uparrow$) drives the survival probability arbitrarily close to one.

It is also illuminating to track the complementary "flip" probability per step. Each interrogation has

$$\varepsilon_{\text{flip}}(\tau) = \sin^2\Big(\tfrac{\Omega\tau}{2}\Big) = \frac{\Omega^2\tau^2}{4} + O(\tau^4),$$

so the expected number of flips over N checks scales like $N\,\varepsilon_{\text{flip}}(\tau) \approx \frac{\Omega^2 T^2}{4N} \to 0$. This connects back to our discussion of two-state Markov dynamics in Example 1: frequent projective "monitoring" makes the effective flip probability per step vanish quadratically in the step size, and the cumulative chance of ever flipping by time T is correspondingly suppressed.

Armed with our basic examples, we next examine some additional mathematical structures in quantum mechanics.

2.3. Additional mathematical structures

Here we will introduce some additional mathematical apparatus which we can view as additional tools for the applications of the axioms of quantum mechanics presented above.

2.3.1. *Tensor products and density matrices*

We now carry the tensor–product technology into the quantum setting and introduce the operator language that lets us handle classical uncertainty and open–system effects in a clean way. When two systems are modeled by Hilbert spaces $\mathcal{H}_A \simeq \mathbb{C}^{N_A}$ and $\mathcal{H}_B \simeq \mathbb{C}^{N_B}$, their composite is described by the tensor product

$$\mathcal{H}_{AB} := \mathcal{H}_A \otimes \mathcal{H}_B \simeq \mathbb{C}^{N_A N_B}.$$

Choose orthonormal bases $\{|i\rangle_A\}_{i=1}^{N_A}$ and $\{|j\rangle_B\}_{j=1}^{N_B}$. The product kets $\{|i\rangle_A \otimes |j\rangle_B\}_{i,j}$ form an orthonormal basis of $\mathcal{H}_{AB}$. As in the classical case, linear maps respect tensoring. If X_A acts on $\mathcal{H}_A$ and Y_B acts on $\mathcal{H}_B$, then

$$(X_A \otimes Y_B)\big(|\psi\rangle_A \otimes |\phi\rangle_B\big) = (X_A|\psi\rangle_A) \otimes (Y_B|\phi\rangle_B).$$

Operations on a single part are written $X_A \otimes \mathbb{1}_B$ or $\mathbb{1}_A \otimes Y_B$.

A pure state $|\Psi\rangle \in \mathcal{H}_{AB}$ is called a **product state** if it factors as $|\Psi\rangle = |\psi\rangle_A \otimes |\phi\rangle_B$. Otherwise it is **entangled**. The following normal form is indispensable.

Theorem 32 (Schmidt decomposition)**.** *For any unit vector $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ there exist orthonormal sets $\{|k\rangle_A\}$ and $\{|k\rangle_B\}$ together with nonnegative numbers $\{\lambda_k\}$ that sum to one such that*

$$|\Psi\rangle = \sum_{k=1}^{r} \sqrt{\lambda_k}\, |k\rangle_A \otimes |k\rangle_B, \quad r \leq \min\{N_A, N_B\}\,.$$

*The number r is uniquely defined and is called the **Schmidt rank**.*

This is really just another way of stating the linear algebraic fact that every $N_A \times N_B$ matrix (in this case the entries of $|\Psi\rangle$ reshaped into such a matrix) has a singular value decomposition, so we defer the proof until a bit later.

Up to this point, our description of a single system has used a unit vector $|\psi\rangle$. That choice corresponds to maximal information. In many situations there is additional classical uncertainty. Perhaps a device prepares $|\psi_j\rangle$ with probability r_j. It is convenient to package such ensembles into a single object, the **density operator** (or **density matrix**)

$$\rho := \sum_j r_j\, |\psi_j\rangle\langle\psi_j| \in \mathcal{S}(\mathcal{H}), \tag{12}$$

which is Hermitian, positive semidefinite, and satisfies $\mathrm{tr}(\rho) = 1$. In fact, any operator which is Hermitian, positive semidefinite, and satisfies $\mathrm{tr}(\rho) = 1$ can be written in the form (12), and so we define:

Definition 33 (Density operator)**.** *A **density operator** $\rho \in \mathcal{S}(\mathcal{H})$ is a linear operator on $\mathcal{H}$ which satisfies $\rho = \rho^\dagger$, $\mathrm{tr}(\rho) = 1$, and $\rho \succeq 0$.*

We say that a state is **pure** when $\rho = |\psi\rangle\langle\psi|$, equivalently $\rho^2 = \rho$ and $\mathrm{tr}(\rho^2) = 1$, and otherwise it is **mixed**. A pure state corresponds to a rank 1 density matrix, and a mixed state corresponds to rank greater than 1. The Born rule extends linearly. Specifically, for a projector P,

$$\Pr[\text{“yes” on } P \text{ given } \rho] = \mathrm{tr}(P\rho),$$

and for an observable A,

$$\mathbb{E}_\rho[A] = \mathrm{tr}(A\rho).$$

Upon a projective measurement with projectors P_j, two kinds of updates occur. If we condition on the outcome j, then

$$\rho \longmapsto \frac{P_j \rho P_j}{\mathrm{tr}(P_j \rho)}.$$

If the outcome is forgotten, then

$$\rho \longmapsto \sum_j P_j \rho P_j,$$

which removes coherences between the corresponding subspaces.

Joint states admit a notion of marginalization that mirrors our classical $\vec{1}^T$ trick. Given ρ_{AB} on $\mathcal{H}_A \otimes \mathcal{H}_B$, the state of A alone is the **partial trace** over B:

$$\rho_A := \mathrm{tr}_B(\rho_{AB}) \in \mathcal{S}(\mathcal{H}_A).$$

In coordinates with respect to any orthonormal basis $\{|j\rangle_B\}$,

$$\mathrm{tr}_B(\rho_{AB}) = \sum_j \bigl(\mathbb{1}_A \otimes \langle j|\bigr)\, \rho_{AB}\, \bigl(\mathbb{1}_A \otimes |j\rangle\bigr). \tag{13}$$

The map tr_B is characterized by the identity

$$\mathrm{tr}\bigl[(X_A \otimes \mathbb{1}_B)\rho_{AB}\bigr] = \mathrm{tr}\bigl[X_A\, \mathrm{tr}_B(\rho_{AB})\bigr] \quad \text{for all } X_A\,,$$

so it really is the quantum version of taking a marginal. If ρ_{AB} is diagonal in the product basis, (13) reduces exactly to summing out the B index. The identity $\mathrm{tr}_B(\rho_{AB}) = \rho_A$ is the quantum sibling of marginalization by dotting probability vectors with with $\vec{1}^T$, as appeared in our earlier discussion.

Two corollaries are immediate from the Schmidt decomposition. First, if $|\Psi\rangle$ is a pure vector on AB and $\rho_{AB} = |\Psi\rangle\langle\Psi|$, then $\rho_A = \mathrm{tr}_B(\rho_{AB})$ and $\rho_B = \mathrm{tr}_A(\rho_{AB})$ share the same nonzero eigenvalues. The state $|\Psi\rangle$ is entangled if and only if either reduced state is mixed, equivalently if and only if the Schmidt rank is strictly greater than 1. Second, every mixed state can be realized as the marginal of a pure state on a larger space. Given a decomposition $\rho_A = \sum_k \lambda_k |k\rangle\langle k|$, the vector

$$|\Phi\rangle_{AR} = \sum_k \sqrt{\lambda_k}\, |k\rangle_A \otimes |k\rangle_R$$

on an auxiliary space $\mathcal{H}_R$ satisfies $\mathrm{tr}_R(|\Phi\rangle\langle\Phi|) = \rho_A$. This construction is called a **purification**.

With the above notations at hand, we can finally give a proof of the Schmidt decomposition. As mentioned above, it is really just a repackaging of the singular value decomposition, but it is instructive to go through the argument in the quantum language above.

Proof of Theorem 32. Let $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ be a unit vector. Form the rank-one projector

$$\rho_{AB} := |\Psi\rangle\langle\Psi|$$

and the reduced state on A

$$\rho_A := \mathrm{tr}_B(\rho_{AB}) \in \mathcal{S}(\mathcal{H}_A).$$

Then ρ_A is Hermitian, positive semidefinite, and satisfies $\mathrm{tr}(\rho_A) = 1$. By the spectral theorem there exist an orthonormal set $\{|k\rangle_A\}_{k=1}^r$ and numbers $\lambda_k \geq 0$ with $\sum_{k=1}^r \lambda_k = 1$ such that

$$\rho_A = \sum_{k=1}^r \lambda_k\, |k\rangle_A\langle k|\,,$$

where $r = \mathrm{rank}(\rho_A) \leq N_A$.

For each k with $\lambda_k > 0$ define a vector in $\mathcal{H}_B$ by

$$|\tilde{k}\rangle_B := \frac{1}{\sqrt{\lambda_k}}(\langle k|_A \otimes \mathbb{1}_B)\,|\Psi\rangle\,.$$

We first check orthonormality. For k, ℓ with $\lambda_k, \lambda_\ell > 0$ we compute

$$\begin{aligned}\langle\tilde{k}|\tilde{\ell}\rangle &= \frac{1}{\sqrt{\lambda_k\lambda_\ell}}\langle\Psi|\bigl(|k\rangle\langle\ell|_A \otimes \mathbb{1}_B\bigr)|\Psi\rangle \\ &= \frac{1}{\sqrt{\lambda_k\lambda_\ell}}\,\langle k|\,\rho_A\,|\ell\rangle = \frac{1}{\sqrt{\lambda_k\lambda_\ell}}\,\lambda_\ell\,\delta_{k\ell} = \delta_{k\ell}\,,\end{aligned}$$

so $\{|\tilde{k}\rangle_B\}_{k=1}^r$ is an orthonormal set in $\mathcal{H}_B$. Hence $r \le N_B$ as well.

Next we claim that $|\Psi\rangle = \sum_{k=1}^r \sqrt{\lambda_k}\,|k\rangle_A \otimes |\tilde{k}\rangle_B$. Let us define

$$|\Phi\rangle := \sum_{k=1}^{r} \sqrt{\lambda_k}\,|k\rangle_A \otimes |\tilde{k}\rangle_B\,,$$

and compare the two vectors by projecting onto A. For any m in an orthonormal basis of $\mathcal{H}_A$ that extends $\{|k\rangle_A\}_{k=1}^r$ we have

$$(\langle m|_A \otimes \mathbb{1}_B)\,|\Psi\rangle = \begin{cases} \sqrt{\lambda_m}\,|\tilde{m}\rangle_B & \text{if } \lambda_m > 0 \\ 0 & \text{if } \lambda_m = 0 \end{cases}$$

by construction. The same identities hold with $|\Psi\rangle$ replaced by $|\Phi\rangle$. Therefore

$$(\langle m|_A \otimes \langle\phi|_B)\,\big(|\Psi\rangle - |\Phi\rangle\big) = 0$$

for every m and every $|\phi\rangle \in \mathcal{H}_B$. Since such product bras span $(\mathcal{H}_A \otimes \mathcal{H}_B)^*$, it follows that $|\Psi\rangle = |\Phi\rangle$.

Finally observe the reduced state on B,

$$\rho_B := \mathrm{tr}_A(\rho_{AB}) = \sum_{k=1}^{r} \lambda_k\,|\tilde{k}\rangle_B\langle\tilde{k}|\,,$$

so the nonzero spectra of ρ_A and ρ_B agree and equal $\{\lambda_k\}$. The number r is therefore the common rank of ρ_A and ρ_B, which gives $r \le \min\{N_A, N_B\}$.

We have produced orthonormal sets $\{|k\rangle_A\}$ and $\{|\tilde{k}\rangle_B\}$ and nonnegative numbers $\{\lambda_k\}$ that sum to one such that

$$|\Psi\rangle = \sum_{k=1}^{r} \sqrt{\lambda_k}\,|k\rangle_A \otimes |\tilde{k}\rangle_B\,,$$

which is the desired form. □

Remark 34 (Uniqueness and degeneracies)**.** *The multiset of nonzero coefficients $\{\lambda_k\}$ is uniquely determined by $|\Psi\rangle$ since it is the spectrum of ρ_A and also of ρ_B. The orthonormal families $\{|k\rangle_A\}$ and $\{|\tilde{k}\rangle_B\}$ are unique up to phases when the λ_k are distinct. Within a degenerate eigenspace one may apply a unitary rotation on A and the same conjugate rotation on the corresponding span on B without changing the state $|\Psi\rangle$.*

Now we turn to some examples.

Example 11 (Embedding classical probability into quantum states). Fix the computational basis $\{|i\rangle\}_{i=1}^N$ of $\mathbb{C}^N$. A classical distribution $\vec{p} = (p_1, \dots, p_N) \in \Delta_N$ is encoded as the diagonal density matrix

$$\rho_{\mathrm{cl}}(\vec{p}) = \sum_{i=1}^{N} p_i\,|i\rangle\langle i|.$$

A measurement in this basis with projectors $P_i = |i\rangle\langle i|$ returns outcome i with probability $\mathrm{tr}(P_i\rho_{\mathrm{cl}}) = p_i$, matching the classical rule.

Example 12 (Bell state, reduced states, and entanglement). Consider two qubits with computational basis $|0\rangle, |1\rangle$. We will write $|00\rangle$ as a shorthand for $|0\rangle \otimes |0\rangle$, and similarly for $|11\rangle$. The maximally entangled vector

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right), \quad \rho_{AB} = |\Phi^+\rangle\langle\Phi^+|,$$

has reduced states

$$\rho_A = \mathrm{tr}_B(\rho_{AB}) = \frac{1}{2}\,\mathbb{1}, \qquad \rho_B = \mathrm{tr}_A(\rho_{AB}) = \frac{1}{2}\,\mathbb{1}\,.$$

Each qubit by itself looks completely random, yet the pair together sits in a definite pure state. Local mixedness together with global purity is a signature of entanglement and has no classical analogue.

In summary, tensor products allow us to assemble composite systems, while density matrices enable us to represent both quantum superposition and classical randomization within a single calculus. The partial trace serves as the quantum marginalization operator, mirroring our earlier $\vec{1}^T$ trick. Together, these tools provide a unified linear-algebraic framework for handling open systems, correlations, and measurements on subsystems.

2.3.2. *POVMs and channels*

We now broaden the two pillars introduced so far, namely unitary time evolution and projective (yes/no) measurements, into the general language of **quantum channels** and **POVMs** (positive operator-valued measures). This framework cleanly captures **open-system** dynamics (interaction with an environment) and the most general measurement statistics allowed by quantum mechanics. The picture to keep in mind is simple: attach an ancilla (the "apparatus" or "environment"), evolve unitarily on the larger space, and then either (i) forget the ancilla (a channel), or (ii) read the ancilla (a measurement). Everything that happens to a system can be modeled this way.

First we consider dynamics in the form of quantum channels. Fix a system Hilbert space $\mathcal{H}_S \simeq \mathbb{C}^d$ (here the subscript 'S' stands for 'system'). In practice a system rarely evolves in isolation; it can interact with an external register $\mathcal{H}_E \simeq \mathbb{C}^{d'}$ prepared in some state ρ_E (here the subscript 'E' stands for 'environment'). If the joint closed dynamics is unitary U_{SE}, then any initial system state ρ_S evolves as

$$\rho_S \longmapsto \mathcal{E}[\rho_S] := \mathrm{tr}_E\big(U_{SE}(\rho_S \otimes \rho_E)U_{SE}^\dagger\big).$$

From the cyclicity of trace and $\mathrm{tr}(\rho_E) = 1$, we immediately get $\mathrm{tr}(\mathcal{E}[\rho_S]) = \mathrm{tr}(\rho_S)$, i.e. *trace preservation*. Moreover, tensoring with an arbitrary ancilla and applying the above form shows *complete positivity*: $(\mathrm{Id}_A \otimes \mathcal{E})[X] \succeq 0$ for every positive X on $\mathcal{H}_A \otimes \mathcal{H}_S$.[9]

Definition 35 (Quantum channel)**.** *A **quantum channel** (or **quantum process**) on $\mathcal{H}_S$ is a linear map $\mathcal{E} : \mathcal{S}(\mathcal{H}_S) \to \mathcal{S}(\mathcal{H}_S)$ that is completely positive and trace-preserving (CPTP).*

The dilation form above is not just an example; it is universal:

[9]Positivity alone would require $\mathcal{E}[X] \succeq 0$ whenever $X \succeq 0$ on $\mathcal{H}_S$; *complete* positivity demands the same after adjoining *any* spectator system A. Physically, this guarantees the map never creates negative probabilities even on half of an entangled state.

Theorem 36 (Stinespring dilation)**.** *Every CPTP map $\mathcal{E}$ on $\mathcal{H}_S \simeq \mathbb{C}^d$ admits a representation of the above form for some environment dimension d', environment state ρ_E, and unitary U_{SE} that are fixed independently of the input ρ_S.*

We defer the proof of this since we need an additional structural result about quantum channels.

A convenient "matrix-element" form drops out when ρ_E is pure, say $\rho_E = |0\rangle\langle 0|$. Expanding U_{SE} in an orthonormal basis $\{|i\rangle_E\}$ and defining the **Kraus operators**

$$K_i := \langle i|U_{SE}|0\rangle \in \mathbb{C}^{d\times d},$$

we obtain the **operator-sum** (Kraus) representation

$$\mathcal{E}[\rho] = \sum_i K_i\,\rho\,K_i^\dagger, \qquad \sum_i K_i^\dagger K_i = \mathbb{1}.$$

Conversely, any family $\{K_i\}$ obeying the completeness relation defines a CPTP map. The Kraus representation is nonunique: $\{K_i\}$ and $\{\sum_j u_{ij} K_j\}$ (with u unitary) describe the same channel. These facts are formalized and proved in the following theorem:

Theorem 37 (Kraus decomposition)**.** *Let $\mathcal{E} : \mathcal{S}(\mathcal{H}_S) \to \mathcal{S}(\mathcal{H}_S)$ be CPTP on a d–dimensional Hilbert space $\mathcal{H}_S \simeq \mathbb{C}^d$. Then there exist operators $K_1, \ldots, K_r$ on $\mathcal{H}_S$ with*

$$\mathcal{E}[X] = \sum_{i=1}^r K_i\,X\,K_i^\dagger \quad \textit{for all } X, \quad \sum_{i=1}^r K_i^\dagger K_i = \mathbb{1},$$

where $r \le d^2$. Conversely, any finite family $\{K_i\}$ obeying $\sum_i K_i^\dagger K_i = \mathbb{1}$ defines a CPTP map by the same formula. The representation is nonunique: if U is any unitary and $K_i' := \sum_j U_{ij} K_j$, then $\{K_i'\}$ yields the same channel.

Proof. To begin, recall the Choi-Jamiołkowski isomorphism. Fix an orthonormal basis $\{|j\rangle\}_{j=1}^d$ of $\mathcal{H}_S$ and define the (unnormalized) maximally entangled vector

$$|\Omega\rangle := \sum_{j=1}^d |j\rangle \otimes |j\rangle \ \in\ \mathcal{H}_S \otimes \mathcal{H}_S.$$

The **Choi matrix** of $\mathcal{E}$ is

$$J_{\mathcal{E}} := (\mathrm{Id} \otimes \mathcal{E})(|\Omega\rangle\langle\Omega|) = \sum_{j,k=1}^d |j\rangle\langle k| \otimes \mathcal{E}\big(|j\rangle\langle k|\big).$$

By complete positivity we know that $J_{\mathcal{E}} \succeq 0$. Moreover, one can check that for any X on $\mathcal{H}_S$ we have

$$\mathcal{E}[X] = \mathrm{tr}_1\Big[(X^T \otimes \mathbb{1})\,J_{\mathcal{E}}\Big], \tag{14}$$

where tr_1 is the partial trace over the first tensor factor. This "reconstruction identity" follows by expanding X in the basis $\{|j\rangle\langle k|\}$.

Next observe that since $J_{\mathcal{E}} \succeq 0$ it admits a decomposition into rank–one projectors,

$$J_{\mathcal{E}} = \sum_{i=1}^r |v_i\rangle\langle v_i|$$

where $|v_i\rangle \in \mathcal{H}_S \otimes \mathcal{H}_S$, and $r = \mathrm{rank}(J_{\mathcal{E}}) \leq d^2$. Each vector $|v_i\rangle$ can be viewed as defining an operator $K_i : \mathcal{H}_S \to \mathcal{H}_S$ via the canonical "vectorization" correspondence: if $|v_i\rangle = \sum_{a,b} v_{ab}^{(i)} |a\rangle \otimes |b\rangle$, then

$$K_i = \sum_{a,b} v_{ab}^{(i)} \, |b\rangle\langle a|.$$

One can verify directly that for every X,

$$\mathrm{tr}_1\Big[(X^T \otimes \mathbb{1}) \, |v_i\rangle\langle v_i|\Big] = K_i X K_i^\dagger.$$

Combining this with (14), we find

$$\mathcal{E}[X] = \sum_{i=1}^{r} K_i X K_i^\dagger,$$

which is precisely the operator–sum form.

It remains to check the normalization. Since $\mathcal{E}$ is trace–preserving, for all ρ we have

$$\mathrm{tr}(\rho) = \mathrm{tr}(\mathcal{E}[\rho]) = \sum_{i=1}^{r} \mathrm{tr}(K_i \rho K_i^\dagger) = \mathrm{tr}\Big(\rho \sum_{i=1}^{r} K_i^\dagger K_i\Big).$$

Because this holds for all density operators ρ, it follows that $\sum_i K_i^\dagger K_i = \mathbb{1}$.

Conversely, suppose we start with any collection of operators $\{K_i\}$ satisfying $\sum_i K_i^\dagger K_i = \mathbb{1}$. The map

$$\mathcal{E}[X] = \sum_i K_i X K_i^\dagger$$

is clearly linear. Trace preservation follows from the same computation above, and complete positivity is immediate: for any ancilla system A and any positive operator Z on $\mathcal{H}_A \otimes \mathcal{H}_S$, we have

$$(\mathrm{Id}_A \otimes \mathcal{E})[Z] = \sum_i (\mathbb{1}_A \otimes K_i) \, Z \, (\mathbb{1}_A \otimes K_i)^\dagger \ \succeq \ 0.$$

Finally, note that the Kraus representation is not unique. If $u = (u_{ij})$ is any unitary matrix and we define $K_i' = \sum_j u_{ij} K_j$, then

$$\sum_i K_i' X K_i'^\dagger = \sum_j K_j X K_j^\dagger, \qquad \sum_i K_i'^\dagger K_i' = \sum_j K_j^\dagger K_j = \mathbb{1},$$

so $\{K_i\}$ and $\{K_i'\}$ describe the same channel. □

Remark 38 (Minimal Kraus number)**.** *The number r of Kraus operators can always be chosen as $r = \mathrm{rank}(J_{\mathcal{E}}) \leq d^2$. This number is minimal; any other representation can be obtained by enlarging the list with zero operators and applying a unitary rotation among them.*

Having established the Kraus decomposition, we can now establish Stinespring dilation:

PROOF OF THEOREM 36. By the Kraus decomposition, choose operators $K_1, \dots, K_r$ on $\mathcal{H}_S$ with $r \le d^2$ such that

$$\mathcal{E}[\rho] = \sum_{i=1}^{r} K_i \, \rho \, K_i^\dagger \quad \text{and} \quad \sum_{i=1}^{r} K_i^\dagger K_i = \mathbb{1}.$$

Let us introduce an environment Hilbert space $\mathcal{H}_E \simeq \mathbb{C}^r$ with orthonormal basis $\{|i\rangle_E\}_{i=1}^r$ and define an isometry

$$V : \mathcal{H}_S \longrightarrow \mathcal{H}_S \otimes \mathcal{H}_E, \quad V|\psi\rangle \ := \ \sum_{i=1}^{r} K_i|\psi\rangle \otimes |i\rangle_E.$$

Because $\sum_i K_i^\dagger K_i = \mathbb{1}$, we have $V^\dagger V = \mathbb{1}$; indeed, for all $|\phi\rangle, |\psi\rangle \in \mathcal{H}_S$,

$$\langle\phi|V^\dagger V|\psi\rangle = \sum_{i=1}^{r} \langle\phi|K_i^\dagger K_i|\psi\rangle = \langle\phi|\psi\rangle.$$

Taking the partial trace over E then recovers the channel:

$$\mathrm{tr}_E\big(V\rho V^\dagger\big) = \mathrm{tr}_E\left(\sum_{i,j} K_i\rho K_j^\dagger \otimes |i\rangle\langle j|\right) = \sum_i K_i\rho K_i^\dagger = \mathcal{E}[\rho].$$

To express V using a unitary on system plus environment with a fixed environment state, fix a distinguished vector $|0\rangle_E \in \mathcal{H}_E$ and identify $\mathcal{H}_S$ with the d–dimensional subspace $\mathcal{H}_S \otimes |0\rangle_E \subset \mathcal{H}_S \otimes \mathcal{H}_E$. Define U_{SE} on this subspace by

$$U_{SE}\big(|\psi\rangle \otimes |0\rangle_E\big) := V|\psi\rangle$$

for all $|\psi\rangle$ in $\mathcal{H}_S$. Since V is an isometry, this prescription maps an orthonormal basis of $\mathcal{H}_S \otimes |0\rangle_E$ to an orthonormal set in $\mathcal{H}_S \otimes \mathcal{H}_E$. Extend that partial isometry to a unitary U_{SE} on all of $\mathcal{H}_S \otimes \mathcal{H}_E$ by completing orthonormal bases on the domain and codomain and defining U_{SE} to map one basis to the other. Consequently,

$$\mathcal{E}[\rho] = \mathrm{tr}_E\big(V\rho V^\dagger\big) = \mathrm{tr}_E\Big(U_{SE}\,(\rho \otimes |0\rangle\langle 0|)\,U_{SE}^\dagger\Big),$$

which is precisely the stated dilation with environment state $\rho_E = |0\rangle\langle 0|$ and environment dimension $d' = r$. The unitary U_{SE} and the state ρ_E are determined by the chosen Kraus family for $\mathcal{E}$ and therefore are fixed independently of the input ρ. This completes the proof. □

Remark 39 (Minimal and nonunique dilations)**.** *If the Kraus family is chosen to be minimal (with $r = \mathrm{rank}(J_{\mathcal{E}})$), then $d' = r$ is the minimal environment dimension. Any two Kraus representations $\{K_i\}$ and $\{K_i'\}$ related by a unitary mixing $K_i' = \sum_j u_{ij} K_j$ yield dilations whose isometries differ by a unitary on the environment: $V' = (\mathbb{1} \otimes u)\,V$. Allowing a mixed ρ_E entails no extra generality, since any mixed state can be purified by enlarging the environment.*

Next we make some additional remarks about quantum channels.

Remark 40 (Composition and randomized control)**.** *Channels are closed under composition and convex combination. If $\mathcal{E}$ and $\mathcal{F}$ are channels, then so is $\mathcal{F} \circ \mathcal{E}$. And if with classical probabilities r_j you apply $\mathcal{E}_j$, the average map $\sum_j r_j\, \mathcal{E}_j$ is again a channel. Thus the set of channels is a convex monoid under composition.*

Remark 41 (Heisenberg picture)**.** *The adjoint map $\mathcal{E}^*$ acts on observables and satisfies*

$$\mathrm{tr}\big(\mathcal{E}[\rho]\,A\big) = \mathrm{tr}\big(\rho\,\mathcal{E}^*[A]\big), \quad \mathcal{E}^*[\mathbb{1}] = \mathbb{1}\,.$$

In Kraus form, $\mathcal{E}^*[A] = \sum_i K_i^\dagger A K_i$. We will use this duality to shuttle between "state evolution" and "observable evolution."

Having discussed general dynamics, we now turn out attension to general measurements. Projective measurements are special cases of more general procedures obtained by attaching an apparatus, evolving unitarily, and reading an outcome on the apparatus. Let $\{|i\rangle_A\}_{i=1}^N$ be an orthonormal basis for the apparatus and let U act on system+apparatus. If the apparatus is initialized in $|0\rangle_A$ and we measure it in the $\{|i\rangle_A\}$ basis, the probability of outcome i on input ρ is

$$p(i) = \mathrm{tr}(F_i\,\rho)\,, \qquad F_i := M_i^\dagger M_i, \qquad M_i := \langle i|U|0\rangle,$$

with $\sum_i F_i = \mathbb{1}$ by unitarity.

Definition 42 (POVM)**.** *A **positive operator-valued measure** (POVM) on $\mathcal{H}_S$ is a finite collection of positive semidefinite operators $\{F_i\}_{i=1}^N$ obeying $\sum_i F_i = \mathbb{1}$. Given a state ρ, the Born rule assigns outcome probabilities $p(i) = tr(F_i\rho)$.*

The operators F_i are sometimes called **effects**. When $F_i = P_i$ are orthogonal projectors that sum to $\mathbb{1}$ we recover the projective measurements from the axioms. In general, many distinct physical procedures can realize the same POVM statistics. One convenient realization chooses **measurement operators** (one set among many)

$$M_i \quad \text{with} \quad M_i^\dagger M_i = F_i,$$

and then the post-measurement state conditioned on outcome i is

$$\rho \;\longmapsto\; \frac{M_i\,\rho\,M_i^\dagger}{\mathrm{tr}(F_i\rho)}.$$

The family $\{\mathcal{I}_i\}_i$ with $\mathcal{I}_i[\rho] := M_i\rho M_i^\dagger$ is called a **quantum instrument**; it records both the probabilities and the (normalized) output states. Forgetting the outcome yields the average channel $\sum_i \mathcal{I}_i$.

As with channels, there is a universal dilation theorem for POVMs:

Theorem 43 (Naimark dilation)**.** *Every POVM $\{F_i\}$ on $\mathcal{H}_S$ can be realized as a projective measurement on a larger space: there exist an auxiliary Hilbert space $\mathcal{H}_A$, an isometry $V : \mathcal{H}_S \to \mathcal{H}_S \otimes \mathcal{H}_A$, and orthogonal projections $\{\Pi_i\}$ on $\mathcal{H}_A$ such that*

$$F_i = V^\dagger(\mathbb{1}\otimes\Pi_i)V \quad \textit{and} \quad p(i) = \mathrm{tr}\big(F_i\rho\big) = \mathrm{tr}\big[(\mathbb{1}\otimes\Pi_i)\,V\rho V^\dagger\big].$$

Remark 44 (Rank-one refinement)**.** *Every POVM admits a refinement to rank-one effects. Diagonalize each $F_i = \sum_j \lambda_{ij}\,|v_{ij}\rangle\langle v_{ij}|$ and regard the collection $\{F_{i,j} := \lambda_{ij}|v_{ij}\rangle\langle v_{ij}|\}_{i,j}$ as a new POVM. Coarse-graining its outcomes by summing over j reproduces the original statistics:*

$$\sum_j \mathrm{tr}(F_{i,j}\rho) = \mathrm{tr}(F_i\rho).$$

Thus, without loss of generality, one may work with rank-one POVMs when convenient.

To concretize the formalism, we record two examples.

Example 13 (Unsharp qubit measurement). For a qubit with Pauli vector $\vec{\sigma} = (X, Y, Z)$ and a unit vector $\hat{n} \in \mathbb{R}^3$, the two-outcome effects

$$F_{\pm}^{(\eta,\hat{n})} = \frac{\mathbb{1} \pm \eta\, \hat{n}\cdot\vec{\sigma}}{2}, \qquad 0 \le \eta \le 1,$$

form a POVM. The parameter η is a *sharpness*: $\eta = 1$ gives the projective measurement along $\hat{n}$, while smaller η yields noisy readout with probabilities

$$p(\pm) = \operatorname{tr}\big(F_{\pm}^{(\eta,\hat{n})}\rho\big) = \tfrac{1}{2}\big(1 \pm \eta\, \hat{n}\cdot\vec{r}\big),$$

where $\vec{r} = (\langle X\rangle, \langle Y\rangle, \langle Z\rangle)$ is the Bloch vector of ρ.

Example 14 (Embedding classical dynamics into a channel). Classical column-stochastic matrices are naturally realized as quantum channels that act classically on the computational basis and erase coherence. Fix an orthonormal basis $\{|i\rangle\}_{i=1}^N$ and let $M = (M_{ij})$ be column-stochastic ($M_{ij} \ge 0$ and $\sum_i M_{ij} = 1$ for each j). Define Kraus operators

$$K_{i|j} = \sqrt{M_{ij}}\, |i\rangle\langle j|.$$

Then

$$\mathcal{E}_M[\rho] = \sum_{i,j} K_{i|j}\, \rho\, K_{i|j}^{\dagger}, \qquad \sum_{i,j} K_{i|j}^{\dagger} K_{i|j} = \sum_j \Big(\sum_i M_{ij}\Big) |j\rangle\langle j| = \mathbb{1},$$

so $\mathcal{E}_M$ is CPTP. On diagonal inputs $\rho_{\rm cl}(\vec{p}) = \sum_j p_j |j\rangle\langle j|$ we recover the classical update

$$\mathcal{E}_M\big[\rho_{\rm cl}(\vec{p})\big] = \sum_{i,j} M_{ij} p_j\, |i\rangle\langle i| = \rho_{\rm cl}(M\cdot\vec{p}),$$

while for $j \ne k$ the coherence $|j\rangle\langle k|$ is sent to 0 because each Kraus term carries the same input label on both sides. Thus $\mathcal{E}_M$ is a "classicalizing" channel: it dephases in the computational basis and then applies the Markov update to the resulting distribution.

We have seen that the familiar tools of unitary evolution and projective measurements represent only the simplest quantum operations. Real quantum systems demand a richer framework: we enlarge the Hilbert space with ancillary systems, apply unitary evolution to the combined system, then either trace out the ancilla (yielding quantum channels) or measure it (yielding POVMs). This procedure generates the most general dynamics and measurement statistics that quantum mechanics allows. We have explained that quantum channels are completely positive trace-preserving (CPTP) linear maps, characterized by the Kraus representation or Stinespring dilation. POVMs are sets of positive operators that sum to the identity, understood through Naimark's theorem. But the conceptual heart is simple: we compose systems, evolve them unitarily, and then selectively forget or record information.

This unified framework will prove essential for understanding real quantum devices; indeed, in the real world, noise is inevitable, information is incomplete, and systems interact with environments beyond our control. Rather than limitations to work around, these general operations become the natural language for describing quantum processes in practice.

3. A taste of quantum many-body physics

We now turn to many–body systems built from n qubits. The ambient Hilbert space is the n–fold tensor product

$$\mathcal{H} := (\mathbb{C}^2)^{\otimes n} \simeq \mathbb{C}^{2^n}.$$

It is convenient to fix the computational basis $\{|0\rangle, |1\rangle\}$ on each site and to use the Pauli operators X, Y, Z discussed above. To streamline notation, we introduce a shorthand: for $1 \leq i \leq n$ we write

$$X_i := \mathbb{1}^{\otimes(i-1)} \otimes X \otimes \mathbb{1}^{\otimes(n-i)},$$

and similarly for Y_i and Z_i. Products such as $Z_i Z_j$ are understood to mean $Z_i \otimes Z_j$ with identities on all other sites, which we will not display explicitly.

More generally, a **Pauli string** on n qubits is a tensor product

$$P = \sigma_{a_1} \otimes \cdots \otimes \sigma_{a_n}, \qquad \sigma_{a_k} \in \{\, \mathbb{1}, X, Y, Z \,\},$$

and its **weight** is the number of non-identity factors,

$$w(P) := |\{\, k : \sigma_{a_k} \neq \mathbb{1} \,\}|,$$

while its **support** is the set $\mathrm{supp}(P)$ of sites where $\sigma_{a_k} \neq \mathbb{1}$. Two elementary commutation facts will be used repeatedly: Pauli matrices on different sites commute, while distinct Pauli matrices on the same site anticommute. Equivalently, Pauli strings P and Q either commute or anticommute, with

$$PQ = (-1)^{N_{\mathrm{anti}}(P,Q)}\, QP,$$

where $N_{\mathrm{anti}}(P, Q)$ counts the number of sites where both act nontrivially with different Pauli matrices.

With this notation in hand, we can define Hamiltonians. A **Hamiltonian** on $\mathcal{H}$ is a Hermitian operator $H = H^\dagger$. In units $\hbar \equiv 1$, the closed–system time evolution is

$$U(t) = e^{-\,iHt}, \qquad |\Psi(t)\rangle = U(t)\,|\Psi(0)\rangle.$$

Since H is Hermitian, its spectrum is real. We denote its smallest eigenvalue by E_0 (the **ground energy**) and the corresponding eigenspace by the **ground space**. A Hamiltonian is called k–**local** if it decomposes as

$$H = \sum_a H_a\,, \quad w(H_a) \leq k \;\text{ for every term } H_a,$$

i.e. each interaction acts nontrivially on at most k sites. In the qubit setting one often expands H in the Pauli–string basis,

$$H = \sum_\alpha h_\alpha\, P_\alpha\,, \quad w(P_\alpha) \leq k,$$

with real coefficients h_α. To express **geometric locality**, we can place the n qubits on the vertices V of a graph $G = (V, E)$. A geometrically k–local Hamiltonian has each H_a supported on a connected region of at most k vertices (for $k = 2$, typically on edges $(i,j) \in E$). For example, on a line $G = \{1,\dots,n\}$ with edges $(i,i+1)$, a nearest–neighbor two–local Hamiltonian has the form

$$H = \sum_{i=1}^{n-1} H_{i,i+1} + \sum_{i=1}^{n} H_i,$$

with $H_{i,i+1}$ acting only on sites $i, i+1$ and H_i acting on site i.

The canonical playground for these ideas is the (ferromagnetic) **transverse–field Ising model** (TFIM) on a graph $G = (V,E)$:

$$H_{\mathrm{TFIM}}(J,h) = -\,J \sum_{(i,j)\in E} Z_i Z_j \;-\; h \sum_{i\in V} X_i, \quad J \geq 0,\, h \geq 0.$$

The first term lowers the energy when neighboring Z–spins align, while the second term lowers the energy for qubits pointing in the x–direction (the $|+\rangle$ eigenstate of X). Thus the two terms compete, and since Z and X do not commute, the model is genuinely quantum. The model is 2–local and geometrically local on G.

Two limiting regimes are exactly solvable and already illustrative. In the classical limit $h = 0$, all terms commute. Ground states maximize each $Z_i Z_j$, so for $J > 0$ they are the two fully aligned product states $|0\cdots 0\rangle$ and $|1\cdots 1\rangle$, with two–fold degeneracy. Excitations are domain walls: a bond with anti–aligned neighbors costs energy $2J$ (on an open chain; with periodic boundary conditions, domain walls come in pairs costing $4J$ total). In the opposite paramagnetic limit $J = 0$, each site independently minimizes $-hX_i$, with a unique ground state $|+\rangle^{\otimes n}$ where $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. A single spin flip to $|-\rangle$ costs energy $2h$.

Between these limits, the terms fail to commute, which is the source of genuinely quantum behavior. The model enjoys a $\mathbb{Z}_2$ symmetry generated by the global "spin–flip" operator

$$\mathcal{P} := \prod_{i\in V} X_i,$$

under which $Z_i \mapsto -Z_i$ while $X_i \mapsto X_i$. Since $[\mathcal{P}, H_{\mathrm{TFIM}}] = 0$, the Hamiltonian preserves this symmetry. For small h/J the ground space on large graphs approximately breaks the symmetry, exhibiting long–range Z–order. For large h/J the unique ground state is the symmetric paramagnet. On a one–dimensional chain the model is exactly solvable (via Jordan–Wigner fermionization), and at zero temperature there is a quantum phase transition in the thermodynamic limit ($n \to \infty$) at $h = J$ where the energy gap between the lowest and second lowest eigenvalues of H go to zero. While we will not derive this here, a two–site analysis already captures the competition of the two terms.

Example 15 (Two–site TFIM). On two qubits,

$$H_2(J,h) = -\,J\, Z_1 Z_2 - h\,(X_1 + X_2).$$

Diagonalizing (for instance in the joint eigenbasis of the parity $X_1 X_2$) yields four eigenvalues

$$E \in \left\{-\sqrt{J^2 + 4h^2},\ -J,\ +J,\ +\sqrt{J^2 + 4h^2}\right\}.$$

For $J \geq 0$ the ground energy is $E_0 = -\sqrt{J^2 + 4h^2}$, and the gap to the first excited level is

$$\Delta(J, h) = \sqrt{J^2 + 4h^2} - J.$$

We recover the limits discussed above: $\Delta(0, h) = 2h$ and $\Delta(J, 0) = 0$ (reflecting the two–fold degeneracy at $h = 0$). Already at two sites we see how the transverse field h lifts the classical degeneracy and stabilizes a unique paramagnet, while the interaction J favors ferromagnetic order.

On longer chains, the low–energy excitations can be understood in terms of order and disorder. In the $h = 0$ limit, excitations are domain walls that can move freely; turning on a small h allows them to hop and to be created or annihilated in pairs. In the opposite $J = 0$ limit, the excitations are independent spin flips. The $\mathbb{Z}_2$ symmetry generated by $\mathcal{P}$ forbids a nonzero $\langle Z_i \rangle$ expectation value in any exact eigenstate on a finite chain; nevertheless, in the ferromagnetic phase ($h/J \ll 1$) the ground space is nearly two–fold degenerate and exhibits robust long–range correlations $\langle Z_i Z_j \rangle \approx 1$ for distant i, j.