Talks for Quantum Error Correction REU, 2022

Sanchayan Dutta
dutta@ucdavis.edu

Contents

1 8 July, 2022

1.1 Bloch sphere and pure states

I mentioned last day that any qubit state $|\psi\rangle$ can be written as $a|0\rangle + b|1\rangle$ where $a, b \in \mathbb{C}$ and $|0\rangle, |1\rangle \in \mathbb{C}^2$. To represent this in real space would take 4 coordinates! But let's see if we can make it more economical.

As only the relative phases between coefficients of $|0\rangle$ and $|1\rangle$ matter in specifying a quantum state and since global phases are irrelevant, we can WLOG take the coefficient of $|0\rangle$ to be real and non-negative. This cuts down one degree of freedom.

We can cut down another degree of freedom by noticing that $\langle\psi|\psi\rangle = 1$ as the total probability of the quantum system has to be 1. This gives us a possible parameterization $\cos(\theta)|0\rangle + \sin(\theta)e^{i\phi}|1\rangle$ where θ ranges from 0 to $\pi/2$ and ϕ runs from 0 to 2π.

Well, actually we use half-angles $\cos(\theta/2)$ and $\sin(\theta/2)$ in the parameterization where θ runs from 0 to π. Why? Because once we move to the density operator formalism $|\psi\rangle\langle\psi|$, we get a nice parameterization of pure states in terms of polar coordinates. Moreover, we get to visualize Pauli X, Y and Z operators as rotations by 90 degrees of the Bloch sphere. The Pauli matrices form a basis for the real vector space of 2×2 Hermitian matrices.

You might at first feel uneasy about the fact that perpendicular things like $|0\rangle$ and $|1\rangle$ are 180 degrees apart on the Bloch sphere. But the major advantage of this representation is the simple interpretation of Pauli X, Y and Z gates as rotations about the x, y and z axes of the Bloch sphere. This will get more clear once we discuss the density operator formalism.

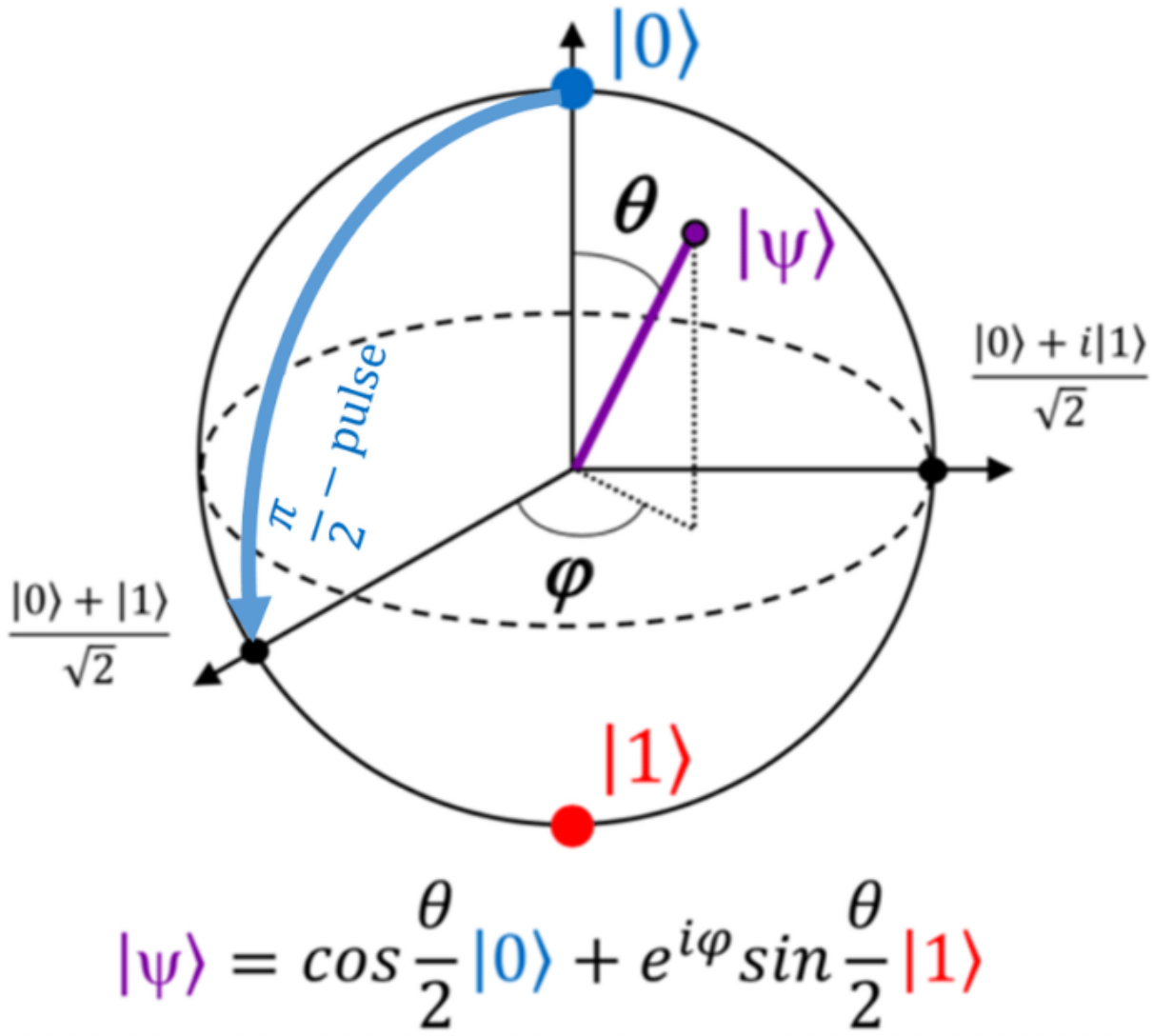
By the way, why are Pauli matrices so super special and have such nice commutation relations? Because there is an isomorphism between the algebra of quaternions and Pauli matrices (upon adding a phase factor of i). Quaternions were traditionally discover to describe rotations in 3 dimensional space by Sir William Rowan Hamilton in 1843, Dublin.

(Wikipedia) The great breakthrough in quaternions finally came on Monday 16 October 1843 in Dublin, when Hamilton was on his way to the Royal Irish Academy where he was going to preside at a council meeting. As he walked along the towpath of the Royal Canal with his wife, the concepts behind quaternions were taking shape in his mind. When the answer dawned on him, Hamilton could not resist the urge to carve the formula for the quaternions,

$$i^2 = j^2 = k^2 = ijk = -1$$

into the stone of Brougham Bridge as he paused on it. Although the carving has since faded away, there has been an annual pilgrimage since 1989 called the Hamilton Walk for scientists and mathematicians who walk from Dunsink Observatory to the Royal Canal bridge in remembrance of Hamilton's discovery.

Equivalently, the Pauli matrices satisfy $(iX)^2 = (iY)^2 = (iZ)^2 = -I$.



(Image taken from https://www.researchgate.net/publication/335028508.)

By the way, why do we talk so much about Hermitian or self-adjoint operators in quantum mechanics? Physically observable quantities like position, momentum and energy yield real valued measurements. The Hilbert space formalism accounts for this by associating observables with Hermitian operators. The eigenvalues of the operator are the allowed values of the observable. And we know Hermitian operators have a real spectrum of eigenvalues.

Like, an energy observable looks like $E_1|0\rangle\langle 0| + E_2|1\rangle\langle 1|$. Any Hermitian operator has a diagonal spectral decomposition of this form and the eigenbasis of the observable is the measurement basis corresponding to it. In this example, the measurement basis is $\{|0\rangle, |1\rangle\}$ and the eigenvalues are E_1 and E_2. You might still wonder though, that aren't there non=Hermitian operators that also have real eigenvalues? Why can't those represent observables in quantum mechanics?

The quick answer is that non-Hermitian matrices, even if they have a real spectrum of eigenvalues, that their eigenspaces are not orthogonal. So it doesn't really make sense to say that one has measured the system to be in some eigenspace but not in the others, because there might be a non-zero overlap with another eigenspace.

Theorem: An operator is Hermitian if and only if it is diagonalizable in an orthonormal basis with real eigenvalues.

The other natural question is about unitaries. Why are evolutions of closed quantum systems described by unitary transformations? There's also a quick answer for this: $||A|\psi\rangle|| = |||\psi\rangle||$ so that total probability is conserved and A should be invertible so that information is conserved. Putting these together, we get that A must be a unitary matrix U satisfying $UU^\dagger = U^\dagger U = I$ (*theorem*).

1.2 Density matrices and mixed states

The pure state of a 2-qudit system can be written in the Schmidt decomposition basis as $|\psi\rangle = \sum_k d_k|a_k\rangle|b_k\rangle$. The advantage of this basis is that there is only a single index k to sum over, instead of something like $\sum_{ij} a_{ij}|i\rangle|j\rangle$ where there are two different indices. This kind of a diagonalization follows from the singular value decomposition of (a_{ij}).

The expectation value of an observable O is given by $\langle\psi|(O|\psi\rangle)$. Why? First, Hermitian observables have a complete set of orthonormal eigenvectors. Let's the eigenvalues of O are given as follows in some eigenbasis $|\psi_q\rangle$.

$$O|\phi_q\rangle = q|\phi_q\rangle \tag{1}$$

Let's say that you have some quantum state, $|\psi\rangle$ and, because you're interested in the operator O, you decide to express this state in terms of the eigenvectors of O such that

$$|\psi\rangle = \sum_q c_q|\phi_q\rangle \tag{2}$$

where the coefficients $c_q = \langle\phi_q|\psi\rangle$ and $\sum_q |c_q|^2 = 1$. We can interpret $|c_q|^2$ as the probability that we measure $|\psi\rangle$ in the state $|\phi_q\rangle$. Expectation values are the average value of q we would expect to measure for a quantum state $|\psi\rangle$.The expectation value is given as

$$\begin{aligned}
\langle O\rangle &= \langle\psi|O|\psi\rangle \\
&= \sum_{q,q'} c^*_{q'} c_q \langle\phi_{q'}|O|\phi_q\rangle \\
&= \sum_{q,q'} c^*_{q'} c_q q \langle\phi_{q'}|\phi_q\rangle \\
&= \sum_{q,q'} c^*_{q'} c_q q \delta_{q,q'} \\
&= \sum_q q|c_q|^2
\end{aligned} \tag{3}$$

which is indeed, the weighted average value we measure for O.

Getting back to where we started, suppose we have two qubit system and we express its pure state as $|\psi\rangle = \sum_k d_k|a_k\rangle|b_k\rangle$. Let A be an *arbitrary* observable on the system of qudit 1 and let B be an *arbitrary* observable on qudit 2. Then say we want to find $\langle A \otimes \mathbb{I}\rangle_\psi$, where $\mathbb{I}$ is the identity observable. This can be written as:

$$\langle A \otimes \mathbb{I}\rangle = \sum_{k,l} d_k d_l \langle a_k|A|a_l\rangle\langle b_k|b_l\rangle$$

$$= \sum_k d_k^2 \langle a_k|A|a_k\rangle$$

$$= \mathrm{Tr}[A(\sum_k d_k^2 |a_k\rangle\langle a_k|)]$$

This last relation follows from the fact that $\mathrm{Tr}|a\rangle\langle b| = \langle b|a\rangle$.

What does $\sum_k d_k^2 \langle a_k|A|a_k\rangle$? Weighted average of expectations of A in the eigenbasis of qudit 1!

The term $\sum_k d_k^2 |a_k\rangle\langle a_k|)$ is called the density matrix of qudit 1 written as ρ_1. The advantage is that $\langle A \otimes I\rangle$ can be concisely written as $\mathrm{Tr}((A \otimes \mathbb{I})\rho) = \mathrm{Tr}(\rho(A \otimes \mathbb{I}))$ for any arbitrary observable A on the first qudit.

Well, there's also a deeper mathematical reason: In full generality, quantum states ρ are *linear functionals* on operators in $\mathcal{L}(H)$ rather and it's a theorem that all symmetrizing functionals on $\mathcal{L}(H)$ are proportional to trace. The trace functional has nice symmetry properties and some nice algebraic properties, but it's chosen mainly for symmetry reasons.

It follows that $\rho \geq 0$ (positive semidefiniteness, meaning the eigenvalues are all non-negative), $\rho = \rho^\dagger$ (self-adjointness or Hermitian) and $\mathrm{Tr}(\rho) = 1$ (probabilities add up to 1).

Now let's talk about mixed states. **There are two ways in which mixed states can originate but they're in effect the same thing.**

If $|\psi\rangle = \sum_k d_k |a_k\rangle|b_k\rangle$ then $\rho_{AB} = |\psi\rangle\langle\psi|$. Taking partial trace over B, we get the state $\rho_A = \mathrm{Tr}_B(\rho_{AB}) = \sum_k d_k^2 |a_k\rangle\langle a_k|$. You can calculate this from $|\psi\rangle\langle\psi| = \sum_{kl}(d_k|a_k\rangle|b_k\rangle)(d_l\langle a_l|\langle b_l|) = \sum_{kl} d_k d_l |a_k\rangle\langle a_l| \otimes |b_k\rangle\langle b_l|$. We know $\mathrm{Tr}_B(A \otimes B) = \mathrm{Tr}_A(A \otimes B) = A\mathrm{Tr}(B)$ and $\mathrm{Tr}_A(A \otimes B) = \mathrm{Tr}(A)B$.

So you see that by taking the partial trace over the pure state $\rho = |\psi\rangle\langle\psi|$ of a joint system AB we got a mixed state of the subsystem A. A density matrix represents a pure state if and only if it can be written as $|\psi\rangle\langle\psi|$ for some quantum state $|\psi\rangle$. Equivalently it's a rank 1 projection matrix. Mixed states are convex combinations of pure states of this form, but are not rank 1 projection matrices in general.

The other way in which mixed states can originate, is due to classical statistical uncertainity in an experiment. Suppose in an experiment you get the state $|0\rangle\langle 0|$ which 50% probability and the state $|1\rangle\langle 1|$ with another 50%. Then this classical superposition can be written as $\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}\mathbb{I}$. where $\mathbb{I}$ is the identity matrix.

However, this same state $\frac{1}{2}\mathbb{I}$ can also originate from a quantum superposition like $|\Omega\rangle = \sqrt{\frac{1}{2}}|00\rangle + \sqrt{\frac{1}{2}}|11\rangle$ which is the Bell state. You will notice that $\mathrm{Tr}_B(|\Omega\rangle\langle\Omega|) = \frac{1}{2}\mathbb{I}$. There is *no way* to distinguish between this preparation of the quantum state and previous preparation which occured due to classical probabilistic uncertainity in an experiment!

Another point is that, non-zero off-diagonal elements of denote that the system features a *quantum superposition* between the eigenbasis elements that you chose to represent ρ. In other words, your state is not only a statistical mixture of your basis states (which can be understood as representing *your* ignorance about the system), but an actual quantum superposition in which those basis elements are "coexisting".

Of course, there can be a combination between those two situations (different degrees of coherence/mixture), and all this is basis dependent (something that looks like a quantum superposition in one basis, can be described in terms of a single basis element in a different basis).

There's another interesting point to be made here. Corresponding to each pure state $|\psi\rangle\langle|\psi\rangle$ there is a corresponding boolean observable $|\psi\rangle\langle|\psi\rangle$ such that $\mathrm{Tr}((|\psi\rangle\langle\langle|\psi\rangle)(|\psi\rangle\langle\psi|)) = 1$. This fact $\rho(x) = \mathrm{Tr}(\rho x) = x(\rho)$ gives a correspondence between states and operators in $\mathcal{M}_d(\mathbb{C})$. For a Bloch

sphere, you can get a dual Bloch sphere of observables – this is a subtle point and I'll talk about it next day. By the way, pure states can also be interpreted as "indecomposable events".

1.3 *-algebra of observables

Okay, I've spoken about various things already today, so I'll just cover this last topic briefly and I can pick up from here next week. This is a subtle topic and will take a lot of care.

Fact 1: Two observables having the same expectation values on all states cannot be distinguished.

Fact 2: An observable A is defined in terms of a concrete experimental apparatus, which yields the numerical results of measurements in any state and each concrete experimental apparatus has inevitable limitations implying a scale bound.

Fact 3: The duality relation between observables and states allows to display and define linear structures in the algebra of observables $\mathcal{O}$.

Fact 4: (Wikipedia) The expectation value of an observable A, which is a Hermitian linear operator, for a given Schrödinger state $|\psi(t)\rangle$, is given by $\langle A\rangle_t = \langle\psi(t)|A|\psi(t)\rangle$.

In the Schrodinger picture, the state $|\psi(t)\rangle$ is related to the state $|\psi(0)\rangle$ at time 0 by a unitary time-evolution operator $U(t)$, i.e., $|\psi(t)\rangle = U(t)|\psi(0)\rangle$.

In the Heisenberg picture, the states remain constant at their initial value, whereas operators evolve according to $A(t) := U^\dagger(t)AU(t)$.

By the Stone–von Neumann theorem, the Heisenberg picture and the Schrödinger picture are unitarily equivalent, just a basis change in Hilbert space. In some sense, the Heisenberg picture is more natural and convenient than the equivalent Schrödinger picture, especially for relativistic theories. Lorentz invariance is manifest in the Heisenberg picture, since the state vectors do not single out the time or space.

Fact 5: In the Heisenberg picture as opposed to the Schrodinger picture, it is virtually always simpler to determine how a system has changed through time. You can select a group of representative observables from the Heisenberg image that span the entire vector space of operators that may act on the system under investigation, including the unitary operators. You can typically determine what unitary operator was applied to that set of observables based on how they change. Therefore, there is more information in the Heisenberg picture observables than in the Schrodinger picture state.

The *-star algebra structure on observables follows from this, but the discussion is quite involved. I will discuss a little bit next day, but you might as well read the textbook by Stocchi.

2 22 July, 2022

2.1 The rotation group SO(3)

Say we want to write down the matrix of the rotation $R(\hat{n}, \psi)$ about an axis $\hat{n}$. You give me an arbitrary axis $\hat{n}$ in the original coordinate system and I rotate by an angle ψ, anywhere from 0 to 2π, and I want to write the 3×3 matrix of this rotation down.

To do this, we can observe the following:

$$R(\hat{e}_3, \psi) = \begin{pmatrix} \cos(\psi) & \sin(\psi) & 0 \\ -\sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where $e_3 = e_1 \times e_2$.

Now let's see what an infinitesimal rotation does, to first order:

$$R(\hat{e}_3, \delta\psi) = I + \delta\psi \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The matrix on the right is called one of the "generators" of rotation. In quantum mechanics, it even turns out that the generators of rotations are physical quantities, namely the components of angular momentum. That you would like to be represented by a Hermitian operator, because you would like to have its eigenvalues real. So let's pull out a factor of i in the definition:

$$R(\hat{e}_3, \delta\psi) = I + i\delta\psi \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = I + iJ_3\delta\psi$$

That's what an infinitesimal rotation in the x-y plane looks like.

Now we want to show that repeated infinitesimal rotations indeed lead to a complete rotation by angle ψ about $\hat{e}_3$. Say we a succession of n rotations, each by $\delta\psi$, such that in the limit $n \to \infty$, the quantity $n\delta\psi \to \psi$. Then it immediately implies that $R(\hat{e}_3, \psi)$ should be equal to repeatedly acting with the matrix J_3.

$$R(\hat{e}_3, \psi) = (I + iJ_3\delta\psi)^n = \lim_{n\to\infty} (I + \frac{iJ_3\psi}{n})^n = \exp(iJ_3\psi).$$

You can get a similar matrix J_i about each of the coordinate axes $\hat{e}_i$.

$$R(\hat{e}_1, \delta\psi) = I + iJ_1\delta\psi$$

where $J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}$. Similarly, we have

$$R(\hat{e}_2, \delta\psi) = I + iJ_2\delta\psi$$

where $J_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}$.

So we have gotten the infinitesimal generators of rotations in each of these planes.

Now the interesting thing is that these matrices form a Lie algebra. In other words, they form a linear vector space among themselves such that the commutator of any two of them is again an element of the same set of elements. $[J_i, J_k] = i\epsilon_{ijk}J_l$. So the commutator of any two of these matrices is the third of these matrices, in cylic permutation order. The ϵ_{ijk}'s are called the structure constants of the Lie algebra SO(3), called $\mathfrak{so}(3)$.

J_1, J_2, J_3 are the infinitesimal generators of the Lie algebra $\mathfrak{so}(3)$. This "infinitesimal generators" is actually bit of a misnomer. There is nothing infinitesimal about them. Rather, technically, they are generators of "infinitesimal transformations" as we saw in the example above. Anyway, J_1, J_2, J_3 in a sense, define the rotation group in 3 dimensions. Meaning if you give me all the properties of these J_i's, then by exponentiation, I can find the general element of this group.

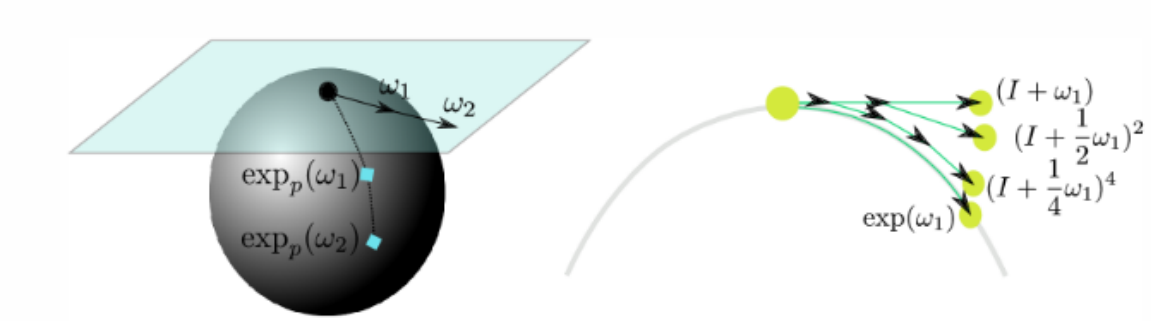
Oh well, now what about our original target of finding $R(\hat{n}, \psi)$ for a general axis of rotation $\hat{n}$? That has a bit of a problem, as it stands. The general rotation $R(\hat{n}, \psi) = e^{i(J_1 n_1 + J_2 n_2 + J_3 n_3)\psi}$. So you first find an infinitesimal transformation about that new axis, which is generated by J_1, J_2, J_3 themselves, but $\mathbf{J}.\hat{n}$. But the difficulty is that the J_i's don't commute with each other. And when they don't commute with each other then $e^{A+B} \neq e^A e^B$. This non-commutativity is the root of all drama in quantum physics!

Fun Fact (from Alain Connes): *The outcome of this still fascinates me now: The noncommutativity, which was discovered by people in quantum mechanics, in fact is a generator of time. I am still thinking about the fact that the passage of time, or the way we feel that time is going on and we cannot stop it, is in fact exactly the consequence of the noncommutativity of quantum mechanics, or more explicitly of the inherent randomness of quantum mechanics.* The fact that our universe has an arrow of time, is precisely because of the non-commutativity of quantum mechanics!

The non-commutativity also is relevant to the fact that our spacetime is curved. If you vary "some quantity" in this direction and then that, or vice versa, they are not the same. This simple fact (well, not so simple) helps us detect curvature in spacetime.

In contrast to SO(3), SO(2) is an abelian group. $R(\psi)R(\psi') = R(\psi + \psi') = R(\psi')R(\psi)$.

The Lie algebra so(3) has a lot of practical applications in robotics. Suppose you have a camera controlled by a mechanical handle of fixed length which can rotate in $\mathbb{R}_3$.



Visual representation of the relationship between the linear tangent space so(3) and the non-linear Lie group SO(3).
Image source: **Handa et al.**

Here in the image you can think of ω_1 and ω_2 as J_1 and J_2, and J_3 as an infinitesimal generator of tilts of the camera face. It's clearly easier to write down the coordinate vector $\hat{n}$ and then $\mathbf{J}.n$ instead of dealing with matrices. We already saw that J_1, J_2, J_3 form a basis of the tangent space at the identity and hence have nice linear algebraic properties. "*This tangent space is a linear space with three degrees of freedom and no additional constraints which makes it much easier to work with than its corresponding Lie group for motion estimation and optimization problems. This allows us to model rotations as skew-symmetric matrices with three free parameters.*"

There are otherwise 9 parameters in the rotation matrix, constrained by $R^T R = I$ and $\det(R) = +1$, which would normally have to be solved as a constraint optimization problem, which can be tedious to implement and program.

This is once again a good time to recall that: (Wikipedia) *Lie algebras are closely related to Lie groups, which are groups that are also smooth manifolds: any Lie group gives rise to a Lie algebra, which is its tangent space at the identity. Conversely, to any finite-dimensional Lie algebra over real or complex numbers, there is a corresponding connected Lie group unique up to finite coverings (Lie's third theorem).*

2.2 Quantum Operations and the Geometric Interpretation of the Knill-Laflamme-Viola Error Detection and Correction Conditions

For this it is first useful to understand the notion of quantum operations or TPCP maps. Any quantum operation on the state of a system ρ can be expressed as $\mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger$ where the terms E_k are maps between two Hilbert spaces $E_k \colon \mathcal{H}_A \to \mathcal{H}_B$. This is a very general setting, and in fact, all realistic maps on quantum states are of this form (**theorem**: Stinespring-Kraus). Thus, noisy quantum channels or operations or systems, can be modelled using these TPCP maps or quantum operations.

The main point relevant to error correction is the following theorem from Nielsen & Chuang:

Theorem 10.2 (N & C): *Suppose C is a quantum code and $\mathcal{R}$ is the error-correction operation constructed in the proof of Theorem 10.1 to recover from a noise process $\mathcal{E}$ with the operation elements $\{E_i\}$. Suppose $\mathcal{F}$ is a quantum operation with operation elements $\{F_j\}$ which are linear combinations of the E_i, that is $F_j = \sum_i m_{ji} E_i$ for some matrix m_{ji} of complex numbers. Then the error-correction operation $\mathcal{R}$ also corrects for the effects of the noise process $\mathcal{F}$ on the code C.*

This is a powerful viewpoint, and is at the root of why quantum error correction works, even though there is a continuum of possible errors. Suppose $\mathcal{E}$ is a TPCP map acting on a single qubit. Then its operation elements can be written as a linear combination of the Pauli matrices I, X, Y, Z. This means, it's possible to discretize quantum errors, and that to fight against the continuum of errors possible on a single qubit, it's enough to win the war against the four Paulis X, Y, Z and I. This is in stark constrast to errors in analog systems because there are infinitely many possible error syndromes. This is why digital error correction is preferred to analog error correction. It also follows from here that for an n-qubit system $(\mathbb{C}^2)^{\otimes n}$ it suffices to correct multi-Pauli errors.

KLV Error Detection

Now suppose our system S has n qubits and the Hilbert space $\mathcal{S}$ has a code subspace $\mathcal{C} \subset \mathcal{S}$. We know $S = \mathcal{C} \oplus \mathcal{C}^\perp$. There is an associated projector Π_C onto the code subspace $\mathcal{C}$. An error E is detected by $\mathcal{C}$ if the following protocol works:

1. Prepare a state $|\psi\rangle$ in $\mathcal{C}$.
2. Allow the TPCP error E to occur, so that the new state is $E|\psi\rangle$.
3. Make a measurement to detect whether the state is in $\mathcal{C}$ or in the orthogonal complement; the outcome is either $\Pi_C E|\psi\rangle$ or $(I - \Pi_C)E|\psi\rangle$.
4. Accept the state in the former case and reject it otherwise. The protocol is correct if the accepted states are proportional to the initial state, i.e., formally:

$$\Pi_C E \Pi_C = \alpha_E \Pi_C$$

for scalars α_E. Or $\Pi_C E|\psi\rangle = \alpha_E|\psi\rangle$ for all $|\psi\rangle \in \mathcal{C}$. It has to be proportional to identity, because otherwise one codeword in $\mathcal{C}$ will be transformed to another in $\mathcal{C}$ and it will be a non-recoverable error. In error correction, we only deal with detecting recoverable errors, and not the hopeless cases.

KLV Error Correction

From **Niel de Beaudrap** (https://quantumcomputing.stackexchange.com/a/7119):

We can try to geometrically interpret the Knill-Laflamme conditions on a code-space $\mathcal{C}$, as follows.

Images of the code-space under an error operation

First, consider any *individual* error operator E_j: this is a Kraus operator of some transformation of the system, and so is not necessarily even unitary.

For any input pure state $|\psi\rangle \in C$ and for an arbitrary operator E_j, the vector $E_j|\psi\rangle$ will represent some distorted image of the input state. You might wonder whether different states $|\psi\rangle \in \mathcal{C}$ could be affected in different ways: specifically, whether some vectors $E_j|\psi\rangle$ will have different norms for various states $|\psi\rangle$. The Knill-Laflamme condition indicates that this isn't possible: as $P_\mathcal{C} E_j^\dagger E_j P_\mathcal{C} = \alpha_{j,j} P_\mathcal{C}$ for some scalar $\alpha_{j,j}$ (which we can show is a non-negative real), it follows that $\| E_j|\psi\rangle \|^2 = \alpha_{j,j}$ for all $|\psi\rangle \in \mathcal{C}$.

So: for any set of errors $\{E_j\}_j$ satisfying the Knill-Laflamme conditions, each error E_j produces some faithful image of the original code-space $\mathcal{C}$, i.e., proportional to a unitary transformation on $\mathcal{C}$, even if E_j is not itself proportional to a unitary.

Images of the code-space under different error operations

Next, consider how these images of the code-space compare for two different error operators E_j and E_k, and under what conditions we can hope to recover a state which have been affected by one of the errors.

* In some cases –; as with distinct single-qubit Pauli operators on a code with distance 3 or greater –; the subspace $E_j\mathcal{C}$ and $E_k\mathcal{C}$ will be orthogonal to one another, and therefore perfectly distinguishable. In particular, we would then have $E_j|\psi\rangle \perp E_k|\varphi\rangle$ for any $|\psi\rangle, |\varphi\rangle \in \mathcal{C}$, so that $P_\mathcal{C} E_j^\dagger E_k P_\mathcal{C} = 0$. If all of the error operators $\{E_j\}_j$ are 'orthogonal' in this sense, there would then be no problem to correct the errors: we could simply perform a measurement of which of the subspaces $E_r\mathcal{C}$ we were in, and then perform a unitary transformation which maps $E_r\mathcal{C} \mapsto \mathcal{C}$. This corresponds to the case where $\alpha_{j,k} = 0$ for any $j \neq k$.

* However, it is also possible (for instance, in a degenerate code as [Patrick Fuentes indicates in his answer](https://quantumcomputing.stackexchange.com/a/7103/124)) to have errors E_j and E_k in which $E_j\mathcal{C}$ and $E_k\mathcal{C}$ are not orthogonal: the unit vectors in those subspaces may have non-zero overlap, or the two subspaces may even be the same. In this case, the only hope for us to be able to correct the error is that –; to the extent that the effect of E_j and E_k on the code-space cannot be distinguished –; the two images in a sense *agree*.

What I mean by this is: suppose that you had a state $|\psi\rangle \in \mathcal{C}$ which (unbeknownst to you) was affected by the error E_j, yielding the state $|\psi'\rangle = E_j|\psi\rangle$. Suppose that you then tried to determine whether it was affected instead by the error E_k, by measuring whether the state was in the space $E_k\mathcal{C}$. In general, this would disturb the state $|\psi'\rangle$. However, if $|\psi'\rangle$ collapsed to the state $|\psi''\rangle = E_k|\psi\rangle$, there would be no problem: you could simply correct the state as though E_k had happened, and thereby recover $|\psi\rangle$.

This is what is guaranteed by the Knill-Laflamme conditions: that for any state $|\psi\rangle \in \mathcal{C}$, the projection of $E_j|\psi\rangle$ onto $E_k\mathcal{C}$ will be proportional to $E_k|\psi\rangle$. Furthermore, that the constant of proportionality $\alpha_{j,k}$ involved is the same for all $|\psi\rangle \in \mathcal{C}$.

Thus, the Knill-Laflamme conditions imply that, to the extent that we cannot perfectly distinguish the effects of E_j and E_k on the entire subspace $\mathcal{C}$, the effects of E_j and E_k on *individual states in* $\mathcal{C}$ are also indistinguishable, so that we cannot confuse one error E_j for another error E_k plus some transformation on the original encoded state.

In summary

The Knill-Laflamme conditions on a set of error operators $\{E_j\}_j$ on a code-space $\mathcal{C}$ indicate that the effect of any one error E_j affects the entire space a uniform way (rescaling it uniformly and transforming it unitarily otherwise); and that to the extent that the effects of two errors E_j and E_k cannot be operationally distinguished from one another, their effect on the code-space is the same

as one another, so that you can measure which error occurred and then correct the state as though that error actually is the one that did occur.

This is the end of Niel's answer. I'll just comment that you can think of the KLV error detection condition as analogous to $d-1$ radii balls in the classical error detection case and the KLV error correction as analogous to two $\lceil \frac{d-1}{2} \rceil$ balls in the classical error correction case in a Hamming space. Of course, it's obvious from here that error correction has more stringent requirements than error detection!

2.3 Quantum Metric Space

You already saw the filtration of errors in Greg's class and his paper with Nik. I also told you the earthquake analogy. There's a correspondence between the error filtration and the quantum metric space given by:

$$d(x,x) = 0 \leftrightarrow I \in \mathcal{V}_0$$
$$d(x,y) = d(y,x) \leftrightarrow I \in \mathcal{V}_t^* = \mathcal{V}_t$$
$$d(x,z) \leq d(x,y) + d(y,z) \leftrightarrow \mathcal{V}_s\mathcal{V}_t \subseteq \mathcal{V}_{s+t}$$

The filtration of a quantum Hamming metric is

$$\mathcal{V}_t = \text{span}\{A_1 \otimes \cdots \otimes A_n \colon A_i \in M_2(\mathbb{C})) \text{ and } A_i = I_2 \text{ for all but at most t values of i}\}$$

This filtration models the error operators that corrupt at most t qubits for some t. Note that an operator $C \in \mathcal{V}_t$ can be a linear combination, or quantum superposition, of operators that corrupt different sets of t or fewer qubits. We saw earlier that it suffices to consider errors of the form where the A_i's are Paulis on individual qubits. X is a bit flip, Z is a phase flip and Y is a bit flip and a phase flip.

If we were just consider $G = SU(2)$, where $U = (U_1, U_2, \ldots, U_n)$ then

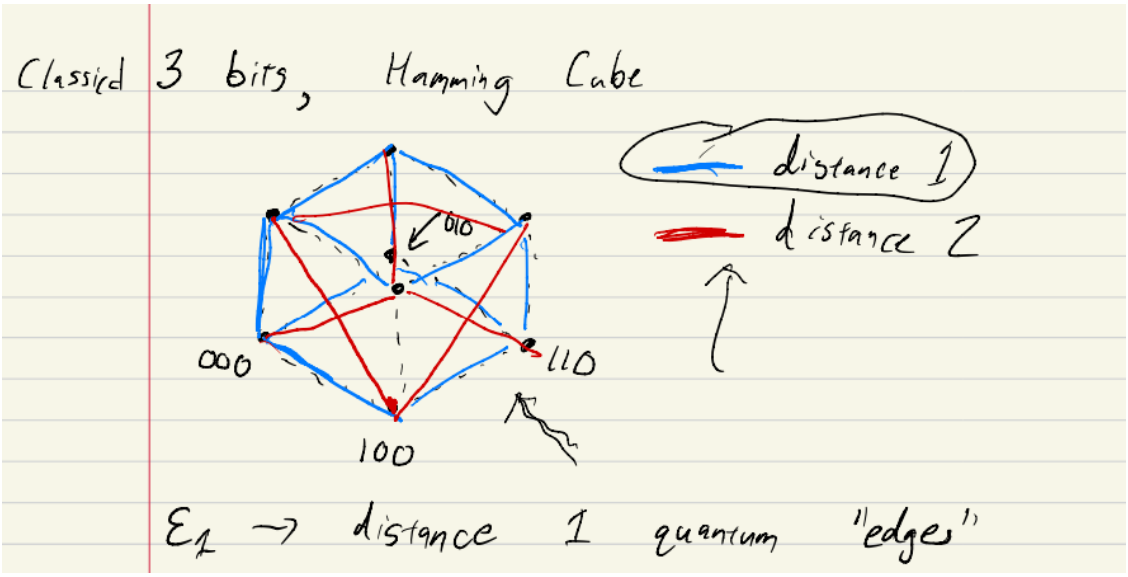
$$U|\psi_1\rangle|\psi_2\rangle \cdots |\psi_n\rangle = U_1|\psi_1\rangle U_2|\psi_2\rangle \cdots U_n|\psi_n\rangle$$

Well, this is nasty to deal with. Instead, it's easier to think of the Lie algebra action first and then deduce the rest by exponentiation (or what quantum mechanics folks call time evolution). For an $E \in \mathfrak{g} = \mathfrak{s}u(2)^{\otimes n}$ an each $E_i \in \mathfrak{s}u(2)$ then

$$E|\psi_1\rangle|\psi_2\rangle \ldots |\psi_n\rangle = (E_1|\psi_1\rangle)|\psi_2\rangle \ldots |\psi_n\rangle + (|\psi_1\rangle)(E_2|\psi_2\rangle) \ldots (|\psi_n\rangle) + \cdots$$

$$\mathcal{V}_1 = \mathbb{C}I_H \oplus \mathfrak{su}(2)$$
$$\mathcal{V}_t = (V_1)^t = \{\text{Span of products of at most } t \text{ elements of } \mathcal{V}_1\}$$

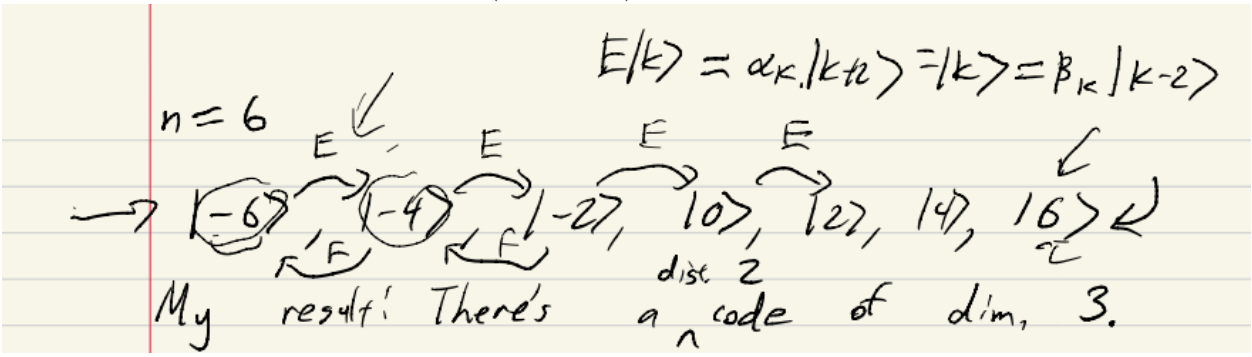
Greg: This is the concept of a quantum metric of Lie type. If L is the Lie algebra of a subgroup of the unitary group $U(d)$, then it has a complexification $L_{\mathbb{C}}$, and you can let that be the distance 1 space V_1 of a quantum metric. (If perhaps $L_{\mathbb{C}}$ does not contain I, then you can throw that in to make V_1.) Then the rest of the quantum metric is of quantum graph type, i.e., V_t is the linear span of the t-fold set product of V_1.



2.4 Rui Okada's work

If $H = V_n$, an $\mathfrak{su}(2)$ irrep. Then for $d = 2$ a quantum linear programming bound says that the dimension of a code must be $K \leq \frac{n}{2}$ (this is an existence bound – no code with dimension has been found yet). If $\dim(V_n) = n + 1$ then this can be interpreted as a (n, K, d) code. This is also called the Delsarte linear programming approach, which is similar to the approach Shor-Laflamme used in their paper Quantum Macwilliams Bound. There is also some work done on this topic by Chris Bumgardner, who was Greg's earlier grad student. The thesis is titled "$\mathfrak{g}$-metric spaces: Theory and Examples".

By the way, for representation theory of $\mathfrak{su}(2)$ look into Brian C Hall's textbook or more conveniently his arXiv notes on An Elementary Introduction to Groups and Representations. There's also a nice introduction by Ivan Vukovic (Google it!).



4. The Irreducible Representations of $\mathsf{su}(2)$

In this section we will compute (up to equivalence) all the finite-dimensional irreducible complex representations of the Lie algebra $\mathsf{su}(2)$. This computation is important for several reasons. In the first place, $\mathsf{su}(2) \cong \mathsf{so}(3)$, and the representations of $\mathsf{so}(3)$ are of physical significance. (The computation we will do here is found in every standard textbook on quantum mechanics, under the heading "angular momentum.") In the second place, the representation theory of $\mathsf{su}(2)$ is an illuminating example of how one uses commutation relations to determine the representations of a Lie algebra. In the third place, in determining the representations of general semisimple Lie algebras (Chapter 6), we will explicitly use the representation theory of $\mathsf{su}(2)$.

Now, every finite-dimensional complex representation π of $\mathsf{su}(2)$ extends by Prop. 5.5 to a complex-linear representation (also called π) of the complexification of $\mathsf{su}(2)$, namely $\mathsf{sl}(2;\mathbb{C})$.

PROPOSITION 5.8. *Let π be a complex representation of $\mathsf{su}(2)$, extended to a complex-linear representation of $\mathsf{sl}(2;\mathbb{C})$. Then π is irreducible as a representation of $\mathsf{su}(2)$ if and only if it is irreducible as a representation of $\mathsf{sl}(2;\mathbb{C})$.*

PROOF. Let us make sure we are clear about what this means. Suppose that π is a complex representation of the (real) Lie algebra $\mathsf{su}(2)$, acting on the complex space V. Then saying that π is irreducible means that there is no non-trivial invariant *complex* subspace $W \subset V$. That is, even though $\mathsf{su}(2)$ is a real Lie algebra, when considering complex representations we are interested only in complex invariant subspaces.

Now, suppose that π is irreducible as a representation of $\mathsf{su}(2)$. If W is a (complex) subspace of V which is invariant under $\mathsf{sl}(2;\mathbb{C})$, then certainly W is invariant under $\mathsf{su}(2) \subset \mathsf{sl}(2;\mathbb{C})$. Therefore $W = \{0\}$ or $W = V$. Thus π is irreducible as a representation of $\mathsf{sl}(2;\mathbb{C})$.

On the other hand, suppose that π is irreducible as a representation of $\mathsf{sl}(2;\mathbb{C})$, and suppose that W is a (complex) subspace of V which is invariant under $\mathsf{su}(2)$. Then W will also be invariant under $\pi(X+iY) = \pi(X)+i\pi(Y)$, for all $X, Y \in \mathsf{su}(2)$. Since every element of $\mathsf{sl}(2;\mathbb{C})$ can be written as $X + iY$, we conclude that in fact W is invariant under $\mathsf{sl}(2;\mathbb{C})$. Thus $W = \{0\}$ or $W = V$, so π is irreducible as a representation of $\mathsf{su}(2)$. □

Theorem 1: For any $l = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ there exists an irreducible representation of $\mathfrak{so}(3)$ of dimension $2l+1$ and any two irreducible representations of $\mathfrak{so}(3)$ of dimension $2l+1$ are isomorphic.

Theorem 2: *Proposition 5.8* in the image.

Greg's Note: A complex Lie algebra may have more than one pre-complexification, and that $\mathfrak{sl}(2,\mathbb{C})$ has two of them, namely $\mathfrak{su}(2)$ and $\mathfrak{sl}(2,\mathbb{R})$. However, theorem, $\mathfrak{sl}(2,\mathbb{R})$ has no non-trivial finite-dimensional unitary representations, although it does have infinite-dimensional ones.

Rui's idea for the code construction

Suppose we have 3 orthonormal code vectors $|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle \in V_n$ (irrep of $\mathfrak{su}(2)$ of dimension $n+1$). From the KLV error detection condition we have

$\langle\psi_i|E|\psi_i\rangle = \alpha_E$ for all $|\psi_i\rangle \in \mathcal{C}$

and

$\langle\psi_i|E|\psi_j\rangle = 0$ for all $|\psi_i\rangle, |\psi_j\rangle \in \mathcal{C}$ s.t. $i \neq j$.

A natural basis of V_n is given by the eigenvectors of H action, s.t. $H|k\rangle = |k\rangle$.

Where $\mathfrak{sl}_2 = \text{span}\{E = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, F = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\}$.

It can be deduced from the KLV condition that one possible option for $|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle$ is:

$|\psi_1\rangle = a|-6\rangle + b|4\rangle$

$|\psi_2\rangle = c|-4\rangle + d|6\rangle$

$|\psi_3\rangle = |0\rangle$

A code $\mathcal{C}$ detects an error if after E occurs then the word is no longer in $\mathcal{C}$.

2.5 Isometries

Isometries (in this context) are unitaries $U \in U(\mathcal{H})$ where if $\mathcal{E}_t$ is a degree t error then $UE_TE^\dagger$ is also degree t. These unitaries form a group $G \subseteq U(\mathcal{H})$, which is $U(2)^n \ltimes S_n$.

These are products of two types of operators: (1) $g \in U(2)^{\otimes n}$ and (2) g that permutes the tensor factors of a vector in H.

If $g = g_1 \otimes g_2 \otimes \cdots \otimes g_n$ and $E_t = E_1 \otimes E_2 \otimes \cdots \otimes E_n$

$$gE_tg^\dagger = (g_1E_1g_1^\dagger) \otimes (g_2E_2g_2^\dagger) \otimes \cdots \otimes (g_nE_ng_n^\dagger)$$

The happy coincidence is that the real vector space $\mathfrak{g} = \mathfrak{su}(2)$ is almost the same as $\mathcal{V}_1$ (complex), which are the distance 1 errors in quantum Hamming space.