We know look at two detailed applications of group theory in classical and quantum error correction.

1 Classical Error Correcting codes(Hamming Codes)

We would like to send k bits $m_1, m_2 \dots m_k$ over a noisy communication channel. The potential 2^k bits live in a k dimensional vector space $F_2^k = F_2 \otimes F_2 \dots F_2$ over the finite field F_2. What is done is that these 2^k codes are padded with n-k *parity* bits to make an n bit encoded message in an 2^n dimensional vector space. These extra parity bits are set so that they obey n-k linearly independent constraints known as *parity checks*. Each of these parity checks can be thought of as a vector in F_2^n: $H(i) = H_1(i) \dots H_n(i)$. The constraint is that all codewords must have a vanishing scalar product with the parity check $H(i)$ i.e $H(i).c = \sum_{j=1}^{j=n} H_j(i) c_j = 0$ with $i = 1, \dots n-k$. Theses parity checks can be assembled into a matrix called the *parity check matrix* as follows

$$H = \begin{pmatrix} H_1(1) & H_2(1) & H_3(1) \dots & H_n(1) \\ \vdots & & \dots & \vdots \\ H_1(n-k) & H_2(n-k) & H_3(n-k) \dots & H_n(n-k) \end{pmatrix} \quad (1)$$

We can then write our parity condition as merely

$$H.c = 0 \quad (2)$$

The linear code $\mathcal{C}$ with $(n-k) \times n$ parity check matrix H consists of the 2^k vectors $c \in F_2^n$ that satisfy the parity check condition. The vectors c are referred to as codewords. Note that the linear code C also forms a vector space as can be easily verified.

Since linear code C for a k dimensional subspace there is a basis for the codewords $b_1, \dots b_k$ i.e

$$c = \sum_{j=1}^{j=k} m_j b_j \quad (3)$$

It is useful to introduced what is called the generator matrix G which is formed by assembling it from the basis vectors as columns i,e

$$G = \begin{pmatrix} b_1 & \dots & b_k \\ \vdots & \dots & \vdots \end{pmatrix} \quad (4)$$

The codespace is then the column space of G. Thus the linear transformation G can be thought of as a map $G : F_2^k \to F_2^n$ that encodes the message, m, into the codeword c:

$$c = Gm \quad (5)$$

where

$$m = \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_k \end{pmatrix} \tag{6}$$

Combining 1.2 and 1.5 we find a different way of defining the linear error correcting code C

$$0 = HGm \tag{7}$$

Since the above equation is true for all $m \in F_2^k$ we have that

$$0 = HG \tag{8}$$

Now, note that the (n-k) columns of the H^T are linearly independent and from 1.8 we have that

$$0 = G^T H^T \tag{9}$$

So we can define a *dual* code space $C_\perp$ that has as its generator matrix H^T and its parity check matrix is G^T we thus re-write 1.9 as

$$H_\perp G_\perp = 0 \tag{10}$$

Since the rows of G^T are a basis of C 1.10 says that any basis for C is orthogonal to any basis for $C_\perp$. So any codeword in $C_\perp$ is perpendicular to any code word in C.

1.1 Errors, Hamming Weight and Distance

Because we are assuming a noisy communication channel, the codeword c, we send will in general not be the message received at the other end. Let us call the received message y. Having done this we can define an error vector e as such:

$$e = y - c = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} \tag{11}$$

If $e_i = 0$ fo tall i then no errors occurred. We further make the assumption that any errors that accumulate are not correlated and occur with probability p_i for all i. We would like the decoder to look at the received message y and guess what error $e_\star$ occurred and return the best guess $\tilde{c}$ which is defined as follows:

$$\tilde{c} = y - e_\star \tag{12}$$

Definition The Hamming weight wt(x) of a vector x is equal to the number of its non-zero components x_i

Definition The Hamming distance d(x,y) is the number of places that two vectors x and y differ.

Theorem 1.1 *d(x,y) = wt(x-y)*

Proof The places in which the two vectors are both zero when one takes the difference will not contribute to $wt(x-y)$. Thus the places in which they differ will contribute to $wt(x-y)$ but these are precisely the places that also contribute to $d(x, y)$,

An important property for a linear code C is the minimum distance d between any codewords.

$$d_{min} = \min d(c, c')_{c,c' \in C} \tag{13}$$

One can find a quick way of the finding d_{min} by noting the following: since $d(x, y) = wt(x - y)$ and $x - y \in C$ it then follows that there is a codeword in the linear code C that is equal to d_{min} so all one needs to do in order to find d_{min} is to find the codeword with the least hamming weight.

A linear code C with length n, dimension k and d_{min} is referred to as a [n,k,d] code.

Theorem 1.2 *A linear code C with minimum distance d can correct $t = [\frac{d-1}{2}]$ bit errors*

Proof Imagine a space in which points in that space are codewords in F_2^n and the distance between those points is equal to the hamming distance between the codewords. Let s $\in F_2^n$ and a sphere of radius r between the set of all vectors v such that $d(s, v) \leq r$. IF spheres of radius t are placed around each point then none of the spheres will overlap since $2t \leq d - 1$. Thus if in a received word y, no more than t errors occur, that word will lie in one and only one sphere and nearest neighbor decoding will correctly identify original codeword. If spheres of radius of t+1 are put around the points then some of the spheres will overlap since the diameter will be d+1. Thus if t+1 errors occur in a received word ,y,then it will lie in the overlap region and if $d(y, c') < t + 1$ then nearest neighbor decoding will give c' as the original codeword which will be wrong.

1.2 Error Detection and Correction

The parity check matrix H produces a linear transformation from $F_2^n \rightarrow F_2^{n-k}$. The image of the parity check matrix is called the error syndrome. It is important to note at this point that kernel of this mapping is our linear code C i.e

$$ker(H) = C \tag{14}$$

We can know define cosets of the linear code C as

$$x + C = \{x + c | \in C\} \tag{15}$$

for $x \in F_2^n$

Definition Let g+ C be a coset of$C$. The vector I of minimum weight in this coset is called the coset leader. If there exists more than one then randomly pick one.

The cosets allow us to define the quotient group F_2^n/C since F_2^n is an abelian group and C is a normal subgroup. By the first Isomorphism theorem $F_2^{n-k} \simeq F_2^n/C$. We thus have a one to one correspondence between the possible error syndromes and the available cosets. This allows one to define a maximum likelihood detection scheme. Let y be the received vector and let $g + C$ be the coset it belongs to. Thus y = g+x for some x in C. Let c be the transmitted codeword so that the error e = y-c = g + (x-x) so $e \in g + C$. Therefore the most probable error e_p is the vector in $g + C$ that has the minimum weight. The decoder thus returns $\tilde{c} = y - I$ as the most probable transmitted codeword.

In this application we used an abelian group as a setting for the codewords and identified the possible errors with elements of the quotient group.

2 Quantum Error Correction (Stabilizer formalism)

A quantum error correcting code(QECC) t that encodes k qubits into n qubits is defined through an encoding map ζ from the k-qubit Hilbert space H^k onto a 2^k dimensional subspace C_q of the n-qubit Hilbert space H^n. It is required to be unitary. We choose the single-qubit computational basis states(CB) to be the eigenstates of σ_z^j i.e

$$\sigma_z^j |\delta_j\rangle = (-1)^{\delta_j} |\delta_j\rangle \tag{16}$$

where j labels the qubits. The CB states for H^k are formed by talking all possible direct products of the single-qubits CB states:

$$|\delta\rangle \equiv |\delta_1 \ldots \delta_k\rangle = |\delta_1\rangle \otimes \ldots \otimes |\delta_k\rangle \tag{17}$$

This establishes a one to one correspondence between the unencoded states $|\delta\rangle = |\delta_1 \ldots \delta_k\rangle$ and encoded states $\overline{|\delta\rangle} = \overline{|\delta_1 \ldots \delta_k\rangle}$ so we have that:

$$\overline{|\delta_1 \ldots \delta_k\rangle} = \zeta |\delta_1 \ldots \delta_k\rangle \tag{18}$$

Also we have that $\sigma_z^j \to Z_j = \zeta \sigma_z^j \zeta^\dagger$

Quantum stabilizer codes, C_q is identified with a unique subspace that is fixed by elements of an abelian subgroup S known as the stabilizer group. More specifically we have that for all $|c\rangle \in C_q$

$$s |c\rangle = |c\rangle \tag{19}$$

The stabilizer group is a subgroup of the Pauli group which is a group consisting of n-fold distinct tensor product of the Pauli operators $\sigma_z, \sigma_y, \sigma_x$ and the identity operator

2.1 Stabilizer Group

The stabilizer group S is constructed from a set of n-k operators $g_1, \ldots g_{n-k}$ known as the generators of S because each element in the stabilizer group can be each element can be written as a unique product of powers of the generators

$$s = g_1^{p_1} \ldots g_{n-k}^{p_{n-k}} \tag{20}$$

Because the stabilizer group is abelian the generators are mutually commuting operators, Hermitian and unitary operators and of order 2. As a consequence of their order their eigenvalues are merely ± 1.

As the parent space H^n is 2^n dimensional we need n commuting operators to specify a unique state in the Hilbert space. In fact these n operators can be chosen to be the following: $g_1 \ldots g_{n-k}; Z_1 \ldots Z_k$ and the 2^n simultaneous eigenstates of these operators can be chosen to be the basis for H^n.These eigenstates can be labeled by strings $l = l_1, \ldots l_{n-k}; \delta = \delta_1 \ldots \delta_k$ so that

$$g_i \left|l, \delta\right\rangle = (-1)^{l_i} \left|l, \delta\right\rangle \tag{21}$$

$$Z_j \left|l, \delta\right\rangle = (-1)^{\delta_j} \left|l, \delta\right\rangle \tag{22}$$

where $i = 1 \ldots n-k$, $j = 1 \ldots k$ and l_i and $\delta_j = 0, 1$

Note that for a given string $l = l_1 \ldots l_{n-k}$ the set of 2^k eigenstates $|l; \delta\rangle$ span a subspace of H^n which can be labelled by the string l i.e $C_q(l_1, l_2 \ldots l_{n-k}) \equiv C_q(l)$ and provide a partition of H^n and the subspace determined by the stabilizer group is labeled as $C_q(00\ldots0)$. In other words the subspace determined by the stabilizer group are those elements in the Hilbert that are left invariant under the action of the stabilizer group elements so $s\left|c\right\rangle = \left|c\right\rangle \forall c \in C_q$

Modeling the noise in a quantum setting is more complicated than in the classical case. It is known to be exponentially hard to exactly simulate the noise acting for example in a quantum circuit. Thus if we are to study the noise in a quantum circuit and how to apply QECC we have to assume a model of the errors that can be efficiently simulated on a classical computer. For this discussion we will model the noise as simply the random application of σ_x, σ_y and σ_z with probabilities p_x, p_y and p_z respectively. The σ_x operator flips a qubit, σ_z potentially changes the phase of the qubit and σ_y does a combination of both. To prove the above behavior apply the operators to the state $|+\rangle = \dfrac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle = \dfrac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ which form a basis for the one qubit hilbert space.

So what we have is that elements of the stabilizer group can be labelled by bit strings of length $n-k$ $p = p_1 \ldots p_{n-k} \in F_2^{n-k}$

Theorem 2.1 *Let E be an error operator and C_q a quantum stabilizer code with generators $g_1, \ldots g_{n-k}$. The image $E(C_q)$ of C_q under E is $C(l)$ with $l = l_i \ldots l_{n-k}$.*

$$l_i = \begin{cases} 0 \ \textit{if} \ [E, g_i] = 0 \\ 1 \ \textit{if} \ \{E, g_i\} = 0 \end{cases}$$

Proof We assume what will be proved later, that is E either commutes or anti-commutes with g_i. So $g_i E |c\rangle = (-1)^{l_i} E g_i |c\rangle = (-1)^{l_i} E |c\rangle$ where $l_i = \{0, 1\}$. This means that corrupted state is an eigenvector of the generators. Now $\{|l; \delta\rangle : l \in H_2^{n-k}, \delta \in H_2^k\}$ span H_2^n and so $E |c\rangle = \sum_l \sum_\delta a(l; \delta) |l; \delta\rangle$. Because E commutes or anti-commutes with all the generators and $|l; \delta\rangle$ are eigenvectors of the generator, this all implies that sum over l does not exist and we only have one particular value of l. Therefore $E |c\rangle = \sum_\delta a(l; \delta) |l; \delta\rangle$. This means that the error taes the element of the code space to a specific subspace of $C(l)$. Therefore $E(C_q) \subset C(l)$. But these vector spaces are the same dimension so they are in fact equal. □

The lesson to take is that for each E we can attach a syndrome measurement $S(E) = l_1 \ldots l_{n-k}$

Example

Quantum Stabilizer Code for Phase Flip Channel

$\eta : H_s^1 \longmapsto C_q \subset H_2^3$

We need 2 generators $\{g_1, g_2\}$. There 8 possible errors. We want to protect our state against a phase flip. Thus there are three possible errors $E_1 = \sigma_z^1, E_2 = \sigma_z^2, E_3 = \sigma_z^3$. Depending on whether we have $|0\rangle, |1\rangle$ we will have a phase flip or not. The phase flip will show up as a change in the relative phase. So we choose eigenstates of $\sigma_x^1 \sigma_x^2$, $\sigma_x^2 \sigma_x^3$ as elements of our code subspace.

Error Syndrome
$S(E) = (1, 0)$

$$\{\sigma_z^1, \sigma_x^1 \sigma_x^2\} = 0$$
$$[\sigma_z^1, \sigma_x^2 \sigma_x^3] = 0$$

$S(E) = (1, 1)$

$$\{\sigma_z^2, \sigma_x^1 \sigma_x^2\} = 0$$
$$[\sigma_z^2, \sigma_x^2 \sigma_x^3] = 0$$

$S(E) = (0, 1)$

$$[\sigma_z^3, \sigma_x^1 \sigma_x^2] = 0$$
$$\{\sigma_z^3, \sigma_x^2 \sigma_x^3\} = 0$$

Elements of the stabilzer group are of the form $s(p) = g_1^{p_1} g^{p_2}\ p_1, p_2 \in \{0, 1\}$
So $S = \{I, \sigma_x^1 \sigma_x^3, \sigma_x^1 \sigma_x^2, \sigma_x^1 \sigma_x^3\}$

2.2 Deeper Study of Stabilizer Formalism

The Pauli group G_n has elements written as $e = i^\lambda \sigma_{j_1}^1 \otimes \cdots \otimes \sigma_{j_n}^n$ where $\lambda = \{0, 1, 2, 3\}, j_k = \{I, x, y, z\}$. Since $\sigma_y^k = -i\sigma_x^k \sigma_z^k$ we can always write the elements of the Pauli group as $e = i^\lambda \sigma_x(a)\sigma_z(b)$ where $a = a_1 \dots a_n$ and $b = b_1 \dots b_n$. a and b are n bit strings. We will worth with the quotient group $G_n/C\ C = \{\pm I, \pm iI\}$.

Theorem 2.2 *1. The order of G_n and G_n/C are $2^{2n+2} and 2^n$ respectively.*
2. $\forall e \in G_n$, $e^2 = \pm I, e^\dagger = \pm e$, $e^{-1} = e^\dagger$
3. $\forall e, f \in G_n$ either $[e, f] = 0$ or $\{e, f\} = 0$

Proof The order of the group is a trivial exercise in combinatorics. We move on the second claim. $e^2 = i^{2\lambda} \sigma_{j_1}^1 \otimes \cdots \otimes \sigma_{j_n}^n \sigma_{j_1}^1 \otimes \cdots \otimes \sigma_{j_n}^n = (-1)^\lambda (-1)^{a \cdot b} \sigma_x(a)^2 \sigma_x(b)^2 = (-1)^{\lambda + a \cdot b} I \implies \pm I$.
$e^\dagger = (-i)^\lambda \sigma_z(b)^\dagger \sigma_x(a)^\dagger = (-1)^{a \cdot b}(-1)^\lambda (i)^\lambda \sigma_x(a)\sigma_z(b) = \pm e$. Lastly, we prove that the elements either commute or anti-commute. Let $e = i^{\lambda_e} \sigma_x(a_e)\sigma_z(b_e)$, $f = i^{\lambda_f} \sigma_x(a_f)\sigma_z(b_f)$. The $ef = (i)^{\lambda_e + \lambda_f} \sigma_x(a_e)\sigma_z(b_e)\sigma_x(a_f)\sigma_z(b_f) = (i)^{\lambda_e + \lambda_f} (-1)^{b_e \cdot a_f} \sigma_x(a_e)\sigma_x(a_f)\sigma_z(b_e)\sigma$ $(i)^{\lambda_e + \lambda_f}(-1)^{b_e \cdot a_f}(-1)^{a_e \cdot b_f} \sigma_x(a_f)\sigma_z(b_f)\sigma_x(a_e)\sigma_z(b_e) = (-1)^{b_e \cdot a_f + a_e \cdot b_f} fe\ \Gamma = b_e \cdot a_f + a_e \cdot b_f \in \mathbb{Z}$ if Γ is even then the elements commute and if Γ is odd then the elements anti-commute $\square$

2.2.1 Errors

Errors having a vanishing syndrome $S(E) = 0$ commute wil all generators of the stabilizer group. Let $C(S)$ be the set of error $e \in G_n$ that commute with $e \in S$. $C(S)$ is called the centralizer. The centralizer is a subgroup of G_n ? 1. The centralizer has the identity since $[I, S] = 0 \forall s \in S$, 2. For $g \in C(S)$ we have $[g, s] = 0 \forall s \in S$ so $gs = sg$ so $g^{-1}ggs = sggg^{-1}$ but $g^2 = I \implies g^{-1}a = sg^{-1} \implies [g^{-1}, s = 0]$ so $g^{-1} \in C(S)$. Since the stabilizer grup is Abelian $S \subset C(S)$. If an even $e \in C(S)$ is in S the it needs no error correction if it is in $C(S - S)$ it will not be detectable. Further more $C(S)$ is a normal subgroup of G_n. The proof proceeds as follows.

Let $c \in C(S), s \in S, g \in G_n$. We have $csc^{-1} \in S$. Now look at $gc(g^{-1}sg)g^{-1} = g(g^{-1}sg)cg^{-1} = s(gcg^{-1}) \in C(S)$. Therefore $C(S)$ is normal. A slightly more abstract way is to notice that G_n acts on S by conjugation with the kernel being $C(S)$. Now a kernel is a group and moreover a normal subgroup. This all means we can create the quotient group $G_n/C(S)$.

Theorem 2.3 *Two elements e_1, e_2 are in the same coset iff they have the same error syndrome.*

Proof If they are in the same coset then they differ by an element c of $C(S)$. Let $e_2 = e_1 c$. Consider $e_1 g_i |d\rangle$ with $|d\rangle$ being a code word. So $e_2 c^{-1} g_i |d\rangle = e_2 g_i c^{-1} |d\rangle = e_2 g_i |d\rangle = e_1 g_i |d\rangle$. In other direction, $e_2 e_1 g = g e_2 e_1$ so the product belongs in the centralizer. Product commutes because e_1, e_2 have the same syndrome measurement. Therefore $\exists c \in C(S)$ such that $e_2 e_1 = c \implies e_1 = c e_2^{-1}$. Proving the theorem

The result is that different error syndromes are in 1-1 relation with the cosets of the centralizer.

Theorem 2.4 $|G_n : C(S)|$ *= number of cosets is* 2^{n-k}.

Proof Each coset corresponds to a unique syndrome measurement. The number of syndrome measurements is 2^{n-k}

Theorem 2.5 *The number of elements in* $C(S) = 2^{n+k+2}$

Proof By Lagrange's theorem $\frac{|G_n|}{C(S)} = |G_n : C(S)|$. So $\frac{G_n}{|G_n : C(S)|} = |C(S)| = \frac{2^{n+2}}{2^{n-k}} = 2^{n+k+2}$

Now $C(S)/S$ has elements that commute with the stabilizer generators but change the codeword. The turn out to be the logical operators.