Stabilizer formalism for qubits
2014 September 7

These are my evolving notes about the stabilizer formalism for qubits.

1. Pauli group and subgroups: Elementary properties

The *Pauli group* on N qubits (under matrix multiplication) is the set

$$\mathcal{P}_N = \left\{ \begin{pmatrix} \pm 1 \\ \pm i \end{pmatrix} \sigma_{\alpha_1} \otimes \cdots \otimes \sigma_{\alpha_N} \;,\;\; \alpha_k = 0,1,2,3 \right\} . \tag{1}$$

1. $|\mathcal{P}_N| = 4^{N+1} = 2^{2(N+1)}$.
2. Elements of $\mathcal{P}_N$ either commute or anticommute.
3. Only the four phasings of the identity operator have nonzero trace:

$$\mathrm{tr}(g) = 2^N \delta_{gI} - 2^N \delta_{g,-I} + i2^N \delta_{g,iI} - i2^N \delta_{g,-iI} \;.$$

4. All $g \in \mathcal{P}_N$ are unitary, i.e., $gg^\dagger = I$. If $g^2 = I$, g has eigenvalues ± 1, and if $g^2 = -I$, g has eigenvalues $\pm i$:

$$g^2 = I \quad \Longleftrightarrow \quad g = \pm \sigma_{\alpha_1} \otimes \cdots \otimes \sigma_{\alpha_N} \quad \Longleftrightarrow \quad g = g^{-1} = g^\dagger \;,$$

$$g^2 = -I \quad \Longleftrightarrow \quad g = \pm i \sigma_{\alpha_1} \otimes \cdots \otimes \sigma_{\alpha_N} \quad \Longleftrightarrow \quad g = -g^{-1} = -g^\dagger \;.$$

For any state $|\psi\rangle$, $|\langle\psi|g|\psi\rangle| \le 1$, with equality iff $|\psi\rangle$ is an eigenstate of g; moreover, $\langle\psi|g|\psi\rangle = 1$ iff $g|\psi\rangle = |\psi\rangle$.
5. Trivial, but important abelian and normal* subgroups of $\mathcal{P}_N$ are $J = \{\pm I\}$ and $K = \{\pm I, \pm iI\}$. The coset $gK = Kg = \{\pm g, \pm ig\}$, and thus the quotient group $\mathcal{P}_N/K$ is the abelian group whose product is Pauli matrix multiplication, but with the phases ignored.
6. Every element of $\mathcal{P}_N$ conjugates to itself or its negative under all elements of $\mathcal{P}_N$: $hgh^{-1} = hgh^\dagger = \pm g$, with the upper (lower) sign holding if g commutes (anticommutes) with h.
7. We can make the following trivial statement about any subgroup S of $\mathcal{P}_N$: $-I \in S$ iff there exists $g \in S$ such that $-g \in S$ (and then for all $g \in S$, $-g \in S$). [Proof: $\Longrightarrow$ $-I \in S$ implies $I \in S$ and $-I \in S$; $\Longleftarrow$ If $g^2 = -I$, then $-I \in S$, and if $g^2 = I$, then $g(-g) = -I \in S$.] We can summarize as follows:

$$\Big(-I \in S \quad \Longleftrightarrow \quad \exists\, g \in S \text{ such that } -g \in S \Big) \quad \Longrightarrow \quad \forall\, g \in S,\, -g \in S \;. \tag{2}$$

or, contrapositively,

$$-I \notin S \quad \Longleftrightarrow \quad g \in S \text{ implies } -g \notin S \;. \tag{3}$$

---

* H is a normal (invariant) subgroup of G if $ghg^{-1} \in H \;\forall\, g \in G$ and $h \in H$. Normal subgroups are important because left and right cosets are the same, and one can define a group operation for cosets and thus the quotient group G/H. All subgroups of an abelian group are normal.

8. If $\exists\, g \in S$ satisfying $g^2 = -I$, then $-I \in S$. The converse of this statement is not true, as is shown by the subgroup $S = \{\pm I, \pm X\}$. Contrapositively, if $-I \notin S$, then S contains no elements g satisfying $g^2 = -I$ and, hence, all the elements of S satisfy $g^2 = I$ (and $g \neq -I$).
9. If S contains anticommuting elements g and h, then $-I \in S$. [Proof: If $g^2 = -I$ or $h^2 = -I$, then $-I \in S$, and if $g^2 = I$ and $h^2 = I$, then $(gh)^2 = -g^2h^2 = -I$, implying again that $-I \in S$.] The converse of this statement is not true, as is shown by the abelian subgroup $S = \{\pm I, \pm X\}$. Contrapositively, if $-I \notin S$, then S is abelian.
10. Combining 7, 8, and 9, we can elaborate Eq. (3) as follows:

$$\begin{aligned} &\Big(-I \notin S \quad \Longleftrightarrow \quad g \in S \text{ implies } -g \notin S \Big) \\ &\qquad\qquad\qquad \Longrightarrow \quad S \text{ is abelian and } g^2 = I \text{ and } g \neq -I \ \forall\, g \in S. \end{aligned} \tag{4}$$

11. A subgroup of $\mathcal{P}_N$ is normal iff it contains $-I$.

2. Stabilizer subgroups and stabilized subspaces

A vector $|\psi\rangle$ is *stabilized* by $g \in \mathcal{P}_N$ if $g|\psi\rangle = |\psi\rangle$. If $|\psi\rangle$ is stabilized by g, then it is also stabilized by g^2. The vectors stabilized by all the elements of a subgroup S form a subspace V_S. S is called the *stabilizer* of V_S. V_S is the intersection of the subspaces stabilized by each of the elements of S. We let P_S denote the projector onto V_S.

1. The only vector stabilized by $-I$ is the zero vector.
2. The only vector stabilized by g satisfying $g^2 = -I$ is the zero vector.
3. The only vector stabilized by g and $-g$ is the zero vector.
3. The only vector stabilized by anticommuting g and h is the zero vector. (Proof: $|\psi\rangle = gh|\psi\rangle = -hg|\psi\rangle = -|\psi\rangle$.)
4. If S stabilizes a nontrivial subspace, then $-I \notin S$, which implies that S is an abelian subgroup such that $\forall\, g \in S$, $g^2 = I$ and $-g \notin S$.
5. If $g^2 = I$, g has eigenvalues ± 1, giving $g = P_{g,+1} - P_{g,-1}$, where $P_{g,\pm} = \frac{1}{2}(I \pm g)$ is the projector onto the ± 1 eigensubspace.
6. The projector onto the stabilized subspace V_S is

$$P_S = \frac{1}{|S|} \sum_{g \in S} g \; . \tag{5}$$

Proof: First note that if S contains an element that squares to $-I$, then $-I \in S$ and $g \in S$ implies $-g \in S$; in this case, S stabilizes only the zero vector, i.e., $P_S = 0$, which is what the sum gives. Now we can suppose that S contains only elements such that $g^2 = I$; in this case, P_S is a projector since it is a Hermitian operator that squares to itself:

$$P_S^2 = \frac{1}{|S|^2} \sum_{g,h \in S} gh = \frac{1}{|S|^2} \sum_{g \in S} \underbrace{\sum_{h \in S} gh}_{= \sum_{h \in S} h} = \frac{1}{|S|} \sum_{h \in S} h = P_S \; .$$

Then, to show that P_S projects onto V_S, notice that for any vector $|\psi\rangle$,

$$\begin{aligned} P_S|\psi\rangle = |\psi\rangle \iff |S| = |S|\langle\psi|P_S|\psi\rangle = \left|\sum_{g\in S}\langle\psi|g|\psi\rangle\right| \le \sum_{g\in S}|\langle\psi|g|\psi\rangle| \le |S| \\ \iff \langle\psi|g|\psi\rangle = 1 \ \ \forall \ g\in S \\ \iff g|\psi\rangle = |\psi\rangle \ \ \forall \ g\in S \\ \iff |\psi\rangle \in V_S \ . \end{aligned} \tag{6}$$

7. If $-I\in S$, then since $g\in S$ implies $-g\in S$, Eq. (5) gives $P_S=0$, consistent with what we already know. If $-I\notin S$, then the only traceful operator g in the sum (5) is $g=I$, so

$$\begin{pmatrix}\text{dimension}\\ \text{of } V_S\end{pmatrix} = \mathrm{tr}(P_S) = \frac{2^N}{|S|}\ , \tag{7}$$

thus establishing that $-I\notin S$ implies that S stabilizes a nontrivial subspace of dimension $2^N/|S|$.

8. Using 4 and 7, we can elaborate Eq. (4) yet further:

$$\begin{aligned} \Big(-I\notin S \quad \iff \quad g\in S \text{ implies } -g\notin S \quad \iff \quad S \text{ stabilizes a nontrivial subspace}\Big) \\ \implies \quad S \text{ is abelian and } g^2=I \text{ and } g\neq -I\ \forall\ g\in S. \end{aligned} \tag{8}$$

The subgroups that stabilize a nontrivial subspace are the ones we are mainly interested in.

8. The subgroup generated by elements $g_1,\ldots,g_l$ is denoted $\langle g_1,\ldots,g_l\rangle = S$. Some examples for one and two qubits, with the projector onto V_S (and basis vectors for V_S when it is nontrivial) also listed, are the following:

$$\begin{array}{rll}
\text{(i)} & \langle X\rangle = \{I,X\} & P_S=\frac{1}{2}(I+X) \qquad |+\rangle \\
\text{(ii)} & \langle X,-X\rangle = \{\pm I,\pm X\} & P_S=0 \\
\text{(iii)} & \langle iX\rangle = \{\pm I,\pm iX\} & P_S=0 \\
\text{(iv)} & \langle X,Z\rangle = \{\pm I,\pm X,\pm Z,\pm iY\} & P_S=0 \\
\text{(v)} & \langle X,Y,Z\rangle = \mathcal{P}_1 & P_S=0 \\
\text{(vi)} & \langle XX\rangle = \{II,XX\} & P_S=\frac{1}{2}\big(II+XX\big) \\
 & & \quad (|00\rangle+|11\rangle)/\sqrt{2},\ (|01\rangle+|10\rangle)/\sqrt{2} \\
\text{(vii)} & \langle XX,ZZ\rangle = \{II,XX,ZZ,-YY\} & P_S=\frac{1}{4}\big(II+XX+ZZ-YY\big) \\
 & & \quad =\frac{1}{2}\big(II+XX\big)\frac{1}{2}(II+ZZ) \\
 & & \quad (|00\rangle+|11\rangle)/\sqrt{2} \\
\text{(viii)} & \langle -XX\rangle = \{II,-XX\} & P_S=\frac{1}{2}\big(II-XX\big) \\
 & & \quad (|00\rangle-|11\rangle)/\sqrt{2},\ (|01\rangle-|10\rangle)/\sqrt{2} \\
\text{(ix)} & \langle -XX,ZZ\rangle = \{II,-XX,ZZ,YY\} & P_S=\frac{1}{4}\big(II-XX+ZZ+YY\big) \\
 & & \quad =\frac{1}{2}\big(II-XX\big)\frac{1}{2}(II+ZZ) \\
 & & \quad (|00\rangle-|11\rangle)/\sqrt{2} \\
\text{(x)} & \langle XX,ZZ,YY\rangle = \{\pm II,\pm XX,\pm ZZ,\pm YY\} & P_S=0 \\
\text{(xi)} & \langle XI,YI,ZI,IX,IZ\rangle = \mathcal{P}_2 & P_S=0
\end{array} \tag{9}$$

By convention, we never include the identity operator in the set of generators, and we say that the subgroup generated by an empty generator set is the trivial subgroup consisting of the identity operator. The entire group $\mathcal{P}_N$ is generated by the single-qubit Pauli operators for each qubit. Indeed, as the last example illustrates, an independent generator set consists of X, Y, and Z for one qubit and X and Z for the other qubits, for a total of $2N+1$ generators.

3. Check-vector representation

A useful representation of a Pauli group element g is in terms of a $2N$-dimensional *check vector* $r(g)$. We develop this representation in this section.

The check vector is based on the binary representation of the numbers 0,1,2, and 3, which are used to specify the Pauli matrices: $0 \to 0 = 00$, $3 = z \to 1 = 01$, $1 = x \to 2 = 10$, $2 = y \to 3 = 11$. This representation can be written formally as $\alpha \to r_1(\alpha)r_2(\alpha) = r(\alpha)$, where

$$\begin{aligned} r_1(\alpha) &= \delta_{x\alpha} + \delta_{y\alpha} = \begin{cases} 0, & \alpha = 0 \\ 0, & \alpha = 3 = z \\ 1, & \alpha = 1 = x \\ 1, & \alpha = 2 = y \end{cases} , \\ r_2(\alpha) &= \delta_{z\alpha} + \delta_{y\alpha} = \begin{cases} 0, & \alpha = 0 \\ 1, & \alpha = 3 = z \\ 0, & \alpha = 1 = x \\ 1, & \alpha = 2 = y \end{cases} . \end{aligned} \tag{10}$$

We can reconstruct a Pauli matrix from its representation using

$$\begin{aligned} \sigma_\alpha &= i^{r_1(\alpha)r_2(\alpha)} X^{r_1(\alpha)} Z^{r_2(\alpha)} \\ &= i^{r_1(\alpha)r_2(\alpha)} (-1)^{r_1(\alpha)r_2(\alpha)} Z^{r_2(\alpha)} X^{r_1(\alpha)} \\ &= (-i)^{r_1(\alpha)r_2(\alpha)} Z^{r_2(\alpha)} X^{r_1(\alpha)} . \end{aligned} \tag{11}$$

We could use the indices r_1 and r_2 to label Pauli matrices, thus using a quadruple-valued representation in which r_1 and r_2 take on values $0, 1, 2, 3$ (i.e., their values are taken mod 4). This is the standard approach when talking about discrete displacement operators in arbitrary dimensions. For qubits this representation works in the following way:

$$||\,\sigma_{r_1 r_2}|| = \begin{pmatrix} I & Z & I & Z \\ X & Y & -X & -Y \\ I & -Z & I & -Z \\ X & -Y & -X & Y \end{pmatrix} . \tag{12}$$

In situations where we can afford to let the phases fall where they may, this representation is certainly the most convenient. In the considerations here, however, we are concerned with the actual phase in front of a tensor product of Pauli operators. We can't afford to let a Pauli operator have different phases multiplying it depending on the circumstances. Thus we insist throughout that in the representation (11), r_1 and r_2 are either 0 or 1 (i.e., their values are taken mod 2).

Multiplication of Pauli matrices corresponds to bitwise mod-2 addition of the two-bit representations, plus some additional phase information:

$$\begin{aligned}\sigma_\alpha\sigma_\beta &= i^{\,r_1(\alpha)r_2(\alpha)+r_1(\beta)r_2(\beta)}X^{r_1(\alpha)}Z^{r_2(\alpha)}X^{r_1(\beta)}Z^{r_2(\beta)}\\ &= (-1)^{r_2(\alpha)r_1(\beta)}i^{\,r_1(\alpha)r_2(\alpha)+r_1(\beta)r_2(\beta)}X^{r_1(\alpha)+r_1(\beta)}Z^{r_2(\alpha)+r_2(\beta)}\\ &= i^{\,-r_1(\alpha)r_2(\beta)+r_2(\alpha)r_1(\beta)}i^{\,(r_1(\alpha)+r_1(\beta))(r_2(\alpha)+r_2(\beta))}X^{r_1(\alpha)+r_1(\beta)}Z^{r_2(\alpha)+r_2(\beta)}\\ &= i^{\,-r(\alpha)\wedge r(\beta)}i^{\,(r_1(\alpha)+r_1(\beta))(r_2(\alpha)+r_2(\beta))}X^{r_1(\alpha)+r_1(\beta)}Z^{r_2(\alpha)+r_2(\beta)}\;,\end{aligned}\tag{13}$$

The two-bit representations are added bitwise mod 2 automatically because of their appearance in the exponents of X and Z, but the phases involve mod-4 arithmetic in the exponents of i. The phases in Eq. (13) are given by a symmetric product $\big(r_1(\alpha)+r_1(\beta)\big)\big(r_2(\alpha)+r_2(\beta)\big)$ and an antisymmetric (skew) product

$$\begin{aligned}r(\alpha)\wedge r(\beta) &= r_1(\alpha)r_2(\beta)-r_2(\alpha)r_1(\beta)\\ &= (\delta_{x\alpha}+\delta_{y\alpha})(\delta_{z\beta}+\delta_{y\beta}\big)-(\delta_{z\alpha}+\delta_{y\alpha})(\delta_{x\beta}+\delta_{y\beta}\big)\\ &= \delta_{x\alpha}\delta_{z\beta}+\delta_{x\alpha}\delta_{y\beta}+\delta_{y\alpha}\delta_{z\beta}-\delta_{z\alpha}\delta_{x\beta}-\delta_{z\alpha}\delta_{y\beta}-\delta_{y\alpha}\delta_{x\beta}\\ &= \begin{cases}0, & \text{if } \sigma_\alpha \text{ and } \sigma_\beta \text{ are not different “spatial” Pauli matrices,}\\ +1, & \text{if } \alpha\beta=xy,\ yz, \text{ or } xz,\\ -1, & \text{if } \alpha\beta=yx,\ zy, \text{ or } zx.\end{cases}\end{aligned}\tag{14}$$

The skew product, calculated mod 2,

$$r(\alpha)\wedge r(\beta) \bmod 2 = \begin{cases}0, & \text{if } \sigma_\alpha \text{ and } \sigma_\beta \text{ are not different “spatial” Pauli matrices,}\\ 1, & \text{if } \sigma_\alpha \text{ and } \sigma_\beta \text{ are different “spatial” Pauli matrices,}\end{cases}\tag{15}$$

determines whether σ_α and σ_β commute or anticommute. Once we're calculating the skew product mod 2, the minus sign in the definition can be changed to a plus sign without any effect. If we let σ_γ denote the Pauli operator produced by the product, then we have $r_1(\gamma)=\big(r_1(\alpha)+r_1(\beta)\big) \bmod 2$ and $r_2(\gamma)=\big(r_2(\alpha)+r_2(\beta)\big) \bmod 2$, i.e., $r(\gamma)=r(\alpha)+r(\beta) \bmod 2$, where the addition is done bitwise.

To extract the phase in front of σ_γ, we note that

$$\begin{aligned}\sigma_\gamma = i^{\,r_1(\gamma)r_2(\gamma)}X^{r_1(\gamma)}Z^{r_2(\gamma)} &= i^{\,[(r_1(\alpha)+r_1(\beta)) \bmod 2][(r_2(\alpha)+r_2(\beta)) \bmod 2]}X^{r_1(\alpha)+r_1(\beta)}Z^{r_2(\alpha)+r_2(\beta)}\\ &= i^{\,(r_1(\alpha)+r_1(\beta))(r_2(\alpha)+r_2(\beta)) \bmod 2}X^{r_1(\alpha)+r_1(\beta)}Z^{r_2(\alpha)+r_2(\beta)}\;,\end{aligned}\tag{16}$$

where we use the fact that

$$(a \bmod 2)(b \bmod 2) = ab \bmod 2\;,\tag{17}$$

a property that is special to mod-2 arithmetic. What modular arithmetic implies is that

$$\begin{aligned}(a \bmod n + b \bmod n) \bmod n &= (a+b) \bmod n\;,\\ [(a \bmod n)(b \bmod n)] \bmod n &= ab \bmod n\;.\end{aligned}\tag{18}$$

What these say is that once you mod the result of a set of arithmetic operations, you are entitled to do the entire set of operations mod n. On the other hand, it is not generally true that

$$\begin{aligned}a \bmod n + b \bmod n &= (a+b) \bmod n\;,\\ (a \bmod n)(b \bmod n) &= ab \bmod n\;.\end{aligned}\tag{19}$$

The only case where these properties hold is for the case of multiplication mod 2, as in Eq. (17). In particular, it is not true that $a \bmod 2 + b \bmod 2 = (a+b) \bmod 2$. We need Eq. (17) to simplify Eq. (16) to the last line, because arithmetic in the exponent of i is done mod 4, not mod 2.

Plugging Eq. (16) into Eq. (13) gives

$$\begin{aligned}\sigma_\alpha\sigma_\beta &= i^{-r(\alpha)\wedge r(\beta)} i^{(r_1(\alpha)+r_1(\beta))(r_2(\alpha)+r_2(\beta)) - [(r_1(\alpha)+r_1(\beta))(r_2(\alpha)+r_2(\beta)) \bmod 2]}\sigma_\gamma \\ &= i^{-r(\alpha)\wedge r(\beta)} i^{(r_1(\alpha)+r_1(\beta))(r_2(\alpha)+r_2(\beta)) - \overline{(r_1(\alpha)+r_1(\beta))(r_2(\alpha)+r_2(\beta))}}\sigma_\gamma \;.\end{aligned} \tag{20}$$

Here we introduce the notation $\overline{a} = a \bmod 2$, which we use throughout the following whenever it is convenient. The phase of the product is the factor in front of σ_γ. Whereas the first phase factor can have any of the four possible values, the second factor is always ± 1. To see this, define the map

$$E(n) \equiv n - \overline{n} \;, \tag{21}$$

which takes the even part of n, i.e., subtracts 1 if n is odd and 0 if n is even, thereby mapping n to the nearest even integer less than or equal to n. Notice that

$$\sum_j E(n_j) = \sum_j n_j - \sum_j \overline{n}_j = \sum_j n_j - (\#\text{ of odd } n_j\text{'s}) \;. \tag{22}$$

Using this notation, we can write

$$\sigma_\alpha\sigma_\beta = i^{-r(\alpha)\wedge r(\beta)} i^{E[(r_1(\alpha)+r_1(\beta))(r_2(\alpha)+r_2(\beta))]}\sigma_\gamma \;. \tag{23}$$

The multiplication table for $(r_1(\alpha)+r_1(\beta))(r_2(\alpha)+r_2(\beta))$ is

$$\begin{array}{cccccc} & \beta & 0 & 3=z & 1=x & 2=y \\ & r_1(\beta)r_2(\beta) & 00 & 01 & 10 & 11 \\ & & & & & \\ \alpha & r_1(\alpha)r_2(\alpha) & & & & \\ 0 & 00 & 0 & 0 & 0 & 1 \\ 3=z & 01 & 0 & 0 & 1 & 2 \\ 1=x & 10 & 0 & 1 & 0 & 2 \\ 2=y & 11 & 1 & 2 & 2 & 4 \end{array} \;. \tag{24}$$

Notice that $(r_1(\alpha)+r_1(\beta))(r_2(\alpha)+r_2(\beta))$ is odd iff $\sigma_\gamma = Y$; this property generalizes in ways that we return to repeatedly below. The multiplication table gives us

$$E[(r_1(\alpha)+r_1(\beta))(r_2(\alpha)+r_2(\beta))] = \begin{cases} 2\;, & \alpha\beta = xy,\ yx,\ yz, \text{ or } zy, \\ 4\;, & \alpha\beta = yy, \\ 0\;, & \text{otherwise.} \end{cases} \tag{25}$$

It is clear from Eqs. (14) and (25) how this method manages to get the phase right: the skew-product term gets the right phase except for giving the wrong sign in the cases $\alpha\beta = xy$, yx, yz, and zy, a sign that is reversed by the term involving the symmetric product. At this point this is an absurdly complicated way of writing the phase, but the method becomes more useful when we consider arbitrary products of Pauli group elements.

Going one step further to a triple product of Pauli operators, we get

$$
\begin{aligned}
\sigma_\alpha\sigma_\beta\sigma_\gamma = \, & i^{-r(\alpha)\wedge r(\beta)-r(\alpha)\wedge r(\gamma)-r(\beta)\wedge r(\gamma)} i^{(r_1(\alpha)+r_1(\beta)+r_1(\gamma))(r_2(\alpha)+r_2(\beta)+r_2(\gamma))} \\
& \times X^{r_1(\alpha)+r_1(\beta)+r_1(\gamma)} Z^{r_2(\alpha)+r_2(\beta)+r_2(\gamma)} ,
\end{aligned}
\tag{26}
$$

and at this point we can use induction to show that the general product is given by

$$
\begin{aligned}
\sigma_{\alpha_1}\cdots\sigma_{\alpha_n} &= i^{-\sum_{j<k} r(\alpha_j)\wedge r(\alpha_k)} i^{\sum_{j,k} r_1(\alpha_j) r_2(\alpha_k)} \underbrace{X^{r_1(\alpha_1)+\cdots+r_1(\alpha_n)} Z^{r_2(\alpha_1)+\cdots+r_2(\alpha_n)}}_{= i^{-\overline{\sum_{j,k} r_1(\alpha_j) r_2(\alpha_k)}}\sigma_\beta} \\
&= i^{-\sum_{j<k} r(\alpha_j)\wedge r(\alpha_k)} i^{\sum_{j,k} r_1(\alpha_j) r_2(\alpha_k) - \overline{\sum_{j,k} r_1(\alpha_j) r_2(\alpha_k)}} \sigma_\beta \\
&= i^{-\sum_{j<k} r(\alpha_j)\wedge r(\alpha_k)} i^{E\left(\sum_{j,k} r_1(\alpha_j) r_2(\alpha_k)\right)} \sigma_\beta ,
\end{aligned}
\tag{27}
$$

where

$$
r(\beta) = \overline{\sum_j r(\alpha_j)} . \tag{28}
$$

The phase in front of σ_β now has an explicit formula in terms of the two-bit representations. The product

$$
\sum_{j,k} r_1(\alpha_j) r_2(\alpha_k) = \sum_j r_1(\alpha_j) \sum_k r_2(\alpha_k) \tag{29}
$$

is odd iff both $\sum_j r_1(\alpha_j)$ and $\sum_j r_2(\alpha_j)$ are odd, which is equivalent to saying that

$$
r_1(\beta) = \overline{\sum_j r_1(\alpha_j)} = 1 = \overline{\sum_j r_2(\alpha_j)} = r_2(\beta) , \tag{30}
$$

i.e., $\sigma_\beta = Y$.

To define the check vector, we write a group element as

$$
g = i^s \sigma_{\alpha_1} \otimes \cdots \otimes \sigma_{\alpha_N} .
$$

Ignoring the phase i^s and using the two-bit representation, we now define two N-dimensional row vectors, $r_1(g)$ and $r_2(g)$, whose kth components are the two bits of the representation of σ_{α_k}:

$$
\begin{aligned}
[r_1(g)]_k &= r_1(\alpha_k) = \delta_{x\alpha_k} + \delta_{y\alpha_k} , \\
[r_2(g)]_k &= r_2(\alpha_k) = \delta_{z\alpha_k} + \delta_{y\alpha_k} .
\end{aligned}
\tag{31}
$$

The check vector is simply the $2N$-dimensional row vector formed by stringing $r_1(g)$ and $r_2(g)$ together along a single row:

$$
r(g) = \begin{pmatrix} r_1(g) & r_2(g) \end{pmatrix} = \begin{pmatrix} [r_1(g)]_1 & \cdots & [r_1(g)]_N & [r_2(g)]_1 & \cdots & [r_2(g)]_N \end{pmatrix} . \tag{32}
$$

Using Eq. (11), we can write the arbitrary Pauli group element g in terms of its check vector:

$$g = i^s i^{\sum_k r_1(\alpha_k) r_2(\alpha_k)} \bigotimes_{k=1}^{N} X^{r_1(\alpha_k)} Z^{r_2(\alpha_k)} . \tag{33}$$

The relation between Pauli matrix multiplication and addition in the two-bit representation gives us immediately that

$$r(gh) = (r(g) + r(h)) \bmod 2 = \overline{r(g) + r(h)} , \tag{34}$$

where the addition is done bitwise mod 2. With this natural kind of addition, the check-vector space is a $2N$-dimensional vector space over the binary field $\{0,1\}$, which we call R_N; the 2^{2N} vectors in this vector space are in 1-1 correspondence with the 4^N products of Pauli matrices, i.e., with the 4^N elements of the quotient group $\mathcal{P}_N/K$. The check vectors make up a group under vector addition that is isomorphic to the (abelian) quotient group $\mathcal{P}_N/K$. They allow us to analyze products in $\mathcal{P}_N$, modulo the phases, in terms of linear algebra in the check-vector space.

To get the phase of a product, we have to be more careful. The phase of the product of $g = i^s \sigma_{\alpha_1} \otimes \cdots \otimes \sigma_{\alpha_N}$ and $h = i^t \sigma_{\beta_1} \otimes \cdots \otimes \sigma_{\beta_N}$ can be extracted from the following general expression:

$$\begin{aligned} gh &= i^{\,s+t} i^{\,-r(g)\wedge r(h)} i^{\,(r_1(g)+r_1(h))\bullet(r_2(g)+r_2(h))} \bigotimes_{k=1}^{N} X^{r_1(\alpha_k)+r_1(\beta_k)} Z^{r_2(\alpha_k)+r_2(\beta_k)} \\ &= i^{\,s+t} i^{\,-r(g)\wedge r(h)} i^{\,(r_1(g)+r_1(h))\bullet(r_2(g)+r_2(h))} \bigotimes_{k=1}^{N} X^{[r_1(g)]_k+[r_1(h)]_k} Z^{[r_2(g)]_k+[r_2(h)]_k} . \end{aligned} \tag{35}$$

Here the symmetric part of the phase is given by the ordinary dot product

$$r_1(g) \bullet r_2(h) = r_1(g) r_2^T(h) = \sum_{k=1}^{N} [r_1(g)]_k [r_2(h)]_k = \sum_{k=1}^{N} r_1(\alpha_k) r_2(\beta_k) , \tag{36}$$

and the skew product is the symplectic product

$$\begin{aligned} r(g) \wedge r(h) \equiv r(g) \Lambda r^T(h) &= \begin{pmatrix} -r_2(g) & r_1(g) \end{pmatrix} \begin{pmatrix} r_1(h) \\ r_2(h) \end{pmatrix} \\ &= r_1(g) \bullet r_2(h) - r_2(g) \bullet r_1(h) \\ &= \sum_{k=1}^{N} [r_1(g)]_k [r_2(h)]_k - [r_2(g)]_k [r_1(h)]_k \\ &= \sum_{k=1}^{N} r_1(\alpha_k) r_2(\beta_k) - r_2(\alpha_k) r_1(\beta_k) \\ &= \sum_{k=1}^{N} r(\alpha_k) \wedge r(\beta_k) \end{aligned} \tag{37}$$

where we introduce the fundamental $2N \times 2N$ symplectic matrix

$$\Lambda = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} . \tag{38}$$

Calculated mod 2, the skew product determines whether the number of anticommuting elements in the Pauli products of g and h is even or odd and thus determines whether g and h commute:

$$\overline{r(g) \wedge r(h)} = \begin{cases} 0, & \text{if } g \text{ and } h \text{ commute,} \\ 1, & \text{if } g \text{ and } h \text{ anticommute.} \end{cases} \tag{39}$$

Once we're calculating the skew product mod 2, we can change the minus sign in the symplectic matrix to a plus sign without changing the skew product.

Now we're ready to get the phase in front of the product by writing

$$\begin{aligned} gh &= i^{s+t} i^{-r(g)\wedge r(h)} i^{(r_1(g)+r_1(h))\circ(r_2(g)+r_2(h))} \\ &\quad \times \bigotimes_{k=1}^{N} i^{\overline{[r_1(g)]_k+[r_1(h)]_k)([r_2(g)]_k+[r_2(h)]_k}} X^{[r_1(g)]_k+[r_1(h)]_k} Z^{[r_2(g)]_k+[r_2(h)]_k} , \end{aligned} \tag{40}$$

where we define a special sort of dot product, called the *circle product*,

$$\begin{aligned} a \circ b \equiv \sum_{k=1}^{N} E(a_k b_k) &= \sum_{k=1}^{N} a_k b_k - \overline{a_k b_k} \\ &= a \bullet b - \underbrace{\sum_{k=1}^{N} \overline{a_k}\,\overline{b_k}}_{=\,\overline{a} \bullet \overline{b}} \\ &= a \bullet b - (\#\text{ of } k \text{ such that } \overline{a_k} = 1 = \overline{b_k}) , \end{aligned} \tag{41}$$

which is clearly symmetric. Notice that $a \circ b \neq a \bullet b - \overline{a \bullet b}$; this is precisely the step we cannot take in an exponent of i, where arithmetic is done mod 4. Obviously, the circle product is always even. Since the circle product is symmetric, it is obvious that the commutation of g and h is determined by whether the skew product is even or odd, as already noted.

We can write the circle product of Eq. (40) as

$$\begin{aligned} (r_1(g) + r_1(h)) \circ (r_2(g) + r_2(h)) &= (r_1(g) + r_1(h)) \bullet (r_2(g) + r_2(h)) \\ &\quad - \overline{r_1(g) + r_1(h)} \bullet \overline{r_2(g) + r_2(h)} \\ &= (r_1(g) + r_1(h)) \bullet (r_2(g) + r_2(h)) - r_1(gh) \bullet r_2(gh) \\ &= (r_1(g) + r_1(h)) \bullet (r_2(g) + r_2(h)) \\ &\quad - (\#\text{ of } k \text{ such that } [r_1(gh)]_k = 1 = [r_2(gh)]_k) \\ &= (r_1(g) + r_1(h)) \bullet (r_2(g) + r_2(h)) - (\#\text{ of } Y\text{s in } gh) . \end{aligned} \tag{42}$$

We can use Eq. (27) to write an expression for an arbitrary product of Pauli group elements. Letting $g = g_1 \cdots g_n$, with $g_j = i^{s_j}\sigma_{\alpha_{j1}} \otimes \cdots \otimes \sigma_{\alpha_{jN}}$, we have

$$
\begin{aligned}
g &= i^{\sum_{j=1}^n s_j} i^{-\sum_{j<k} r(g_j)\wedge r(g_k)} i^{(r_1(g_1)+\cdots+r_1(g_n))\bullet(r_2(g_1)+\cdots+r_2(g_n))} \\
&\quad \times \bigotimes_{k=1}^{N} X^{r_1(\alpha_{1k})+\cdots+r_1(\alpha_{nk})} Z^{r_2(\alpha_{1k})+\cdots+r_2(\alpha_{nk})} \\
&= i^{\sum_{j=1}^n s_j} i^{-\sum_{j<k} r(g_j)\wedge r(g_k)} i^{(r_1(g_1)+\cdots+r_1(g_n))\bullet(r_2(g_1)+\cdots+r_2(g_n))} \\
&\quad \times \bigotimes_{k=1}^{N} X^{[r_1(g_1)]_k+\cdots+[r_1(g_n)]_k} Z^{[r_2(g_1)]_k+\cdots+[r_2(g_n)]_k} \\
&= i^{\sum_{j=1}^n s_j} i^{-\sum_{j<k} r(g_j)\wedge r(g_k)} i^{(r_1(g_1)+\cdots+r_1(g_n))\circ(r_2(g_1)+\cdots+r_2(g_n))} \\
&\quad \times \bigotimes_{k=1}^{N} i^{\overline{([r_1(g_1)]_k+\cdots+[r_1(g_n)]_k)([r_2(g_1)]_k+\cdots+[r_2(g_n)]_k)}} X^{[r_1(g_1)]_k+\cdots+[r_1(g_n)]_k} Z^{[r_2(g_1)]_k+\cdots+[r_2(g_n)]_k} \\
&= i^{\sum_{j=1}^n s_j} i^{-\sum_{j<k} r(g_j)\wedge r(g_k)} i^{(r_1(g_1)+\cdots+r_1(g_n))\circ(r_2(g_1)+\cdots+r_2(g_n))} \sigma_{\beta_1} \otimes \cdots \otimes \sigma_{\beta_k} \ .
\end{aligned}
\tag{43}
$$

We have

$$
r(g) = \left(\sum_{j=1}^{N} r(g_j)\right) \bmod 2 = \overline{\sum_{j=1}^{N} r(g_j)} \ , \tag{44}
$$

with $[r(g)]_k = r(\beta_k)$ and with the three factors in front of the tensor product giving the phase in front of the product of Pauli operators in g. Once again, we can re-write the circle product in a variety of forms:

$$
\begin{aligned}
&(r_1(g_1) + \cdots + r_1(g_n)) \circ (r_2(g_1) + \cdots + r_2(g_n)) \\
&\qquad = (r_1(g_1) + \cdots + r_1(g_n)) \bullet (r_2(g_1) + \cdots + r_2(g_n)) \\
&\qquad\quad - \overline{r_1(g_1) + \cdots + r_1(g_n)} \bullet \overline{r_2(g_1) + \cdots + r_2(g_n))} \\
&\qquad = (r_1(g_1) + \cdots + r_1(g_n)) \bullet (r_2(g_1) + \cdots + r_2(g_n)) - \underbrace{r_1(g) \bullet r_2(g)}_{=(\#\text{ of }Y\text{s in }g)} \\
&\qquad = \sum_{j,k} r_1(g_j) \bullet r_2(g_k) - r_1(g) \bullet r_2(g) \\
&\qquad = \sum_{j<k} r_1(g_j) \bullet r_2(g_k) + \sum_{j>k} r_1(g_j) \bullet r_2(g_k) + \sum_{j} r_1(g_j) \bullet r_2(g_j) - r_1(g) \bullet r_2(g) \ .
\end{aligned}
\tag{45}
$$

This result allows us to manipulate the phase factor in Eq. (43),

$$
F = i^{\sum_{j=1}^n s_j} i^{-\sum_{j<k} r(g_j)\wedge r(g_k)} i^{(r_1(g_1)+\cdots+r_1(g_n))\circ(r_2(g_1)+\cdots+r_2(g_n))} \ , \tag{46}
$$

into other, often more useful forms. Noting that

$$
\sum_{j<k} r(g_j) \wedge r(g_k) = \sum_{j<k} r_1(g_j) \bullet r_2(g_k) - r_2(g_j) \bullet r_1(g_k) = \sum_{j<k} r_1(g_j) \bullet r_2(g_k) - \sum_{j>k} r_1(g_j) \bullet r_2(g_k) \ , \tag{47}
$$

we have

$$
\begin{aligned}
-\sum_{j<k} r(g_j) \wedge r(g_k) &+ (r_1(g_1) + \cdots + r_1(g_n)) \circ (r_2(g_1) + \cdots + r_2(g_n)) \\
&= 2\sum_{j>k} r_1(g_j) \bullet r_2(g_k) + \sum_j r_1(g_j) \bullet r_2(g_j) - r_1(g) \bullet r_2(g)\;.
\end{aligned}
\tag{48}
$$

Thus the phase factor becomes

$$
F = i^{\sum_{j=1}^n s_j} i^{2\sum_{j>k} r_1(g_j)\bullet r_2(g_k)} i^{\sum_j r_1(g_j)\bullet r_2(g_j)} i^{-r_1(g)\bullet r_2(g)}\;. \tag{49}
$$

In this form, which reverses the steps we took to put the phase factor in terms of symmetric and skew products, we can do mod-2 arithmetic in the exponent of $i^2 = -1$, which turns out to be useful in some contexts.

Throughout the remainder of this document, we generally assume that check-vector arithmetic is done mod 2, without explicitly noting it. The exception occurs when we are dealing with phases, where we do not do modular arithmetic except where it is noted explicitly.

4. Stabilizer formalism. Basics

Our objective now is to characterize the subgroups that stabilize nontrivial subspaces in terms of properties of an independent set of generators or, better yet, in terms of the check vectors of an independent set of generators.

We can immediately rephrase Eq. (4) in terms of generators as follows:

$$
\Big(-I \notin S = \langle g_1, \ldots, g_l \rangle \quad \Longleftrightarrow \quad -g_j \notin S\ \forall\ g_j\Big) \quad \Longrightarrow \quad \begin{array}{c}\text{The generators commute}\\ \text{and}\\ g_j \neq -I \text{ and } g_j^2 = I\ \forall\ g_j.\end{array} \tag{50}
$$

This is a trivial rephrasing, which is not very useful, because to use it one must verify for each generator that $-g_j$ is not in the entirety of S. It is not sufficient just to verify that $-g_j$ is not among the generators, as examples (iii), (iv), (v), (x), and (xi) in Eq. (9) show. Moreover, example (x) also shows that adding the further conditions on the right of Eq. (50) does not make it sufficient to verify that $-g_j$ is not among the generators. Nonetheless, this rephrasing corrects Nielsen and Chuang's Exercise 10.34, which is manifestly false, as is shown by examples (ii), (iv), (v), (x), and (xi) in Eq. (9).

For any subgroup, a set of generators is *independent* if removing any generator changes the subgroup generated (by making it smaller). Independence means that no generator can be written as a product of the others. In the examples of Eq. (9), the generators are independent.

An element g of any subgroup $S = \langle g_1, \ldots, g_l \rangle$ can be written as $g = \pm g_1^{a_1} \cdots g_l^{a_l}$, where $a_j = 0, 1$, since any product of generators can be permuted into the standard order by using the commutation and anticommutation properties of the generators, at the expense of introducing a minus signs at the anticommutations. This gives $r(g) = \sum_j a_j r(g_j)$, showing that the check vector $r(g)$ for any subgroup element can be expanded in terms of the generator check vectors $r(g_j)$.

For any subgroup, if the check vectors corresponding to a set of generators g_j are linearly independent, then the generators are independent. Suppose the generators are not independent.

This means that we can generate one generator, call it g_1, from the others, i.e., $g_1 = g_2^{a_2} \cdots g_l^{a_l}$ with at least one $a_j \neq 0$. This gives $r(g_1) = \sum_{j=2}^{l} a_j r(g_j)$, contradicting the linear independence of the generators. The converse is not true, i.e., independence of a set of generators does not imply linear independence of the corresponding check vectors, as one can see from examples (ii), (v), (x), and (xi) in Eq. (9). For the subgroups that stabilize a nontrivial subspace, however, i.e., those such that $-I \notin S$, the converse does hold. This is the content of Nielsen and Chuang's Proposition 10.3, which we now get to by a somewhat different route.

Let's specialize to subgroups $S = \langle g_1, \ldots, g_l \rangle$ whose generators commute, satisfy $g_j \neq -I$ and $g_j^2 = I$ (notice that none of the generators can have a zero check vector), and have linearly independent check vectors. We can conclude immediately that the generators are independent, giving us the following result:

$$\begin{array}{c}\text{Generators commute,}\\ g_j \neq -I \text{ and } g_j^2 = I \;\forall\; g_j, \text{ and}\\ \text{generator check vectors } r(g_j)\\ \text{are linearly independent}\end{array} \quad\Longrightarrow\quad \text{Generators } g_j \text{ are independent} \tag{51}$$

We can simplify the hypothesis here, since linear independence of the generator check vectors implies that none of the generators is $-I$ (or I for that matter). Thus we are left with

$$\begin{array}{c}\text{Generators commute,}\\ g_j^2 = I \;\forall\; g_j, \text{ and}\\ \text{generator check vectors } r(g_j)\\ \text{are linearly independent}\end{array} \quad\Longrightarrow\quad \text{Generators } g_j \text{ are independent} \tag{52}$$

The converse of this statement is not true, of course, as is demonstrated by examples (ii), (iii), (iv), (v), (x), and (xi) in Eq. (9).

Now let's suppose instead that $-I \in S$, so that we can write $-I = g_1^{a_1} \cdots g_l^{a_l}$ for some set of a_j with at least one $a_j \neq 0$ (for if all $a_j = 0$, then the product gives I, not $-I$). This gives $0 = r(-I) = \sum_j a_j r(g_j)$, implying that the generator check vectors are not linearly independent. We can rewrite this as the contrapositive:

$$\begin{array}{c}\text{Generators commute,}\\ g_j^2 = I \;\forall\; g_j, \text{ and}\\ \text{generator check vectors } r(g_j)\\ \text{are linearly independent}\end{array} \quad\Longrightarrow\quad -I \notin S \tag{53}$$

The converse of this statement is also not true, since we can always use an overcomplete set of generators whose check vectors are thus not linearly independent.

Putting together the necessary conditions in Eqs. (52) and (53) does give sufficient conditions:

$$\begin{array}{c}\text{Generators commute,}\\ g_j^2 = I \;\forall\; g_j, \text{ and}\\ \text{generator check vectors } r(g_j)\\ \text{are linearly independent}\end{array} \quad\Longleftrightarrow\quad \begin{array}{c}-I \notin S\\ \text{and}\\ \text{generators } g_j \text{ are independent}\end{array} \tag{54}$$

Proof in reverse direction: By Eq. (50), $-I \notin S$ gets us the first two implications. Now suppose the check vectors are not linearly independent. This implies that one check vector, say, for g_1, can be expanded in terms of the others, i.e., $r(g_1) = \sum_{j=1}^{l} a_j r(g_j)$. Now consider the subgroup element $h = \prod_{j=2}^{l} g_j^{a_j}$, which satisfies $r(h) = r(g_1)$, implying that $h = \pm g_1$ or $h = \pm i g_1$. All but $h = g_1$ are ruled out by Eq. (4): the latter possibilities are ruled out by $h^2 = I$ for all elements of S, and $h = -g_1$ is ruled out by $-h \notin S$ for all elements of S. Thus we have $g_1 = h$, contradicting our assumption that the generators are independent.

In view of Eq. (50), we can also write

$$-I \notin S = \langle g_1, \ldots, g_l \rangle \quad \Longrightarrow \quad \left(\begin{array}{c} \text{Generators } g_j \text{ are independent} \\ \Longleftrightarrow \\ \text{Generator check-vectors } r(g_j) \text{ are linearly independent} \end{array} \right) \tag{55}$$

This is Nielsen and Chuang's Proposition 10.3, but Eq. (54) is strictly stronger in the forward direction.

We already know that S stabilizes a nontrivial subspace iff $-I \notin S$. These are thus the interesting subgroups. We now understand from Eq. (54) that to generate a subgroup with $-I \notin S$, we should use commuting Hermitian generators $g_1, \ldots, g_l$ whose check vectors are linearly independent, for this is equivalent to having an independent set of generators that generate a subgroup such that $-I \notin S$. To specify an interesting stabilizer, we thus have to give l generators, each of the form $\pm\sigma_{\alpha_1} \otimes \cdots \otimes \sigma_{\alpha_N}$. This requires 1 bit for the $\pm$ and 2 bits for each Pauli matrix, for a total of $2N+1$ bits per generator and $l(2N+1)$ for all the generators and, hence, the entire stabilizer.

Each element of such an S can be written as a product of generators, $g = g_1^{a_1} \cdots g_l^{a_l}$, with $a_j = 0, 1$. All such products are distinct, as we can easily see in the following way: if two different products yield the same element, i.e., $g = g_1^{a_1} \cdots g_l^{a_l} = g_1^{b_1} \cdots g_l^{b_l}$, then $I = g_1^{c_1} \cdots g_l^{c_l}$, where $c_j = a_j - b_j$, with at least one c_j being nonzero; we then have $0 = r(I) = \sum_j c_j r(g_j)$, contradicting the linear independence of the generator check vectors. This means that $|S| = 2^l$, with each element of S specified by an l-bit string $a_1 \ldots a_l$, and thus also establishes that

$$\left(\begin{array}{c} \text{dimension} \\ \text{of } V_S \end{array} \right) = \frac{2^N}{|S|} = 2^{N-l} \ . \tag{56}$$

Notice that these conclusions are consistent with our convention that an empty generator set ($l = 0$) generates the trivial subgroup consisting of the identity operator, which stabilizes the entire Hilbert space. Each independent generator halves the size of the stabilized subspace.

If we let R_S be the l-dimensional subspace spanned by the check vectors $r(g)$ for $g \in S$, these considerations tells us that the check vectors for a set of independent generators are a linearly independent basis for R_S, and the 2^l check vectors for all $g \in S$ exhaust all the vectors in R_S.

We can also rewrite the projector (5) for such an S as the product of the $+1$ projectors $P_{g_j} = \frac{1}{2}(I + g_j)$ associated with the generators:

$$P_S = \frac{1}{|S|} \sum_{g \in S} g = \frac{1}{2^l} \sum_{a_1, \ldots, a_l = 0,1} g_1^{a_1} \cdots g_l^{a_l} = \prod_{j=1}^{l} \frac{1}{2}(I + g_j) = \prod_{j=1}^{l} P_{g_j} \ . \tag{57}$$

This is illustrated in a nontrivial way by examples (vii) and (ix) of Eq. (9).

All of the information about a set of generators, except the $\pm$ in front of the Pauli products, is encoded in the *check matrix*,

$$G = \begin{pmatrix} r(g_1) \\ \vdots \\ r(g_l) \end{pmatrix} = \left(\begin{array}{c|c} r_1(g_1) & r_2(g_1) \\ \vdots & \vdots \\ r_1(g_l) & r_2(g_l) \end{array}\right) = \left(G_1 \mid G_2 \right) , \tag{58}$$

i.e, the $l \times 2N$ matrix whose rows are the check vectors of the generators. The conditions for a suitable set of generators are that the rows of G be linearly independent and that the mod-2 skew product (39) of different rows be zero (commuting generators), i.e.,

$$G\Lambda G^T = \begin{pmatrix} G_1 & G_2 \end{pmatrix} \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} G_1^T \\ G_2^T \end{pmatrix} = \begin{pmatrix} -G_2 & G_1 \end{pmatrix} \begin{pmatrix} G_1^T \\ G_2^T \end{pmatrix} = 0 \bmod 2 . \tag{59}$$

To get the generators, one constructs the corresponding products of Pauli operators and puts ± 1 in front of each one. Notice that for an arbitrary subgroup element $g = g_1^{a_1} \cdots g_l^{a_l}$, we have

$$r(g) = \sum_{j=1}^{l} a_j r(g_j) = \begin{pmatrix} a_1 & \cdots & a_l \end{pmatrix} G . \tag{60}$$

We occasionally find it useful to refer to the subgroup $\hat{S} = \{\pm g, \pm ig \mid g \in S\}$, of order $|\hat{S}| = 2^{l+2}$, which is just S augmented by all the rephasings of its elements (formally, S is isomorphic to the quotient group $\hat{S}/K$). Clearly the check vectors of all the elements of $\hat{S}$ lie in R_S, since rephasings do not affect the check vectors. Moreover, if $g \notin \hat{S}$, then $r(g)$ does not lie in R_S. [Proof: Suppose $r(g) \in R_S$, so that we can write it as $r(g) = \sum_{j=1}^{l} a_j r(g_j)$, giving $g = \pm g_1^{a_1} \ldots g_l^{a_l} \in \hat{S}$ or $g = \pm i g_1^{a_1} \ldots g_l^{a_l} \in \hat{S}$.]

Let's get a little more mileage out of what we can say about stabilizer subgroups and their eigensubspaces. Let $S = \langle g_1, \ldots, g_l \rangle$ be a stabilizer with $-I \notin S$, generated by independent generators $g_1, \ldots, g_l$. The subspace stabilized by these generators is the 2^{N-l}-dimensional subspace V_S. The generators $(-1)^{c_1} g_1, \ldots, (-1)^{c_l} g_l$, obtained by rephasing the original generators, generate a group that stabilizes the subspace $V_S^{(c_1 \ldots c_l)}$ of simultaneous $+1$ eigenstates of these new generators. We let $P_S^{(c_1 \ldots c_l)}$ denote the projector onto $V_S^{(c_1 \ldots c_l)}$. An equivalent way of describing $V_S^{(c_1 \ldots c_l)}$ is that it is the subspace of simultaneous eigenstates of the original generators $g_1, \ldots, g_l$ with eigenvalues $(-1)^{c_1}, \ldots, (-1)^{c_l}$, respectively. The 2^l subspaces $V_S^{(c_1 \cdots c_l)}$ are the simultaneous eigensubspaces of the original generators and, hence, of the whole stabilizer S; these subspaces are orthogonal, each having 2^{N-l} dimensions. Each l-dimensional subspace R_S, consisting of check vectors whose pairwise skew products vanish, corresponds to a way of decomposing the N-qubit Hilbert space into 2^l orthogonal subspaces $V_S^{(c_1 \ldots c_l)}$, each of dimension 2^{N-l}.

5. Stabilizer formalism. Further considerations

The centralizer* of a subgroup S in $\mathcal{P}_N$ is the set of elements in $\mathcal{P}_N$ that commute with all

* The *centralizer* $\mathcal{Z}(H|G)$ of a subgroup H of G is defined as the set of elements of G that commute with all elements of H. The centralizer is a subgroup of G; if H is abelian, H is a (normal) subgroup of the centralizer. We generally omit the group designation in the notation for the centralizer, allowing context to fill it in.

elements of S:

$$\mathcal{Z}(S) = \{z \in \mathcal{P}_N \mid zh = hz \;\forall\; h \in S\} \,. \tag{61}$$

For any S, the centralizer $\mathcal{Z}(S)$ is a normal subgroup of $\mathcal{P}_N$, because for any $g \in \mathcal{P}_N$ and $z \in \mathcal{Z}(S)$, $h \in S$ commutes with z and either commutes or anticommutes with g and g^{-1}, so $gzg^{-1}h = hgzg^{-1}$ and hence $gzg^{-1} \in \mathcal{Z}(S)$.

For a subgroup S with $-I \notin S$, notice that $\mathcal{Z}(\hat{S}) = \mathcal{Z}(S)$ and also that S is a (normal) subgroup of $\mathcal{Z}(S)$. A group element z is in $\mathcal{Z}(S)$ iff it commutes with all the generators of S, so according to Eq. (39), z is in $\mathcal{Z}(S)$ iff it satisfies $r(h)\Lambda r^T(z) \bmod 2 = 0 \;\forall\; h \in S$ or, equivalently,

$$G\Lambda r^T(z) \bmod 2 = 0 \,, \quad \text{where } G\Lambda = \big(\, -G_2 \;\big|\; G_1 \,\big), \tag{62}$$

i.e., $\Lambda r^T(z)$ is in the null subspace of G. The rows of $G\Lambda$ are linearly independent iff the rows of G are linearly independent, so these are l linearly independent conditions on $r(z)$. Thus the check vectors $r(z)$ span a $(2N-l)$-dimensional subspace, containing 2^{2N-l} vectors. Taking into account the four phases for each check vector, the number of elements in the centralizer is

$$|\mathcal{Z}(S)| = 2^{2N+2-l} = 4^{N+1}2^{-l} \,. \tag{63}$$

In particular, this means that (i) half of $\mathcal{P}_N$ commutes with any particular group element $g \neq \pm I, \pm iI$ (the other half anticommutes), (ii) adding an independent generator cuts the size of the centralizer in half, and (iii) when there are N independent generators, the size of the centralizer is 2^{N+2}, which means that $\mathcal{Z}(S) = \hat{S}$ is just S augmented by the rephasings of each element in S.

Now we're in a position to calculate the number of distinct stabilizer subgroups and also the number of distinct independent generators sets for each stabilizer (I learned how to do this from Jim Harrington of Caltech when we were both at the PASI on the Physics of Information in Buzios, Brazil, for the first two weeks of 2003 December). Suppose we have $S_k = \langle g_1, \ldots, g_k \rangle$ generated from independent generators, and we want to know how many choices there are for adding an additional independent generator. The new generator must be chosen from the $g^2 = I$ part of $\mathcal{Z}(S_k)$, which makes available $|\mathcal{Z}(S_k)|/2 = 2^{2N+1-k}$ possibilities, but we have to exclude elements of S_k or their negatives, thus reducing the number of possibilities by $2|S_k| = 2^{k+1}$. Thus the number of possibilities for the $(k+1)$th generator is $|\mathcal{Z}(S_k)|/2 - 2|S_k| = 2^{2N+1-k} - 2^{k+1} = 2^{k+1}\big(2^{2(N-k)} - 1\big)$. This means that the number of distinct (independent) l-element generator sets is

$$\frac{1}{l!}\prod_{k=0}^{l-1} 2^{k+1}\big(2^{2(N-k)} - 1\big) = \frac{2^{l(l+1)/2}}{l!}\prod_{k=0}^{l-1}\big(2^{2(N-k)} - 1\big) = \frac{2^{l(l+1)/2}}{l!}\prod_{k=0}^{l-1}\big(2^{N-k} + 1\big)\big(2^{N-k} - 1\big) \,, \tag{64}$$

where the $l!$ takes into account the fact that the above counting includes all permutations of a given set of generators. Equation (64) is also 2^l times the number of sets of l linearly independent check vectors in R_N whose pairwise skew products vanish (the factor of 2^l takes into account the two possible signs in going from check vectors to generators).

Using the same sort of counting, we can find the number of (independent) l-element generator sets for a particular stabilizer subgroup S. If we've already accumulated k generators, the number of choices for the $(k+1)$th (independent) generator is the number of elements in S, $|S| = 2^l$, minus the number of elements, 2^k, in the group generated by the first k generators. Thus the number of distinct sets of independent generators for S is

$$\frac{1}{l!}\prod_{k=0}^{l-1} 2^k\big(2^{l-k}-1\big) = \frac{2^{l(l-1)/2}}{l!}\prod_{k=0}^{l-1}\big(2^{l-k}-1\big)\;, \tag{65}$$

where again the $l!$ removes the overcounting of different permutations of the same generator set. This same number is, of course, the number of linearly independent basis sets in R_S.

Dividing Eq. (64) by Eq. (65) gives the number Ω_l of distinct stabilizer groups of size 2^l (which is also the 2^l times the number of subspaces R_S of dimension l):

$$\Omega_l = 2^l\prod_{k=0}^{l-1}\frac{2^{2(N-k)}-1}{2^{l-k}-1} = 2^l\prod_{k=0}^{l-1}\frac{\big(2^{N-k}+1\big)\big(2^{N-k}-1\big)}{2^{l-k}-1}\;. \tag{66}$$

For $l=N$, this simplifies to

$$\Omega_N = 2^N\prod_{k=0}^{N-1}\big(2^{N-k}+1\big) = 2^N\prod_{k=1}^{N}\big(2^k+1\big) = 2^{N(N+3)/2}\prod_{k=1}^{N}\big(1+2^{-k}\big)\;, \tag{67}$$

which is also the number of distinct states stabilized by stabilizer groups. The first few values of Ω_N are $\Omega_1 = 6$, $\Omega_2 = 60$, $\Omega_3 = 1{,}080$, $\Omega_4 = 36{,}720$, and $\Omega_5 = 2{,}423{,}520$. Since the product on the far right in Eq. (67) is an increasing function of k, we also have $1.5 = \Omega_1/2^{N(N+3)/2} \le \Omega_N/2^{N(N+3)/2} < \Omega_\infty/2^{N(N+3)/2} = 2.38423$.

Now recall that any element $g\in\mathcal{P}_N$ either commutes or anticommutes with a generator g_j: $g_jg = (-1)^{c_j}gg_j$. From this we get $g_jg|\psi\rangle = (-1)^{c_j}gg_j|\psi\rangle = (-1)^{c_j}g|\psi\rangle$ for $|\psi\rangle\in V_S$, meaning that $g|\psi\rangle \in V_S^{(c_1\ldots c_l)}$. This is a property of a coset of $\mathcal{Z}(S)$, since all elements of a coset commute or anticommute with the same generators in S [$g_j(gz) = (-1)^{c_j}(gz)g_j$ for $z\in\mathcal{Z}(S)$] and thus all of which map V_S to the same subspace $V_S^{(c_1\ldots c_l)}$. If $z\in\mathcal{Z}(S)$, then $z|\psi\rangle\in V_S$, i.e., the centralizer preserves V_S.

For any choice of the binary variables c_j, $j=1,\ldots,l$, we can find an element $g\in\mathcal{P}_N$ and, hence, a coset of $\mathcal{Z}(S)$, such that $g_jg=(-1)^{c_j}gg_j$, $j=1,\ldots,l$. To show this, let $c = \begin{pmatrix} c_1 & \cdots & c_l\end{pmatrix}$ be the l-dimensional row vector given by the binary variables. Since the l rows of the check matrix G are linearly independent, there exist 2^{2N-l} column vectors x^T satisfying $G\Lambda x^T = c^T$. (This is not a subspace, but the 2^{2N-l} solutions of $G\Lambda x^T = c^T$, each with four possible phasings, produce the 2^{2N+2-l} elements of the coset.) Pick one solution, and let g be such that $r(g) = x$, giving $G\Lambda r^T(g) = c^T$. This is equivalent to $r(g_j)\wedge r(g) = c_j$, implying that $gg_j = (-1)^{c_j}g_jg$. In terms of check vectors, the coset property is that the general solution of $G\Lambda x^T = c^T$ can be written as $x = r(g)+r(z)$, where g is a particular member of the appropriate coset of the centralizer [$r(g)$ is a particular solution of $G\Lambda x^T = c^T$], and $z\in\mathcal{Z}(S)$ [$G\Lambda r^T(z) = 0$].

What we have shown is that the 2^l cosets of $\mathcal{Z}(S)$, each containing $|\mathcal{Z}(S)| = 2^{2N+2-l}$ elements, are in 1-1 correspondence with the binary variables $c_1,\ldots,c_l$; the coset elements are those whose check vectors satisfy

$$G\Lambda r^T(g) = c^T = \begin{pmatrix} c_1\\ \vdots\\ c_l\end{pmatrix}\;, \tag{68}$$

which is equivalent to saying that $gg_j = (-1)^{c_j} g_j g$, i.e., that g maps V_S to $V_S^{(c_1 \ldots c_l)}$. We can write this as $gP_S g^\dagger = P_S^{(c_1 \ldots c_l)}$.

Notice that we can find representatives for all the cosets by starting with check vectors for the cosets that have only one 1 in the vector c. Let $r(h_j)$ be a solution of Eq. (68) for the case where c has a 1 at the jth site. Then a solution for an arbitrary c is $r(g) = \sum_j c_j r(h_j)$, which means that a coset representative for the coset c is $g = h_1^{c_1} \cdots h_l^{c_l}$.

The normalizer* of a stabilizer subgroup S in $\mathcal{P}_N$ is the set of elements of $\mathcal{P}_N$ that conjugate all elements of S to elements of S:

$$\mathcal{N}(S|\mathcal{P}_N) = \{ g \in \mathcal{P}_N \mid gSg^{-1} = S \} \; . \tag{69}$$

Generally, the normalizer is different from the centralizer, but in the case of stabilizer subgroups of $\mathcal{P}_N$, they are the same. This is obvious because any $g \notin \mathcal{Z}(S)$ anticommutes with at least one generator g_j, giving $gg_jg^{-1} = -g_j \notin S$, meaning that $g \notin \mathcal{N}(S)$.

Now suppose that we switch from an initial generator set, $g_1, \ldots, g_l$, to a new generator set, $g'_1, \ldots, g'_l$, where

$$g'_k = g_1^{a_{k1}} \cdots g_l^{a_{kl}} = \prod_{j=1}^{l} g_j^{a_{kj}} \; . \tag{70}$$

The requirement that the new generators be independent is that the new check vectors,

$$r(g'_k) = \sum_{j=1}^{l} a_{kj} r(g_j) \; , \tag{71}$$

be linearly independent, which means that the $l \times l$ matrix A, whose matrix elements are $A_{kj} = a_{kj}$, must have linearly independent rows and thus be invertible. We can write the check matrix for the new generators as

$$G' = AG \; . \tag{72}$$

The independent generator sets are thus in 1-1 correspondence with invertible $l \times l$ matrices A, and hence the number of such matrices is given by Eq. (65) multiplied by $l!$ (we remove the $l!$ from the denominator because the matrices include permutation matrices, which simply reorder the generators).

6. Canonical generator sets

We now consider independent generator sets for the entire Pauli group. For this purpose, consider $2N$ Pauli-group elements g_j, $j = 1, \ldots, 2N$, whose check vectors $r(g_j)$ are linearly independent, and let $S = \langle g_1, \ldots, g_{2N} \rangle$ be the subgroup these elements generate. We want to show that S consists of the elements $\pm g_1^{a_1} \cdots g_{2N}^{a_{2N}}$ and that all of these elements are distinct, thus making $|S| = 2^{2N+1}$. First, all the products $g_1^{a_1} \cdots g_{2N}^{a_{2N}}$ are clearly in S, and they are all distinct because the linear independence

* The *normalizer* $\mathcal{N}(H|G)$ of a subgroup H of G is the set of elements of G that conjugate all elements of H to elements of H. The normalizer is a subgroup of G, and H and $\mathcal{Z}(H|G)$ are normal subgroups of the normalizer.

of the generator check vectors guarantees that their check vectors $\sum_{j=1}^{2N} a_j r(g_j)$ are different. Second, since there are anticommuting generators, we have $-I \in S$, which implies that $-g_1^{a_1} \cdots g_{2N}^{a_{2N}} \in S$; these elements with a minus sign in front of the product are all distinct and different from the elements with a plus sign. Finally, any product of the generators can be brought into the standard order at the expense of introducing minus signs at each anticommutation, so there are no other elements in S than those listed. What all this means is that S contains two of the four phasings of each Pauli product and thus has half of the elements of $\mathcal{P}_N$. To get the entire Pauli group, we have to add one generator, which can be taken to be i times any element of S. An example of this procedure is the set of generators discussed at the end of Sec. 2: start with X and Z for each qubit, and then add $iXZ = Y$ for one qubit.

Now suppose the $2N$ generators fall into two sets, $g_1, \ldots, g_N$ and $h_1, \ldots, h_N$, each of which is a set of independent (Hermitian) generators for a stabilizer group. Moreover, suppose that $[g_j, h_k] = 0$ except when $j = k$. In terms of the check vectors, we have that the $2N$ vectors $r(g_j)$ and $r(h_k)$ are linearly independent and $r(g_j) \wedge r(g_k) \bmod 2 = 0$, $r(h_j) \wedge r(h_k) \bmod 2 = 0$, and $r(g_j) \wedge r(h_k) \bmod 2 = \delta_{jk}$. We define the $2N \times 2N$ matrix of check vectors for both generator sets:

$$\mathcal{G} = \begin{pmatrix} G \\ H \end{pmatrix} = \begin{pmatrix} G_1 & G_2 \\ H_1 & H_2 \end{pmatrix} . \tag{73}$$

Throughout we generally use upper-case script letters for matrices of check vectors that do not correspond to generators for a stabilizer subgroup. The conditions on $\mathcal{G}$ are that the rows be linearly independent and that

$$\begin{aligned} \mathcal{G}\Lambda\mathcal{G}^T &= \begin{pmatrix} G_1 & G_2 \\ H_1 & H_2 \end{pmatrix} \Lambda \begin{pmatrix} G_1^T & H_1^T \\ G_2^T & H_2^T \end{pmatrix} \\ &= \begin{pmatrix} -G_2 & G_1 \\ -H_2 & H_1 \end{pmatrix} \begin{pmatrix} G_1^T & H_1^T \\ G_2^T & H_2^T \end{pmatrix} \\ &= \begin{pmatrix} G_1 G_2^T - G_2 G_1^T & G_1 H_2^T - G_2 H_1^T \\ H_1 G_2^T - H_2 G_1^T & H_1 H_2^T - H_2 H_1^T \end{pmatrix} \\ &= \Lambda \bmod 2 . \end{aligned} \tag{74}$$

We call $2N$-element generator sets of this sort *canonical* because the corresponding matrix $\mathcal{G}$ preserves the fundamental simplectic matrix Λ. An example of a canonical generator set is the set of Xs and Zs for all qubits; we call this the *fiducial canonical generator set*.

What we want to do now is to count the number of (ordered) canonical generator sets. In counting the number of stabilizer-generator sets, we have already done the first part of this counting, i.e., the part that assembles the generators $g_1, \ldots, g_N$, but we repeat this counting here in check-vector language. Suppose we have assembled generators $g_1, \ldots, g_k$ and want to add the generator g_{k+1}. The linear independence of $r(g_{k+1})$ says that we cannot use any of the 2^k vectors in the k-dimensional subspace spanned by $r(g_1), \ldots, r(g_k)$, and the commutation condition on g_{k+1} is that $G_k \Lambda r^T(g_{k+1}) \bmod 2 = 0$, which is satisfied by the 2^{2N-k} vectors in a $(2N-k)$-dimensional subspace, which includes all of the vectors in the k-dimensional subspace spanned by $r(g_1), \ldots, r(g_k)$. Thus the number of check vectors available at this point is $2^{2N-k} - 2^k$, which when one takes into account the two possible signs for the corresponding group element, gives $2(2^{2N-k} - 2^k)$ available group elements.

Thus the number of ordered generator sets is

$$
\begin{aligned}
\prod_{k=0}^{N-1} 2(2^{2N-k} - 2^k) &= 2^N \prod_{k=0}^{N-1} 2^k \prod_{k=0}^{N-1} 2^{2(N-k)} \prod_{k=0}^{N-1} (1 - 2^{-2(N-k)}) \\
&= 2^N \prod_{k=0}^{N-1} 2^k \left(\prod_{k=1}^{N} 2^k \right)^2 \prod_{k=1}^{N} (1 - 2^{-2k}) \\
&= 2^{3N} \left(\prod_{k=0}^{N-1} 2^k \right)^3 \prod_{k=1}^{N} (1 - 2^{-2k}) \\
&= 2^{3N(N+1)/2} \prod_{k=1}^{N} (1 - 2^{-2k}) \;.
\end{aligned}
\tag{75}
$$

The difference between this result and $l = N$ version of Eq. (64) is that here we are interested in ordered sets, so there is no factorial in the denominator.

We're ready now to assemble the second set of generators. Suppose we have assembled the generators $h_1, \ldots, h_k$ and want to add the generator h_{k+1}. Requiring that h_{k+1} anticommute with g_{k+1} guarantees that h_{k+1} is not in the group generated by $g_1, \ldots, g_N, h_1, \ldots, h_k$, because all these generators commute with g_{k+1}. Thus the only condition we need to impose is the commutation condition

$$
\begin{pmatrix} G \\ H_k \end{pmatrix} \Lambda r^T(h_{k+1}) = \begin{pmatrix} e_k^T \\ 0_k \end{pmatrix} \;, \tag{76}
$$

where e_k is a N-dimensional vector that has zeroes at all sites except for a 1 at the kth site and 0_k is the k-dimensional zero vector. Equation (76) imposes $N + k$ linearly independent conditions, which are satisfied by $2^{2N-N-k} = 2^{N-k}$ check vectors; this gives 2^{N-k+1} available group elements. Notice that we can characterize h_{k+1} as a Hermitian element of the centralizer of $\langle h_1, \ldots, h_k \rangle$ and of the coset of $\langle g_1, \ldots, g_N \rangle$ corresponding to the vector e_k.

The result is that the number of canonical generator sets is

$$
\begin{aligned}
\Upsilon_N &= 2^{3N(N+1)/2} \prod_{k=1}^{N} (1 - 2^{-2k}) \prod_{k=0}^{N-1} 2^{N-k+1} \\
&= 2^{3N(N+1)/2} \prod_{k=1}^{N} (1 - 2^{-2k}) 2^N \prod_{k=1}^{N} 2^k \\
&= 2^{2N^2+3N} \prod_{k=1}^{N} (1 - 2^{-2k})
\end{aligned}
\tag{77}
$$

It is easy to calculate the first few values of Υ_N: $\Upsilon_1 = 24$, $\Upsilon_2 = 11{,}520$, $\Upsilon_3 = 92{,}897{,}280$, and $\Upsilon_4 = 12{,}128{,}668{,}876{,}800$. Since the final product in Eq. (77) is a decreasing function of k, we also have $0.688538 = \Upsilon_\infty / 2^{2N^2+3N} < \Upsilon_N / 2^{2N^2+3N} \leq \Upsilon_1 / 2^{2N^2+3N} = 0.75$.

7. Stabilizer formalism. Gates

A particularly important role is played by the unitary operators that preserve the Pauli group under unitary conjugation. These unitary operators make up a group, which is the normalizer of the Pauli group in the group of unitaries on N qubits,

$$\mathcal{N}\big(\mathcal{P}_N|\mathrm{U}(2^N)\big) = \{U \in \mathrm{U}(2^N) \mid UgU^\dagger \in \mathcal{P}_N \ \forall\ g \in \mathcal{P}_N\}\ . \tag{78}$$

This normalizer is also called the *Clifford group*. Clifford-group elements take stabilizer groups to stabilizer groups and thus stabilizer states to stabilizer states. A Clifford unitary can be multiplied by a phase without changing how it conjugates operators. We don't care about these phases, so throughout our discussion of the Clifford group, we ignore them, regarding two unitaries that differ by a phase as the same unitary. What we are actually talking about is thus the quotient group with the phases mod'ed out.

The Clifford group is generated by one- and two-qubit unitaries. One generator set is the following: (i) the Hadamard gate,

$$H = ie^{-i[(X+Z)/\sqrt{2}](\pi/2)} = \frac{1}{\sqrt{2}}(X+Z) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\ , \tag{79}$$

and the S gate,

$$S = e^{i\pi/4}e^{-iZ\pi/4} = \frac{e^{i\pi/4}}{\sqrt{2}}(I - iZ) = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}\ , \tag{80}$$

for each qubit; and (ii) the controlled-NOT (or controlled-PHASE) for each pair of qubits. In this section, we will use the notation $C_{j,k}$ for a controlled-NOT where qubit j is the control and qubit k is the target.

To show that these unitaries generate the Clifford group, we use the notion of canonical generator sets introduced at the end of the preceding section. Our strategy is to show that the elements of the Clifford group are in one-to-one correspondence with the canonical generator sets. This establishes that the number of elements of the Clifford is given by Υ_N. As part of this demonstration, it emerges that any element of the Clifford group can be generated from the set of generators above.

Since unitary conjugations preserve commutators, it is clear that any Clifford unitary takes canonical generator sets to canonical generator sets. In particular, any Clifford unitary U transforms the fiducial canonical generator set $\{X_j, Z_j \mid j = 1,\ldots,N\}$ to some other canonical generator set $\{g_j, h_j \mid j = 1,\ldots,N\}$, i.e.,

$$UX_jU^\dagger = g_j \quad \text{and} \quad UZ_jU^\dagger = h_j\ , \quad j = 1,\ldots,N\ . \tag{81}$$

Since the fiducial set generates all Pauli products, these transformation equations specify how U transforms any Pauli product. Since the Pauli products are a complete set of operators, a linear map on operators is specified by how it maps Pauli products. This means that all elements of the Clifford group map the fiducial canonical generator set to different canonical sets, thus establishing that the number of canonical generator sets is an upper bound on the number of Clifford unitaries.

What remains to be shown is that there is a Clifford unitary that gives every transformation of the form (81). To show this, we use a method devised by Bryan Eastin, which shows that every transformation of the the form (81) is generated by a product of the generators listed above. This establishes that the Clifford unitaries are in one-to-one correspondence with the canonical generator sets, while also showing that the generators listed above generate the entire Clifford group.

The first fact we need is that the desired result holds for one qubit, i.e., that the 24 transformations of the form (81) for a single qubit are generated by H and S. These transformations are rotations that preserve the six cardinal directions. Any such transformation is represented by a positive-determinant 3×3 permutation matrix, with arbitrary signs. There are 6 permutation matrices and 8 sets of signs for each, giving the 48 elements of the octahedral group. We want only the half of these that have positive determinant, which are the required 24 transformations. Here we give the Clifford unitary, generated by H and S, for each of these transformations:

U	$UXU^\dagger$	$UYU^\dagger$	$UZU^\dagger$	Description
I	X	Y	Z	Identity
$X = HZH$	X	$-Y$	$-Z$	180° about $\hat{\mathbf{x}}$
Z	$-X$	$-Y$	Z	180° about $\hat{\mathbf{z}}$
$iY = ZX = ZHZH$	$-X$	Y	$-Z$	180° about $\hat{\mathbf{y}}$
SHS	X	$-Z$	Y	90° about $\hat{\mathbf{x}}$
$XSHS = HZHSHS$	X	Z	$-Y$	90° about $-\hat{\mathbf{x}}$
$ZSHS$	$-X$	$-Z$	$-Y$	180° about $(\hat{\mathbf{z}} - \hat{\mathbf{y}})/\sqrt{2}$
$iYSHS = ZHZHSHS$	$-X$	Z	Y	180° about $(\hat{\mathbf{z}} + \hat{\mathbf{y}})/\sqrt{2}$
H	Z	$-Y$	X	180° about $(\hat{\mathbf{z}} + \hat{\mathbf{x}})/\sqrt{2}$
$XH = HZ$	$-Z$	Y	X	90° about $\hat{\mathbf{y}}$
ZH	Z	Y	$-X$	90° about $-\hat{\mathbf{y}}$
$iYH = ZHZ$	$-Z$	$-Y$	$-X$	180° about $(\hat{\mathbf{z}} - \hat{\mathbf{x}})/\sqrt{2}$
SH	Z	X	Y	120° about $(-\hat{\mathbf{z}} - \hat{\mathbf{x}} - \hat{\mathbf{y}})/\sqrt{3}$
$XSH = HZHSH$	$-Z$	X	$-Y$	120° about $(-\hat{\mathbf{z}} + \hat{\mathbf{x}} + \hat{\mathbf{y}})/\sqrt{3}$
ZSH	Z	$-X$	$-Y$	120° about $(\hat{\mathbf{z}} + \hat{\mathbf{x}} - \hat{\mathbf{y}})/\sqrt{3}$
$iYSH = ZHZHSH$	$-Z$	$-X$	Y	120° about $(\hat{\mathbf{z}} - \hat{\mathbf{x}} + \hat{\mathbf{y}})/\sqrt{3}$
S	Y	$-X$	Z	90° about $\hat{\mathbf{z}}$
$XS = HZHS$	$-Y$	$-X$	$-Z$	180° about $(\hat{\mathbf{x}} - \hat{\mathbf{y}})/\sqrt{2}$
ZS	$-Y$	X	Z	90° about $-\hat{\mathbf{z}}$
$iYS = ZHZHS$	Y	X	$-Z$	180° about $(\hat{\mathbf{x}} + \hat{\mathbf{y}})/\sqrt{2}$
HS	$-Y$	$-Z$	X	120° about $(-\hat{\mathbf{z}} - \hat{\mathbf{x}} + \hat{\mathbf{y}})/\sqrt{3}$
$XHS = HZS$	Y	Z	X	120° about $(\hat{\mathbf{z}} + \hat{\mathbf{x}} + \hat{\mathbf{y}})/\sqrt{3}$
ZHS	Y	$-Z$	$-X$	120° about $(\hat{\mathbf{z}} - \hat{\mathbf{x}} - \hat{\mathbf{y}})/\sqrt{3}$
$iYHS = ZHZS$	$-Y$	Z	$-X$	120° about $(-\hat{\mathbf{z}} + \hat{\mathbf{x}} - \hat{\mathbf{y}})/\sqrt{3}$. (82)

This listing shows explicitly that, given a pair of anticommuting Pauli operators, including signs, there is a Clifford unitary (generated by H and S) that transforms the first to X and the second to Z and, given a pair of commuting Pauli operators, there is a Clifford unitary (generated by H and S) that transforms the two so that each becomes an X or an I.

For the second fact, we consider four pairs of two-qubit Pauli products, XI and ZI, XX and ZI, XI and ZX, and XX and ZX, and display explicitly for each pair that there is a Clifford unitary (generated by Hs and controlled-NOTs) that transforms the pair to XI and ZI:

$$
\begin{array}{ccccccccc}
XI & & & & & & & & \\
ZI & & & & & & & & \\
XX & & XI & & & & & & \\
ZI & \xrightarrow[C_{1,2}]{} & ZI & & & & & & \\
XI & & ZI & & ZI & & XI & & \\
ZX & \xrightarrow[H_1]{} & XX & \xrightarrow[C_{1,2}]{} & XI & \xrightarrow[H_1]{} & ZI & & \\
XX & & XI & & ZI & & ZI & & XI \\
ZX & \xrightarrow[C_{1,2}]{} & ZX & \xrightarrow[H_1]{} & XX & \xrightarrow[C_{1,2}]{} & XI & \xrightarrow[H_1]{} & ZI
\end{array}
\;. \tag{83}
$$

For the third fact, we consider three qubits and display explicitly a Clifford unitary (generated by controlled-NOTs) that transforms XXX to XII and ZZZ to ZII:

$$
\begin{array}{c} XXX \\ ZZZ \end{array} \xrightarrow[C_{2,3}]{} \begin{array}{c} XXI \\ ZIZ \end{array} \xrightarrow[C_{3,1}]{} \begin{array}{c} XXI \\ ZII \end{array} \xrightarrow[C_{1,2}]{} \begin{array}{c} XII \\ ZII \end{array} \;. \tag{84}
$$

These three facts are all we need to get to the main result. Consider first two Pauli products (up to sign), g and h, that anticommute (and thus that anticommute at an odd number of sites, which means there is at least one such site). We want to show that there is a Clifford unitary V (generated by Hs, Ss, and controlled-NOTs) such that $VgV^\dagger = X_1$ and $VhV^\dagger = Z_1$. We can associate any minus sign in front of the Pauli product in g or h with one of the anticommuting sites.

The first fact then shows that there is a Clifford unitary that transforms g and h so that at anticommuting sites, the transformed g has an X and the transformed h has a Z and at commuting sites, g and h have Xs or Is. If the first site is not anticommuting, choose an anticommuting site and swap it with the first site (using a sequence of three controlled-NOTs). The second fact shows that there is a Clifford unitary that transforms the first site and each commuting site to the required form. There being an even number of anticommuting sites other than the first, we can partition these anticommuting sites into pairs, and then the third fact shows that there is a Clifford unitary that transforms the first site and each such pair of anticommuting sites to the required form. This demonstrates the existence of the required Clifford unitary V.

Consider now an arbitrary canonical generator set $\{g_j, h_j \mid j = 1, \ldots, N\}$. As we have just shown, there is a Clifford unitary that transforms g_1 to X_1 and h_1 to Z_1, since g_1 and h_1 anticommute. Since unitaries preserve commutators, the transformed versions of all the other generators commute with X_1 and Z_1, which means they must all have I_1 at the first site. We can now forget about the first site and apply the same procedure to the second site in the transformed versions of g_2 and h_2. Proceeding in this way, we end up with a Clifford unitary that transforms an arbitrary canonical generator set to the fiducial canonical generating set. The inverse Clifford unitary transforms in the opposite direction and provides an explicit construction of a Clifford unitary, generated by Hs, Ss, and controlled-NOTs, that demonstrates our main result.

8. Stabilizer formalism. Measurements

Suppose we measure an element g of the Pauli group. We must have $g^2 = I$ so that g is an observable. In this discussion of measurements, we will assume that S is generated by N independent generators so that there is a single stabilized state $|\psi\rangle$ on which we are making measurements. There are two possibilities: (i) g commutes with all the generators, i.e., $g \in \mathcal{Z}(S)$; (ii) g anticommutes with at least one of the generators, i.e., $g \notin \mathcal{Z}(S)$. We consider each possibility in turn.

- If g commutes with all the generators, i.e., $g \in \mathcal{Z}(S)$, $g|\psi\rangle \in V_S$ and thus is equal to $|\psi\rangle$ up to a phase. Since the phase can only be ± 1, we have $g|\psi\rangle = \pm|\psi\rangle$, so that g (upper sign) or $-g$ (lower sign) is an element of S. The result of the measurement is predictable, but we need, of course, a way of determining whether to predict $+1$ or -1. We know that $g = \pm g_1^{a_1} \cdots g_N^{a_N}$ and, hence, that $r(g) = \sum_{j=1}^{N} a_j r(g_j)$. The linear independence of the check vectors implies immediately that the coefficients a_j are unique. Formally, we can write a matrix equation

$$r(g) = aG \,, \tag{85}$$

where G is the generator *check matrix* (58) and

$$a = \begin{pmatrix} a_1 & \cdots & a_N \end{pmatrix} \,. \tag{86}$$

Since the rows of G are linearly independent, we can invert it to give a $N \times 2N$ right matrix inverse G^{-1}. Although this inverse is not unique, the product rG^{-1} is unique for any row vector r that lies in the subspace spanned by the rows of G. Thus a is uniquely determined by

$$a = r(g)G^{-1} \,. \tag{87}$$

Given a, we simply calculate $g_1^{a_1} \cdots g_N^{a_N}$ and determine whether it is equal to g or $-g$. The post-measurement state is $P_{\pm g}|\psi\rangle = \frac{1}{2}(I \pm g)|\psi\rangle = |\psi\rangle$, and thus the measurement does not change the stabilizer generators.

- If g anticommutes with at least one generator, call it g_1, then

$$\langle\psi|g|\psi\rangle = \langle\psi|gg_1|\psi\rangle = -\langle\psi|g_1 g|\psi\rangle = -\langle\psi|g|\psi\rangle \,, \tag{88}$$

which implies that $\langle\psi|g|\psi\rangle = 0$ and thus that the probability for getting either $+1$ or -1 is $1/2$, which can be simulated by a coin flip with a fair coin. The post-measurement state is $P_{\pm g}|\psi\rangle = \frac{1}{2}(1 \pm g)|\psi\rangle$, the upper (lower) sign applying if the result of the measurement is $+1$ (-1). We can get at these results—and also find generators for the post-measurement state—by writing $g_j g = (-1)^{c_j} g g_j$ ($c_1 = 1$ by our assumed ordering of the generators, although other c_j's might also be 1), which implies, according to the discussion above, that $g|\psi\rangle \in V_S^{(c_1 \ldots c_N)}$. Since $V_S^{(c_1 \ldots c_N)}$ is orthogonal to V_S, we get immediately that $\langle\psi|g|\psi\rangle = 0$. It is clear that the $N-1$ commuting generators $g_1^{c_2} g_2, \ldots, g_1^{c_N} g_N$ are independent; moreover, g_1 and g commute with these generators and are independent of them. Thus these generators stabilize the two-dimensional subspace spanned by the states $P_{\pm g_1}|\psi\rangle = \frac{1}{2}(1 \pm g_1)|\psi\rangle$ or by the states $P_{\pm g}|\psi\rangle = \frac{1}{2}(1 \pm g)|\psi\rangle$. Adding g_1 to these generators gives a set of generators that stabilizes the original state $|\psi\rangle$. Adding $\pm g$ to this set gives an independent set of generators that stabilizes the post-measurement state $P_{\pm g}|\psi\rangle = \frac{1}{2}(1 \pm g)|\psi\rangle$, so a set of generators for the post-measurement state is $\pm g, g_1^{c_2} g_2, \ldots, g_1^{c_N} g_N$.

9. Gottesman-Knill theorem

This section needs to be finalized.

Assumptions:

- N qubits initially in a product state in the Z basis.
- Allowed gates: Pauli operators X, Y, and Z, plus H, S and C-NOT.
- Allowed measurements: Products of Pauli operators.

There is an efficient (nonlocal) simulation of the states, dynamics, and measurements.

States: We have to give N stabilizer generators, each of the form $\pm\sigma_{\alpha_1}\otimes\cdots\otimes\sigma_{\alpha_N}$. This requires 1 bit for the $\pm$ and 2 bits for each Pauli matrix, for a total of $2N+1$ bits per generator and $N(2N+1)$ for all the generators and, hence, the entire stabilizer. This is actually an overestimate, since the number of possible states is, according to Eq. (67), $\sim 2^{N(N+3)/2}$, but either way, about $O(N^2)$ bits are required to specified a stabilized state.

Dynamics:

Measurements: which of the possibilities, commuting with all generators or anticommuting with at least one generator, holds can be determined by calculating $G\Lambda r^T(g)$, requiring $O(N^2)$ operations, and determining if the result is zero.

In practice, we could compute the inverse in $O(N^3)$ operations using Gaussian elimination on the columns. Knowing a, we can compute the product $g_1^{a_1}\cdots g_N^{a_N}$ [using $O(N^2)$ operations] and determine whether the product is equal to $+g$ or to $-g$, thus settling which is the predictable result of the measurement.

Since the check-vector formalism typically requires us to deal with nonsquare matrices, it is useful here to review properties of such matrices. We want to allow for vector spaces over finite fields, so in the following we consider matrices over arbitrary fields.

Let

$$A=\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}=\begin{pmatrix}a_{11}&\cdots&a_{1m}\\ \vdots&\ddots&\vdots\\ a_{n1}&\cdots&a_{nm}\end{pmatrix}\tag{89}$$

be an $n\times m$ matrix ($n\le m$), with the vectors a_j, $j=1,\dots,n$, being m-dimensional row vectors. We suppose throughout the following that the rows of A, i.e., the vectors a_j, are linearly independent. Let V be the n-dimensional subspace spanned by the vectors a_j, and let V_k be the $(n-1)$-dimensional subspace spanned by the vectors a_j with a_k left out. The linear independence of the vectors is equivalent to saying that a_k does not lie in V_k for $k=1,\dots,n$. There are two kinds of equations involving A that we want to solve:

$$Ax^T=y^T\qquad\text{and}\qquad yA=x\;,\tag{90}$$

where x is an m-dimensional row vector and y is an n-dimensional row vector.

The first of these equations can be rewritten as

$$\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}=\sum_j y_je_j^T=y^T=Ax^T=\begin{pmatrix}a_{11}x_1+\cdots+a_{1m}x_m\\ \vdots\\ a_{n1}x_1+\cdots+a_{nm}x_m\end{pmatrix}=\begin{pmatrix}a_1\bullet x\\ \vdots\\ a_n\bullet x\end{pmatrix}\;,\tag{91}$$

where e_j is the n-dimensional row vector that has zeroes in each position except the jth, where there is a 1. In this form A defines a linear map from an m-dimensional space to an n-dimensional space. Using Gaussian elimination on the linearly independent rows of A, it is easy to show that Eq. (91) always has a solution.

The kernel (or null subspace) of A is the subspace of vectors v satisfying

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = 0 = Av^T = \begin{pmatrix} a_1 \bullet v \\ \vdots \\ a_n \bullet v \end{pmatrix} . \tag{92}$$

The kernel is thus the subspace $V_\perp$ of vectors orthogonal to V. It is not true that $V \cap V_\perp = 0$, because for vector spaces over finite fields, there are generally vectors that are self-orthogonal. Two vectors x and x' map to the same vector under A, i.e., $Ax^T = Ax'^T$, iff they differ by an element of $V_\perp$.

Consider now a special case of Eq. (91),

$$e_k^T = Ab_k^T = \begin{pmatrix} a_1 \bullet b_k \\ \vdots \\ a_n \bullet b_k \end{pmatrix} \qquad \Longleftrightarrow \qquad a_j \bullet b_k = \delta_{jk} . \tag{93}$$

The m-dimensional row vectors b_k, $k = 1, \ldots, n$, are determined only up to addition of an element of $V_\perp$. If we pick a particular solution b_k for each k, these solutions are linearly independent $\big(0 = \sum_j r_j b_j \Longrightarrow 0 = \sum_j r_j A b_j^T = \sum_j r_j e_j^T \Longrightarrow r_j = 0,\ j = 1, \ldots, n\big)$ and thus span an n-dimensional subspace W. Since nontrivial elements of W do not map to zero, it is clear that $W \cap V_\perp = 0$.

Now consider any vector x. We can write

$$Ax^T = \sum_j y_j e_j^T = \sum_j y_j A b_j^T = A \sum_j y_j b_j^T , \tag{94}$$

which implies that $x = \sum_j y_j b_j + v$, where v is a unique element of $V_\perp$. This means that the entire vector space is the direct sum of W and $V_\perp$, and this further implies that $V_\perp$ is m-dimensional. Finally, we now see that the general solution of Eq. (91) can be written as

$$x = \sum_j y_j b_j + v , \tag{95}$$

where v is an arbitrary element of $V_\perp$.

The second of the equations in Eq. (90) takes the form

$$x = yA = \sum_{j=1}^{n} y_j a_j , \tag{96}$$

which has a solution for y only if $x \in V$, in which case the solution is unique because of the linear independence of the vectors a_j.

We can write the solutions explicitly in terms of matrix inverses. Let

$$B = \begin{pmatrix} b_1^T & \cdots & b_n^T \end{pmatrix} \tag{97}$$

be an $m \times n$ matrix whose columns are m-dimensional row vectors that satisfy Eq. (93). It is easy to see that

$$(AB)_{jk} = a_j \bullet b_k = \delta_{jk} \qquad \Longleftrightarrow \qquad AB = I_n \; ; \tag{98}$$

i.e., B is a right inverse of A. The columns of B are only determined up to addition of an element of $V_\perp$. Nonetheless, because $x \in V$, the equation $x = yA$ has a unique solution

$$\begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix} = y = xB = \begin{pmatrix} x \bullet b_1 & \cdots & x \bullet b_n \end{pmatrix} \; . \tag{99}$$

Now let the vectors w_j, $j = n+1, \ldots, m$, be added to the vectors a_j, $j = 1, \ldots, n$, to make a basis for the entire vector space. The vectors w_j span a subspace $W_\perp$; clearly the direct sum of V and $W_\perp$ is the entire vector space. The matrix

$$\tilde{A} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \\ w_{n+1} \\ \vdots \\ w_m \end{pmatrix} \tag{100}$$

obviously has linearly independent rows and thus is invertible. The inverse,

$$\tilde{A}^{-1} = \begin{pmatrix} b_1^T & \cdots & b_n^T & v_{n+1}^T & \cdots & v_m^T \end{pmatrix} \; , \tag{101}$$

satisfies

$$\tilde{A}\tilde{A}^{-1} = I_m \qquad \Longleftrightarrow \qquad \begin{aligned} a_j \bullet b_k &= \delta_{jk} \\ a_j \bullet v_k &= 0 \\ w_j \bullet b_k &= 0 \\ w_j \bullet v_k &= \delta_{jk} \end{aligned} \; . \tag{102}$$

Each choice of the vectors w_j leads to a unique choice for the vectors b_k. The subspace $W_\perp$ is the subspace of vectors orthogonal to W. The freedom in choosing the vectors w_j is that we can add to each w_j an arbitrary element of V. This freedom corresponds to the freedom to add elements of $V_\perp$ to the vectors b_k.

The equation

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \tilde{A}x^T = \begin{pmatrix} a_1 \bullet x \\ \vdots \\ a_n \bullet x \\ w_{n+1} \bullet x \\ \vdots \\ w_m \bullet x \end{pmatrix} \tag{103}$$

has the unique solution

$$x^T = \tilde{A}^{-1} \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \sum_{j=1}^{n} y_j b_j^T + \sum_{j=n+1}^{m} y_j v_j^T \ . \tag{104}$$

The second sum on the right is the arbitrary element of $V_\perp$ that appears in the solution (95).