Ann. Inst. H. Poincaré D
Comb. Phys. Interact. (Online first)
DOI 10.4171/AIHPD/160

Quantum error-correcting codes and their geometries

Simeon Ball, Aina Centelles, and Felix Huber

Abstract. This is an expository article aiming to introduce the reader to the underlying mathematics and geometry of quantum error correction. Information stored on quantum particles is subject to noise and interference from the environment. Quantum error-correcting codes allow the negation of these effects in order to successfully restore the original quantum information. We briefly describe the necessary quantum-mechanical background to be able to understand how quantum error correction works. We go on to construct quantum codes: firstly qubit stabilizer codes, then qubit non-stabilizer codes, and finally codes with a higher local dimension. We will delve into the geometry of these codes. This allows one to deduce the parameters of the code efficiently, deduce the inequivalence between codes that have the same parameters, and presents a useful tool in deducing the feasibility of certain parameters. We also include sections on quantum maximum distance separable codes and the quantum MacWilliams identities.

Contents

We have used various sources in the preparation of this article, principally Gottesman [8, 9], Glynn et al. [7] and Ketkar et al. [14]. The most original parts of these notes are Sections 4 and 6. Section 5 is based on Ketkar et al. [14] but massaged so that appears as a straightforward generalisation of the qubit case of Section 2. Although the main results of Section 3 are from Glynn et al. [7], in a deviation from their approach, we have chosen to prove these results without using the $\mathbb{F}_4$ trick, which we do not

2020 Mathematics Subject Classification. Primary 81P73; Secondary 94B60.
Keywords. Quantum error-correcting codes.

consider until later in Section 5.5. The interested reader is referred to the books by Sakurai [19] and Nielsen and Chuang [16] for standard treatments of quantum mechanics and quantum information theory, to the book by Haroche and Raimond [11] for a thorough treatment of current experiments in quantum mechanics, and to the book by Aaronson [1] for further connections to mathematics, computer science, physics, and philosophy. For those uninitiated in quantum mechanics or quantum computing, we strongly recommend the delightful mnemonic essay on quantum computing by Matuschak and Nielsen [15].

1. Quantum codes

1.1. Introduction

A *qubit* is a two-state or two-level quantum-mechanical system. For example, the intrinsic angular momentum (*spin*) of an electron is such a system. It can only take two values when measured in arbitrary spatial direction, say by measuring the electrons deflection when passing by an inhomogeneous magnetic field. The two corresponding spin-states are commonly referred to as "spin-up" and "spin-down" states with respect to that direction. Another example is the polarization of light. Here the two states can be taken to be vertically and horizontally polarized light; another choice is light that is left circularly and right circularly polarized. In general, a continuum of different photon polarizations is possible. Yet only two distinct states are observed when, e.g., putting beamsplitters or polarization filters in the path of a light beam.

This raises the question: why are only ever two discrete values corresponding to two discrete states observed, if electrons and photons can take on a continuum of possible spin-directions or polarizations? The answer lies with what measurements on quantum systems reveal. It turns out that for a two-state quantum-mechanical system, any individual measurements can only ever reveal the answer to a binary question. In other words, the measurement indicates in which of two mutually exclusive states the qubit can be found after the measurement. Thus while qubits can take on a continuity of states and a continuity of measurements can be performed, only two-valued results can ever be obtained. Thus the notion of a qubit as a *quantum bit*. We will not dwell on the strangeness of quantum mechanics further, the interested reader is referred to discussions of the Stern–Gerlach and double-slit experiments such as found in the books by Sakurai [19] and Haroche and Raimond [11].[1]

[1]For a visualisation of these experiments, see http://toutestquantique.fr/en/spin/ and http://toutestquantique.fr/en/duality/.

In mathematical terms, a qubit is represented by a unit vector in $\mathbb{C}^2$. The spin-up and spin-down (or any other choice of a pair of physically completely distinguishable states) are represented by an orthonormal basis $|0\rangle$ and $|1\rangle$. The notation $|0\rangle$ is a shorthand for the vector $[\begin{smallmatrix}1\\0\end{smallmatrix}]$, and $|1\rangle$ stands for $[\begin{smallmatrix}0\\1\end{smallmatrix}]$. The two *kets* $|0\rangle$ and $|1\rangle$ are also known as the *computational basis* vectors.

Consider now the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}.$$

While $|\psi\rangle \in \mathbb{C}^2$ represents a physically unique state, it is, upon measurement in the spin-up–spin-down direction, found in either of these two directions with equal probability. Sometimes this situation is referred to as the system being "in two states simultaneously". A more accurate description is that the system is "in *superposition* of spin-up and spin-down", or in other words, the system is correctly described as a *linear combination* of spin-up and spin-down.

A typical qubit reads

$$|\alpha\rangle = \alpha_0\,|0\rangle + \alpha_1\,|1\rangle.$$

As usual, $\overline{z}$ is the complex conjugate of the complex number z. When measured, the qubit is with probability $\overline{\alpha_0}\alpha_0$ found in state $|0\rangle$ (spin-up) and with probability $\overline{\alpha_1}\alpha_1$ found in state $|1\rangle$ (spin-down). Since the sum of these two probabilities must be one, we have that for a qubit

$$\overline{\alpha_0}\alpha_0 + \overline{\alpha_1}\alpha_1 = 1. \tag{1.1}$$

The "ket" notation $|\alpha\rangle$ is used for a column vector, whilst the "bra" notation $\langle\alpha|$ is used for a row vector whose coordinates are the complex conjugates of the coordinates of $|\alpha\rangle$. Thus, the "bra" $\langle\alpha|$ is a linear form. The *inner product* or "bra-ket" on $\mathbb{C}^2$ is defined as

$$\langle\alpha|\beta\rangle = \overline{\alpha_0}\beta_0 + \overline{\alpha_1}\beta_1.$$

The normalisation condition in equation (1.1) then reads as $\langle\alpha|\alpha\rangle = 1$, and qubits are represented by complex vectors in $\mathbb{C}^2$ of unit length.

A *unitary transformation* of $\mathbb{C}^2$ is given by a non-singular 2×2 matrix U which preserves this inner product, so

$$\langle U\alpha|U\beta\rangle = \langle\alpha|\beta\rangle,$$

for all $\langle\alpha|$ and $|\beta\rangle$. The set of such unitaries forms the special unitary group SU(2).

In particular,

$$\langle U\alpha|U\alpha\rangle = \langle\alpha|\alpha\rangle = 1.$$

The matrix

$$U = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

is an example of a unitary transformation since

$$\begin{aligned} \langle U\alpha | U\beta \rangle &= (\overline{-i\alpha_1}\,\langle 0| + \overline{i\alpha_0}\,\langle 1|)(-i\beta_1\,|0\rangle + i\beta_0\,|1\rangle) \\ &= \overline{i\alpha_0}(i\beta_0) + \overline{-i\alpha_1}(-i\beta_1) = \langle \alpha | \beta \rangle. \end{aligned}$$

Note that $\{|0\rangle, |1\rangle\}$ is an orthonormal basis, so

$$\langle 0|0\rangle = \langle 1|1\rangle = 1 \quad \text{and} \quad \langle 0|1\rangle = \langle 1|0\rangle = 0.$$

The *Hermitian conjugate* $M^\dagger$ of the linear operator M is the operator which satisfies

$$\langle M\psi\,|\,\phi\rangle = \langle \psi\,|\,M^\dagger\phi\rangle.$$

An operator M is *Hermitian* if $M = M^\dagger$. In matrix terms, this is equivalent to the conjugate transpose being the same as the matrix itself. For example,

$$\begin{pmatrix} 1 & 2+i \\ 2-i & 2 \end{pmatrix}$$

defines a Hermitian operator on $\mathbb{C}^2$.

Let M be a linear operator defined on a complex space with orthonormal basis B. The *trace* of M is defined as

$$\operatorname{tr}(M) = \sum_{|\psi\rangle \in B} \langle \psi\,|\,M\,|\psi\rangle.$$

We can easily prove that the trace of an operator does not depend on the basis chosen. Firstly, note that

$$\operatorname{tr}(MN) = \sum_{|\psi\rangle \in B} \langle \psi\,|\,MN\,|\psi\rangle = \sum_{|\psi\rangle, |\phi\rangle \in B} \langle \psi\,|\,M\,|\phi\rangle\,\langle \phi\,|\,N\,|\psi\rangle$$

and

$$\sum_{|\psi\rangle, |\phi\rangle \in B} \langle \phi\,|\,N\,|\psi\rangle\,\langle \psi\,|\,M|\phi\rangle = \sum_{|\phi\rangle \in B} \langle \phi\,|\,NM\,|\phi\rangle = \operatorname{tr}(NM),$$

hence

$$\operatorname{tr}(PMP^{-1}) = \operatorname{tr}(P^{-1}PM) = \operatorname{tr}(M).$$

In matrix terms, the trace is equal to the sum of the elements on the principal diagonal.

The *Pauli matrices*,

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},$$

are unitary linear transformations of $\mathbb{C}^2$ which form a basis for the space of 2×2 matrices. In general, any error (also those which are not unitary) affecting a single qubit can be written as a linear combination of the Pauli matrices. We sometimes denote $\sigma_0, \sigma_x, \sigma_y, \sigma_z$ simply by I, X, Y, Z, respectively. Note that the Pauli matrices are both unitary and Hermitian. They are also mutually orthogonal under the *Hilbert–Schmidt inner product*

$$\langle A, B \rangle = \operatorname{tr}(A^\dagger B).$$

A *measurement* or *observable* is represented by a Hermitian operator. For example, the spin-up – spin-down measurement $\widehat{\sigma}_z$ is represented by the Pauli matrix σ_z.[2]

The outcome of an individual measurement can only take two values. These correspond to the eigenvalues of σ_z which are $+1$ and -1. After the measurement, the state is then found in the corresponding eigenstate: in $|0\rangle$ if the outcome $+1$ was obtained, and in $|1\rangle$ if the outcome -1 was obtained. These occur with probabilities

$$p_0 = |\langle \alpha | 0 \rangle|^2 \quad \text{and} \quad p_1 = |\langle \alpha | 1 \rangle|^2,$$

respectively.

An *expectation value* is obtained by the repeated measurement of identically prepared spin particles. Measuring the spin value of $\widehat{\sigma}_z$ on a qubit

$$|\alpha\rangle = \alpha_o |0\rangle + \alpha_1 |1\rangle$$

yields the expectation value

$$\langle \widehat{\sigma}_z \rangle = \langle \alpha | \sigma_z | \alpha \rangle = \operatorname{tr}(\sigma_z |\alpha\rangle\langle\alpha|) = \alpha_0^2 - \alpha_1^2.$$

One can check that this leads to the correct expectation value of

$$\langle \widehat{\sigma}_z \rangle = p_0 \cdot (+1) + p_1 \cdot (-1) = \alpha_0^2 - \alpha_1^2 = \langle \alpha | \sigma_z | \alpha \rangle.$$

The above treatment can be generalised. Denote by $\widehat{A}$ an observable which is represented by a Hermitian matrix A. Let m_i and $|m_i\rangle$ be its eigenvalues and corresponding eigenvectors. Measuring an observable $\widehat{A}$ on a quantum state $|\alpha\rangle$ yields the values m_i with probability $p_i = |\langle \alpha | m_i \rangle|^2$. The state is found in the corresponding eigenstates afterwards.

[2] This direction is commonly referred to as the "z-direction" in the x-y-z axis scheme.

This leads to the expectation value

$$\langle \widehat{A} \rangle = \langle \alpha | \, A \, | \alpha \rangle = \mathrm{tr}(A \, | \alpha \rangle \langle \alpha |).$$

The description of multiple quantum systems takes place in the tensor product space of the individual Hilbert spaces. Thus a system of n qubits is described in the n-fold tensor product space of the one-qubit spaces. One arrives at the 2^n-dimensional Hilbert space $(\mathbb{C}^2)^{\otimes n} = \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$ (n times).

A *density matrix* is used to describe a classical probability distribution (also called a statistical mixture or statistical ensemble) over quantum states. Suppose that some source emits the quantum state $|\phi_i\rangle$ with probability p_i. One requires that $p_i \geq 0$ and $\sum_i p_i = 1$. From the discussion in the previous section, it is clear that the measurement of an observable $\widehat{A}$ must yield an expectation value of

$$\langle \widehat{A} \rangle = \sum_i p_i \, \langle \phi_i | \, A \, | \phi_i \rangle.$$

By linearity, this can be rewritten as

$$\langle \widehat{A} \rangle = \mathrm{tr} \left(A \sum_i p_i \, | \phi_i \rangle \langle \phi_i | \right).$$

Indeed, the operator

$$\rho = \sum_{i=1}^{r} p_i \, | \phi_i \rangle \langle \phi_i |$$

captures all there is to know about a quantum system, and ρ is known as the density matrix describing it.

For a complex matrix ρ to represent a quantum state, it is required that $\rho = \rho^\dagger$, $\langle \psi | \, \rho \, | \psi \rangle \geq 0$ for all $|\psi\rangle$ (positive-semidefinite) and $\mathrm{tr}(\rho) = 1$. Compared with classical probability theory, this corresponds to a real-valued, non-negative, and normalized probability distribution. The density matrix formalism can indeed be seen as a generalization of classical probability theory, and quantum mechanics can be taken to be the study of the cone formed by complex positive-semidefinite matrices, and transformations thereof. This is an analogy to the probability simplex encountered in classical probability theory.

Now we can state what we left out in the preceding discussion about measurements: consider the case when some eigenvalues of the measurement operator $A = \sum m_i \, | m_i \rangle \langle m_i |$ are equal, i.e., the spectrum of A is degenerate. What is the probability for obtaining outcome i and what is the post-measurement state? Let P_j be the projector onto the eigenspace with eigenvalue m_j of A. Then a measurement yields

outcome m_j with probability $p_j = \mathrm{tr}(P_j \rho)$ and the density operator immediately after the measurement reads

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

The *time evolution* of an isolated qubit is given by a unitary operator in SU(2):

$$|\alpha\rangle \mapsto U(t)\,|\alpha\rangle.$$

On a closed quantum system of n qubits, the time evolution is given by unitary operators on $\mathcal{H}_{\text{system}} = (\mathbb{C}^2)^{\otimes n}$. In the case of a quantum system interacting with its environment, such unitaries can also act on a larger system

$$\mathcal{H}_{\text{system}} \otimes \mathcal{H}_{\text{environment}}.$$

A unitary on such a larger system can be represented on $\mathcal{H}_{\text{system}}$ in the (non-unique) operator-sum or Kraus decomposition as

$$|\alpha\rangle \mapsto \sum_i K_i\,|\alpha\rangle\langle\alpha|\,K_i^\dagger \quad \text{with the constraint } \sum_i K_i^\dagger K_i = \mathbb{1}.$$

Throughout the paper, $\mathbb{1}$ will denote the identity map. The operators K_i are also known as *Kraus operators*.

More generally, this reads for a density matrix as

$$\rho \mapsto \sum_i K_i \rho K_i^\dagger \quad \text{with the constraint } \sum_i K_i^\dagger K_i = \mathbb{1}.$$

The above map is also known as a *quantum channel* or *completely positive map* and represents the most general form of physical change a quantum state can undergo. In the case of a classical (conventional) bit, an error is represented by the bit-flip $0 \leftrightarrows 1$. For qubits, we regard any non-identity unitary transformation or non-identity quantum channel as an *error*. We can decompose any unitary or quantum channel in terms of a matrix basis.

A good choice is the *Pauli group*: it is generated by all possible tensor products of the 4 Pauli matrices together with phases ± 1 or $\pm i$. Observe that σ_x, σ_z and σ_y anti-commute. That is,

$$\sigma_x\sigma_y = -\sigma_y\sigma_x, \quad \sigma_x\sigma_z = -\sigma_z\sigma_x, \quad \sigma_y\sigma_z = -\sigma_z\sigma_y$$

and then

$$\sigma_x\sigma_y = i\sigma_z, \quad \sigma_y\sigma_z = i\sigma_x, \quad \sigma_z\sigma_x = i\sigma_y.$$

Thus, the Pauli group $\mathcal{P}_n$ is a non-abelian group consisting of the 4^n tensor products of σ_0, σ_x, σ_z and σ_y, which together with the four phases is a group of size 4^{n+1}.

A *quantum error-correcting code* is a linear subspace Q of $(\mathbb{C}^2)^{\otimes n}$ into which a number of logical qubits can be encoded such that all errors of a certain type can be detected and/or corrected. The question we ask is thus: given a noisy channel $\mathcal{E}$, does there exist a recovery channel $\mathcal{R}$ such that every density matrix ρ, for which the support of ρ is contained in Q, can be recovered? In other words, for all density matrices ρ with spectral decomposition

$$\rho = \sum_i p_i \, |\phi_i\rangle \langle\phi_i|,$$

where $|\phi_i\rangle \in Q$, we require that

$$\mathcal{R} \circ \mathcal{E}(\rho) = \rho.$$

1.2. A 1-qubit error-correcting quantum code

A classical *code* is a subset of A^n, where A is a finite set, called the *alphabet*, and n is the *length* of the code. The repetition code is the simplest type of code in which each element $a \in A$ is encoded as $(a, a, \ldots, a)$, an n-tuple of a's. For example, the binary repetition code of length 3 is $\{(000), (111)\}$ and we encode

$$0 \mapsto 000 \quad \text{and} \quad 1 \mapsto 111.$$

This encoding allows us to correct up to one error by taking a majority decision. In other words, we decode the codewords

$$000, 001, 010, 100 \text{ as } 0 \quad \text{and} \quad 111, 011, 110, 101 \text{ as } 1.$$

Can we apply the same strategy to obtain a quantum code? Not quite. A quantum repetition code (on three qubits for example) does not exist, since we cannot map

$$|\alpha\rangle \mapsto |\alpha\rangle \otimes |\alpha\rangle \otimes |\alpha\rangle.$$

It would contradict the following (no-cloning) theorem.

Theorem 1.1 (No-cloning). *There is no linear map which takes* $|\alpha\rangle$ *to* $|\alpha\rangle \otimes |\alpha\rangle$ *for all* $|\alpha\rangle \in (\mathbb{C}^2)^{\otimes n}$.

Proof. Suppose there is such a map. Then

$$|\alpha\rangle \mapsto |\alpha\rangle \otimes |\alpha\rangle, \quad |\beta\rangle \mapsto |\beta\rangle \otimes |\beta\rangle.$$

Such a map however is not linear, as

$$|\alpha\rangle + |\beta\rangle \mapsto (|\alpha\rangle + |\beta\rangle) \otimes (|\alpha\rangle + |\beta\rangle) \neq |\alpha\rangle \otimes |\alpha\rangle + |\beta\rangle \otimes |\beta\rangle. \qquad \blacksquare$$

However, we could try the following repetition-type code:

$$\alpha_0 |0\rangle + \alpha_1 |1\rangle \mapsto \alpha_0 |000\rangle + \alpha_1 |111\rangle.$$

Above and from now on, we simplify the notation $|0\rangle \otimes |0\rangle$ as $|00\rangle$, etc.

Suppose now a "bit-flip" σ_x happens on the second position. This gives

$$\sigma_0 \otimes \sigma_x \otimes \sigma_0(\alpha_0 |000\rangle + \alpha_1 |111\rangle) = \alpha_0 |010\rangle + \alpha_1 |101\rangle.$$

One can correct such an error by majority decision,

$$\alpha_0 |010\rangle + \alpha_1 |101\rangle \quad \text{decodes as } \alpha_0 |000\rangle + \alpha_1 |111\rangle.$$

One needs a measurement that indicates exactly where the bit-flip has occurred. This can be done, as will be explained in Example 2.8.

However, we cannot correct a single σ_z error, since

$$\alpha_0 |000\rangle - \alpha_1 |111\rangle$$

is also a possible state of our code.

Shor [20] was the first to introduce a quantum code which can correct any single-qubit error. He circumvented this apparent problem by introducing a majority decision on the signs to correct a σ_z error.

Example 1.2 (Shor code). The coding space for the Shor code is $(\mathbb{C}^2)^{\otimes 9}$, and a qubit is encoded as

$$|\alpha\rangle \mapsto |\alpha_L\rangle$$

according to

$$|0_L\rangle = (|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle)$$

and

$$|1_L\rangle = (|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle).$$

Hence, by linearity,

$$\begin{aligned}\alpha_0 |0\rangle + \alpha_1 |1\rangle \mapsto\ & \alpha_0(|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle) \\ & + \alpha_1(|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle).\end{aligned}$$

Suppose that we have a σ_x error (bit-flip) occurring on the 4-th bit. Then the α_0 term would change to

$$(|000\rangle + |111\rangle) \otimes (|100\rangle + |011\rangle) \otimes (|000\rangle + |111\rangle),$$

which we would detect and correct by taking the majority decision as with the classical error-correcting code, so we decode

$$|100\rangle + |011\rangle \quad \text{as } |000\rangle + |111\rangle.$$

Now suppose we have a σ_z error (phase error) occurring on the 7-th bit. Then the α_0 term would be

$$(|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle) \otimes (|000\rangle - |111\rangle),$$

which we would detect and correct by taking the majority decision on the signs.

Since $\sigma_y = i\sigma_x\sigma_z$, we can also correct σ_y errors, since the two decisions we made above are independent of each other. Note that the scalar i does not play a role in the decoding.

1.3. The orthogonal projection onto a subspace

Let Q be a subspace of $(\mathbb{C}^2)^{\otimes n}$, and let $Q^\perp$ be its orthogonal subspace with respect to the standard inner product defined on $(\mathbb{C}^2)^{\otimes n} \cong \mathbb{C}^{2^n}$. Any vector $|\psi\rangle$ can be written (uniquely) as the sum of vectors $P|\psi\rangle \in Q$ and $P^\perp|\psi\rangle \in Q^\perp$. The map

$$|\psi\rangle \to P|\psi\rangle$$

is a linear map, called the *orthogonal projection* onto Q.

Lemma 1.3. *If $\{|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_k\rangle\}$ is an orthonormal basis for Q, then*

$$P = \sum_{i=1}^{k} |\psi_i\rangle\langle\psi_i|.$$

Proof. For any $j \leqslant k$,

$$P|\psi_j\rangle = \sum_{i=1}^{k} |\psi_i\rangle\langle\psi_i|\psi_j\rangle = |\psi_j\rangle,$$

so $P|\psi\rangle = |\psi\rangle$ for all $|\psi\rangle \in Q$.

Furthermore,

$$P|\psi\rangle = \sum_{i=1}^{k} |\psi_i\rangle\langle\psi_i|\psi\rangle = 0$$

for all $|\psi\rangle \in Q^\perp$. ∎

Clearly, by definition, $P^2 = P$. By Lemma 1.3, P is Hermitian since it is the sum of Hermitian operators. The following lemma implies that this is enough to characterise P.

Lemma 1.4. *If P is a linear Hermitian operator for which $P^2 = P$ and whose image is Q, then P is the orthogonal projection onto Q.*

Proof. The operator P is Hermitian, so it is diagonalisable with real eigenvalues. Since $P^2 = P$, its eigenvalues are 0 and 1. By the spectral decomposition theorem,

$$P = \sum_{i=1}^{k} |\psi_i\rangle\,|\psi_i\rangle,$$

where $\{|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_k\rangle\}$ is an orthonormal basis for its eigenspace with eigenvalue 1. Since

$$P\,|\psi_j\rangle = |\psi_j\rangle$$

for all $j = 1, \ldots, k$, the eigenspace with eigenvalue 1 contains $\mathrm{im}(P)$, the image of P. The eigenspace with eigenvalue 0 is $\mathrm{im}(P)^{\perp}$. Thus, P is the orthogonal projection onto $\mathrm{im}(P)$. ∎

1.4. Error-detection and correction

For the reliable transmission of an (unknown) quantum system over a noisy channel, we are now faced with three major challenges:

(1) Measurement disturbance. As explained in Section 1.1, measurements induce an "update" of the state that is measured. Thus, when obtaining error syndromes in order to understand what error has occurred, the underlying quantum state may be altered.

(2) Continuous set of errors. The set of errors is continuous and not discrete. How can we distinguish and correct for an error set this large?

(3) No-cloning. Unknown quantum states cannot be copied. Thus an approach of adding redundancy, as done for a classical repetition code, is bound to fail.

How can these challenges be overcome? Firstly, the syndrome measurements are chosen such that they stabilise the set of quantum states that consist of the code. In this way, all code states remain unchanged when extracting the syndromes, while erroneous states are changed in reversible fashion. Secondly, the linearity of quantum mechanics implies that when some discrete set of errors can be corrected, then one can correct all errors which lie in their span. We shall not show a proof of this here, but one can be found in [9, Theorem 2] and [5]. Lastly, the encoded quantum information is distributed amongst many systems and thus "hidden" from any noisy channel. In this way, the state does not have to be copied and no redundancy is added. This not only gives rise to the below Knill–Laflamme conditions on error correction, but also provides an information-theoretic interpretation of quantum error correction.

In quantum error correction, one is faced with the following task. Let

$$\mathcal{N}(\cdot) = \sum_{\mu} E_{\mu}(\cdot)E_{\mu}^{\dagger}, \quad \text{where } \sum_{\mu} E_{\mu}^{\dagger}E_{\mu} = \mathbb{1},$$

be a quantum channel. Given the channel $\mathcal{N}$, for which codes Q does there exist a recovery channel $\mathcal{R}$ such that $\mathcal{R} \circ \mathcal{N}(\rho) = \rho$ for all

$$\rho = \sum_{i} p_i \, |\phi_i\rangle \, \langle\phi_i|,$$

where $|\phi_i\rangle \in Q$?

It turns out that the set of correctable states form subspaces. The following theorem gives a necessary and sufficient condition for a recovery channel to exist.

Theorem 1.5 (Knill–Laflamme conditions). *Let $\mathcal{Q}$ be a subspace of $(\mathbb{C}^d)^{\otimes n}$. The channel $\mathcal{N}(\cdot) = \sum_{\mu} E_{\mu}(\cdot)E_{\mu}^{\dagger}$ can be corrected by a code $\mathcal{Q}$ if and only if for all $|\phi\rangle$, $|\psi\rangle$ in $\mathcal{Q}$ and errors E_{μ}, E_{ν},*

$$\langle\phi| \, E_{\mu}^{\dagger}E_{\nu} \, |\psi\rangle = c_{\mu\nu}\langle\phi|\psi\rangle$$

for some $c_{\mu\nu} \in \mathbb{C}$.

This condition implies the following two essential properties:

(1) Orthogonal code states remain orthogonal under the action of errors,

$$\text{if } \langle\phi|\psi\rangle = 0 \quad \text{then } \langle\phi| \, E_{\mu}^{\dagger}E_{\nu} \, |\psi\rangle = 0,$$

and thus orthogonal codewords remain orthogonal under the noise.

(2) The expectation value of $E_{\mu}^{\dagger}E_{\nu}$ is constant when $|\phi\rangle$ ranges over the set of code states. In other words, for all quantum states $|\phi\rangle, |\psi\rangle \in Q$,

$$\operatorname{tr}\big[|\phi\rangle\langle\phi| \, E_{\mu}^{\dagger}E_{\nu}\big] = \langle\phi| \, E_{\mu}^{\dagger}E_{\nu}|\phi\rangle = \langle\psi| \, E_{\mu}^{\dagger}E_{\nu} \, |\psi\rangle = c_{\mu\nu}.$$

In this way, the encoded quantum information is "hidden" from the noisy channel.

Lastly, a set of errors $\mathcal{E}$ is said to be *detectable* if and only if all errors $E_{\mu}^{\dagger}E_{\nu}$ with $E_{\mu}, E_{\nu} \in \mathcal{E}$ are correctable.

1.5. Error weights

We define the *weight* $\operatorname{wt}(M)$ of an operator M in the Pauli group $\mathcal{P}_n$ to be the number of tensor factors which are not equal to σ_0. For example,

$$M = \sigma_x \otimes \sigma_z \otimes \sigma_0 \otimes \sigma_y \otimes \sigma_0$$

has weight three.

In classical codes, the distance between any two elements of A^n is the number of coordinates in which they differ. If the minimum distance of a code C is at least $2t + 1$, then C is a t-error-correcting code (i.e., we can correct errors if up to t coordinates of a codeword change). In quantum codes the same holds, if a quantum code can detect all errors of weight less than $2t + 1$, then it is a t-error-correcting code.

2. Qubit stabilizer codes

2.1. Definition and examples

Most of the quantum codes presently known are stabilizer codes, and their usefulness lies partially in the fact that their connection with classical codes allows for them to be described in an efficient way. Here, we will mainly deal with stabilizer codes, although we will also see examples of quantum codes in Section 4 which are not stabilizer codes.

A *qubit stabilizer code* $Q(S)$ is the joint eigenspace with eigenvalue 1 of the elements of an abelian subgroup S of $\mathcal{P}_n$ not containing $-\mathbb{1}$. The subgroup S is also known as the *stabilizer*.

We will often define S as being generated by a set of $n - k$ commuting independent generators $M_1, \ldots, M_{n-k}$ of $\mathcal{P}_n$. By independent we mean that $M_1, \ldots, M_{n-k}$ generate S,

$$\langle M_1, \ldots, M_{n-k} \rangle = \left\{ \prod M_1^{\alpha_1} \cdots M_{n-k}^{\alpha_{n-k}} \mid \alpha_1, \ldots, \alpha_{n-k} \in \{0, 1\} \right\} = S$$

while any smaller subset does not.

It is important to note that we require $-\mathbb{1} \notin S$, since otherwise $Q(S) = \{0\}$. We also assume that there is no coordinate in which every element of S has a σ_0 in that coordinate, as we could simply delete this coordinate and this would not affect the error-correcting capabilities of the code.

Note that the phase of any element in S is ± 1, since if

$$M = \pm i \sigma_1 \otimes \cdots \otimes \sigma_n,$$

then

$$M^2 = -\mathbb{1} \in S,$$

which, as mentioned above, implies that $Q(S) = \{0\}$.

Example 2.1. Suppose $n = 2$ and S is generated by a single Pauli operator $M = \sigma_x \otimes \sigma_z$.

Let $|\alpha\rangle \in (\mathbb{C}^2)^{\otimes 2}$. Then $|\alpha\rangle$ can be written as

$$|\alpha\rangle = \alpha_{00}\,|00\rangle + \alpha_{01}\,|01\rangle + \alpha_{10}\,|10\rangle + \alpha_{11}\,|11\rangle$$

for some $\alpha_{ij} \in \mathbb{C}$. Now,

$$M\,|\alpha\rangle = \alpha_{00}\,|10\rangle - \alpha_{01}\,|11\rangle + \alpha_{10}\,|00\rangle - \alpha_{11}\,|01\rangle.$$

Thus, $|\alpha\rangle$ is in the eigenspace of M with eigenvalue 1 if and only if

$$\alpha_{00} = \alpha_{10}, \quad \alpha_{01} = -\alpha_{11}.$$

We note that the dimension of $Q(S)$ is 2.

We often use the shorthand notation

$$\sigma_0 = I, \quad \sigma_x = X, \quad \sigma_y = Y \quad \text{and} \quad \sigma_z = Z,$$

so in the previous example we might write $M = XZ$.

Example 2.2. Suppose $n = 3$ and S is generated by M_1, M_2, M_3, where

$$M_1 = \sigma_0 \otimes \sigma_x \otimes \sigma_z, \quad M_2 = \sigma_0 \otimes \sigma_y \otimes \sigma_x, \quad M_3 = \sigma_x \otimes \sigma_z \otimes \sigma_y.$$

In the shorthand notation, we would write that S is defined by

$$M_1 = I\ X\ Z, \quad M_2 = I\ Y\ X, \quad M_3 = X\ Z\ Y.$$

Observe that $M_i M_j = M_j M_i$ for all i and $j \in \{1, 2, 3\}$. For example,

$$\begin{aligned} M_2 M_1 &= (\sigma_0 \otimes \sigma_y \otimes \sigma_x)(\sigma_0 \otimes \sigma_x \otimes \sigma_z) = \sigma_0 \otimes (-i\sigma_z) \otimes (-i\sigma_y) \\ &= -\sigma_0 \otimes \sigma_z \otimes \sigma_y \end{aligned}$$

and

$$\begin{aligned} M_1 M_2 &= (\sigma_0 \otimes \sigma_x \otimes \sigma_z)(\sigma_0 \otimes \sigma_y \otimes \sigma_x) = \sigma_0 \otimes i\sigma_z \otimes i\sigma_y \\ &= -\sigma_0 \otimes \sigma_z \otimes \sigma_y. \end{aligned}$$

This can be checked quickly by verifying that different Pauli matrices $\{\sigma_x, \sigma_y, \sigma_z\}$ coincide in the same position in M_i and M_j $(i \neq j)$ an even number of times.

To find a basis for the stabilizer code, suppose that

$$|\alpha\rangle = \sum_{ijk} \alpha_{ijk}\,|ijk\rangle$$

is in the code space, i.e. that α is in the $+1$-eigenspace of all M_i.

Since

$$M_1 |\alpha\rangle = \sum_{j=0}^{1} (\alpha_{j00} |j10\rangle - \alpha_{j01} |j11\rangle + \alpha_{j10} |j00\rangle - \alpha_{j11} |j01\rangle),$$

we have that $|\alpha\rangle$ is in the $+1$-eigenspace $\widetilde{M}_1 = \mathrm{im}(I + M_1)$ of M_1 if and only if

$$\alpha_{j00} = \alpha_{j10} \quad \text{and} \quad \alpha_{j01} = -\alpha_{j11}.$$

Similarly,

$$M_2 |\alpha\rangle = i \sum_{j=0}^{1} (\alpha_{j00} |j11\rangle + \alpha_{j01} |j10\rangle - \alpha_{j10} |j01\rangle - \alpha_{j11} |j00\rangle).$$

Thus, $|\alpha\rangle$ is in the $+1$-eigenspace $\widetilde{M}_2$ if and only if

$$i\alpha_{j00} = \alpha_{j11} \quad \text{and} \quad \alpha_{j01} = -i\alpha_{j10}.$$

Finally,

$$\begin{aligned} M_3 |\alpha\rangle = i(&\alpha_{000} |101\rangle - \alpha_{001} |100\rangle - \alpha_{010} |111\rangle + \alpha_{011} |110\rangle \\ &+ \alpha_{100} |001\rangle - \alpha_{101} |000\rangle - \alpha_{110} |011\rangle + \alpha_{111} |010\rangle), \end{aligned}$$

so $|\alpha\rangle$ is in the $+1$-eigenspace $\widetilde{M}_3$ if and only if

$$i\alpha_{000} = \alpha_{101}, \quad \alpha_{100} = -i\alpha_{001}, \quad \alpha_{111} = -i\alpha_{010}, \quad \alpha_{110} = i\alpha_{011}.$$

Thus,

$$Q(S) = \widetilde{M}_1 \cap \widetilde{M}_2 \cap \widetilde{M}_3$$

is the one-dimensional subspace spanned by

$$|000\rangle - i |001\rangle + |010\rangle + i |011\rangle - |100\rangle + i |101\rangle - |110\rangle - i |111\rangle.$$

In fact, we seldom actually calculate a basis this way, as it is not necessary in practice. We have only calculated this previous example so one gets a feel of how laborious this is even for small parameters. From a practical point of view, it is enough to know the orthogonal projection P for the subspace Q.

2.2. The dimension and minimum distance of a stabilizer code

Let S be an abelian subgroup of $\mathcal{P}_n$. Let $Q(S)$ be the subspace defined as the joint eigenspace of eigenvalue 1 of the elements of S. Let $P = P(S)$ be the orthogonal projection onto the subspace $Q(S)$.

Lemma 2.3. *The orthogonal projection is*

$$P = \frac{1}{|S|} \sum_{E \in S} E.$$

Proof. Since S is an abelian subgroup, one has

$$MP = PM = P$$

for all $M \in S$.

Suppose that $|\psi\rangle \in Q(S)$. Then, $P|\psi\rangle = |\psi\rangle$ and therefore $|\psi\rangle \in \mathrm{im}(P)$.

Vice versa, if $|\psi\rangle \in \mathrm{im}(P)$, then for all $M \in S$,

$$M|\psi\rangle = MP|\phi\rangle = P|\phi\rangle = |\psi\rangle,$$

so $|\psi\rangle \in Q(S)$. Thus, $Q(S) = \mathrm{im}(P)$.

Since $E^\dagger = E$ for all $E \in \mathcal{P}_n$, we have that $P^\dagger = P$. Moreover,

$$P^2 = P \frac{1}{|S|} \sum_{M \in S} M = \frac{1}{|S|} \sum_{M \in S} PM = \frac{1}{|S|} \sum_{M \in S} M = P.$$

By Lemma 1.4, $P = P(S)$. ∎

For the proof of the next theorem, it is worth noting that

$$\mathrm{tr}(\sigma_1 \otimes \cdots \otimes \sigma_n) = \mathrm{tr}(\sigma_1) \cdots \mathrm{tr}(\sigma_n).$$

Thus, for all $E \in \mathcal{P}_n$ with phase ± 1, where $E \neq \pm \mathbb{1}$, $\mathrm{tr}(E) = 0$ and $\mathrm{tr}(\mathbb{1}) = 2^n$.

Theorem 2.4. *The stabilizer code $Q(S)$, which is the joint $+1$-eigenspace of an abelian subgroup S generated by $n - k$ independent elements, has dimension 2^k.*

Proof. By Lemma 2.3, the orthogonal projection onto $Q(S)$ is

$$P = \frac{1}{|S|} \sum_{M \in S} M.$$

The image of P is its eigenspace of eigenvalue one and also $Q(S)$.

The operator P is Hermitian and thus diagonalisable. Since $P^2 = P$, its eigenvalues are 0 and 1. The trace of P is equal to the sum of its eigenvalues, which in

the case of P is the dimension of the eigenspace of eigenvalue one. Therefore, the dimension of $Q(S)$ is equal to the trace of $P(S)$.

It only remains to note that

$$\mathrm{tr}(M) = 0$$

for all $M \in \mathcal{P}_n$ with the exception of $M = \mathbb{1}$, in which case $\mathrm{tr}(\mathbb{1}) = 2^n$. Thus, $\dim Q = 2^n/|S| = 2^k$. ∎

Having ascertained the dimension of a stabilizer code, we go on to determine its minimum distance.

Let Centraliser(S) denote the set of elements of $\mathcal{P}_n$ that commute with all elements of S, i.e., the centraliser of S in the group $\mathcal{P}_n$.

Lemma 2.5. *E is an undetectable error for the qubit stabilizer code $Q(S)$ if and only if $E \in$ Centraliser$(S) \setminus S$.*

Proof. We proceed by contradiction.

($\Rightarrow$) Suppose that E is undetectable but $E \notin$ Centraliser$(S) \setminus S$. Since any two elements of $\mathcal{P}_n$ either commute or anti-commute, $E \notin$ Centraliser(S) implies there is a $M \in S$ such that

$$EM = -ME.$$

Take any $|\psi\rangle, |\phi\rangle \in Q(S)$ with $\langle\psi|\,\phi\rangle = 0$. Then

$$\langle\psi|\,E|\phi\rangle = \langle\psi|\,ME|\phi\rangle = -\langle\psi|\,EM|\phi\rangle = -\langle\psi|\,E|\phi\rangle,$$

which implies $\langle\psi|\,E|\phi\rangle = 0$.

If $E \in S$, then

$$\langle\psi|\,E|\phi\rangle = \langle\psi|\,\phi\rangle.$$

Hence, by Theorem 1.5, E is detectable, a contradiction.

($\Leftarrow$) Suppose that E is detectable with $E \in$ Centraliser$(S) \setminus S$. Let $|\psi\rangle \in Q(S)$. Since $E \in$ Centraliser(S),

$$ME\,|\psi\rangle = EM\,|\psi\rangle = E\,|\psi\rangle$$

holds for all $M \in S$, which implies that $E\,|\psi\rangle \in Q$.

Extend $\{|\psi\rangle\}$ to an orthonormal basis B for Q. Since E is detectable,

$$\langle\phi|\,E\,|\psi\rangle = 0$$

for all $|\phi\rangle \in B \setminus \{|\psi\rangle\}$. This implies that $E\,|\psi\rangle$ is in the subspace $(B \setminus \{|\psi\rangle\})^{\perp}$. Since this subspace has as a basis $\{|\psi\rangle\}$,

$$E\,|\psi\rangle = \lambda_\psi\,|\psi\rangle$$

for some $\lambda_\psi \in \mathbb{C}$. Hence, $|\psi\rangle$ is an eigenvector of E.

By Theorem 1.5,

$$\langle\phi|\, E|\phi\rangle = \lambda_E$$

for all $|\phi\rangle \in B$. Since $\langle\psi|\psi\rangle = 1$, this implies that $\lambda_\psi = \lambda_E$.

The same argument as made above for $|\psi\rangle$ holds for all $|\phi\rangle \in Q(S)$. Thus, for all $|\phi\rangle \in Q(S)$,

$$E|\phi\rangle = \lambda_E|\phi\rangle.$$

Since $E \notin S$, then $\lambda_E \neq 1$.

The subgroup generated by S and $\lambda_E^{-1}E$ defines a smaller stabilizer code, so there is a $|\psi\rangle \in Q$ such that

$$\lambda_E^{-1}E\,|\psi\rangle \neq |\psi\rangle,$$

contradicting the above. Hence, E is not detectable. ∎

In the case $k = 0$, we have that $Q(S)$ is a one-dimensional subspace so cannot be used to store quantum information, and all errors are correctable according to the definition. However, we do not rule out considering such codes, since for any proper subgroup S' of S, the code $Q(S')$ will be of interest. Since the elements of $S \setminus S'$ will be in Centraliser$(S') \setminus S'$, Theorem 2.6 indicates that it makes sense to define the minimum distance of $Q(S)$ to be equal to the minimum weight of the non-identity elements of S. These codes are called *self-dual*, for reasons that will become clear in Theorem 2.12.

Theorem 2.6. *If $k \geqslant 1$, then the minimum distance of the 2^k-dimensional stabilizer code $Q(S)$ with stabilizer group S is equal to the minimum weight of the errors in* Centraliser$(S) \setminus S$.

Proof. According to Lemma 2.5, $Q(S)$ can detect all errors which are not elements of Centraliser$(S) \setminus S$. In particular, it can also detect all errors of weight less than the minimum weight of an error in Centraliser$(S) \setminus S$. ∎

If there are elements of S whose weight is less than the minimum distance of $Q(S)$, then the code is called *impure*. If this is not the case, then the code is called *pure*.

We should mention that there is also the concept of a degenerate code. According to Calderbank et al. [6], a non-degenerate code is one for which different errors produce linearly independent results when applied to elements of the code. Whereas a code is pure if distinct errors produce orthogonal results. It is straightforward to verify that, for additive codes, 'pure' and 'non-degenerate' coincide. In general, however, a pure code is non-degenerate, but the converse need not be true.

We use the shorthand notation $(\!(n, K, d)\!)$ to denote a quantum code of $(\mathbb{C}^2)^{\otimes n}$ of dimension K and minimum distance d. The notation $[\![n, k, d]\!]$ denotes a quantum

code of dimension 2^k. If it is a stabilizer code $Q(S)$, then d is equal to the minimum weight of the elements in Centraliser$(S) \setminus S$.

We now rewrite the Shor code from Example 1.2 as a stabilizer code.

Example 2.7 (An $[\![9, 1, 3]\!]$ code). Let S be the subgroup generated by the following elements of $\mathcal{P}_9$:

$$\begin{aligned}
M_1 &= \sigma_z \otimes \sigma_z \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0,\\
M_2 &= \sigma_0 \otimes \sigma_z \otimes \sigma_z \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0,\\
M_3 &= \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_z \otimes \sigma_z \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0,\\
M_4 &= \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_z \otimes \sigma_z \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0,\\
M_5 &= \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_z \otimes \sigma_z \otimes \sigma_0,\\
M_6 &= \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_z \otimes \sigma_z,\\
M_7 &= \sigma_x \otimes \sigma_x \otimes \sigma_x \otimes \sigma_x \otimes \sigma_x \otimes \sigma_x \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_0,\\
M_8 &= \sigma_0 \otimes \sigma_0 \otimes \sigma_0 \otimes \sigma_x \otimes \sigma_x \otimes \sigma_x \otimes \sigma_x \otimes \sigma_x \otimes \sigma_x.
\end{aligned}$$

In shorthand notation, this would be in the following way:

$$\begin{aligned}
M_I &= Z\ Z\ I\ I\ I\ I\ I\ I\ I,\\
M_2 &= I\ Z\ Z\ I\ I\ I\ I\ I\ I,\\
M_3 &= I\ I\ I\ Z\ Z\ I\ I\ I\ I,\\
M_4 &= I\ I\ I\ I\ Z\ Z\ I\ I\ I,\\
M_5 &= I\ I\ I\ I\ I\ I\ Z\ Z\ I,\\
M_6 &= I\ I\ I\ I\ I\ I\ I\ Z\ Z,\\
M_7 &= X\ X\ X\ X\ X\ X\ I\ I\ I,\\
M_8 &= I\ I\ I\ X\ X\ X\ X\ X\ X.
\end{aligned}$$

One can check that M_i and M_j commute for any i and j. Suppose that E is an error of weight at most 2. We want to prove that $E \in S$ or E does not commute with some M_i.

We proceed with a case-by-case analysis. If E has weight one and a single X or Y, then it does not commute with one of $M_1, \ldots, M_6$. If E has weight one and a single Z, then it does not commute with one of M_7, M_8.

Suppose E has an X error and, without loss of generality, suppose E has an X error in the first system. Then E must have a X or Y in the second system so that it commutes with M_1. But then it must also have a X or Z in the third system so that it commutes with M_2, contradicting the fact that it has weight two.

We leave the case-by-case analysis as an exercise but conclude that the only errors of weight two which commute with all the M_i are precisely those which are in S, i.e., $M_1, \ldots, M_6, M_1M_2, M_3M_4, M_5M_6$.

We will prove that the minimum distance of this code is 3 in a very simple manner once we have determined its geometry.

An important observation here is that the Shor code is *impure* since S contains errors of weight 2, whereas the minimum distance is 3.

We can store the same amount of information on fewer qubits with the following code.

Example 2.8 (An $[\![5, 1, 3]\!]$ code). Let S be the subgroup generated by the following elements of $\mathcal{P}_5$:

$$\begin{aligned} M_I &= X\ Z\ Z\ I\ X, \\ M_2 &= Z\ X\ I\ Z\ X, \\ M_3 &= I\ Z\ X\ Z\ Y, \\ M_4 &= Z\ I\ Z\ X\ Y. \end{aligned}$$

This representation makes the task of checking that $M_iM_j = M_jM_i$ fairly quick. We will prove that the minimum distance is 3 by considering its geometry in Example 3.15.

Let us see how we can use this example to correct errors of weight one. We perform measurements $\widehat{M}_i$ on $E|\phi\rangle$. This will return a value ± 1 (the eigenvalues of M_i). This gives us a "syndrome", a 4-tuple of signs for each error E. These are given in the following table:

	M_1	M_2	M_3	M_4		M_1	M_2	M_3	M_4		M_1	M_2	M_3	M_4
$XIIII$	+	−	+	−	$ZIIII$	−	+	+	+	$YIIII$	−	−	+	−
$IXIII$	−	+	−	+	$IZIII$	+	−	+	+	$IYIII$	−	−	−	+
$IIXII$	−	+	+	−	$IIZII$	+	+	−	+	$IIYII$	−	+	−	−
$IIIXI$	+	−	−	+	$IIIZI$	+	+	+	−	$IIIYI$	+	−	−	−
$IIIIX$	+	+	−	−	$IIIIZ$	−	−	−	−	$IIIIY$	−	−	+	+

Since each syndrome is distinct, we can use this look-up table to identify the error and correct it. An important observation here is that when we perform the measurement $\widehat{M}_i$, only the sign of the state can possibly change. Since

$$M_iE|\phi\rangle = \pm EM_i|\phi\rangle = \pm E|\phi\rangle,$$

$E|\phi\rangle$ is an eigenvector of M_i, so after measuring we will be in the state $\pm E|\phi\rangle$. Thus, we can measure consecutively each measurement $\widehat{M}_i$ for $i = 1, \ldots, n-k$.

2.3. Qubit stabilizer codes as binary linear codes

In this section, we introduce a connection between qubit stabilizer codes and classical binary linear codes. We will go on to exploit this connection to construct qubit quantum codes and then to realise a more general connection between stabilizer codes and classical codes.

Let $\mathbb{F}_q$ denote the finite field with q elements. Consider the map

$$\tau\colon \{\sigma_0, \sigma_x, \sigma_y, \sigma_z\} \to \mathbb{F}_2^2$$

defined by

$$\tau\colon \begin{cases} \sigma_0 \mapsto (0|0), \\ \sigma_x \mapsto (1|0), \\ \sigma_z \mapsto (0|1), \\ \sigma_y \mapsto (1|1). \end{cases}$$

We extend the map τ to $\mathcal{P}_n$ by applying τ to an element of $\mathcal{P}_n$ coordinatewise, where the image of the j-th position of M is the j and $(j+n)$-th coordinate in $\tau(M)$. For example,

$$\tau(\sigma_x \otimes \sigma_y \otimes \sigma_0 \otimes \sigma_x \otimes \sigma_z) = (11010 \mid 01001).$$

We draw the line between the n-th and the $(n+1)$-st coordinates, for readability sake. We ignore the phase, so $\tau(\lambda M) = \tau(M)$ for all $\lambda \in \{\pm 1, \pm i\}$. Effectively, this defines the domain of the map τ as $\mathcal{P}_n/\{\pm 1, \pm i\}$.

Lemma 2.9. *For all $M, N \in \mathcal{P}_n/\{\pm 1, \pm i\}$,*

$$\tau(MN) = \tau(M) + \tau(N).$$

Proof. Observe that the multiplicative structure, up to a phase factor (for example, we ignore the i in $\sigma_y = i\sigma_x\sigma_z$), is isomorphic to the additive structure of $\mathbb{F}_2^2$. ∎

We have established a bijection between the elements of $\mathcal{P}_n/\{\pm 1, \pm i\}$ and $\mathbb{F}_2^{2n}$. The above lemma implies that a subgroup S of $\mathcal{P}_n$ is in bijective correspondence with a subspace of $\mathbb{F}_2^{2n}$. We now wish to ascertain what property this subspace has if S is a subgroup generated by commuting elements of $\mathcal{P}_n$.

To this end, we define an alternating form for $u, w \in \mathbb{F}_2^{2n}$,

$$(u, w)_a = \sum_{j=1}^{n} (u_j w_{j+n} - u_{j+n} w_j).$$

Lemma 2.10. *For $M, N \in \mathcal{P}_n/\{\pm 1, \pm i\}$,*

$$MN = NM \quad \textit{if and only if} \quad (\tau(M), \tau(N))_a = 0.$$

Proof. Suppose $u = \tau(M)$ and $w = \tau(N)$. One can check directly that

$$u_j w_{j+n} - w_j u_{j+n} = 0$$

if and only if the Pauli matrices in the j-th position of M and N commute, and we have ± 1 otherwise.

The operators M and N commute if and only if there is an even number of positions where the Pauli matrices do not commute. This is the case if and only if there is an even number of coordinates j for which

$$u_j w_{j+n} - w_j u_{j+n} = 1,$$

a condition equivalent to $(\tau(M), \tau(N))_a = 0$. ∎

The symplectic weight of a vector $v \in \mathbb{F}_2^{2n}$ is defined as

$$|\{i \in \{1, \ldots, n\} \mid (v_i, v_{i+n}) \neq (0, 0)\}|.$$

Lemma 2.11. *The weight of $M \in \mathcal{P}_n$ is equal to the symplectic weight of $\tau(M)$.*

Proof. We have that $n - \mathrm{wt}(M)$ is equal to the number of σ_0's in M which is equal to n minus the symplectic weight of $\tau(M)$. ∎

For a subspace $C \leqslant \mathbb{F}_2^{2n}$, we define $\perp_a$ by

$$C^{\perp_a} = \left\{u \in \mathbb{F}_2^{2n} \mid (u, w)_a = 0 \text{ for all } w \in C\right\}.$$

Theorem 2.12. *S is a subgroup of $\mathcal{P}_n$ generated by $n - k$ independent mutually commuting elements if and only if $C = \tau(S)$ is an $(n - k)$-dimensional subspace of $\mathbb{F}_2^{2n}$ for which $C \leqslant C^{\perp_a}$. If $k \neq 0$, then the minimum distance of $Q(S)$ is equal to the minimum symplectic weight of the elements of $C^{\perp_a} \setminus C$. If $k = 0$, then the minimum distance of $Q(S)$ is equal to the minimum symplectic weight of the non-zero elements of $C = C^{\perp_a}$.*

Proof. The fact that the subspace $C = \tau(S)$ is contained in $C^{\perp_a}$ follows from Lemmas 2.9 and 2.10. By Theorem 2.6, for $k \neq 0$, the minimum distance is equal to the minimum weight of the images of the elements of Centraliser(S) under τ, which are not elements of the image of S. Since $C = \tau(S)$ and $C^{\perp_a} = \tau(\text{Centraliser}(S))$, the theorem follows for $k \neq 0$.

For $k = 0$, by definition, the minimum distance is equal to the minimum weight of the images of the elements of S under τ, which are the non-zero elements of C. ∎

We can construct a generator matrix $\mathrm{G}(S)$ for $C = \tau(S)$ by taking the $(n - k) \times 2n$ matrix whose i-th row is $\tau(M_i)$.

Lemma 2.13. *S is a subgroup of $\mathcal{P}_n$ generated by $n-k$ independent elements if and only if the matrix $\mathrm{G}(S)$ has rank $n-k$.*

Proof. There is a proper subset $J \subseteq \{1,\dots,n-k\}$ such that

$$\sum_{j\in J} \tau(M_j) = 0$$

if and only if the rank of $\mathrm{G}(S)$ is not equal to $n-k$. By Lemma 2.9, this is true if and only if

$$\prod_{j\in J} M_j = \mathbb{1}.$$

■

The following table serves as a useful reference:

$\mathcal{P}_n$	the Pauli group, given by n-fold tensor products of Pauli matrices $\sigma_0, \sigma_x, \sigma_y, \sigma_z$ with phases $\{\pm i, \pm 1\}$
$M_1,\dots,M_{n-k}$	the generators, a set of independent elements of $\mathcal{P}_n$ that generate S
S	the stabilizer, an abelian subgroup of $\mathcal{P}_n$
$Q(S)$	the quantum code obtained as the joint intersection of the eigenspaces of eigenvalue 1 of the operators in S
Centraliser(S)	the centraliser of S in $\mathcal{P}_n$
C	the subspace of $\mathbb{F}_2^{2n}$ obtained from the image of S under τ
$C^{\perp_a}$	the subspace of $\mathbb{F}_2^{2n}$ obtained as the image of Centraliser(S) under τ
$\mathrm{G}(S)$	the $(n-k)\times 2n$ generator matrix for C whose i-th row is $\tau(M_i)$

Example 2.14 (An $[\![5,0,3]\!]$ stabilizer code). Let S be the subgroup of $\mathcal{P}_5$ generated by the following pairwise commuting elements:

$$\begin{aligned} M_1 &= X\ Z\ I\ \ I\ Z, \\ M_2 &= Z\ X\ Z\ I\ \ I, \\ M_3 &= I\ Z\ X\ Z\ I, \\ M_4 &= I\ \ I\ Z\ X\ Z, \\ M_5 &= Z\ I\ \ I\ Z\ X. \end{aligned}$$

The matrix $\mathrm{G}(S)$ for this code is

$$\left(\begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \end{array}\right).$$

One can check directly that $(u, v)_a = 0$ for any two rows u, v of $\mathrm{G}(S)$. Alternatively, it is enough to observe that A is symmetric and that

$$(I \mid A)\begin{pmatrix} A^t \\ I \end{pmatrix} = A^t + A = A + A = 0.$$

We will prove in Example 3.15 that the minimum distance of $Q(S)$ is 3.

Observe that any $n \times n$ symmetric matrix A gives an $[\![n, 0, d]\!]$ code, where $G(S) = (I \mid A)$. The difficulty lies in choosing A so that the symplectic weight of the code generated by G (and hence d) is large.

3. The geometry of additive, linear and stabilizer codes

3.1. Additive and linear codes over a finite field

We recall that a *code* of length n is a subset C of A^n, where A is a finite set called the *alphabet*. An element of C is called a *codeword*.

The *distance* between any two elements of A^n is the number of coordinates in which they differ. The *minimum distance* of C is the minimum distance between any two codewords of C.

Suppose A is a finite abelian group with identity element 0. If $u + v \in C$ for all $u, v \in C$, then we say that C is *additive*.

The *weight* of an element (codeword) u of an additive code is the number of non-zero coordinates that it has.

Lemma 3.1. *If C is an additive code over an alphabet which is a finite abelian group, then the minimum distance d of C is equal to the minimum non-zero weight w.*

Proof. Summing $u \in C$ enough times will eventually give the n-tuple of all zeros, hence $0 = (0, \ldots, 0) \in C$. Note that this implies $-u \in C$ too.

Suppose that u is a codeword of minimum weight w. Then since $0 \in C$, we have $w \geqslant d$. Suppose that u and v are two codewords which differ in exactly d coordinates. Then $u - v$ is a codeword in C of weight d and so $d \geqslant w$. ∎

Suppose that $A = \mathbb{F}_q$, the finite field with $q = p^h$ elements, p prime. If C is additive, then $\lambda u \in C$ for all $\lambda \in \mathbb{F}_p$, so C is a subspace over $\mathbb{F}_p$. If C has the additional property that $\lambda u \in C$ for all λ in $\mathbb{F}_q$, then we say C is *linear*. A linear code of length n is a subspace of $\mathbb{F}_q^n$.

We use the notation $(n, K, d)_q$ code to denote a code over an alphabet of size q of length n, size K and minimum distance d. The notation $[n, k, d]_q$ code denotes a k-dimensional linear code over $\mathbb{F}_q$ of length n and minimum distance d.

3.2. The geometry of linear codes

We will begin our geometrical study of codes by considering linear codes over $\mathbb{F}_q$.

Let G be a $k \times n$ matrix. We recall that when a^t is a row vector in $\mathbb{F}_q^k$, the expression a^tG yields a linear combination of the rows of G. Likewise, when b is a column vector in $\mathbb{F}_q^n$, the expression Gb yields a linear combination of the columns of G.

Let C be a k-dimensional linear code over $\mathbb{F}_q$ of length n; in other words, C is a k-dimensional subspace of $\mathbb{F}_q^n$. We describe C by a $k \times n$ matrix G whose row space is C, i.e., the rows of G are a basis for C. Thus, for each $u \in C$, there is an $a^t = (a_1, \ldots, a_k) \in \mathbb{F}_q^k$ such that

$$u = a^t \mathrm{G}.$$

In other words, the *generator matrix* G acts as a linear encoding matrix for the message a, yielding the codeword u ready to be sent over a noisy channel.

The geometry of C is seen by considering the set of columns of the generator matrix G. Let $\mathcal{X}$ be the set of columns of G, so $\mathcal{X}$ is a set (possibly multiset) of n vectors of $\mathbb{F}_q^k$. The codeword $u = a^t$G has a zero in its i-th coordinate if and only if

$$a \cdot z = a_1 z_1 + \cdots + a_k z_k = 0,$$

where $z = (z_1, \ldots, z_k)$ is the i-th column of G. This property is unaffected if we replace z by a non-zero scalar multiple of z, so it is natural to consider $\mathcal{X}$ as a set (possibly multiset) of n points of PG$(k-1, q)$, the $(k-1)$-dimensional projective space over $\mathbb{F}_q$.

The projective space PG$(k-1, q)$ is obtained from the vector space $\mathbb{F}_q^k$ by identifying the vectors which are scalar multiples of each other. In this way, the *points* of PG$(k-1, q)$ are the one-dimensional subspaces of $\mathbb{F}_q^k$ and, more generally, the $(i-1)$-dimensional subspaces of PG$(k-1, q)$ are the i-dimensional subspaces of $\mathbb{F}_q^k$. The *lines, planes* and *hyperplanes* of PG$(k-1, q)$ are the one-dimensional, two-dimensional and co-dimension 1 subspaces, respectively. Note that in PG$(k-1, q)$ familiar geometric properties hold. For example, two points are joined by a line; the intersection of two planes in a three-dimensional subspace is a line. If a point x is contained in a subspace π, we say that x is *incident* with π. If two subspaces π_1 and π_2 have an empty intersection (i.e., their corresponding subspaces in $\mathbb{F}_q^k$ intersect in the zero vector), then we say that they are *skew*.

A set of points $x_1, \ldots, x_r$ of a projective space are *independent* if they span an $(r-1)$-dimensional (projective) subspace. If they are not independent, then they are *dependent*.

The number of r-tuples of linearly independent vectors of $\mathbb{F}_q^k$ is

$$(q^k - 1)(q^{k-1} - 1) \cdots (q^{k-r+1} - 1).$$

Hence, the number of r-dimensional subspaces of $\mathbb{F}_q^k$ is

$$\begin{bmatrix} k \\ r \end{bmatrix}_q := \frac{(q^k-1)(q^{k-1}-1)\cdots(q^{k-r+1}-1)}{(q^r-1)(q^{r-1}-1)\cdots(q-1)}.$$

Thus, the number of points of $\mathrm{PG}(k-1,q)$ is

$$\frac{q^k-1}{q-1} = q^{k-1}+q^{k-2}+\cdots+q+1.$$

There is a natural duality between the points of $\mathrm{PG}(k-1,q)$ and the hyperplanes of $\mathrm{PG}(k-1,q)$. A point $(a_1,\ldots,a_k)$ is mapped to the hyperplane defined as the kernel of the linear form

$$a_1X_1+\cdots+a_kX_k.$$

For example, the point $(1,-1,0)$ is mapped to the hyperplane $X_1-X_2=0$,

Thus, the number of hyperplanes of $\mathrm{PG}(k-1,q)$ is also

$$q^{k-1}+q^{k-2}+\cdots+q+1,$$

which can be checked directly by calculating $\begin{bmatrix} k \\ k-1 \end{bmatrix}_q$.

The number of lines of $\mathrm{PG}(3,q)$ is

$$\frac{(q^4-1)(q^3-1)}{(q^2-1)(q-1)} = (q^2+1)(q^2+q+1).$$

The number of points in $\mathrm{PG}(k-1,2)$ is 2^k-1, and the number of lines of $\mathrm{PG}(k-1,2)$ is $(2^k-1)(2^{k-1}-1)/3$.

Lemma 3.2. *The number of $(r-1)$-dimensional subspaces of $\mathrm{PG}(k-1,q)$ containing a fixed $(s-1)$-dimensional subspace is*

$$\begin{bmatrix} k-s \\ r-s \end{bmatrix}_q.$$

Proof. For any s-dimensional subspace U of the space $\mathbb{F}_q^k$, the quotient space $\mathbb{F}_q^k/U$ is a $(k-s)$-dimensional vector space. An r-dimensional subspace containing U is an $(r-s)$-dimensional subspace in the quotient space. Thus, the lemma holds, taking into account the dimension shift when considering the projective space. ∎

The following theorem explains what the minimum distance d of a linear code implies for the set of points $\mathcal{X}$.

Theorem 3.3. *An $[n, k, d]$ linear code over $\mathbb{F}_q$ is equivalent to a set (possibly multiset) of points $\mathcal{X}$ in $\mathrm{PG}(k-1, q)$ in which every hyperplane of $\mathrm{PG}(k-1, q)$ contains at most $n-d$ points of $\mathcal{X}$ and some hyperplane contains exactly $n-d$ points of $\mathcal{X}$.*

Proof. Let G be a $k \times n$ matrix whose row space is an $[n, k, d]$ linear code C. Let $\mathcal{X}$ be the set of columns of G viewed as points of $\mathrm{PG}(k-1, q)$.

Recall that the codeword $u = a^t\mathrm{G}$ has a zero in its i-th coordinate if and only if

$$a \cdot z = a_1 z_1 + \cdots + a_k z_k = 0,$$

where $z = (z_1, \ldots, z_k)$ is the i-th column of G.

The kernel of the linear form

$$a_1 X_1 + \cdots + a_k X_k$$

defines a hyperplane π_a of $\mathrm{PG}(k-1, q)$. The codeword $u = a^t\mathrm{G}$ has weight w if and only if u has exactly $n-w$ zero coordinates. This is the case if and only if π_a is incident with $n-w$ points of $\mathcal{X}$.

By Lemma 3.1, the minimum distance of a linear code is equal to its minimum weight. Hence, the maximum number of points of $\mathcal{X}$ on a hyperplane of $\mathrm{PG}(k-1, q)$ is $n-d$, where d is the minimum distance of C. ∎

3.3. The geometry of additive codes

An additive code C over $\mathbb{F}_q$ is linear over $\mathbb{F}_p$, where $q = p^h$ for some prime p. Therefore, $|C| = p^r$ for some r. The following theorem is the additive version of Theorem 3.3; the set of points $\mathcal{X}$ is replaced by a set of subspaces.

Theorem 3.4. *An (n, p^r, d) additive code over $\mathbb{F}_q$ with $q = p^h$ is equivalent to a set (possibly multiset) $\mathcal{X}$ of $\leqslant (h-1)$-dimensional subspaces in $\mathrm{PG}(r-1, p)$ in which every hyperplane of $\mathrm{PG}(r-1, p)$ contains at most $n-d$ subspaces of $\mathcal{X}$ and some hyperplane contains exactly $n-d$ subspaces of $\mathcal{X}$.*

Proof. Let G be an $r \times n$ matrix which is a basis for C over $\mathbb{F}_p$. As in the case of linear codes, we consider the set (possibly multiset) $\mathcal{X}$ of columns of G. However, we should not consider the elements of $\mathcal{X}$ as points of $\mathrm{PG}(r-1, q)$, since we obtain C from G by taking the row span over $\mathbb{F}_p$ and not over $\mathbb{F}_q$. Thus, we consider the elements of $\mathcal{X}$ as subspaces of $\mathrm{PG}(r-1, p)$. Suppose that $e \in \mathbb{F}_q$ is such that $\{1, e, e^2, \ldots, e^{h-1}\}$ is a basis for $\mathbb{F}_q$ over $\mathbb{F}_p$. Then, up to scalar factor, we can write $x \in \mathcal{X}$ as

$$\sum_{j=0}^{h-1} e^j x_j,$$

where $x_j \in \mathbb{F}_p^r$. We associate x with the subspace that is spanned by $x_0, \ldots, x_{h-1}$ in $\mathrm{PG}(r-1, p)$, which we denote by ℓ_x. The subspace ℓ_x has dimension at most $h-1$.

Suppose that x is the i-th column of G, so $x \in \mathcal{X}$. The non-zero codeword $u = a^t \mathrm{G}$, where $a \in \mathbb{F}_p^r$, has a zero in its i-th coordinate if and only if the hyperplane of $\mathrm{PG}(r-1, p)$, which is the kernel of the linear form

$$a_1 X_1 + \cdots + a_r X_r,$$

contains the subspace ℓ_x. ∎

Observe that a linear code over $\mathbb{F}_q$ necessarily has size q^k, so if we wish to obtain an additive code with the same parameters as a linear code, then $r = kh$ for some k.

3.4. The geometry of qubit quantum codes

For the moment, we restrict to the case $q = 2$ and consider the geometrical consequences of Theorem 2.12, which describes the connection between stabilizer codes and binary linear codes.

A qubit stabilizer code $Q(S)$ is equivalent to a binary linear code $C = \tau(S)$ of length $2n$ which is contained in its alternating dual $C^{\perp_a}$. According to Theorem 2.12, the minimum distance of $Q(S)$ is the minimum symplectic weight of $C^{\perp_a} \setminus C$.

Consider once again the Shor code from Example 1.2.

Example 3.5 (Shor code). Applying the map τ to the elements in Example 1.2, we have that $C = \tau(S)$ is the $\mathbb{F}_2$ row span of the matrix

$$G(S) = \left(\begin{array}{ccccccccc|ccccccccc}
0&0&0&0&0&0&0&0&0&1&1&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&1&1&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&1&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1\\
1&1&1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&1&1&1&1&1&1&0&0&0&0&0&0&0&0&0
\end{array}\right).$$

Since there are two columns which are linearly dependent, there are elements of $C^{\perp_a}$ of symplectic weight two; these are images under τ of Pauli operators of Centraliser(S) of weight two.

To see this, recall that the alternating form is defined as

$$(u, w)_a = \sum_{j=1}^{n} (u_j w_{j+n} - u_{j+n} w_j),$$

so the dependency of the first two columns implies that

$$(0,0,0,0,0,0,0,0,0 \mid 1,1,0,0,0,0,0,0,0)$$

is an element of $C^{\perp_a}$. However, this element is an element of C since it is the first row of the matrix. Recall that the minimum distance is equal to the minimum symplectic weight of $C^{\perp_a} \setminus C$. Therefore, although $C^{\perp_a}$ contains elements of symplectic weight 2, the minimum symplectic weight of $C^{\perp_a} \setminus C$ is in fact 3. We will prove this in Example 3.9.

Given a subgroup S generated by $n-k$ commuting elements $M_1, \ldots, M_{n-k}$ of $\mathcal{P}_n$, we obtain a set $\mathcal{X}$ of n lines or possibly points in $\mathrm{PG}(n-k-1,2)$ in the following way. For each $i \in \{1,\ldots,n\}$, we get a line (or a point) by considering the span of the i-th and $(i+n)$-th column of the generator matrix $\mathrm{G}(S)$. Vice versa, given a set of n lines in $\mathrm{PG}(n-k-1,2)$, we construct an $(n-k) \times 2n$ matrix, from which we obtain $M_1, \ldots, M_{n-k}$ by applying τ^{-1} to the rows of the matrix.

On first sight it may seem that there is a certain amount of freedom when we reconstruct the code from a given quantum set of lines. Each line is incident with three points, and we can choose which pair of points on the line to use to construct the i-th and the $(i+n)$-th column of G. This choice is equivalent to invoking a permutation of $\{\sigma_x, \sigma_y, \sigma_z\}$ on the i-th position of each of the $M_1, \ldots, M_{n-k}$. This does not affect the property that these elements pairwise commute, so we define all quantum codes that can be obtained from each other in this way to be equivalent.

Thus, in Example 2.14, invoking the permutation σ which takes

$$X \to Z \to Y \to X$$

on the M_i in the first, second and fourth positions gives

$$\begin{aligned}
\sigma(M_1) &= Z\ Y\ I\ I\ Z,\\
\sigma(M_2) &= Y\ Z\ Z\ I\ I,\\
\sigma(M_3) &= I\ Y\ X\ Y\ I,\\
\sigma(M_4) &= I\ I\ Z\ Z\ Z,\\
\sigma(M_5) &= Y\ I\ I\ Y\ X.
\end{aligned}$$

The matrix whose i-th row is $\tau(M_i)$ is

$$\left(\begin{array}{ccccc|ccccc}
0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1\\
1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0\\
0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1\\
1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0
\end{array}\right).$$

Comparing this to the matrix

$$\mathrm{G}(S) = \left(\begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \end{array}\right)$$

from Example 2.14, we see that the set of lines $\mathcal{X}$ remains unchanged.

There is also a choice between the scalar factor of M when we apply τ^{-1} to a row of the matrix G. We will always assume that this factor to be 1. However, changing the sign of some of the generators of a subgroup S can be useful, as we shall see in Section 4.

Lemma 3.6. *The span of the i-th and $(i+n)$-th column of the generator matrix* $\mathrm{G}(S)$ *is a line of* $\mathrm{PG}(n-k-1,2)$ *for all $i = 1,\dots,n$ if and only if the minimum non-zero weight of* $\mathrm{Centraliser}(S)$ *is at least two.*

Proof. We fail to obtain a line of $\mathrm{PG}(n-k-1,2)$ if and only if either the i-th and $(i+n)$-th columns of the matrix $\mathrm{G}(S)$ are the same non-zero vector, or one (or both of them) is the zero vector. This implies that in the i-th position of all the Pauli operators in S, there is either σ_0 or a fixed element $\sigma \in \{\sigma_x, \sigma_y, \sigma_z\}$. This occurs if and only if there is an element of $\mathrm{Centraliser}(S)$ of weight 1. ∎

If a quantum code $Q(S)$ is pure, then the condition that the minimum non-zero weight of $\mathrm{Centraliser}(S)$ is at least 2 can be replaced by the condition that the minimum distance of $Q(S)$ is at least 2. However, this does not need to hold for impure codes. Indeed, it could be that there are elements of $\mathrm{Centraliser}(S) \cap S$ of weight one. Yet, if the stabilizer of an $[\![n,k,d]\!]$ code $Q(S)$ contains an element of weight one, then it is easy to see that one can construct an $[\![n-1,k,d]\!]$ stabilizer code by deleting that position.

We would like to give a geometrical interpretation of the fact that the code $C = \tau(S)$ is contained in $C^{\perp_a}$.

Recall that we say two subspaces of $\mathrm{PG}(k-1,q)$ are *skew* if they do not intersect.

Theorem 3.7. *The following are equivalent:*

(1) *There is an $[\![n,k,d]\!]$ stabilizer code $Q(S)$, where S is a subgroup generated by $n-k$ independent commuting elements of $\mathcal{P}_n$ and whose centraliser contains no element of weight one.*

(2) *There is a set of n lines $\mathcal{X}$ spanning* $\mathrm{PG}(n-k-1,2)$ *with the property that every co-dimension 2 subspace is skew to an even number of the number of lines of $\mathcal{X}$.*

Proof. (1 $\Rightarrow$ 2) Let $C = \tau(S)$ and let $\mathrm{G} = \mathrm{G}(S)$ be an $(n-k) \times 2n$ generator matrix for C. From Lemma 2.13, the matrix G has rank $n-k$. Thus, its columns span PG(n − k − 1, 2). Let $\mathcal{X}$ be the set of n lines obtained for $i = 1, \ldots, n$ as the span of the i-th and $(i+n)$-th column of $\mathrm{G}(S)$.

Let $u, w \in C$, so $u = (a_1, \ldots, a_{n-k})\mathrm{G}$ and $w = (b_1, \ldots, b_{n-k})\mathrm{G}$ for some $a = (a_1, \ldots, a_{n-k}) \in \mathbb{F}_2^{n-k}$ and $b = (b_1, \ldots, b_{n-k}) \in \mathbb{F}_2^{n-k}$. One has $C \subseteq C^{\perp_a}$ if and only if

$$(u, w)_a = \sum_{j=1}^{n} (u_j w_{n+j} - w_j u_{n+j}) = 0$$

for all $u, w \in C$. We want to deduce the geometrical meaning of $(u, w)_a = 0$.

Consider a single term in the sum first. Let x and y be the j-th and the $(n+j)$-th column of G, respectively. Then

$$u_j w_{n+j} - u_{n+j} w_j = (a \cdot x)(b \cdot y) - (a \cdot y)(b \cdot x).$$

The right-hand side is zero if and only if the matrix

$$\begin{pmatrix} a \cdot x & a \cdot y \\ b \cdot x & b \cdot y \end{pmatrix}$$

has zero determinant, i.e., it has rank 1.

This is if and only if there exists $\lambda, \mu \in \mathbb{F}_2$ such that

$$a \cdot (\lambda x + \mu y) = 0 \quad \text{and} \quad b \cdot (\lambda x + \mu y) = 0.$$

Recall that we define π_a as the hyperplane which is the kernel of the linear form

$$a \cdot X = a_1 X_1 + \cdots + a_{n-k} X_{n-k}.$$

We can thus rewrite the above conditions as the requirement that the point $\lambda x + \mu y$ is contained in both π_a and π_b. In other words, there is a point on the line ℓ, spanned by x and y, which is incident with the intersection of the two hyperplanes π_a and π_b; see Figure 1.

Returning to the condition $(u, v)_a = 0$, we must therefore get an even number of ones in the sum

$$\sum_{j=1}^{n} (u_j w_{n+j} - u_{n+j} w_j).$$

All lines of $\mathcal{X}$ that are skew to $\pi_a \cap \pi_b = \ker(a \cdot X) \cap \ker(b \cdot X)$ contribute; for any given a and b, there must in total be an even number of such lines.

We note that *every* co-dimension 2 subspace of PG$(n-k-1, 2)$ can be realised in this way (as the intersection of some $a \cdot X = 0$ and $b \cdot X = 0$). This proves the forward implication.

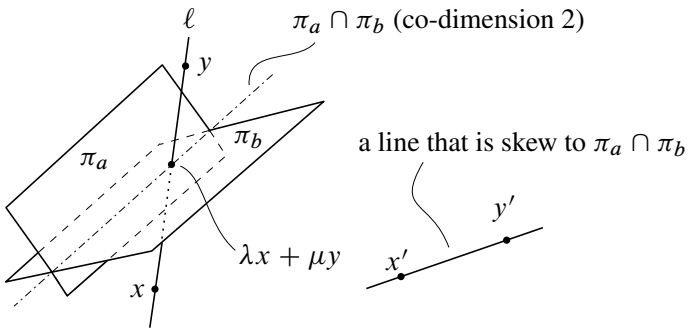

Figure 1. A point $\lambda x + \mu y$ on the intersection of the hyperplanes π_a and π_b.

($1 \Leftarrow 2$) Let $\mathcal{X}$ be a set of lines spanning $\mathrm{PG}(n-k-1,2)$ with the property that every co-dimension 2 subspace of $\mathrm{PG}(n-k-1,2)$ is skew to an even number of lines of $\mathcal{X}$. Let G be the matrix whose i-th and $(i+n)$-th columns are points which span the i-th line of $\mathcal{X}$. Let C be the code generated by G. Since $\mathcal{X}$ spans $\mathrm{PG}(n-k-1,2)$, the code C is $(n-k)$-dimensional. As we proved in the forward implication, the property that every co-dimension 2 subspace is skew to an even number of lines of $\mathcal{X}$ implies that for any two codewords u and v of C, $(u,v)_a = 0$ holds. By Lemma 2.10, the image under τ^{-1} of C is an abelian subgroup S of $\mathcal{P}_n$ and by Lemma 2.13, it is generated by $n-k$ pairwise commuting elements of $\mathcal{P}_n$. ∎

Let $\mathcal{X}$ be a set of lines and let $\Theta(\mathcal{X})$ be the space spanned by the lines of $\mathcal{X}$. We say that $\mathcal{X}$ is a *quantum set of lines* if it has the property that every co-dimension 2 subspace of $\Theta(\mathcal{X})$ is skew to an even number of lines of $\mathcal{X}$. To deduce the minimum distance of the corresponding stabilizer code, we introduce the parameter $d(\mathcal{X})$.

Recall that r points are *independent* if they span an $(r-1)$-dimensional subspace; they are *dependent* otherwise.

Consider first the case in which $\dim \Theta(\mathcal{X}) \neq |\mathcal{X}| - 1$. By Theorem 3.7, $\mathcal{X}$ will give a quantum $[\![n,k,d]\!]$ code with $k \neq 0$. We define the parameter $d(\mathcal{X})$ as the minimum number of dependent points that can be found on distinct lines of $\mathcal{X}$; not including the dependencies for which there is a hyperplane of $\Theta(\mathcal{X})$ which both

(a) contains all the lines of $\mathcal{X}$ which do not contain the dependent points,

(b) contains all the dependent points.[3]

Thus, $d(\mathcal{X}) = r$, where r is minimal such that there exists a set of dependent points $\{x_1, \ldots, x_r\}$, where each x_i is incident with a line $\ell_i \in \mathcal{X}$ and the lines

[3] In the original definition of Glynn et al. [7], condition (b) does not appear.

$\ell_1, \ldots, \ell_r$ are distinct, but for which there is no hyperplane containing the lines $\mathcal{X} \setminus \{\ell_1, \ldots, \ell_r\}$ and the points $\{x_1, \ldots, x_r\}$.

In the case in which $\dim \Theta(\mathcal{X}) = |\mathcal{X}| - 1$, Theorem 3.7 implies that $\mathcal{X}$ will give a quantum $[\![n, k, d]\!]$ code with $k = 0$. We define the parameter $d(\mathcal{X})$ as the minimum d for which there is a hyperplane of $\Theta(\mathcal{X})$ containing $|\mathcal{X}| - d$ lines of $\mathcal{X}$. Equivalently, it is the minimum number of dependent points that can be found on distinct lines of $\mathcal{X}$. This definition and the equivalence will be justified in the proof of Theorem 3.8.

From now on, we assume that the centraliser of the stabilizer S contains no elements of weight one. By Lemma 3.6, this assumption guarantees that there is a quantum set of lines associated with the stabilizer code. As mentioned before, this is equivalent to assuming that the minimum distance is at least 2 in the case of pure codes.

Theorem 3.8. *There is an $[\![n, k, d]\!]$ stabilizer code if and only if there is a quantum set of lines $\mathcal{X}$ for which $d(\mathcal{X}) = d$ and $\Theta(\mathcal{X}) = \mathrm{PG}(n - k - 1, 2)$.*

Proof. We only have to prove the part about the minimum distance since Theorem 3.7 covers the rest.

($\Rightarrow$) Let $Q(S)$ be an $[\![n, k, d]\!]$ stabilizer code given by some stabilizer S. Let $C = \tau(S)$. As in the proof of Theorem 3.7, let $\mathrm{G} = \mathrm{G}(S)$ be the $(n-k) \times 2n$ generator matrix with entries from $\mathbb{F}_2$ whose row space forms the code C. Define a set of lines

$$\mathcal{X} = \{\ell_j \mid j = 1, \ldots, n\},$$

where ℓ_j is the line that corresponds to the span of the j-th and $(j + n)$-th column of G.

Consider the case $k \neq 0$. By Theorem 2.12, the parameter d is the minimum symplectic weight of $C^{\perp_a} \setminus C$. Suppose now that $v \in C^{\perp_a}$ has symplectic weight w, and let W denote the set of positions that contribute to the weight

$$W = \{j \in \{1, \ldots, n\} \mid (v_j, v_{n+j}) \neq (0, 0)\}.$$

Clearly, $|W| = w$.

Denote by x_j the j-th column of G. Since $v = (v_1, \ldots, v_{2n})$ is in $C^{\perp_a}$, one has

$$\sum_{j \in W} (v_{n+j} x_j - x_{n+j} v_j) = 0.$$

Each summand corresponds to some point of ℓ_j. Thus, there are $w = |W|$ points on distinct lines $\{\ell_j \mid j \in W\}$ which are dependent. However, since the minimum distance d is the minimum symplectic weight of $C^{\perp_a} \setminus C$, we have to disregard this

dependency if $v \in C$. A vector v is in C if and only if $v = a\mathrm{G}$ for some $a \in \mathbb{F}_2^{n-k}$. As a consequence, $v_j = a \cdot x_j$ for all $j = 1, \ldots, 2n$.

First, consider those positions j of v that *do not* contribute to its symplectic weight, that is, $j \notin W$. For each $j \notin W$, one has that $v_j = a \cdot x_j = 0$ and $v_{n+j} = a \cdot x_{n+j} = 0$ if and only if the line l_j is contained in the hyperplane π_a described by $a \cdot X = 0$. So the lines of $\{\ell_j \mid j \in \{1, \ldots, n\} \setminus W\}$ are contained in π_a.

Second, consider those positions j of v that contribute to its symplectic weight, $j \in W$. Then

$$a \cdot (v_{n+j} x_j - x_{n+j} v_j) = v_{n+j}(a \cdot x_j) - (a \cdot x_{n+j}) v_j = v_{n+j} v_j - v_{n+j} v_j = 0$$

since $v_j = a \cdot x_j$ and $v_{n+j} = a \cdot x_{n+j}$. Hence, the dependent points are also contained in the hyperplane $a \cdot X = 0$. This exactly coincides with our definition of $d(\mathcal{X})$.

Now, consider the case $k = 0$. By Theorem 2.12, the parameter d is the minimum non-zero symplectic weight of C. Let $v \in C$ be of minimum non-zero symplectic weight. Since $v \in C$, $v = a\mathrm{G}$ for some $a \in \mathbb{F}_2^{n-k}$. Thus, $v_j = a \cdot x_j$ for all $j = 1, \ldots, 2n$. Let W denote the set of positions that contribute to the symplectic weight of v, i.e.,

$$W = \{j \in \{1, \ldots, n\} \mid (v_j, v_{n+j}) \neq (0, 0)\}.$$

Then, for $j \in W$, $a \cdot x_j = a \cdot x_{n+j} = 0$ which is equivalent to the line $\ell_j \in \mathcal{X}$ being contained in the hyperplane $a \cdot X = 0$. Therefore, there is a hyperplane of $\Theta(\mathcal{X})$ containing $|\mathcal{X}| - d$ lines of $\mathcal{X}$, which coincides with our definition of $d(\mathcal{X})$ in this case.

Alternatively, since $C = C^{\perp_a}$, the parameter d is the minimum non-zero symplectic weight of $C^{\perp_a}$. As in the case $k \neq 0$, a vector $v = (v_1, \ldots, v_{2n}) \in C^{\perp_a}$ of symplectic weight d will give a dependency of d points of $\mathcal{X}$, which coincides with our alternative definition of $d(\mathcal{X})$ in this case.

($\Leftarrow$) Vice-versa, suppose that $\mathcal{X}$ is a quantum set of lines for which $d(\mathcal{X}) = d$ and $\Theta(\mathcal{X}) = \mathrm{PG}(n-k-1, 2)$. Let $\mathrm{G} = \mathrm{G}(S)$ be the $(n-k) \times 2n$ generator matrix for a code C, whose i-th and $(i+n)$-th column span the i-th line of $\mathcal{X}$. Let $S = \tau^{-1}(C)$ and let $Q(S)$ be the stabiliser code. By Theorem 3.7 and the fact that $\Theta(\mathcal{X}) = \mathrm{PG}(n-k-1, 2)$, $Q(S)$ is an $[\![n, k, d]\!]$ stabilizer code for some d. The fact that $d = d(\mathcal{X})$ follows from the same arguments as in the forward implication, observing that if

$$a \cdot (v_{n+j} x_j - x_{n+j} v_j) = 0,$$

then

$$v_{n+j}(a \cdot x_j) - (a \cdot x_{n+j}) v_j = 0,$$

which implies $v_j = a \cdot x_j$ and $v_{n+j} = a \cdot x_{n+j}$, assuming $(a \cdot x_j, a \cdot x_{n+j}) \neq (0, 0)$. This is precisely the assumption that ℓ_j is not contained in the hyperplane π_a. ∎

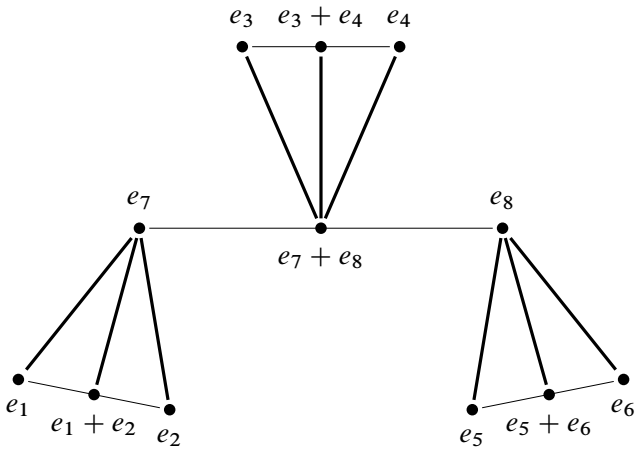

Figure 2. The set of nine (thick) lines describing the geometry of the Shor code.

Example 3.9 (Shor code). As we saw in Example 3.5, the Shor code has the generator matrix

$$\mathrm{G}(S) = \left(\begin{array}{ccccccccc|ccccccccc} 0&0&0&0&0&0&0&0&0&1&1&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&1&1&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0&1&1&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&0\\ 0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1\\ 1&1&1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&1&1&1&1&1&1&0&0&0&0&0&0&0&0&0 \end{array}\right).$$

Let e_i denote the i-th vector in the canonical basis of $\mathbb{F}_2^8$. The quantum set of lines $\mathcal{X}$ is

$$\{\langle e_1, e_7\rangle, \langle e_1 + e_2, e_7\rangle, \langle e_2, e_7\rangle, \langle e_3, e_7 + e_8\rangle, \langle e_3 + e_4, e_7 + e_8\rangle,$$
$$\langle e_4, e_7 + e_8\rangle, \langle e_5, e_8\rangle, \langle e_5 + e_6, e_8\rangle, \langle e_6, e_8\rangle\},$$

which is drawn in Figure 2. Here, $\langle e_i, e_j\rangle$ denotes the line spanned by points e_i and e_j.

Note that the point e_7 is on the two lines $\langle e_1, e_7\rangle$ and $\langle e_1 + e_2, e_7\rangle$, and thus e_7 is "dependent with itself". So at first sight it seems that $d(\mathcal{X}) = 2$. However, the remaining seven lines span a six-dimensional subspace since the two planes $\langle e_3, e_4, e_7 + e_8\rangle$ and $\langle e_5, e_6, e_8\rangle$ span a five-dimensional subspace, while the line $\langle e_2, e_7\rangle$ extends this to a six-dimensional subspace that also contains the point e_7 (i.e., contains all dependent points). Following Theorem 3.8, we do not count this dependency and conclude that $d(\mathcal{X}) \geqslant 3$. The dependency of e_7 with itself implies that the Shor code is impure. The

dependent points $\{e_1, e_2, e_1 + e_2\}$ imply that $d(\mathcal{X}) = 3$. Although the six lines not containing these points are contained in a hyperplane, there is no hyperplane containing the six lines and the dependent points, thus we do not disregard this dependency. Thus, we see that condition (b) is essential in the definition of $d(\mathcal{X})$.

Let us generalize one feature of the Shor code further: a *planar pencil of lines* in a projective space is a set of lines which are all contained in some plane and are all the lines incident with a point in that plane. As illustrated in Figure 2, the Shor code is the union of three planar pencils.

Observe that a planar pencil of lines is itself a quantum set of lines. Our aim is to show that a quantum set of lines is nothing more than the union modulo two of planar pencils of lines. We first prove a few lemmas.

Lemma 3.10. *The union modulo two of two quantum sets of lines is a quantum set of lines.*

Proof. Let $\mathcal{X}$ and $\mathcal{Y}$ be two quantum sets of lines. Recall that $\Theta(\mathcal{X})$, $\Theta(\mathcal{Y})$, and $\Theta(\mathcal{X} \cup \mathcal{Y})$ are the spaces spanned by $\mathcal{X}$, $\mathcal{Y}$, and both sets of lines, respectively. A co-dimension 2 subspace π intersects $\Theta(\mathcal{X})$ in either a co-dimension 2 subspace, in a hyperplane, or in $\Theta(\mathcal{X})$. In the first case, it is skew to an even number of the lines of $\mathcal{X}$; in the latter two cases it is skew to none (which is even).

Let $\overline{\mathcal{X}}$ be the subset of $\mathcal{X}$ of lines skew to π. Likewise, let $\overline{\mathcal{Y}}$ be the subset of $\mathcal{Y}$ of lines skew to π. Then π is skew to $|\overline{\mathcal{X}}| + |\overline{\mathcal{Y}}| - 2|\overline{\mathcal{X}} \cap \overline{\mathcal{Y}}|$ lines of the union modulo two of $\mathcal{X}$ and $\mathcal{Y}$.

Since both $|\overline{\mathcal{X}}|$ and $|\overline{\mathcal{Y}}|$ are even, every co-dimension 2 subspace is skew to an even number of lines of $\mathcal{X} \cup \mathcal{Y}$. This proves the lemma. ∎

An *r-sputnik* is a set of $(r + 1)$ concurrent lines (they are all incident with some point) in an r-dimensional subspace π with the property that any r of them span π. In Figure 4, a 3-sputnik is illustrated, compare this to Figure 3.

Our aim will be to prove that a quantum set of lines is the union modulo two of planar pencils of lines. Firstly, we will prove that this claim is true for an r-sputnik.

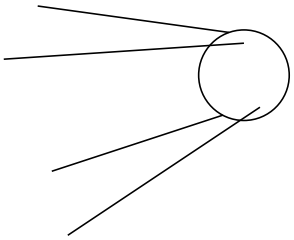

Figure 3. A 3-sputnik looks quite like a Soviet radio satellite from 1957.

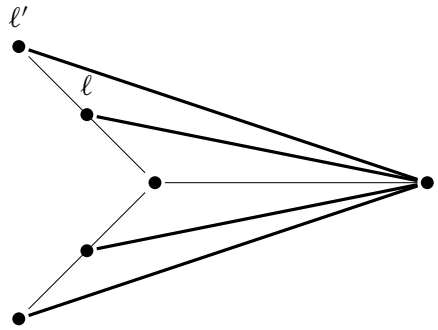

Figure 4. A 3-sputnik seen as the union modulo two of two planar pencils of lines.

Lemma 3.11. *An r-sputnik is the union modulo two of planar pencils of lines. In particular, an r-sputnik is a quantum set of lines.*

Proof. Let $\mathcal{X}$ be an r-sputnik and take any two lines ℓ and $\ell' \in \mathcal{X}$. The $r-1$ lines of $\mathcal{X} \setminus \{\ell, \ell'\}$ span an $(r-1)$-dimensional subspace which intersects the plane spanned by ℓ and ℓ' in a line ℓ''. The line ℓ'' is the third line in the planar pencil of lines spanned by ℓ and ℓ'. Thus, adding (modulo 2) this pencil of lines to $\mathcal{X}$, we get an $(r-1)$-sputnik. Now continue adding planar pencils of lines in this way until we get a 2-sputnik. Since a 2-sputnik is a planar pencil of lines, it is a quantum set of lines. We can then reverse the process adding planar pencils of lines to recover the r-sputnik which, by Lemma 3.10, is also a quantum set of lines. ∎

Lemma 3.12. *Let $\mathcal{X}$ be a quantum set of lines. There is a set D of dependent points such that each point of D is incident with a different line of $\mathcal{X}$.*

Proof. Let $\pi = \Theta(\mathcal{X})$ be the subspace spanned by the lines of $\mathcal{X}$ and let $\ell \in \mathcal{X}$. Let $\pi' = \Theta(\mathcal{X} \setminus \{\ell\})$ be the subspace spanned by the lines of $\mathcal{X} \setminus \{\ell\}$. The subspace π' is either a co-dimension 2 subspace of π, a hyperplane of π, or π itself. The first case is ruled out since $\mathcal{X}$ is a quantum set of lines and, by definition, any co-dimension 2 subspace is skew to an even number of lines of $\mathcal{X}$. Therefore, there is a point of x of ℓ incident with π'. Any point of π' is the sum of points incident with the lines of $\mathcal{X} \setminus \{\ell\}$. Thus, we obtain a set of dependent points, each incident with a line of $\mathcal{X}$. If in this set there are two points y and z incident with same line ℓ' of $\mathcal{X}$, then we can replace y and z by $\ell' \setminus \{y, z\}$. Hence, we obtain a set of dependent points, each incident with a distinct line of $\mathcal{X}$. ∎

Lemma 3.13. *A quantum set of three lines is a planar pencil of lines.*

Proof. Suppose that the quantum set of three lines $\mathcal{X} = \{\ell_1, \ell_2, \ell_3\}$ span PG$(4, 2)$ or PG$(5, 2)$, respectively. Then there is a point $x \in \ell_2$ such that the co-dimension 2

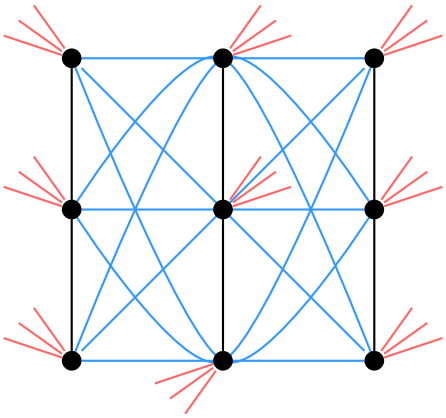

Figure 5. Configuration of the lines in PG(3, 2).

subspace spanned by ℓ_1 and x (resp. ℓ_1 and ℓ_2) is skew to ℓ_3. This contradicts the definition of a quantum set of lines.

Suppose that the quantum set of three lines $\mathcal{X} = \{\ell_1, \ell_2, \ell_3\}$ span PG(3, 2). If ℓ_1 and ℓ_2 intersect, then the co-dimension 2 subspace ℓ_1 (as well as ℓ_2) must also intersect ℓ_3. Since they span PG(3, 2), the three lines must be concurrent (and not co-planar). Taking the union modulo 2 of the planar pencil of lines spanned by ℓ_2 and ℓ_3, we obtain, by Lemma 3.10, a quantum set of two lines, which does not exist. Thus we have three pairwise skew lines ℓ_1, ℓ_2, ℓ_3 with the property that any line incident with two of them is incident with the third. This implies there are nine lines which are all incident with exactly one point of each of ℓ_1, ℓ_2, ℓ_3, see Figure 5. By Lemma 3.2, a point of PG(3, 2) is incident with seven lines of PG(3, 2), so in all we have that there are (at least)

$$9(7-4)+3+9=39$$

lines of PG(3, 2), when in fact, by Lemma 3.2, there are 35.

Therefore, the quantum set of three lines span a PG(2, 2). A co-dimension 2 subspace is just a point, so a quantum set of lines must be incident with every point of the plane. Hence, $\mathcal{X}$ is a planar pencil of lines. ∎

The following theorem is due to Glynn, Gulliver, Maks and Gupta [7]. It is important to note that if the qubit stabilizer code has minimum distance 2, then it is possible that the quantum set of lines $\mathcal{X}$ contains repeated lines. This occurs, for example, in the $[\![5, 2, 2]\!]$ code.

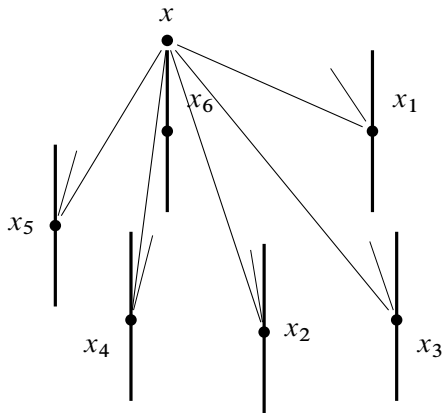

Figure 6. The thick lines are in $\mathcal{X}$, the medium-thick lines are in $\mathcal{X}'$, and the thin lines make up the planar pencils at each point $x_1, \ldots, x_r$.

Theorem 3.14. *A qubit stabilizer code with minimum distance at least three is equivalent to a quantum set of lines which is generated by the union modulo two of planar pencils of lines.*

Proof. Let $\mathcal{X}$ be a quantum set of lines. We will prove that there is an r-sputnik $\mathcal{X}'$ such that the union modulo 2 of $\mathcal{X}$, $\mathcal{X}'$ and $r-1$ planar pencils of lines is a quantum set of $|\mathcal{X}|-1$ lines. Since, by Lemma 3.11, $\mathcal{X}'$ is the union modulo 2 of planar pencils of lines, this implies that, by iteration, we can take the union modulo 2 of $\mathcal{X}$ and some planar pencils of lines and obtain a quantum set of three lines, by Lemma 3.10. By Lemma 3.13, this set of three lines is a planar pencil of lines, and we are done.

By Lemma 3.12, there is a set $x_1, \ldots, x_{r+1}$ of minimally dependent points incident with the lines $\ell_1, \ldots, \ell_{r+1}$ of $\mathcal{X}$, respectively. Let $x \in \ell_{r+1} \setminus \{x_{r+1}\}$. Let ℓ'_j be the line spanned by the points x and x_j for $j = 1, \ldots, r$. Let $\mathcal{X}'$ be the r-sputnik,

$$\mathcal{X}' = \{\ell'_j \mid j = 1, \ldots, r\} \cup \{\ell_{r+1}\}.$$

Let $\mathcal{L}_j$ be the planar pencil of lines spanned by ℓ_j and ℓ'_j. In Figure 6, $r = 5$, the lines ℓ_j are the thick lines, the ℓ'_j are the medium thickness lines, and the thin lines are the third line in the planar pencil of lines spanned by ℓ_j and ℓ'_j.

By Lemma 3.10, the union modulo two of

$$\left(\bigcup_{j=1}^{r} \mathcal{L}_j\right) \cup \mathcal{X} \cup \mathcal{X}'$$

is a quantum set of lines and, on inspection, it is a set of $|\mathcal{X}|-1$ lines. ∎

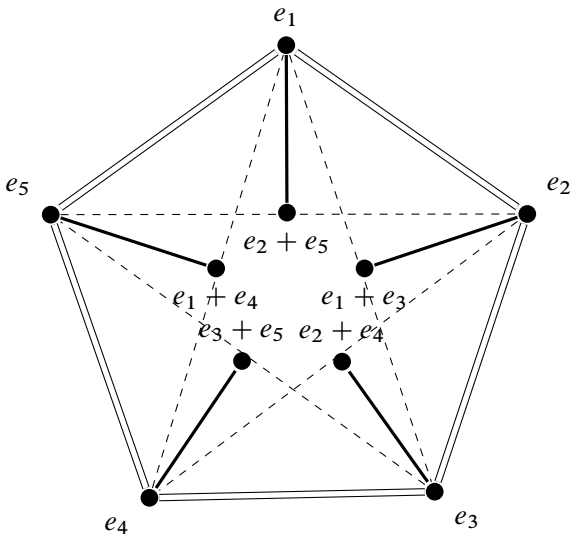

Figure 7. The $[\![5, 0, 3]\!]$ code as the union modulo two of planar pencils of lines.

Example 3.15. Consider again the $[\![5, 0, 3]\!]$ code constructed in Example 2.14. As a quantum set of lines $\mathcal{X}$, this is the union modulo two of pencils of lines drawn in Figure 7.

Since $k = 0$, $d(\mathcal{X})$ is the minimum d for which there is a hyperplane of $\mathrm{PG}(4, q)$ containing $|\mathcal{X}| - d = 5 - d$ lines of $\mathcal{X}$. Since any three lines span the whole space, we have that $d = 3$. Thus, this is an $[\![5, 0, 3]\!]$ code.

We can also construct the $[\![5, 1, 3]\!]$ code from Figure 7. We only have to replace e_5 with $e_1 + e_2 + e_3 + e_4$ and check that the five (thick) lines are then pairwise skew. This can be done by writing down the 15 points and checking we get every point of $\mathrm{PG}(3, 2)$. Then, since any two of the thick lines are pairwise skew, we have that the minimum distance is 3.

Example 3.16. The $[\![6, 0, 4]\!]$ code is the sum modulo 2 of 16 planar pencils of lines, see Figure 8. The cyclic structure allows one to check quickly that there are no three collinear points intersecting distinct lines of the six lines of the quantum set of lines. Indeed, the points of weight two obtained by summing two points incident with the quantum lines are cyclic shifts of 26, 36, 46 and the points of weight three obtained by summing two points incident with the quantum lines are cyclic shifts of 134 and 146. Therefore, the minimum distance of the code is at least 4. The points $e_{126}, e_{34}, e_{16}, e_{234}$ are four dependent points, implying that the minimum distance of the code is 4.

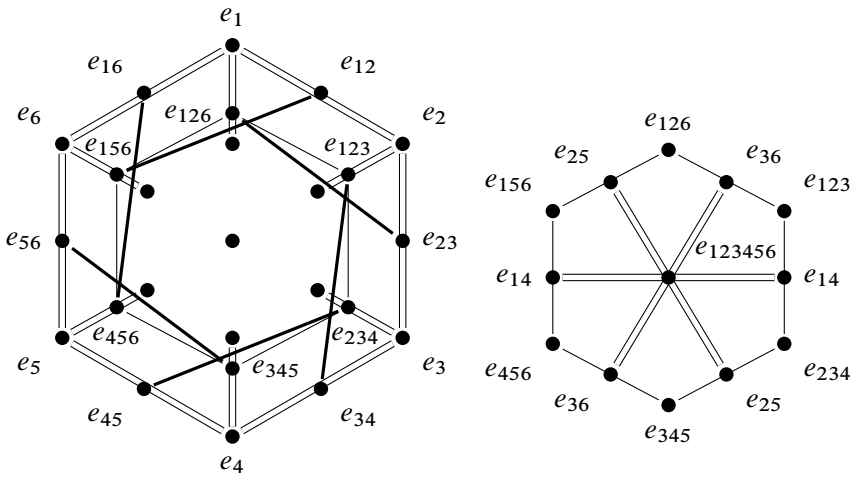

Figure 8. The quantum set of lines (the thicker lines) giving an $[\![6, 0, 4]\!]$ code.

Research Problem 1. The parameters $[\![14, 3, 5]\!]$ are the smallest for which it is unknown whether there exists a qubit stabilizer code or not [10]. To construct such a code, one should look for a union modulo two of planar pencils of lines that give 14 lines in PG(10, 2) such that for any four points on 4 of the 14 lines that also lie on a common plane, the remaining 10 lines are contained in a hyperplane which also contains those four dependent points.

Theorem 3.14 can also be used to rule out the existence of quantum codes with certain parameters sets. For example, were an $[\![4, 0, 3]\!]$ stabilizer code to exist, then $\mathcal{X}$ would be a set of four skew lines in PG(3, 2) with the property that any line is skew to an even number of lines of $\mathcal{X}$. However, the lines of $\mathcal{X}$ themselves are skew to the other three lines of $\mathcal{X}$, which is an odd number. A more interesting exercise is to prove that an $[\![7, 0, 4]\!]$ code does not exist. To prove this, show that there are at least five three-dimensional subspaces which intersect all of the 7 lines of PG(6, 2) in the quantum set of lines and prove that these pairwise intersect in a point.

4. Non-additive qubit quantum codes

4.1. Direct sum of stabilizer codes

As discussed in the previous sections, a stabilizer code is defined as the common $(+1)$-eigenspace of a set of pairwise commuting Pauli operators $M_1, \ldots, M_{n-k}$; this is the generator of the code. In other words, these codes are completely characterized by an abelian subgroup $S = \langle M_1, \ldots, M_{n-k} \rangle \subset \mathcal{P}_n$.

The aim of this section is to construct quantum codes that are the *direct sum* of stabilizer codes. Technically speaking, any subspace can be regarded as a quantum code, and naturally we want to make sure to obtain a large minimum distance when taking this direct sum of subspaces. Thus, we seek for some additional structure amongst them. While each individual subspace will again be defined by a set of generators $M_1, \ldots, M_{n-k}$, we will now not simply take the joint eigenspace with eigenvalue 1 as our code space.

We have already observed that to avoid constructing a trivial code, one restricts the stabilizer not to contain a non-trivial multiple of the identity, $-\mathbb{1} \notin S$. This implies that each generator can only have an overall phase of $+1$ or -1 and they are of the form

$$M_j = \pm\sigma_1 \otimes \cdots \otimes \sigma_n$$

for some $\sigma_1, \ldots, \sigma_n \in \mathcal{P}_1$. Now observe that when $M_1, \ldots, M_{n-k}$ commute, then so do

$$\pm M_1, \ldots, \pm M_{n-k}.$$

Thus for all $t = (t_1, \ldots, t_{n-k}) \in \{0,1\}^{n-k}$, one can define a corresponding stabilizer code $Q(S_t)$ as the joint $(+1)$-eigenspace of

$$(-1)^{t_1} M_1, \ldots, (-1)^{t_{n-k}} M_{n-k}.$$

For distinct t and $t' \in T$, there is a j such that $t_j \neq t'_j$. Without loss of generality, suppose that $t_j = 1$. For all $|v\rangle \in Q(S_t)$ and $|w\rangle \in Q(S_{t'})$, one has $\langle v|w\rangle = \langle v|M_j w\rangle = \langle M_j v|w\rangle = -\langle v|w\rangle = 0$. Consequently, $Q(S_t)$ and $Q(S_{t'})$ are orthogonal.

For any $T \subset \{0,1\}^m$, we define a *direct sum stabilizer code* (confusingly also known as a *union stabilizer code*) as

$$Q(S_T) = \bigoplus_{t \in T} Q(S_t).$$

To be able to determine the minimum distance of this quantum code, we first determine the errors which are not detectable.

As before, let G be the generator matrix whose row space is $C = \tau(S)$. Let $t, u \in T \setminus \{0\}$, and let $\mathrm{A}_{t,u}$ be an $(n-k) \times (n-k)$ non-singular matrix whose first two columns are t and u. Then $\mathrm{A}_{t,u}^{-1}\mathrm{G}$ is also a generator matrix for C, and we can find another set

$$\{M'_i \mid i = 1, \ldots, n-k\}$$

of generators of S, where M'_i is obtained from the i-th row of $A_{t,u}^{-1}G$ by applying τ^{-1}, in other words, reversing the construction above.

Let $S_{t,u}$ be the subgroup of S generated by $M'_3, \ldots, M'_{n-k}$.

Lemma 4.1. *Suppose* $|\psi^t\rangle \in Q_t(S)$ *and* $|\psi^u\rangle \in Q_u(S)$. *Then, for all* $M \in S_{t,u}$,

$$M\,|\psi^t\rangle = |\psi^t\rangle \quad \textit{and} \quad M\,|\psi^u\rangle = |\psi^u\rangle\,.$$

Proof. Observe that $Q_t(S)$ depends on the set of generators chosen for S. If we use the set of generators $M'_1, \ldots, M'_{n-k}$ for S, then $Q_t(S)$ becomes $Q_{(1,0,0,\ldots,0)}(S)$ and $Q_u(S)$ becomes $Q_{(0,1,0,\ldots,0)}(S)$. Thus, $M'_j\,|\psi^t\rangle = |\psi^t\rangle$ and $M'_j\,|\psi^u\rangle = |\psi^u\rangle$ for all $j \in \{3, \ldots, n-k\}$. ∎

Lemma 4.2. *Suppose* $Q(S_T)$ *is unable to detect an error* E. *Then there is a pair* $t, u \in T$ *such that* $E \in \text{Centraliser}(S_{t,u})$.

Proof. Suppose there is no such pair. Then, for all $t, u \in T$, there is a $M_{t,u} \in S_{t,u}$ for which E anti-commutes with $M_{t,u}$.

Suppose $|\psi^t\rangle \in Q_t(S)$ and $|\psi^u\rangle \in Q_u(S)$ are in an orthogonal basis for $Q(S_T)$. By Lemma 4.1,

$$M_{t,u}\,|\psi^t\rangle = |\psi^t\rangle \quad \text{and} \quad M_{t,u}\,|\psi^u\rangle = |\psi^u\rangle$$

and so

$$\langle\psi^t|\,E\,|\psi^u\rangle = \langle\psi^t|EM_{t,u}\,|\psi^u\rangle = -\langle\psi^t|M_{t,u}E\,|\psi^u\rangle = -\langle\psi^t|E\,|\psi^u\rangle.$$

Hence, $\langle\psi^t|E\,|\psi^u\rangle = 0$, and by Theorem 1.5, E is detectable. ∎

Thus, according to Lemma 4.2, we only need to concern ourselves with the errors which are in $\text{Centraliser}(S_{t,u})$ for any $t, u \in T$. This motivates the definition

$$d_T = \min\{d_{t,u} \mid t, u \in T\}, \tag{4.1}$$

where $d_{t,u}$ is the minimum weight of a Pauli operator in $\text{Centralise}(S_{t,u})$.

Theorem 4.3. *The subspace* $Q(S_T)$ *is an* $((n, |T|2^k, d_T))$ *quantum code.*

Proof. If error E is undetectable, then it is an element of $\text{Centraliser}(S_{t,u})$ for some $t, u \in T$. ∎

4.2. The Rains, Hardin, Shor, Sloane non-additive quantum code

The notion of a non-additive quantum code first appeared in [18], although the geometric observation given here appears to be new.

Example 4.4 (Rains, Hardin, Shor, Sloane). Consider the following elements of $\mathcal{P}_5$:

$$M_1 = Z\ X\ Y\ Y\ X,$$
$$M_2 = X\ Z\ X\ Y\ Y,$$
$$M_3 = Y\ X\ Z\ X\ Y,$$
$$M_4 = Y\ Y\ X\ Z\ X,$$
$$M_5 = X\ Y\ Y\ X\ Z.$$

The corresponding matrix whose i-th row is $\tau(M_i)$ is

$$\left(\begin{array}{ccccc|ccccc} 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \end{array}\right).$$

Observe that deleting any two rows of this matrix, we obtain a 3×10 matrix whose 5 pairs of columns define a quantum set of lines in PG$(2, 2)$. This quantum set of lines defines a stabilizer code whose minimum distance is 2. Therefore, if we set

$$T = \{\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}\},$$

then, by Theorem 4.3, $Q(S_T)$ is an $((5, 6, 2))$ quantum code.

4.3. The geometry of direct sum stabilizer codes

Let $\mathcal{X}$ be the quantum set of lines of PG$(n-k-1, 2)$ associated with the $[\![n, k, d]\!]$ quantum stabilizer code $Q(S)$, where S is the subgroup generated by $M_1, \ldots, M_{n-k}$. Let $P = \{e_1, \ldots, e_r\}$ be a set of points of PG$(n-k-1, 2)$ chosen so that the projection from any two points $e_i, e_j \in P$ of the lines of $\mathcal{X}$ is a set of lines of PG$(n-k-3, 2)$. If this projection is a set of lines, then it is necessarily a quantum set of lines, which we denote by $\mathcal{X}_{ij}$. The set T is the set of supports of the elements of P.

The parameter $d(\mathcal{X}_{ij})$ is the size of the smallest set of dependent points incident with distinct lines of $\mathcal{X}_{ij}$. Thus, the definition in (4.1) will be

$$d_T = \min\{d(\mathcal{X}_{ij}) \mid i, j \in \{1, \ldots, r\}\}.$$

Hence, we have a purely geometric way to construct direct sum stabilizer codes with parameters $((n, (r+1)2^k, d_T))$.

This is taken much further in [4], where the geometrical construction is generalised to prime alphabets.

Research Problem 2. Find quantum sets of lines $\mathcal{X}$ for which there are points with the property that the projection of the lines of $\mathcal{X}$ from any pair is onto a quantum set of lines $\mathcal{X}'$ with relatively large $d(\mathcal{X}')$. It should be possible to make direct sum stabilizer codes with good parameters from this geometrical construction. It would be of great interest if one could construct codes with parameters for which stabilizer codes could feasibly exist but none are known to exist.

5. Stabilizer codes for larger alphabets

5.1. The higher-dimensional Pauli group

When a quantum system has D levels, we speak of a quDit. In this section, we will consider quantum codes over such larger subsystems. Consequently, these codes are subspaces of the Hilbert space $(\mathbb{C}^D)^{\otimes n}$.

We will consider $(\mathbb{C}^q)^{\otimes n}$, where $q = p^h$, is the power of a prime p. The restriction to prime powers allows us to use the structure of the finite field for their construction. In the case when D is not a prime power, one can use the ring $\mathbb{Z}/D\mathbb{Z}$, but then most of the constructions that we will consider here will not work. We label the coordinates of $\mathbb{C}^q$ with elements of $\mathbb{F}_q$, where $\mathbb{F}_q$ denotes the finite field with q elements. In this way, a basis for the space of endomorphisms of $\mathbb{C}^q$ can be indexed by the elements of $\mathbb{F}_q \times \mathbb{F}_q$.

For each $a \in \mathbb{F}_q$, we define a $q \times q$ matrix $X(a)$ to be matrix obtained from the linear map which permutes the coordinates of $\mathbb{C}^q$ by adding a to the index. In other words, with basis $\{|x\rangle \mid x \in \mathbb{F}_q\}$ of $\mathbb{C}$,

$$X(a)\,|x\rangle = |x + a\rangle.$$

For example, if $q = 3$ and the elements of $\mathbb{F}_q$ are $\{0, 1, 2\}$, then

$$X(0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad X(1) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad X(2) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

For each $b \in \mathbb{F}_q$, we define a $q \times q$ matrix $Z(b)$ to be the diagonal matrix whose i-th diagonal entry is $w^{\mathrm{tr}(ib)}$. Here, $w = e^{2\pi i/p}$ is a primitive p-th root of unity, and tr is the trace map from $\mathbb{F}_q$ to its prime subfield $\mathbb{F}_p$,

$$\mathrm{tr}(a) = \sum_{j=0}^{h-1} a^{p^j}.$$

As in the previous case, if we take say $q = 3$, then

$$Z(0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad Z(1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}, \quad Z(2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega^2 & 0 \\ 0 & 0 & \omega \end{pmatrix},$$

where ω is a primitive complex third root of unity. Recall that the rows and columns of the matrix are indexed by elements of $\mathbb{F}_q$, so $i \in \mathbb{F}_q$. Thus,

$$Z(b)\,|x\rangle = \omega^{\mathrm{tr}(xb)}\,|x\rangle.$$

We define the Pauli group for q odd as

$$\mathcal{P}_1 = \left\{\omega^c X(a)Z(b) \mid a, b \in \mathbb{F}_q,\ c \in \mathbb{Z}/p\mathbb{Z}\right\}$$

and for q even, that is when $p = 2$, as

$$\mathcal{P}_1 = \left\{i^f \omega^c X(a)Z(b) \mid a, b \in \mathbb{F}_q,\ c \in \mathbb{Z}/2\mathbb{Z},\ f \in \mathbb{Z}/2\mathbb{Z}\right\}.$$

The reason that we accommodate this slightly larger group for q even is due to Lemma 5.2 below. One can check that this definition coincides with our definition of the Pauli group for $q = 2$.

More generally, we define the group of Pauli operators on $(\mathbb{C}^q)^{\otimes n}$ to be the n-fold direct product $\mathcal{P}_n = \mathcal{P}_1 \times \cdots \times \mathcal{P}_1$ (n times). Thus

$$\mathcal{P}_n = \{\sigma_1 \otimes \cdots \otimes \sigma_n \mid \sigma_j \in \mathcal{P}_1\}.$$

The size of $\mathcal{P}_n$ is pq^{2n} for q odd and $4q^{2n}$ for q even.

The weight of an element $c\sigma_1 \otimes \cdots \otimes \sigma_n$, where $\sigma_i = X(a_i)Z(b_i)$, is the number of $i \in \{1, \ldots, n\}$ such that $\sigma_i \neq X(0)Z(0)$.

Lemma 5.1. *For all $a, b \in \mathbb{F}_q^n$,*

$$\omega^{\mathrm{tr}(a\cdot b)} X(a)Z(b) = Z(b)X(a).$$

Proof. We have

$$X(a)Z(b)\,|x\rangle = \omega^{\mathrm{tr}(b\cdot x)} X(a)\,|x\rangle = \omega^{\mathrm{tr}(b\cdot x)}\,|x + a\rangle.$$

Meanwhile,

$$Z(b)X(a)\,|x\rangle = Z(b)\,|x + a\rangle = \omega^{\mathrm{tr}(b\cdot(x+a))}\,|x + a\rangle. \qquad \blacksquare$$

The following lemma implies that non-identity elements of the Pauli group have order p for q odd. Note that for q even this is not the case; there are elements of order four. However, we extend the Pauli group as above (defining $\sigma_y = i\sigma_x\sigma_z$), and in this way we introduce more elements of order two. We do this so that we have more options for M_i in our set of pairwise commuting operators which will generate the abelian subgroup S.[4]

Lemma 5.2. *For all $a, b \in \mathbb{F}_q^n$ and $r \in \mathbb{N}$,*

$$(X(a)Z(b))^r = \omega^{\binom{r}{2}\operatorname{tr}(a\cdot b)} X(a)^r Z(b)^r.$$

Proof. By induction on r, we have

$$\begin{aligned}(X(a)Z(b))^r &= (X(a)Z(b))^{r-1} X(a)Z(b)\\ &= \omega^{\binom{r-1}{2}\operatorname{tr}(a\cdot b)} X(a)^{r-1} Z(b)^{r-1} X(a)Z(b).\end{aligned}$$

By Lemma 5.1, this is equal to

$$\omega^{\binom{r-1}{2}\operatorname{tr}(a\cdot b)} X(a)^{r-1}\omega^{(r-1)\operatorname{tr}(a\cdot b)} X(a)Z(b)^{r-1}Z(b) = \omega^{\binom{r}{2}\operatorname{tr}(a\cdot b)} X(a)^r Z(b)^r. \quad \blacksquare$$

As in the case of qubit codes, we will again be looking to construct stabilizer codes and for this reason it will be of interest to know when elements $M, N \in \mathcal{P}_n$ commute or not. For this reason, the following lemma is fundamental.

Lemma 5.3. *For all $a, b, a', b' \in \mathbb{F}_q^n$,*

$$X(a)Z(b)X(a')Z(b') = \omega^{\operatorname{tr}(a\cdot b' - b\cdot a')} X(a')Z(b')X(a)Z(b).$$

Proof. $X(a)$ and $X(a')$ commute, likewise $Z(b)$ and $Z(b')$, so the lemma follows from Lemma 5.1. $\blacksquare$

5.2. Error detection and correction

As in the case of qubit codes, it suffices to consider errors from the group $\mathcal{P}_n$ of Pauli-errors which are unitary operators of the form

$$E = \sigma_1 \otimes \cdots \otimes \sigma_n,$$

where $\sigma_i = X(a)Z(b)$ for some $a, b \in \mathbb{F}_q$.

[4] This was overlooked in the seminal paper of Ketkar et al. [14] on stabilizer codes over finite fields. They do not accommodate the larger Pauli group when q is even, or include any version of Lemma 5.2. However, this larger group is necessary for all the examples of qubit stabiliser codes we have included here.

Let Q be a quantum error-correcting code of $(\mathbb{C}^q)^{\otimes n}$, i.e., a subspace of $(\mathbb{C}^q)^{\otimes n}$.

Then again, as in the case of qubit codes, Q detects an error $E \in \mathcal{P}$ if for all $|\phi\rangle, |\psi\rangle \in Q$ with $\langle\phi|\psi\rangle = 0$, we have that

$$\langle\phi| E |\psi\rangle = 0 \quad \text{and} \quad \langle\phi| E |\phi\rangle = c_E$$

for some constant c_E which depends only on E.

A quantum code Q has minimum distance d if one can detect Pauli-errors with up to $d-1$ non-identity matrices and correct Pauli-errors with up to $\lfloor \frac{d-1}{2} \rfloor$ non-identity matrices. We say that a quantum code of $(\mathbb{C}^q)^{\otimes n}$ of dimension K and minimum distance d is an $((n, K, d))_q$ code. If the code has dimension $K = q^k$, then we say that the code is an $[\![n, K, d]\!]_q$ code. Note that some authors reserve the latter notation $[\![n, K, d]\!]_q$ for stabilizer codes only.

5.3. Stabilizer codes

A *stabilizer code* is the intersection of the eigenspaces with eigenvalue one of the elements of an abelian subgroup S of $\mathcal{P}_n$. As before, we denote the code by $Q(S)$. We insist that $\lambda\mathbb{1} \notin S$ whenever $\lambda \neq 1$, since otherwise $Q(S)$ is trivial.

As in the qubit case, a stabilizer code $Q(S)$ with stabilizer S can detect all Pauli-errors that are scalar multiples of elements in S or that do not commute with some element of S. We denote by Centraliser(S) the elements of $\mathcal{P}_n$ that commute with all elements of S. A non-detectable Pauli-error must be in Centraliser(S).

Commuting elements are characterised as follows.

By Lemma 5.3, two elements $M = cX(a)Z(b)$ and $N = c'X(a')Z(b')$ satisfy

$$MN = \omega^{\mathrm{tr}(b\cdot a' - b'\cdot a)} NM.$$

Therefore, M and N commute if and only if the trace symplectic form

$$\mathrm{tr}(b \cdot a' - b' \cdot a) \tag{5.1}$$

is zero.

As in the case for qubit codes, we introduce the map τ which maps elements of $\mathcal{P}_n$ to $\mathbb{F}_q^{2n}$ by

$$\tau(X(a)Z(b)) = (a|b).$$

For elements $u, w \in \mathbb{F}_q^{2n}$, the trace symplectic form is

$$(u, w)_a = \sum_{j=1}^{n} \mathrm{tr}(u_j w_{j+n} - w_j u_{j+n}). \tag{5.2}$$

Then with $u = (a|b)$ and $w = (a'|b')$, this is the trace symplectic form (5.1).

5.4. Stabiliser codes as additive codes over $\mathbb{F}_q$

Let τ be the map that maps $cX(a)Z(b)$ to $(a|b) \in \mathbb{F}_q^{2n}$. The group S is mapped to an additive code $C = \tau(S)$. The symplectic weight of $(a|b) \in \mathbb{F}_q^{2n}$ is the number of $i \in \{1, \ldots, n\}$ such that $(a_i, b_i) \neq (0, 0)$. Thus, an element $cX(a)Z(b)$ of weight w is mapped to a vector of symplectic weight w.

The elements of Centraliser(S) are mapped to the dual code of C, namely

$$C^{\perp_a} = \{w \in \mathbb{F}_q^{2n} \mid (u, w)_a = 0 \text{ for all } u \in C\}.$$

Here the dual $\perp_a$ is taken with respect to the trace symplectic form (5.2).

We have the following important theorem.

Theorem 5.4. *An $((n, K, d))_q$ stabilizer code exists if and only if there exists an additive code $C \leqslant \mathbb{F}_q^{2n}$ of size $|C| = q^n/K$ such that $C \leqslant C^{\perp_a}$. If $K \neq 1$, then d is the minimum symplectic weight of an element of $C^{\perp_a} \setminus C$, otherwise d is the minimum symplectic weight of an element of $C^{\perp_a} = C$.*

Proof. Let S be an abelian subgroup of $\mathcal{P}_n$ not containing non-trivial multiples of the identity. Let $Q(S)$ be the corresponding $((n, K, d))_q$ stabilizer code and let

$$P = \frac{1}{|S|} \sum_{M \in S} M.$$

Then, as in Lemma 2.3, P is the orthogonal projection onto $Q(S)$. For any element $M = X(a)Z(b)$, we have that $M^\dagger M = \mathbb{1}$, so $M \in S$ if and only if $M^\dagger \in S$. Hence, $P^\dagger = P$.

Thus, since P is Hermitian and $P^2 = P$, the dimension of its image $Q(S)$ is equal to the trace of P. Since $\operatorname{tr}(M) = 0$ for all $M \in \mathcal{P}_n$, $M \neq \mathbb{1}$ and $\operatorname{tr}(\mathbb{1}) = q^n$, one has $\operatorname{tr}(P) = q^n/|S|$ and so $|S| = q^n/K$ since $\dim Q(S) = K$.

We note that $C = \tau(S)$ is an additive code since S is an abelian subgroup and has size $|S| = q^n/K$. Since $\tau(\text{Centraliser}(S)) = C^{\perp_a}$, we have $C \leqslant C^{\perp_a}$. For $K \neq 1$, the minimum symplectic weight of any element of $C^{\perp_a} \setminus C$ is d since the minimum distance of $Q(S)$ is the minimum weight of the Pauli operators in Centraliser$(S) \setminus S$. As in the qubit case, if $K = 1$, then we define the minimum distance of $Q(S)$ to be the minimum weight of the Pauli operators in Centraliser$(S) = S$, which is equal to the minimum symplectic weight of any element of $C^{\perp_a} = C$.

The backwards implication is similar. Let $S = \tau^{-1}(C)$ and define the stabilizer code to be $Q(S)$. Then the dimension follows as above. If $K \neq 1$, then the minimum distance of $Q(S)$ corresponds as above to the minimum symplectic weight of an element of $C^{\perp_a} \setminus C$ since Centraliser(S) is equal to $\tau^{-1}(C^{\perp_a})$ up to a scalar factor. If $K = 1$, then the minimum distance of $Q(S)$ corresponds to the minimum non-zero symplectic weight of the elements of $C^{\perp_a} = C$. ∎

5.5. Constructions

The following theorem is known as the Calderbank–Shor–Steane construction. The $\perp$ refers to the standard inner product on $\mathbb{F}_q^n$ given by

$$u \cdot v = u_1 v_1 + \cdots + u_n v_n.$$

Theorem 5.5. *Suppose there are linear codes C_1 and C_2 with parameters $[n, k_1, d_1]_q$ and $[n, k_2, d_2]_q$, such that $C_1^\perp \leqslant C_2$. Then there is an $[\![n, k_1 + k_2 - n, d]\!]_q$ code, where d is the minimum weight of the elements in $(C_1 \setminus C_2^\perp) \cup (C_2 \setminus C_1^\perp)$ if $k_1 + k_2 \neq n$ and d is the minimum non-zero weight of the elements in $C_1 \cup C_2$ if $k_1 + k_2 = n$.*

Proof. Let $C = C_1^\perp \times C_2^\perp \leqslant \mathbb{F}_q^{2n}$. Then C is a linear code over $\mathbb{F}_q$, and for all $v = (v_1|v_2)$ and $w = (w_1|w_2)$ in C,

$$(v, w)_a = \operatorname{tr}(v_1 \cdot w_2 - v_2 \cdot w_1) = \operatorname{tr}(0 - 0) = 0.$$

In the above, the first term vanishes since $v_1 \in C_1^\perp \leqslant C_2$ and $w_2 \in C_2^\perp$. Likewise, the second term vanishes since $v_2 \in C_2^\perp$ and $w_1 \in C_1^\perp \leqslant C_2$. Hence, $C \leqslant C^{\perp_a}$ and Theorem 5.4 applies.

To determine the minimum distance, first note that $C^{\perp_a} \geqslant C_2 \times C_1$, since for all $v = (v_1|v_2) \in C_1^\perp \times C_2^\perp$ and $w = (w_2|w_1) \in C_2 \times C_1$,

$$(v, w)_a = \operatorname{tr}(v_1 \cdot w_1 - v_2 \cdot w_2) = \operatorname{tr}(0 - 0) = 0.$$

The dimension of $C_2 \times C_1$ is $k_1 + k_2$ and the dimension of $C^{\perp_a}$ is $2n - (n - k_1) - (n - k_2) = k_1 + k_2$, so

$$C^{\perp_a} = C_2 \times C_1.$$

Thus, by Theorem 5.4, if $k_1 + k_2 \neq n$, then the minimum distance of the stabilizer code $\tau^{-1}(C)$ is the minimum weight of the elements in $(C_1 \setminus C_2^\perp) \cup (C_2 \setminus C_1^\perp)$. If $k_1 + k_2 = n$, then the minimum distance of the stabilizer code $\tau^{-1}(C)$ is the minimum non-zero weight of the elements in $C_2 \times C_1 = C_1^\perp \times C_2^\perp$, which is equal to the minimum non-zero weight of the elements in $C_1 \cup C_2 = C_1^\perp \cup C_2^\perp$. ∎

Example 5.6. The ternary extended Golay code C_1 is a $[12, 6, 6]_3$ code for which $C_1 = C_1^\perp$. Applying Theorem 5.5, this implies that there is an $[\![12, 0, 6]\!]_3$ quantum stabilizer code.

The code C_1 has a generator matrix

$$\mathrm{G} = \begin{pmatrix} 1 & 0 & 2 & 1 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 & 1 & 2 & 2 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 2 & 1 & 2 & 2 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 2 & 1 & 2 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 2 & 1 & 2 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 2 & 1 & 2 & 2 & 1 \end{pmatrix},$$

so $C = C_1 \times C_1$ has generator matrix, a 12×24 matrix

$$\left(\begin{array}{c|c} 0 & \mathrm{G} \\ \hline \mathrm{G} & 0 \end{array}\right).$$

The 12 Pauli operators generating the stabilizer group S are

$$\begin{pmatrix}
Z(1) & 1 & Z(2) & Z(1) & Z(2) & Z(2) & 1 & 1 & 1 & 1 & 1 & Z(1) \\
1 & Z(1) & 1 & Z(2) & Z(1) & Z(2) & Z(2) & 1 & 1 & 1 & 1 & Z(1) \\
1 & 1 & Z(1) & 1 & Z(2) & Z(1) & Z(2) & Z(2) & 1 & 1 & 1 & Z(1) \\
1 & 1 & 1 & Z(1) & 1 & Z(2) & Z(1) & Z(2) & Z(2) & 1 & 1 & Z(1) \\
1 & 1 & 1 & 1 & Z(1) & 1 & Z(2) & Z(1) & Z(2) & Z(2) & 1 & Z(1) \\
1 & 1 & 1 & 1 & 1 & Z(1) & 1 & Z(2) & Z(1) & Z(2) & Z(2) & Z(1) \\
X(1) & 1 & X(2) & X(1) & X(2) & X(2) & 1 & 1 & 1 & 1 & 1 & X(1) \\
1 & X(1) & 1 & X(2) & X(1) & X(2) & X(2) & 1 & 1 & 1 & 1 & X(1) \\
1 & 1 & X(1) & 1 & X(2) & X(1) & X(2) & X(2) & 1 & 1 & 1 & X(1) \\
1 & 1 & 1 & X(1) & 1 & X(2) & X(1) & X(2) & X(2) & 1 & 1 & X(1) \\
1 & 1 & 1 & 1 & X(1) & 1 & X(2) & X(1) & X(2) & X(2) & 1 & X(1) \\
1 & 1 & 1 & 1 & 1 & X(1) & 1 & X(2) & X(1) & X(2) & X(2) & X(1)
\end{pmatrix}.$$

The next construction is called the $\mathbb{F}_{q^2}$ trick (for qubit codes, this is the $\mathbb{F}_4$ trick). It is not really a trick at all, but it is a quick and effective way to construct quantum codes. These codes are a very special type of stabilizer code in which we impose more structure on the additive code C.

For any two vectors u, v in $\mathbb{F}_{q^2}^n$, we define the Hermitian form

$$u \circ v = u^q \cdot v, \tag{5.3}$$

and for an $\mathbb{F}_{q^2}$-linear code E, we define

$$E^{\perp_h} = \left\{u \in \mathbb{F}_{q^2}^n \mid u \circ v = 0 \text{ for all } v \in E\right\}.$$

Theorem 5.7. *If there exists a linear $[n, n-k, d]_{q^2}$ code D for which $D^{\perp_h} \leqslant D$, then there is an $[\![n, n-2k, \geqslant d]\!]_q$ stabilizer code.*

Proof. The code $D^{\perp_h}$ is a $[n, k, d']_{q^2}$ code for some d'. Fix a basis $\{e, e^q\}$ for $\mathbb{F}_{q^2}$ over $\mathbb{F}_q$, where $e^{2q} \neq e^2$. Let θ be the map from $\mathbb{F}_{q^2}^n$ to $\mathbb{F}_q^{2n}$ defined by

$$\theta((a_1 e + b_1 e^q, \ldots, a_n e + b_n e^q)) = (a_1, \ldots, a_n | b_1, \ldots, b_n).$$

Let $C = \theta(D^{\perp_h})$, a $2k$-dimensional linear code over $\mathbb{F}_q$ of length $2n$.

For $u \in D^{\perp_h}$ and $u' \in D$,

$$0 = u^q \cdot u' = \sum_{i=1}^{n} (a_i e + b_i e^q)^q (a_i' e + b_i' e^q).$$

This implies that

$$0 = \sum_{i=1}^{n} (a_i' b_i e^2 + b_i' a_i e^{2q} + (a_i a_i' + b_i b_i') e^{q+1}).$$

Applying the $x \mapsto x^q$ map, we get

$$0 = \sum_{i=1}^{n} (a_i' b_i e^{2q} + b_i' a_i e^2 + (a_i a_i' + b_i b_i') e^{q+1}).$$

Subtracting the last two equations, we have

$$0 = (e^{2q} - e^2) \sum_{i=1}^{n} (a_i b_i' - b_i a_i').$$

Hence,

$$(\theta(u), \theta(u'))_a = 0,$$

and so $\theta(D) \leqslant C^{\perp_a}$. Since $|D| = |C^{\perp_a}| = q^{2(n-k)}$, we have that $\theta(D) = C^{\perp_a}$.

Moreover, $C = \theta(D^{\perp_h})$ and $D^{\perp_h} \leqslant D$, so $C \leqslant C^{\perp_a}$. The symplectic weight of an element of $\theta(u)$ is equal to the weight of u, so the minimum symplectic weight of $C^{\perp_a} \setminus C$ is the minimum weight of $D \setminus D^{\perp_h}$.

The theorem then follows from Theorem 5.4. ∎

We will use the construction of Theorem 5.7 to obtain quantum MDS codes in the next section.

Research Problem 3. If k is small enough, one can multiply the columns of a generator matrix for $D^{\perp_h}$ with non-zero scalars to obtain an equivalent code for which $D^{\perp_h} \leqslant D$ holds. It would be interesting to calculate the combinatorial threshold for codes, when this can always be done, and then deduce properties of codes which surpass this threshold.

5.6. The geometry of qubit codes

In the case $q = p^h$, Theorem 5.4 implies that the existence of an $((n, q^n / p^r, d))_q$ stabilizer code $Q(S)$ is equivalent to the existence of an additive code $C \leqslant C^{\perp_a}$ of length $2n$ such that C is generated by r vectors of $\mathbb{F}_q^{2n}$ that are linearly independent over $\mathbb{F}_p$. Thus, the code C is generated by an $r \times 2n$ matrix $\mathrm{G}(S)$ over $\mathbb{F}_p$ and its columns are vectors in $\mathbb{F}_q^r$. We have seen in Section 3.3 that when $h > 1$, we should consider those columns as subspaces of $\mathrm{PG}(r-1, p)$ and not as points of $\mathrm{PG}(r-1, q)$.

Let x_i be the i-th column of the matrix $\mathrm{G}(S)$, and let e be an element of $\mathbb{F}_q$ with the property that $\{1, e, e^2, \ldots, e^{h-1}\}$ is a basis for $\mathbb{F}_q$ over $\mathbb{F}_p$.

Then there are vectors $x_{i,j} \in \mathbb{F}_p^r$ such that

$$x_i = \sum_{j=0}^{h-1} x_{i,j} e^j.$$

Let ℓ_i be the subspace

$$\ell_i = \langle x_{i,0}, \ldots, x_{i,h-1}, x_{i+n,0}, \ldots, x_{i+n,h-1} \rangle \tag{5.4}$$

as a subspace of $\mathrm{PG}(r-1, p)$.

The following lemma can be considered as a generalisation of Lemma 3.6.

Lemma 5.8. *The subspace ℓ_i is a $(2h-1)$-dimensional subspace for all $i = 1, \ldots, n$ if and only if the minimum non-zero weight of* Centraliser(S) *is at least two.*

Proof. Suppose that ℓ_i is a $(2h-1)$-dimensional subspace for all $i = 1, \ldots, n$ and that $E \in$ Centraliser(S) has weight one. Suppose that E has an $X(a)Z(b) \neq X(0)Z(0)$ in its i-th position. Consider any $M \in S$ and suppose that in the i-th coordinate M has the Pauli matrix $X(a')(Z(b')$. Since M and E commute,

$$\mathrm{tr}(a'b - b'a) = 0.$$

Thus, (a', b') is in the kernel of the linear (over $\mathbb{F}_p$) form

$$\mathrm{tr}(bX - aY).$$

The kernel of a linear form is a hyperplane of $\mathrm{PG}(2h-1, p)$, so ℓ_i has dimension at most $2h-2$, a contradiction.

Suppose that the minimum non-zero weight of Centraliser(S) is at least two and that ℓ_i is not a $(2h-1)$-dimensional subspace for some $i = 1, \ldots, n$. Since ℓ_i does not span the whole of $\mathrm{PG}(2h-1, p)$, there is an element $(a, b) \in \mathbb{F}_q^2$ such that

$$\mathrm{tr}(a'b - b'a) = 0$$

for all $X(a')Z(b')$ occurring in the i-th position of some $M \in S$. This implies that the Pauli operator of weight one E with an $X(a)Z(b)$ commutes with all $M \in S$, contradicting the fact that the minimum non-zero weight of Centraliser(S) is at least two. ∎

Thus, by Lemma 5.8, the geometry of the stabilizer code $Q(S)$ for which the minimum non-zero weight of Centraliser(S) is at least two is given by a set $\mathcal{X}$ of $(2h-1)$-dimensional subspaces of $\mathrm{PG}(r-1, p)$ of size n. The following lemma

allows us to deduce the minimum distance of $Q(S)$, at least in the case that $Q(S)$ is pure.

Lemma 5.9. *There are w dependent points incident with distinct subspaces of $\mathcal{X}$ if and only if there is an element of* Centraliser(S) *of weight w.*

Proof. Suppose that there is an element in Centraliser(S) of weight w. Then the image under τ of this element is a vector $v \in C^{\perp_a}$ with symplectic weight w. Let D be the support of v restricted to the first n coordinates. As before, let x_i be the i-th column of the matrix $\mathrm{G}(S)$, and define x_{ij} as in (5.4). Since $v \in C^{\perp_a}$,

$$\sum_{i \in D} \operatorname{tr}(v_{i+n} x_i - x_{i+n} v_i) = 0.$$

This implies

$$\sum_{i \in D} \sum_{j=0}^{h-1} (x_{ij} \operatorname{tr}(v_{i+n} e^j) - x_{i+n} \operatorname{tr}(v_i e^j)) = 0.$$

The summand is a point of the subspace ℓ_i and there are $|D| = w$ such points. This proves the backwards implication.

Suppose there are w dependent points incident with distinct subspaces of $\mathcal{X}$. Then there are a subset $D \subseteq \{1, \ldots, n\}$ of size w and $\lambda_{i,j}, \lambda_{i+n,j} \in \mathbb{F}_p$, such that

$$\sum_{i \in D} \sum_{j=0}^{h-1} (\lambda_{i,j} x_{i,j} - \lambda_{i+n,j} x_{i+n,j}) = 0.$$

Recall that

$$x_i = \sum_{j=0}^{h-1} x_{i,j} e^j.$$

Since ℓ_i is a $(2h-1)$-dimensional subspace, the points $x_j, x_j^p, \ldots, x_j^{p^{h-1}}$ are h linearly independent points, which implies there are $\mu_{i,r} \in \mathbb{F}_q$ such that

$$x_{i,j} = \sum_{r=0}^{h-1} \mu_{i,r} x_i^{p^r}.$$

Since $x_{i,j} \in \mathbb{F}_p^r$, we have that $\mu_{i,r} = \mu_i^{p^r}$ for some μ_i. Substituting in the above gives

$$\sum_{i \in D} \sum_{j=0}^{h-1} \sum_{r=0}^{h-1} (\lambda_{i,j} (\mu_i x_i)^{p^r} - \lambda_{i+n,j} (\mu_{i+n} x_{i+n})^{p^r}) = 0.$$

If we define

$$v_i = \sum_{j=0}^{h-1} \lambda_{i,j}\mu_i,$$

this equation becomes

$$\sum_{i\in D} \mathrm{tr}(v_{i+n}x_i - v_i x_{i+n}) = 0. \qquad \blacksquare$$

The property that defines $\mathcal{X}$ as a quantum set of lines for $p = 2$ does not carry over to the case $p \geqslant 3$. This is because we can scale any column of G by an element of $\mathbb{F}_q \setminus \{0, 1\}$ and not alter the set of lines $\mathcal{X}$. This will alter the value of $(u, v)_a$, so the geometric interpretation of $C \leqslant C^{\perp_a}$ will not be so clean as in the qubit case. Moreover, it is difficult to deduce the pureness of the code directly from the geometry. To see this, suppose that $v \in C^{\perp_a}$ has symplectic support D and for simplicity sake assume that q is prime. Then

$$\sum_{i\in D}(v_{i+n}x_i - v_i x_{i+n}) = 0.$$

Now, $v \in C$ if and only if there is an $a \in \mathbb{F}_p^r$ such that $v_i = a \cdot x_i$. This implies that the lines not incident with the dependent points are once again contained in a hyperplane, but we cannot deduce that the points of the dependencies are contained in the hyperplane $a \cdot X = 0$. Indeed, the fact that

$$a \cdot (v_{i+n}x_i - v_i x_{i+n}) = 0$$

implies that $(v_i, v_{i+n}) = \lambda_i(x_i, x_{i+n})$ for some non-zero scalar $\lambda_i \in \mathbb{F}_q$. Since this λ_i depends on i, we cannot deduce that $v_i = a \cdot x_i$ for all $i = 1, \ldots, 2n$.

However, this also means that when $p \geqslant 3$, we have some flexibility in choosing a basis for ℓ_i and this choice will affect whether $C \leqslant C^{\perp_a}$. Consider the set of n $(2h-1)$-dimensional subspaces of $\mathrm{PG}(4n-1, p)$ associated with a pure $[\![n, n-4, 3]\!]_q$ stabilizer code. By Lemma 5.9, these subspaces are pairwise skew. In geometrical language, this is called a *partial spread*. To construct such a code, according to Theorem 5.7, it suffices to construct a $[n, n-2, 3]_{q^2}$ linear code D for which $D^{\perp_h} \leqslant D$. Such a code is has a generator matrix

$$\begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ y_1 & y_2 & \ldots & y_n \end{pmatrix},$$

where $x_i y_j \neq x_j y_i$ and

$$\sum_{i=1}^{n} x_i^{q+1} = \sum_{i=1}^{n} y_i^{q+1} = \sum_{i=1}^{n} x_i^q y_i = 0. \tag{5.5}$$

For any $n \leqslant q^2 + 1$, such a matrix can be found by scaling the first three columns so that the equations in (5.5) are satisfied.

Research Problem 4. Glynn et al. [7] developed the geometry of qubit stabilizer codes, introducing the concept of a quantum set of lines. This led them to prove Theorem 3.14, which gives a beautiful geometric classification of qubit stabilizer codes. Here, we have generalised the concept of quantum set of lines to non-qubit stabilizer codes. Although we have seen that the existence of non-identity non-zero scalars means we cannot hope for such a clean geometric classification, one can certainly expect some geometric classification for larger q.

6. Quantum MDS codes

6.1. Stabiliser MDS codes

Let C be a code of length n and minimum distance d over an alphabet of size q. If we consider any $n - (d - 1)$ coordinates, then any two codewords must be different on these coordinates (if not the distance between them is at most $d - 1$), so there are at most q^{n-d+1} codewords in the code. This is the *Singleton bound*

$$|C| \leqslant q^{n-d+1}.$$

A code which attains the Singleton bound is called a *maximum distance separable code* or simply an MDS code.

Recall that if C is an additive code over $\mathbb{F}_q$, where $q = p^h$ for some prime p, then C is linear over $\mathbb{F}_p$ and so necessarily $|C| = p^r$ for some r, see Section 3.3. Thus, if C is also an MDS code, then h divides r and $|C| = q^k$, where $k = n - d + 1$.

Theorem 5.4 states that an $[\![n, k, d]\!]_q$ stabilizer code exists if and only if there exists an additive code $C \leqslant \mathbb{F}_q^{2n}$ of size $|C| = q^{n-k}$ such that $C \leqslant C^{\perp_a}$ and the minimum symplectic weight of an element of $C^{\perp_a} \setminus C$ is d. Considering $C^{\perp_a}$ as a code over the alphabet $\mathbb{F}_q \times \mathbb{F}_q$, then $C^{\perp_a}$ has minimum weight d, so

$$|C^{\perp_a}| \leqslant q^{2n-2d+2}.$$

Since $|C| = q^{n-k}$, we have that $|C^{\perp_a}| = q^{n+k}$, which implies that for an $[\![n, k, d]\!]_q$ stabilizer code to exist, we must have the condition $k \leqslant n - 2(d - 1)$. Compare this with the Singleton bound above $k \leqslant n - (d - 1)$ for codes of size q^k.

What is perhaps surprising is that this bound holds for all $[\![n, k, d]\!]_q$ quantum codes. The *quantum Singleton bound* states that

$$n \geqslant k + 2(d - 1).$$

Consequently, codes reaching equality are called *quantum maximum distance separable codes*, or QMDS codes for short. We will prove this bound in Section 6.3.

6.2. Reed–Solomon codes

The classical example of an MDS code is the following linear code over $\mathbb{F}_q$. Denote by $\{a_1, \ldots, a_q\}$ the elements of $\mathbb{F}_q$. The *Reed–Solomon code* is

$$C = \{(f(a_1), \ldots, f(a_q), f_{k-1}) \mid f \in \mathbb{F}_q[X],\ \deg f \leqslant k-1\},$$

where f_{k-1} denotes the coefficient of X^{k-1} in $f(X)$. If $k \leqslant q$, then each polynomial f defines a different codeword, so the dimension of C is k. A non-zero codeword has weight at least $n-k+1$ since a polynomial of degree at most $k-1$ has at most $k-1$ zeros. Lemma 3.1 then implies that the minimum distance $d = n-k+1$ and so the code is MDS.

We can use Theorem 5.7 to construct quantum stabilizer codes from Reed–Solomon codes over $\mathbb{F}_{q^2}$, but only if we can scale the coordinates of C so that $C \leqslant C^{\perp_h}$. Then $D = C_h^{\perp}$ is a $[n, n-k, k+1]_{q^2}$ linear MDS code with the property that $D_h^{\perp} \leqslant D$. Observe that replacing the i-th coordinate $f(a_i)$ by $\lambda_i f(a_i)$ does not alter the parameters of the code. Such a code is then called a *generalised Reed–Solomon code*. This can only be done for $k \leqslant q$, and we obtain an $[\![q^2+1, q^2+1-2k, k+1]\!]_q$ stabilizer code. For case $k = q$, one can check that the Reed–Solomon code

$$\{(f(a_1), \ldots, f(a_{q^2}), f_{q-1}) \mid f \in \mathbb{F}_{q^2}[X],\ \deg f \leqslant q-1\}$$

is contained in its Hermitian dual, so there is no need to scale in this case.

6.3. Quantum Singleton bound

To prove the quantum Singleton bound we will need some technical tools.

1. *Bloch decomposition.* Let $\{e_i\}$ be a basis for the space of complex $D \times D$ matrices such that $\operatorname{tr}(e_i^\dagger e_j) = D\delta_{ij}$. For qubits, take for example the Pauli matrices. Every one-quDit density matrix can then be expanded as

$$\rho = \frac{1}{D} \sum_i \operatorname{tr}(e_i^\dagger \rho) e_i,$$

where we recall that the trace of a matrix is given by the sum of its diagonal elements, $\operatorname{tr}(M) = \sum_i m_{ii}$ for any square matrix $M = (m_{ij})$.

Consider now an n-partite system in the space $(\mathbb{C}^D)^{\otimes n}$. Denote by $\{E_\alpha\}$, with a multi-index $\alpha = (\alpha_1, \ldots, \alpha_n)$, the matrix basis formed by tensor products of the e_i's

$$E_\alpha = e_{\alpha_1} \otimes \cdots \otimes e_{\alpha_n}.$$

For tensor products, such as say $E \otimes F$, one has $\mathrm{tr}(E \otimes F) = \mathrm{tr}(E) \cdot \mathrm{tr}(F)$. In other words, the trace of a tensor product factorizes. Consequently, $\mathrm{tr}(E_\alpha^\dagger E_\beta) = D^n \delta_{\alpha\beta}$, and the matrix basis formed by $\{E_\alpha\}$ is orthogonal.

Denote by $\mathrm{wt}(E_\alpha)$ the number of non-identity terms in the tensor-decomposition, and by $\mathrm{supp}(E_\alpha)$ the collection of sites where the non-identity terms act on. Naturally, $\mathrm{wt}(E_\alpha) = |\mathrm{supp}(E_\alpha)|$.

We can expand an n-partite state as

$$\rho = \frac{1}{D^n} \sum_E \mathrm{tr}(E^\dagger \rho) E.$$

As above, we from now on omit the index α for readability. This is the Bloch decomposition of ρ.

2. *Partial trace.* Consider the linear function tr_j which maps

$$\mathrm{tr}_j \colon e_{\alpha_1} \otimes \cdots \otimes e_{\alpha_n} \mapsto \mathrm{tr}(e_{\alpha_j}) \cdot e_{\alpha_1} \otimes \cdots \otimes e_{\alpha_{j-1}} \otimes e_{\alpha_{j+1}} \otimes \cdots \otimes e_{\alpha_n}.$$

The function tr_j is called the *partial trace* and its action can be understood as that of removing the j-th tensor component.

The partial trace does not depend on the basis. Its coordinate-free definition is the following: Let V and W be two vector spaces and denote by I_W the identity matrix on W. The partial trace tr_W is the unique operator, which for all M acting on $V \otimes W$ and all N acting on V satisfies

$$\mathrm{tr}(M \cdot (N \otimes I_W)) = \mathrm{tr}(\mathrm{tr}_W(M) \cdot N).$$

Considering the Hilbert–Schmidt inner product $\langle M, N \rangle = \mathrm{tr}(M^\dagger N)$, the partial trace can be seen as the adjoint to the map $V \to V \otimes I_W$. Note that partial traces over different subsystems commute, $\mathrm{tr}_j \mathrm{tr}_i = \mathrm{tr}_i \mathrm{tr}_j$, and one has that

$$\mathrm{tr}(M_1 \otimes M_2 \otimes \cdots \otimes M_n) = \mathrm{tr}(M_1)\,\mathrm{tr}(M_2) \cdots \mathrm{tr}(M_n).$$

3. *Purification.* A density matrix ρ on $\mathcal{H}_A$ can always be diagonalized as

$$\rho = \sum_{i=1}^{\dim(\mathcal{H}_A)} \lambda_i \, |\lambda_i\rangle\langle\lambda_i|_A \,,$$

where $\{|\lambda_i\rangle_A\}$ is its set of eigenvectors and $\{\lambda_i\}$ is its set of corresponding eigenvalues.

The density matrix ρ acting on some Hilbert space $\mathcal{H}_A$ can always be represented as the reduction or marginal of a pure state on $\mathcal{H}_A \otimes \mathcal{H}_B$ with $\dim(\mathcal{H}_B) \geq \dim(\mathcal{H}_A)$.

This works as follows: choose an orthonormal basis $\{|\lambda_i\rangle^B\}$ for an arbitrary $\dim(\mathcal{H}_A)$-dimensional subspace of $\mathcal{H}_B$. We then write

$$|\phi\rangle = \sum_{i=1}^{\dim(\mathcal{H}_A)} \sqrt{\lambda_i}\, |\lambda_i\rangle_A \otimes |\lambda_i\rangle^B .$$

It can be checked that $\mathrm{tr}_B(|\phi\rangle\langle\phi|) = \rho$ and the state $|\phi\rangle$ is known as a *purification* of ρ.

4. *Von Neumann entropy.* Consider a classical probability distribution represented by a set of probabilities $p_i \geq 0$ with $\sum_i p_i = 1$. Its *Shannon entropy* is

$$S(p) = -\sum_i p_i \log(p_i).$$

We can introduce a similar quantity for quantum states. Given a density matrix ρ, its von Neumann entropy is defined as

$$S(\rho) = -\operatorname{tr} \rho \log(\rho).$$

Such matrix functions of Hermitian operators can be evaluated on their eigenvalues $\{\lambda_i\}$. Then the von Neumann entropy evaluates as

$$S(\rho) = -\sum_i \lambda_i \log(\lambda_i).$$

Let us now write $S_A = S(\mathrm{tr}_B[\rho_{AB}])$ and so on. For a state ρ on $\mathcal{H}_A$ with purification $|\phi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, we have that $S_A = S_B$.

The von Neumann entropy satisfies *subadditivity* and *strong subadditivity*,

$$S_{AB} \leq S_A + S_B,$$
$$S_{ABC} + S_B \leq S_{AB} + S_{BC}.$$

We are now in position to prove the quantum Singleton bound.

Theorem 6.1 (Quantum Singleton bound). *Any* $[\![n, k, d]\!]_q$ *code with* $k \geq 1$ *satisfies*

$$n \geq k + 2(d-1).$$

Proof. The distance must be bounded by $2(d-1) < n$, as otherwise $n - (d-1) < (d-1)$ and we could recover the encoded state from two disjoint subsystems, violating the no-cloning theorem.

Let $\Pi_{\mathcal{Q}} = \sum_{i=1}^{q^k} |v_i\rangle\langle v_i|$ be the projector onto the code space. A purification with a reference system R leads to

$$|\psi_{QR}\rangle = \frac{1}{\sqrt{q^k}} \sum_{i=1}^{q^k} |v_i\rangle \otimes |i_R\rangle,$$

where $|i_R\rangle$ is any orthonormal basis for R. Let us partition the code into the three subsystems A, B, C such that $|A| = |B| = d-1$ and $|C| = n - 2(d-1)$. Then $S_R = \log(q^k)$. As the code has distance d, any subsystem of size strictly smaller than d cannot reveal anything about the reference system R: indeed, the condition of $\rho_{RA} = \rho_R \otimes \rho_A$ is known to be a necessary and sufficient condition for the subsystem A to be correctable [16]; this is also equivalent to $S_{RA} = S_R + S_A$. With the subadditivity of the von Neumann entropy this leads to

$$\begin{aligned} S_R + S_A = S_{RA} = S_{BC} \leq S_B + S_C, \\ S_R + S_B = S_{RB} = S_{AC} \leq S_A + S_C, \end{aligned}$$

where we used that the entropies of complementary subsystems are equal for a pure state. The combination of the above two inequalities yields

$$\log q^k = S_R \leq S_C \leq \log \dim(\mathcal{H}_C) = \log q^{n-2(d-1)}. \qquad \blacksquare$$

Similar to classical MDS codes, quantum MDS are, in a certain sense, extremal. We have the following interesting properties:

(a) If an $[\![n, n-2d+2, d]\!]$ quantum MDS code exists, then there is also an $[\![n-s, n-2d+2+s, d-s]\!]$ codes for all $0 \leq s \leq d$, see Figure 9.

(b) For every subset $S \subset \{1, \ldots, n\}$ with $|S| \leq \frac{n+k}{2}$, we have that $\mathrm{tr}_{S^c}(P) \propto \mathbb{1}$, where P is the orthogonal projection onto the quantum MDS code.

Let us discuss these properties: (a) states that QMDS codes form families of codes where $n+k$ is constant. Within each family, only the member with the highest distance has to be determined, as its descendants can be obtained by a partial trace: tracing out over a single particle, one has $n \mapsto n-1$, $k \mapsto k+1$, $d \mapsto d-1$. This works because QMDS codes are *pure codes*, that is, all their $(d-1)$-party marginals are maximally mixed. For general quantum codes, this method of making new codes from old is not necessarily possible.

Property (b) states that for all pure states $|v\rangle$ in the code, the marginals of size less than d are maximally mixed. This implies that every vector in the code space shows maximal bipartite entanglement across any bipartition of $d-1$ vs. $n-d+1$ parties. Thus QMDS codes form subspaces that show high bipartite entanglement. We relate this to similar property of classical MDS codes: the parity check matrix H

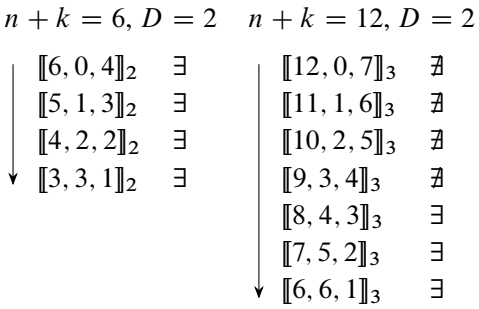

Figure 9. Two families of quantum MDS codes. Once the topmost existing parent code is known, (here: $[\![6,0,4]\!]_2$ and $[\![8,4,3]\!]_3$), its descendants can be obtained by partial traces.

of a classical $[n,k,d]$ code has the property that every set of $n-k$ columns are linearly independent.

A necessary condition for QMDS to exist is the following bound.

Proposition 6.2 ([13]). *If there is an $[\![n,n-2d+2,d]\!]_q$ quantum MDS code, then*

$$n \leqslant q^2 + d - 2.$$

This should be compared to the "trivial" upper bound for MDS codes. If there is a $(n,q^k,n-k+1)_q$ MDS code, then

$$n \leqslant q + k - 1.$$

The MDS conjecture states that if $4 \leqslant k \leqslant q$ and there is a $(n,q^k,n-k+1)_q$ MDS code, then $n \leqslant q+1$. This is known to hold for linear codes if q is a prime, see [3].

For quantum MDS codes, the MDS conjecture states that if $5 \leqslant d \leqslant q^2-1$ and there is a linear $[\![n,n-2d+2,d]\!]_q$ MDS code, then

$$n \leqslant q^2 + 1.$$

Ketkar [14, Corollary 65] claims that if the classical MDS conjecture holds for linear codes, then quantum MDS conjecture holds for stabilizer codes. This is not the case. By Theorem 5.4, the existence of a stabilizer code is equivalent to the existence of an additive code, so [14, Corollary 65] should state that the quantum MDS conjecture holds for stabilizer codes if the MDS conjecture holds for additive codes.

Research Problem 5. Prove the MDS conjecture for linear codes with q non-prime.

Research Problem 6. Prove the MDS conjecture for additive codes over $\mathbb{F}_q$, starting with $q = p^2$ for some prime p.

Research Problem 7. Find all inequalities that relate the von Neumann entropies of the marginals of multipartite systems.

Research Problem 8. Show that all QMDS codes are either stabilizer codes or the direct sum of stabilizer codes.

7. Weight enumerators

7.1. MacWilliams identity for linear codes

Let C be an $[n, k, d]_q$ code and define A_i to be the number of codewords of C of weight i, i.e., the number of codewords of C which have i non-zero coordinates. Since the zero codeword is in C, $A_0 = 1$ and since the minimum distance is d, $A_i = 0$ for all $i = 1, \ldots, d-1$. Let B_i denote the number of codewords of $C^\perp$ of weight i. The MacWilliam's identities relate the polynomials

$$A(x, y) = \sum_{i=1}^{n} A_i x^{n-i} y^i \quad \text{and} \quad B(x, y) = \sum_{i=1}^{n} B_i x^{n-i} y^i.$$

Specifically, we have that

$$|C|B(x, y) = A(y + (q-1)x, y - x),$$

and dually

$$|C^\perp|A(x, y) = B(y + (q-1)x, y - x).$$

Let G be a $k \times n$ generator matrix for C and let $\mathcal{X}$ be the set or multi-set of columns of G, viewed as points of $\mathrm{PG}(k-1, q)$. In Section 3.2, we saw that a non-zero codeword $u = a\mathrm{G}$ corresponds to a hyperplane π_a of $\mathrm{PG}(k-1, q)$ and that $\pi_a = \pi_{\lambda a}$ for any $\lambda \in \mathbb{F}_q$. The number of points of $\mathcal{X}$ incident with the hyperplane π_a is n minus the weight of the codeword u. Thus, for $i \neq 0$, there are $A_i/(q-1)$ hyperplanes which are incident with $n - i$ points of $\mathcal{X}$.

7.2. MacWilliams identity for quantum codes

As for classical codes, weight enumerators can be defined for quantum codes, which again are useful to deduce the error-correcting properties of codes and to obtain bounds on their existence.

Let Q be a quantum code and let P be the orthogonal projection onto Q. The weights of the primary and secondary *Shor–Laflamme enumerators* are

$$A_j = \sum_{\mathrm{wt}(E)=j} \mathrm{tr}(EP)\,\mathrm{tr}(E^\dagger P), \quad B_j = \sum_{\mathrm{wt}(E)=j} \mathrm{tr}(EPE^\dagger P),$$

where the sum is over Pauli operators E of weight j and phase 1.

The enumerator polynomials are given by

$$A(x,y)=\sum_{j=0}^{n}A_j x^{n-j}y^j,\quad B(x,y)=\sum_{j=0}^{n}B_j x^{n-j}y^j.$$

Lemma 7.1. *For a stabilizer code, A_j is $q^{2n}/|S|^2$ times the number of elements in the stabilizer subgroup S that have weight j. Similarly, B_j is $q^n/|S|$ times the number of elements in the normaliser of S of weight j.*

Proof. By Lemma 2.3,

$$P=\frac{1}{|S|}\sum_{M\in S}M.$$

The map tr is linear and $\operatorname{tr}(M)=0$ unless $M=\mathbb{1}$ and $\operatorname{tr}(\mathbb{1})=q^n$.

Hence, if $E\notin S$,

$$\operatorname{tr}(EP)\operatorname{tr}(E^\dagger P)=0$$

and if $E\in S$, then

$$\operatorname{tr}(EP)\operatorname{tr}(E^\dagger P)=q^{2n}/|S|^2.$$

Thus, A_j is $q^{2n}/|S|^2$ times the number of elements in the stabilizer subgroup S that have weight j.

We leave the result for B_j as an exercise. ∎

The geometrical interpretation of A_j for stabilizer codes is as follows. Suppose that $\mathcal{X}$ is a quantum set of lines in $\mathrm{PG}(n-k-1,q)$. Then A_j is $(q-1)$ times number of hyperplanes containing $n-j$ lines of $\mathcal{X}$.

The *quantum MacWilliams identity* states that

$$q^n B(x,y)=A(x+(q^2-1)y,x-y)$$

and, respectively, that

$$q^n A(x,y)=B(x+(q^2-1)y,x-y).$$

Before proving the quantum MacWilliams identity, consider the next example.

Example 7.2 (Self-dual hexacode). Consider the $[6,3,4]_4$ code D generated by the matrix

$$\begin{pmatrix}1&0&0&1&1&1\\0&1&0&1&e&e^2\\0&0&1&1&e^2&e\end{pmatrix},$$

where $e^2=e+1$. One can prove that the minimum distance is 4 by checking that all 3×3 submatrices are non-singular. By verifying that the Hermitian inner product (5.3)

between any two rows is zero, one quickly concludes that $D = D^{\perp_h}$. Theorem 5.7 implies that we can construct an $[\![6, 0, 4]\!]_2$ stabilizer code $Q(S)$ from D. By writing out the entries in the matrix over $\mathbb{F}_2$ and considering the $\mathbb{F}_2$ span, we obtain the matrix $\mathrm{G}(S)$ for this quantum code.

Consider the $[\![6, 0, 4]\!]_2$ code that can be constructed from the code D. The code $\tau(S)$ is spanned by the generator matrix

$$\mathrm{G}(S) = \left(\begin{array}{cccccc|cccccc} 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \end{array}\right).$$

Thus, the stabilizer subgroup has generators

$$\begin{aligned} M_1 &= X\ 1\ 1\ X\ X\ X, \\ M_2 &= Z\ 1\ 1\ Z\ Z\ Z, \\ M_3 &= 1\ X\ 1\ X\ Z\ Y, \\ M_4 &= 1\ Z\ 1\ Z\ Y\ X, \\ M_5 &= 1\ 1\ X\ X\ Y\ Z, \\ M_6 &= 1\ 1\ Z\ Z\ X\ Y. \end{aligned}$$

By Lemma 5.9, the quantum set of six lines $\mathcal{X}$ we get from the matrix $\mathrm{G}(S)$ has the property that any three lines of $\mathcal{X}$ span the whole space $\mathrm{PG}(5, 2)$. Therefore, any two span a three-dimensional subspace which is contained in three hyperplanes which contain no further line of $\mathcal{X}$. Thus, there are 45 hyperplanes which contain exactly two lines of $\mathcal{X}$. Let ℓ be a line of $\mathcal{X}$. There are 15 hyperplanes containing ℓ, so counting pairs (ℓ, π) where $\ell \in \mathcal{X}$ and π is a hyperplane containing ℓ, we conclude that any hyperplane containing a line of $\mathcal{X}$ contains two lines of $\mathcal{X}$.

Thus, we work out the weight distribution. For codes with $k = 0$ (that is, pure states), both weight distributions coincide; this can be checked from the definition. From before, we have that A_j is the $(q-1)$ times number of hyperplanes containing $n - j$ lines of $\mathcal{X}$. Thus, we have proved that the weight distribution for the quantum hexacode is

$$(A_0, \ldots, A_6) = (1, 0, 0, 0, 45, 0, 18).$$

The corresponding enumerator polynomials are

$$A(x, y) = B(x, y) = x^6 + 45x^2y^4 + 18y^6.$$

This polynomial is indeed invariant under the quantum MacWilliams transform since

$$\begin{aligned} 64B(x,y) &= (x+3y)^6 + 45(x+3y)^2(x-y)^4 + 18(x-y)^6 \\ &= 64(x^6 + 45x^2y^4 + 18y^6). \end{aligned}$$

Research Problem 9. For stabilizer codes, A_j and B_j count the number of terms in the stabilizer S and its normaliser $N(S)$, respectively; there is no such combinatorial interpretation for general quantum codes. Although A_j can interpreted as the Hilbert–Schmidt norms of the j-body correlations that appear in the code, we would like to determine what objects B_j is counting for non-stabilizer codes.

We return to the proof of the quantum MacWilliams identity.

Quantum MacWilliams identity. We will only state a proof sketch; the rather tedious combinatorial details can be found in [12, 17].

Let S be a collection of subsystems, and denote by tr_S the partial trace the systems in S. Denote by S^c the complement of S in $\{1,\dots,n\}$. Consider now how the partial trace tr_S followed by a "padding" with the identity acts on an operator P:

$$\begin{aligned} \mathrm{tr}_S(P) \otimes \mathbb{1}_S &= \mathrm{tr}_S\left(\frac{1}{q^n}\sum_E \mathrm{tr}(E^\dagger P)E\right) \otimes \mathbb{1}_S \\ &= \frac{1}{q^{n-|S|}} \sum_{\mathrm{supp}(E)\subseteq S^c} \mathrm{tr}(E^\dagger P)E. \end{aligned} \tag{7.1}$$

It can be shown (cf. [12, Appendix A]) that this can also be written as

$$\begin{aligned} \mathrm{tr}_S(P) \otimes \mathbb{1}_S &= \int_{\substack{U(q^n) \text{ s.t.} \\ \mathrm{supp}(U)\subseteq S}} UPU^\dagger \, dU \\ &= \frac{1}{q^{|S|}} \sum_{\mathrm{supp}(E)\subseteq S} EPE^\dagger, \end{aligned} \tag{7.2}$$

where the integration is over the unitarily invariant Haar measure of unitary matrices that act trivially on the subsystem S^c. The second equality follows from the fact that any complete orthonormal matrix basis $\{E_\alpha\}$ containing the identity forms a unitary 1-design.[5]

The quantum MacWilliams identity now essentially follows from equating equations (7.1) and (7.2), summing over all subsystems of size $|S| = m$, multiplying

[5] t-designs replace the integration over some compact group by a finite sum. A unitary t-design is a set of unitaries U_i, $i = 1,\dots,K$, acting on $\mathbb{C}^q$ such that $\int_{U(D)} P_{t,t}(U)\, dU = \frac{1}{K}\sum_{i=1}^K P_{t,t}(U_i)$ holds for every homogeneous polynomial $P_{t,t}$ that has degree t in the matrix elements of U and degree t in the matrix elements of U^*.

by P and taking the trace. This yields terms of the form $\sum \operatorname{tr}(E^\dagger P)\operatorname{tr}(EP)$ and $\sum \operatorname{tr}(E^\dagger PEP)$, corresponding to the two types of weights A_j and B_j.

Proceeding in this manner, equation (7.1) gives

$$\begin{aligned}\sum_{|S|=m} \operatorname{tr}(\operatorname{tr}_S(P)\otimes \mathbb{1}_S \cdot P) &= \sum_{|S|=m} \operatorname{tr}\left(q^{m-n} \sum_{\operatorname{supp}(E)\subseteq S^c} \operatorname{tr}(E^\dagger P)E\cdot P\right)\\ &= q^{m-n}\sum_{|S|=m}\sum_{\operatorname{supp}(E)\subseteq S^c} \operatorname{tr}(E^\dagger P)\,(EP)\\ &= q^{m-n}\sum_{j=0}^{n-m}\binom{n}{n-m}\binom{n-m}{j}\binom{n}{j}^{-1} A_j\\ &= q^{m-n}\sum_{j=0}^{n-m}\binom{n-j}{m} A_j.\end{aligned}$$

Meanwhile, equations (7.2) gives

$$\begin{aligned}\sum_{|S|=m} \operatorname{tr}(\operatorname{tr}_S(P)\otimes \mathbb{1}_S \cdot P) &= \sum_{|S|=m} \operatorname{tr}\left(q^{-m} \sum_{\operatorname{supp}(E)\subseteq S} E^\dagger PE\cdot P\right)\\ &= q^{-m}\sum_{|S|=m}\sum_{\operatorname{supp}(E)\subseteq S} \operatorname{tr}(E^\dagger PEP)\\ &= q^{-m}\sum_{j=0}^{m}\binom{n}{m}\binom{m}{j}\binom{n}{j}^{-1} B_j\\ &= q^{-m}\sum_{j=0}^{m}\binom{n-j}{n-m} B_j.\end{aligned}$$

Thus for every operator P and $0 \le m \le n$ one has that

$$q^{m-n}\sum_{j=0}^{n-m}\binom{n-j}{m} A_j = q^{-m}\sum_{j=0}^{m}\binom{n-j}{n-m} B_j.$$

Using generating functions, in other words, the weight enumerator polynomials $A(x,y)$ and $B(x,y)$, as well as Krawtchouk polynomials, we obtain the MacWilliams identity

$$q^n B(x,y) = A(x+(q^2-1)y, x-y).$$

This ends the proof sketch. ∎

The enumerators and their weights have a couple of interesting properties. Let $K = \dim(\operatorname{im} P)$.

(a) The weights A_j and B_j are invariant under the local choice of basis and are so-called local unitary invariants (LU-invariants). That is,

$$A_j(P) = A_j(P') \quad \text{and} \quad B_j(P) = B_j(P'),$$

where $P' = (U_1 \otimes \cdots \otimes U_n)P(U_1^\dagger \otimes \cdots \otimes U_n^\dagger)$, and $U_1, \ldots, U_n$ are unitary $q \times q$ matrices.

(b) $A_0 = \dim(P)$ and $KB_j \geq A_j \geq 0$.

(c) A projection operator P with $K = \dim(\mathrm{im}(P))$ is a code of distance d if and only if it satisfies $KB_j = A_j$ for $0 \leq j < d$.

(d) One can check that for codes with $K = 1$, the enumerator polynomial is invariant under the quantum MacWilliams transform, and one has $B(x, y) = A(x, y)$. When such a code is of stabilizer type, it corresponds to a classical self-dual code.

Some comments are in order. The weights must be LU-invariant – the properties of the code should not depend on the way one sets up the local coordinate system for each spin particle. The last two properties are useful to obtain weights of hypothetical codes and to apply the machinery of linear programming bounds [2]. That is, one sets up a system of linear equalities and inequalities in the variables $A_0, \ldots, A_n$ making use of (a), (b) and the quantum MacWilliams identity.

For example, it is a longstanding open problem if a (pure) code with the parameters $[\![24, 0, 10]\!]_2$ exists. It is known that such code must have even weights only, and using linear programming, one can fix the weight distribution to be

$$[A_{10}, A_{12}, A_{14}, \ldots, A_{24}] = [18216, 156492, 1147608, 3736557, 6248088, 4399164, 1038312, 32778].$$

Indeed this is also the weight distribution of hypothetical $[24, 12, 10]$ self-dual additive code over GF(4) (see OEIS http://oeis.org/A030331).

Research Problem 10. Either find a quantum code with parameters $[\![24, 0, 10]\!]_2$, or show that no such code can exist.

We refer to the tables by M. Grassl [10] for more existence results.

Funding. The first author acknowledges the support of the Spanish Ministry of Science and Innovation grants MTM2017-82166-P and PID2020-113082GB-I00. The third author acknowledges the support of the Spanish MINECO (Severo Ochoa SEV-2015-0522), Fundació Cellex and Mir-Puig, Generalitat de Catalunya (SGR 1381 and CERCA Programme), and the European Union under Horizon2020 (PROBIST 754510).

References

[1] S. Aaronson, *Quantum computing since Democritus*. Cambridge University Press, Cambridge, 2013 Zbl 1353.68003 MR 3058839

[2] A. Ashikhmin and S. Litsyn, Upper bounds on the size of quantum codes. In *Proceedings of IEEE International Symposium on Information Theory (Cambridge, MA, USA)*, pp. 351–371, IEEE, 1998

[3] S. Ball, On sets of vectors of a finite vector space in which every subset of basis size is a basis. *J. Eur. Math. Soc. (JEMS)* **14** (2012), no. 3, 733–748 Zbl 1241.15002 MR 2911882

[4] S. Ball and P. Puig, The geometry of non-additive stabiliser codes. 2021, arXiv:2107.11281

[5] T. A. Brun and D. E. Lidar, *Quantum error correction*. Cambridge University Press, Cambridge, 2013

[6] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, Quantum error correction via codes over GF(4). *IEEE Trans. Inform. Theory* **44** (1998), no. 4, 1369–1387 Zbl 0982.94029 MR 1665774

[7] D. G. Glynn, T. A. Gulliver, J. G. Maks, and M. K. Gupta, The geometry of additive quantum codes. Unpublished manuscript, 2004, available at https://www.academia.edu/17980449/

[8] D. Gottesman, *Stabilizer codes and quantum error correction*. Ph.D. thesis, California Institute of Technology, 1997, arXiv:quant-ph/9705052

[9] D. Gottesman, An introduction to quantum error correction and fault-tolerant quantum computation. In *Quantum information science and its contributions to mathematics*, pp. 13–58, Proc. Sympos. Appl. Math. 68, American Mathematical Society, Providence, RI, 2010 Zbl 1211.81043 MR 2762145

[10] M. Grassl, Bounds on the minimum distance of linear codes and quantum codes. Available at http://www.codetables.de

[11] S. Haroche and J.-M. Raimond, *Exploring the quantum: atoms, cavities and photons*. Oxf. Grad. Texts, Oxford University Press, Oxford, 2006 Zbl 1118.81001 MR 2271425

[12] F. Huber, C. Eltschka, J. Siewert, and O. Gühne, Bounds on absolutely maximally entangled states from shadow inequalities, and the quantum MacWilliams identity. *J. Phys. A* **51** (2018), no. 17, paper no. 175301 Zbl 1390.81074 MR 3787259

[13] F. Huber and M. Grassl, Quantum codes of maximal distance and highly entangled subspaces. *Quantum* **4** (2020), paper no. 284

[14] A. Ketkar, A. Klappenecker, S. Kumar, and P. K. Sarvepalli, Nonbinary stabilizer codes over finite fields. *IEEE Trans. Inform. Theory* **52** (2006), no. 11, 4892–4914 Zbl 1242.94045 MR 2300363

[15] A. Matuschak and M. A. Nielsen, Quantum computing for the very curious. San Francisco, 2019, available at https://quantum.country/qcvc

[16] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*. Cambridge University Press, Cambridge, 2000 Zbl 1049.81015 MR 1796805

[17] E. M. Rains, Quantum weight enumerators. *IEEE Trans. Inform. Theory* **44** (1998), no. 4, 1388–1394 Zbl 0982.94030 MR 1665778

[18] E. M. Rains, R. H. Hardin, P. W. Shor, and N. J. A. Sloane, A nonadditive quantum code. *Phys. Rev. Lett.* **79** (1997), no. 4, 953–954

[19] J. J. Sakurai, *Modern quantum mechanics*. Addison-Weyley, Reading, 1994

[20] P. W. Shor, Scheme for reducing decoherence in quantum memory. *Phys. Rev. A* **52** (1995), no. 4, R2493–R2496

Communicated by Gil Kalai

Received 7 July 2020.

Simeon Ball
Departament de Matemàtiques, Universitat Politècnica de Catalunya, Mòdul C3,
Campus Nord, Carrer Jordi Girona 1-3, 08034 Barcelona, Spain; simeon@ma4.upc.edu

Aina Centelles
Facultat de Matemàtiques, Universitat Politècnica de Catalunya, Carrer de Pau Gargallo 14,
08028 Barcelona, Spain; aina.centelles@estudiant.upc.edu

Felix Huber
ICFO – The Institute of Photonic Sciences, Mediterranean Technology Park,
Avinguda Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain;
current address: Faculty of Physics, Astronomy and Applied Computer Science,
Institute of Theoretical Physics, Jagiellonian University, ul. Łojasiewicza 11, 30-348 Kraków,
Poland; felix.huber@uj.edu.pl