Convolutional Codes: MacWilliams Identity and Hulls

A Thesis

submitted to
Indian Institute of Science Education and Research Pune
in partial fulfillment of the requirements for the
BS-MS Dual Degree Programme

by

Prateek Dwivedi

Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.

April, 2025

Supervisor: Prof. Julia Lieb, Prof. Krishna Kaipa

Certificate

This is to certify that this dissertation entitled 'Convolutional Codes: MacWilliams Identity and Hulls' towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Prateek Dwivedi at Indian Institute of Science Education and Research under the supervision of Prof. Julia Lieb, Junior Professor, Fakultät Mathematik und Naturwissenschaften, TU Ilmenau, and Prof. Krishna Kaipa, Professor, Department of Mathematics, IISER Pune, during the academic year 2024-2025.

Julia Lieb

Prof. Julia Lieb

K Vinod

Prof. Krishna Kaipa

Vivek Mohan Mallick

Prof. Vivek Mohan Mallick

Committee:

Prof. Julia Lieb

Prof. Krishna Kaipa

Prof. Vivek Mohan Mallick

Declaration

I hereby declare that the matter embodied in the report entitled 'Convolutional Codes: MacWilliams Identity and Hulls' are the results of the work carried out by me at the Fakultät Mathematik und Naturwissenschaften, Technische Universität Ilmenau and Indian Institute of Science Education and Research, Pune, under the supervision of Prof. Julia Lieb and Prof. Krishna Kaipa , and the same has not been submitted elsewhere for any other degree.

Prateek Dwivedi

Acknowledgments

I would like to take this opportunity to thank my supervisors, Prof. Kaipa and Prof. Lieb. Especially Prof. Kaipa, who introduced me to the field of coding theory and has been a mentor for the past 2 years. This would not be possible without them. I would also like to thank my parents, Jyoti and Atulya Dwivedi, for providing unconditional support at all times and my friends, Rik Sarkar, Shreya Kulkarni, Jeevan Kumar Papaganthi, and Khushi Datri, for uplifting me during the bleakest of times.

Abstract

This thesis provides a succinct summary of some of the research work in the field of Coding Theory with regard to convolutional coding. Specifically, it explores the MacWilliams Identity and its derivation for the case of convolutional coding. It also extends the idea of LCD codes to the convolutional case while verifying important results that were known to hold for block codes in the convolutional setting.

Contents

Chapter 1

Preliminaries

Coding theory is the study of error-correcting codes. These are methods of encoding data for transmission such that the original message can be decoded despite noise from the medium of transmission. One adds redundancy to the data in such a way that errors induced by noise from the channel of transmission have little to no effect on the data received post-decoding of the transmitted message. This process is known as encoding. The encoded data is then transmitted through a noisy channel, which induces errors in the transmitted data. This corrupt message is then received, and the best possible estimate of the original data is recovered through a process called decoding.
Let F be a set of cardinality q called the alphabet. Words formed by the combination of elements of the alphabet of length m are assumed to be "vectors" of length m. The coordinates are also sometimes called bits. Assume that all messages are words of length k. Thus, the set of messages can be treated as a subset of F^k, let it be denoted by $\mathcal{M}$.

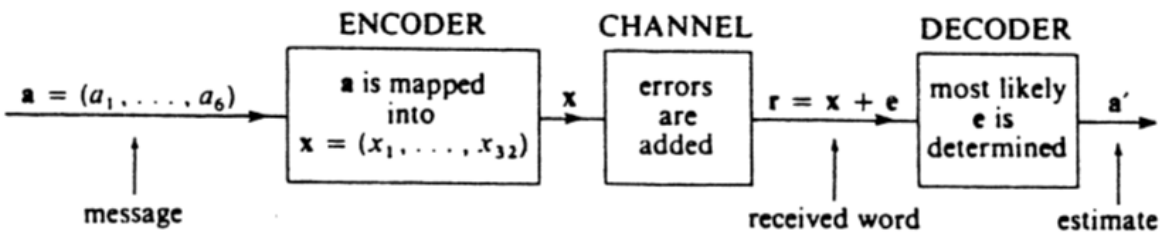

Figure 1

Figure 1.1: This figure is from [7]

1.1 Block Codes

The code $\mathscr{C}$ is a subset of F^n, i.e. each codeword is a vector of length n. Every message is individually mapped to a codeword, also known as a block. The mapping is done through an injective function called the encoder, $\iota : \mathcal{M} \to \mathscr{C}$. Assume that the size of the code, i.e. $\#\mathscr{C}$, is M, then the code is known as a $(n, M)_q$-code.

1.1.1 Weights and Distances in F^m

In coding theory, the most commonly used notion of weight of a vector is the Hamming weight, which is the number of non-zero coordinates of a vector. The Hamming distance of two vectors is then the number of coordinates where they do not coincide. It is easy to see that this notion of distance is a metric if F is a field $\mathbb{F}$:

1. for any $x, y \in \mathbb{F}^m$, $d(x, y) \geq 0$ holds because the number of coordinates can not be negative. $d(x, y) = 0$ implies that x and y differ in 0 coordinates, i.e. $x = y$. On the other hand, $x = y$ implies that they do not differ at any coordinate, i.e. $d(x, y) = 0$.
2. if $d(x, y) = d$ then x and y differ at d coordinates which also implies $d(y, x) = d$.
   Therefore, $d(x, y) = d(y, x)$.
3. for any $x, y, z \in \mathbb{F}^m$, let $D(x, y) = \{i | x_i \neq y_i\}$, where x_i is the i^{th} coordinate of x.
   Therefore, $|D(x, y)| = d(x, y)$. $D(x, y) = D(x, z) \Delta D(z, y) \subset D(x, z) \cup D(z, y)$, thus $d(x, y) \leq d(x, z) + d(z, y)$, satisfying the triangle inequality.

For any set S, we define the weight enumerator polynomial as:

$$\mathrm{we}_S(z) = \sum_{i \geq 0} \#\{u \in S | \mathrm{w}_H(u) = i\} z^i$$

Where w_H is the Hamming weight.
If for an $(n, M)_q$-code $\mathscr{C}$, the least distance between a pair of non-equal codewords is denoted by d, then we say that the code $\mathscr{C}$ is an $(n, M, d)_q$-code.

1.1.2 Linear Codes

A linear code is a code that has a subspace structure to it. In this case, the code is a k-dimensional subspace of $\mathbb{F}_q^n$. For linear codes, we can call a $(n, M)_q$ code a $(n, k)_q$ code where $M = q^k$ since $\mathscr{C}$

is a k-dimensional space over $\mathbb{F}_q$.
For a linear code, the least distance between a pair of non-equal codewords coincides with the least weight of non-zero codewords of $\mathscr{C}$. This is because for any pair of codewords c_1 and c_2, $c := c_1 - c_2$ is also a codeword. Assume that the least distance for this code is d. We can then call the code $\mathscr{C}$ an $(n, k, d)_q$-code. We can drop q from the above notation when the base field is obvious.
Since the code is a k-dimensional space over $\mathbb{F}_q$, we have a k element basis for the code $\mathscr{C}$. Let it be $\{g_i\}_{i=1}^k$. Consider the $k \times n$ matrix

$$G = \begin{pmatrix} -g_1- \\ \vdots \\ -g_k- \end{pmatrix}$$

It is immediate that the code is the $\mathbb{F}_q$ rowspace of the matrix G. Any such matrix is known as a generator matrix for the code $\mathscr{C}$, and the encoding function can be seen as matrix multiplication with G. Since the code is k-dimensional, $\text{rank}(G) = k$ for any generator matrix G of the code $\mathscr{C}$.
Any matrix L of rank $n - k$ such that for any codeword c, $cL = 0$ is known as a parity check matrix for the code $\mathscr{C}$.

1.1.3 Dual of a Code

On $\mathbb{F}_q^n$, we define the bilinear form

$$(\cdot, \cdot) : \mathbb{F}_q^n \times \mathbb{F}_q^n \to \mathbb{F}_q : (x = (x_1, ..., x_n), y = (y_1, ..., y_n)) \mapsto \sum_{i=1}^{n} x_i y_i$$

Since we are working in a field, this form is symmetric. It is also bilinear, which is evident to see. We are now ready to define the dual of a code.

Definition 1.1.1. *For a code $\mathscr{C}$, its dual, $\mathscr{C}^\perp$, is defined as*

$$\mathscr{C}^\perp = \{u \in \mathbb{F}_q^n | (u, c) = 0 \forall c \in \mathscr{C}\}$$

If $\mathscr{C}$ is a (n, k, d)-linear code, then its dual $\mathscr{C}^\perp$ is a $(n, n-k, d)$-linear code, generated by the parity check matrix of $\mathscr{C}$. [7]
In the finite case, it is not necessary that the dual of a space and the space are orthogonal complements. Thus, their intersection is not necessarily empty. It is possible to have a vector that is orthogonal to itself; consider the field $\mathbb{F}_3$ and the vector $(1, 1, 1)$ in $\mathbb{F}_3^3$. Clearly, $(1, 1, 1)(1, 1, 1)^T = 3 = 0$.

Thus, given a code $\mathscr{C}$, we define the following [9]:

Definition 1.1.2. *The hull of a code is the space* $\mathscr{C} \cap \mathscr{C}^{\perp}$.

1.1.4 MacWilliams Identity

The MacWilliams Identity, discovered by FJ MacWilliams in 1963 [8], is a linear transformation that relates the weight enumerator of a code to its dual. Given a linear code $\mathscr{C}$ and its dual $\mathscr{C}^{\perp}$, we have

$$\mathrm{we}_{\mathscr{C}^{\perp}}(z) = \frac{(1+(q-1)z)^n}{q^k}\mathrm{we}_{\mathscr{C}}\left(\frac{1-z}{1+(q-1)z}\right)$$

Which can then be written as $\mathrm{we}_{\mathscr{C}^{\perp}}(z) = q^{-k}H(\mathrm{we}_{\mathscr{C}}(z))$, where H is defined as follows:

$$H : \mathbb{C}[z] \to \mathbb{C}[z] : h(z) \mapsto (1+(q-1)z)^n h\left(\frac{1-z}{1+(q-1)z}\right) \tag{1.1}$$

When $\mathscr{C}$ is a non-linear code, there is no concept of a dual and the weight distribution, denoted by $\mathrm{we}'_{\mathscr{C}}(z)$, is derived as follows:
Define $A_i(x;\mathscr{C}) := \#\{c \in \mathscr{C} | d(x,c) = i\}$, and $A_i(\mathscr{C}) := \frac{1}{|\mathcal{M}|}\sum_{x\in\mathscr{C}} A_i(x;\mathscr{C})$. Then

$$\mathrm{we}'_{\mathscr{C}}(z) = \sum_{i=0}^{n} A_i(\mathscr{C})z^i$$

The above weight enumerator and the usual definition coincide when applied to a linear code (or any "distance-invariant" code).
When applying the MacWilliams transfer to the weight enumerator for a non-linear code, we get some interesting combinatorial properties, in spite of the dual not being well defined for it.

1.1.5 Linear Code with a Complementary Dual (LCD Code)

The code and its dual are not to be seen as orthogonal complements [7]. The linear codes whose duals are indeed orthogonal complements are known as Linear Codes with Complementary Duals, or LCD codes for short.[9]

Definition 1.1.3. *Let* $\mathscr{C}$ *be a linear code and* $\mathscr{C}^{\perp}$ *be its dual.* $\mathscr{C}$ *is an LCD code if and only if* $\mathscr{C} \cap \mathscr{C}^{\perp} = \{0\}$.

Another way of characterising LCD codes in terms of their generator matrix is (from [9]):

Theorem 1.1.1. *Let G be a generator matrix for the code $\mathscr{C}$. Then $\mathscr{C}$ is an LCD Code if and only if the matrix GG^T is invertible.*
Moreover, the linear operator $\Pi_{\mathscr{C}} = G^T(GG^T)^{-1}G$ is the orthogonal projection from $\mathbb{F}^n$ to $\mathscr{C}$ for an LCD code $\mathscr{C}$.

Proof. If we assume that GG^T is non-invertible, then there exists a non-zero $u \in \mathbb{F}^k$ such that $uGG^T = 0$. Since u is non-zero, uG is a non-zero codeword. Any non-zero v in $\mathscr{C}$ can be written as $u'G$ for some non-zero vector u' in $\mathbb{F}^k$. Thus,

$$\begin{aligned} uGv^T &= uG(u'G)^T \\ &= uGG^Tu' \\ &= 0 \end{aligned}$$

Therefore, uG lies in $\mathscr{C}^\perp \cap \mathscr{C}$ and is non-zero. Hence, $\mathscr{C}$ is not an LCD code.
If we conversely assume that GG^T is invertible, then consider the linear transform

$$T : \mathbb{F}^n \to \mathbb{F}^n; u \mapsto uG^T(GG^T)^{-1}G$$

For any non-zero codeword, $v = uG$,

$$Tv = uGG^T(GG^T)^{-1}G = uG = v$$

Moreover for any vector v in $\mathscr{C}^\perp$,

$$Tv = vG^T(GG^T)^{-1}G = 0$$

Any vector v in $\mathbb{F}^n$ goes to $\mathscr{C}$ as $vG^T(GG^T)^{-1}$ is a vector in $\mathbb{F}^k$. Hence, T is an orthogonal projection for $\mathscr{C}$, and $\mathscr{C}$ is an LCD code. □

1.2 Convolutional Codes

These differ from standard block codes in the sense that the encoder for a convolutional code takes a sequence of messages $u = (u(0), u(1), ...)$, each $u(i)$ from the set of messages $\mathcal{M} \subset \mathbb{F}^k$ to a sequence of codewords $x = (x(0), x(1), ...)$ each $x(i) \in \mathbb{F}^n$, but the i^{th} codeword depends on previous codewords as well. This is done through a state vector $s(i)$, of dimension m. We can describe the encoder as follows: $s(0) = 0$ and $\forall i \geq 0$

$$s(i+1) = s(i)\mathcal{A} + u(i)\mathcal{B}$$

$$x(i) = s(i)\mathcal{C} + u(i)\mathcal{D}$$

Where $\mathcal{A} \in \mathbb{F}^{m\times m}$, $\mathcal{B} \in \mathbb{F}^{k\times m}$, $\mathcal{C} \in \mathbb{F}^{m\times n}$, and $\mathcal{D} \in \mathbb{F}^{k\times n}$. This is known as the State Space description of a convolutional code. One then takes the code to be the set of all possible outputs x. One could also take the code to be the set of generating functions. Given a sequence, $(a_0, a_1, ...)$, the generating function associated with the given sequence is the power series $\sum_i a_i z^i$ where z is an indeterminate, i.e. a variable.
In terms of the generating functions, $S(z) = \sum_{i\geq 0} s(i)z^i$, $U(z) = \sum_{i\geq 0} u(i)z^i$, and $X(z) = \sum_{i\geq 0} x(i)z^i$, the state space description is

$$z^{-1}S(z) = \sum_{i\geq 0} s(i+1)z^i + 0 = S(z)\mathcal{A} + U(z)\mathcal{B}$$

$$X(z) = S(z)\mathcal{C} + U(z)\mathcal{D}$$

Solving these equations,

$$\begin{aligned} z^{-1}S(z) - S(z)\mathcal{A} &= U(z)\mathcal{B} \\ \implies S(z) = U(z)\left(\mathcal{B}(z^{-1}I - \mathcal{A})^{-1}\right) &= U(z)E(z) \end{aligned}$$

and

$$\begin{aligned} X(z) &= U(z)E(z)\mathcal{C} + U(z)\mathcal{D} \\ \implies X(z) = U(z)\left(E(z)\mathcal{C} + \mathcal{D}\right) &= U(z)G(z) \end{aligned}$$

Thus, by enlarging the possible sequence of messages to come from $\mathbb{F}((z))$, it is easy to see that the convolutional code associated with $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is just the $\mathbb{F}((z))$-rowspace of $G(z)$, which from the preceding equation has rational functions as entries. $G(z)$ is therefore known as the generator matrix of the code. Any $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ that satisfies the above is known as a realisation of the generator matrix G. This means that the basis of $\mathscr{C}$ is entirely in $\mathbb{F}(z)$, which is a subfield of $\mathbb{F}((z))$; thus we can just consider the rational subcode, i.e. $\mathscr{C} = \mathscr{C} \cap \mathbb{F}(z)$ without giving up too much.

Definition 1.2.1 (Convolutional Code)**.** *An (n, k) convolutional code is a k-dimensional subspace of $\mathbb{F}(z)^n$, where $\mathbb{F}(z)$ is the field of Rational functions with coefficients in $\mathbb{F}$.*

Certain definitions consider convolutional codes as submodules of n-dimensional modules over the ring $\mathbb{F}_q[z]$. This treatment of convolutional codes is used prominently in Chapter 2.

1.2.1 Polynomial Generator Matrices (PGM)

A polynomial generator matrix is a generator matrix for $\mathscr{C}$ with all entries in $\mathbb{F}[z]$. All codes have a polynomial generator matrix: consider $G(z)$ to be a generator matrix with entries in $\mathbb{F}(z)$, then one can multiply the entire matrix by the lcm of the denominators of the entries and produce a matrix $G'(z)$ with entries in $\mathbb{F}[z]$. The row spaces of the two matrices are the same, as they differ only by a scalar factor. Thus, they generate the same code.
For any row vector $a(z) = (a_1(z), ..., a_m(z)) \in \mathbb{F}[z]^m$ for any m, we define the degree of $a(z)$ as

$$\deg(a(z)) = \max_i \{\deg(a_i(z))\}.$$

Let $G(z)$ be a $k \times n$ PGM for $\mathscr{C}$, with rows $g_i(z) \in \mathbb{F}[z]^n$. Let $\deg(g_i(z)) = h_i$, then we define the external degree of $G(z)$ as

$$\text{extdeg}(G(z)) = \sum_{i=1}^{k} h_i$$

and the internal degree of $G(z)$ as

$$\text{intdeg}(G(z)) = \max\{\deg(\text{full size minors of } G(z))\}$$

Clearly, for any polynomial matrix $G(z)$, $\text{intdeg}(G(z)) \leq \text{extdeg}((G(z))$, as any $k \times k$ minor will select only one entry from each row and column and the degree of this entry is $\leq h_i$.

The following result proves to be instrumental in deriving multiple results about PGMs:

Lemma 1.2.1. *[10] Let G(z) be a $k \times n$ polynomial matrix and T(z) be any $k \times k$ non-singular polynomial matrix, then*

$$intdeg(T(z)G(z)) = \quad intdeg(T(z)) + \quad intdeg(G(z))$$

Thus, $intdeg(T(z)G(z) \geq \ intdeg(T(z))$, with equality if and only if $intdeg(T(z)) = 0$, which implies that $\Delta(T(z)) \in (\mathbb{F}[z])^\times = \mathbb{F} \setminus \{0\}$ ($T(z)$ is unimodular).

Proof. The set of $k \times k$ minors of $T(z)G(z)$ is exactly the set of $k \times k$ minors of $G(z)$ each multiplied by $\Delta(T(z))$. Thus, the degree of each full-sized minor of $T(z)G(z)$ is the degree of the corresponding full-sized minor of $G(z)+$ intdeg$(T(z))$. Therefore, intdeg$(T(z)G(z)) =$ intdeg$(G(z))+$ intdeg$(T(z))$. The second part follows trivially. □

Definition 1.2.2 (Basic Matrices)**.** *A $k \times n$ polynomial matrix $G(z)$ is known as* basic *if it satisfies any of the following equivalent conditions [10]:*

1. *Among all polynomial generator matrices of the form $T(z)G(z)$, where $T(z)$ is a non-singular $k \times k$ matrix, $G(z)$ has minimal internal degree*
2. *The Smith Normal Form of $G(z)$ is $\begin{pmatrix} \mathbb{I}_{k\times k} & 0 \end{pmatrix}$*
3. *The gcd of the $k \times k$ minors is 1*
4. *$G(\alpha)$ has full rank $\forall \alpha \in \bar{\mathbb{F}}$*
5. *$G(z)$ has a right inverse in $\mathbb{F}[z]^{n\times k}$*
6. *If $x(z) = u(z)G(z)$ is in $\mathbb{F}[z]^n$, then $u(z) \in \mathbb{F}[z]^k$*
7. *$G(z)$ is the submatrix of an unimodular matrix.*

Definition 1.2.3. *$[G(z)]_{hr}$ is the matrix in which each row has the coefficients of the highest degree term in that row, i.e.*

$$[G(z)]_{hr\,i,j} = coeff_{D^{h_i}} G(z)_{i,j}$$

Definition 1.2.4 (Reduced Matrices)**.** *A $k \times n$ matrix is said to be reduced if any of the following hold [10]*

1. *Among all PGMs of the form $T(z)G(z)$, where $T(z)$ is a unimodular $k \times k$ matrix, $G(z)$ has minimal external degree*
2. *$[G]_{hr}$ has rank k*
3. *$extdeg(G(z)) = intdeg(G(z))$*
4. *(Predictable Degree Property). For any polynomial vector $u(z) \in \mathbb{F}[z]^k$,*

$$\deg\left(u(z)G(z)\right) = max_i\{\deg(u_i(z)) + h_i\}$$

Definition 1.2.5 (Canonical PGM)**.** *Among all PGMs for a convolutional code $\mathscr{C}$, a canonical PGM is one with a minimal external degree.*

Definition 1.2.6 (Degree of a Code)**.** *The degree of a code $\mathscr{C}$ is the minimal external degree of all PGMs of the code, denoted by $deg(\mathscr{C})$.*

Theorem 1.2.2. *[10] Let G(z) be a PGM for $\mathscr{C}$; G(z) is a canonical matrix if and only if it is both reduced and basic.*

Proof. ($\implies$) Let $G_0(z)$ be a basic PGM for $\mathscr{C}$ such that among all basic PGMs, it has minimal external degree. Then,

$$\text{intdeg}(G_0(z)) \leq \text{ intdeg}(G(z)) \leq \text{ extdeg}(G(z)) \leq \text{ extdeg}(G_0(z)) \tag{1.2}$$

For any unimodular $k \times k$ matrix $T(z)$, we have

$$\text{intdeg}(G_0(z)) = \text{ intdeg}(T(z)G_0(z))$$

i.e. $T(z)G_0(z)$ is another basic PGM. Thus, by our assumption about $G_0(z)$, we have

$$\begin{aligned} \text{extdeg}(G_0(z)) &\leq \text{ extdeg}(T(z)G_0(z)) \\ &\implies G_0(z) \text{ is reduced} \\ \implies \text{ intdeg}(G_0(z)) &= \text{ extdeg}(G_0(z)) \end{aligned}$$

Combining this and 1.2, we get

$$\text{intdeg}(G_0(z)) = \text{ intdeg}(G(z)) = \text{ extdeg}(G(z))$$

$\implies G(z)$ is basic and reduced.
($\impliedby$) Let $G(z)$ be a basic and reduced PGM for $\mathscr{C}$, then for any PGM for $\mathscr{C}$, $G_0(z)$,

$$\text{extdeg}(G(z)) = \text{ intdeg}(G(z)) \leq \text{ intdeg}(G_0(z)) \leq \text{ extdeg}(G_0(z))$$

thus $G(z)$ has minimal external degree among all PGMs for $\mathscr{C}$, i.e. $G(z)$ is a canonical PGM for $\mathscr{C}$. □

Corollary 1.2.3. *[10] If $G(z)$ is a basic PGM for $\mathscr{C}$, then $deg(\mathscr{C}) = intdeg(G(z))$, which is also the minimal internal degree among all PGMs of $\mathscr{C}$.*

Theorem 1.2.4. *Let $h_1 \leq h_2 ... \leq h_k$ be the row degrees of a canonical $k \times n$ PGM of $\mathscr{C}$, $G(z)$, and let $F(z)$ be a $k \times n$ PGM of $\mathscr{C}$ with the row degrees $f_1 \leq f_2 ... \leq f_k$, then $h_i \leq f_i \forall i \in [k] = \{1, 2, ..., k\}$*

Proof. Assume the contrary. Then $\exists j \in [k]$ such that$f_j < h_j$ but $f_i \geq h_i \forall i < j$. Each row of $F(z)$ is in $\mathscr{C}$ and thus is a linear combination of rows of $G(z)$. Since $G(z)$ is a basic PGM (1.2.2) and

rows of $F(z)$ are polynomial vectors, the coefficients of the rows are also polynomials, and since $G(z)$ is a reduced PGM, by predictable degree property, the first $j+1$ rows of $F(z)$ have to be linear combinations of the first j rows of $G(z)$ which implies that they are not linearly independent, thus has rank $< k$, and thus cannot generate $\mathscr{C}$ which is a contradiction.
Therefore, $h_i \leq f_i \forall i \in [k]$ □

From this, it follows that any two canonical matrices have the same row degrees. These common row degrees are known as Forney indices. Throughout this paper, they are denoted by $\{\gamma_i\}_{i=1}^{k}$. The maximum row degree is known as the memory of the code, it is usually denoted by m. The sum of the Forney indices, $\sum \gamma_i = \gamma$, gives the external degree of a canonical matrix and, thus, the degree of the code. An (n, k) convolutional code with degree γ is called an (n, k, γ) convolutional code.
Given a PGM $G(z)$, one can write it as

$$G(z) = \sum_{i=0}^{m} G_i z^i$$

Where G_i are constant matrices. A generator matrix is known as delay-free if G_0 has full rank.

1.2.2 Duals

Analogously to the bilinear form defined on $\mathbb{F}_q^n$, we define the following bilinear form on $\mathbb{F}_q^n(z)$

$$(\cdot,\cdot) : \mathbb{F}_q(z)^n \times \mathbb{F}_q(z)^n \to \mathbb{F}_q(z) : (x(z) = (x_1(z), ..., x_n(z)), y(z) = (y_1(z), ..., y_n(z))) \mapsto \sum_{i=1}^{n} x_i(z) y_i(z)$$

We can then define the dual $\mathscr{C}^\perp$ of any code $\mathscr{C}$ as follows:

$$\mathscr{C}^\perp = \{y(z) \in \mathbb{F}_q(z)^n | (c(z), y(z)) = 0 \forall c(z) \in \mathscr{C}\}$$

If $\mathscr{C}$ is an (n, k) convolutional code, then $\mathscr{C}^\perp$ is an $(n, n-k)$ convolutional code.[10]

We can also define another bilinear form on $\mathbb{F}_q^n(z)$, which is related to the sequence interpretation of the codewords, as we see in the system theoretic description at the beginning of this section,

$$< \cdot, \cdot >: \mathbb{F}_q^n(z) \times \mathbb{F}_q^n(z) \to \mathbb{F}_q : \left(u = \sum_i u_i z^i, v = \sum_i v_i z^i \right) = \sum_i (u_i, v_i)$$

Where (u_i, v_i) is the bilinear form on $\mathbb{F}_q^n$ as defined in the section regarding block codes.
The Dual concerning this bilinear form, called the trellis dual, is then defined as the set

$$\mathscr{C}_t^{\perp} = \{u = \sum_i u_i z^i \in \mathbb{F}_q^n(z) | < u, c >= 0 \forall c \in \mathscr{C}\}$$

This treatment of the dual can be seen in [6].

1.2.3 Free Distance Bounds

Let $c(z)$ be a codeword; then we can write each component of $c(z)$ as a Laurent series and thereby write $c(z)$ as a vector Laurent series, $\sum c_i z^i$, $c_i \in \mathbb{F}^n$. We define the weight of $c(z)$ as the sum of Hamming weights of each c_i. Similar to the block code case, we define the distance between two code words c and c^* as $dist(c, c^*) = w(c - c^*)$. This concept of weight allows for vectors of infinite weight.
The free distance d of a convolutional code $\mathscr{C}$ is defined as the minimum distance between a pair of non-equal codewords. Due to the linear nature of $\mathscr{C}$, this is equivalent to the minimum weight of any non-zero codeword. If an (n, k, γ) code has free distance d, it is referred to as an (n, k, γ, d) code. One can also denote the free distance of a code $\mathscr{C}$ as $d_{\text{free}}(\mathscr{C})$.

Definition 1.2.7 (Optimal Code)**.** *An (n, k, γ) code that has maximal free distance among all (n, k, γ) codes is known as an optimal code.*

By polynomial codewords, we refer to $x(z) \in \mathscr{C}$ such that all entries of $x(z)$ are polynomials. For any $L \geq 0$, we define $\mathscr{C}_L$ to be the set of polynomial codewords of $\mathscr{C}$ with degree $\leq L$. This is a vector space over $\mathbb{F}$, and we set $\dim_{\mathbb{F}}(\mathscr{C}_L) = \delta_L$.

Theorem 1.2.5.

$$\delta_L = \sum_{i=1}^{k} \max(L + 1 - h_i, 0) \tag{1.3}$$

[10]

Proof. Let $G(z)$ be a canonical PGM for $\mathscr{C}$ with row indices $h_1 \leq h_2 ... \leq h_k$, and let its rows be $\{g_i(z)\}$. Let $x(z) \in \mathscr{C}_L$, then $x(z) = u(z)G(z)$. Since $G(z)$ is basic, $u(z)$ has all polynomial entries. By the predictable degree property (applies since $G(z)$ is reduced), we have $\max_i\{h_i + deg(u_i)\} \leq L \implies deg(u_i) + h_i \leq L \ \forall i \in [k]$. Thus, the $\mathbb{F}$-span of $\mathcal{X} = \{z^j g_i(z) | j + h_i \leq L\}$ is $\mathscr{C}_L$. For any $\mathbb{F}$-linear combination of elements of $\mathcal{X}$, one can collect the terms with each $g_i(z)$, and write

it as a $\mathbb{F}(z)$-linear combination of $g_i(z)_{i=1}^k$, which form a $\mathbb{F}(z)$-basis of $\mathscr{C}$ and hence are linearly independent. Thus, $\mathcal{X}$ is a $\mathbb{F}$-basis of $\mathscr{C}_L$.

$$\begin{aligned}
\delta_L &= \#\mathcal{X} \\
&= \#\left\{\bigcup_{i=1}^{k}\{z^j g_i(z) | j + h_i \leq L, i \text{ is fixed}\}\right\} \\
&= \sum_{i=1}^{k} \#\{z^j g_i(z) | j \leq L - h_i, i \text{ is fixed}\} \\
&= \sum_{i=1}^{k} \#\{j \in \mathbb{N} \cup \{0\} | j \leq L - h_i\} \\
&\implies \delta_L = \sum_{i=1}^{k} \max(L - h_i + 1, 0)
\end{aligned}$$

□

Corollary 1.2.6.

$$\sum \delta_L t^L = \frac{\sum_{i=1}^{k} t^{h_i}}{(1-t)^2}$$

Note: *This series is known as the Hilbert series associated with the code.*

Proof.

$$\begin{aligned}
\sum \delta_L t^L &= \sum_{L \geq 0} \sum_{i=1}^{k} \max(L - h_i + 1, 0)\, t^L \\
&= \sum_{i=1}^{k} (0 + \sum_{L \geq h_i} (L - h_i + 1) t^L)
\end{aligned}$$

Setting $j = L - h_i$,

$$\begin{aligned}
\sum \delta_L t^L &= \sum_{i=1}^{k} t^{h_i} \sum_{j \geq 0} (j+1) t^j \\
&\implies \sum \delta_L t^L = \frac{\sum_{i=1}^{k} t^{h_i}}{(1-t)^2}
\end{aligned}$$

□

Corollary 1.2.7. *Let $\mathscr{C}$ be an (n, k, γ)-convolutional code with Forney indices $h_1 \leq h_2 ... \leq h_k$,*

then

$$\delta_L \geq \max((L+1)k - \gamma, 0) \forall L \in \mathbb{N} \cup 0 \tag{1.4}$$

We have equality in (1.4) $\forall L$ iff $h_k - h_1 \leq 1$. A code for which this is true is called a compact code.

Proof. For any $L \in \mathbb{N} \cup \{0\}$, the i^{th} term of the series that is equal to δ_L is greater than or equal to $L + 1 - h_i$. Therefore,

$$\delta_L \geq \sum_{i=1}^{k} (L + 1 - h_i) = k(L+1) - \gamma$$

and we know that δ_L is non-negative. Therefore,

$$\delta_L \geq \max((L+1)k - \gamma, 0) \forall L \in \mathbb{N} \cup 0$$

If either $L + 1 \leq h_1$ or $L + 1 \geq h_k$, all the terms in (1.3) are 0 if $L + 1 \leq h_1$ and $L + 1 - h_i$ if $L + 1 \geq h_k$ and equality holds in (1.4).
If $h_k - h_1 \leq 1$, one of the above two cases holds $\forall L \in \mathbb{N} \cup \{0\}$, thus equality holds in (1.4) $\forall L$.
If $h_k - h_1 > 1$, $\exists$ at least one $L \in \mathbb{N}$ such that $h_k > L + 1 > h_1$. There is a largest index j such that $h_j \leq L + 1 \leq h_j + 1$, for any $i \in [k] \geq j + 1$, the terms in (1.3) are 0, as $L + 1 \leq h_i$, but for any $i \in [k] \leq j$, the terms of (1.3) are $L + 1 - h_i$, and so $\delta_L = j(L+1) - \sum_{i=1}^{j} h_i$, which is not equal to $\max((L+1)k - \gamma, 0)$. This establishes the second part of the corollary. □

For a compact code, $h_k - h_1 \leq 1$, i.e. they are either equal or differ by one, as they are integers. Since they add up to γ, if they are equal, we get $kh_i = \gamma$ implies that $h_i = \frac{\gamma}{k} \forall i \in [k]$. If they differ by one, the Forney indices are all equal to h_1 or $h_1 + 1 = h_k$. Say j of them are equal to $h_1 + 1$, then $kh_1 + j = \gamma$. Since k is the total number of indices, $j < k$, which implies that $j = \gamma \mod k$, $h_1 = \lfloor \frac{\gamma}{k} \rfloor$ and $h_k = \lceil \frac{\gamma}{k} \rceil$.

Since $\forall L$, $\mathscr{C}_L$ forms a subcode over $\mathbb{F}$, the minimum distance of $\mathscr{C}_L$ has to be greater than or equal to $d_{free}(\mathscr{C})$. Since $\delta_L \geq (L+1)k - \gamma$ and the fact that each code of lower dimension can be written as a subcode, we have $\Delta_q((L+1)n, (L+1)k - \gamma) \geq \Delta_q((L+1)n, \delta_L)$, where $\Delta_q(n,k)$ is the maximum possible minimum distance of a $(n,k)_q$ block code.

1.2.4 Physical Realisation of Encoders

Definition 1.2.8 (Abstract Linear Encoder)**.** *A linear onto function that maps $\mathbb{F}(z)^k$ to $\mathscr{C}$ is known as an abstract linear encoder for $\mathscr{C}$.*

$$\Phi : \mathbb{F}(z)^k \hookrightarrow \mathscr{C}$$

There exists a bijective correspondence between the set of abstract linear encoders and the set of generator matrices for $\mathscr{C}$. This is illustrated by the following.
Let Φ be an abstract linear encoder. $e_i(z) := (0, ..., 0, 1, 0, ..., 0) \in \mathbb{F}(z)^k$ such that 1 is at the i^{th} coordinate. This forms a basis of $\mathbb{F}(z)^k$. Let $g_i(z) = \Phi(e_i(z))$. Thus, $g_i(z) \in \mathbb{F}(z)^n$ and there are k of them.
$G(z) := \begin{pmatrix} g_1(z) \\ \vdots \\ g_k(z) \end{pmatrix}$ is a $k \times n$ matrix.
Let $u(z) \in \mathbb{F}(z)^k$, then $u(z)$ can be written as an $\mathbb{F}(z)$-linear combination of the h_i's, let it be $\sum_{i=1}^{k} h_i u_i(z)$. Then for any $x(z) \in \mathscr{C}$,

$$x(z) = \Phi(u(z)) = \Phi(\sum_{i=1}^{k} h_i u_i(z)) = \sum_{i=1}^{k} u_i(z)\Phi(h_i) = \sum_{i=1}^{k} u_i(z) g_i(z) \implies$$

the row space of $G(z)$ is $\mathscr{C}$. In fact, the rows form a basis for $\mathscr{C}$, as they are images of a set of linearly independent elements of $\mathbb{F}(z)^k$ under a linear function.
On the other hand, given a generator matrix, $G(z)$, we define the function as $\Phi : \mathbb{F}(z)^k \to \mathscr{C}$, $u(z) \mapsto u(z)G(z)$. This is clearly onto and linear; thus, Φ is a valid abstract encoder function for $\mathscr{C}$.

Definition 1.2.9 (Physical Realization)**.** *Given an encoder Φ for an (n, k)-convolutional code $\mathscr{C}$, its physical realisation E is a k-input n-output circuit such that when one feeds in the coefficients of $u(z)$ at each time step expressed as a Laurent series, one gets as the output the coefficients of $x(z) = \Phi(u(z))$ at each time step.*

It can be made using the following elements: Memory Elements, Delays, Adders, and Multipliers. The degree of a physical realisation is the number of delay elements in it.
Circuits in which the intermediate wires carry the signal from only one input, with or without delays, are known as direct form realisations (or controller canonical form).
Let $G(z)$ be the generator matrix associated with Φ. Φ has multiple physical realisations. The McMillan degree of $G(z)$ is the lowest degree of any physical realization of Φ.

Any rational function $\frac{p(z)}{q(z)}$, with $q(z) \neq 0$, can be written as a Laurent series, $\frac{p(z)}{q(z)} = \sum_{i=l}^{\infty} a_i z^i$. The causal part of a ration function is thus $\sum_{i=\max(0,l)}^{\infty} a_i z^i$. We can then define the causal part of a codeword $x(z)$ by taking the causal part of each coordinate. We denote the causal part of $x(z)$ by $x^*(z)$. A causal codeword is one that is equal to its causal part, whereas an anti-causal codeword is one with causal part 0.
Given an (n, k, m) convolutional code $\mathscr{C}$ with Forney indices $(h_1, .., h_k)$, we define $\mathscr{C}^* := \{x(z) \in \mathscr{C} | x^*(z) \in \mathscr{C}\}$, i.e. the set of codewords whose causal part is also a codeword. $\mathscr{C}^*$ is an infinite dimensional vector space over $\mathbb{F}$. It is also a subspace of $\mathscr{C}$ over $\mathbb{F}$.

Definition 1.2.10 (Abstract State Space)**.** *The quotient space $\Sigma_{\mathscr{C}} = \mathscr{C}/\mathscr{C}^*$ over $\mathbb{F}$ is known as the abstract state space of $\mathscr{C}$.*

Theorem 1.2.8. *If $\Sigma_{\mathscr{C}}$ is the abstract state space for a convolutional code $\mathscr{C}$, then*

$$\deg \mathscr{C} = \dim_{\mathbb{F}}(\Sigma_{\mathscr{C}}) = \gamma \tag{1.5}$$

[10]

Proof. Let $G(z)$ be a generator matrix for $\mathscr{C}$ with rows $\{g_i\}_{i=1}^k$ and $\deg(g_i(z)) = h_i$.

$$b_{i,j} := z^j g_i(z) \; i \in [k], j \in \{-h_i, ..., -1\}$$

Let $x(z)$ be any codeword in $\mathscr{C}$. Then, $\exists u(z) \in \mathbb{F}(z)^k$ such that $x(z) = u(z)G(z) = \sum_{i=1}^k u_i(z)g_i(z)$. Let $\sum_j u_{i,j} z^j$ be the Laurent series expansion of $u_i(z)$. Define $u'(z) := \sum_{j=-h_i}^{-1} u_{i,j} z^j$ and $x'(z) := \sum_{i=1}^k u_i'(z)g_i(z) = \sum_{i=1}^k \sum_{j=-h_i}^{-1} u_{i,j} z^j g_i(z) = \sum_{i=1}^k \sum_{j=-h_i}^{-1} u_{i,j} b_{i,j}$. Therefore, $x'(z)$ lies in $span_{\mathbb{F}}\{b_{i,j}\}$. Now,

$$x(z) - x'(z) = \sum_{i=1}^{k} \sum_{j=-h_i}^{0} u_{i,j} z^j g_i(z) + \sum_{i=1}^{k} \sum_{j \geq 0} u_{i,j} z^j g_i(z)$$

Which is the sum of a causal codeword and an anticausal codeword. Thus, $x(z) - x'(z) \in \mathscr{C}^*$, i.e. $x(z) \in x'(z) + \mathscr{C}^*$. Since $x(z)$ is arbitrary, we see that $\Sigma_{\mathscr{C}}$ is spanned by cosets of $\{b_{i,j}\}$.
Given a $\mathbb{F}$-linear combination of $b_{i,j}$, $\sum_{i=1}^k \sum_{j=-h_i}^{-1} u_{i,j} b_{i,j} = \sum_{i=1}^k u_i(z) g_i(z)$, where $u_i(z) = \sum_{j=-h_i}^{-1} u_{i,j} z^j$ and not all $u_{i,j}$ are 0, we have to show that this isn't 0 in $\Sigma_{\mathscr{C}}$, i.e. $\notin \mathscr{C}^*$. $\deg(u_i(z)g_i(z)) \leq h_i - 1$ as $\deg(u_i(z)) \leq -1$ as if any $u_{i,j}$ is non-zero, then it is the coefficient of z^{-j}, and there is no positive power in $u_i(z)$, if all are 0 for some i, then $\deg(u_i(z)) = \deg(0) = -\infty$ for that value of i. Assume to the contrary, then the causal part of $x(z)$, $x^*(z)$ is a codeword as well. Then $\exists v(z) \in \mathbb{F}(z)^k$ such that $x(z) = v(z)G(z)$. Since $x(z)$ has finite terms with positive power, $x^*(z)$ is a polynomial vector, therefore by the basic nature of $G(z)$, $v(z) = (v_1(z), ..., v_k(z))$ is also a

polynomial vector. Looking at the anticausal part of $x(z)$,

$$x(z) - x^*(z) = \sum (u_i(z) - v_i(z))g_i(z)$$

Since $u_i(z)$ only has terms of negative power, if any coordinate of $v(z)$ is non-zero, then, by the predictable degree property, the degree of the anticausal part of $x(z)$ is non-negative, which is a contradiction by the definition of anticausal. Thus the causal part of $x(z)$ has to be 0. but if this is the case, $\deg(x(z)) < 0$, which is only possible if all $u_{i,j}$ are 0, thus contradicting the fact that not all $u_{i,j}$ are 0. Therefore, $x(z) \notin \mathscr{C}^*$, therefore $\{b_{i,j}\}$ are linearly independent.
This means that $\{b_{i,j}\}$ forms an $\mathbb{F}$-basis for $\Sigma_{\mathscr{C}}$ and there are exactly m of them, thus $\dim_{\mathbb{F}}(\Sigma_{\mathscr{C}}) = m = \deg(\mathscr{C})$. □

Corollary 1.2.9. *For any physical encoder K of $\mathscr{C}$,* $\deg(K) \geq \gamma$ *[10]*

We can now describe the state space diagram, as in [2], of a Convolutional code.
The state space diagram is a labelled directed graph with the vertex set $\Sigma_{\mathscr{C}}$ and edge set $\{(X, Y) \in \Sigma_{\mathscr{C}} \times \mathbb{F}_q^{\gamma} | \exists$ at least one $u \in \mathbb{F}^k$ such that $Y = X\mathcal{A} + u\mathcal{B}\}$ where (X, Y) is an ordered pair indicating the direction of the edge. The edge (X, Y) is labelled by $\binom{u}{v}$ where u satisfies

$$Y = X\mathcal{A} + u\mathcal{B}$$

and v is the corresponding output:

$$v = X\mathcal{C} + u\mathcal{D}$$

It is clear that the labels of the graph are in no way unique, especially when not all the Forney indices of the code are non-zero.
As we shall see later, the concept of weight distribution used for convolutional codes is closely related to the state space diagram.

We shall now describe a method for constructing a minimal realisation given a polynomial generator matrix

$$G = (g_{i,j}(z))_{i,j};\ g_{i,j}(z) = \sum_{k=0}^{h_i} g_{i,j}^{(k)} z^k$$

with row degrees $h_1, h_2, ..., h_k$ and setting h as sum of row degrees, a minimal realisation is one in which the dimensions of $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ are $(h \times h, k \times h, h \times n, k \times n)$. One can now construct a state space description of the code known as the controller canonical form. As shown in [2], this is achieved as follows:

$\forall i \in \{1, 2, 3, .., k\}$

$$A_i := \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}_{h_i \times h_i} ; B_i := \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix} \in \mathbb{F}^{h_i};$$

$$C_i := (g_{i,j}^{(k)})_{k,j} \in \mathbb{F}^{h_i \times n} \text{ where k lies in } \{1, 2, ..., h_i\}$$

Where k varies from 1 to h_i. When the h_i are 0, the corresponding matrices are taken to be 0. We then define the following matrices:

$$\mathcal{A} := \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{pmatrix}; \mathcal{B} := \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_k \end{pmatrix}; \mathcal{C} := \begin{pmatrix} C_1 \\ \vdots \\ C_k \end{pmatrix}$$

As well as

$$\mathcal{D} := \begin{pmatrix} g_{i,j}^{(0)} \end{pmatrix}$$

One can then easily calulate that $G(z) = \mathcal{B}(I - \mathcal{A}D^{-1})\mathcal{C} + \mathcal{D}$, i.e. $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is a valid state space description of $\mathscr{C}$. Clearly, given a canonical matrix, the row degrees are replaced by the Forney indices $\{\gamma_i\}$.

Assume that G is a canonical PGM with row degrees $\{\gamma_i\}$ such that $\gamma_1, \gamma_2, ...\gamma_r > 0 = \gamma_{r+1} = ... = \gamma_k$. Let $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ be the controller canonical form of G as described above. We define the following index sets that are used frequently in what follows.

$$\mathcal{J} = \{\gamma_1, \gamma_1 + \gamma_2, ..., \gamma\}$$

$$\mathcal{I} = \{1, 1 + \gamma_1, 1 + \gamma_1 + \gamma_2, ..., 1 + \sum_{i=1}^{r-1} \gamma_i\}$$

Keeping in mind the above conditions, a few important properties of this realisation that are used extensively in Chapter 2 are:

1. $\mathcal{A}\mathcal{B}^T = 0$
2. $\mathcal{B}\mathcal{B}^T = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$
3. $\mathcal{A}^T\mathcal{A}_{i,j} = \begin{cases} 1 & i = j \notin \mathcal{I} \\ 0 & \text{else} \end{cases}$

4. $\mathcal{B}^T\mathcal{B}_{i,j} = \begin{cases} 1 & i = j \in \mathcal{I} \\ 0 & \text{else} \end{cases}$

5. $\{u\mathcal{B}|u \in \mathbb{F}_q^k\} = Sp_{\mathbb{F}_q}(\{e_i|i \in \mathcal{I}\})$ and $\ker \mathcal{B} = \left\{u \begin{pmatrix} 0_{(k-r)\times r} & I_{k-r} \end{pmatrix}\right\}$, which can also be written as $Sp_{\mathbb{F}_q}(\{e_i|i \in \{k-r+1, ..., k\}\})$

Where $Sp_F(\cdot)$ denotes the space generated by $\cdot$ over F and e_i denotes the vector that has all 0's except at the i^{th} coordinated, where it is 1. The first four can easily be confirmed by simple matrix multiplication, while the 5^{th} is obvious by taking a look at the matrix $\mathcal{B}$.

1.2.5 Catastrophic Generator Matrices

Definition 1.2.11. *A generator matrix $G(z)$ is said to be catastrophic if $\exists u(z) \in \mathbb{F}(z)^k$ such that $wt(u(z)) = \infty$, but $x(z) = u(z)G(z)$ is of finite weight.*

Theorem 1.2.10 (Massey-Sain)**.** *Let $G(z)$ be a PGM for $\mathscr{C}$, a convolutional code. The following are equivalent:*

1. *No infinite weight input $u(z)$ can produce a finite weight codeword $x(z) = u(z)G(z)$*
2. *The gcd of $k \times k$ minors of $G(z)$ is a power of D*
3. *$G(z)$ has a finite weight right inverse.*

If a PGM satisfies any of the above conditions, it is known as a non-catastrophic PGM. Otherwise, it is known as a catastrophic PGM.

Chapter 2

MacWilliams Identity for Convolutional Codes

For this chapter, let $\mathscr{C}$ be an $(n,k)-$convolutional code and G be a canonical generator matrix for $\mathscr{C}$ with forney indices $\{\gamma_1,...,\gamma_k\}$ such that the first r are positive and the rest are 0 and $\deg(\mathscr{C}) = \sum_{i=1}^{r}\gamma_i = \gamma$. We also fix $(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ to be the controller canonical realisation for G. Let the dual of $\mathscr{C}$ be denoted by $\hat{\mathscr{C}}$. We also chose to work in the framework of the polynomial ring $\mathbb{F}_q[z]$ instead of the field of rational functions $\mathbb{F}_q(z)$ over the finite field $\mathbb{F}_q$, where $q = p^s$, making the code a k-dimensional submodule of the module $\mathbb{F}[z]^n$. For any object χ related to the code, let $\hat{\chi}$ denote the analogous object for the dual code. This is not to be confused with the objects $\chi^\perp$ when χ is a set associated with the code, as $\chi^\perp$ denotes the orthogonal to said set.

Definition 2.0.1 (Concatenated Codewords)**.** *A polynomial codeword $v \in \mathscr{C}$ is said to be concatenated at l if it can be written as*

$$v = v^* + \hat{v}$$

where v^* and $\hat{v}$ are both non-zero codewords such that $\deg v^* = l-1$ and $z^{-l}\hat{v}$ is a polynomial codeword as well.

Definition 2.0.2 (Atomic Codewords)**.** *A polynomial codeword that is not concatenated for any non-negative integer l is known as an atomic codeword.*

Since, in the convolutional case, the weight of a codeword is just the sum of the weights of the coefficient vectors when written as a Laurent Series, there can be infinite vectors of a given weight. Instead, the weight enumerators, in this case, only count the atomic atomic codewords. The reason for this becomes apparent with the following lemma from [2].

Lemma 2.0.1. *Let the maximum Forney index for G be m.*

1. *Let $v = uG$ be a codeword where $u = \sum_{i=0}^{L} u_i z^i$ such that $u_0 \neq 0 \neq u_L$. If for m consectutive indices $u_i = 0$ then v is concatenated.*
2. *For all non-negative integers α, the number of atomic codewords with weight lesser than or equal to alpha is finite.*

We can now define the weight distribution for convolutional codes

Definition 2.0.3. *For all positive integers α and l, let $\omega_{l,\alpha} = \#\{v \in \mathscr{C} | v \text{ is atomic, } wt(v) = \alpha, \deg(v) = l - 1\}$. The weight distribution of the convolutional code $\mathscr{C}$ is the generating function:*

$$\Omega(W, L) = \sum_{l=1}^{\infty} \sum_{\alpha=1}^{\infty} \omega_{l,\alpha} W^{\alpha} L^{l}$$

Unfortunately, as shown by Shearer and McEliece in [11], a MacWilliam's type identity does not exist for the weight distribution as given above. Thus, we look at another way of enumerating the weights, known as the Weight Adjacency Matrices.

2.1 Weight Adjacency Matrices

We now describe an alternative way to enumerate the codewords by weight first introduced in [1].

Definition 2.1.1. *For all X, Y in the state space, and $\alpha \in \{0, 1, 2, ..., n\}$ let*

$$\lambda_{X,Y}^{(\alpha)} = \#\{u \in \mathbb{F}^k | Y = X\mathcal{A} + u\mathcal{B}, wt(X\mathcal{C} + u\mathcal{D}) = \alpha\}$$

The adjacency matrix is then defined as:

$$\Lambda_G = \left(\sum_{\alpha=0}^{n} \lambda_{X,Y}^{(\alpha)} W^{\alpha} \right)_{X,Y} \in \mathbb{Z}[W]^{q^{\gamma} \times q^{\gamma}}$$

We also define

$$\lambda_{X,Y} = \sum_{\alpha=0}^{n} \lambda_{X,Y}^{(\alpha)} W^{\alpha}$$

One can see that this is just the weight enumerator for the set of labels on the edge (X, Y) in the

state space diagram for the code related to the above realisation.
Clearly, the adjacency matrix, unlike the Weight Distribution given above, depends on the generator matrix used and the realisation of said generator matrix.
Since any two minimal realisations, $(\mathcal{A}', \mathcal{B}', \mathcal{C}', \mathcal{D}')$ and $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ of a given generator matrix for a code are related as follows:

$$(\mathcal{A}', \mathcal{B}', \mathcal{C}', \mathcal{D}') = (T\mathcal{A}T^{-1}, \mathcal{B}T^{-1}, T\mathcal{C}', \mathcal{D}')$$

for some invertible $\gamma \times \gamma$ matrix T, we see that for all X, Y in the state space:

$$\lambda'^{(\alpha)}_{X,Y} = \#\{u \in \mathbb{F}^k | Y = X\mathcal{A}' + u\mathcal{B}', \mathrm{wt}(X\mathcal{C}' + u\mathcal{D}') = \alpha\}$$
$$= \#\{u \in \mathbb{F}^k | YT = (XT)\mathcal{A} + u\mathcal{B}T^{-1}, \mathrm{wt}((XT)\mathcal{C} + u\mathcal{D}) = \alpha\} = \lambda^{(\alpha)}_{XT,YT}$$

Thus, for any minimal representation of a fixed canonical generator matrix, the adjacency matrices are just "reorderings" of each other.
One can recover the weight enumerator from the matrix through a process described by McEliece and summarised in [4].
The problem with the adjacency matrices is that they are not an invariant for a code since they depend on the chosen generator and its state space description. To get around this, we define 2 matrices M and M' in $\mathbb{Z}[W]^{q^\gamma \times q^\gamma}$ to be related ($M \sim M'$) if there exists an $\gamma \times \gamma$ invertible matrix T such that

$$M_{X,Y} = M'_{XT,YT}$$

for all X and Y in $\mathbb{F}^\gamma$.
This relation is clearly symmetric, reflexive, and transitive; hence, it is a valid equivalence relation. Let $\bar{M}$ be the equivalence class of M in $\mathbb{Z}[W]^{\gamma^m \times \gamma^m}$. Clearly, given a canonical generator matrix G, the adjacency matrices of every minimal realisation of G lie in the same equivalence class. In fact, this equivalence class is an invariant of the code[2]. This class is known as the generalised adjacency matrix for the code $\mathscr{C}$ and is denoted by $\bar{\Lambda}(\mathscr{C})$.

Definition 2.1.2. *By $\mu_{X,Y}$, we denote the set of inputs $u \in \mathbb{F}_q^k$ such that $Y = X\mathcal{A} + u\mathcal{B}$, i.e.*

$$\mu_{X,Y} := \{u \in \mathbb{F}_q^k | Y = X\mathcal{A} + u\mathcal{B}\}$$

Definition 2.1.3. *We use μ to denote the subset of $\mathbb{F}^\gamma \times \mathbb{F}^\gamma$ consisting of pairs of states (X, Y) such that the set $\mu_{X,Y}$ is non-empty. Such a pair of states is known as a connected pair. Thus,*

$$\mu := \{(X, Y) \in \mathbb{F}^\gamma \times \mathbb{F}^\gamma | \mu_{X,Y} \neq \emptyset\}$$

Proposition 2.1.1. *[4]*

$$\mu = \left\{ (X,u) \begin{pmatrix} I_\gamma & \mathcal{A} \\ 0 & \mathcal{B} \end{pmatrix} | (X,u) \in \mathbb{F}_q^\gamma \times \mathbb{F}_q^k \right\} \subset \mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma$$

as vector space over $\mathbb{F}_q$, and therefore, it has rank $\gamma + r$

This is easily verified by looking at the state space description of the code.

Proposition 2.1.2. *[4] The orthogonal of μ is given by*

$$\mu^\perp = \{(X, -X\mathcal{A}) | X_j = 0 \forall j \in \mathcal{J}\}$$

Proof. Let (A, B) be in the orthogonal. Then,

$$(A,B)(X,Y)^T = 0$$

for all connected pairs of states (X, Y). Therefore,

$$(A,B)\left(\begin{pmatrix} I_\gamma & \mathcal{A} \\ 0 & \mathcal{B} \end{pmatrix}^T (X,u)^T \right) = 0$$

From this, we arrive at the following:

$$\begin{aligned} & A + B\mathcal{A}^T = 0 \\ \implies & B\mathcal{A}^T\mathcal{A} = -A\mathcal{A} \\ \implies & B = -A\mathcal{A} \end{aligned}$$

And the equation

$$B\mathcal{B}^T = 0$$

This implies that $B = -A\mathcal{A}$ is zero at all the indices coinciding with 1 in $\mathcal{B}$, i.e. the indices in $\mathcal{J}$. Thus, A is also zero at the indices in $\mathcal{J}$. Therefore, any (A, B) orthogonal to μ is of the form $(A, -A\mathcal{A})$ where A is 0 at the indices in $\mathcal{J}$. □

2.1.1 Useful Block Codes

We consider the two block codes:

1. $\mathscr{C}_C = \left\{ u \begin{pmatrix} \mathcal{C} \\ \mathcal{D} \end{pmatrix} | u \in \mathbb{F}_q^{\gamma+k} \right\}$

2. $\mathscr{C}_{const} = Sp_{\mathbb{F}_q}(\{g_i(z)\}_{i=r+1}^{k})$, where $Sp_F()$ denotes the space generated by the object inside the parentheses over F. In this case, the space is generated over $\mathbb{F}_q$

As is obvious, for a delay-free code, as we assume from here on, the dimension of $\mathscr{C}_C$ is at least k. Let it be $k + \hat{r}$, where $\hat{r}$ lies in $\{0, ..., n-k\}$. Similarly, the dimension of $\mathscr{C}_{const}$ is $k - r$.
Since we are working with the polynomial ring, we see that $\mathscr{C}_{const}$ is the same as $\mathbb{F}_q \cap \mathscr{C}$. Also, we can clearly see that $\mathscr{C}_{const}$ is contained in $\mathscr{C}_C$ since the generator of $\mathscr{C}_{const}$ is just the last $k-r$ rows of $\mathcal{D}$. From the properties of the controller canonical form, we know that

$$\ker \mathcal{B} = Sp_{\mathbb{F}_q}(\{e_i | i \in \{r+1, ..., k\}\}) = \left\{ u \begin{pmatrix} 0 & I_{k-r} \end{pmatrix} \right\}$$

Thus, we see that the rowspace of the last $k - r$ rows of $\mathcal{D}$ is exactly $(\ker \mathcal{B})\mathcal{D}$.
The following lemma from [4] gives a glimpse into the utility of these codes when it comes to MacWilliams Identities:

Lemma 2.1.3. *For information as in this chapter,*

$$(\mathscr{C}_C)^{\perp} = \hat{\mathscr{C}}_{const}$$

Proof. Recall the notation as in the definition of the controller canonical form for the canonical generator matrix G of $\mathscr{C}$. We define $g_i^{(k)}$ as the vector of coefficients of z^k in row g_i of the generator matrix.
For any $c \in (\mathscr{C}_C)^{\perp}$

$$\begin{aligned} & (c, g_i^{(k)}) = 0 \forall i, k \\ \iff & (c, g_i) = 0 \forall i \\ \iff & c \in \hat{\mathscr{C}} \cap \mathbb{F}_q = \hat{\mathscr{C}}_{\text{const}} \end{aligned}$$

Where the 2nd equality holds since c is a constant vector. □

From the above proposition, we can conclude that the dimension of $\hat{\mathscr{C}}_{const}$ is $n - k - \hat{r}$. Since $\hat{\mathscr{C}}$ is an (n,n-k) convolutional code, we can conclude that the number of non-zero Forney indices of $\hat{\mathscr{C}}$ is $\hat{r}$.

2.1.2 Relation between $\mathscr{C}_{const}$ and The Adjacency Matrix

The nontrivial entries of Λ_G can be described in terms of the weight enumerator of cosets of $\mathscr{C}_{const}$. For any two states X and Y, define

$$\sigma(X,Y) : \mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma \to \mathbb{F}_q^n \text{ such that} (X,Y) \mapsto X\mathcal{C} + Y\mathcal{B}^T\mathcal{D}$$

This is a vector space homomorphism.
We can now prove the following,

Lemma 2.1.4. *For any pair of states (X,Y) in μ,*

$$\lambda_{X,Y} = we(\sigma(X,Y) + \mathscr{C}_{const}) \tag{2.1}$$

Proof. For any pair of states (X,Y) in μ, there exists atleast one $u \in \mathbb{F}_q^k$ such that

$$\begin{aligned} Y &= X\mathcal{A} + u\mathcal{B} \\ \implies Y\mathcal{B}^T - X\mathcal{A}\mathcal{B}^T &= u\mathcal{B}\mathcal{B}^T \end{aligned}$$

Through the definitions of the controller canonical form, it is easy to see that $\mathcal{A}\mathcal{B}^T = 0$, and that

$$\mathcal{B}\mathcal{B}^T = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

Thus, for any $u \in \mathbb{F}_q^K$ satisfying the above, the first r coordinates are determined by the state Y, while the other $k-r$ can be anything, as the last $k-r$ rows of $\mathcal{B}$ are 0. Therefore,

$$\{u \in \mathbb{F}_q^k | Y - X\mathcal{A} = u\mathcal{B}\} = Y\mathcal{B}^T + \left\{ v \begin{pmatrix} 0 & I_{k-r} \end{pmatrix} | v \in \mathbb{F}_q^{k-r} \right\} = Y\mathcal{B}^T + \ker \mathcal{B}$$

Let $M = \{u \in \mathbb{F}_q^k | Y - X\mathcal{A} = u\mathcal{B}\}$ Now, $\lambda_{X,Y}$ can be written as the weight enumerator polynomial for the set of outputs with starting state X and ending state Y, which becomes the set $\{X\mathcal{C} + u\mathcal{D} | u \in M\}$

$$\begin{aligned} \{X\mathcal{C} + u\mathcal{D} | u \in M\} &= X\mathcal{C} + (Y\mathcal{B}^T + \ker \mathcal{B})\mathcal{D} \\ &= (X\mathcal{C} + Y\mathcal{B}^T\mathcal{D}) + (\ker \mathcal{B})\mathcal{D} \\ &= \sigma(X,Y) + \mathscr{C}_{\text{const}} \end{aligned}$$

Thus, $\lambda_{X,Y} = \text{we}(\sigma(X,Y) + \mathscr{C}_{\text{const}})$, as we claimed. □

From this, we can conclude that for any pair of states (X, Y) in μ such that $\sigma(X, Y) \in \mathscr{C}_{\text{const}}$

$$\lambda_{X,Y} = \text{we}(\mathscr{C}_{\text{const}})$$

If $\sigma(X, Y)$ is not in $\mathscr{C}_{\text{const}}$, since for all $\zeta \in \mathbb{F}_q^\times$, $w(\zeta u) = w(u)$, we have for any $\zeta \in \mathbb{F}_q^\times$

$$\text{we}(\sigma(X, Y) + \mathscr{C}_{\text{const}}) = \text{we}(\zeta(\sigma(X, Y) + \mathscr{C}_{\text{const}})) = \text{we}(\zeta(\sigma(X, Y)) + \mathscr{C}_{\text{const}})$$

Since $\sigma(X, Y) \notin \mathscr{C}_{\text{const}}$, these sets are disjoint for all values of $\zeta \in \mathbb{F}_q$. Therefore, we can write the space generated by $\sigma(X, Y)$ and $\mathscr{C}_{\text{const}}$, $Sp_{\mathbb{F}_q}(\mathscr{C}_{\text{const}}, \sigma(X, Y))$ as a disjoint union of the sets $\zeta(\sigma(X, Y)) + \mathscr{C}_{\text{const}}$ over all $\zeta \in \mathbb{F}_q$. Thus,

$$\begin{aligned}
\text{we}(Sp(\mathscr{C}_{\text{const}}, \sigma(X, Y))) &= \sum_{\zeta \in \mathbb{F}_q} \text{we}(\zeta(\sigma(X, Y)) + \mathscr{C}_{\text{const}}) \\
&= \text{we}(\mathscr{C}_{\text{const}}) + \sum_{\zeta \in \mathbb{F}_q^\times} \text{we}(\zeta(\sigma(X, Y)) + \mathscr{C}_{\text{const}}) \\
&= \text{we}(\mathscr{C}_{\text{const}}) + (q-1)\text{we}((\sigma(X, Y)) + \mathscr{C}_{\text{const}}) \\
\text{we}((\sigma(X, Y)) + \mathscr{C}_{\text{const}}) &= \frac{1}{q-1} \left(\text{we}(Sp(\mathscr{C}_{\text{const}}, \sigma(X, Y)) - \text{we}(\mathscr{C}_{\text{const}}))\right)
\end{aligned}$$

We can then conclude that for any pair of states (X, Y),

$$\lambda_{X,Y} = \begin{cases} 0 & (X, Y) \notin \mu \\ \text{we}(\mathscr{C}_{\text{const}}) & (X, Y) \in \mu; \sigma(X, Y) \in \mathscr{C}_{\text{const}} \\ \frac{1}{q-1} \left(\text{we}(Sp(\mathscr{C}_{\text{const}}, \sigma(X, Y)) - \text{we}(\mathscr{C}_{\text{const}}))\right) & (X, Y) \in \mu; \sigma(X, Y) \notin \mathscr{C}_{\text{const}} \end{cases} \tag{2.2}$$

To simplify notation, we define the set Ξ as

$$\Xi := \{(X, Y) \in \mu | \sigma(X, Y) \in \mathscr{C}_{\text{const}}\} \tag{2.3}$$

An alternative way of characterising this set from [3] is:

Proposition 2.1.5.

$$\Xi = \{(X, Y) \in \mu | \exists u \in \mu_{X,Y} \textit{ such that } 0 = X\mathcal{C} + u\mathcal{D}\}$$

Proof. Let (X, Y) be in Ξ.
Since $\Xi \subset \mu$, we can write Y as $X\mathcal{A} + u\mathcal{B}$ for some $u \in \mathbb{F}^k$.

Now,

$$\begin{aligned}\sigma(X,Y) &= X\mathscr{C} + Y\mathcal{B}^T\mathcal{D} \\ &= X\mathcal{C} + X\mathcal{A}\mathcal{B}^T\mathcal{D} + u\mathcal{B}\mathcal{B}^T\mathcal{D} \\ &= X\mathcal{C} + u\mathcal{B}\mathcal{B}^T\mathcal{D} \in \mathscr{C}_{\text{const}}\end{aligned}$$

From the previous subsection, we know that $\mathscr{C}_{\text{const}} = (\ker \mathcal{B})\mathcal{D}$, thus for some $u \in \ker \mathcal{B}$

$$\begin{aligned}x\mathcal{C} + u\mathcal{B}\mathcal{B}^T\mathcal{D} &= u'\mathcal{D} \\ \implies X\mathcal{C} + (u\mathcal{B}\mathcal{B}^T - u')\mathcal{D} &= 0\end{aligned}$$

$(u\mathcal{B}\mathcal{B}^T - u')\mathcal{B} = u\mathcal{B}$, since from the proof of lemma 2.1.4, we know that

$$\mathcal{B}\mathcal{B}^T = \begin{pmatrix} 0 & I_r \\ 0 & 0 \end{pmatrix}$$

and $u' \in \ker \mathcal{B}$, hence $u\mathcal{B}\mathcal{B}^T - u' \in \mu_{X,Y}$ and therefore (X,Y) lies in the set on the right.
To prove the reverse inclusion, let (X,Y) lie in the set on the right, i.e. for some $u \in \mu_{X,Y}$,

$$Y = x\mathcal{A} + u\mathcal{B}$$

$$0 = X\mathcal{C} + u\mathcal{D} \implies X\mathcal{C} = -u\mathcal{D}$$

$$\begin{aligned}\sigma(X,Y) = X\mathcal{C} + X\mathcal{A}\mathcal{B}^T\mathcal{D} + u\mathcal{B}\mathcal{B}^T\mathcal{D} \\ = -u\mathcal{D} + u\mathcal{B}\mathcal{B}^T\mathcal{D} \\ = u(I_r - I_k)\mathcal{D}\end{aligned}$$

By the properties of $\ker \mathcal{B}$, we can see that $u(I_r - I_k)$ is just r 0's followed by an element in $\ker \mathcal{B}$, hence the element above is in $\mathscr{C}_{\text{const}}$, i.e. $\sigma(X,Y) \in \mathscr{C}_{\text{const}}$ which implies $(X,Y) \in \Xi$.
We can now conclude

$$\Xi = \{(X,Y) \in \mu | \exists u \in \mu_{X,Y} \text{ such that } 0 = X\mathcal{C} + u\mathcal{D}\}$$

□

Since $\mathscr{C}_{const}$ is a subspace of $\mathscr{C}_C$ and

$$\sigma(X,Y) = (X, Y\mathcal{B}^T)\begin{pmatrix}\mathcal{C}\\ \mathcal{D}\end{pmatrix}$$

also lies in $\mathscr{C}_C$, we see that the function

$$\varrho(X,Y) : \mu \to \mathscr{C}_C/\mathscr{C}_{\text{const}} : \sigma(X,Y) + \mathscr{C}_{\text{const}}$$

as defined in [4], is well defined. From the definition of Ξ (2.3), it is obvious that

$$\ker \varrho = \Xi$$

It also has the following properties,

Lemma 2.1.6. *ϱ is a surjective homomorphism that satisfies*

$$\dim(\Xi) = \gamma - \hat{r}$$

Proof. The fact that ϱ is a homomorphism follows directly from the homomorphism nature of σ. As for the surjectivity, we notice that for any $v \in \sigma(\mu)$,

$$\begin{aligned} v &= u\begin{pmatrix} I & \mathcal{A}\\ 0 & \mathcal{B}\end{pmatrix}\begin{pmatrix}\mathcal{C}\\ \mathcal{B}^T\mathcal{D}\end{pmatrix}\\ &= u\begin{pmatrix}\mathcal{C} + \mathcal{A}\mathcal{B}^T\mathcal{D}\\ \mathcal{B}\mathcal{B}^T\mathcal{D}\end{pmatrix}\\ &= u\begin{pmatrix}\mathcal{C}\\ \mathcal{B}\mathcal{B}^T\mathcal{D}\end{pmatrix}\end{aligned}$$

Thus, clearly, any vector of the form

$$u\begin{pmatrix}\mathcal{C}\\ \mathcal{B}\mathcal{B}^T\mathcal{D}\end{pmatrix}$$

is in $\sigma(\mu)$. Observe that the matrix

$$\begin{pmatrix}\mathcal{C}\\ \mathcal{B}\mathcal{B}^T\mathcal{D}\end{pmatrix}$$

is just the rows of the matrix

$$\begin{pmatrix}\mathcal{C}\\ \mathcal{D}\end{pmatrix}$$

excluding the last $k-r$ rows which generate $\mathscr{C}_{\text{const}}$; since $\varrho(\mu) = \sigma(\mu) + \mathscr{C}_{\text{const}}$, we can then conclude

that $\varrho(\mu) = \mathscr{C}_C/\mathscr{C}_{\text{const}}$, and hence ϱ is surjective.
As for the dimension of the kernel, by the first isomorphism theorem and the fact that ϱ is surjective, we have that

$$\begin{aligned} \dim(\mu) - \dim(\ker \varrho) &= \dim(\mathscr{C}_C) - \dim(\mathscr{C}_{\text{const}}) \\ \implies \dim(\ker \varrho) &= \gamma + r - (k + \hat{r}) + (k - r) \\ \dim \Xi &= \gamma - \hat{r} \end{aligned}$$

□

From this, one can conclude that $\varrho(\mu) + \mathscr{C}_{\text{const}} = \mathscr{C}_C$.

Definition 2.1.4. *The set μ^- is defined as*

$$\mu^- := Sp_{\mathbb{F}_q}(\{(0, e_i) \in \mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma | i \notin \mathcal{I}\}$$

Where e_i is the unit vector with 1 at the i^{th} coordinate and zero elsewhere.

Proposition 2.1.7. *Let μ^* be any subspace of μ such that $\mu = \mu^* \oplus \Xi$. Then $\mu^* \oplus \Xi \oplus \mu^- = \mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma$*

Proof. $\mu \cap \mu^- = \{0\}$ since none of the generators of μ^- can be written as

$$u \begin{pmatrix} I & \mathcal{A} \\ 0 & \mathcal{B} \end{pmatrix}$$

for some u in $\mathbb{F}_q^{\gamma+k}$ since none of the e_i for $i \notin \mathcal{I}$ are in the rowspace of $\mathcal{B}$ and we know that this matrix generates μ from Proposition 2.1.1
$\dim \mu^-$ is $\gamma - r$ and thus, $\dim \mu^- + \dim \mu = 2\gamma$. We can then conclude that

$$\mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma = \mu \oplus \mu^- = \mu^* \oplus \Xi \oplus \mu^- \tag{2.4}$$

Where the second equality holds from the assumption in the statement of the proposition. □

For any pair of states (X, Y) in μ^-, $(X, Y) = (0, \sum_{i \notin \mathcal{I}} a_i e_i)$. Thus,

$$\sigma(X, Y) = X\mathcal{C} + Y\mathcal{B}^T\mathcal{D} = 0\mathcal{C} + \sum_{i \notin \mathcal{I}} a_i (e_i \mathcal{B}^T)\mathcal{D} = 0$$

The last equality follows from the fact that $e_i \mathcal{B}^T = 0$ for any $i \notin \mathcal{I}$. Therefore, $\sigma|_{\mu^-} = 0$.

Lemma 2.1.8 (Invariance of λ). *Let (X, Y) be in μ^* and (X', Y') be in Ξ, then*

$$\lambda_{X,Y} = \lambda_{X+X',Y+Y'}$$

Proof. From the linearity of σ, we have

$$\sigma(X + X', Y + Y') = \sigma(X, Y) + \sigma(X', Y')$$

By (2.3), $\sigma(X', Y') \in \mathscr{C}_{\text{const}}$, $\sigma(X + X', Y + Y') + \mathscr{C}_{\text{const}} = \sigma(X, Y) + \mathscr{C}_{\text{const}}$. Therefore, by (2.1),

$$\begin{aligned} \lambda_{X+X',Y+Y'} &= \text{we}(\sigma(X + X', Y + Y') + \mathscr{C}_{\text{const}}) \\ &= \text{we}(\sigma(X, Y) + \mathscr{C}_{\text{const}}) \\ \implies \lambda_{X+X',Y+Y'} &= \lambda_{X,Y} \end{aligned}$$

□

2.2 MacWilliams Matrices

Let $q = p^s$ and η be a primitive p^{th} root of unity in $\mathbb{C}$. Then the group of complex valued characters on $\mathbb{F}_q^\gamma$ is given by

$$\{\tau_X := \eta^{\text{Tr}((X,\cdot))} | X \in \mathbb{F}_q^\gamma\}$$

Where $\text{Tr} : \mathbb{F}_q \to \mathbb{F}_p : x \mapsto \sum_{i=0}^{s-1} x^{p^i}$ is the trace defined on $\mathbb{F}_q$ and $(\cdot, \cdot)$ is the bilinear form on $\mathbb{F}_q^\gamma$

Definition 2.2.1. *For any invertible $\gamma \times \gamma$ matrix P over $\mathbb{F}_q$, we define the P-MacWilliams Matrix as*

$$\mathscr{M}(P) = q^{\frac{-\gamma}{2}} (\tau_{(XP)}(Y))_{X,Y} \;\; X, Y \in \mathbb{F}^\gamma \tag{2.5}$$

Thus, the MacWilliams matrices are $q^\gamma \times q^\gamma$ complex-valued matrices. Let $\mathscr{M}(I)$ be denoted by $\mathscr{M}$.

For any invertible matrix P in $GL_\gamma(\mathbb{F}_q)$, define the complex-valued $q^\gamma \times q^\gamma$ invertible matrix $\mathcal{P}(P)$ as

$$\mathcal{P}(P)_{X,Y} = \begin{cases} 1 & Y = XP \\ 0 & \text{else} \end{cases} \tag{2.6}$$

Let Π denote the set of all such permutation matrices, i.e. $\{\mathcal{P}(P) | P \in \text{GL}_\gamma(\mathbb{F}_q)\}$.
Clearly, $\mathcal{P}$ is a group homomorphism from $\text{GL}_\gamma(\mathbb{F}_q)$ to Π.

Thus, for any $P \in \mathrm{GL}_\gamma(\mathbb{F}_q)$, through simple matrix multiplication,

$$\begin{aligned}(\mathcal{P}(P)\mathscr{M})_{X,Y} &= \sum_Z \mathcal{P}(P)_{X,Z}\mathscr{M}_{Z,Y} \\ &= \mathscr{M}_{XP,Y} + 0 \\ &= \mathscr{M}(P)_{X,Y} \\ \implies (\mathcal{P}(P)\mathscr{M}) &= \mathscr{M}(P)\end{aligned}$$

From the following result, which appears in [4], along with the above conclusion, we can calculate the inverse of any P-Macwilliams Matrix.

Theorem 2.2.1.

$$\mathscr{M}^2 = \mathcal{P}(-I)$$

And thus, $\mathscr{M}^4 = I$

Proof.

$$\begin{aligned}q^\gamma \mathscr{M}^2_{X,Y} &= \sum_{Z\in\mathbb{F}_q^\gamma} \tau_X(Z)\tau_Z(Y) \\ &= \sum_{Z\in\mathbb{F}_q^\gamma} \tau_X(Z)\tau_Y(Z) \\ &= \sum_{Z\in\mathbb{F}_q^\gamma} \tau_{X+Y}(Z) \\ &= \begin{cases} q^\gamma & X = -Y \\ 0 & \text{else} \end{cases} \\ \implies \mathscr{M}^2 &= \mathcal{P}(-I)_{X,Y}\end{aligned}$$

Hence, $\mathscr{M}^4 = \mathcal{P}(-I)^2 = \mathcal{P}(I) = I$. □

From this, it is easy to see that the inverse of a MacWilliams Matrix is another MacWilliams Matrix.

As we shall see, the matrix $\mathscr{M}\Lambda(G)\mathscr{M}$ will prove to be useful in a MacWilliams type identity for convolutional codes. We define

$$f_{(X,Y)} := (\mathscr{M}\Lambda(G)\mathscr{M})_{X,Y} \tag{2.7}$$

The following two theorems from [4] help set up the framework for the main result

Theorem 2.2.2. *[4] For any (X,Y) in $\mathbb{F}^\gamma \times \mathbb{F}^\gamma$,*

$$f_{(X,Y)} = \begin{cases} 0 & (X,Y) \notin \Xi^\perp \\ q^{-\hat{r}} we(\mathscr{C}_C) & (X,Y) \in \mu^\perp \\ \frac{1}{q^\gamma(q-1)} \left(q \sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^\perp} \lambda_{Z_1,Z_2} - q^{\gamma-\hat{r}} we(\mathscr{C}_C) \right) & (X,Y) \in \Xi^\perp \backslash \mu^\perp \end{cases} \tag{2.8}$$

Proof. By definition (2.7),

$$\begin{aligned} f_{(X,Y)} &= (\mathscr{M}\Lambda(G)\mathscr{M})_{X,Y} \\ \implies q^\gamma f_{(X,Y)} &= \sum_{(Z_1,Z_2) \in \mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma} \tau_X(Z_1)\lambda_{Z_1,Z_2}\tau_{Z_2}(Y) \\ &= \sum_{(Z_1,Z_2) \in \mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma} \tau_X(Z_1)\lambda_{Z_1,Z_2}\tau_Y(Z_2) \\ &= \sum_{(Z_1,Z_2) \in \mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma} \tau_{(X,Y)}((Z_1,Z_2))\lambda_{Z_1,Z_2} \end{aligned}$$

Where, $\tau_{(X,Y)}$ is a character on $\mathbb{F}^{2\gamma}$. If $(X,Y) = (0,0)$, the character is trivial and we get

$$q^\gamma f_{(0,0)} = \sum_{(Z_1,Z_2) \in \mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma} \lambda_{Z_1,Z_2}$$

For $(X,Y) \neq (0,0)$, let $(V_1,V_2) \in \mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma$ such that $\mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma = (Sp((X,Y)))^\perp \oplus Sp((V_1,V_2))$ Using above to further simplify the expression for $f_{(X,Y)}$,

$$\begin{aligned} q^\gamma f_{(X,Y)} &= \sum_{\nu \in \mathbb{F}_q} \sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^\perp} \tau_{(X,Y)}((Z_1+\nu V_1, Z_2+\nu V_2))\lambda_{Z_1+\nu V_1, Z_2+\nu V_2} \\ &= \sum_{\nu \in \mathbb{F}_q} \tau_{(X,Y)}((\nu V_1, \nu V_2)) \sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^\perp} \lambda_{Z_1+\nu V_1, Z_2+\nu V_2} \\ = \sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^\perp} \lambda_{Z_1,Z_2} &+ \sum_{\nu \in \mathbb{F}_q^\times} \tau_{(X,Y)}((\nu V_1, \nu V_2)) \sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^\perp} \lambda_{Z_1+\nu V_1, Z_2+\nu V_2} \end{aligned}$$

Since the last term is summing over all (Z_1,Z_2) in a particular space, we could just write it as summing over $(\nu Z_1, \nu Z_2)$, which allows us to write $\lambda_{(Z_1+V_1,Z_2+V_2)}$ since for any $\nu \mathbb{F}_q^\times$, $\lambda_{(X,Y)} = \lambda_{(\nu X,\nu Y)}$. Thus,

$$q^\gamma f_{(X,Y)} = \sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^\perp} \lambda_{Z_1,Z_2} + \sum_{\nu \in \mathbb{F}_q^\times} \tau_{(X,Y)}((\nu V_1, \nu V_2)) \sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^\perp} \lambda_{Z_1+V_1, Z_2+V_2}$$

Since $(X,Y) \neq (0,0)$, the function $\nu \mapsto \tau_{(X,Y)}((\nu V_1, \nu V_2))$ is a valid non-trivial character on $\mathbb{F}_q$. We

now have

$$q^{\gamma} f_{(X,Y)} = \sum_{(Z_1,Z_2)\in\ (Sp((X,Y)))^{\perp}} \lambda_{Z_1,Z_2} - \sum_{(Z_1,Z_2)\in\ (Sp((X,Y)))^{\perp}} \lambda_{Z_1+V_1,Z_2+V_2} \tag{2.9}$$

$Sp((V_1,V_2))$ forms the complement of $(Sp((X,Y)))^{\perp}$ and therefore, the second term is just

$$\frac{1}{q-1} \sum_{(Z_1,Z_2)\notin\ (Sp((X,Y)))^{\perp}} \lambda_{Z_1,Z_2}$$

We conclude

$$q^{\gamma} f_{(X,Y)} = \sum_{(Z_1,Z_2)\in\ (Sp((X,Y)))^{\perp}} \lambda_{Z_1,Z_2} - \frac{1}{q-1} \sum_{(Z_1,Z_2)\notin\ (Sp((X,Y)))^{\perp}} \lambda_{Z_1,Z_2} \tag{2.10}$$

If $(X,Y) \notin \Xi^{\perp}$, $\Xi \not\subseteq (Sp((X,Y)))^{\perp}$, and thus in (2.9), we can choose (V_1,V_2) such that it lies in Ξ, and hence $\lambda_{Z_1+V_1,Z_2+V_2} = \lambda_{Z_1,Z_2}$ and hence for $(X,Y) \notin \Xi^{\perp}$,

$$f_{(X,Y)} = 0$$

If $(X,Y) \in \mu^{\perp}$, $\mu \subseteq (Sp((X,Y)))^{\perp}$ which means that none of the elements (Z_1,Z_2) not in $(Sp((X,Y)))^{\perp}$ are in μ and hence $\lambda_{Z_1,Z_2} = 0$. Therefore, from (2.10)

$$\begin{aligned} q^{\gamma} f_{(X,Y)} = \sum_{(Z_1,Z_2)\in\ (Sp((X,Y)))^{\perp}} \lambda_{Z_1,Z_2} - \frac{1}{q-1} \sum_{(Z_1,Z_2)\notin\ (Sp((X,Y)))^{\perp}} \lambda_{Z_1,Z_2} \\ = \sum_{(Z_1,Z_2)\in\mu} \lambda_{Z_1,Z_2} \\ \implies f_{(X,Y)} = q^{-\hat{r}} \mathrm{we}(\mathscr{C}_C) \end{aligned}$$

Finally, if $(X,Y) \in \Xi^{\perp} \backslash \mu^{\perp}$, from (2.10), we have,

$$f_{(X,Y)} = \frac{1}{q^{\gamma}(q-1)} \left((q-1) \sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^{\perp}} \lambda_{Z_1,Z_2} - \sum_{(Z_1,Z_2) \notin\ (Sp((X,Y)))^{\perp}} \lambda_{Z_1,Z_2} \right)$$

$$= \frac{1}{q^{\gamma}(q-1)} \left((q) \sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^{\perp}} \lambda_{Z_1,Z_2} - \left(\sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^{\perp}} \lambda_{Z_1,Z_2} + \sum_{(Z_1,Z_2) \notin\ (Sp((X,Y)))^{\perp}} \lambda_{Z_1,Z_2} \right) \right)$$

$$= \frac{1}{q^{\gamma}(q-1)} \left((q) \sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^{\perp}} \lambda_{Z_1,Z_2} - \sum_{(Z_1,Z_2) \in \mathbb{F}_q^{\gamma} \times \mathbb{F}_q^{\gamma}} \lambda_{Z_1,Z_2} \right)$$

$$= \frac{1}{q^{\gamma}(q-1)} \left((q) \sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^{\perp}} \lambda_{Z_1,Z_2} - q^{\gamma - \hat{r}} \mathrm{we}(\mathscr{C}_C) \right)$$

□

Let $(U,V) \in \mu^{\perp} \subseteq \Xi^{\perp}$, for the data as in the theorem above, in the first two cases for $f_{(X,Y)}$, $f_{(X+U,Y+V)} = f_{(X,Y)}$ immediately. As for the third case, we have

$$\sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^{\perp}} \lambda_{Z_1,Z_2} = \sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^{\perp} \cap \mu} \lambda_{Z_1,Z_2}$$

$$= \sum_{(Z_1,Z_2) \in\ (Sp((X+U,Y+V)))^{\perp} \cap \mu} \lambda_{Z_1,Z_2}$$

Since $(U,V) \in \mu^{\perp}$. Hence, we can conclude that for any $(U,V) \in \mu^{\perp}$,

$$f_{(X+U,Y+V)} = f_{(X,Y)} \tag{2.11}$$

We are now ready to apply the MacWilliams Transform on these polynomials, and this is valid since these are complex coefficient polynomials of degree at most n. Applying the transform, we get values related to the dual code $\hat{\mathscr{C}}$.

Theorem 2.2.3. *[4] For $(X,Y) \in \mathbb{F}_q^{\gamma} \times \mathbb{F}_q^{\gamma}$ and H is the MacWilliams Transform as in* (1.1)

$$q^{-k} H(f_{(X,Y)}) = \begin{cases} 0 & (X,Y) \notin \Xi^{\perp} \\ we(\hat{\mathscr{C}}_{const}) & (X,Y) \in \mu^{\perp} \\ \frac{1}{q-1} \left(we(Sp_{\mathbb{F}_q}(\hat{\mathscr{C}}_{const}, c(X,Y)) - we(\hat{\mathscr{C}}_{const})) \right) & (X,Y) \in \Xi^{\perp} \backslash \mu^{\perp} \end{cases} \tag{2.12}$$

Where $c(X,Y)$ is any element in $[\sigma((Sp_{\mathbb{F}_q}(X,Y)))^{\perp} \cap \mu^*) + \mathscr{C}_{const}]^{\perp} \backslash \hat{\mathscr{C}}_{const}$

Proof. The first two cases are immediately obvious from applying the MacWilliams Transform to (2.8). As for the third case, let (X, Y) lie in $\Xi^{\perp} \backslash \mu^{\perp}$.
Thus, $\Xi \subseteq Sp_{\mathbb{F}_q}(X,Y)^{\perp}$ and $\mu \not\subseteq Sp_{\mathbb{F}_q}(X,Y)^{\perp}$. This implies $\Xi \subseteq \mu \cap Sp_{\mathbb{F}_q}(X,Y)^{\perp}$. Since $\mu = \mu^* \oplus \Xi$, we can write

$$\mu \cap Sp_{\mathbb{F}_q}(X,Y)^{\perp} = (\mu^* \cap Sp_{\mathbb{F}_q}(X,Y)^{\perp}) \oplus \Xi$$

From the above discussion and the dimension of Ξ as calculated in Lemma 2.1.6,

$$\begin{aligned}
\sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^{\perp}} \lambda_{Z_1,Z_2} &= \sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^{\perp} \cap \mu} \lambda_{Z_1,Z_2} \\
&= (q^{\gamma - \hat{r}}) \sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^{\perp} \cap \mu^*} \lambda_{Z_1,Z_2} \\
&= (q^{\gamma - \hat{r}}) \sum_{(Z_1,Z_2) \in\ (Sp((X,Y)))^{\perp} \cap \mu^*} \mathrm{we}(\sigma(Z_1,Z_2) + \mathscr{C}_{\mathrm{const}}) \\
&= (q^{\gamma - \hat{r}}) \mathrm{we} \left(\bigcup_{(Z_1,Z_2) \in\ (Sp((X,Y)))^{\perp} \cap \mu^*} \sigma(Z_1,Z_2) + \mathscr{C}_{\mathrm{const}} \right)
\end{aligned}$$

Let

$$S(X,Y) := \bigcup_{(Z_1,Z_2) \in\ (Sp((X,Y)))^{\perp} \cap \mu^*} \sigma(Z_1,Z_2) + \mathscr{C}_{\mathrm{const}} \tag{2.13}$$

By definition of μ^* as any direct complement to Ξ in μ, we have $\ker \sigma \cap \mu^* = \{0\}$ since $\ker \sigma \subseteq \ker \varrho = \Xi$. Therefore,

$$\dim\{\sigma((Sp((X,Y)))^{\perp} \cap \mu^*)\} = \dim\{(Sp((X,Y)))^{\perp} \cap \mu^*\} = \dim \mu^* - 1$$

It is easy to see that

$$\sigma(\mu^*) \cap \mathscr{C}_{\mathrm{const}} = \{0\}$$

and from (2.4), $\dim \mu^* + \dim \Xi = \dim \mu$ implies that $\dim \mu^* = r + \hat{r}$ and we know from the definition of $\mathscr{C}_{\mathrm{const}}$ that $\dim \mathscr{C}_{\mathrm{const}} = k - r$.
Since $S(X,Y)$ can also be written as $((Sp((X,Y)))^{\perp} \cap \mu^*) + \mathscr{C}_{\mathrm{const}}$, we see that

$$\dim S(X,Y) = \dim \mu^* - 1 + \dim \mathscr{C}_{\mathrm{const}} = k + \hat{r} - 1$$

Therefore, $\dim(S(X,Y))^{\perp} = n + 1 - k - \hat{r} = \dim \hat{\mathscr{C}}_{\mathrm{const}} + 1$. We also observe that $S(X,Y) \subset \mathscr{C}_C$ which, upon taking the dual on either side becomes $\hat{\mathscr{C}}_{\mathrm{const}} \subset S(X,Y)^{\perp}$.
By the dimension and the inclusion, we can see that there exists some $c(X,Y) \in S(X,Y)^{\perp} \backslash \hat{\mathscr{C}}_{\mathrm{const}}$ such that

$$S(X,Y)^{\perp} = Sp_{\mathbb{F}_q}(\hat{\mathscr{C}}_{\mathrm{const}}, c(X,Y))$$

Now, applying the MacWilliams Transform to $f_{(X,Y)}$ in this case yields:

$$\begin{aligned} q^{-k}H(f_{(X,Y)}) &= q^{-k}H\left(\frac{1}{q^{\gamma}(q-1)}\left(q\sum_{(Z_1,Z_2)\in\mathbb{F}_q^{\gamma}\times\mathbb{F}_q^{\gamma}}\lambda_{Z_1,Z_2} - q^{\gamma-\hat{r}}\mathrm{we}(\mathscr{C}_C)\right)\right) \\ &= \frac{q^{-\hat{r}-k}}{q-1}\left(qH(\mathrm{we}(S(X,Y))) - H(\mathrm{we}(\mathscr{C}_C))\right) \\ &= \frac{q^{-\hat{r}-k}}{q-1}\left(q(q^{k+\hat{r}-1})\mathrm{we}((S(X,Y))^{\perp}) - q^{k+\hat{r}}\mathrm{we}(\hat{\mathscr{C}}_{\mathrm{const}})\right) \\ &= \frac{1}{q-1}\left(\mathrm{we}(Sp_{\mathbb{F}_q}(\hat{\mathscr{C}}_{\mathrm{const}}, c(X,Y)) - \mathrm{we}(\hat{\mathscr{C}}_{\mathrm{const}})\right) \end{aligned}$$

□

2.3 MacWilliams Identity for Convolutional Codes

To reiterate, for any object χ related to the code, let $\hat{\chi}$ denote the analogous object for the dual code. This is not to be confused with the objects $\chi^{\perp}$ when χ is a set associated with the code, as $\chi^{\perp}$ denotes the orthogonal to said set.

Define the $2\gamma \times 2\gamma$ matrix

$$W_0 := \begin{pmatrix} \hat{\mathcal{C}}\mathcal{C}^T & \hat{\mathcal{C}}\mathcal{D}^T\mathcal{B} \\ \hat{\mathcal{B}}\hat{\mathcal{D}}\mathcal{C}^T & 0 \end{pmatrix} \tag{2.14}$$

Some properties of this matrix[4]:

1. Rowspace of W_0 is a subset of $\Xi^{\perp}$. Let it be denoted by $\mathcal{W}_0$
2. $$\hat{\Xi} \oplus \hat{\mu}^{-} \subseteq \ker W_0 \tag{2.15}$$
3. $\{uW_0\} \cap \mu^{\perp} = \{0\}$
4. W_0 is injective on $\hat{\mu}^*$
5. $\mathrm{rank}(W_0) = r + \hat{r}$ and $\mathcal{W}_0 \oplus \mu^{\perp} = \Xi^{\perp}$

Lemma 2.3.1. *[4] For W_0 as defined above and any $(X,Y) \in \hat{\mu}$,*

$$\hat{\lambda}_{X,Y} = q^{-k}H(f_{(X,Y)W_0}) \tag{2.16}$$

Where H is the MacWilliams Transform used in the block code case.

Proof. For $(X, Y) = (0, 0)$, we can immediately see that the above holds by Theorem 2.2.8 and (2.2). So we assume that $(X, Y) \neq (0, 0)$.
We have $\hat{\mu} = \hat{\mu}^* \oplus \hat{\Xi}$, thus, every element of $\hat{\mu}$ can be written as $(X, Y) + (X' + Y')$, such that $(X, Y) \in \hat{\mu}^*$ and $(X' + Y') \in \hat{\Xi}$. From the invariance of λ(Lemma 2.1.8), we have

$$\hat{\lambda}_{X+X',Y+Y'} = \hat{\lambda}_{X,Y}$$

and from the invariance of f(2.11), we have

$$f_{(X,Y)W_0+(X',Y')W_0} = f_{(X,Y)W_0}$$

Due to these results, it suffices to prove this lemma for $(X, Y) \in \hat{\mu}^*$.
From the above discussion regarding the properties of W_0, we can see that $(X, Y)W_0 \in \Xi^{\perp} \backslash \mu^{\perp}$. Therefore, from Theorem 2.2.3,

$$q^{-k} H(f_{(X,Y)W_0}) = \frac{1}{q-1} \left(\mathrm{we}(Sp_{\mathbb{F}_q}(\hat{\mathscr{C}}_{\mathrm{const}}, c((X, Y)W_0)) - \mathrm{we}(\hat{\mathscr{C}}_{\mathrm{const}})) \right)$$

Claim: $\hat{\sigma}(X, Y)$ is a suitable candidate for $c((X, Y)W_0)$.
Clearly, $\hat{\sigma}(X, Y) \in \hat{\mathscr{C}_C} = \mathscr{C}_{\mathrm{const}}^{\perp}$, which follows from Lemma 2.1.3.
Since $(X, Y) \in \hat{\mu}^* \backslash 0$, $\hat{\sigma}(X, Y) \notin \hat{\mathscr{C}}_{\mathrm{const}}$.
Finally, from the properties of W_0, we see that for $(X', Y') \in (Sp_{\mathbb{F}_q}((X, Y)W_0)) \cap \mu^*$

$$\hat{\sigma}(X, Y) \in ((Sp_{\mathbb{F}_q}(X, Y)))^{\perp} \cap \mu^*)^{\perp}$$

And therefore,

$$\hat{\sigma}(X, Y) \in [\sigma((Sp_{\mathbb{F}_q}(X, Y)))^{\perp} \cap \mu^*) \cap \mathscr{C}_{\mathrm{const}}]^{\perp} / \hat{\mathscr{C}}_{\mathrm{const}}$$

and thus, along with (2.2) for the dual, we can conclude that

$$q^{-k} H(f_{(X,Y)W_0}) = \frac{1}{q-1} \left(\mathrm{we}(Sp_{\mathbb{F}_q}(\hat{\mathscr{C}}_{\mathrm{const}}, \hat{\sigma}(X, Y)) - \mathrm{we}(\hat{\mathscr{C}}_{\mathrm{const}})) \right) = \hat{\lambda}_{X,Y}$$

□

Definition 2.3.1. *Let $\mathcal{G}$ be a direct complement of $\Xi^{\perp}$ in $\mathbb{F}_q^{\gamma} \times \mathbb{F}_q^{\gamma}$, i.e.*

$$\mathcal{G} \oplus \Xi^{\perp} = \mathbb{F}_q^{\gamma} \times \mathbb{F}_q^{\gamma}$$

The following weak version of the MacWilliams identity was proven in [4]. Given the two ways of decomposing $\mathbb{F}_q^{\gamma} \times \mathbb{F}_q^{\gamma}$ and the dimensions of the components, one can set up a diagram as follows where each arrow's label is a function mapping the space near the tail end of the arrow to the space

near the head of the arrow.

$$\begin{array}{ccccccc} \mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma = & & \hat{\mu}^* & \oplus & \hat{\Xi} & \oplus & \hat{\mu}^- \\ \downarrow g & & \downarrow g_0 & & \downarrow g_1 & & \downarrow g_2 \\ \mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma = & & \mathcal{W}_0 & \oplus & \mu^\perp & \oplus & \mathcal{G} \end{array} \tag{2.17}$$

Theorem 2.3.2. *Assume information as in* (2.17)*, and let the isomorphism g_0 be the one induced by the matrix W_0 as defined above. Fix any isomorphisms g_1 and g_2, which exist due to the dimensions being the same over the same base field, and let $g = \bigoplus_{i=0}^{2} g_i$ be the automorphism on $\mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma$. Then,*

$$\hat{\lambda}_{X,Y} = q^{-k}(H(\mathscr{M}\Lambda\mathscr{M}))_{g(X,Y)}$$

For all $(X, Y) \in \mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma$.

Proof. By definition of f,

$$f_{(X,Y)} = (\mathscr{M}\Lambda\mathscr{M})_{(X,Y)}$$

There are the following three cases to consider:

1. $(X, Y) \in \hat{\mu}^-$, then $\hat{\lambda}_{X,Y} = 0$ and $g(X, Y) \in \mathcal{G}$, more importantly, $g(X, Y) \notin \Xi^\perp$, hence $q^{-k}H(f_{g(X,Y)}) = 0$ by (2.12)
2. $(X, Y) \in \hat{\Xi}$, then $\hat{\sigma}(X, Y) \in \hat{\mathscr{C}}_{\text{const}}$, which implies $\hat{\lambda}_{X,Y} = \text{we}(\hat{\sigma}(X, Y) + \hat{\mathscr{C}}_{\text{const}}) = \text{we}(\hat{\mathscr{C}}_{\text{const}})$. We also have $g(X, Y) \in \mu^\perp$, and thus by (2.12), $q^{-k}H(f_{g(X,Y)}) = \text{we}(\hat{\mathscr{C}}_{\text{const}}) = \hat{\lambda}_{X,Y}$.
3. Finally, for $(X, Y) \in \hat{\mu}^*$, the result follows directly from (2.16).

□

We use this as a starting point to prove a stronger result, as in [3].

Since the weaker version of the identity leaves the isomorphisms g_1 and g_2 arbitrary, as well as the spaces μ^* and $\mathcal{G}$, we try to fix them to get an expression for the identity which is easy to calculate.
Define the projection mapping

$$\pi_1 : \mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma \to \mathbb{F}_q^\gamma : (X, Y) \mapsto X$$

Lemma 2.3.3. *The map π_1 restricted to Ξ is injective.*

Proof. Let $(X, Y) \in \Xi$. If $(X, Y) \in \ker \pi_1$, then $X = 0$. Since (X, Y) lies in Ξ, we have $\sigma(X, Y)$ lies in $\mathscr{C}_{\text{const}}$. Therefore, $Y\mathcal{B}^T\mathcal{D} \in \mathscr{C}_{\text{const}}$. For some u, $Y = u\mathcal{B}$. Thus, we have $u\mathcal{B}\mathcal{B}^T\mathcal{D} \in \mathscr{C}_{\text{const}}$. Now, $\mathscr{C}_{\text{const}} = (\ker \mathcal{B})\mathcal{D}$ along with the previous discussion implies that $u\mathcal{B}\mathcal{B}^T$ lies in $\ker \mathcal{B}$.
Since

$$\mathcal{B}^T\mathcal{B}_{i,j} = \begin{cases} 1 & i = j \in \mathcal{I} \\ 0 & \text{else} \end{cases}$$

We can see that $\mathcal{B}\mathcal{B}^T\mathcal{B} = \mathcal{B}$. Hence,

$$u\mathcal{B} = u\mathcal{B}\mathcal{B}^T\mathcal{B} = (u\mathcal{B}\mathcal{B}^T)\mathcal{B} = 0$$

Thus, $(X, Y) = (0, 0)$ Since (X, Y) is arbitrary, we can conclude that $\ker \pi_1|_\Xi = \{0\}$, i.e. $\pi_1|_\Xi$ is injective. □

A few more useful properties:

Proposition 2.3.4. *For the controller canonical forms of the code and its dual,*

1. *rank* $(\mathcal{C}\hat{\mathcal{D}}^T\hat{\mathcal{B}}) = \hat{r}$
2. $\ker(\mathcal{C}\hat{\mathcal{D}}^T\hat{\mathcal{B}}) = \pi_1(\Xi)$

Corollary 2.3.5. *Let* $E := \ker \mathcal{C}\hat{\mathcal{D}}^T\hat{\mathcal{B}}$*, then the function*

$$\psi : E \to \mathbb{F}_q^\delta : X \mapsto Y \text{ such that } (X, Y) \in \Xi$$

is a well-defined linear injective map. Moreover, E doesn't contain a non-zero subspace that is invariant under ψ

Define the Matrices:

$$J_0 := \mathcal{B}^T\mathcal{D}; J_i := \mathcal{B}^T\mathcal{B}\mathcal{A}^{i-1}\mathcal{C} i \geq 1$$

and analogously, $\hat{J}$. We also define:

$$K := \sum_{m\geq 2}\sum_{i=1}^{m-1}\sum_{j=0}^{i-1}(\hat{\mathcal{A}}^T)^{i-1}\hat{J}_j J_{m-j}^T A^{m-i+1}$$

And analogously, $\hat{K}$ associated with the dual code. The following straightforward properties of these matrices are proved in [3]

1. $K + \hat{K}^T = -\hat{\mathcal{C}}\mathcal{C}^T$

2. $\hat{\mathscr{C}} J_0^T + \hat{J}_0 \mathscr{C}^T = K\mathcal{A} + \hat{\mathcal{A}}^T \hat{K}^T$

3. $K\mathcal{A}\mathcal{A}^T = K$

We define the matrices

$$W_1 := \begin{pmatrix} K & -K\mathcal{A} \\ 0 & 0 \end{pmatrix} \tag{2.18}$$

$$W_2 = \begin{pmatrix} \hat{K}^T & 0 \\ -\hat{\mathcal{A}}^T K^T & 0 \end{pmatrix} \tag{2.19}$$

Moreover,

$$\begin{aligned} W_1 \begin{pmatrix} I & 0 \\ \mathcal{A}^T & \mathcal{B}^T \end{pmatrix} &= \begin{pmatrix} K & -K\mathcal{A} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I & 0 \\ \mathcal{A}^T & \mathcal{B}^T \end{pmatrix} \\ &= \begin{pmatrix} K - K\mathcal{A}\mathcal{A}^T & -K(\mathcal{A}\mathcal{B}^T) \\ 0 & 0 \end{pmatrix} \\ &\implies W_1 \begin{pmatrix} I & 0 \\ \mathcal{A}^T & \mathcal{B}^T \end{pmatrix} = 0 \end{aligned}$$

From this, we can conclude that

$$\begin{aligned} \{uW_1\} &\subseteq \ker \begin{pmatrix} I & \mathcal{A} \\ 0 & \mathcal{B} \end{pmatrix} \\ &\implies \{uW_1\} \subseteq \mu^\perp \end{aligned} \tag{2.20}$$

Which upon taking the orthogonals and writing in terms of the dual implies

$$\hat{\mu} \subseteq \ker W_2 \tag{2.21}$$

From the definitions and properties above, we can see that $W_2 = \hat{W}_1^T$ and that $W = W_0 + W_1 + W_2$ is

$$\begin{pmatrix} 0 & P \\ -P & 0 \end{pmatrix}$$

Where $P := \hat{\mathcal{C}} J_0^T - K\mathcal{A}$. We claim that P lies in $GL_\gamma(q)$ and that the functions g_i induced by W_i have the desired properties for the diagram and that for all $(X, Y) \in \mathbb{F}_q^\gamma$

$$\hat{\Lambda} = q^{-k} H((\mathscr{M} \Lambda^T \mathscr{M}^{-1})_{XP,YP})$$

Thus, we can conclude that

$$\bar{\Lambda}(\hat{\mathscr{C}}) = q^{-k} H(\mathscr{M} \bar{\Lambda}(\mathscr{C})^T \mathscr{M}^{-1})$$

Where $\bar{\Lambda}(\mathscr{C})$ is the Generalised Adjacency Matrix for the code $\mathscr{C}$.

Proof of $P \in GL_\gamma(q)$. [3] For any $X \in \ker P$,

$$\begin{aligned} XP &= 0 \\ \implies X\hat{\mathcal{C}}J_0^T &= XK\mathcal{A} \\ \implies XK\mathcal{A} &= X\hat{\mathcal{C}}\mathcal{D}^T\mathcal{B} \end{aligned}$$

Therefore, $XK\mathcal{A} \in \{u\mathcal{A}\} \cap \{u\mathcal{B}\} = \{0\}$. This implies, $X\hat{\mathcal{C}}\mathcal{D}^T\mathcal{B} = 0$, i.e. $X \in \hat{E}$. Therefore, $\ker P \subseteq \hat{E}$.
Thus, $(X, \hat{\psi}(X)) \in \hat{\Xi}$.
From (2.21), $\hat{\mu} \subseteq \ker W_2$ and $\hat{\Xi} \subseteq \ker W_0$, $(X, \hat{\psi}(X)) \in \ker W_0 \cap \ker W_2$.
Moreover, $(X, \hat{\psi}(X))W_1 = (XK, -XK\mathcal{A}) = (XK, 0)$. Since $K\mathcal{A}\mathcal{A}^T = K$ and $XK\mathcal{A}\mathcal{A}^T = (XK\mathcal{A})\mathcal{A}^T = 0$, $XK = 0$, i.e. $(X, \hat{\psi}(X)) \in \ker W_1$.
Therefore,

$$(X, \hat{\psi}(X)) \in \ker W$$

By the definition of W and P, we can conclude that $\hat{\psi}(X) \in \ker P$, which implies that $\ker P$ is invariant under $\hat{\psi}$, which means that it is either $\hat{E}$ or $\{0\}$. It can't be the whole space since P is non-zero, hence $\ker P = \{0\}$, thus $P \in GL_\gamma(q)$ □

The following properties of the map g, along with the fact that $\ker W_0 = \hat{\Xi} \oplus \hat{\mu}^-$ allow for the description of the MacWilliams Identity as above.

Proposition 2.3.6. *[3] Let* $\hat{\mu}^* := \ker W_1 \cap \hat{\mu}$ *and* $\mathcal{G} := \{uW_2\}$*, then:*

1. $\ker W_1 = \hat{\mu}^* \oplus \hat{\mu}^-$
2. $\ker W_2 = \hat{\mu}$
3. $\{uW_1\} = \mu^\perp$ *and* $\mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma = \Xi^\perp \oplus \mathcal{G}$

Proof.

1. The sum on the right is clearly a direct sum by definitions of the sets involved. Since

$$\hat{\mu}^* \subseteq \ker W_1$$

and looking at the structure of $\hat{\mu}^-$ from Definition 2.1.4 and W_1 from (2.18),

$$\hat{\mu}^- \subseteq \ker W_1$$

Hence, $\hat{\mu}^* \oplus \hat{\mu}^-$ is contained in $\ker W_1$.
Now, we have that $\hat{\Xi} \subseteq \ker W_0$ from (2.15) and from (2.21) $\hat{\Xi} \subseteq \hat{\mu} \subseteq \ker uW_2$,

$$\hat{\Xi} \subseteq \ker W_0 \cap \ker W_2$$

Hence, for any $(X, Y) \in \hat{\Xi}$, $(X, Y)W = (X, Y)W_1$. Since W is invertible (follows from the invertibility of P, which is shown above), we have to have $(X, Y)W_1 \neq 0$ for any non-zero (X, Y) in $\hat{\Xi}$, i.e. $\hat{\Xi} \cap \ker W_1 = \{0\}$, which implies that $\hat{\Xi} \subseteq \{uW_1\}$. Hence,

$$\gamma - r = \dim \hat{\Xi} \leq \text{rank } W_1 \leq \dim \mu^{\perp} = \gamma - r$$

Therefore, rank$W_1 = \gamma - r$ and $\{uW_1\} = \mu^{\perp}$.

$$\dim \hat{\mu}^* = \dim \ker W_1 + \dim \hat{\mu} - 2\gamma = r + \hat{r} = \dim \hat{\mu} - \dim \hat{\Xi}$$

This along with $\hat{\mu}^* \cap \hat{\Xi} = \{0\}$ and the fact $\hat{\mu}^* \subseteq \hat{\mu}$ as well $\hat{\Xi} \subset \hat{\mu}$, we see that $\hat{\mu} = \hat{\Xi} \oplus \hat{\mu}^*$.
This lets us conclude that $\mathbb{F}_q^{\gamma} \times \mathbb{F}_q^{\gamma}$ decomposes as in diagram (2.17).
Now,

$$\dim \ker W_1 = \gamma + r = 2\gamma - \dim \hat{\Xi} = \dim(\hat{\mu}^* \oplus \hat{\mu}^-)$$

Which allows us to conclude that

$$\ker W_1 = \hat{\mu}^* \oplus \hat{\mu}^-$$

2. $W_2 = \hat{W}_1^T$, hence rank$W_2 = \gamma - \hat{r}$. Thus, $\dim \ker W_2 = \gamma + hatr = \dim \hat{\mu}$. Since we know that $\ker W_2 \subset \hat{\mu}$, we can then conclude that $\ker W_2 = \hat{\mu}$.

3. The first part is shown in proof 1. As for the other part, one can observe that $\dim \mathcal{G} = \gamma - \hat{r}$ and $\dim \Xi^{\perp} = \gamma + \hat{r}$.
Assume that $(X, Y)W_2 \in \Xi^{\perp}$ for some (X, Y). Since we don't want $(X, Y)W_2$ to be 0, we can assume that $(X, Y) \in \hat{\mu}^-$, hence $(X, Y)W = (X, Y)W_2$.
$\Xi^{\perp} = \mu^{\perp} \oplus \mathcal{W}_0 = \{uW_1\} \oplus \mathcal{W}_0$, therefore for some (X_1, Y_1) and (X_0, Y_0), we have

$$(X, Y)W_2 = (X_1, Y_1)W_1 + (X_0, Y_0)W_0)$$

Again, we can assume that $(X_0, Y_0) \notin \ker W_0$ which implies $(X_0, Y_0) \in \hat{\mu}^*$ and $(X_1, Y_1) \notin$

$\ker W_1$ which implies $(X_1, Y_1) \in \hat{\Xi}$

Therefore,

$$(X, Y)W = ((X_0, Y_0) + (X_1, Y_1))W$$

Which implies $(X, Y) = (X_0, Y_0) + (X_1, Y_1) \in \hat{\mu}^- \cap \hat{\mu}$, thus $(X, Y)W_2 = 0$, which lets us conclude that $\mathcal{G} \cap \Xi^\perp = \{0\}$, and $\mathbb{F}_q^\gamma \times \mathbb{F}_q^\gamma = \Xi^\perp \oplus \mathcal{G}$.

□

This proposition fixes μ^* and $\mathcal{G}$ instead of just leaving them as any complement of previously defined sets as in Proposition 2.1.7 and Definition 2.3.1.

Chapter 3

Convolutional LCDs and other Hull Properties

Definition 3.0.1. *The (Euclidean) Hull of a code is the intersection of the code with its dual. Let it be denoted by H, then*

$$\mathscr{H} = \mathscr{C} \cap \mathscr{C}^{\perp}$$

Analogously to Massey's definition of LCD codes in the block code case [9], one can define CLCD codes as follows:

Definition 3.0.2. *A convolutional code is said to be a convolutional linear code with complementary dual if the intersection of the code with its dual (Hull of the Code) is trivial, i.e.* $\{0\}$

Definition 3.0.3. *A convolutional code* $\mathscr{C}$ *is said to be self-orthogonal if* $\mathscr{C}$ *is a subset of* $\mathscr{C}^{\perp}$*. It is said to be self-dual if* $\mathscr{C} = \mathscr{C}^{\perp}$*. In either of these cases, the hull coincides with the code.*

The alternative characterisation provided by Massey in [9] is also valid for the convolutional case without any change in the argument.

Proposition 3.0.1. *For any Convolutional Code with a Complementary Dual, and G any non-catastrophic Generator for* $\mathscr{C}$*, the matrix* GG^T *is of full rank.*

Proposition 3.0.2. *A convolutional code, $\mathscr{C}$ with generator matrix G is self orthogonal if and only if*

$$GG^T = 0$$

Proof. ($\implies$) Obvious since each row is in the code and its dual.
($\impliedby$) Any codeword c is a linear combination of the rows of G. Since we define the dual using a bilinear form, (c, c) can be written as a linear combination of $\{(g_i(z), g_j(z)\}_{i,j}$ where $g_i(z)$ denotes the rows of G. Since GG^T is 0, all $\{(g_i(z), g_j(z)\}_{i,j}$ are 0, and we get that for any code word c, $(c, c) = 0$. Thus, $\mathscr{C}$ is self-orthogonal. □

Theorem 3.0.3. *For any k-dimensional code $\mathscr{C}$, let G be any $k \times n$ generator matrix of the code and $\mathscr{H}$ be the hull of the code. Then,*

$$\dim\{\mathscr{H}\} = k - \text{rank}\{GG^T\}$$

Proof. Let the dimension of the code $\mathscr{C}$ be m. Then $m \leq k$. Since the hull is the intersection of 2 vector spaces, it is a vector space over $\mathbb{F}(z)^n$. Therefore, there exists a basis

$$\{l_i(z)\}_{i=1}^m$$

for the Hull consisting of m linearly independent vectors. Define $L(Z)$ as:

$$L(z) = \begin{pmatrix} -l_1(z)- \\ \vdots \\ -l_m(z)- \end{pmatrix}$$

Therefore,

$$L(z)L^T(z)_{i,j} = (l_i(z), l_j(z)) = 0$$

as each row lies in the dual of the code and the code itself. Since the hull is a subspace of the code, we can extend the basis of the hull to a basis of the code, say $\{l_i(z)\}_{i=1}^k$. Define $G(z)$ as

$$G(z) = \begin{pmatrix} -l_1(z)- \\ \vdots \\ -l_k(z)- \end{pmatrix}$$

Thus, G is a generator matrix for the code. Looking at the matrix $G(z)G^T(z)$,

$$GG^T = \begin{pmatrix} 0_{m\times m} & 0 \\ 0 & M \end{pmatrix}$$

Where M is such that none of the rows are identically 0.
Consider the code generated by the basis elements that do not lie in the hull. Since any vector in this code is a linear combination of the non-hull basis elements, one can see that none of the vectors in this code annihilate the code and, thus, are not in the dual of this code. Therefore, this code has to be an LCD code. Hence, for any generator matrix K of this code, KK^T is non-singular, and hence so is M. Thus, the rank of the matrix GG^T is $k-m$. Any generator matrix of $\mathscr{C}$ is then of the form $G' = WG$, where W is an invertible $k \times k$ matrix. Hence,

$$G'(G')^T = WGG^TW^T$$

Since W and W^T are full rank matrices, $\text{rank}\big(G'(G')^T\big) = \text{rank}\big(GG^T\big) = k-m$.

□

From the above, we can conclude that the dimension of the hull is exactly the nullity of the matrix GG^T where G is any generator matrix for the code. A similar result with regards to the block code case can be found in [12]

We continue to show a method of construction of Convolutional LCD codes, which is based on a method for the construction of Block LCD codes in [9]:
For this, we first need Lagrange's 4 square theorem, that is, any positive integer p can be written as the sum of 4 squares, allowing for repetitions and 0s,

$$p = \sum_{i=1}^{4} \alpha_i^2$$

For a Proof of the same, please refer to chapter 20 in [5].

Theorem 3.0.4. *For a code $\mathscr{C}$ over $\mathbb{F}_q(z)$, with a generator matrix of the form*

$$G = \begin{pmatrix} I_k & P \end{pmatrix}$$

such that the characteristic of the field is $p = \sum_{i=1}^{4} \alpha_i^2$*, the matrix*

$$G' = \begin{pmatrix} I_k & \alpha_1 P & \alpha_2 P & \alpha_3 P & \alpha_4 P \end{pmatrix}$$

generates a convolutional LCD code.

Proof. The fact that G' generates a convolutional LCD code is immediately obvious from the fact that $G'(G')^T = I_k + \sum_{i=1}^{4} \alpha_i^2 (PP^T) = I_k + p(PP^T) = I_k$ □

We end with an example of Convolutional LCD codes:
Consider the matrix

$$G = \begin{pmatrix} 1 & 0 & z \\ 0 & 1 & 1+z^2 \end{pmatrix}$$

over the field $\mathbb{F}_2(z)$.
Then $2 = 1^2 + 1^2 + 0 + 0$, hence the matrix

$$G' = \begin{pmatrix} 1 & 0 & z & z \\ 0 & 1 & 1+z^2 & 1+z^2 \end{pmatrix}$$

generates a $(4, 2)$-Convolutional code with a complementary dual.

Chapter 4

Conclusion

In this thesis, we have looked at the definition of Convolutional Codes and associated properties. We then turned our focus onto the properties of the dual of a convolutional code, in particular, the MacWilliams Identity for convolutional codes and properties of the hull of the code.

Further work includes examining the MacWilliams Identity for the trellis duality. The presence of a MacWilliams Identity also allows to investigate the self-duality through the use of Invariant Theory.

The translation of a few fundamental properties of Hulls and LCDs into the convolutional case provides hope that we can expect a similar "extension" of other properties into the convolutional case.

Bibliography

[1] Khaled A. S. Abdel-Ghaffar. "On Unit Constraint-Length Convolutional Codes". In: *IEEE Transactions on Information Theory* 38.1 (1992), pp. 200–204. DOI: 10.1109/18.108260. URL: https://ieeexplore.ieee.org/document/108260.

[2] Heide Gluesing-Luerssen. "On the weight distribution of convolutional codes". In: *Linear Algebra and its Applications* 408 (2005), pp. 298–326. ISSN: 0024-3795. DOI: 10.1016/j.laa.2005.06.023.

[3] Heide Gluesing-Luerssen and Gert Schneider. "A MacWilliams Identity for Convolutional Codes: The General Case". In: *Information Theory, IEEE Transactions on* 55 (Aug. 2009), pp. 2920–2930. DOI: 10.1109/TIT.2009.2021302.

[4] Heide Gluesing-Luerssen and Gert Schneider. "On the MacWilliams Identity for Convolutional Codes". In: *IEEE Transactions on Information Theory* 54.4 (2008), pp. 1536–1547. DOI: 10.1109/TIT.2008.917664.

[5] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers*. 6th. Oxford University Press, 2008. ISBN: 978-0199219865.

[6] R. Johannesson, P. Stahl, and E. Wittenmark. "A note on type II convolutional codes". In: *IEEE Transactions on Information Theory* 46.4 (2000), pp. 1510–1514. DOI: 10.1109/18.850684.

[7] Jacobus H. van Lint. *Introduction to Coding Theory*. Third Revised and Expanded Edition. Vol. 86. Graduate Texts in Mathematics. Springer-Verlag Berlin Heidelberg, 1999. ISBN: 978-3-642-63653-0. DOI: 10.1007/978-3-642-58575-3.

[8] Jessie Macwilliams. "A theorem on the distribution of weights in a systematic code". In: *The Bell System Technical Journal* 42.1 (1963), pp. 79–94. DOI: 10.1002/j.1538-7305.1963.tb04003.x.

[9] James L. Massey. "Linear Codes with Complementary Duals". In: *Discrete Mathematics* 106/107 (1992), pp. 337–342. DOI: 10.1016/0012-365X(92)90563-U.

[10] Robert McEliece. "The Algebraic Theory of Convolutional Codes". In: *Handbook of Coding Theory Vol. I.* Ed. by V. Pless, W.C. Huffman, and R.A. Brualdi. Amsterdam: Elsevier Science, 1998, pp. 1065–1138. ISBN: 9780444814722.

[11] J. Shearer and R. McEliece. "There is no MacWilliams identity for convolutional codes (Corresp.)" In: *IEEE Transactions on Information Theory* 23.6 (1977), pp. 775–776. DOI: 10.1109/TIT.1977.1055785.

[12] Satanan Thipworawimon and Somphong Jitman. "Hulls of linear codes revisited with applications". In: *Journal of Applied Mathematics and Computing* 62 (2020), pp. 325–340. DOI: 10.1007/s12190-019-01286-7.