On the Hamming distance of linear codes over a finite chain ring *

Graham H. Norton Ana Salagean
Algebraic Coding Research Group
Centre for Communications Research
University of Bristol, U.K.

July 19, 2001

Abstract

Let R be a finite chain ring (e.g. a Galois ring), K its residue field and C a linear code over R. We prove that $\mathrm{d}(C)$, the Hamming distance of C, is $\mathrm{d}(\overline{(C:\alpha)})$, where $(C:\alpha)$ is a submodule quotient, α is a certain element of R and $\overline{}$ denotes projection to K. These two codes also have the same set of minimal codeword supports. We explicitly construct a generator matrix/polynomial of $\overline{(C:\alpha)}$ from the generator matrix/polynomials of C. We show that in general $\mathrm{d}(C) \leq \mathrm{d}(\overline{C})$ with equality for free codes (i.e. for free R-submodules of R^n) and in particular for Hensel lifts of cyclic codes over K. Most of the codes over rings described in the literature fall into this class.

We characterise MDS codes over R and prove several analogues of properties of MDS codes over finite fields. We compute the Hamming weight enumerator of a free MDS code over R.

Keywords: Finite chain ring, Galois ring, Hamming distance, MDS code.

1 Introduction

It is now known that important families of binary non-linear codes are in fact images under a Gray map of linear codes over Z_4; see [6] and the references cited there. Consequently codes over finite rings have received renewed attention in the recent literature. For codes over Z_4, it is usually the Lee distance which is studied, due to the fact that it coincides with the Hamming distance of the image of the code under a Gray map.

However, the Hamming distance of (linear) codes over finite rings is still important for a number of reasons. For codes over Z_{2^a} with $a > 2$ the Lee distance is no longer equal to the Hamming distance of the image of the code under a Gray map. (For example the elements of Z_8 have Lee weights ranging from 0 to 4, whereas their images under any Gray map are elements of Z_2^3 which can have Hamming weight at most 3.) Consequently it is not clear which metric is the most appropriate in this case. Most of the well-known algebraic decoding algorithms for codes over finite fields use Hamming distance. Some of these algorithms can be generalised to codes over finite rings. For example analogues of Berlekamp-Massey algorithm were devised for Z_m in [17], for Galois rings in [7] and more generally for any finite chain ring in [12, 13]. There are many results on the exact

*Research supported by the U.K. Engineering and Physical Sciences Research Council under Grant L07680. A preliminary version of this paper was presented at the Workshop on Coding and Cryptography in Paris, January 1999. Current addresses: G.H.Norton, Dept. Mathematics, University of Queensland, Brisbane 4072; ghn@maths.uq.edu.au; A. Salagean, Dept. Mathematics, Nottingham Trent University, UK; ana.salagean@ntu.ac.uk.

value or lower bounds for the Hamming distance of codes over finite fields. Thus it is useful to have a simple mechanism to transfer all these results to codes over finite rings. Finally, let us note that the Hamming distance is obviously a lower bound for the Lee distance of the code.

We work with codes over finite chain rings. A finite chain ring is a finite ring whose ideals can be linearly ordered by inclusion or equivalently, a finite local ring with principal maximal ideal. Examples of finite chain rings are Galois rings, and in particular Z_{p^a} where p is a prime and $a \geq 1$. Section 2 reviews finite chain and Galois rings and recalls several basic results from [14]. Galois rings are a natural setting for Reed-Solomon and generalised Reed-Muller codes. BCH codes can also be defined over finite chain rings ([13]), in analogy to BCH codes over Galois fields ([8, Chapter 7]).

The results of this paper (in particular Sections 4 and 5) depend on the structure theorems for linear and cyclic codes over a finite chain ring which are proved in [14]. These structure theorems generalise and extend results of [4] and are recalled in Subsection 3.2 for the convenience of the reader.

Let R be a finite chain ring, K its residue field, γ a fixed generator of the maximal ideal of R and ν the nilpotency index of γ. We put $\alpha = \gamma^{\nu-1}$. The canonical projections from $R[X]$ to $K[X]$ and from R^n to K^n will be denoted by $\overline{}$.

The main result of the paper is in Subsection 4.1, where we show that the (Hamming) distance of a linear code C over R is equal to the distance of $\overline{(C:\alpha)}$, where for $r \in R$, $(C:r)$ is the submodule quotient $\{e \in R^n |\ re \in C\}$. These two codes also have the same set of minimal codeword supports. We show that in general the distance of a linear code C over R is at most the distance of $\overline{C}$. Hence we cannot increase distance by working over finite chain rings rather than over finite fields. More precisely, for a given length and distance, the best rate of linear codes over R is the same as the best rate of linear codes over K; see Corollary 4.7. For free codes (i.e. codes which are free R-submodules of R^n) the distance of C is the same as the distance of $\overline{C}$. In particular, the (extended) Hensel lift of a cyclic code has the same distance as the original code over the finite field K. Hence the classical BCH, Hartmann-Tzeng, Roos etc. bounds for cyclic codes over a finite field also hold for their Hensel lifts. The BCH bound was stated in [18, Theorem 4] with an incorrect proof; see Remark 4.4(ii).

In Subsection 4.2 we construct a generator matrix/polynomial for the code $\overline{(C:\alpha)}$, given a generator matrix/polynomials of C. In Subsection 4.3 we examine a number of codes over Galois rings described in the literature and apply our results to either determine or give lower bounds for their distance.

Finally, Section 5 deals with MDS codes over a finite chain ring. We characterise these codes: a code C over R is MDS if and only if $\overline{(C:\alpha)}$ is an MDS code over K. For free codes, this means that C is MDS if and only if $\overline{C}$ is MDS. We prove a number of properties of MDS codes over R analogous to properties of MDS codes over finite fields. We determine the weight enumerator of a free MDS code.

When R is a Galois field our results are either straightforward or reduce to classical results.

2 Preliminaries

2.1 Finite chain rings and Galois rings

We begin with the definition and some properties of finite chain rings and continue with Galois rings, following mainly [10]. For more details and proofs we refer the reader to [14].

DEFINITION 2.1 *A finite commutative ring with* $1 \neq 0$ *is called a* finite chain ring *if its ideals are linearly ordered by inclusion.*

A simple example of a finite chain ring is the ring Z_{p^a} of integers modulo p^a, for some prime p and $a \geq 1$. A finite chain ring is a local ring. It is well known, and not difficult to prove, that a ring is a finite chain ring if and only if it is a finite local principal ideal ring. Let γ be a fixed generator of the maximal ideal of R. Then γ is nilpotent and let ν be its nilpotency index i.e. the smallest positive integer such that $\gamma^\nu = 0$. Denote by K the residue field $R/\gamma R$, which is finite. The cardinality of R is $|R| = |K|^\nu$; see for example [14, Lemma 2.4]. *Throughout this paper, R denotes a finite chain ring with $1 \neq 0$, γ a fixed generator of the maximal ideal of R, ν the nilpotency index of γ and K the residue field of R. We set $\alpha = \gamma^{\nu-1}$.*

The ideals of R are $\gamma^i R$, $i = 0, \ldots, \nu$. All the elements of the maximal ideal γR are zero-divisors and the elements of $R \setminus \gamma R$ are units. There is a form of unique factorisation in R:

LEMMA 2.2 ([10, p. 340]) *For any $r \in R \setminus \{0\}$ there is a unique integer i, $0 \leq i < \nu$ such that $r = u\gamma^i$, with u a unit. The unit u is unique modulo $\gamma^{\nu-i}$ only.*

The following result will be used throughout the paper:

COROLLARY 2.3 *If $1 \leq i < j \leq \nu$ and $\gamma^i c \in \gamma^j R$, then $c \in \gamma^{j-i} R$. In particular, if $\gamma^i c = 0$ then $c \in \gamma^{\nu-i} R$.*

There is a canonical projection homomorphism from R to K which extends naturally to a projection $R[X] \to K[X]$; for any $f \in R[X]$ we denote by $\overline{f}$ its image under this projection; also, for a set $C \subseteq R[X]$ we define $\overline{C} = \{\overline{f} |\ f \in C\}$.

For any $l \geq 1$, one can construct a Galois extension ring of Z_{p^a} as $GR(p^a, l) = Z_{p^a}[X]/(f)$ with p a prime number, $a \geq 1$ and $f \in Z_{p^a}[X]$ a monic basic irreducible polynomial of degree l. (Recall that a non-unit $f \in R[X]$ is called *basic irreducible* if f and $\overline{f}$ are irreducible.) There are basic irreducible polynomials of degree l for all $l \geq 1$, see e.g. [14, Lemma 2.5]. The ring $GR(p^a, l)$ is called a Galois ring. For $l = 1$ we obtain the ring Z_{p^a}. For $a = 1$ we obtain the finite field with p^l elements, $GF(p^l)$. For any multiple m of l there is an inclusion homomorphism $GR(p^a, l) \subseteq GR(p^a, m)$.

A Galois ring $GR(p^a, l)$ is a finite chain ring. We fix p as the generator of its maximal ideal. The nilpotency index of p is a and the residue field is $GF(p^l)$. In fact, any finite chain ring is a certain homomorphic image of a polynomial ring $GR(p^a, l)[X]$, see [10, Theorem XVII.5].

For any two elements a and b of a ring, we will write $a|b$ for "a divides b".

We also extend the projection of R to K to a projection of R^n to K^n. For any element $c \in R^n$ we denote by $\overline{c}$ its image under this projection. For a set $C \subseteq R^n$, we define $\overline{C} = \{\overline{c} |\ c \in C\}$.

For any constant $r \in R$ and any $c \in R^n$ we denote by rc the usual multiplication of a vector by a scalar. Also, for a set $C \subseteq R^n$ we write rC for the set $\{rc |\ c \in C\}$. We will say that a vector $c \in R^n$ is divisible by a constant $r \in R$, and write $r|c$, if all entries of c are divisible by r. Lemma 2.2 implies that for any $c \in R^n$ there is a unique i such that $c = \gamma^i e$, $0 \leq i \leq \nu - 1$, $e \in R^n$ and $\gamma \not| e$.

The R-submodule αR^n also has the structure of K-vector space, with multiplication of a vector $\alpha c \in \alpha R^n$ by $b \in K$ defined to be $a\alpha c$ where $a \in R$ is an element for which $\overline{a} = b$. As in [14, Lemma 2.9] we obtain the following:

LEMMA 2.4 *The map $\varphi : \alpha R^n \to K^n$ given by $\varphi(\alpha c) = \overline{c}$ is an isomorphism of K-vector spaces.*

2.2 Linear algebra over R

For solving linear systems over a commutative ring, the McCoy rank of a matrix ([9, p. 159]) plays a role similar to that of rank in the case of fields. It is not hard to see that the McCoy rank of a matrix over a finite chain ring R is equivalent to the following:

DEFINITION 2.5 *The* McCoy rank *of a matrix M over R is the largest number t such that at least one of the $t \times t$ minors of M is a unit. If none of the minors of M is a unit the McCoy rank is 0.*

Clearly the McCoy rank of an $m \times n$ matrix is at most $\min\{m, n\}$.

We examine now linear dependence in R^n. Note that if $v \in R^n$ and $\gamma | v$ then v is linearly dependent, as $\alpha v = 0$.

THEOREM 2.6 *Let $v_1, \ldots, v_m \in R^n \setminus \gamma R^n$. Then $v_1, \ldots, v_m$ are linearly dependent if and only if $\overline{v}_1, \ldots, \overline{v}_m$ are linearly dependent.*

PROOF. Let $v_1, \ldots, v_m$ be linearly dependent i.e. suppose that there are $\mu_1, \ldots, \mu_m \in R$, not all zero, such that $\sum_{i=1}^m \mu_i v_i = 0$. Let j be maximal such that $\gamma^j | \mu_i$ for $1 \leq i \leq m$. Write $\mu_i = \gamma^j \mu_i'$. Then $\gamma^j \sum_{i=1}^m \mu_i' v_i = 0$ implies $\gamma | \sum_{i=1}^m \mu_i' v_i$. Hence $\sum_{i=1}^m \overline{\mu_i'} \overline{v}_i = 0$, which means that $\overline{v}_1, \ldots, \overline{v}_m$ are linearly dependent, since by the maximality of j, at least one of the μ_i' is not divisible by γ.

Let $\overline{v}_1, \ldots, \overline{v}_m$ be linearly dependent i.e. $\sum_{i=1}^m \beta_i \overline{v}_i = 0$, for some $\beta_i \in K$, not all zero. Let $\mu_i \in R$ be such that $\overline{\mu}_i = \beta_i$. Then $\gamma | \sum_{i=1}^m \mu_i v_i$, hence $\sum_{i=1}^m \alpha \mu_i v_i = \alpha \sum_{i=1}^m \mu_i v_i = 0$. At least one of the μ_i is a unit i.e. at least one of the $\alpha \mu_i$ is not zero. Therefore $v_1, \ldots, v_m$ are linearly dependent. $\square$

COROLLARY 2.7 *The following assertions are equivalent:*

(i) the McCoy rank of M is t.

(ii) the rank of $\overline{M}$ is t.

(iii) M has t linearly independent rows and any $t+1$ rows are linearly dependent.

(iv) M has t linearly independent columns and any $t+1$ columns are linearly dependent.

A system of linear equations for which the McCoy rank of the matrix equals the number of equations can be solved in the usual way:

PROPOSITION 2.8 (CRAMÉR'S RULE, [11, P. 80]) *Let M be an $n \times n$ matrix with McCoy rank equal to n. Then a system of equations $Mx^{\mathrm{tr}} = b$ has the unique solution $x_i = \det(M_i)/\det(M)$, for $i = 1, \ldots, n$, where M_i is the matrix M with the i-th column replaced by b.*

COROLLARY 2.9 *Let M be an $m \times n$ matrix over R, with $m \leq n$. If the McCoy rank of M is m then $|\{x \in R^n |\ Mx^{\mathrm{tr}} = 0\}| = |R|^{n-m}$.*

3 Codes over a finite chain ring

3.1 Definitions

Let $n \geq 1$ be a fixed natural number. By a *(block) code* of length n over R we will mean a non-empty subset of R^n. We will only consider codes different from $\{0\}$ and n will always denote the length of the code. The code is called *linear* if it is an R-submodule of R^n.

A code is called *cyclic* if it is linear and invariant with respect to cyclic shifts. As usual, by identifying the entries of a vector in R^n with the coefficients of a polynomial in $R[X]$ of degree less than n, cyclic codes are precisely the ideals of $R[X]/(X^n - 1)$. We will denote by $\mathrm{id}(S)$ the ideal generated by a set $S \subseteq R[X]/(X^n - 1)$ and write $\mathrm{id}(f_1, \ldots, f_s)$ for $\mathrm{id}(\{f_1, \ldots, f_s\})$.

When working with cyclic codes we will always assume that n is not divisible by the characteristic of K, so that $X^n - 1$ has no multiple factors in $K[X]$. Then for any monic factor $f \in K[X]$ of $X^n - 1$ there is a unique monic $g \in R[X]$ such that $\overline{g} = f$ and $g | X^n - 1$. This is a consequence of Hensel lifting, see for example [14, Theorem 2.7].

DEFINITION 3.1 (HENSEL LIFT OF A CYCLIC CODE) *Let $f \in K[X]$ be monic such that $f|X^n - 1$ and let $g \in R[X]$ be the unique monic polynomial such that $g|X^n - 1$ and $\overline{g} = f$. Then the cyclic code* $\mathrm{id}(g)$ *is called the* Hensel lift *of the cyclic code* $\mathrm{id}(f)$.

It is easy to check that the Hensel lift of a cyclic code E over K is a free cyclic code over R whose projection is E.

3.2 The structure of linear and cyclic codes over R

In this subsection we recall some results on the structure of linear and cyclic codes over R. We will use the approach introduced for the ring Z_{p^a} in [4] and developed in greater detail for finite chain rings in [14].

For $k > 0$, I_k denotes the $k \times k$ identity matrix.

DEFINITION 3.2 (GENERATOR MATRIX) *Let C be a linear code over R. A generator matrix G for C is said to be in* standard form *if after a suitable permutation of the coordinates,*

$$G = \begin{pmatrix} I_{k_0} & A_{01} & A_{02} & A_{03} & \dots & A_{0,\nu-1} & A_{0,\nu} \\ 0 & \gamma I_{k_1} & \gamma A_{12} & \gamma A_{13} & \dots & \gamma A_{1,\nu-1} & \gamma A_{1,\nu} \\ 0 & 0 & \gamma^2 I_{k_2} & \gamma^2 A_{23} & \dots & \gamma^2 A_{2,\nu-1} & \gamma^2 A_{2,\nu} \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \gamma^{\nu-1} I_{k_{\nu-1}} & \gamma^{\nu-1} A_{\nu-1,\nu} \end{pmatrix} \tag{1}$$

say, where the columns are grouped into blocks of sizes $k_0, k_1, \dots, k_{\nu-1}, n - \sum_{i=0}^{\nu-1} k_i$ with $k_i \geq 0$. We associate to G the matrix

$$A = \begin{pmatrix} I_{k_0} & A_{01} & A_{02} & A_{03} & \dots & A_{0,\nu-1} & A_{0,\nu} \\ 0 & I_{k_1} & A_{12} & A_{13} & \dots & A_{1,\nu-1} & A_{1,\nu} \\ 0 & 0 & I_{k_2} & A_{23} & \dots & A_{2,\nu-1} & A_{2,\nu} \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & I_{k_{\nu-1}} & A_{\nu-1,\nu} \end{pmatrix}. \tag{2}$$

The following results generalize [4, p. 22–23]. For proofs, see [14, Section 3].

THEOREM 3.3 *Any non-zero linear code C has a generator matrix in standard form. All generator matrices in standard form for a code C have the same parameters $k_0, \dots, k_{\nu-1}$ and $|C| = |K|^{\sum_{i=0}^{\nu-1}(\nu-i)k_i}$.*

This theorem justifies the following notation.

DEFINITION 3.4 *Let C be a linear code. We denote by $k(C)$ the number of rows of a generating matrix G in standard form for C, and for $i = 0, \dots, \nu-1$ we denote by $k_i(C)$ the number of rows of G that are divisible by γ^i but not by γ^{i+1}.*

Clearly, $k(C) = \sum_{i=0}^{\nu-1} k_i(C)$.

THEOREM 3.5 *Let C be a code with generator matrix G in standard form as in (1). Then*

(i) If for $0 \leq i < j \leq \nu$, $B_{i,j} = -\sum_{k=i+1}^{j-1} B_{i,k} A^{tr}_{\nu-j,\nu-k} - A^{tr}_{\nu-j,\nu-i}$, then

$$H = \begin{pmatrix} B_{0,\nu} & B_{0,\nu-1} & \dots & B_{0,1} & I_{n-k(C)} \\ \gamma B_{1,\nu} & \gamma B_{1,\nu-1} & \dots & \gamma I_{k_{\nu-1}(C)} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ \gamma^{\nu-1} B_{\nu-1,a} & \gamma^{\nu-1} I_{k_1(C)} & \dots & 0 & 0 \end{pmatrix} = \begin{pmatrix} B_0 \\ \gamma B_1 \\ \vdots \\ \gamma^{\nu-1} B_{\nu-1} \end{pmatrix} \tag{3}$$

is a generator matrix for $C^\perp$ and a parity check matrix for C.

(ii) For $i = 0, \ldots, \nu - 1$, $\overline{(C^\perp : \gamma^i)} = (\overline{(C : \gamma^{\nu-1-i})})^\perp$. We have $k(C^\perp) = n - k_0(C)$, $k_0(C^\perp) = n - k(C)$ and $k_i(C^\perp) = k_{\nu-i}(C)$, for $i = 1, \ldots, \nu - 1$.

(iii) $(C^\perp)^\perp = C$.

A code is called *free* if it is a free R-submodule.

COROLLARY 3.6 *Let C be a linear code. The following assertions are equivalent:*

(i) C is a free code.

(ii) Any generator matrix in standard form for C is of the form $(I_{k(C)}\ M)$ for some matrix M.

(iii) $k(C) = k_0(C)$.

(iv) $C^\perp$ is free.

DEFINITION 3.7 (GENERATING SET IN STANDARD FORM) *We say that $S = \{\gamma^{a_0} g_{a_0}, \gamma^{a_1} g_{a_1}, \ldots, \gamma^{a_s} g_{a_s}\}$ is a* generating set in standard form *for the cyclic code $C = \mathrm{id}(S)$ if*

(i) $0 \le s < \nu$, $0 \le a_0 < a_1 < a_2 < \ldots < a_s < \nu$,

(ii) $g_{a_i} \in R[X]$ are monic for $i = 0, \ldots, s$.

(iii) $g_{a_s} | g_{a_{s-1}} | \ldots | g_{a_0} | X^n - 1$ and

(iv) $\deg(g_{a_i}) > \deg(g_{a_{i+1}})$ for $i = 0, \ldots, s - 1$.

The following is [14, Theorem 3.18] and generalizes [4, Theorem 6].

THEOREM 3.8 *Any non-zero cyclic code C over R has a unique generating set in standard form. In the notation of Definitions 3.4 and 3.7, we have $k(C) = n - \deg(g_{a_s})$.*

COROLLARY 3.9 *The code C is the Hensel lift of a cyclic code if and only if there is a monic $g \in R[X]$ such that $\{g\}$ is the generating set in standard form for C.*

Recall that the reciprocal of a non-zero polynomial f is $f^*(X) = X^{\deg(f)} f(1/X)$. For a polynomial $f \in R[X]$ whose constant term is a unit, we define $f^\#$ to be equal to f^* divided by the leading coefficient of f^*. Clearly, the constant term of any divisor of $X^n - 1$ is a unit.

THEOREM 3.10 *Let C be a cyclic code over R with $\{\gamma^{a_0} g_{a_0}, \gamma^{a_1} g_{a_1}, \ldots, \gamma^{a_s} g_{a_s}\}$ a generating set in standard form. Put $a_{s+1} = \nu$ and for $j = -1$, $g_{a_j} = X^n - 1$. For $j = 0, \ldots, s + 1$, let $b_j = \nu - a_{s+1-j}$ and $h_{b_j} = ((X^n - 1)/g_{a_{s-j}})^\#$. Then $\{\gamma^{b_0} h_{b_0}, \gamma^{b_1} h_{b_1}, \ldots, \gamma^{b_{s+1}} h_{b_{s+1}}\}$ is the generating set in standard form for $C^\perp$.*

3.3 Related code constructions

For any code $C \subseteq R^n$ and $r \in R$ we define the code $(C : r)$ by

$$(C : r) = \{e \in R^n | re \in C\}.$$

When C is linear, $(C : r)$ is a submodule quotient and it is a linear code. When C is cyclic, $(C : r)$ is an ideal quotient and it is a cyclic code.

The following properties will provide a characterization of free codes.

PROPOSITION 3.11 *Let C be a code. The following assertions are equivalent:*

(i) $C \cap \gamma^i R^n = \gamma^i C$ *for all* $i = 0, \ldots, \nu - 1$.

(ii) For all $i = 0, \ldots, \nu - 1$ *and for all* $e \in (C : \gamma^i)$ *there is a* $c \in C$ *such that* $e \equiv c \bmod \gamma^{\nu - i}$.

(iii) There is a $D \subseteq R^n \setminus \gamma R^n$ *such that* $C = D \cup \gamma D \cup \ldots \cup \gamma^{\nu-1} D$.

PROOF. (ii)⇒ (i) Let $c \in C \cap \gamma^i R^n$ and $e \in R^n$ be such that $c = \gamma^i e$. Then $e \in (C : \gamma^i)$ and there is a $c' \in C$ such that $e \equiv c' \bmod \gamma^{\nu - i}$ i.e. $c = \gamma^i e = \gamma^i c'$. So $c \in \gamma^i C$.

(i)⇒ (iii) Take $D = C \setminus \gamma R^n$. If $c \in C$ and i is maximal such that $\gamma^i | c$, then $c \in C \cap \gamma^i R^n = \gamma^i C$. Hence $c = \gamma^i e$ for some $e \in C$. By the maximality of i, $\gamma \not| e$, so $e \in D$ and $c \in \gamma^i D$.

The other parts are left as an exercise. □

Note that the property (ii) of Proposition 3.11 implies that $\overline{(C : \gamma^i)} = \overline{C}$, for $i = 0, \ldots, \nu - 1$.

COROLLARY 3.12 *A code is free if and only if it is linear and it satisfies any of the equivalent properties in Proposition 3.11.*

PROOF. Let C be a free code and let $G = (I_{k(C)} \ M)$ be a generator matrix in standard form for C. We show that for $0 \leq i \leq \nu - 1$, $C \cap \gamma^i R^n \subseteq \gamma^i C$. Let $c = vG = (v \ vM) \in C \cap \gamma^i R^n$ for some $v \in R^{k(C)}$. Then $v = \gamma^i u$ for some $u \in R^{k(C)}$ and $c = \gamma^i uG \in \gamma^i C$, as required.

Let C be a linear code satisfying Proposition 3.11(ii). Then $\overline{C} = \overline{(C : \gamma^i)}$ for $i = 0, \ldots, \nu - 1$ and so $k_0(C) = \dim(\overline{C}) = \dim(\overline{(C : \gamma^i)}) = k_0(C) + \ldots + k_i(C)$ by [14, Lemma 3.4]. Thus $k_1(C) = \ldots = k_{\nu-1}(C) = 0$ i.e. C is free. □

Note that codes satisfying any of the equivalent properties of Proposition 3.11 are closed under multiplication by γ, but they need not be linear. Examples of such non-linear codes can easily be constructed using part (iii) of Proposition 3.11 .

For any code C we will denote by C^+ the code obtained by extending C by an overall parity check symbol.

PROPOSITION 3.13 *For any code* C *and any* $r \in R$, *we have* $\overline{(C^+ : r)} = \overline{(C : r)}^+$.

The proof is straightforward. In particular, the case $r = 1$ yields $\overline{C^+} = \overline{C}^+$.

4 The Hamming distance of linear codes over R

4.1 The main result

For $c \in R^n$ we denote by wt (c) the *(Hamming) weight* of c. The (minimum) distance of a code C will be denoted by $\mathrm{d}(C)$. The *support* of $c = (c_1, \ldots, c_n) \in R^n$ is the set $\mathrm{supp}(c) = \{i |\ c_i \neq 0\}$. For a code C we define $\mathrm{supp}(C) = \{\mathrm{supp}(c) |\ c \in C, c \neq 0\}$. We will denote by $\mathrm{S}(C)$ the set of minimum supports of elements in C i.e. the elements of $\mathrm{supp}(C)$ which are minimal with respect to inclusion. Obviously $\mathrm{d}(C) = \min\{|A| \mid A \in \mathrm{S}(C)\}$ so if C and D are codes such that $\mathrm{S}(C) = \mathrm{S}(D)$ then $\mathrm{d}(C) = \mathrm{d}(D)$.

LEMMA 4.1 *The isomorphism* $\varphi : \alpha R^n \to K^n$ *defined in Lemma 2.4 preserves the support and weight of codewords i.e.* $\mathrm{supp}(\alpha c) = \mathrm{supp}(\varphi(\alpha c)) = \mathrm{supp}(\overline{c})$ *and* wt $(\alpha c) =$ wt $(\overline{c})$.

PROOF. The codeword αc has zero entries exactly on those positions where c has entries divisible by γ. These are, on the other hand, exactly the positions where $\overline{c}$ has zero entries. □

THEOREM 4.2 *Let C be a linear code over R. Then:*

(i) $\mathrm{S}(C) = \mathrm{S}(C \cap \gamma^i R^n)$ *and* $\mathrm{d}(C) = \mathrm{d}(C \cap \gamma^i R^n)$ *for all* $0 \leq i \leq \nu - 1$.

(ii) $\mathrm{S}(C) = \mathrm{S}(\overline{(C : \alpha)})$ *and* $\mathrm{d}(C) = \mathrm{d}(\overline{(C : \alpha)})$ *for all* $0 \leq i \leq \nu - 1$.

(iii) If $\overline{C} \neq \{0\}$ *then* $\mathrm{d}(C) \leq \mathrm{d}(\overline{C})$.

(iv) $\mathrm{S}(C^+) = \mathrm{S}((\overline{(C : \alpha)})^+)$ *and* $\mathrm{d}(C^+) = \mathrm{d}((\overline{(C : \alpha)})^+)$.

PROOF. (i) We prove first the equality $\mathrm{S}(C) = \mathrm{S}(D)$ where $D = C \cap \gamma^{\nu-1} R^n$. Let $c \in C$ be such that $\mathrm{supp}(c) \in \mathrm{S}(C)$. We have $\mathrm{supp}(c) \supseteq \mathrm{supp}(\gamma c) \supseteq \mathrm{supp}(\gamma^2 c) \supseteq \ldots \supseteq \mathrm{supp}(\gamma^{\nu-1} c)$. If j is maximal such that $\gamma^j c \neq 0$, then $\mathrm{supp}(c) = \mathrm{supp}(\gamma^j c)$, due to the minimality of $\mathrm{supp}(c)$. By Corollary 2.3, $\gamma^{j+1} c = 0$ implies $\gamma^j c \in D$. Since $D \subseteq C$ and $\mathrm{supp}(\gamma^j c)$ is minimal in $\mathrm{supp}(C)$, it is also minimal in $\mathrm{supp}(D)$. Hence $\mathrm{S}(C) \subseteq \mathrm{S}(D)$. For the converse inclusion, let $c \in D$ be such that $\mathrm{supp}(c)$ is minimal in $\mathrm{supp}(D)$. Since $D \subseteq C$, we have $\mathrm{supp}(c) \in \mathrm{supp}(C)$. Assume for a contradiction that $\mathrm{supp}(c)$ is not minimal in $\mathrm{supp}(C)$. Then there is an $e \in C$ such that $\mathrm{supp}(e) \in \mathrm{S}(C)$ and $\mathrm{supp}(e) \subset \mathrm{supp}(c)$ with strict inclusion. We showed $\mathrm{S}(C) \subseteq \mathrm{S}(D)$, so there is an $e' \in D$ such that $\mathrm{supp}(e) = \mathrm{supp}(e') \subset \mathrm{supp}(c)$ contradicting the fact that $c \in \mathrm{S}(D)$.

For $0 \leq j \leq \nu - 2$ we apply the preceding result to the code $C \cap \gamma^j R^n$ i.e. $\mathrm{S}(C \cap \gamma^j R^n) = \mathrm{S}(C \cap \gamma^j R^n \cap \gamma^{\nu-1} R^n) = \mathrm{S}(C \cap \gamma^{\nu-1} R^n) = \mathrm{S}(C)$.

(ii) Let φ be the isomorphism defined in Lemma 2.4. One can easily verify that $C \cap \alpha R^n = \underline{\alpha}(C : \alpha)$ and therefore $\varphi(C \cap \alpha R^n) = \overline{(C : \alpha)}$. By Lemma 4.1, $\mathrm{S}(C \cap \alpha R^n) = \mathrm{S}(\varphi(C \cap \alpha R^n)) = \mathrm{S}(\overline{(C : \alpha)})$.

(iii) Using (ii), $\overline{C} \neq \{0\}$ and $\overline{(C : \gamma^i)} \subseteq \overline{(C : \gamma^{i+1})}$ for $i = 0, \ldots, \nu - 2$, we have $\mathrm{d}(C) = \mathrm{d}(\overline{(C : \alpha)}) = \mathrm{d}(\overline{(C : \gamma^{\nu-1})}) \leq \mathrm{d}(\overline{(C : \gamma^{\nu-2})}) \leq \ldots \leq \mathrm{d}(\overline{(C : \gamma^0)}) = \mathrm{d}(\overline{C})$.

(iv) Use the fact that $\overline{(C^+ : \alpha)} = (\overline{(C : \alpha)})^+$ by Proposition 3.13. □

COROLLARY 4.3 *If C is a free code (in particular, if C is the Hensel lift of a cyclic code) then* $\mathrm{S}(C) = \mathrm{S}(\overline{C})$, $\mathrm{d}(C) = \mathrm{d}(\overline{C})$, $\mathrm{S}(C^+) = \mathrm{S}(\overline{C}^+)$ *and* $\mathrm{d}(C^+) = \mathrm{d}(\overline{C}^+)$.

PROOF. Apply Theorem 4.2 (ii) and (iv) using the fact that for a free code $\overline{C} = \overline{(C : \alpha)}$, by Corollary 3.12. □

By Corollary 4.3 above, all the classical lower bounds for the distance of cyclic codes over finite fields (BCH, Hartmann-Tzeng, Roos *etc.*) also apply to their Hensel lifts.

REMARKS 4.4 *(i) Theorem 4.2 holds more generally for non-linear codes closed under multiplication by γ, as can be seen from the proof. Corollary 4.3 is also valid for non-linear codes that satisfy any of the properties of Proposition 3.11.*

(ii) The inequality $\mathrm{d}(C) \geq \mathrm{d}(\overline{C})$ *contained in Corollary 4.3 was stated for Hensel lifts of cyclic codes in [18, Proof of Theorem 4] but the proof given there is incorrect. Namely for a word $c \in C$ of minimum weight it is inferred that* $\mathrm{d}(C) = \mathrm{wt}\ (c) \geq \mathrm{wt}\ (\overline{c}) \geq \mathrm{d}(\overline{C})$, *but the last inequality fails when* $\overline{c} = 0$.

4.2 Determining $\overline{(C : \alpha)}$ and its distance

We saw in Theorem 4.2(ii) that finding $\mathrm{d}(C)$ reduces to finding $\mathrm{d}(\overline{(C : \alpha)})$. We will determine the latter code using the generator matrix/polynomials of the code C described in Subsection 3.2.

THEOREM 4.5 *(i) Let C be a linear code. Let G be a generator matrix in standard form for C as in (1), A the matrix associated to G as in (2) and B_0 as in Theorem 3.5. Then $\overline{A}$ is a generator matrix and $\overline{B_0}$ a parity check matrix for the code $\overline{(C : \gamma^{\nu-1})}$.*

(ii) If C is a cyclic code with generating set in standard form $\{\gamma^{a_0} g_{a_0}, \gamma^{a_1} g_{a_1}, \ldots, \gamma^{a_s} g_{a_s}\}$, then the generator polynomial of $\overline{(C : \gamma^{\nu-1})}$ is $\overline{g}_{a_s}$.

(iii) $(\overline{(C : \alpha)})^{\perp} = \overline{C^{\perp}}$.

PROOF. Part (i) follows from Theorem 3.5 and [14, Lemma 3.4], part (ii) from [14, Lemma 3.17] and part (iii) from Theorem 3.5(ii). □

COROLLARY 4.6 *(i) Let C be a linear code with generator matrix in standard form G as in (1) and A associated to G as in (2). If D is the linear code over K generated by $\overline{A}$, then* $\mathrm{S}(C) = \mathrm{S}(D)$ *and* $\mathrm{d}(C) = \mathrm{d}(D)$.

(ii) $\mathrm{d}(C) = d$ *iff any $d-1$ columns of $\overline{B}_0$ are linearly independent and some d columns of $\overline{B}_0$ are linearly dependent.*

(iii) If C is a cyclic code with generating set in standard form $\{\gamma^{a_0} g_{a_0}, \gamma^{a_1} g_{a_1}, \ldots, \gamma^{a_s} g_{a_s}\}$ and $D = \mathrm{id}(\overline{g}_{a_s}) \subset K[X]/(X^n - 1)$, then $\mathrm{S}(C) = \mathrm{S}(D)$ *and* $\mathrm{d}(C) = \mathrm{d}(D)$. *Also,* $\mathrm{S}(C^+) = \mathrm{S}(D^+)$ *and* $\mathrm{d}(C^+) = \mathrm{d}(D^+)$.

PROOF. Apply Theorem 4.5 to Theorem 4.2. □

Recall that the rate of a code C over R is $\rho(C) = (\log_{|R|} |C|)/n$. If C is a linear code, using the expression for $|C|$ given in Theorem 3.3 and the fact that $|R| = |K|^{\nu}$, we obtain $\rho(C) = (\sum_{i=0}^{\nu-1} (\nu - i) k_i(C))/(n\nu)$. If C is free then $\rho(C) = k(C)/n$, by Corollary 3.6.

COROLLARY 4.7 *(i) Let C be a linear code which is not free. There is a free code D of the same length such that* $\mathrm{d}(D) = \mathrm{d}(C)$, $k(D) = k(C)$ *and* $|D| > |C|$.

(ii) Let C be a linear code. Then $\rho(C) \leq \rho(\overline{(C : \alpha)})$, with equality if and only if C is free. When C is free, $\rho(C) = \rho(\overline{C})$.

(iii) $\max\{\rho(C) | C \subseteq R^n, C \text{ linear}, \mathrm{d}(C) = d\} = \max\{\rho(C) | C \subseteq K^n, C \text{ linear}, \mathrm{d}(C) = d\}$.

PROOF. (i) Let G be a generator matrix in standard form for C and A the matrix associated to G as in (2). Take D to be the free code generated by A. By Corollary 4.6, $\mathrm{d}(C) = \mathrm{d}(D)$, and by Theorem 3.3, $|D| > |C|$, as $k_0(D) = k(D) = k(C)$. (We remark that if C is a cyclic code given by a generating set in standard form $\{\gamma^{a_0} g_{a_0}, \gamma^{a_1} g_{a_1}, \ldots, \gamma^{a_s} g_{a_s}\}$, this amounts to taking $D = \mathrm{id}(g_{a_s})$; see [14, Theorem 3.19].)

(ii) Use the fact that $\rho(\overline{(C : \alpha)}) = \dim(\overline{(C : \alpha)})/n = k(C)/n$, as $\dim(\overline{(C : \alpha)}) = k(C)$, by Theorem 4.5(i).

(iii) The "$\leq$ part" follows from (ii). For proving the "$\geq$ part", let E be a linear code over K with $\mathrm{d}(E) = d$ and $\rho(E)$ maximal. Let G be a matrix over R such that $\overline{G}$ is a generator matrix of E. The code C over R generated by G is free and it has the same distance and rate as $\overline{C} = E$. □

Hence, from the point of view of Hamming distance, it is always better to work with free codes (which includes Hensel lifts of cyclic codes).

REMARKS 4.8 *(i) All the results on cyclic codes we proved so far in this paper also hold when we replace $X^n - 1$ by any polynomial $f \in R[X]$ such that $\overline{f}$ has no multiple factors in $K[X]$. The codes thus obtained have a structure similar to that of cyclic codes, see [14, Remark 3.26]. For the particular case $f = X^n + 1$ we obtain negacyclic codes and for $f = X^n + c$, $c \in R \setminus \{0\}$ we obtain constacyclic codes.*

(ii) There are codes C of length $n \geq 2$ for which the difference $\mathrm{d}(\overline{C}) - \mathrm{d}(C)$ is as large as the distance between two non-zero codes of length n can be i.e. $n-1$. For example if C has generator matrix

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ 0 & & \gamma I_{n-1} & \end{pmatrix}$$

then $\mathrm{d}(C) = 1$ but $\mathrm{d}(\overline{C}) = n$.

(iii) In [4, Corollary to Theorem 6] it is proved that any ideal C in $Z_{p^a}[X]/(X^n-1)$ is principal. One might be tempted to apply Corollary 4.6(iii) to the unique generator g of C and conclude (wrongly) that $\mathrm{d}(C) = \mathrm{d}(\mathrm{id}(g)) = \mathrm{d}(\mathrm{id}(\overline{g})) = \mathrm{d}(\overline{C})$ for any cyclic code C. For the unique generator of C does not necessarily divide X^n-1, as can be seen from [4, loc. cit.*] so it is not necessarily a generating set in standard form for C. Indeed $\{g\}$ cannot be a generating set in standard form for C unless C is the Hensel lift of a cyclic code over K, see Corollary 3.9. This situation is illustrated in Example 4.10 below.*

Cyclic codes over Galois rings can also be described in terms of roots of $X^n - 1$. The following theorem determines their distance in this case.

THEOREM 4.9 *Let $R = GR(p^a, l)$ and let $m \in \mathbb{N}$ be such that $l|m$ and $n|p^m-1$. Let $\xi \in GR(p^a, m)$ be a primitive root of X^n-1 such that $\overline{\xi}$ is primitive as well and let U be a system of representatives for the conjugacy classes of the roots of ξ. Finally, let $0 \leq s \leq a-1$, $0 \leq a_0 < \ldots < a_s \leq a-1$ and let L_j, for $j = 0, \ldots, s$ be pairwise disjoint subsets of U. Then*

$$\mathrm{d}(\{c \in R_n |\ p^{a_j} c(\xi^{i_j}) = 0, j = 0, \ldots, s, i_j \in L_j\}) = \mathrm{d}(\{c \in K_n |\ c(\overline{\xi}^{i_0}) = 0, i_0 \in L_0\}),$$

where $R_n = R[X]/(X^n-1)$ and $K_n = K[X]/(X^n-1)$.

PROOF. Use Corollary 4.6(iii) and [14, Theorem 3.28]. □

4.3 Examples

Using the results of the previous two subsections we will determine or give lower bounds for the distance of several codes over finite rings described in the literature.

EXAMPLE 4.10 ([4, EXAMPLE 5]) *Let $R = Z_4$ and $n = 7$. The factorisation of $X^7 - 1$ over Z_2:*

$$X^7 - 1 = (X-1)(X^3 + X + 1)(X^3 + X^2 + 1)$$

lifts to $Z_4[X]$ giving

$$X^7 - 1 = (X-1)(X^3 + 2X^2 + X + 3)(X^3 + 3X^2 + 2X + 3).$$

Let $g_1 = X^3 + 2X^2 + X + 3$ and $g_0 = (X-1)g_1$. Consider the cyclic code $C = \mathrm{id}(g_0, 2g_1)$, which is code number 22 in Table 1, loc.cit. By Corollary 4.6(iii), $\mathrm{d}(C) = \mathrm{d}(\mathrm{id}(\overline{g_1})) = \mathrm{d}(\mathrm{id}(X^3+X+1)) = 3$. On the other hand, by [4, Corollary to Theorem 6], the ideal C is principal with generator $g = g_0 + 2g_1$. The reader might be tempted to apply Corollary 4.6(iii) to this generator and conclude that $\mathrm{d}(C) = \mathrm{d}(\overline{C}) = \mathrm{d}(\mathrm{id}(\overline{g})) = \mathrm{d}(\mathrm{id}(\overline{g_0})) = 4$. However, $g \not| X^7 - 1$, so it is not a generating set in standard form for C. (See also Remark 4.8(iii).)

Similarly we can determine the distance for the other codes of [4, Example 5], obtaining distance 1 for the codes 8,9,11,14,18,25, distance 2 for codes 7,10,20, distance 3 for codes 5,12,22,26, distance 4 for codes 3,16 and distance 7 for codes 2,24. Note that the distances in [4, Table 1] are Lee distances, whereas we obtain Hamming distances.

Many of the codes over rings described in the literature are (extended) Hensel lifts of cyclic codes. By Corollary 4.3, these codes have the same distance as the original (extended) cyclic codes over K. We now present some examples of this fact.

EXAMPLE 4.11 (THE [8,4] HAMMING CODE) *In [4, Example 1] an extended Hamming code of length 8 over Z_{2^a} is constructed by extending the Hensel lift of the code generated by X^3+X+1 over Z_2. It is proved,* loc. cit. *that this code has distance 4, which also follows from Corollary 4.3.*

EXAMPLE 4.12 (BCH CODES, [18]) *The BCH code over $R = Z_{p^a}$ of designed distance d is the Hensel lift of the BCH code over Z_p of designed distance d. By Corollary 4.3, the two codes have the same distance.*

EXAMPLE 4.13 (KERDOCK AND PREPARATA CODES, [6]) *In [6] it is proved that both Kerdock and Preparata codes are the image under a Gray map of certain extended cyclic codes over Z_4, denoted $\mathcal{K}$ and $\mathcal{P}$, respectively. We will show that $\mathrm{d}(\mathcal{K}) = 2^{m-1}$ and $\mathrm{d}(\mathcal{P}) = 4$.*

Let $R = Z_4$, m be odd, $m \geq 3$ and $n = 2^m - 1$. Let $h \in Z_4[X]$ be a primitive basic irreducible polynomial of degree m such that $h|X^n - 1$. Let g be the reciprocal of $(X^n-1)/((X-1)h)$. Note that $g|X^n - 1$. As in [6], let $\mathcal{K} = \mathrm{id}(g)^+$. By Corollary 4.6(iii), $\mathrm{d}(\mathcal{K}) = \mathrm{d}(\mathrm{id}(\overline{g})^+)$. Without loss of generality we may assume that h is the minimal polynomial of a primitive n-th root of unity $\xi \in GR(4,m)$ and that $\overline{\xi}$ is a primitive n-th root of unity in $GF(2^m)$. The code $\mathrm{id}(\overline{g})$ is a cyclic code defined by the roots $(\overline{\xi})^{2^{m-1}+1}, (\overline{\xi})^{2^{m-1}+2}, \ldots, (\overline{\xi})^{2^m-2}$, hence it is a BCH code with designed distance $2^{m-1}-1$. By [8, Ch. 9, Theorem 5], this is the actual distance of the code. The extended code has distance 2^{m-1}. Hence $\mathrm{d}(\mathcal{K}) = 2^{m-1}$.

Now let $\mathcal{P} = \mathrm{id}(h)^+$. By Corollary 4.6(iii), $\mathrm{d}(\mathcal{P}) = \mathrm{d}(\mathrm{id}(\overline{h})^+)$. The code $\mathrm{id}(\overline{h})$ is the Hamming code with distance 3. Hence $\mathrm{d}(\mathcal{P}) = 4$.

EXAMPLE 4.14 (GENERALISED REED-MULLER CODES) *It is well-known that for $0 \leq r \leq m(p^l - 1)$ the generalised Reed-Muller code $GRM(r,m)$ over $GF(p^l)$ is an extended BCH code [1, Section 5.5]. It follows that we can define $GRM(r,m)$ over $GR(p^a,l)$ as an extension of a Hensel lift, and by Corollary 4.3 that these codes will have distance equal to the distance of their projections over $GF(p^l)$ given in [1, Corollary 5.5.4]. In fact Example 4.13 is the case $p = a = 2$ and $l = 1$ with $r = 1$ for $\mathcal{K}$ and $r = m-1$ for $\mathcal{P}$ since $\mathcal{K}^\perp = \mathcal{P}$, [6].*

EXAMPLE 4.15 (GOETHALS, GOETHALS-DELSARTE CODES [6]) *Let m be odd, $n = 2^m - 1$ and $1 \leq r \leq (m-1)/2$. In [6, Theorem 24] it is shown that the Goethals-Delsarte code $GD(m+1, r+2)$ (and for the particular case $r = 1$, the Goethals code) has the same weight distribution as the image under a Gray map of the code $\mathcal{GD} = \{c \in Z_4[X]/(X^n-1) \mid c(\xi) = 2c(\xi^3) = \ldots = 2c(\xi^{1+2^r}) = 0\}^+$, where ξ is a primitive n-th root of unity such that $\overline{\xi}$ is primitive as well. By Theorem 4.9, $\mathrm{d}(\mathcal{GD}) = \mathrm{d}(\{c \in Z_2[X]/(X^n-1) \mid c(\overline{\xi}) = 0\}^+) = 4$.*

EXAMPLE 4.16 (QUADRATIC RESIDUE CODES) *Quadratic residue codes over Z_{p^a} are constructed in [4, 3, 16, 2] as Hensel lifts of quadratic residue codes over Z_p. By Corollary 4.3, their distance is the same as the distance of the original codes over Z_p.*

EXAMPLE 4.17 (GOLAY CODES) *The Golay codes of length 24 over Z_{2^a} and of length 12 over Z_{3^a} are constructed in [4, Example 2 and 3] by lifting the generator polynomial of the respective Golay codes over Z_2 and Z_3, then appending a 1 to each row of the generator matrix. By Corollary 4.3 the lifted codes have the same distance (8 and 6, respectively), as the original Golay codes. This result is stated (without proof) in loc. cit.*

EXAMPLE 4.18 *In [5] a cyclic code which is not a Hensel lift is studied, namely $C = D^+$ with $D = \{c \in Z_4[X]/(X^n-1) \mid c(\xi) = c(\xi^3) = 2c(\xi^5) = 0\}$, where $n = 2^m - 1$ for some m and ξ is*

a primitive n-th root of unity such that $\overline{\xi}$ is primitive as well. By Theorem 4.9, D has the same distance as the code $\{c \in Z_2[X]/(X^n-1) |\ c(\overline{\xi}) = c(\overline{\xi}^3) = 0\}$, a BCH code of designed distance 5. Hence $\mathrm{d}(C) \geq 6$.

5 MDS codes over R

5.1 The Singleton bound over R

We begin by recalling three possible proofs of the Singleton bound for a linear code C over a field with q elements. The first proof consists of deleting $\mathrm{d}(C)-1$ coordinates in the code and obtaining $|C|$ distinct words, which gives $|C| \leq q^{n-\mathrm{d}(C)+1}$(for any code, not necessarily linear). If C is linear, $|C| = q^{\dim(C)}$ implies that $\mathrm{d}(C) \leq n - \dim(C) + 1$, as required. The second proof looks at a generator matrix of C in standard form. Any row of this matrix is a codeword of weight $n - \dim(C) + 1$. The third proof uses the fact that any $\mathrm{d}(C)-1$ columns of a parity check matrix for C are linearly independent, otherwise there would be a word of weight $\mathrm{d}(C)-1$ in the code. On the other hand the parity check matrix has $n - \dim(C)$ rows, so it can have no more than $n - \dim(C)$ linearly independent columns.

Let us apply similar arguments to a linear code C over the finite chain ring R, omitting the trivial case $C = R^n$. The first argument gives $|C| \leq |R|^{n-\mathrm{d}(C)+1}$ (for any code). Recall that $|R| = |K|^{\nu}$, and that $|C| = |K|^{\sum_{i=0}^{\nu-1}(\nu-i)k_i(C)}$ by Theorem 3.3. Hence, for a linear code

$$\mathrm{d}(C) \leq n - \frac{1}{\nu}\sum_{i=0}^{\nu-1}(\nu-i)k_i(C) + 1. \tag{4}$$

Secondly, in a generator matrix in standard form, the last row is a codeword of weight $n-k(C)+1$, hence

$$\mathrm{d}(C) \leq n - k(C) + 1. \tag{5}$$

Thirdly, any $\mathrm{d}(C)-1$ columns of the parity check matrix are linearly independent. Only the rows which are not divisible by γ in the standard form of the parity check matrix may be linearly independent, i.e. at most $k_0(C^{\perp}) = n - k(C)$ rows. By Corollary 2.7 this is also the maximum number of linearly independent columns. We obtain again the inequality (5).

Note that (5) implies (4). The two inequalities coincide if and only if C is a free code, by Corollary 3.6. As in [4], we will call inequality (5) the Singleton bound over R.

5.2 MDS codes

DEFINITION 5.1 ([4, P.28]) *A linear code C for which $\mathrm{d}(C) = n - k(C) + 1$ is called an* MDS (maximum distance separable) *code.*

Clearly when R is a finite field we have $k(C) = \dim(C)$ and we obtain the usual definition of an MDS code. By Corollary 4.7(i) we could restrict our attention to free MDS codes. However, whenever possible, the results of this section are proved for MDS codes which are not necessarily free.

For the rest of this section, we put $k = k(C)$ and $d = \mathrm{d}(C)$.

Since any codeword of C provides a linear dependency between the columns of any parity check matrix, the following is immediate.

THEOREM 5.2 *Let C be a linear code over R. Then C is MDS if and only if any $n-k$ columns of a parity check matrix of C are linearly independent.*

Note that for codes over finite fields, C is MDS if and only if any k columns of the generator matrix of C are linearly independent (see [8, Corollary 3, Ch. 11]). This assertion does not hold over R. For if C is an MDS code which is not free and k columns of G are linearly independent, then all k rows of G are linearly independent, which is impossible since some of them are divisible by γ.

The following theorem gives an important characterisation of MDS codes over finite chain rings.

THEOREM 5.3 *A linear code C is MDS over R if and only if $\overline{(C:\alpha)}$ is MDS over K.*

PROOF. By Theorem 4.5(i) we have $\dim(\overline{(C:\alpha)}) = k(C)$ and by Theorem 4.2(ii), $\mathrm{d}(\overline{(C:\alpha)}) = \mathrm{d}(C)$. □

For codes over finite fields, a code is MDS if and only if its dual is MDS. (See for example [8, Theorem 2, Ch. 11].) This is not, in general, the case over R. There is however the following characterization:

THEOREM 5.4 *A code C is MDS if and only if $\overline{C^\perp}$ is MDS.*

PROOF. By Theorem 5.3, C is MDS if and only if $\overline{(C:\alpha)}$ is MDS. From Theorem 4.5(iii), $(\overline{(C:\alpha)})^\perp = \overline{C^\perp}$, hence $\overline{(C:\alpha)}$ is MDS if and only if $\overline{C^\perp}$ is MDS. □

COROLLARY 5.5 *Let C be a free code. The following assertions are equivalent.*

(i) C is MDS.

(ii) $\overline{C}$ is MDS.

(iii) $C^\perp$ is MDS.

(iv) Any k columns of a generator matrix G of C are linearly independent.

(v) $\gamma^j C$ is MDS for all $j \in \{0,\ldots,\nu-1\}$.

(vi) $\gamma^j C$ is MDS for some $j \in \{0,\ldots,\nu-1\}$.

PROOF. For a free code we have $\overline{C} = \overline{(C:\alpha)}$ and $C \cap \gamma^j R^n = \gamma^j C$ by Proposition 3.11 and Corollary 3.12.

For (i) $\Leftrightarrow$ (ii) use Theorem 5.3.

For (i) $\Leftrightarrow$ (iii) use the fact that, by Theorem 5.4, C is MDS if and only if $\overline{C^\perp}$ is MDS. Applying the the already proved equivalence (i) $\Leftrightarrow$ (ii) to the free code $C^\perp$, we have that $\overline{C^\perp}$ is MDS if and only if $C^\perp$ is MDS.

For (ii) $\Leftrightarrow$ (iv) use the fact that since C is free, $\overline{G}$ has k non-zero rows and $\overline{C}$ is MDS if and only if any k columns of $\overline{G}$ are linearly independent. Then apply Theorem 2.6.

For (i) $\Leftrightarrow$ (v) and (i) $\Leftrightarrow$ (vi) use the fact that $\mathrm{d}(C) = \mathrm{d}(C \cap \gamma^j R^n)$ by Theorem 4.2(i), and $k(\gamma^j C) = k(C)$. □

The only known non-trivial MDS codes over finite fields are Reed-Solomon codes and their extensions (see [8, §5, Ch. 11]). Using Theorem 5.3 and Corollary 5.5 we will construct MDS codes over Galois rings.

EXAMPLE 5.6 (REED-SOLOMON CODES AND THEIR EXTENSIONS) *Reed-Solomon codes over Galois rings were introduced in [18] as Hensel lifts of Reed-Solomon codes. In this example $R =$*

$GR(p^a, l)$. Take $n|p^l - 1$ and let $\alpha \in R$ be a primitive n-th root of unity such that $\overline{\alpha}$ is a primitive n-th root of unity in K. Then the Reed-Solomon code over R with distance d is generated by $g = (X-\alpha)(X-\alpha^2)\dots(X-\alpha^{d-1})$ and has parity check matrix:

$$H = \begin{pmatrix} 1 & \alpha & \alpha^2 & \dots & \alpha^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \alpha^{d-1} & (\alpha^{d-1})^2 & \dots & (\alpha^{d-1})^{n-1} \end{pmatrix}.$$

This is a free MDS code over R, by Corollary 5.5. We can extend this code as in [8, §5, Ch.11] and obtain other MDS codes. Namely, the codes of length $n+1$ and $n+2$ with distance $d+1$ and $d+2$ respectively, defined by the parity check matrices

$$H' = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & \alpha & \alpha^2 & \dots & \alpha^{n-1} & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & \alpha^{d-1} & (\alpha^{d-1})^2 & \dots & (\alpha^{d-1})^{n-1} & 0 \end{pmatrix}$$

and

$$H'' = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 & 1 & 0 \\ 1 & \alpha & \alpha^2 & \dots & \alpha^{n-1} & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 1 & \alpha^{d-1} & (\alpha^{d-1})^2 & \dots & (\alpha^{d-1})^{n-1} & 0 & 0 \\ 1 & \alpha^{d} & (\alpha^{d})^2 & \dots & (\alpha^{d})^{n-1} & 0 & 1 \end{pmatrix}$$

are MDS. This can be easily checked using Corollary 5.5 since the codes are free and $\overline{H'}$ and $\overline{H''}$ are parity check matrices of MDS codes over K. Also for $p = 2$ the extended code with parity check matrix:

$$H''' = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 & 1 & 0 & 0 \\ 1 & \alpha & \alpha^2 & \dots & \alpha^{n-1} & 0 & 1 & 0 \\ 1 & \alpha^2 & (\alpha^2)^2 & \dots & (\alpha^2)^{n-1} & 0 & 0 & 1 \end{pmatrix}$$

and its dual code are MDS.

Actually any other free code over R whose parity check matrix coincides modulo p with H, H', H'' or H''' above is MDS. Also, any code C which is not necessarily free, but for which $\overline{B}_0$ equals $\overline{H}, \overline{H'}, \overline{H''}$ or $\overline{H'''}$ (where B_0 is as defined in Theorem 3.5) is MDS by Theorems 4.5 and 5.3.

In the following example we construct an MDS code C such that neither $\overline{C}$ nor $C^\perp$ is MDS. This means Corollary 5.5 fails if we drop the assumption that C is free.

EXAMPLE 5.7 *Let $R = Z_{49}$ and $n = 6$. Then $\alpha = 31$ is a primitive sixth root of unity in R. Also, $\overline{\alpha} = 3$ is a primitive sixth root of unity in $K = Z_7$. Put $g_0 = (X-\alpha)(X-\alpha^2)(X-\alpha^4)$ and $g_1 = (X-\alpha)(X-\alpha^2)$. Define the cyclic code $C = \mathrm{id}(g_0, 7g_1)$ (which is not free). Since $\overline{(C:7)} = \mathrm{id}(\overline{g_1})$ is an MDS (Reed-Solomon) code, C is an MDS code over R, by Theorem 5.3. The code $\overline{C}$ is generated by $\overline{g_0} = (X-\overline{\alpha})(X-\overline{\alpha}^2)(X-\overline{\alpha}^4) = (X-3)(X-2)(X-4) = X^3+5X^2+5X+4 \in Z_7[X]$. Now the codeword $(X+2)\overline{g_0} = X^4+X^2+1$ has weight 3 and $\dim(\overline{C}) = n - \deg(\overline{g_0}) = 3$, so $\overline{C}$ is not MDS.*

By Theorem 3.10, $C^\perp = \mathrm{id}(h_0, 7h_1)$ where $h_0 = ((X^6-1)/g_1)^\# = (X-1)(X-\alpha)(X-\alpha^2)(X-\alpha^3)$ and $h_1 = ((X^6-1)/g_0)^\# = (X-1)(X-\alpha)(X-\alpha^3)$. The code $\overline{(C^\perp:7)}$ is generated by $\overline{h_1} = X^3+4X^2+6X+3$ and contains the codeword $(X+3)\overline{h}_1 = X^4+4X^2+2$ of weight 3. Hence $\overline{(C^\perp:7)}$ and consequently $C^\perp$ are not MDS codes. Note however that $\overline{C^\perp} = \mathrm{id}(\overline{h}_0)$ is MDS.

THEOREM 5.8 (CF. [8, THEOREM 4, CH. 11]) *An MDS code C has a word of weight d in any d coordinates.*

PROOF. Let H be a parity check matrix for C. The codewords $c \in C$ must satisfy the system of equations: $\gamma^i B_i c^{tr} = 0$ for $i = 0, \ldots, a-1$, where the B_i are as in Theorem 3.5.

The code $\overline{(C:\alpha)}$ is MDS, by Theorem 5.3, and $\overline{B}_0$ is a parity check matrix for this code, by Theorem 4.5. So any $n-k$ columns of $\overline{B}_0$ are linearly independent. By Theorem 2.6 this means any $n-k$ columns of B_0 are linearly independent.

Let $i_1, \ldots, i_d \in \{1, \ldots, n\}$ be d distinct coordinates. We put $c_j = 0$ for $j \notin \{i_1, \ldots, i_d\}$ and $c_{i_d} = 1$ in the system $B_0 c^{tr} = 0$ and solve for the $d-1 = n-k$ unknowns $c_{i_1}, \ldots, c_{i_{d-1}}$ using Proposition 2.8. The solution c we obtain need not be a solution of the whole system $Hc^{tr} = 0$, but $\gamma^j c$ with j maximal such that $\gamma^j c \neq 0$ will be, since $\gamma^i B_i \gamma^j c^{tr} = 0$ trivially for $i \geq 1$. Hence $\gamma^j c \in C$. Since $\emptyset \neq \mathrm{supp}(\gamma^j c) \subseteq \mathrm{supp}(c) \subseteq \{i_1, \ldots, i_d\}$ and wt $(\gamma^j c) \geq d$, we conclude $\mathrm{supp}(\gamma^j c) = \{i_1, \ldots, i_d\}$ as required. □

For free MDS codes, we can determine the weight enumerator, as for MDS codes over a finite field. Our proof follows [15, §3.9].

LEMMA 5.9 *Let C be a free MDS code, $1 \leq i_1 < \ldots < i_s \leq n$ and $Z(i_1, \ldots, i_s) = \{c \in C |\ c_j = 0 \text{ for } j = i_1, \ldots, i_s\}$. Then*

$$|Z(i_1, \ldots, i_s)| = \begin{cases} |R|^{k-s} & s \leq k-1 \\ 1 & s \geq k. \end{cases}$$

PROOF. If $s \geq k$ and $c \in Z(i_1, \ldots, i_s)$ then wt $(c) \leq n - s < d$, hence $c = 0$. Now let $s \leq k-1$. Since C is free, H has $n - k_0(C) = n - k$ rows. Denote by H' the $(n-k) \times (n-s)$ matrix obtained from H by deleting columns $i_1, \ldots, i_s$. For any $c \in R^n$ denote by c' the vector obtained from c by deleting the coordinates $i_1, \ldots, i_s$. Then $c \in Z(i_1, \ldots, i_s)$ if and only if $H'(c')^{tr} = 0$ and $c_{i_1} = \ldots = c_{i_s} = 0$. Since C is MDS, by Theorem 5.2 any $n-k$ columns of H are linearly independent, hence by Corollary 2.7 the McCoy rank of H' is $n-k$. By Corollary 2.9, the number of solutions of the system $H'(c')^{tr} = 0$ is $|R|^{(n-s)-(n-k)} = |R|^{k-s}$. □

THEOREM 5.10 *Let C be a free MDS code over R. For $d \leq w \leq n$, denote by A_w the number of words of weight w in C. Then*

$$A_w = \binom{n}{w} \sum_{i=0}^{w-d} (-1)^i \binom{w}{i} (|R|^{w+1-d-i} - 1).$$

PROOF. The proof is similar to [15, §3.9]. It uses Lemma 5.9 and combinatorial arguments which are independent of the algebraic structure of the finite alphabet. □

COROLLARY 5.11 (CF. [8, COROLLARY 7, CHAPTER 11]) *Let C be an MDS code over R. If $k \geq 2$ then $|R| \geq n - k + 1$. If $k \leq n-2$ then $|R| \geq k+1$.*

PROOF. If C is free then from Theorem 5.10, the condition $A_{n-k+2} \geq 0$ and a similar condition for the weight distribution of $C^{\perp}$, imply the desired inequalities as in the proof of [8, Corollary 7, Chapter 11]. If C is not free, then using Corollary 4.7(i) we construct a free MDS code D with the same parameters d and k. We then apply the previous argument to D. □

Acknowledgement. The authors gratefully acknowledge financial support from the U.K. Engineering and Physical Sciences Research Council (EPSRC). The second author was supported by EPSRC Grant L07680.

References

[1] E.F. Assmus and J. Key. *Designs and Their Codes.* Cambridge University Press, 1992.

[2] A. Bonnecaze, P. Solé, C. Bachoc, and B. Mourrain. Type II codes over Z_4. *IEEE Trans. Inform. Theory*, 43(3):969–976, 1997.

[3] A. Bonnecaze, P. Solé, and A.R. Calderbank. Quaternary quadratic residue codes and unimodular lattices. *IEEE Trans. Inform. Theory*, 41(2):366–377, 1995.

[4] A. R. Calderbank and N. J. A. Sloane. Modular and p-adic codes. *Designs, Codes and Cryptography*, 6:21–35, 1995.

[5] A.R. Calderbank, G. McGuire, P.V. Kumar, and T. Helleseth. Cyclic codes over Z_4, locator polynomials, and Newton's identities. *IEEE Trans. Inform. Theory*, 42(1):217–226, 1996.

[6] A. R. Hammons, Jr., P. V. Kumar, A. R. Calderbank, N. J. A. Sloane, and P. Solé. The Z_4 linearity of Kerdock, Preparata, Goethals and related codes. *IEEE Trans. Inform. Theory*, 40:301–319, 1994.

[7] J. C. Interlando, R. Palazzo, and M. Elia. On the decoding of Reed-Solomon and BCH codes over integer residue rings. *IEEE Trans. Inform. Theory*, 43(3):1013–1021, 1997.

[8] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-correcting Codes.* North Holland, Amsterdam, 1977.

[9] N. H. McCoy. *Rings and Ideals.* Number 8 in Carus Mathematical Monographs. Mathematical Association of America, 1962.

[10] B. R. McDonald. *Finite Rings with Identity.* Marcel Dekker, New York, 1974.

[11] B. R. McDonald. *Linear Algebra over Commutative Rings.* Marcel Dekker, New York, 1984.

[12] G. H. Norton. On minimal realization over a finite chain ring. *Designs, Codes and Cryptography*, 16:161–178, 1999.

[13] G. H. Norton and A. Salagean. On the key equation over a commutative ring. *Designs, Codes and Cryptography*, 20:125–141, 2000.

[14] G. H. Norton and A. Salagean. On the structure of linear and cyclic codes over finite chain rings. *Applicable algebra in engineering, communication and computing*, 10:489–506, 2000.

[15] W. W. Peterson and W.J. Weldon. *Error-Correcting Codes.* MIT Press, 1972.

[16] V. S. Pless and Z. Qian. Cyclic codes and quadratic residue codes over Z_4. *IEEE Trans. Inform. Theory*, 42(5):1594–1600, 1996.

[17] J. A. Reeds and N. J. A. Sloane. Shift-register synthesis (modulo m). *SIAM J. Computing*, 14:505–513, 1985.

[18] P. Shankar. On BCH codes over arbitrary integer rings. *IEEE Trans. Inform. Theory*, 25(4):480–483, 1979.