

MACWILLIAMS IDENTITIES OVER SOME SPECIAL POSETS*

SEDA AKBİYİK AND İRFAN SİAP

ABSTRACT. In this paper we introduce a level weight enumerator for linear binary codes whose index set is a forest. This weight enumerator gives most of the weight enumerators as a special case by specializing its variables. We prove a MacWilliams identity for this weight enumerator over this special family of posets which also generalizes the previous results in literature. Further, both the code and its dual are considered over this family of posets using the definition of this weight enumerator which was not possible before. We conclude by an illustrative example and some remarks.

1. INTRODUCTION

Coding theory has found a well recognised place in the digital era that we are in. It has found applications in transmitting and restoring the digital messages. Encoding and decoding these messages in an efficient way depends on the structure of the codes. To accomplish this goal codes are defined as linear structures i.e. vector subspaces and endowed with a particular metric that serves as measuring the distances between the vectors in order to detect and correct errors. Linear codes first and mainly are considered with the Hamming metric [4]. Later, codes over different metrics due to their applications and purposes are considered. The problem of determining the minimum distance d of codes, i.e. error correcting capacity of codes, was generalized by Neiderreiter [7, 8]. A metric which is called poset (partially ordered set) metric on codes is first considered by Brualdi et. al. [6]. This metric is a very important generalization of the metrics and especially it generalizes the well known and most important metrics such as the Hamming and Rosenbloom-Tsfasman (RT) [5] metrics. Due to this generalization studying codes over this metric has attracted the researchers. However, since it is a generalization the problems are difficult to solve with respect to this metric. One of the main problems

Received by the editors Nov. 30, 2012; Accepted: June 14, 2013.

2000 *Mathematics Subject Classification.* Primary 05C38, 15A15; Secondary 05A15, 15A18.

Key words and phrases. Linear P -codes, P -complete weight enumerator, MacWilliams identity, hierarchical poset, discrete chain poset, tree, forest.

The main results of this paper were presented in part at the conference Algerian-Turkish International Days on Mathematics 2012 (ATIM' 2012) to be held October 9–11, 2012 in Annaba, Algeria at the Badji Mokhtar Annaba University.

is to establish a MacWilliams identity with respect to this metric. This identity enables us to explicitly determine the weight enumerator of its dual algebraically by applying a specific change of variables to the weight enumerator of the original code. The importance becomes more evident when the dimension of the code is too large and hence the dimension of its dual can be very small if the length of the code has a reasonable size which is the case in general. The problem of establishing a MacWilliams identity with respect to a poset metric has been a challenging problem at first. The first attempts of establishing such identities have failed when the researchers considered the same metric for both the code and its dual which has been the case with all previous metrics. In order to overcome this difficulty on poset metrics, the dual of the code is considered over the dual poset and hence a different but a similar metric for the dual space is introduced [2]. Even with this modification it has also shown that not all posets are suitable for obtaining a MacWilliams identity. Some more work on posets and MacWilliams identity is done in [6, 8, 10]. It is proven that the family of hierarchical posets which is a very small family of posets is the only one suitable for this purpose [2]. The authors have introduced a new and more detailed weight enumerator called P-complete weight enumerator to overcome this problem very recently [1]. Therein it is shown that if such a weight enumerator is defined then MacWilliams identity can be obtained and further the dual code is considered over the same metric. The work in [1] is done over a special family of posets, so called discrete chain poset, and this family is different from hierarchical posets. Here, the authors introduce a new level complete weight enumerator which is defined over posets that are represented by forests and the previous results are obtained as a corollary.

The main advantage of defining level complete weight enumerator for codes over posets is that not only we obtain the MacWilliams identity over a considerably large family of posets but further we use the same metric for both the code and its dual which is a new contribution to the literature.

In order to prove the MacWilliams Identity for codes over forests, in the following section we present the basics for binary codes and graph theory that is needed to define the posets presented by the forests. In the next section, we present the definition of level complete weight enumerator over posets represented by forests and present some well known auxiliary lemmas that play an important role in the proof of the main theorem. Next we present a moderate example that illustrates the main theorem. We finalize the paper by some concluding remarks.

2. PRELIMINARIES

Let $\mathbb{Z}_2 = \{0, 1\}$ denote the set of integers modulo 2, which is well known to be a finite field with 2 elements, and $V = \mathbb{Z}_2^n$. The set V is a \mathbb{Z}_2 -vector space. A \mathbb{Z}_2 -vector subspace of V is called a linear code of length n . The inner product of two vectors $v = (v_1, \dots, v_n)$ and $u = (u_1, \dots, u_n)$ defined over V is a \mathbb{Z}_2 -valued function which is $\langle v, u \rangle = \sum_{i=1}^n v_i u_i$. To each linear code C of length n a linear code

$C^\perp = \{v \in V | \langle v, u \rangle = 0, \text{ for all } u \in C\}$ can be associated. The linear code C^\perp is called the dual of C . The Hamming distance between two vectors $v = (v_1, \dots, v_n)$ and $u = (u_1, \dots, u_n)$ is defined by $d_H(v, u) = |\{i | v_i \neq u_i\}|$. It is well known that d_H is a metric on V . Another important notation for codes is the Hamming weight of a vector $v \in V$ which is $w_H(v) = |\{i | v_i \neq 0\}|$. The minimum Hamming distance of a linear code C is $d_{\min}(C) = d_H(C) = \min\{d_H(u, v) | \forall u, v \in C, u \neq v\}$. Also the Hamming weight of a code C is $w_H(C) = \min\{w_H(u) | \forall u \in C, u \neq 0\}$. When the code is linear which is the case in this article, $d_H(C) = w_H(C)$. A linear binary code of length n , dimension k and minimum Hamming distance d is simply denoted by $[n, k, d]$. These three parameters play a crucial role for defining a linear code. Especially the Hamming distance d of a code reveals the quality of the code as shown in the following theorem.

Theorem 2.1. [4] *If C is a linear code of Hamming distance $d = 2t+1$ or $d = 2t+2$, then C can correct up to t errors.*

The interested readers for a more detailed treatment of this subject are welcome to refer to [4, 9, 3].

Let (P, \leq) be a partially ordered set of cardinality n . A subset I of P is called an *ideal* if $x \in I$ and $y \leq x$ imply that $y \in I$. For a subset A of the poset P , $\langle A \rangle$ will denote the smallest ideal of P containing A . We assume that $P = \{1, 2, 3, \dots, n\}$ and the coordinate positions of vectors in \mathbb{Z}_2^n are in one-to-one correspondence with the elements of P . Let $x = (x_1, x_2, \dots, x_n)$ be a vector in \mathbb{Z}_2^n . The P -weight of x is defined as the cardinality $w_P(x) = |\langle \text{supp}(x) \rangle|$ of the smallest ideal of P containing the support of x , where $\text{supp}(x) = \{i \in P : x_i \neq 0\}$. The P - (poset) distance of the elements $x, y \in \mathbb{Z}_2^n$ is defined as $d_P(x, y) = w_P(x - y)$.

If P is an antichain in which no two elements are comparable, then the P -weight and the P -distances reduce to the Hamming weight and the Hamming distance, respectively. If P consists of a single chain, then P -weight and P -distance are Rosenbloom Tsfasman (RT) weight and RT distance. It is known that the P -distance $d_P(.,.)$ is a metric on \mathbb{Z}_2^n . The metric $d_P(.,.)$ on \mathbb{Z}_2^n is called a poset-metric. If \mathbb{Z}_2^n is endowed with a poset-metric, then a subset C of \mathbb{Z}_2^n is called a poset-code. If the poset-metric corresponds to a poset P , then C is called a P -code.

Next, we present some basic definitions from graph theory that will be needed in the next section.

Definition 2.2. A graph $G = (T, E)$ is defined by a finite nonempty set T which is called the set of vertices and a finite set E which is called the set of edges which is a subset of $V \times V$. In general a graph is represented by a diagram consisting of points (vertices) joined by lines (edges).

If $(a, b) \in E$, then we say that an edge between the vertices a and b exists. If we employ a direction from a to b , then the graph is called a directed otherwise undirected. In the case of the directed graphs, geometrically when drawing them,

along the edges an arrow that points the direction is used. In this paper all graphs are assumed to be undirected.

Definition 2.3. If all vertices in a graph are connected to each other by at least one edge, then the graph is called a connected graph. Otherwise it is called non-connected.

Some families of graphs induce posets by defining a natural relation between the vertices. An important family of connected graphs is the hierarchical poset (Figure 1). A graph that consists of disjoint union of chains (Figure 2) induces a poset called a discrete chain poset which is an example for a non connected graph.

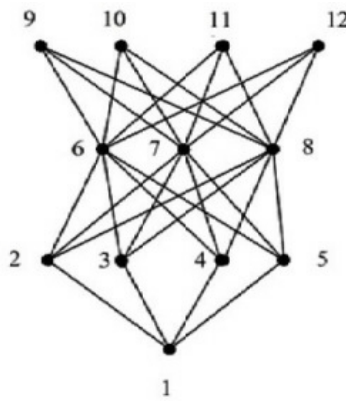


FIGURE 1. A poset

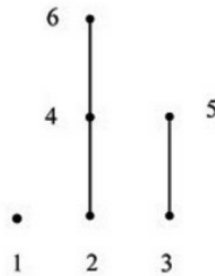


FIGURE 2. A discrete chain poset.

Definition 2.4. In a graph G a sequence of k connected edges is called a walk of length k . If starting and the final vertex are the same, then the walk is called a closed walk. In a walk if every edge is different, then the walk is a trace. Also, if all vertices are different too, then the trace is called a road. In a closed walk if all edges are different, then the walk is called a closed trace. Also, if all vertices are

different too, then it is called a cycle. A connected graph, without a cycle is called a tree. An example of a tree is given in Figure 3.

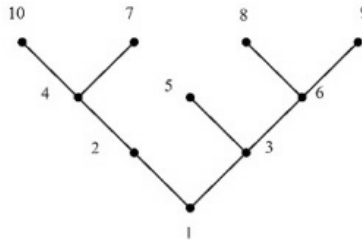


FIGURE 3. A tree

Definition 2.5. A forest is a disjoint union of trees (Figure 4). The levels on the forests are defined as the number of edges (distance) from the root (upside down-preferred for convenience here). So, in each tree we have level one vertices that are one edge of distance from the roots and level two ones that are two edges apart from the roots, and so on.

In Figure 4, the level one vertices are $\{1, 2, 3\}$. The level two vertices are $\{4, 5, 6, 7\}$. The level three vertices are $\{8, 9\}$.

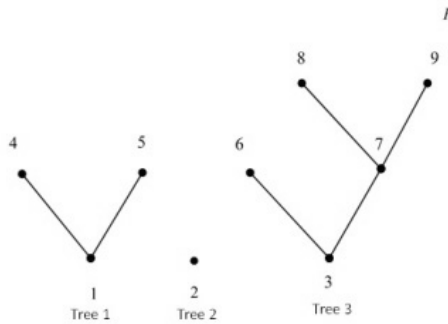


FIGURE 4. A forest

3. LEVEL WEIGHT ENUMERATOR AND THE MACWILLIAMS IDENTITY

In this section we define the level weight enumerator and prove a MacWilliams identity. First we define necessary notations and terms and present some auxiliary statements.

Definition 3.1. [2, 5, 10] Let C be a linear P-code of length n . The poset weight enumerator of C is defined by $W_{C,P}(x) = \sum_{u \in C} x^{w_P(u)} = \sum_{i=0}^n A_{i,P} x^i$, where $A_{i,P} = |\{u \in C \mid w_P(u) = i\}|$.

Example 3.2. [2] Let $P = \{1, 2, 3\}$ be a poset with order relation $1 < 2 < 3$. Consider the binary linear P- codes $C_1 = \{000, 001\}$ and $C_2 = \{000, 111\}$.

Then the poset weight enumerator of C_1 and C_2 is given by $W_{C_1,P}(x) = 1 + x^3 = W_{C_2,P}(x)$. The dual codes of C are $C_1 = \{000, 100, 010, 110\}$ and $C_2 = \{000, 110, 101, 011\}$, respectively. The P- weight enumerators of the dual codes are given by $W_{C_1^\perp,P}(x) = 1 + x + 2x^2$ and $W_{C_2^\perp,P}(x) = 1 + x^2 + 2x^3$.

Definition 3.3. Let P be a poset which has n vertices, s levels, and C be a binary linear code defined on the poset P . Such a code is referred to as a P - code. Then the level complete weight enumerator of C is defined as

$$W_{C,P}(z_1, z_2, \dots, z_s) = \sum_{u \in C} \prod_{i=1}^s z_i^{w_H(u_i)}$$

where u_i denotes the index part of the codeword which is in the i th level of the code.

Example 3.4. Consider the poset codes C_1 and C_2 in Example 3.2 with the same poset P . According to the Definition 3.3 the P- level weight enumerator of these codes are given by $W_{C_1,P}(z_1, z_2, z_3) = 1 + z_3$ and $W_{C_2,P}(z_1, z_2, z_3) = 1 + z_1 z_2 z_3$. The level weight enumerators of the dual codes are given by $W_{C_1^\perp,P}(z_1, z_2, z_3) = 1 + z_1 + z_2 + z_1 z_2$, and $W_{C_2^\perp,P}(z_1, z_2, z_3) = 1 + z_1 z_2 + z_1 z_3 + z_2 z_3$.

Definition 3.5. Let F be a forest which has k trees, n vertices and s levels, and C be a binary linear code defined on the forest F . Such a code is referred to as a P - code. Then the level complete weight enumerator of C is defined as

$$W_{C,F}(z_1^{(1)} \dots, z_s^{(1)} \dots, z_1^{(k)} \dots, z_s^{(k)}) = \sum_{u \in C} \prod_{j=1}^k \prod_{i=1}^s (z_i^{(j)})^{w_H(u_i^{(j)})}$$

where $z_i^{(j)}$ denotes the index part of the codeword which is in the j th level of the i th tree.

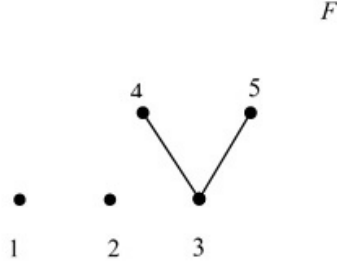


FIGURE 5. The forest F

Example 3.6. Let $C = \{00000, 10110, 01011, 11101\}$ be a P -code on forest F shown in Figure 5 which has two 1-leveled trees and a 2-leveled tree. Then the level complete weight polynomial of C on forest F is

$$W_{C,F}(z_1^{(1)}, z_2^{(1)}, z_3^{(1)}, z_3^{(2)}) = 1 + z_1^{(1)} z_3^{(1)} z_3^{(2)} + z_2^{(1)} (z_3^{(2)})^2 + z_1^{(1)} z_2^{(1)} z_3^{(1)} z_3^{(2)}.$$

To prove the main theorem, we present the following auxiliary lemmas whose proofs can be found in [4].

Lemma 3.7. [4] *Let C be a binary linear code of length n and $\chi_u(v) = (-1)^{\langle u,v \rangle}$ for every $u, v \in C$. For a fixed v , if $v \notin C^\perp$, then*

$$\sum_{u \in C} \chi_u(v) = 0$$

and if $v \in C^\perp$, then

$$\sum_{u \in C} \chi_u(v) = |C|.$$

Lemma 3.8. [4] *Let C be a binary linear code of length n and $f : Z_2^n \rightarrow C[z_1, z_2, \dots, z_s]$ be a function. Then,*

$$\sum_{v \in C^\perp} f(v) = \frac{1}{|C|} \sum_{u \in C} \tilde{f}(u),$$

where

$$\tilde{f}(u) = \sum_{v \in Z_2^n} (-1)^{\langle u,v \rangle} f(v)$$

for all $u \in Z_2^n$.

Theorem 3.9. *If C is a P -code on a forest F composed by k trees, n vertices and s levels and C^\perp is the dual code of C , then*

$$\begin{aligned} & W_{C^\perp, F}(z_1^{(1)}, \dots, z_s^{(1)}, \dots, z_1^{(k)}, \dots, z_s^{(k)}) \\ &= \frac{1}{|C|} \prod_{j=1}^k \prod_{i=1}^s (1 + z_i^{(j)})^{n_i^{(j)}} W_{C, F} \left(\frac{1 - z_1^{(1)}}{1 + z_1^{(1)}}, \dots, \frac{1 - z_s^{(1)}}{1 + z_s^{(1)}}, \dots, \frac{1 - z_1^{(k)}}{1 + z_1^{(k)}}, \dots, \frac{1 - z_s^{(k)}}{1 + z_s^{(k)}} \right) \end{aligned}$$

where the length of each part of $u \in C$ in i th tree and level j is denoted by $n_i^{(j)}$.

Proof. In order to apply Lemma 3.8 we first define a function f that represents the terms in the level weight enumerator such that

$$\begin{aligned} f(v) &= (z_1^{(1)})^{w_H(v_1^{(1)})} (z_2^{(1)})^{w_H(v_2^{(1)})} \dots (z_s^{(1)})^{w_H(v_s^{(1)})} \dots (z_1^{(k)})^{w_H(v_1^{(k)})} (z_2^{(k)})^{w_H(v_2^{(k)})} \\ &\quad \dots (z_s^{(k)})^{w_H(v_s^{(k)})} \\ &= \prod_{j=1}^k \prod_{i=1}^s (z_i^{(j)})^{w_H(v_i^{(j)})}. \end{aligned}$$

Then by Lemma 3.8,

$$\begin{aligned} \tilde{f}(u) &= \sum_{v \in Z_2^n} (-1)^{\langle u, v \rangle} f(v) \\ &= \sum_{v \in Z_2^n} (-1)^{\langle u, v \rangle} \prod_{j=1}^k \prod_{i=1}^s (z_i^{(j)})^{w_H(v_i^{(j)})} \\ &= \sum_{v_1^{(1)} \in Z_2^{n_1^{(1)}}} \dots \sum_{v_s^{(k)} \in Z_2^{n_s^{(k)}}} (-1)^{\sum_{i=1}^k \sum_{j=1}^{n_i} u_i^{(j)} v_i^{(j)}} \prod_{j=1}^k \prod_{i=1}^s (z_i^{(j)})^{w_H(v_i^{(j)})} \\ &= \sum_{v_1^{(1)} \in Z_2^{n_1^{(1)}}} \dots \sum_{v_s^{(k)} \in Z_2^{n_s^{(k)}}} \prod_{j=1}^k \prod_{i=1}^s (-1)^{\sum_{i=1}^k \sum_{j=1}^{n_i} u_i^{(j)} v_i^{(j)}} (z_i^{(j)})^{w_H(v_i^{(j)})} \\ &= \prod_{j=1}^k \prod_{i=1}^s \left(\sum_{v_i^{(j)} \in Z_2^{n_i^{(j)}}} (-1)^{\sum_{i=1}^k \sum_{j=1}^{n_i} u_i^{(j)} v_i^{(j)}} (z_i^{(j)})^{w_H(v_i^{(j)})} \right) \\ &= \prod_{j=1}^k \prod_{i=1}^s (1 + z_i^{(j)})^{n_i^{(j)}} \cdot \left(\frac{1 - z_i^{(j)}}{1 + z_i^{(j)}} \right)^{w_H(u_i^{(j)})} \end{aligned}$$

Again by using Lemma 3.8, we find that

$$\sum_{u \in C^\perp} f(u) = \frac{1}{|C|} \sum_{u \in C} \tilde{f}(u) = \frac{1}{|C|} \prod_{j=1}^k \prod_{i=1}^s (1 + z_i^{(j)})^{n_i^{(j)}} \left(\frac{1 - z_i^{(j)}}{1 + z_i^{(j)}} \right)^{w_H(u_i^{(j)})}.$$

So we obtain

$$\begin{aligned} W_{C^\perp, F}(z_1^{(1)} \dots, z_s^{(1)} \dots, z_1^{(k)} \dots, z_s^{(k)}) \\ = \frac{1}{|C|} \prod_{j=1}^k \prod_{i=1}^s (1 + z_i^{(j)})^{n_i^{(j)}} W_{C, F} \left(\frac{1 - z_1^{(1)}}{1 + z_1^{(1)}} \dots, \frac{1 - z_s^{(1)}}{1 + z_s^{(1)}} \dots, \frac{1 - z_1^{(k)}}{1 + z_1^{(k)}} \dots, \frac{1 - z_s^{(k)}}{1 + z_s^{(k)}} \right). \end{aligned}$$

□

Example 3.10. Let C be the linear code defined in Example ?? . Then, by applying Theorem 3.9 we can find the level complete weight enumerator of the dual code C^\perp on forest F as follows:

$$\begin{aligned} &W_{C,F}(z_1^{(1)}, z_2^{(1)}, z_3^{(1)}, z_3^{(2)}) \\ &= \frac{1}{4}(1 + z_1^{(1)})(1 + z_2^{(1)})(1 + z_3^{(1)})(1 + z_3^{(2)})^2 W_{C,F} \left(\frac{1 - z_1^{(1)}}{1 + z_1^{(1)}}, \frac{1 - z_2^{(1)}}{1 + z_2^{(1)}}, \frac{1 - z_3^{(1)}}{1 + z_3^{(1)}}, \frac{1 - z_3^{(2)}}{1 + z_3^{(2)}} \right) \\ &= 1 + z_1^{(1)} z_3^{(1)} + z_1^{(1)} z_2^{(1)} z_3^{(2)} + z_2^{(1)} z_3^{(2)} + z_2^{(1)} z_3^{(1)} z_3^{(2)} + z_1^{(1)} z_2^{(1)} z_3^{(1)} z_3^{(2)} + z_1^{(1)} (z_3^{(2)})^2 \\ &\quad + z_3^{(1)} (z_3^{(2)})^2. \end{aligned}$$

3.1. MacWilliams Identity on Trees. Now by taking $k = 1$ in a forest we obtain a tree and similarly in (3.5), we obtain the level complete weight enumerator for trees.

Definition 3.11. Let T be a tree which has n vertices and s levels, and C be a binary linear poset code defined on a tree T . Then the level complete weight enumerator of C is defined as

$$W_{C,T}(z_1, z_2, \dots, z_s) = \sum_{u \in C} \prod_{i=1}^s z_i^{w_H(u_i)}.$$

Corollary 1. If C is a P -code on n vertices and s levels of a tree T and C^\perp be the dual code of C , then

$$W_{C^\perp,T}(z_1, z_2, \dots, z_s) = \frac{1}{|C|} \prod_{i=1}^s (1 + z_i)^{n_i} W_{C,T} \left(\frac{1 - z_1}{1 + z_1}, \dots, \frac{1 - z_s}{1 + z_s} \right)$$

where the length of parts of $u \in C$ in level i is denoted by n_i .

Proof. Simply follows by taking $k = 1$ in Definition (3.5). □

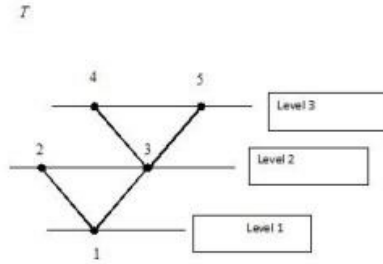


FIGURE 6. The tree T

Example 3.12. Let $C = \{00000, 10110, 01001, 11111\}$ be a P -code on tree T in figure 6 which has length 5 and 3 levels. Then the P -complete weight polynomial of C on tree T is

$$W_{C,T}(z_1, z_2, \dots, z_s) = 1 + z_1 z_2 z_3 + z_2 z_3 + z_1 z_2^2 z_3^2.$$

So by applying Theorem 3.9 we can find the level complete weight enumerator of C^\perp on the tree T ,

$$\begin{aligned} W_{C^\perp, T}(z_1, z_2, z_3) &= \frac{1}{4}(1+z_1)(1+z_2)^2(1+z_3)^2 W_{C,T}\left(\frac{1-z_1}{1+z_1}, \frac{1-z_2}{1+z_2}, \frac{1-z_3}{1+z_3}\right) \\ &= 1 + z_1 z_2 + z_1 z_3 + 2z_2 z_3 + z_1 z_2^2 z_3 + z_1 z_2 z_3^2 + z_2^2 z_3^2. \end{aligned}$$

4. CONCLUSION

Here we define a level weight enumerator for binary codes whose index set is over a forest which falls into family of poset codes. This definition enabled us to establish the MacWilliams Identity for both the code and its dual code over the same metric. This was shown to be impossible if the weight enumerator is defined in a different way by researchers in the literature [2]. Poset codes in general are more difficult to study because they generalize many metrics including the most important ones such as Hamming and Rosenbloom-Tsfasman. This new approach is believed that will attract many researchers to study it further.

REFERENCES

- [1] S. Akbiyik, I. Siap, *A P-Complete weight enumerator with respect to poset metric and its MacWilliams identity*, (Turkish), Adiyaman University Journal of Science, **1** (1) (2011) 28–39.
- [2] H.K. Kim, D.Y. Oh, *A classification of posets admitting the MacWilliams identity*, IEEE Transactions on Information Theory **51** (4) (2005) 1424–1431.
- [3] W.C. Huffman, V.Pless, *Fundamentals of Error-Correcting Codes*, Cambridge University Press, 2003.
- [4] F.J. MacWilliams, N.J. Sloane, *The Theory of Error-Correcting Codes*, Amsterdam, The Netherlands: North-Holland, 1977.
- [5] S.T. Dougherty, M.M. Skriganov, *MacWilliams Duality and the Rosenbloom-Tsfasman Metric*, Mosc. Math. J. **2** (1) (2002) 81–97.
- [6] R.A. Brualdi, J.S. Graves, K.M. Lawrence, *Codes with a poset metric*, Discrete Math. **147** (1995) 57–72.
- [7] H. Niederreiter, *Points Sets and Sequences with Small Discrepancy*, Monatsh. Math. **104** (1987) 221–228.
- [8] H. Niederreiter, *A combinatorial problem for vector spaces over finite fields*, Discrete Math. **96** (1991) 273–337.
- [9] S. Roman, *Introduction to Coding and Information Theory*, Springer, New York, Berlin, Heidelberg, 1997.
- [10] J. N. Gutierrez, H. Tapia-Recillas, *A MacWilliams identity for poset-codes*, Congr. Numer. **133** (1998) 63–73.

Current address: Seda Akbıyık and İrfan Siap; Yıldız Technical University, Department of Mathematics, İstanbul, TURKEY

E-mail address: akbiyiks@yildiz.edu.tr, isiap@yildiz.edu.tr

URL: <http://communications.science.ankara.edu.tr/index.php?series=A1>