and

$$\left|\mathrm{Im}\left\{\sum_{x\in\mathcal{T}} e(f(x))\right\}\right| \le (D-1)\sqrt{q/2}.$$

This leads to minor improvements in (4)–(6) in the coefficient of the term arising from the case $k = 0$.

For $P(m, 2)$ and the small set of Kasami sequences of length $L-1$ we have approximately equal maximum even correlation. The Kasami set has considerably fewer sequences, however, the best known upper bound (see [7]) for their maximum aperiodic correlation has $6\sqrt{2}/\pi$ as the coefficient of $\sqrt{q}\ln(4L/\pi)$ where we have $8/\pi$.

REFERENCES

[1] A. Barg, "On small families of sequences with low periodic correlation," in *Lecture Notes in Computer Science*, vol. 781. Berlin, Germany: Springer-Verlag, 1994, pp. 154–158.

[2] S. Boztas, R. Hammons, and P. V. Kumar, "4-phase sequences with near optimum correlation properties," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1101–1113, May 1992.

[3] T. Helleseth and P. V. Kumar, "Sequences with low correlation," in *Handbook of Coding Theory*, R. Brualdi, C. Huffman, and V. Pless, Eds., preprint.

[4] T. Helleseth, P. V. Kumar, O. Moreno, and A. G. Shanbhag, "Improved estimates for the minimum distance of weighted degree Z_4 trace codes," in *Proc. 1995 IEEE Int. Symp. on Information Theory* (Whistler, B.C., Canada, Sept. 17–22, 1995).

[5] S. M. Krone and D. V. Sarwate, "Quadriphase sequences for spread-spectrum multiple-access communication," *IEEE Trans. Inform. Theory*, vol. 30, pp. 520–529, May 1984.

[6] P. V. Kumar, T. Helleseth, and A. R. Calderbank, "An upper bound for Weil exponential sums over Galois rings and applications," *IEEE Trans. Inform. Theory*, vol. 41, pp. 456–468, Mar. 1995.

[7] J. Lahtonen, "On the odd and the aperiodic correlation properties of the Kasami sequences," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1506–1508, Sept. 1995.

[8] S. Litsyn and A. Tietäväinen, "Character sum constructions of constrained error-correcting codes," *Appl. Algebra in Eng., Commun. and Comp.*, vol. 5, pp. 45–51, 1994.

[9] A. A. Nechaev, "Kerdock code in a cyclic form," *Discr. Math. Appl.*, vol. 1, pp. 365–384, 1991.

[10] D. V. Sarwate, "An upper bound on the aperiodic autocorrelation function for a maximal-length sequence," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 685–687, July 1984.

[11] A. G. Shanbhag, P. V. Kumar, and T. Helleseth, "An upper bound for the aperiodic correlation of weighted-degree CDMA sequences," in *Proc. 1995 IEEE Int. Symp. on Information Theory* (Whistler, B.C., Canada, Sept. 17–22, 1995).

[12] ——, "Improved binary codes and sequence families from Z_4-linear codes," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1582–1587, Sept. 1996.

[13] H. Tarnanen, "An elementary proof to the weight distribution formula of the first order shortened Reed–Muller coset code," preprint.

[14] P. Udaya and M. U. Siddiqi, "Optimal biphase sequences with large linear complexity derived from sequences over Z_4," *IEEE Trans. Inform. Theory*, vol. 42, pp. 206–217, Jan. 1996.

[15] I. M. Vinogradov, *Elements of Number Theory.* New York: Dover, 1954.

New Construction for Families of Binary Sequences with Optimal Correlation Properties

Jong-Seon No, Kyeongcheol Yang, *Member, IEEE*, Habong Chung, *Member, IEEE*, and Hong-Yeop Song, *Member, IEEE*

***Abstract*—In this correspondence, we present a construction, in a closed form, for an optimal family of 2^m binary sequences of period $2^{2m}-1$ with respect to Welch's bound, whenever there exists a balanced binary sequence of period 2^m-1 with ideal autocorrelation property using the trace function. This construction enables us to reinterpret a small set of Kasami and No sequences as a family constructed from *m*-sequences. New optimal families of binary sequences are constructed from the Legendre sequences of Mersenne prime period, Hall's sextic residue sequences, and miscellaneous sequences of unknown type. In addition, we enumerate the number of distinct families of binary sequences, which are constructed from a given binary sequence by this method.**

***Index Terms*—Kasami sequences, Legendre sequences, No sequences, optimal correlation property, signature sequences.**

I. INTRODUCTION

Code-division multiple access (CDMA) systems use pseudonoise binary sequences as signature sequences, and several spread-spectrum communication systems also use them as spreading codes for low probability of intercept [18], [20]. Desirable characteristics of a family of binary sequences for such applications include long-period, low out-of-phase autocorrelation values, low crosscorrelation values, low nontrivial partial-period correlation values, large linear span, balance of symbols, large family size, and ease of implementation.

A binary (0 or 1) sequence $\{b(t), t = 0, 1, \cdots, N-1\}$ of period $N = 2^n - 1$ is called *balanced* if the number of 1's is one more than the number of 0's [8]. It is said to have the *ideal autocorrelation property* if its periodic autocorrelation function $R(\tau)$ is given by

$$R(\tau) = \begin{cases} N, & \text{for } \tau \equiv 0 \bmod N \\ -1, & \text{for } \tau \not\equiv 0 \bmod N \end{cases}$$

where $R(\tau)$ is defined as

$$R(\tau) = \sum_{t=0}^{N-1} (-1)^{b(t+\tau)+b(t)}$$

and $t+\tau$ is computed modulo N. Note that $R(\tau)$ is the number of agreements minus the number of disagreements between $\{b(t)\}$ and $\{b(t+\tau)\}$ as t runs from 0 to $N-1$ [7], [8], [21]. It is well known that the ideal autocorrelation property implies the balance property.

Let $\{b(t)\}$ and $\{c(t)\}$ be two binary sequences of period N. Two sequences $\{b(t)\}$ and $\{c(t)\}$ are said to be *cyclically equivalent*

Manuscript received April 26, 1996; revised February 12, 1997. This work was supported in part by the Korean Ministry of Information and Communications. The material in this correspondence was presented in part at the 1996 IEEE International Symposium of Information Theory and Its Applications (ISITA'96), Victoria, BC, Canada, September 17–20, 1996

J.-S. No is with the Department of Electronic Engineering, Konkuk University, Seoul 143-701, Korea.

K. Yang is with the Department of Electronic Communication Engineering, Hanyang University, Seoul 133-791, Korea.

H. Chung is with the Department of Electronic Engineering, Hong-Ik University, Seoul 121-791, Korea.

H.-Y. Song is with the Department of Electronic Engineering, Yonsei University, Seoul 120-749, Korea.

Publisher Item Identifier S 0018-9448(97)05017-7.

if there exists an integer τ such that $c(t) = b(t+\tau)$ for all t. Otherwise, they are said to be *cyclically distinct*. For an integer r, the sequence $\{c(t)\}$ is called the *decimation* by r of the sequence $\{b(t)\}$ if $c(t) = b(rt)$ for any integer t. It is easily checked that the period of $\{c(t) = b(rt)\}$ is given by N divided by $\gcd(r, N)$. It is also well known that if a sequence $\{b(t)\}$ of period N has the ideal autocorrelation property, so does its decimation $\{b(rt)\}$ by r, where r is an integer relatively prime to N. Two sequences $\{b(t)\}$ and $\{c(t)\}$ are said to be *equivalent* if there are some integers r and τ such that $c(t+\tau) = b(rt)$ for all t. They are said to be *inequivalent*, otherwise.

Consider a set of J binary sequences, each with period N, denoted by

$$\{v^{(j)}(t), \qquad t = 0, 1, \cdots, N-1\}, \qquad j = 1, 2, \cdots, J.$$

The periodic crosscorrelation $R_{jk}(\tau)$ at shift τ between two sequences $\{v^{(j)}(t)\}$ and $\{v^{(k)}(t)\}$ from this collection is defined as

$$R_{jk}(\tau) = \sum_{t=0}^{N-1} (-1)^{v^{(j)}(t+\tau)+v^{(k)}(t)}.$$

The maximum out-of-phase periodic autocorrelation magnitude R_A for this signal set is defined as

$$R_A = \max_j \max_{0<\tau<N} |R_{jj}(\tau)|$$

and the maximum crosscorrelation magnitude R_C between sequences in this set is given by

$$R_C = \max_{j\neq k} \max_{0\leq\tau<N} |R_{jk}(\tau)|.$$

The criterion for signal design is to minimize

$$R_{\max} = \max(R_A, R_C).$$

In signal design, the Welch bound and the Sidelnikov bound are used to test the optimality of sequence sets. Some of well-known optimal families of binary sequences include Gold sequences [6], Kasami sequences [18], [20], bent sequences [12], [20], and No sequences [15]. Gold sequences form an optimal set with respect to Sidelnikov's bound [19] which states that for any set of N or more binary sequences of period N

$$R_{\max} > (2N-2)^{1/2}.$$

The small set of Kasami sequences is an optimal collection of binary sequences with respect to Welch's bound [22], which implies that

$$R_{\max} \geq 1 + 2^{n/2}$$

when it is applied to a set of $2^{n/2}$ sequences of period $N = 2^n - 1$ for an even integer n. Bent and No sequences also form an optimal set with respect to Welch's bound, respectively, but they have larger linear spans than Gold sequences and Kasami sequences.

In this correspondence, we show that if a binary sequence of period $2^m - 1$ in a trace expression has the ideal autocorrelation property, it can be used to construct, in a closed form, a family of 2^m binary sequences of period $2^{2m} - 1$ with optimal correlation with respect to Welch's bound. This construction method enables us to reinterpret the small set of Kasami sequences as well as the No sequences as a family constructed from the m-sequences. New optimal families of binary sequences are constructed from the Legendre sequences of Mersenne prime period, Hall's sextic residue sequences, and miscellaneous sequences of unknown type. In addition, we enumerate the number of distinct families of binary sequences, which are constructed from a given binary sequence by this method.

This correspondence is organized as follows. In Section II, we present the main results to construct an optimal family of binary sequences with respect to Welch's bound. In Section III, the small set of Kasami sequences and the No sequences are reinterpreted as a family constructed from the m-sequences. New optimal families of binary sequences are constructed from the Legendre sequences of Mersenne prime period in Section IV. Hall's sextic residue sequences and miscellaneous sequences of unknown type are also considered in Section IV.

II. Construction of a Family of Binary Sequences with Optimal Correlation

Let q be a prime power and F_q be the finite field with q elements. Let $n = em > 1$ for some positive integers e and m. Then the trace function $\mathrm{tr}_m^n(\cdot)$ from F_{2^n} to its subfield F_{2^m} is a mapping [10], [11] given by

$$\mathrm{tr}_m^n(x) = \sum_{i=0}^{e-1} x^{2^{mi}}.$$

No *et al.* [17] presented a closed-form expression of binary sequences of longer period with ideal autocorrelation property in a trace representation, if a given binary sequence with ideal autocorrelation property is described using the trace function. The idea of extension in [17] will be helpful for our further discussion, so it is quoted without proof in the following theorem.

Theorem 1 [17]: Let m and n be positive integers such that $m \mid n$. Let α be a primitive element of F_{2^n} and set $\beta = \alpha^T$ where $T = (2^n-1)/(2^m-1)$. Assume that for an index set I, the sequence $\{b(t_1), t_1 = 0, 1, \cdots, M-1\}$ of period $M = 2^m - 1$ given by

$$b(t_1) = \sum_{a\in I} \mathrm{tr}_1^m(\beta^{at_1})$$

has the ideal autocorrelation property. For any integer r, $1 \leq r \leq M-1$, relatively prime to M, the sequence

$$\{c(t), t = 0, 1, \cdots, N-1\}$$

of period $N = 2^n - 1$ defined by

$$c(t) = \sum_{a\in I} \mathrm{tr}_1^m([\mathrm{tr}_m^n(\alpha^t)]^{ar})$$

also has the ideal autocorrelation property.

Based on the idea of extension in Theorem 1, we will provide a method to construct an optimal family of 2^m binary sequences of period $2^{2m} - 1$ from a given binary sequence of period $2^m - 1$ with ideal autocorrelation property. Throughout the correspondence, we use the following notations. Let m and n be positive integers such that $m \mid n$. Let α be a primitive element in F_{2^n} and set $\beta = \alpha^{(2^n-1)/(2^m-1)}$. Note that β is primitive in F_{2^m}. From now on, we assume that the sequence $\{b(t_1), t_1 = 0, 1, \cdots, M-1\}$ of period $M = 2^m - 1$ given by

$$b(t_1) = \sum_{a\in I} \mathrm{tr}_1^m(\beta^{at_1}) \tag{1}$$

has the ideal autocorrelation property for an index set I.

Theorem 2: Let $n = 2m$, and let $\{s^{(j)}(t)\}$ be the sequence given by

$$s^{(j)}(t) = \sum_{a\in I} \mathrm{tr}_1^m([\mathrm{tr}_m^n(\alpha^{2t}) + \gamma_j\beta^t]^{ar}), \qquad \text{for } \gamma_j \in F_{2^m}$$

where r, $1 \leq r \leq M-1$, is an integer relatively prime to $M = 2^m - 1$, and the index set I is in (1). Define the family $\mathcal{F}$ of 2^m binary sequences of period $N = 2^n - 1$ as

$$\mathcal{F} = \{\{s^{(j)}(t), t = 0, 1, \cdots, N-1\} \mid j = 1, 2, \cdots, 2^m\}.$$

Then the family $\mathcal{F}$ is an optimal set of 2^m binary sequences of period N with respect to Welch's bound. Furthermore, $R_{ij}(\tau)$ takes only a value -1, 2^m-1, or -2^m-1 for any i, j, and τ except for the case where $i=j$ and $\tau \equiv 0 \pmod N$.

Proof: We will show that the possible values of $R_{ij}(\tau)$ are -1, 2^m-1, or -2^m-1 for any i, j, and τ except for the case where $i=j$ and $\tau \equiv 0 \pmod N$. Let $T=2^m+1$. Since $\gcd(2^m-1, T)=1$, any integer t, $0 \le t \le 2^{2m}-2$, can be uniquely written as

$$t = t_1 T + t_2(2^m-1), \qquad 0 \le t_1 \le 2^m-2,\ 0 \le t_2 \le 2^m.$$

Then each sequence $\{s^{(i)}(t),\ t=0,\ 1,\ \cdots,\ 2^n-2\}$ becomes

$$\begin{aligned} s^{(i)}(t) &= \sum_{a\in I} \mathrm{tr}_1^m \left\{ [\mathrm{tr}_m^n (\alpha^{2t_1T+2t_2\{2^m-1\}}) \right. \\ &\qquad \left. + \gamma_i \beta^{t_1T+t_2(2^m-1)}]^{ar} \right\} \\ &= \sum_{a\in I} \mathrm{tr}_1^m \{\beta^{2art_1}[\mathrm{tr}_m^n (\alpha^{2t_2\{2^m-1\}}) + \gamma_i]^{ar}\} \end{aligned}$$

since $\alpha^{2t_1T} \in F_{2^m}$ and $\beta^T = \beta^2$. For short notation, we define

$$f(\gamma_i, t_2) = \mathrm{tr}_m^n [\alpha^{2t_2(2^m-1)}] + \gamma_i.$$

Then we have

$$s^{(i)}(t) = \sum_{a\in I} \mathrm{tr}_1^m \{\beta^{2art_1}[f(\gamma_i, t_2)]^{ar}\}.$$

Similarly, we have

$$s^{(j)}(t+\tau) = \sum_{a\in I} \mathrm{tr}_1^m \{\beta^{2ar(t_1+\tau_1)}[f(\gamma_j, t_2+\tau_2)]^{ar}\}$$

where τ, $0 \le \tau \le 2^{2m}-2$, is also uniquely written as

$$\tau = \tau_1 T + \tau_2(2^m-1), \qquad 0 \le \tau_1 \le 2^m-2,\ 0 \le \tau_2 \le 2^m.$$

Thus we get the equation at the bottom of this page. Note that the inner sum

$$\sum_{t_1=0}^{2^m-2} (-1)^{\sum_{a\in I} \mathrm{tr}_1^m(\beta^{2art_1}([f(\gamma_i,t_2)]^{ar}+[\beta^{2\tau_1}f(\gamma_j,t_2+\tau_2)]^{ar}))}$$

yields 2^m-1 when

$$f(\gamma_i, t_2) = \beta^{2\tau_1} f(\gamma_j, t_2+\tau_2).$$

When

$$f(\gamma_i, t_2) \ne \beta^{2\tau_1} f(\gamma_j, t_2+\tau_2)$$

we claim that the inner sum is -1. If either $f(\gamma_i, t_2)=0$ or $f(\gamma_j, t_2+\tau_2)=0$, the exponent to (-1) in the inner sum is essentially a shift of the sequence $\{b(2rt_1)\}$. Since $\gcd(2r, M)=1$, it is obvious that the sequence $\{b(2rt_1)\}$ is balanced and has the ideal autocorrelation property. This implies that the inner sum gives -1. On the other hand, if $f(\gamma_i, t_2) \ne 0$ and $f(\gamma_j, t_2+\tau_2) \ne 0$, the inner sum is the autocorrelation of the sequence $\{b(2rt_1)\}$ at a nonzero shift $\pmod N$, so it is -1 by the assumption. Thus the inner sum always yields -1 if

$$f(\gamma_i, t_2) \ne \beta^{2\tau_1} f(\gamma_j, t_2+\tau_2).$$

Therefore, it is sufficient to find the size of the set of τ_2's such that the inner sum gives the value 2^m-1 in order to compute $R_{ij}(\tau)$. Let

$$\Lambda = \{t_2 | 0 \le t_2 \le 2^m,\ f(\gamma_i, t_2) = \beta^{2\tau_1} f(\gamma_j, t_2+\tau_2)\}.$$

Then we have

$$\begin{aligned} R_{ij}(\tau) &= (2^m-1)\cdot|\Lambda| + (-1)\cdot(2^m+1-|\Lambda|) \\ &= 2^m|\Lambda| - (2^m+1). \end{aligned} \tag{2}$$

By defining $x = \alpha^{2t_2(2^m-1)}$ and $u = \alpha^{2\tau_2(2^m-1)}$, we have

$$|\Lambda| = |\{x | x + x^{2^m} + \gamma_i = \beta^{2\tau_1}(ux + u^{2^m}x^{2^m} + \gamma_j)\}|.$$

Note that $x \in F_{2^{2m}} \backslash \{0\}$ and $x^{2^{2m}} = x$, so we get

$$x^{2^m} = \alpha^{2t_2(2^m-1)\cdot 2^m} = \alpha^{-2t_2(2^m-1)} = x^{-1}.$$

Similarly, we have $u^{2^m} = u^{-1}$. Thus

$$\begin{aligned} |\Lambda| &= |\{x | x + x^{-1} + \gamma_i = \beta^{2\tau_1}(ux + (ux)^{-1} + \gamma_j)\}| \\ &= |\{x | x^2 + 1 + \gamma_i x = \beta^{2\tau_1}(ux^2 + u^{-1} + \gamma_j x)\}|. \end{aligned}$$

The degree of the polynomial in x is at most 2, which means $|\Lambda| \le 2$. Hence we conclude that

$$R_{ij}(\tau) \in \{-2^m-1, -1, 2^m-1\}$$

from (2). □

Theorem 3: Let $k=2n$, and let γ be a primitive element of F_{2^k}. Set $\beta = \gamma^{(2^k-1)/(2^m-1)}$ and $\alpha = \gamma^{(2^k-1)/(2^n-1)}$. Let $\{s^{(j)}(t)\}$ be the sequence given by

$$s^{(j)}(t) = \sum_{a\in I} \mathrm{tr}_1^m([\mathrm{tr}_m^n([\mathrm{tr}_n^k(\gamma^{2t}) + \delta_j \alpha^t]^u)]^{ar})$$

for $\delta_j \in F_{2^n}$ and the index set I in (1), where r, $1 \le r \le M-1$, is an integer relatively prime to $M=2^m-1$, and u, $1 \le u \le N-1$, is an integer relatively prime to $N=2^n-1$. Define the family $\mathcal{F}$ of 2^n binary sequences of period $K=2^k-1$ as

$$\mathcal{F} = \{\{s^{(j)}(t),\ t=0,\ 1,\ \cdots,\ K-1\} \,|\, j=1,\ 2,\ \cdots,\ 2^n\}.$$

Then the family $\mathcal{F}$ is an optimal set of 2^n binary sequences of period K with respect to Welch's bound, and $R_{ij}(\tau)$ takes only a value -1, 2^n-1, or -2^n-1 for any i, j, and τ except for the case where $i=j$ and $\tau \equiv 0 \pmod K$.

Proof: By Theorem 1, the sequence $\{b(t_1)\}$ in (1) can be extended to a sequence $\{c(t_2),\ t_2=0,\ 1,\ \cdots,\ N-1\}$ of period $N=2^n-1$ with ideal autocorrelation property given by

$$c(t_2) = \sum_{a\in I} \mathrm{tr}_1^m([\mathrm{tr}_m^n(\alpha^{t_2})]^{ar}).$$

Let $T=(2^k-1)/(2^n-1)$. Since $\gcd(2^n-1, T)=1$, any integer t, $0 \le t \le 2^k-2$, can be uniquely written as

$$t = t_2 T + t_3(2^n-1), \qquad 0 \le t_2 \le 2^n-2,\ 0 \le t_3 \le 2^n.$$

$$\begin{aligned} R_{ij}(\tau) &= \sum_{t=0}^{N-1} (-1)^{s^{(i)}(t)+s^{(j)}(t+\tau)} \\ &= \sum_{t_2=0}^{2^m} \sum_{t_1=0}^{2^m-2} (-1)^{\sum_{a\in I} \mathrm{tr}_1^m(\beta^{2art_1}([f(\gamma_i,t_2)]^{ar}+[\beta^{2\tau_1}f(\gamma_j,t_2+\tau_2)]^{ar}))}. \end{aligned}$$

Then each sequence $\{s^{(j)}(t)\}$ becomes

$$s^{(j)}(t) = \sum_{a\in I} \mathrm{tr}_1^m \left([\mathrm{tr}_m^n ([\mathrm{tr}_n^k (\gamma^{2t_2T+2t_3(2^n-1)}) + \delta_j \alpha^{t_2T+t_3(2^n-1)}]^u)]^{ar}\right)$$
$$= \sum_{a\in I} \mathrm{tr}_1^m ([\mathrm{tr}_m^n (\alpha^{2ut_2} [\mathrm{tr}_n^k (\gamma^{2t_3(2^n-1)}) + \delta_j]^u)]^{ar})$$

since $\gamma^{2t_2T} \in F_{2^n}$ and $\alpha^T = \alpha^2$. For short notation, we define

$$f(\delta_j, t_3) = \mathrm{tr}_n^k (\gamma^{2t_3(2^n-1)}) + \delta_j.$$

Then we have

$$s^{(j)}(t) = \sum_{a\in I} \mathrm{tr}_1^m ([\mathrm{tr}_m^n (\alpha^{2ut_2}[f(\delta_j, t_3)]^u)]^{ar}).$$

Note that

$$s^{(j)}(t) = \begin{cases} 0, & \text{if } f(\delta_j, t_3) = 0 \\ c(2u(t_2+l)), & \text{if } f(\delta_j, t_3) = \alpha^{2l}. \end{cases}$$

Since $\{c(2ut_2), t_2 = 0, 1, \cdots, 2^n - 2\}$ is the decimation by $2u$ of $\{c(t_2)\}$, it also has the ideal autocorrelation property. Hence, similar arguments as in the proof of Theorem 2 complete the proof. □

Remark 4: The family $\mathcal{F}$ of sequences $\{s^{(j)}(t)\}$ in Theorem 3 can be obtained by applying Theorem 1 to $\{b(t_1)\}$ in (1) and then Theorem 2. As a first step, $\{b(t_1)\}$ can be extended to a sequence $\{c(t_2), t_2 = 0, 1, \cdots, N-1\}$ of period $N = 2^n - 1$ with ideal autocorrelation property, defined by

$$c(t_2) = \sum_{a\in I} \mathrm{tr}_1^m ([\mathrm{tr}_m^n (\alpha^{t_2})]^{ar})$$

where r, $1 \le r \le M-1$, is an integer relatively prime to M. Writing each trace term in $c(t_2)$ as

$$\mathrm{tr}_1^m ([\mathrm{tr}_m^n (\alpha^{t_2})]^{ar}) = \sum_{i\in J(a)} \mathrm{tr}_1^n (\alpha^{it_2}) \tag{3}$$

for some index set $J(a)$, the sequence $\{c(t_2)\}$ can be expressed as

$$c(t_2) = \sum_{a\in I} \sum_{i\in J(a)} \mathrm{tr}_1^n (\alpha^{it_2}).$$

Applying Theorem 2 to $\{c(t_2)\}$, we have a family

$$\mathcal{F} = \{\{v^{(j)}(t)\} \mid j = 1, 2, \cdots, 2^n\}$$

given by

$$v^{(j)}(t) = \sum_{a\in I} \sum_{i\in J(a)} \mathrm{tr}_1^n ([\mathrm{tr}_n^k (\gamma^{2t}) + \delta_j \alpha^t]^{iu})$$

for an integer u, $1 \le u \le N-1$, relatively prime to N. On the other hand, $s^{(j)}(t)$ in Theorem 3 can be expressed as

$$s^{(j)}(t) = \sum_{a\in I} \mathrm{tr}_1^m ([\mathrm{tr}_m^n ([\mathrm{tr}_n^k (\gamma^{2t}) + \delta_j \alpha^t]^u)]^{ar})$$
$$= \sum_{a\in I} \sum_{i\in J(a)} \mathrm{tr}_1^n (([\mathrm{tr}_n^k (\gamma^{2t}) + \delta_j \alpha^t]^u)^i)$$

using the relation in (3). Hence, $s^{(j)}(t)$ is exactly the same as $v^{(j)}(t)$. □

A general form for $s^{(j)}(t)$ in Theorem 3 can be given as follows:

$$s^{(j)}(t) = \sum_{a\in I} \mathrm{tr}_1^{m_1} ([\mathrm{tr}_{m_1}^{m_2} ([\mathrm{tr}_{m_2}^{m_3} ([\cdots \mathrm{tr}_{m_1}^{m_2} ([\mathrm{tr}_{m_d}^k (\gamma^{2t}) + \delta_j \alpha^t]^{r_d}) \cdots]^{r_3})]^{r_2})]^{ar_1}).$$

Here, $m = m_1|m_2|\cdots|m_d = n$, and r_i is relatively prime to $2^{m_i}-1$ for each $i = 1, 2, \cdots, d$.

In signal design for CDMA, it is desirable to have a lot of distinct families of binary sequences with optimal correlation for a given period. Hence it is an interesting problem to find the number of distinct families of binary sequences constructed from a given binary sequence by Theorem 2.

Two families $\mathcal{F}_A$ and $\mathcal{F}_B$ of sequences of the same period are said to be *equivalent* if each sequence in $\mathcal{F}_A$ is a cyclic shift of some sequence in $\mathcal{F}_B$, and *vice versa*. Otherwise, they are said to be *distinct*. Furthermore, they are said to be *fully distinct* if each sequence in $\mathcal{F}_A$ is cyclically distinct from every sequence in $\mathcal{F}_B$.

For an integer $M = 2^m - 1$, define the cyclotomic coset C_i of an integer i, $0 \le i \le M-1$, by

$$C_i = \{j \mid 0 \le j \le M-1, \; j \equiv i\,2^l \pmod M \text{ for some integer } l \ge 0\}.$$

For the sake of convenience, the cyclotomic coset representative of C_i is often defined as the least integer in C_i. It is easily checked that either $C_i = C_j$ or $C_i \cap C_j = \phi$. Hence the set $\{0, 1, \cdots, M-1\}$ is partitioned into pairwise disjoint cyclotomic cosets, that is,

$$\{0, 1, \cdots, M-1\} = \bigcup_{i\in A} C_i$$

where A is the set of all the cyclotomic coset representatives. Note that

$$\mathrm{tr}_1^m (x^j) = \mathrm{tr}_1^m (x^i)$$

for any integer $j \in C_i$.

For an integer r and an index set I, define rI as

$$rI = \{j | 0 \le j \le M-1, j \equiv ri \pmod M \text{ for } i \in I\}$$

and define the set T_I of cyclotomic cosets associated with I as

$$T_I = \{C_a | a \in I\}.$$

Let N_I be the number of r's relatively prime to M such that $T_I \ne T_{rI}$, i.e.,

$$N_I = |\{r | 1 \le r \le M-1, \gcd(r, M) = 1, \text{ and } T_I \ne T_{rI}\}|. \tag{4}$$

Theorem 5: Let N_{fam} be the number of fully distinct families $\mathcal{F}$ of 2^m binary sequences of period $N = 2^n - 1$ given in Theorem 2. Then we have

$$N_{\mathrm{fam}} = N_I \frac{\varphi(N)}{n}$$

where $\varphi(\cdot)$ is the Euler's phi function and N_I is given in (4).

Proof: In order to evaluate N_{fam}, we need to count the number of choices for α and r. The number of choices for α is $\varphi(N)/n$, since α^i and α^j give the same family for any j in the cyclotomic coset $\bmod N$ containing i. If $r_1 I = r_2 I$, then the family associated with $r = r_1$ is exactly the same as the family associated with $r = r_2$. Thus the number of choices for r is N_I, given by (4). Therefore, $N_{\mathrm{fam}} = N_I \cdot \varphi(N)/n$. □

III. Kasami Sequences and No Sequences

Let m and n be positive integers such that $n = 2m$. Let α be a primitive element of F_{2^n} and set $\beta = \alpha^T$ where $T = 2^m + 1$. Then β is a primitive element of F_{2^m}. Let $\{b(t_1), t_1 = 0, 1, \cdots, M-1\}$ be a binary m-sequence of period $M = 2^m - 1$, given by

$$b(t_1) = \mathrm{tr}_1^m (\beta^{t_1}).$$

Note that the m-sequence $\{b(t_1)\}$ is a binary sequence with ideal autocorrelation property. Applying Theorem 2 to $\{b(t_1)\}$, we get an optimal family $\mathcal{F}$ defined by

$$\mathcal{F} = \{\{s^{(j)}(t), t = 0, 1, \cdots, N-1\} | j = 1, 2, \cdots, 2^m\} \tag{5}$$

where $\{s^{(j)}(t)\}$ is the sequence of period N given by

$$s^{(j)}(t) = \mathrm{tr}_1^m \{[\mathrm{tr}_m^n (\alpha^{2t}) + \gamma_j \beta^t]^r\}$$

for $\gamma_j \in F_{2^m}$. Observe that the family $\mathcal{F}$ in (5) is exactly the family of No sequences [15]. In particular, the family $\mathcal{F}$ becomes the small set of Kasami sequences when $r = 1$ [14], [20]. Hence the small set of Kasami sequences and the No sequences can be reinterpreted as a family constructed from the m-sequences. Similarly, generalized No sequences in [13] and [14] are shown to be families constructed from an m-sequence by applying Theorem 1 successively and then Theorem 2.

Consider the number N_{fam} of fully distinct families $\mathcal{F}$ of 2^m binary sequences of period $N = 2^n - 1$ constructed from an m-sequence by Theorem 2. Since $I = \{1\}$, it is easy to check that $N_I = \varphi(M)/m$. Hence we have

$$N_{\mathrm{fam}} = \frac{\varphi(M)}{m} \cdot \frac{\varphi(N)}{n}$$

which is a known result [15].

IV. New Optimal Families of Binary Sequences

A. New Optimal Families from Legendre Sequences

Let p be an odd prime. The Legendre sequence $\{b(t), t = 0, 1, \cdots p-1\}$ of period p is defined as

$$b(t) = \begin{cases} 1, & \text{if } t = 0 \bmod p \\ 0, & \text{if } t \text{ is a quadratic residue } \bmod p \\ 1, & \text{if } t \text{ is a quadratic nonresidue } \bmod p. \end{cases} \tag{6}$$

It is not difficult to show that $\{b(t)\}$ has the ideal autocorrelation property if and only if $p = 3 (\bmod 4)$ [3], [8]. Recently, a trace representation of the Legendre sequences of period $p = 2^m - 1$ (called Mersenne prime) was derived as follows [16]:

Proposition 6 [16]: Let $M = 2^m - 1$ be a prime for some integer $m \geq 3$ and let u be a primitive element of Z_M, the set of integers $\bmod M$. Then there exists a primitive element α of F_{2^m} such that

$$\sum_{i=0}^{[(M-1)/2m]-1} \mathrm{tr}_1^m (\alpha^{u^{2i}}) = 0$$

and the sequence $\{s(t), t = 0, 1, 2, \cdots, M-1\}$ of period M given by

$$s(t) = \sum_{i=0}^{[(M-1)/2m]-1} \mathrm{tr}_1^n (\alpha^{u^{2i}t}) \tag{7}$$

is exactly the Legendre sequence given in (6).

Consider a decimation $\{s(u^l t)\}$ by u^l of the sequence $\{s(t)\}$ given in (7). Clearly, if l is an even integer, then $\{s(u^l t)\}$ is the Legendre sequence given in (6). It is also easy to show that if l is an odd integer, then $\{s(u^l t)\}$ is the sequence given by

$$s(u^l t) = \begin{cases} 1, & \text{if } t = 0 \bmod M \\ 1, & \text{if } t \text{ is a quadratic residue } \bmod M \\ 0, & \text{if } t \text{ is a quadratic nonresidue } \bmod M. \end{cases}$$

Since $\{s(u^l t)\}$ has the ideal autocorrelation property regardless of l, we will also refer to it as a Legendre sequence hereafter. The following theorem is the consequence of Theorem 2 and Proposition 6.

Theorem 7: Let m be an integer such that $M = 2^m - 1$ is a prime, and let $n = 2m$. Let u be a primitive element of Z_M, the set of integers $\bmod M$. Let α be a primitive element of F_{2^n} and set $\beta = \alpha^T$ where $T = 2^m + 1$. For an integer r, $1 \leq r \leq M-1$, let $\{s^{(j)}(t), t = 0, 1, \cdots, N-1\}$ be the sequence of period $N = 2^n - 1$ given by

$$s^{(j)}(t) = \sum_{i=0}^{[(M-1)/2m]-1} \mathrm{tr}_1^m \{[\mathrm{tr}_m^n (\alpha^{2t}) + \gamma_j \beta^t]^r u^{2i}\}, \quad \text{for } \gamma_j \in F_{2^m}.$$

Then the family $\mathcal{F}$ defined by

$$\mathcal{F} = \{\{s^{(j)}(t)\} | j = 1, 2, \cdots, 2^m\}$$

is an optimal set of 2^m binary sequences of period $N = 2^n - 1$ with respect to Welch's bound.

Consider the number N_{fam} of fully distinct families $\mathcal{F}$ of 2^m binary sequences of period $N = 2^n - 1$ constructed from the Legendre sequence of period $M = 2^m - 1$ by Theorem 7. Since we have

$$I = \left\{ u^{2i} | i = 0, 1, \cdots, \frac{M-1}{2m} - 1 \right\}$$

for a primitive element u in Z_M, it is easy to check that $N_I = 2$. Hence we get

$$N_{\mathrm{fam}} = 2\frac{\varphi(N)}{n}.$$

Remark 8: By Theorem 3 and Remark 4, the Legendre sequences of Mersenne prime period $2^m - 1$ can be used to construct optimal families of period $2^k - 1$, where k is any even multiple of m. □

Example 9: Let $m = 7$ and thus $M = 127 (= 2^7 - 1)$. It is easy to check that $u = 3$ is a primitive element of Z_{127}. Let β be a primitive element of F_{2^7}. The sequence $\{b(t_1), t_1 = 0, 1, \cdots, 126\}$ given by

$$b(t_1) = \sum_{i=0}^{8} \mathrm{tr}_1^7 (\beta^{3^{2i}t_1}) = \sum_{i=0}^{8} \mathrm{tr}_1^7 (\beta^{9^i t_1})$$

is the Legendre sequence of period 127.

Let $n = 2m = 14$. Let α be a primitive element of $F_{2^{14}}$ such that $\beta = \alpha^{129}$. For $\gamma_j \in F_{2^7}$, we define

$$s^{(j)}(t) = \sum_{i=0}^{8} \mathrm{tr}_1^7 \{[\mathrm{tr}_7^{14} (\alpha^{2t}) + \gamma_j \beta^t]^{r u^{2i}}\})$$

where r, $1 \leq r \leq 126$, is an integer. Then the family $\mathcal{F}$ defined by

$$\mathcal{F} = \{\{s^{(j)}(t), t = 0, 1, \cdots, 16382\} | j = 1, 2, \cdots, 128\}$$

is an optimal set of 128 binary sequences of period 16383 with respect to Welch's bound. Note that there are 1512 fully distinct families of binary sequences of period 16383, constructed from the Legendre sequences of period 127. □

B. New Families from Hall's Sextic Residue Sequences

Binary sequences of period $M = 2^m - 1$ with ideal autocorrelation property associated with Hall's difference set appears only when m is 5, 7, and 17 [1], [9]. They are known as the Hall's sextic residue sequences. In the case that $m = 5$, the Hall's sextic residue sequences are exactly the m-sequences of period 31.

Let m be one of 5, 7, or 17, and set $M = 2^m - 1$. Let u be a primitive element in Z_M, and let β be a primitive element of F_{2^m}. From a computer search for a trace representation of the Hall's sextic residue sequence $\{b(t_1), t_1 = 0, 1, \cdots, M-1\}$ of period M, it is found that it can be expressed as

$$b(t_1) = \sum_{i=0}^{[(M-1)/6m]-1} \mathrm{tr}_1^m (\alpha^{u^{6i}t_1}).$$

Note that its decimation by any integer r also has the ideal autocorrelation property. Hence, $\{b(t_1)\}$ and all of its decimations are called the Hall's sextic residue sequences. Applying Theorem 2 to $\{b(t_1)\}$, we have an optimal family with respect to Welch's bound in the following.

Theorem 10: Let $n = 2m$, where m is one of 5, 7, or 17, and let u be a primitive element in Z_M with $M = 2^m - 1$. Let α be a primitive element of F_{2^n} and set $\beta = \alpha^T$ where $T = 2^m + 1$. For any integer r, $1 \le r \le M-1$, let $\{s^{(j)}(t), t = 0, 1, \cdots, N-1\}$ be the sequence of period $N = 2^n - 1$ given by

$$s^{(j)}(t) = \sum_{i=0}^{[(M-1)/6m]-1} \mathrm{tr}_1^m \{[\mathrm{tr}_m^n(\alpha^{2t}) + \gamma_j \beta^t]^{ru^{6i}}\}, \quad \text{for } \gamma_j \in F_{2^m}.$$

Then the family $\mathcal{F}$ defined by

$$\mathcal{F} = \{\{s^{(j)}(t)\} | j = 1, 2, \cdots, 2^m\}$$

is an optimal set of 2^m binary sequences of period $N = 2^n - 1$ with respect to Welch's bound.

Consider the number N_{fam} of fully distinct families of binary sequences of period N constructed from the Hall's sextic residue sequences of period M by Theorem 10. Since

$$I = \{u^{6i} | i = 0, 1, \cdots, [(M-1)/6m] - 1\}$$

in Z_M, it is easy to check that $N_I = 6$. Hence we have

$$N_{\mathrm{fam}} = 6\,\frac{\varphi(N)}{n}.$$

Remark 11: Using Theorem 3 or Remark 4, the Hall's sextic residue sequences of period $M = 2^m - 1$ can be applied to construct optimal families of period $2^k - 1$, where k is any even multiple of m. □

V. New Optimal Families from Miscellaneous Sequences of Unknown Type

To classify and construct balanced binary sequences of period $2^n - 1$ is a very interesting problem in both theory and practice [7], [8]. Especially, the balanced binary sequences of period $2^n - 1$ with ideal autocorrelation property find many applications in spread-spectrum communication systems. A complete search for those sequences was conducted for period 127 by Baumert and Fredrickson [2], 255 by Cheng [4], and 511 by Drier [5].

It is well known that there are six *inequivalent* binary sequences of period 127 with ideal autocorrelation property: an m-sequence, a Legendre sequence, a Hall's sextic residue sequence, and three others called the miscellaneous sequences of unknown type I, II, and III. Let β be a primitive element of F_{2^7}. Then the three miscellaneous sequences are known to have the following trace representation by a computer search:

i) Unknown Type I

$$b_{\mathrm{I}}(t_1) = \mathrm{tr}_1^7(\beta^{t_1}) + \mathrm{tr}_1^7(\beta^{5t_1}) + \mathrm{tr}_1^7(\beta^{7t_1}) + \mathrm{tr}_1^7(\beta^{11t_1}) + \mathrm{tr}_1^7(\beta^{31t_1}).$$

ii) Unknown Type II

$$b_{\mathrm{II}}(t_1) = \mathrm{tr}_1^7(\beta^{t_1}) + \mathrm{tr}_1^7(\beta^{3t_1}) + \mathrm{tr}_1^7(\beta^{7t_1}) + \mathrm{tr}_1^7(\beta^{19t_1}) + \mathrm{tr}_1^7(\beta^{29t_1}).$$

iii) Unknown Type III

$$b_{\mathrm{III}}(t_1) = \mathrm{tr}_1^7(\beta^{t_1}) + \mathrm{tr}_1^7(\beta^{9t_1}) + \mathrm{tr}_1^7(\beta^{13t_1})$$

where t_1 runs from 0 to 126.

By Theorem 2, new optimal families can be constructed from the above sequences of unknown type. For example, consider a family from the sequence $\{b_{\mathrm{III}}(t_1)\}$. Let $n = 2m = 14$. Let α be a primitive element of $F_{2^{14}}$ and set $\beta = \alpha^T$ where $T = 2^7 + 1$. For any integer r, $1 \le r \le 126$, let $\{s^{(j)}(t), t = 0, 1, \cdots, N-1\}$ be the sequence of period $N = 2^{14} - 1$ given by

$$s^{(j)}(t) = \sum_{a \in I} \mathrm{tr}_1^7 \{[\mathrm{tr}_7^{14}(\alpha^{2t}) + \gamma_j \beta^t]^{ar}\}$$

where $\gamma_j \in F_{2^7}$ and $I = \{1, 9, 13\}$. Then the family $\mathcal{F}$ defined by

$$\mathcal{F} = \{\{s^{(j)}(t)\} | j = 1, 2, \cdots, 128\}$$

is an optimal set of 128 binary sequences of period $N = 2^{14} - 1$ with respect to Welch's bound. It is easily checked that $N_I = \varphi(127)/7 = 18$. Hence we have $N_{\mathrm{fam}} = 18\,\varphi(2^{14} - 1)/14 = 13\,608$ optimal families from a binary sequence of each miscellaneous type.

At period 255, it is found that there are four *inequivalent* binary sequences with ideal autocorrelation property: an m-sequence, a GMW sequence, and two others of unknown type. New optimal families can be constructed from a binary sequence of each unknown type. Note that $N_I = \varphi(255)/8 = 16$ in this case.

At period 511, there are five *inequivalent* binary sequences with ideal autocorrelation property: an m-sequence, a GMW sequence, and three others of unknown type. New optimal families can be constructed from a binary sequence of each unknown type. Note that $N_I = \varphi(511)/9 = 48$ in this case.

In the case of period 1023, a computer search found that there is at least one binary sequence $\{b(t_1), t_1 = 0, 1, \cdots, 1022\}$ with ideal autocorrelation property, which is *inequivalent* to any of known binary sequences such as the m-sequences, the GMW sequences, and the extensions of the Legendre sequences. It is given by

$$b(t_1) = \mathrm{tr}_1^{10}(\alpha^{t_1}) + \mathrm{tr}_1^{10}(\alpha^{9t_1}) + \mathrm{tr}_1^{10}(\alpha^{57t_1}) + \mathrm{tr}_1^{10}(\alpha^{73t_1}) + \mathrm{tr}_1^{10}(\alpha^{121t_1})$$

where α is a primitive element of $F_{2^{10}}$. Hence, a new optimal family of 1024 binary sequences of period $2^{20} - 1$ can be constructed from the sequence $\{b(t_1)\}$ described above.

As in the cases of Legendre sequences and Hall's sextic residue sequences, miscellaneous sequences of unknown type of period $M = 2^m - 1$ can be used to construct optimal families of period $2^k - 1$, where k is any even multiple of m, by applying Theorem 3 or Remark 4.

References

[1] L. D. Baumert, *Cyclic Difference Sets* (Lecture Notes in Mathematics). Berlin, Germany: Springer-Verlag, 1971.

[2] L. D. Baumert and H. Fredrickson, "The cyclotomic numbers of order 18 with applications to difference sets," *Math. Comput.*, vol. 21, pp. 204–219, 1967.

[3] D. M. Burton, *Elementary Number Theory.* Newton, MA: Allyn, 1980.

[4] U. Cheng, "Exhaustive construction of $(255, 127, 63)$ cyclic difference sets," *J. Combin. Theory*, vol. A-35, pp. 115–125, 1983.

[5] R. Drier, "$(511, 255, 127)$ cyclic difference sets," *IDA Talk*, July 1992.

[6] R. Gold, "Optimal binary sequences for spread spectrum multiplexing," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 619–621, Oct. 1967.

[7] S. W. Golomb, "On the classification of balanced binary sequences of period $2^n - 1$," *IEEE Trans. Inform. Theory*, vol. IT-26, pp. 730–732, Nov. 1980.

[8] ____, *Shift-Register Sequences.* San Francisco, CA: Holden-Day, 1967; Laguna Hills, CA: Aegean Park, 1982.

[9] D. Jungnickel, "Difference sets," in *Contemporary Design Theory*, J. H. Dinitz and D. R. Stinson, Eds. New York: Wiley, 1992, pp. 241–324.

[10] R. Lidl and H. Niederreiter, *Finite Fields*, vol. 20 of *Encyclopedia of Mathematics and Its Applications.* Reading, MA: Addison-Wesley, 1983.

[11] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes.* Amsterdam, The Netherlands: North-Holland, 1977.
[12] J. D. Olsen, R. A. Scholtz, and L. R. Welch, "Bent-function sequences," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 858–864, Nov. 1982.
[13] J.-S. No, "A new family of binary pseudorandom sequences having optimal periodic correlation properties and large linear span," Ph.D. dissertation, Univ. of Southern California, Los Angeles, CA, May 1988.
[14] ——, "Generalization of GMW sequences and No sequences," *IEEE Trans. Inform. Theory*, vol. 42, pp. 260–262, Jan. 1996.
[15] J.-S. No and P. V. Kumar, "A new family of binary pseudorandom sequences having optimal periodic correlation properties and large linear span," *IEEE Trans. Inform. Theory*, vol. 35, pp. 371–379, Mar. 1989.
[16] J.-S. No, H.-K. Lee, H. Chung, H.-Y. Song, and K. Yang, "Trace representation of Legendre sequences of Mersenne prime period," *IEEE Trans. Inform. Theory*, vol. 42, pp. 2254–2255, Nov. 1996.
[17] J.-S. No, K. Yang, H. Chung, and H.-Y. Song, "Extension of binary sequences with ideal autocorrelation property," in *Proc 1996 IEEE Int. Symp. on Information Theory and Its Applications (ISITA '96)* (Victoria, BC, Canada, Sept. 17–20, 1996), pp. 837–840.
[18] D. V. Sarwate and M. B. Pursley, "Crosscorrelation properties of pseudorandom and related sequences," *Proc. IEEE*, vol. 68, pp. 593–619, May 1980.
[19] V. M. Sidelnikov, "On mutual correlation of sequences," *Sov. Math.–Dokl.*, vol. 12, pp. 197–201, 1971.
[20] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications*, vol. 1. Rockville, MD: Computer Sci. Press, 1985.
[21] H. Y. Song and S. W. Golomb, "On the existence of cyclic Hadamard difference sets," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1266–1268, July 1994.
[22] L. R. Welch, "Lower bounds on the maximum cross-correlation of signals," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 397–399, May 1974.

Is Code Equivalence Easy to Decide?

Erez Petrank and Ron M. Roth, *Member, IEEE*

***Abstract*— We study the computational difficulty of deciding whether two matrices generate equivalent linear codes, i.e., codes that consist of the same codewords up to a fixed permutation on the codeword coordinates. We call this problem Code Equivalence. Using techniques from the area of interactive proofs, we show on the one hand, that under the assumption that the polynomial-time hierarchy does not collapse, Code Equivanence is not NP-complete. On the other hand, we present a polynomial-time reduction from the Graph Isomorphism problem to Code Equivalence. Thus if one could find an efficient (i.e., polynomial-time) algorithm for Code Equivalence, then one could settle the long-standing problem of determining whether there is an efficient algorithm for solving Graph Isomorphism.**

***Index Terms*— Code Equivalence, Graph Isomorphism, interactive proofs, polynomial hierarchy.**

I. Introduction

Let F be a finite field and let G_1 and G_2 be generator matrices of two linear codes C_1 and C_2 over F. We say that G_1 and G_2 are *code-equivalent*, denoted $G_1 \sim G_2$, if the sets C_1 and C_2 are the same, up to a fixed permutation on the coordinates of the codewords of C_1. In other words, $G_1 \sim G_2$ if and only if both matrices have the same order $k \times n$, and there exist an $n \times n$ permutation matrix P and a nonsingular $k \times k$ matrix S over F such that $G_1 = SG_2P$. The problem of deciding whether two generator matrices are code-equivalent will be referred to as the *Code Equivalence* problem.

The purpose of this correspondence is to study the computational difficulty of the Code Equivalence problem. As one application of a related problem, we mention the public-key cryptosystems due to McEliece [9] and Niederreiter [11]. Recall that an *alternant code* over GF(q) is defined by a parity-check matrix of the form $[y_j\alpha_j^i]_{i=0,j=0}^{r-1,n-1}$, where the α_j's are distinct elements in GF(q^m) and the y_j's are nonzero elements in GF(q^m) [8, ch. 12]. Goppa codes are special cases of alternant codes where certain restrictions are imposed on the values y_j's, and generalized Reed–Solomon codes are special cases of alternant codes where $m = 1$. The mentioned cryptosystems are based on the assumption that it is difficult to identify the values α_j and y_j out of an arbitrary generator matrix (or parity-check matrix) of an alternant code. Namely, it is difficult to obtain a code-equivalent matrix of the form $[y_j\alpha_j^i]_{i=0,j=0}^{r-1,n-1}$. On the other hand, as shown in [12], it is easy to extract the values α_j and y_j from any systematic generator matrix of a generalized Reed–Solomon code; hence, cryptosystems based on such a code are breakable. This was pointed out explicitly by Sidelnikov and Shestakov in [13]. For related work, see also the references cited in [10, p. 317].

The significance of the Code Equivalence problem can also be exhibited through the results of Kasami, Lin, and Peterson [6], and Kolesnik and Mironchikov [7], who showed that Reed–Muller codes are equivalent to subcodes of extended Bose–Chaudhuri–Hocquenghem (BCH) codes. Thus it should be interesting to design an efficient algorithm that decides whether two codes are indeed equivalent, and thus infer from the properties that arise from one code representation to the other.

On the positive side, we first show that the Code Equivalence problem is unlikely to be NP-complete. The proof of this assertion relies on techniques developed in the field of *interactive proofs*, which we summarize in Section II. In Section III, we invoke results of Goldwasser, Micali, and Rackoff [4], Goldreich, Micali, and Wigderson [3], Goldwasser and Sipser [5], and Boppana, Håstad, and Zachos [2], to show that if Code Equivalence is NP-complete, then the polynomial hierarchy collapses.

Yet, we do state also a negative result, namely, that Code Equivalence is also unlikely to be too easy. We do this by relating Code Equivalence to the *Graph Isomorphism* problem. Let $\mathcal{G}_1 = (V, E_1)$ and $\mathcal{G}_2 = (V, E_2)$ be two undirected graphs with the same set of vertices V, and with sets of edges E_1 and E_2, respectively. We say that $\mathcal{G}_1$ is isomorphic to $\mathcal{G}_2$ if there exists a permutation (isomorphism) $\pi\colon V \to V$ such that $\{u, v\} \in E_1$ if and only if $\{\pi(u), \pi(v)\} \in E_2$ (we assume here that the graphs have no parallel edges; if they do, then E_1 and E_2 are multisets, in which case isomorphism requires equality of the multiplicities of $\{u, v\}$ and $\{\pi(u), \pi(v)\}$ in E_1 and E_2, respectively). The problem of deciding efficiently (i.e., in polynomial time) whether two graphs are isomorphic is a notoriously open question in Computer Science. The problem has been studied extensively in recent decades, but the state of the art is that there is no known efficient algorithm for determining whether two given graphs are isomorphic.

In Section IV, we show a polynomial-time reduction from Graph Isomorphism to Code Equivalence. This implies that presenting an

Manuscript received November 30, 1995; revised February 12, 1997.

E. Petrank is with the DIMACS Center, P.O. Box 1179, Piscataway, NJ 08855 USA.

R. M. Roth is with Hewlett-Packard Laboratories, Palo Alto, CA 94304 USA, on leave from the Computer Science Department, Technion–Israel Institute of Technology, Haifa 32000, Israel.

Publisher Item Identifier S 0018-9448(97)05216-4.