Trace-Shortened Reed–Solomon Codes

R. J. McEliece
California Institute of Technology, Department of Engineering

G. Solomon[1]

Reed–Solomon (RS) codes have been part of standard NASA telecommunications systems for many years. RS codes are character-oriented error-correcting codes, and their principal use in space applications has been as outer codes in concatenated coding systems [6, Section 5.4.4]. However, for a given character size, say m *bits, RS codes are limited to a length of, at most,* 2^m. *It is known in theory that longer character-oriented codes would be superior to RS codes in concatenation applications, but until recently no practical class of "long" character-oriented codes had been discovered. In 1992, however, Solomon [4,5], discovered an extensive class of such codes, which are now called* trace-shortened Reed–Solomon (TSRS) codes. *In this article, we will continue the study of TSRS codes. Our main result is a formula for the dimension of any TSRS code, as a function of its error-correcting power. Using this formula, we will give several examples of TSRS codes, some of which look very promising as candidate outer codes in high-performance coded telecommunications systems.*

I. Construction Summary

In this section, we will summarize the construction of trace-shortened Reed–Solomon (TSRS) codes, and state our main result (Theorem 1), which is a formula for the binary dimension of an arbitrary TSRS code. In Section II, we will give some numerical examples. In Section III, we will state and prove a theorem (Theorem 2) that is more general than Theorem 1. Finally, in Section IV, we will summarize our results and list several open problems.

We begin with a summary of TSRS codes, as introduced in [4,5]. For any length n of the form $n = 2^m - 1$, any desired minimum distance $d \leq n - 1$, and any positive integer $\mu \leq m$, there is a TSRS code of length n and minimum distance d over the symbol alphabet $V(2^{m-\mu})$, the space of binary $(m - \mu)$-tuples. TSRS codes are constructed using properties of the Galois field $GF(2^m)$. The field $GF(2^{m-\mu})$ does not come into play in the construction, and so TSRS codes are not linear over the symbol field $GF(2^{m-\mu})$. However, they are linear over $GF(2)$, and the symbol-wise cyclic shift of any codeword is also a codeword. Our main result (Theorem 2) is a formula which allows the easy calculation of the binary dimension

[1] Independent consultant to the Communications Systems Research Section.

of any TSRS code. We give several numerical examples, which show that TSRS codes are in some cases "almost maximum distance separable (MDS)," even though they are much longer than any true MDS code can be.

We begin with an ordinary Reed–Solomon (RS) code, as originally defined [3]. That is, with m and k_0 fixed, we consider the set $\mathcal{P}_m^{k_0}$ of polynomials of degree $k_0 - 1$ or less over the field $GF(2^m)$, and for each polynomial $P \in \mathcal{P}_m^{k_0}$, we define a length $n = 2^m - 1$ codeword $C(P)$ as follows:

$$C(P) = (P(1), P(\alpha), \ \ldots, \ P(\alpha^{n-1})) \tag{1}$$

where α is a primitive nth root of unity in $GF(2^m)$. The set of all such codewords, i.e., $\{C(P) : P \in \mathcal{P}_m^{k_0}\}$, is an (n, k_0, d) RS code over $GF(2^m)$, where $n = 2^m - 1$ and $d = n - k_0 + 1$.

The words in the above-constructed RS codes are vectors of length n over the field $GF(2^m)$. The next step in our construction is to "expand" each codeword into a length nm binary vector by representing each codeword symbol as a vector of length m over $GF(2)$. The particular coordinate basis we shall use for this representation is the basis which is dual to the "natural" basis $\{1, \alpha, \ \ldots, \ \alpha^{m-1}\}$. With respect to this dual basis, which we denote by $\{\beta_0, \beta_1, \ \ldots, \ \beta_{m-1}\}$, an element $x \in GF(2^m)$ has representation $x = x_0\beta_0 + \cdots + x_{m-1}\beta_{m-1}$, where the components x_h, which are elements of $GF(2)$, are given by the formula

$$x_h = \mathrm{Tr}_1^m(x\alpha^h) \qquad \text{for } h = 0, 1, \ \ldots, \ m-1 \tag{2}$$

In Eq. (2), and throughout the article, we use the symbol "Tr" to denote the trace operator. The super- and subscripts denote the subfields involved. Thus $\mathrm{Tr}_1^m(\xi)$, the trace of ξ from $GF(2^m)$ to $GF(2^1)$, represents the $GF(2)$-linear mapping from $GF(2^m)$ to $GF(2^1)$, given by

$$\mathrm{Tr}_1^m(\xi) = \xi + \xi^2 + \cdots + \xi^{2^{m-1}} \tag{3}$$

Similarly, if d is a divisor of m, $\mathrm{Tr}_d^m(\xi)$ denotes the trace of ξ from $GF(2^d)$ to $GF(2^m)$, defined as the $GF(2^d)$-linear mapping from $GF(2^m)$ to $GF(2^d)$, given by

$$\mathrm{Tr}_d^m(\xi) = \xi + \xi^{2^d} + \xi^{2^{2d}} + \cdots + \xi^{2^{(f-1)d}} \tag{4}$$

where $f = m/d$. Throughout the article, we will freely use the basic properties of the trace operator and the dual basis representation, which are summarized in [2, Chapter 8].

In any event, we call this "expanded" code the bit-mapped version of the RS code.

Now we can define the trace-shortened Reed–Solomon codes. Given the bit-mapped RS code defined above, for any integer μ in the range $0 \le \mu \le m$, we define the TSRS code of index μ to be the set of bit-mapped RS codewords for which the first μ binary components of each codeword symbol are zero. That is, if $C = (C_0, \ \ldots, \ C_{n-1})$ is a codeword in the original RS code, C is in the corresponding TSRS code of index μ if and only if

$$\mathrm{Tr}_1^m(\alpha^h C_i) = 0 \begin{cases} \text{for } h = 0, 1, \ \ldots, \ \mu - 1 \\ \text{for } i = 0, 1, \ \ldots, \ n - 1 \end{cases} \tag{5}$$

If we delete the μ guaranteed-zero binary components from each bit-mapped RS symbol, the TSRS code defined by Eq. (5) becomes a code of length n whose codeword components lie in the vector space $V(2^{m-\mu})$ of binary $(m-\mu)$-tuples. We note that while $V(2^{m-\mu})$ is not a field, it is nevertheless a group, viz, the elementary abelian group of order $2^{m-\mu}$ [1, Section 3.3]. Since this code is, by definition, a subcode of the parent RS code, its minimum distance is at least d. The bitwise sum of any two codewords satisfying Eq. (5) also satisfies Eq. (5), and so the code is a group code over the elementary abelian group $V(2^{m-\mu})$. Also, any (symbolwise) cyclic shift of any codeword in the TSRS code is also a codeword, so the code is a cyclic group code. The determination of the binary dimension of the code is somewhat challenging, and is the main result of this article. In the next paragraph, we will discuss our formula for this dimension.

As we have seen, the TSRS code of index μ over $GF(2^m)$ is a subgroup of the group $V(2^{m-\mu})^n$. The order of the code, i.e., the number of codewords, need not be a power of $2^{m-\mu}$, but since the code is linear over $GF(2)$, it must be a power of 2, say 2^{K_μ}. The following theorem gives a simple way to calculate K_μ. Before stating it, however, we need to introduce so-called cyclotomic cosets of the field $GF(2^m)$, which are the cycles of the permutation $i \to 2i \bmod (2^m - 1)$ on the set $\{0, 1, \ldots, 2^m - 2\}$ [2, Chapter 7]. We shall denote the size of the jth cyclotomic coset by d_j. For example, if $m = 4$, there are five such cyclotomic cosets:

j	C_j	d_j
0	(0)	1
1	(1, 2, 4, 8)	4
3	(3, 6, 12, 9)	4
5	(5, 10)	2
7	(7, 14, 13, 11)	4

Theorem 1. Denote by e_j the number of integers in the set $\{0, 1, \ldots, k_0 - 1\}$ which lie in the jth cyclotomic coset of $GF(2^m)$. Then the binary dimension of the index μ trace-shortened (n, k_0) RS code is given by the formula

$$K_\mu = \sum_j \max(me_j - \mu d_j, 0) \tag{6}$$

In Section II, we will give two extended numerical examples of TSRS codes. In Section III, we will give our proof of Theorem 1. Finally in Section IV, we will make some concluding remarks, and list several open problems about TSRS codes.

II. Examples

For our first example, we begin with the (15, 9, 7) RS code over the field $GF(16)$. Here $m = 4$, $n = 15$, $k_0 = 9$. In the table below, we list the cyclotomic cosets of $GF(16)$, together with the numbers d_j and e_j, which are needed to apply Theorem 1. (For clarity, we list the numbers in the set $\{0, 1, \ldots, 8\}$ in boldface.)

j	C_j	d_j	e_j
0	**(0)**	1	1
1	**(1, 2, 4, 8)**	4	4
3	**(3, 6**, 12, 9)	4	2
5	**(5**, 10)	2	1
7	**(7**, 14, 13, 11)	4	1

From this table and using Eq. (6), the binary dimension K_μ is given by the formula

$$K_\mu = (4 - \mu)_+ + (16 - 4\mu)_+ + (8 - 4\mu)_+ + (4 - 2\mu)_+ + (4 - 4\mu)_+$$

where x_+ is short for $\max(x, 0)$. Thus we have

$$K_0 = 4 + 16 + 8 + 4 + 4 = 36$$

$$K_1 = 3 + 12 + 4 + 2 + 0 = 21$$

$$K_2 = 2 + 8 + 0 + 0 + 0 = 10$$

$$K_3 = 1 + 4 + 0 + 0 + 0 = 5$$

In other words, by "trace-shortening" the parent (15, 9, 7) RS code, we obtain the codes over $V(8)$, $V(4)$, and $V(2)$, as shown in the table below. (In the following table, we extend the usual (n, k, d) notation for linear codes to the nonlinear TSRS codes by letting k denote the "pseudodimension" of the code, defined as $k_\mu = K_\mu/(m - \mu)$.)

Code parameters	Symbol group
(15, 9, 7)	$V(16)$
(15, 7, 7)	$V(8)$
(15, 5, 7)	$V(4)$
(15, 5, 7)	$V(2)$

Since the Singleton bound implies that any code with $n = 15$ and $d = 7$ must have $k \leq 9$, it follows that the (15, 7, 7) code over $V(8)$ is close to optimal. The (15, 5, 7) code over $V(4)$ is not as good as the (15, 6, 7) generalized Bose–Chaudhari–Hocquenghem (BCH) code, but we could have obtained a code with the same parameters by starting with an alternate RS code. Finally, the (15, 5, 7) code over $V(2)$ has the same parameters as the (15, 5, 7) BCH code.

As our second example, we start with the (31, 27, 5) RS code over $GF(32)$. The distribution of the numbers in the set $\{0, 1, \ldots, 26\}$ is shown in the following table:

j	C_j	d_j	e_j
0	**(0)**	1	1
1	**(1, 2, 4, 8, 16)**	5	5
3	**(3, 6, 12, 24, 17)**	5	5
5	**(5, 10, 20, 9, 18)**	5	5
7	**(7, 14**, 28, **25, 19)**	5	4
11	**(11, 22, 13, 26, 21)**	5	5
15	**(15**, 30, 29, 27, **23)**	5	2

Thus we have from Theorem 1,

$$K_0 = 5 + 25 + 25 + 25 + 20 + 25 + 10 = 135$$

$$K_1 = 4 + 20 + 20 + 20 + 15 + 20 + \ 5 = 104$$

$$K_2 = 3 + 15 + 15 + 15 + 10 + 15 + \ 0 = 73$$

$$K_3 = 2 + 10 + 10 + 10 + \ 5 + 10 + \ 0 = 47$$

$$K_4 = 1 + \ 5 + \ 5 + \ 5 + \ 0 + \ 5 + \ 0 = 21$$

This leads to the following list of codes:

Code parameters	Symbol group
$(31, 27, 5)$	$V(32)$
$(31, 26, 5)$	$V(16)$
$(31, 24.33, 5)$	$V(8)$
$(31, 23.5, 5)$	$V(4)$
$(31, 21, 5)$	$V(2)$

This list contains some interesting codes. In particular, the $(31, 26, 5)$ code over $V(16)$ is optimal, and in fact is "nearly" MDS, despite the fact that it is twice as long as any MDS code over a 16-symbol alphabet.

Let us conclude this section with a few bigger examples, omitting the detailed computations. Beginning with a $(511, 478, 34)$ RS code over $GF(2^9)$, for which the parity-check polynomial is $h(x) = \prod_{j=1}^{478}(1 - \alpha^i x)$, and trace-shortening it with $\mu = 1$, one obtains a $(511, 474, 34)$ TSRS code over the symbol group $V(2^8)$. Similarly, beginning with a $(511, 470, 42)$ RS code over $GF(2^9)$, with $h(x) = \prod_{j=1}^{470}(1 - \alpha^i x)$, again taking $\mu = 1$, one obtains a $(511, 465, 42)$ code over $V(2^8)$. Preliminary calculations performed by Dr. Fabrizio Pollara of the Communications Systems Research Section indicate that when concatenated with the NASA standard $(7, 1/2)$ convolutional code, both of these TSRS codes give an overall performance which is superior to that of the standard $(255, 223)$ RS code.

III. Proof of Theorem 1

In this section we will state and prove a theorem which is slightly more general than Theorem 1.

We begin with the field $F = GF(2^m)$, a positive integer n which is a divisor of $2^m - 1$, and a primitive nth root of unity in F, say α. Let J be a subset of $\{0, 1, \ldots, n-1\}$, with $|J| = k_0$. We then define the code $\mathcal{C}_J$ to be the (n, k_0) cyclic code over F, with check polynomial $h(x) = \prod_{j \in J}(1 - \alpha^j x)$. Equivalently, $\mathcal{C}_J$ consists of all vectors $\mathbf{C} = (C_0, C_1, \ldots, C_{n-1})$ of the form

$$C_i = \sum_{j \in J} c_j \alpha^{ij} \qquad i = 0, 1, \ldots, n-1 \tag{7}$$

where (c_j), for $j \in J$, is an arbitrary set of elements of F, indexed by J. For future reference, we denote the minimum distance of the code $\mathcal{C}_J$ by d_J. Note that if the elements of J form an arithmetic progression modulo n, whose increment is relatively prime to n, then $d_J = n - k_0 + 1$ by the BCH argument, and the code $\mathcal{C}_J$ is a (generalized) Reed–Solomon code. For example, if $n = 2^m - 1$ and $J = \{0, 1, \ldots, k_0 - 1\}$, the code $\mathcal{C}_J$ is the same as the RS code defined in Eq. (1).

We now define the μth-order trace-shortening of $\mathcal{C}_j$, denoted by $\mathcal{C}_J^\mu$, to be the set of codewords $\mathbf{C} \in \mathcal{C}_J$ such that

$$\mathrm{Tr}_1^m(\alpha^h C_i) = 0 \quad \begin{cases} h = 0, 1, \ldots, \mu - 1 \\ i = 0, 1, \ldots, n - 1 \end{cases} \tag{8}$$

In words, $\mathcal{C}_J^\mu$ consists of all codewords of $\mathcal{C}_J$ for which the first μ binary components (with respect to the basis for $GF(2^m)$ over $GF(2)$, which is dual to the basis $\{1, \alpha, \ldots, \alpha^{m-1}\}$) of each codeword symbol are zero. If we ignore these μn guaranteed-zero components, the code $\mathcal{C}_J^\mu$ becomes a code of length n over the set $V(2^{m-\mu})$ of binary $(m - \mu)$-tuples.

We denote the minimum symbol distance of the code $\mathcal{C}_J^\mu$ by d_J^μ. Since every codeword in $\mathcal{C}_J^\mu$ is also a codeword in the parent code $\mathcal{C}_j$, the minimum symbol distance of $\mathcal{C}_J^\mu$ cannot be less than that of $\mathcal{C}_J$, i.e., $d_J^\mu \geq d_J$. This simple bound is all we will have to say about the distance properties of the code $\mathcal{C}_J^\mu$. Our main results concern the *size* of the code.

Because the parent code $\mathcal{C}_J$ is a linear, cyclic code over $GF(2^m)$, it follows that the code $\mathcal{C}_J^\mu$ is closed under the operations of addition of any two codewords, and takes the (symbolwise) cyclic shift of any individual codeword. However, since the field $GF(2^{m-\mu})$ does not come into play, $\mathcal{C}_J^\mu$ cannot be a linear code. Nevertheless, we will define a pseudodimension for $\mathcal{C}_J^\mu$.

Since, as observed, the sum of any two codewords from $\mathcal{C}_J^\mu$ is another codeword, $\mathcal{C}_J^\mu$ is a linear code over $GF(2)$. Let us denote its $GF(2)$-dimension by K_μ. (Thus $K_0 =$

$k_0 \times m$, $K_m = 0$.) Then the pseudodimension of $\mathcal{C}_J^\mu$ is defined as

$$k_\mu = \frac{1}{m-\mu} K_\mu \tag{9}$$

Thus if $|\mathcal{C}_J^\mu|$ denotes the number of codewords in $\mathcal{C}_J^\mu$, we have that

$$|\mathcal{C}_J^\mu| = 2^{(m-\mu)k_\mu}$$

The main result of this article (Theorem 2, below), is the determination of K_μ, and hence also k_μ, for $\mu = 0, 1, \ldots, m$. To state the result, we need to define the modulo n cyclotomic cosets.

Let n be an odd positive integer. If i and j are integers in the range $0 \le i \le n-1$, and if $2^s i \equiv j \pmod{n}$ for some integer s, we say that i and j are conjugate modulo n. It is easy to see that conjugation modulo n is an equivalence relation on the set $\{0, 1, \ldots, n-1\}$, and so the set $\{0, 1, \ldots, n-1\}$ is partitioned into a number of disjoint equivalence classes, which are called the modulo n cyclotomic cosets. Alternatively, the cyclotomic coset containing j, which we will denote by Γ_j, can be described explicitly as the set $\{j, 2j, \ldots, 2^{d-1}j\}$, where d is the least positive integer such that $2^d j \equiv j \pmod{n}$. The integer d is called the degree of j, written $d = \deg(j)$. In what follows, we will denote the cardinality of Γ_j by d_j. It is easy to see that every element of Γ_j has degree d_j, and that d_j is a divisor of n. Finally, we denote by I_n the set consisting of the smallest integers in each cyclotomic coset.

Example 1. Let $n = 15$. A short calculation shows that there are five cyclotomic cosets modulo 15; indeed, we have $I_{15} = \{0, 1, 3, 5, 7\}$, and

$$\begin{array}{ll} \Gamma_0 = (0) & d_0 = 1 \\ \Gamma_1 = (1, 2, 4, 8) & d_1 = 4 \\ \Gamma_3 = (3, 6, 12, 9) & d_3 = 4 \\ \Gamma_5 = (5, 10) & d_5 = 2 \\ \Gamma_7 = (7, 14, 13, 11) & d_7 = 4 \end{array}$$

Thus $\deg(0) = 1$, $\deg(1) = \deg(2) = \deg(4) = \deg(8) = 4$, etc. ❐

Now we can give our formula for the binary dimensions of the codes $\mathcal{C}_J^\mu$. For each $j \in I_n$, we define $J_j = J \cap \Gamma_j$, and $e_j = |J_j|$. (Compare to Theorem 1.)

Theorem 2. The binary dimension K_μ of the code $\mathcal{C}_J^\mu$ is given by the formula

$$K_\mu = \sum_{j \in I} (m e_j - \mu d_j)_+ \tag{10}$$

where $(x)_+ = \max(x, 0)$.

To prove Theorem 2, we need several Lemmas. If n is a divisor of $2^m - 1$, let $P(x)$ be a polynomial of degree $n-1$ in the indeterminate x, with coefficients in $F = GF(2^m)$:

$$P(x) = \sum_{j=0}^{n-1} P_j x^j, \qquad P_j \in F \tag{11}$$

Now define the polynomial $\mathcal{P}(x)$ as follows:

$$\mathcal{P}(x) = \mathrm{Tr}_1^m(P(x)) \bmod (x^n - 1)$$

$$= \sum_{j=0}^{n-1} \mathcal{P}_j x^j \tag{12}$$

where in Eq. (12) it is understood that $\mathrm{Tr}_1^m(\xi) = \xi + \xi^2 + \cdots + \xi^{2^{m-1}}$, as in Eq. (3).

Example 2. Let $m = 4$, $n = 5$. If $P(x) = P_0 + P_1 x + P_2 x^2 + P_3 x^3 + P_4 x^4$, then $\mathcal{P}(x) = P(x) + P(x)^2 + P(x)^4 + P(x)^8 \bmod (x^5 - 1) = 1 \times (P_0 + P_0^2 + P_0^4 + P_0^8) + x \times (P_1 + P_3^2 + P_4^4 + P_2^8) + x^2 \times (P_2 + P_1^2 + P_3^4 + P_4^8) + x^3 \times (P_3 + P_4^2 + P_2^4 + P_1^8) + x^4 \times (P_4 + P_2^2 + P_1^4 + P_3^8)$. Thus,

$$\mathcal{P}_0 = P_0 + P_0^2 + P_0^4 + P_0^8$$

$$\mathcal{P}_1 = P_1 + P_3^2 + P_4^4 + P_2^8$$

$$\mathcal{P}_2 = P_2 + P_1^2 + P_3^4 + P_4^8$$

$$\mathcal{P}_3 = P_3 + P_4^2 + P_2^4 + P_1^8$$

$$\mathcal{P}_4 = P_4 + P_2^2 + P_1^4 + P_3^8$$ ❐

Lemma 1. Let $P(x)$ be a polynomial of degree $n-1$, as defined in Eq. (11). Then $\mathrm{Tr}_1^m(P(x)) = 0$ for all $x \in \{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$ if and only if $\mathcal{P}_j = 0$ for $j = 0, 1, \ldots, n-1$.

Proof: Since $x^n = 1$ for all $x \in \{1, \alpha, \ldots, \alpha^{n-1}\}$, it follows from Eq. (12) that $\mathrm{Tr}_1^m(P(x)) = 0$ for all $x \in \{1, \alpha, \ldots, \alpha^{n-1}\}$ iff $\mathcal{P}(x) = 0$ for all $x \in \{1, \alpha, \ldots, \alpha^{n-1}\}$. But $\deg \mathcal{P}(x) \leq n-1$, so that $\mathcal{P}(x)$ must be identically 0. ❐

The next lemma gives an explicit formula for the coefficients $\mathcal{P}_j$ of $\mathcal{P}(x)$.

Lemma 2. For $j \in \{0, 1, \ldots, n-1\}$, if $d = \deg(j)$, then

$$\mathcal{P}_j = \mathrm{Tr}_d^m \left(\sum_{g \in \Gamma_j} P_g^{2^i} \right) \tag{13}$$

Note: For a fixed j, for each $g \in \Gamma_j$, there exists a unique integer i in the set $\{0, 1, \ldots, d-1\}$ such that $2^i g \bmod n = j$. The summation in Eq. (13) is to be understood to be over all pairs (g, i) such that $g \in \Gamma_j$, $i \in \{0, 1, \ldots, d-1\}$, and $2^i g \bmod n = j$. We will observe this same convention in summations like that in Eq. (13) in the remainder of the article.

Example 3. With $n = 15$ and $j = 10$, we have from Example 1 $\Gamma_{10} = \{5, 10\}$, and so Lemma 2 implies that $\mathcal{P}_{10} = \mathrm{Tr}_2^4(P_5^2 + P_{10}) = (P_5^2 + P_{10}) + (P_5^2 + P_{10})^4 = P_5^2 + P_{10} + P_5^4 + P_{10}^4$. ❐

Proof of Lemma 2: By definition (Eq. (12)),

$$\mathcal{P}(x) = \mathrm{Tr}_1^m(P(x)) \bmod x^n - 1$$

$$= \sum_{g=0}^{n-1} \mathrm{Tr}_1^m(P_g x^g) \bmod x^n - 1$$

Since the exponents appearing in the expansion of $\mathrm{Tr}_1^m(P_g x^g)$ are exactly those which are modulo n conjugates of g, it follows that, in calculating the coefficient of x^j in $\mathcal{P}(x)$, only those indices g which are conjugates of j, i.e, elements of Γ_j, need be considered. In other words, the coefficient of x^j in $\mathcal{P}(x)$ equals the coefficient of x^j in

$$\mathrm{Tr}_1^m \left(\sum_{g \in \Gamma_j} P_g x^g \right) \bmod x^n - 1 \tag{14}$$

Now we need to invoke the fact that, if d is a divisor of m, then

$$\mathrm{Tr}_1^m(\xi) = \mathrm{Tr}_1^d(\mathrm{Tr}_d^m(\xi)) \tag{15}$$

[2, Theorem 8.2]. Combining Eqs. (14) and (15), we find that the coefficient of x^j in $\mathcal{P}(x)$ is in fact the coefficient of x^j in

$$\mathrm{Tr}_1^d \left(\mathrm{Tr}_d^m \sum_{g \in \Gamma_j} P_g x^g \right) \bmod x^n - 1 \tag{16}$$

Now since each element of Γ_j has degree d, it follows that $\mathrm{Tr}_d^m(P_g x^g) = x^g \mathrm{Tr}_d^m(P_g) \bmod x^n - 1$ for each $g \in \Gamma_j$, so that Expression (16) becomes

$$\mathrm{Tr}_1^d \sum_{g \in \Gamma_j} x^g \mathrm{Tr}_d^m(P_g) \tag{17}$$

Recalling that $\mathrm{Tr}_1^d(\xi) = \xi + \xi^2 + \cdots + \xi^{2^{d-1}}$, we see that if i is the unique index in the range $0 \leq i \leq d-1$ such that $2^i g = j$, the gth term in the sum in Eq. (17) contributes exactly

$$\mathrm{Tr}_d^m(P_g)^{2^i} = \mathrm{Tr}_d^m(P_g^{2^i})$$

to the coefficient of x^j. But this is exactly what Eq. (13) says. ❐

Corollary. If j_1 and j_2 are conjugate modulo n, then $\mathcal{P}_{j_1}$ and $\mathcal{P}_{j_2}$ are conjugates in $GF(2^m)$. More precisely, if j has degree d, and if $s \in \{0, 1, \ldots, d-1\}$, then

$$\mathcal{P}_{2^s j \bmod n} = \mathcal{P}_j^{2^s}$$

Proof: If we use Lemma 2 to compute $\mathcal{P}_j^{2^s}$, we find

$$\mathcal{P}_j^{2^s} = \left(\mathrm{Tr}_d^m \sum_{\substack{g \in \Gamma_j \\ 2^i g \bmod n = j}} P_g^{2^i} \right)^{2^s}$$

$$= \mathrm{Tr}_d^m \sum_{\substack{g \in \Gamma_j \\ 2^i g \bmod n = j}} P_g^{2^{i+s}}$$

$$= \mathrm{Tr}_d^m \sum_{\substack{g \in \Gamma_j \\ 2^k g \bmod n = 2^s j}} P_g^{2^k}$$

$$= \mathcal{P}_{2^s j} \qquad \square$$

Example 4. For $m = 4$ and $n = 5$, we have from the Corollary, with $j = 1$ and $s = 0, 1, 2, 3$, that

$$\mathcal{P}_{2^0 \bmod 5} = \mathcal{P}_1 = \mathcal{P}_1$$

$$\mathcal{P}_{2^1 \bmod 5} = \mathcal{P}_2 = \mathcal{P}_1^2$$

$$\mathcal{P}_{2^2 \bmod 5} = \mathcal{P}_4 = \mathcal{P}_1^4$$

$$\mathcal{P}_{2^3 \bmod 5} = \mathcal{P}_3 = \mathcal{P}_1^8$$

These relationships can be verified directly by referring to Example 2. $\square$

Now we are prepared to begin the proof of Theorem 2. In effect, we wish to count the number of sets $(c_j)_{j \in J}$ such that Eq. (8) holds. If we substitute the formula given in Eq. (7) into Eq. (8), we find that $(c_j)_{j \in J}$ defines a word in the TSRS code $\mathcal{C}_j^\mu$ if and only if

$$\mathrm{Tr}_1^m \left(\sum_{j \in J} \alpha^h c_j x^j \right) = 0 \quad \begin{cases} h = 0, 1, \ldots, \mu - 1 \\ x \in \{1, \alpha, \ldots, \alpha^{n-1}\} \end{cases} \tag{18}$$

Now, for $h = 0, 1, \ldots, \mu - 1$, we define the polynomial $P_h(x)$ as

$$P_h(x) = \sum_{j \in J} \alpha^h c_j x^j$$

Thus Eq. (18) holds if and only if $\mathrm{Tr}_1^m(P_h(x)) = 0$ for all $x \in \{1, \alpha, \ldots, \alpha^{n-1}\}$, for all $h = 0, 1, \ldots, \mu - 1$. By Lemma 1, this will be true if and only if $\mathcal{P}_{h,j} = 0$ for all $h = 0, 1, \ldots, \mu - 1$ and all $j \in J$, where $\mathcal{P}_{h,j}$ is the coefficient of x^j in the polynomial $\mathcal{P}_h(x) = \mathrm{Tr}_1^m(P_h(x)) \bmod (x^n - 1)$ (Eq. (12)).

Now by Lemma 2, if $d = \deg(j)$, the coefficient $\mathcal{P}_{h,j}$ is given by the formula

$$\mathcal{P}_{h,j} = \mathrm{Tr}_d^m \left(\sum_{g \in J_j} c_g^{2^i} \alpha^{h 2^i} \right) \tag{19}$$

where in Eq. (19), $J_j = \Gamma_j \cap J$.

In summary, a set $(c_j)_{j \in J}$ of elements from $GF(2^m)$ corresponds to a codeword in $\mathcal{C}_J^\mu$ if and only if $\mathcal{P}_{h,j}$, as defined in Eq. (19), is zero, for all $h = 0, 1, \ldots, \mu - 1$ and all $j \in J$. However, by the Corollary to Lemma 2, conjugate j's correspond to conjugate $\mathcal{P}_{h,j}$'s, and so if $\mathcal{P}_{h,j} = 0$ for *one* element j of a given cyclotomic coset, it will be zero for all other elements of the coset as well. Thus in "solving" the equations $\mathcal{P}_{h,j} = 0$, it is sufficient to restrict j to lie in the set I_n, consisting of the least element of each cyclotomic coset. Thus if we want to count the number of sets $(c_j)_{j \in J}$ corresponding to codewords in the TSRS code $\mathcal{C}_j^\mu$, we get one set of equations of the form $\mathcal{P}_{h,j} = 0$ for $h = 0, 1, \ldots, \mu - 1$ for each modulo n cyclotomic coset, i.e., each $j \in I_n$.

To simplify the notation, for each $g \in J_j$, where $j \in I_n$, we define $x_g = c_g^{2^i}$, where according to our convention i is the unique index such that $2^i g \bmod n = j$. Note that since the mapping $\xi \to \xi^{2^i}$ is one to one, the c_g's can be uniquely recovered from the x_g's. For the remainder of the proof, we shall focus on the problem of determining when a set $(x_g)_{g \in J}$ corresponds to a codeword in $\mathcal{C}_j^\mu$. By the foregoing discussion and Eq. (19), this will be true if and only if

$$\mathrm{Tr}_d^m \left(\sum_{g \in J_j} x_g \alpha^{h 2^i} \right) = 0 \qquad \text{for } h = 0, 1, \ldots, \mu - 1 \tag{20}$$

for all $j \in I_n$. Since a set of equations of the form of Eq. (20) involves only those variables x_g corresponding to g's in a fixed cyclotomic coset, it follows that if the number of solutions to Eq. (20) is denoted by N_j, then the total number of codewords in the code $\mathcal{C}_J^\mu$ is simply $\prod_{j \in I_n} N_j$.

Let Γ be one of the cyclotomic cosets modulo n, with $|\Gamma| = d$, and let E be a subset of Γ, with $|E| = e$. Let $(x_g)_{g\in E}$ be e variables taking values in the field $GF(2^m)$, which satisfy the μ simultaneous linear equations

$$\mathrm{Tr}_d^m\left(\sum_{g\in E} x_g \alpha^{h2^g}\right) = 0 \qquad \text{for } h = 0, 1, \ldots, \mu - 1 \quad (21)$$

Theorem 3. The set of solutions $(x_g)_{g\in E}$ to Eq. (21) is a vector space over $GF(2^d)$, of dimension $(ef-\mu)_+$, where $f = m/d$. The number of solutions is therefore $2^{d(ef-\mu)_+}$.

Theorem 3, combined with the previous discussion, completes the proof of Theorem 2, since it implies that, for each $j \in I_n$, the number of solutions to Eq. (20) is $2^{d_j(e_j f_j - \mu)_+}$, where $f_j = m/d_j$, which is the same as $2^{(me_j - \mu d_j)_+}$, as asserted in Theorem 2.

Proof of Theorem 3: The fact that the set of solutions to Eq. (21) is a vector space over $GF(2^d)$ follows from the fact that Tr_d^m is a linear mapping from $GF(2^m)$ to $GF(2^d)$, i.e, that if x and x' are elements of $GF(2^m)$, and if λ and λ' are elements of $GF(2^d)$, then

$$\mathrm{Tr}_d^m(\lambda x + \lambda' x) = \lambda \,\mathrm{Tr}_d^m(x) + \lambda' \,\mathrm{Tr}_d^m(x) \quad (22)$$

Using Eq. (22), it is easy to see that if (x_g) is one solution to Eq. (21), and if (x'_g) is another, and if λ and λ' are elements of $GF(2^d)$, then $(\lambda x_g + \lambda' x'_g)$ is also a solution to Eq. (21). In the remainder of the proof, we will show that the $GF(2^d)$ dimension of the solution space to Eq. (21) is $(ef-\mu)_+$.

To simplify notation, let $q = 2^d$, so that $2^m = q^f$, and let $\tau = \mathrm{Tr}_d^m$. If now $(x_g)_{g\in E}$ is an arbitrary vector from $GF(2^m)^e \cong GF(q^f)^e$, define, for $h = 0, 1, \ldots, \mu - 1$,

$$y_h = \sum_{g\in E} x_g \alpha^{h2^g} \quad (23)$$

and

$$z_h = \tau(y_h) \quad (24)$$

Now let T be the $GF(q)$-linear mapping from $GF(q^f)^e = GF(q)^{ef}$ to $GF(q)^\mu$, defined by

$$T : (x_g)_{g\in E} \rightarrow (z_h)_{h=0,\ \ldots,\ \mu-1}$$

We will show that the rank of the mapping T is $\min(ef, \mu)$, so that the nullity of T is $(ef-\mu)_+$, and this will complete the proof. We will consider the cases $ef \geq \mu$ and $ef < \mu$ separately.

Case 1: $ef \geq \mu$. We need to show that $\mathrm{rank}(T) = \mu$. This is equivalent to showing that there is no set of nonzero (λ_h)'s such that

$$\sum_{h=0}^{\mu-1} \lambda_h z_h = 0 \quad (25)$$

for all vectors (x_g) in $GF(q^f)^e$. If Eq. (25) *is* true, then from Eq. (24), we have

$$\tau\left(\sum_{h=0}^{\mu-1} \lambda_h y_h\right) = 0 \quad (26)$$

for all (x_g)'s. The inner sum in Eq. (26) is, by Eq. (23),

$$\sum_{h=0}^{\mu-1} \lambda_h y_h = \sum_{h=0}^{\mu-1} \lambda_h \sum_{g\in E} x_g \alpha^{h2^g}$$

$$= \sum_{g\in E} x_g \sum_{h=0}^{\mu-1} \lambda_h \alpha^{h2^g}$$

It follows then that

$$\tau\left(\sum_{g\in E} x_g \beta_g\right) = 0$$

for all (x_g), where $\beta_g = \sum_{h=0}^{\mu-1} \lambda_h \alpha^{h2^g}$. But it is easy to see that this can hold if and only if $\beta_g = 0$ for all $g \in E$. In summary, then, Eq. (25) will be true for all (x_g)'s if and only if

$$\sum_{h=0}^{\mu-1} \lambda_h \alpha^{h2^g} = 0 \qquad \text{for all } g \in E \quad (27)$$

Next, define the polynomial $L(x)$ as

$$L(x) = \sum_{h=0}^{\mu-1} \lambda_h x^h$$

Then Eq. (27) says $L(\alpha^{2^i}) = 0$ for all $g \in E$, so that $L(x)$ has $|E| = e$ distinct roots in the set $\{\alpha, \alpha^2, \ldots, \alpha^{2^{d-1}}\}$. However, since the coefficients of $L(x)$ are in $GF(q)$, then every $GF(q)$ conjugate of a root of $L(x)$ is also a root. But conjugation with respect to $GF(q)$ is the mapping $\beta \rightarrow \beta^q = \beta^{2^d}$. But since a is a primitive root, each element of $GF(2^m) = GF(2^{df})$ of the form α^{2^i} with $0 \leq i \leq d-1$ has exactly f $GF(2^d)$ conjugates, viz, α^{2^k} for $k = i, i+d, \ldots, i+d(f-1)$. Hence by taking conjugates of the original e roots of $L(x)$, we obtain ef roots. But since $ef \geq \mu$ by assumption, and since $\deg L(x) \leq \mu - 1$, it follows that the coefficients of $L(x)$, i.e, the λ_h's, are all zero. This means that no nontrivial relationship of the form of Eq. (25) can hold, which completes the proof in Case 1.

Case 2: $ef < \mu$. In this case we need to show that $\text{rank}(T) = ef$. But clearly $\text{rank}(T) \leq ef$, since the $GF(q)$-dimension of the space of all (x_g)'s is ef. However, by the argument in Case 1, the first ef components of (z_g) are linearly independent, and so $\text{rank}(T) \geq ef$ as well. Hence $\text{rank}(T) = ef$, as asserted. ❐

IV. Summary and Conclusions

In this article, we have introduced an extensive class of symbol-oriented error-correcting codes, which have properties much like those of Reed–Solomon codes, without, however, suffering from the major drawback of RS codes, viz, an intrinsic limitation on codeword length. As subcodes of RS codes, these codes can be decoded by any RS decoding algorithm. However, the study of these codes is in its infancy, and we therefore close with a list of unsolved problems related to TSRS codes.

(1) Our selection of the representation of an element from $GF(2^m)$ as a binary m-tuple was more or less arbitrary. If another representation is used, will a code of larger dimension result?

(2) Devise an efficient encoding algorithm for an arbitrary TSRS code, or at least a large class of them.

(3) Determine the conditions under which TSRS codes are systematic over the symbol alphabet $V(2^{m-\mu})$. (A necessary condition for this to be so is that the pseudodimension k_μ be an integer. But in [4], it was shown by example that this condition is not sufficient.)

(4) Study the combinatorial optimality of TSRS codes. For example, devise bounds on the cardinality of a (n, k, d) code over a q-letter alphabet when q is a fixed fraction of n, say λn, as $n \rightarrow \infty$.

(5) Investigate the relationship between TSRS codes and generalized BCH codes, i.e., BCH codes whose symbol field is $GF(2^d)$ and whose locator field is $GF(2^m)$, where d is a divisor of m.

(6) Compare the distance properties of TSRS codes to algebraic geometry codes with approximately the same values of n, k, and q.

(7) Do TSRS codes meet (or exceed) the Gilbert-Varshamov bound?

References

[1] M. Hall, *The Theory of Groups*, New York: Macmillan, 1959.

[2] R. J. McEliece, *Finite Fields for Computer Scientists and Engineers*, Boston: Kluwer, 1986.

[3] I. S. Reed and G. Solomon, "Polynomial Codes over Certain Finite Fields," *J. Soc. Indus. Appl. Math.*, vol. 8, pp. 300–304, 1960.

[4] G. Solomon, "Nonlinear, Nonbinary Cyclic Group Codes," *The Telecommunications and Data Acquisition Progress Report 42-108, vol. October–December 1991*, Jet Propulsion Laboratory, Pasadena, California, pp. 84–95, February 15, 1992.

[5] G. Solomon, "Nonlinear, Nonbinary Cyclic Group Codes," *Proceedings of the 1993 International Symposium on Information Theory*, San Antonio, Texas, p. 192, January 17–22, 1993.

[6] J. H. Yuen, ed., *Deep Space Telecommunications Engineering*, JPL Publication 82–76, Jet Propulsion Laboratory, Pasadena, California, July 1982.