It is easy to show that this bound is sharp. Let $V = \mathrm{GF}(2)$ and $e_1, \ldots, e_n$ be the coordinate vectors which span $V^n$. The code $C$ spanned by any $n-2$ of these vectors, say $C = \langle e_1, e_2, \ldots, e_{n-2} \rangle$, has covering radius 2. Adding another vector to the code reduces the covering radius to 1.

*Corollary 1:* If $C$ is any code of covering radius $R$, then the covering radius of any subcode $C_0$ of index 2 is at most $2R+1$.

*Corollary 2:* Let $C$ be a code of covering radius $R$ and norm $N$. Then $N \leq 3R+1$.

*Proof:* Let $x$ be any vector such that $d(x, C) = R$. Then, for any subcode $C_0$ of index 2, $d(x, C_0) \leq 2R+1$. Thus, $N \leq 3R+1$. ■

### References

[1] R. L. Graham and N. J. A. Sloane, "On the covering radius of codes," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 385–401, May 1985.

[2] W. Wesley Peterson and E. J. Weldon, Jr., *Error-Correcting Codes*, 2d ed. Cambridge, MA: MIT Press, 1972.

# On the Decoder Error Probability for Reed–Solomon Codes

ROBERT J. McELIECE AND LAIF SWANSON

***Abstract*—Upper bounds on the decoder error probability for Reed-Solomon codes are derived. By definition, "decoder error" occurs when the decoder finds a codeword other than the transitted codeword; this is in contrast to "decoder failure," which occurs when the decoder fails to find any codeword at all. These results imply, for example, that for a $t$ error-correcting Reed-Solomon code of length $q-1$ over GF($q$), if more than $t$ errors occur, the probability of decoder error is less than $1/t$!**

## I. Introduction

Let $C$ be an $(n, k)$ code over GF($q$) with minimum distance $d$. We assume $C$ is being used to correct $t$ errors, where $t$ is a fixed integer satisfying $2t \leq d-1$. We further assume the decoder is a bounded distance decoder, that is, it looks for a codeword within distance $t$ of the received word; if there is such a codeword, the decoder finds it, and if not, the decoder reports "failure."

If the transmitted codeword suffers $t$ or fewer errors, it will be decoded correctly. If, on the other hand, it suffers more than $t$ errors, one of two things can happen. Either the decoder will fail to find a codeword (decoder failure), or it will find a codeword other than the transmitted codeword (decoder error). We denote by $P_F$ and $P_E$ the probabilities of decoder failure and error, respectively. Of course, if the number of errors is $t$ or less, $P_F = P_E = 0$. If the number of errors exceeds $t$, but is less than $d-t$, then $P_F = 1$ and $P_E = 0$, since fewer than $d-t$ errors cannot move the transmitted codeword to within distance $t$ of another codeword.

If $d-t$ or more errors occur, it is in general quite difficult to calculate, or even estimate, $P_F$ and $P_E$, although if the code is being used in a practical communications system, it is important to do so. A useful heuristic estimate can be based on the assumption that if at least $d-t$ errors occur, the error pattern can be treated as if it were completely random. The probability that a completely random error pattern will cause decoder error (i.e., lie within distance $t$ of a nonzero codeword) is given by

$$Q = \frac{(q^k - 1) \cdot V_n(t)}{q^n} = (q^{-r} - q^{-n}) V_n(t), \tag{1}$$

where $r = n - k$ is the code's redundancy and

$$V_n(t) = \sum_{s=0}^{t} \binom{n}{s} (q-1)^s \tag{2}$$

is the volume of a Hamming sphere of radius $t$. This argument leads to the following estimate for $P_E$:

$$P_E \approx Q \cdot \Pr\{\geq d-t \quad \text{errors}\}. \tag{3}$$

It is difficult to justify this estimate in general, but in this paper we will see that if we increase $Q$ slightly by defining $Q'$ as

$$Q' = (q-1)^{-r} V_n(t), \tag{4}$$

then for Reed–Solomon (RS) codes,

$$P_E \leq Q' \cdot \Pr\{\geq d-t \quad \text{errors}\}. \tag{5}$$

In fact (5) will follow from more detailed results, which we now describe.

If $q_u$ denotes the probability that the error pattern has weight $u$, then plainly

$$P_E = \sum_{u=0}^{n} P_E(u) q_u, \tag{6a}$$

$$P_F = \sum_{u=0}^{n} P_F(u) q_u, \tag{6b}$$

where $P_E(u)$ and $P_F(u)$ denote the conditional probabilities of decoder error and failure, respectively, given $u$ channel errors. As mentioned above, we have $P_E(u) = P_F(u) = 0$ for $u \leq t$ and $P_E(u) = 0$, $P_F(u) = 1$ for $t < u < d-t$. For $u \geq d-t$ we have $P_F(u) + P_E(u) = 1$, and so if $P_E(u)$ is known, $P_F(u)$ can be calculated, and vice versa.

Here is our main result. Let $C$ be an $(n, k)$ RS, or any other maximum distance separable (MDS) code, with minimum distance $d = n - k + 1$. We assume as above that the code is being used to correct $t$ errors, for some fixed value of $t$ with $2t \leq d-1$. We further assume that the code is being used on a channel for which all error patterns of the same weight are equiprobable, for example, a $q$-ary symmetric channel. Under these assumptions, we shall prove in Section III that

$$P_E(u) = 0, \qquad \text{for } u \leq d-t-1 \tag{7a}$$

$$P_E(u) \leq (q-1)^{-r} \sum_{s=d-u}^{t} \binom{n}{s} (q-1)^s, \qquad \text{for } d-t \leq u \leq d-1 \tag{7b}$$

$$P_E(u) \leq Q', \qquad \text{for } u \geq d. \tag{7c}$$

Of course (7a) needs no further proof; it is included only to make the bounds in (7) apply to all values of $u$. The bound (7b) actually follows from a slightly sharper, but more complicated bound on $P_E(u)$ that appears in Section III as (15).

We can combine (7b) and (7c), at the cost of weakening (7b) slightly, to obtain an upper bound on $P_E(u)$ which is uniform in

Manuscript received October 22, 1985; revised January 3, 1986. Part of this work was presented at the IEEE International Symposium on Information Theory, Brighton, England, June 1985. This work was supported by the National Aeronautics and Space Administration under contract NAS7-918. R. J. McEliece's contribution was also supported by a grant from IBM.

The authors are with the Department of Electrical Engineering and Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91125.

IEEE Log Number 8608898.

0018-9448/86/0900-0701$01.00 ©1986 IEEE

$u$ for $u \ge d - t$:

$$P_E(u) \le Q', \qquad \text{for } u \ge d - t.^{\dagger} \tag{8}$$

The ratio of this uniform bound $Q'$ to the heuristic estimate $Q$ in (1) is usually very close to 1, and is always less than $(q/(q-1))^n$, which for $n \le q - 1$ cannot exceed $e = 2.718 \cdots$. In any event, combining (6a) with (7a) and (8), we obtain the bound (5).

Although as a practical matter it is not hard to compute the bound $Q'$ numerically, for some applications it may be worthwhile to have a simpler, though weaker, bound. In the Appendix, we show that (8) implies that provided $n \le q - 1$, for all $u \ge d - t$,

$$P_E(u) \le \begin{cases} \dfrac{1}{(q-1)^{r-2}} + \dfrac{1}{(q-1)^{r}}, & \text{if } t = 1 \\[2ex] \dfrac{1}{(q-1)^{r-2t}} \cdot \dfrac{1}{t!}, & \text{if } t \ge 2. \end{cases} \tag{9}$$

Since $r \ge 2t$ in all cases, (9) implies, whenever $n \le q - 1$,

$$P_E(u) \le \frac{1}{t!}, \qquad \text{for all } u \ge t + 1. \tag{10}$$

Kasami and Lin [2] have also studied the problem of decoder error for RS codes. They showed that on a $q$-ary symmetric channel $P_E$ is at most $Q$, that is,

$$\sum_{u=d-t}^{n} P_E(u)\binom{n}{u}\epsilon^u(1-\epsilon)^{n-u} \le Q, \tag{11}$$

where $\epsilon$ is the probability of channel symbol error. They further showed that $P_E = Q$ only when $\epsilon = (q-1)/q$, that is, when the error pattern is completely random. This shows that $Q$ is the tightest possible bound on the sum in (11) which is independent of $\epsilon$. However, except when the probability of $\ge d - t$ errors is very nearly one, our bound (5) will be smaller than Kasami and Lin's bound (11). And since most well-designed systems will have $\Pr\{u \ge d - t\} \ll 1$, we conclude that our bound is likely to be more useful in practice than Kasami and Lin's.

Finally we note that since with $\epsilon = (q-1)/q$ equality holds in (11), the average of the $P_E(u)$'s with respect to one particular probability distribution is $Q$. Since $P_E(u)$ is 0 for $u < d - t$, it follows that for some values of $u$, $P_E(u) > Q$. Thus the conjecture that $P_E(u) \le Q$ for all $u$ is not tenable. (It would be nice to have a uniform lower bound on the $P_E(u)$'s, but we have been unable to find one.)

Here is the plan for the rest of the paper. In Section II we review certain known facts about MDS codes, and in particular obtain a (known) upper bound on the number of words of a given weight in an MDS code. Then in Section III we use this inequality, together with some of the ideas used to obtain it, to prove the results promised in (7). Finally, in the Appendix, we obtain an upper bound on $V_n(t)$ which can be used, together with (8), to prove (9).

## II. Preliminaries

In this section we will review some known results about MDS codes which are needed in our proof. Our remarks will be self-contained, but proofs may also be found in [4, chapter 11].

Let $C$ be a code, not necessarily linear, of length $n$ with $q^k$ codewords over GF$(q)$. If we examine any set of $k - 1$ components of the codewords, we find that there are only $q^{k-1}$ possibilities for the $q^k$ codewords. Thus there must be a pair of codewords that agree on these $k - 1$ components, and so the minimum distance $d$ of the code must satisfy $d \le n - k + 1$. A code for which $d = n - k + 1$ is called a *maximum distance separable* (MDS) code. By this definition, RS codes and cosets of RS codes are MDS codes.

Let $K$ be a subset of $k$ coordinate positions of an MDS code. If two codewords were equal on $K$, the distance between them would be at most $n - k$. But this is impossible, since $d = n - k + 1$. We conclude that all $q^k$ codewords are different on $K$, and so, for any possible $k$-tuple of elements from GF$(q)$, say, $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_k)$, there is a unique codeword that, when restricted to $K$, equals $\alpha$. This important fact we call the the basic combinatorial property of MDS codes.

We now wish to estimate the number of codewords of weight $u$, for $u \ge d$, in an MDS code. A word of weight $u$ must vanish on a set of $v = n - u$ coordinates. Thus let $V$ be an arbitrary subset of $v$ coordinates. We will estimate the number of codewords that vanish on $V$. Since $u \ge d$, then $v \le k - 1$. Thus by the basic combinatorial property, if we specify that the codeword is zero exactly on $V$, we may specify $k - v$ other, nonzero, components arbitrarily. There are $(q-1)^{k-v} = (q-1)^{u-r}$ ways to do this, and so there are at most $(q-1)^{u-r}$ codewords that vanish exactly on $V$. Since there are $\binom{n}{v} = \binom{n}{u}$ possibilities for $V$, if $A_u$ denotes the number of codewords of weight $u$, we have

$$A_u \le \binom{n}{u}(q-1)^{u-r}, \qquad \text{for } u \ge d. \tag{12}$$

(Actually, an exact formula for $A_u$ is known for linear MDS codes [4, Theorem 6, chap. 11]. This formula is however rather complicated, and we have found the estimate in (12) to be more useful in the present application.) Next we let $V$ be a subset of $v$ coordinate positions, where $v \ge k$. If we project the original code onto $V$ (this process is usually called puncturing the original code), the result will be a certain $(v, k)$ code. Since the parent $(n,k)$ code has $d = n - k + 1$, the punctured code must have distance $d' \ge d - (n - v) = v - k + 1$. Since it is impossible for $d'$ to be greater than $v - k + 1$, equality must hold and it follows that the punctured code is a $(v, k)$ MDS code. This simple fact will be referred to in the proof in the next section.

## III. Proof of Results

We call a word, not necessarily a codeword, *decodable* if it lies within distance $t$ of some codeword. If $D_u$ denotes the number of decodable words of weight $u$, then for $u \ge t + 1$ we have, assuming that all error patterns of weight $u$ are equiprobable,

$$P_E(u) = \frac{D_u}{\binom{n}{u}(q-1)^u}. \tag{13}$$

Thus the problem of finding the $P_E(u)$'s is essentially the same as that of finding the weight enumerator for the set of decodable words. For example, (7c) is equivalent to

$$D_u \le \binom{n}{u}(q-1)^{u-r}V_n(t), \qquad \text{for } u \ge d. \tag{14}$$

The plan is to obtain upper bounds on $D_u$ that will imply our various bounds on $P_E(u)$. We need to distinguish two cases, $u \ge d$ and $u \le d - 1$.

First we assume $u \ge d$. Each decodable word can be written uniquely as $C + E$, where $C$ is a codeword and $E$ is a word of weight $\le t$. For a fixed $E$, as $C$ runs through the set of codewords, $\{C + E\}$ is a coset of the RS code. Since any coset of an RS code is an MDS code, by (12) we know that the number of words of weight $u$ is less than or equal to $\binom{n}{u}(q-1)^{u-r}$, since we are assuming $u \ge d$. Since the set of decodable words is the disjoint union of $V_n(t)$ cosets of the RS code, (14), and so (7c), follow.

Now we assume that $u \le d - 1$. A decodable word of weight $u$ will vanish on a set of size $v = n - u$. For each of the $\binom{n}{v}$ subsets $V$ of $v$ coordinates, we will obtain an upper bound on the

[†] After we obtained our proof of (8), Jørn Justesen sent us an unpublished proof of (8) due to V. Zyablov. This proof is quite different from ours, and apparently cannot be used to obtain (7b) or (15). Still, we do wish to acknowledge the existence, and indeed the priority, of Zyablov's proof.

number of decodable words of weight $u$ that vanish on $V$. This upper bound will imply (7b).

As before, we will use the fact that each decodable word is of the form $C+E$, where $C$ is a codeword and $E$ has weight $\le t$. If $C+E$ vanishes on $V$, then $C$ must have weight $\le t$ on $V$, say weight $w$. We note that $w=0$ is not possible, since $u \ge t+1$. By our remarks in Section II, we know that $C$ restricted to $V$ is a linear $(v, k)$ MDS code, and so its minimum weight (distance ) is $d-u$. Thus $w$, the weight of $C$ on $V$, satisfies $d-u \le w \le t$. (If $d-u > t$, there are no such words; this gives another proof of (7a).) By (12), it follows that the number of codewords with weight $w$ on $V$ is at most $\binom{v}{w}(q-1)^{w-r'}$, where $r' = r-u$ is the redundancy of the restricted code.

For each codeword $C$ with weight $w$ in $V$, we must count the number of $E$'s such that $C+E$ vanishes on $V$. Suppose that $E$ has weight $s \ge w$. On $V$, $E$ must match $C$ exactly, but the $s-w$ other nonzero components can be arbitrarily placed outside $V$. Thus the total number of $E$'s, for a given $C$ of weight $w$, is

$$\sum_{s=w}^{t}\binom{u}{s-w}(q-1)^{s-w}.$$

Therefore the total number of decodable words vanishing on $V$ is at most

$$\sum_{w=d-u}^{t}\binom{v}{w}(q-1)^{w-r'}\sum_{s=w}^{t}\binom{u}{s-w}(q-1)^{s-w}$$
$$=(q-1)^{-r'}\sum_{s=d-u}^{t}(q-1)^{s}\sum_{w=d-u}^{s}\binom{v}{w}\binom{u}{s-w}.$$

This is a bound on the number of decodable words of weight $u$ vanishing on $V$. If we multiply it by the number of possible subsets $V$ with $v$ elements, namely, $\binom{n}{v}=\binom{n}{u}$, we obtain a bound on $D_u$, and hence by (13),

$$P_E(u) \le (q-1)^{-r}\sum_{s=d-u}^{t}(q-1)^{s}\sum_{w=d-u}^{s}\binom{v}{w}\binom{u}{s-w}. \quad (15)$$

This bound is a bit clumsy for everyday use, but we note in passing that for $u=d-t$ (the smallest value of $u$ for which $P_E(u)$ is not 0) it simplifies to

$$P_E(d-t) \le (q-1)^{-(d-t-1)}\binom{n-d+t}{t}, \quad (16)$$

which is in fact the exact value of $P_E(u)$ in this case [1].

Finally, we simplify the bound (15) by recalling a well-known combinatorial identity ([3, (1.2.6.21)]):

$$\sum_{w\ge 0}\binom{v}{w}\binom{u}{s-w}=\binom{v+u}{s}.$$

Since $v+u=n$, this means that the inner sum in (15) is at most $\binom{n}{s}$, and so (7b) follows from (15).

## Appendix

In this Appendix we will derive (9) from (8). We wish to thank one of our referees for providing the following proof, which is much simpler than our original one. The key is the inequality

$$V_n(t) \le \frac{n^t}{t!}(q-1)^t, \qquad \text{for } q \ge 4,\ t \ge 2, \text{ and } n \le q-1. \quad \text{(A.1)}$$

(Note that for $q=2$ or 3, the only possible MDS codes of length $\le q-1$ must have $t=0$.) This inequality can be verified directly for $t=2$, and can be proved for $t \ge 3$ by induction, as follows.

$$V_n(t)=\binom{n}{t}(q-1)^t+V_n(t-1)$$
$$\le \binom{n}{t}(q-1)^t+\frac{n^{t-1}}{(t-1)!}(q-1)^{t-1}$$
$$\text{(induction hypothesis)}$$
$$=\left\{\frac{n!}{(n-t)!n^t}+\frac{t}{n(q-1)}\right\}\frac{n^t(q-1)^t}{t!}.$$

The first term in the braces is less than or equal to $(n-1)/n$, and the second term is less than or equal to $1/n$. The inequality (A.1) thus follows.

If $n \le q-1$, the bound (A.1) immediately implies (9) for $t \ge 2$. The case $t=1$ in (9) must be handled separately, and follows from the fact that $V_n(1)=1+n(q-1)$ is less than or equal to $1+(q-1)^2$, if $n \le q-1$.

## References

[1] E. R. Berlekamp and J. L. Ramsey, "Readable erasures improve the performance of Reed-Solomon codes," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 632–633, Sep. 1978.

[2] T. Kasami and S. Lin, "On the probability of undetected error for the maximum distance separable codes," *IEEE Trans. Comm.*, vol. COM-32, pp. 998–1006, Sep. 1984.

[3] D. E. Knuth, *The Art of Computer Programming, vol. 1, Fundamental Algorithms*. Reading, MA: Addison-Wesley, 1968.

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

# Inequivalent Cyclic Codes of Prime Length

DONALD W. NEWHART, ASSOCIATE MEMBER, IEEE

*Abstract*—**A characterization of the equivalence classes of prime-length cyclic codes over any finite field is given, generalizing the binary case solved by Leon, Masely, and Pless. In the special case of cyclic $(p,k)$ codes over GF$(q)$, with $p|(q-1)$, a one-to-one correspondence between the equivelance classes and the orbits of $k$-sets under the affine group, GA$(1,p)$ is established.**

In general, two linear $(n,k)$ codes $C_1, C_2$ over a field $F$ are said to be equivalent if for some $n \times n$ monomial matrix, $M$, with entries in $F$, we have $C_1 \cdot M = C_2$, where the equal sign denotes set equality [1]. When $F$ is GF(2), an equivalence map involves only a permutation of the coordinates of the codewords. For all other fields, we need to consider the combination of coordinate permutations followed by the (nonzero) scalar action on coordinate entries by a diagonal matrix to reach all the equivalents of a given code. In [3] it was shown that if $p$ is prime and $C_1, C_2$ are equivalent binary cyclic codes of length $p$, then they are equivalent under the affine group GA$(1,p)$, that is, there is a permutation on the coordinates $0, 1, \cdots, p-1$ of the form: $i \to ai+b \pmod p$, where $a, b \in \text{GF}(p)$ with $a$ nonzero, which maps $C_1$ to $C_2$. In this correspondence we extend this characterization to include nonbinary codes, although we cast our result in terms of the roots of the generator polynomials of the codes.

Before stating the result, we shall make some observations to simplify the discussion. The automorphism group of a linear

Manuscript received April 23, 1985; revised January 3, 1986. This work was partially presented at the IEEE International Symposium on Information Theory, Brighton, England, June 1985.

The author is with the Department of Defense, 9800 Savage Road, Fort Meade, MD 20755.

IEEE Log Number 8608897.

0018-9448/86/0900-0703$01.00 ©1986 IEEE