Foreword

This chapter is based on lecture notes from coding theory courses taught by Venkatesan Guruswami at University at Washington and CMU; by Atri Rudra at University at Buffalo, SUNY and by Madhu Sudan at MIT.

This version is dated **May 8, 2015**. For the latest version, please go to

http://www.cse.buffalo.edu/ atri/courses/coding-theory/book/

The material in this chapter is supported in part by the National Science Foundation under CAREER grant CCF-0844796. Any opinions, findings and conclusions or recomendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).

Chapter 5

The Greatest Code of Them All: Reed-Solomon Codes

In this chapter, we will study the Reed-Solomon codes. Reed-Solomon codes have been studied a lot in coding theory. These codes are optimal in the sense that they meet the Singleton bound (Theorem 4.3.1). We would like to emphasize that these codes meet the Singleton bound not just asymptotically in terms of rate and relative distance but also in terms of the dimension, block length and distance. As if this were not enough, Reed-Solomon codes turn out to be more versatile: they have many applications outside of coding theory. (We will see some applications later in the book.)

These codes are defined in terms of univariate polynomials (i.e. polynomials in one unknown/variable) with coefficients from a finite field $\mathbb{F}_q$. It turns out that polynomials over $\mathbb{F}_p$, for prime p, also help us define finite fields $\mathbb{F}_{p^s}$, for $s > 1$. To kill two birds with one stone[1], we first do a quick review of polynomials over finite fields. Then we will define and study some properties of Reed-Solomon codes.

5.1 Polynomials and Finite Fields

We begin with the formal definition of a (univariate) polynomial.

Definition 5.1.1. Let $\mathbb{F}_q$ be a finite field with q elements. Then a function $F(X) = \sum_{i=0}^{\infty} f_i X^i, f_i \in \mathbb{F}_q$ is called a polynomial.

For our purposes, we will only consider the finite case; that is, $F(X) = \sum_{i=0}^{d} f_i X^i$ for some integer $d > 0$, with coefficients $f_i \in \mathbb{F}_q$, and $f_d \neq 0$. For example, $2X^3 + X^2 + 5X + 6$ is a polynomial over $\mathbb{F}_7$.

Next, we define some useful notions related to polynomials. We begin with the notion of degree of a polynomial.

[1] No birds will be harmed in this exercise.

Definition 5.1.2. For $F(X) = \sum_{i=0}^{d} f_i X^i$ ($f_d \neq 0$), we call d the *degree* of $F(X)$. We denote the degree of the polynomial $F(X)$ by $\deg(F)$.

For example, $2X^3 + X^2 + 5X + 6$ has degree 3.

Let $\mathbb{F}_q[X]$ be the set of polynomials over $\mathbb{F}_q$, that is, with coefficients from $\mathbb{F}_q$. Let $F(X), G(X) \in \mathbb{F}_q[X]$ be polynomials. Then $\mathbb{F}_q[X]$ has the following natural operations defined on it:

Addition:

$$F(X) + G(X) = \sum_{i=0}^{\max(\deg(F),\deg(G))} (f_i + g_i) X^i,$$

where the addition on the coefficients is done over $\mathbb{F}_q$. For example, over $\mathbb{F}_2$, $X + (1 + X) = X \cdot (1 + 1) + 1 \cdot (0 + 1)1 = 1$ (recall that over $\mathbb{F}_2$, $1 + 1 = 0$).[2]

Multiplication:

$$F(X) \cdot G(X) = \sum_{i=0}^{\deg(F)+\deg(G)} \left(\sum_{j=0}^{\min(i,\deg(F))} p_j \cdot q_{i-j} \right) X^i,$$

where all the operations on the coefficients are over $\mathbb{F}_q$. For example, over $\mathbb{F}_2$, $X(1 + X) = X + X^2$; $(1 + X)^2 = 1 + 2X + X^2 = 1 + X^2$, where the latter equality follows since $2 \equiv 0 \bmod 2$.

Next, we define the notion of a root of a polynomial.

Definition 5.1.3. $\alpha \in \mathbb{F}_q$ is a root of a polynomial $F(X)$ if $F(\alpha) = 0$.

For instance, 1 is a root of $1 + X^2$ over $\mathbb{F}_2$.

We will also need the notion of a special class of polynomials, which are like prime numbers for polynomials.

Definition 5.1.4. A polynomial $F(X)$ is irreducible if for every $G_1(X), G_2(X)$ such that $F(X) = G_1(X)G_2(X)$, we have $\min(\deg(G_1), \deg(G_2)) = 0$

For example, $1 + X^2$ is not irreducible over $\mathbb{F}_2$, as $(1 + X)(1 + X) = 1 + X^2$. However, $1 + X + X^2$ is irreducible, since its non-trivial factors have to be from the linear terms X or $X + 1$. However, it is easy to check that neither is a factor of $1 + X + X^2$. (In fact, one can show that $1 + X + X^2$ is the only irreducible polynomial of degree 2 over $\mathbb{F}_2$– see Exercise 5.1.) A word of caution: if a polynomial $E(X) \in \mathbb{F}_q[X]$ does not have any root in $\mathbb{F}_q$, it does *not* mean that $E(X)$ is irreducible. For example consider the polynomial $(1 + X + X^2)^2$ over $\mathbb{F}_2$– it does not have any root in $\mathbb{F}_2$ but it obviously is not irreducible.

Just as the set of integers modulo a prime is a field, so is the set of polynomials modulo an irreducible polynomial:

Theorem 5.1.1. *Let $E(X)$ be an irreducible polynomial with degree ≥ 2 over $\mathbb{F}_p$, p prime. Then the set of polynomials in $\mathbb{F}_p[X]$ modulo $E(X)$, denoted by $\mathbb{F}_p[X]/E(X)$, is a field.*

[2] This will be a good time to remember that operations over a finite field are much different from operations over integers/reals. For example, over reals/integers $X + (X + 1) = 2X + 1$.

The proof of the theorem above is similar to the proof of Lemma 2.1.2, so we only sketch the proof here. In particular, we will explicitly state the basic tenets of $\mathbb{F}_p[X]/E(X)$.

- Elements are polynomials in $\mathbb{F}_p[X]$ of degree at most $s-1$. Note that there are p^s such polynomials.
- Addition: $(F(X)+G(X)) \mod E(X) = F(X) \mod E(X) + G(X) \mod E(X) = F(X)+G(X)$. (Since $F(X)$ and $G(X)$ are of degree at most $s-1$, addition modulo $E(X)$ is just plain simple polynomial addition.)
- Multiplication: $(F(X)\cdot G(X)) \mod E(X)$ is the unique polynomial $R(X)$ with degree at most $s-1$ such that for some $A(X)$, $R(X)+A(X)E(X) = F(X)\cdot G(X)$
- The additive identity is the zero polynomial, and the additive inverse of any element $F(X)$ is $-F(X)$.
- The multiplicative identity is the constant polynomial 1. It can be shown that for every element $F(X)$, there exists a unique multiplicative inverse $(F(X))^{-1}$.

For example, for $p=2$ and $E(X) = 1+X+X^2$, $\mathbb{F}_2[X]/(1+X+X^2)$ has as its elements $\{0,1,X,1+X\}$. The additive inverse of any element in $\mathbb{F}_2[X]/(1+X+X^2)$ is the element itself while the multiplicative inverses of $1, X$ and $1+X$ are $1, 1+X$ and X respectively.

A natural question to ask is if irreducible polynomials exist. Indeed, they do for every degree:

Theorem 5.1.2. *For all $s \geq 2$ and $\mathbb{F}_p$, there exists an irreducible polynomial of degree s over $\mathbb{F}_p$. In fact, the number of such irreducible polynomials is $\Theta\left(\frac{p^s}{s}\right)$.*[3]

Given any monic [4] polynomial $E(X)$ of degree s, it can be verified whether it is an irreducible polynomial by checking if $\gcd(E(X), X^{q^s} - X) = E(X)$. This is true as the product of all monic irreducible polynomials in $\mathbb{F}_q[X]$ of degree exactly s is known to be the polynomial $X^{q^s} - X$. Since Euclid's algorithm for computing the $\gcd(F(X), G(X))$ can be implemented in time polynomial in the minimum of $\deg(F)$ and $\deg(G)$ and $\log q$ (see Section D.7.2), this implies that checking whether a given polynomial of degree s over $\mathbb{F}_q[X]$ is irreducible can be done in time $\text{poly}(s, \log q)$.

This implies an efficient Las Vegas algorithm[5] to generate an irreducible polynomial of degree s over $\mathbb{F}_q$. Note that the algorithm is to keep on generating random polynomials until it comes across an irreducible polynomial (Theorem 5.1.2 implies that the algorithm will check $O(s)$ polynomials in expectation). Algorithm 7 presents the formal algorithm.

The above discussion implies the following:

[3] The result is true even for general finite fields $\mathbb{F}_q$ and not just prime fields but we stated the version over prime fields for simplicity.

[4] I.e. the coefficient of the highest degree term is 1. It is easy to check that if $E(X) = e_s X^s + e_{s-1}X^{s-1} + \cdots + 1$ is irreducible, then $e_s^{-1} \cdot E(X)$ is also an irreducible polynomial.

[5] A Las Vegas algorithm is a randomized algorithm which always succeeds and we consider its time complexity to be its expected worst-case run time.

Algorithm 7 Generating Irreducible Polynomial

INPUT: Prime power q and an integer $s > 1$
OUTPUT: A monic irreducible polynomial of degree s over $\mathbb{F}_q$

1: $b \leftarrow 0$
2: WHILE $b = 0$ DO
3: $F(X) \leftarrow X^s + \sum_{i=0}^{s-1} f_i X^i$, where each $f_i$is chosen uniformly at random from $\mathbb{F}_q$.
4: IF $\gcd(F(X), X^{q^s} - X) = F(X)$ THEN
5: $b \leftarrow 1$.
6: RETURN $F(X)$

Corollary 5.1.3. *There is a Las Vegas algorithm to generate an irreducible polynomial of degree s over any $\mathbb{F}_q$ in expected time* $\text{poly}(s, \log q)$.

Now recall that Theorem 2.1.3 states that for every prime power p^s, there a unique field $\mathbb{F}_{p^s}$. This along with Theorems 5.1.1 and 5.1.2 imply that:

Corollary 5.1.4. *The field $\mathbb{F}_{p^s}$ is $\mathbb{F}_p[X]/E(X)$, where $E(X)$ is an irreducible polynomial of degree s.*

5.2 Reed-Solomon Codes

Recall that the Singleton bound (Theorem 4.3.1) states that for any $(n, k, d)_q$ code, $k \leq n - d + 1$. Next, we will study Reed-Solomon codes, which meet the Singleton bound, i.e. satisfy $k = n - d + 1$ (but have the unfortunate property that $q \geq n$). Note that this implies that the Singleton bound is tight, at least for $q \geq n$.

We begin with the definition of Reed-Solomon codes.

Definition 5.2.1 (Reed-Solomon code)**.** Let $\mathbb{F}_q$ be a finite field. Let $\alpha_1, \alpha_2, ...\alpha_n$ be distinct elements (also called *evaluation points*) from $\mathbb{F}_q$ and choose n and k such that $k \leq n \leq q$. We define an encoding function for Reed-Solomon code as $RS : \mathbb{F}_q^k \to \mathbb{F}_q^n$ as follows. A message $\mathbf{m} = (m_0, m_1, ..., m_{k-1})$ with $m_i \in \mathbb{F}_q$ is mapped to a degree $k-1$ polynomial.

$$\mathbf{m} \mapsto f_{\mathbf{m}}(X),$$

where

$$f_{\mathbf{m}}(X) = \sum_{i=0}^{k-1} m_i X^i. \tag{5.1}$$

Note that $f_{\mathbf{m}}(X) \in \mathbb{F}_q[X]$ is a polynomial of degree at most $k-1$. The encoding of $\mathbf{m}$ is the evaluation of $f_{\mathbf{m}}(X)$ at all the α_i's :

$$RS(\mathbf{m}) = \left(f_{\mathbf{m}}(\alpha_1), f_{\mathbf{m}}(\alpha_2), ..., f_{\mathbf{m}}(\alpha_n)\right).$$

We call this image Reed-Solomon code or RS code. A common special case is $n = q - 1$ with the set of evaluation points being $\mathbb{F}^* \stackrel{\text{def}}{=} \mathbb{F} \setminus \{0\}$.

For example, the first row below are all the codewords in the $[3,2]_3$ Reed-Solomon codes where the evaluation points are $\mathbb{F}_3$ (and the codewords are ordered by the corresponding messages from $\mathbb{F}_3^2$ in lexicographic order where for clarity the second row shows the polynomial $f_{\mathbf{m}}(X)$ for the corresponding $\mathbf{m} \in \mathbb{F}_3^2$):

(0,0,0),	(1,1,1),	(2,2,2),	(0,1,2),	(1,2,0),	(2,0,1),	(0,2,1),	(1,0,2),	(2,1,0)
0,	1,	2,	X,	X+1,	X+2,	2X,	2X+1,	2X+2

Notice that by definition, the entries in $\{\alpha_1, ..., \alpha_n\}$ are distinct and thus, must have $n \leq q$.

We now turn to some properties of Reed-Solomon codes.

Claim 5.2.1. *RS codes are linear codes.*

Proof. The proof follows from the fact that if $a \in \mathbb{F}_q$ and $f(X), g(X) \in \mathbb{F}_q[X]$ are polynomials of degree $\leq k-1$, then $af(X)$ and $f(X)+g(X)$ are also polynomials of degree $\leq k-1$. In particular, let messages $\mathbf{m}_1$ and $\mathbf{m}_2$ be mapped to $f_{\mathbf{m}_1}(X)$ and $f_{\mathbf{m}_2}(X)$ where $f_{\mathbf{m}_1}(X), f_{\mathbf{m}_2}(X) \in \mathbb{F}_q[X]$ are polynomials of degree at most $k-1$ and because of the mapping defined in (5.1), it is easy to verify that:

$$f_{\mathbf{m}_1}(X) + f_{\mathbf{m}_2}(X) = f_{\mathbf{m}_1+\mathbf{m}_2}(X),$$

and

$$a f_{\mathbf{m}_1}(X) = f_{a\mathbf{m}_1}(X).$$

In other words,

$$RS(\mathbf{m}_1) + RS(\mathbf{m}_2) = RS(\mathbf{m}_1 + \mathbf{m}_2)$$

$$aRS(\mathbf{m}_1) = RS(a\mathbf{m}_1).$$

Therefore RS is a $[n,k]_q$ linear code. □

The second and more interesting claim is the following:

Claim 5.2.2. *RS is a $[n,k,n-k+1]_q$ code. That is, it matches the Singleton bound.*

The claim on the distance follows from the fact that every non-zero polynomial of degree $k-1$ over $\mathbb{F}_q[X]$ has at most $k-1$ (not necessarily distinct) roots, and that if two polynomials agree on more than $k-1$ places then they must be the same polynomial.

Proposition 5.2.3 ("Degree Mantra")**.** *A nonzero polynomial $f(X)$ of degree t over a field $\mathbb{F}_q$ has at most t roots in $\mathbb{F}_q$*

Proof. We will prove the theorem by induction on t. If $t = 0$, we are done. Now, consider $f(X)$ of degree $t > 0$. Let $\alpha \in \mathbb{F}_q$ be a root such that $f(\alpha) = 0$. If no such root α exists, we are done. If there is a root α, then we can write

$$f(X) = (X - \alpha)g(X)$$

where $\deg(g) = \deg(f) - 1$ (i.e. $X - \alpha$ divides $f(X)$). Note that $g(X)$ is non-zero since $f(X)$ is non-zero. This is because by the fundamental rule of division of polynomials:

$$f(X) = (X - \alpha)g(X) + R(X)$$

where $\deg(R) \leq 0$ (as the degree cannot be negative this in turn implies that $\deg(R) = 0$) and since $f(\alpha) = 0$,

$$f(\alpha) = 0 + R(\alpha),$$

which implies that $R(\alpha) = 0$. Since $R(X)$ has degree zero (i.e. it is a constant polynomial), this implies that $R(X) \equiv 0$.

Finally, as $g(X)$ is non-zero and has degree $t-1$, by induction, $g(X)$ has at most $t-1$ roots, which implies that $f(X)$ has at most t roots. □

We are now ready to prove Claim 5.2.2

Proof of Claim 5.2.2. We start by proving the claim on the distance. Fix arbitrary $\mathbf{m}_1 \neq \mathbf{m}_2 \in \mathbb{F}_q^k$. Note that $f_{\mathbf{m}_1}(X), f_{\mathbf{m}_2}(X) \in \mathbb{F}_q[X]$ are distinct polynomials of degree at most $k-1$ since $\mathbf{m}_1 \neq \mathbf{m}_2 \in \mathbb{F}_q^k$. Then $f_{\mathbf{m}_1}(X) - f_{\mathbf{m}_2}(X) \neq 0$ also has degree at most $k-1$. Note that $wt(RS(\mathbf{m}_2) - RS(\mathbf{m}_1)) = \Delta(RS(\mathbf{m}_1), RS(\mathbf{m}_2))$. The weight of $RS(\mathbf{m}_2) - RS(\mathbf{m}_1)$ is n minus the number of zeroes in $RS(\mathbf{m}_2) - RS(\mathbf{m}_1)$, which is equal to n minus the number of roots that $f_{\mathbf{m}_1}(X) - f_{\mathbf{m}_2}(X)$ has among $\{\alpha_1, ..., \alpha_n\}$. That is,

$$\Delta(RS(\mathbf{m}_1), RS(\mathbf{m}_2)) = n - |\{\alpha_i \mid f_{\mathbf{m}_1}(\alpha_i) = f_{\mathbf{m}_2}(\alpha_i)\}|.$$

By Proposition 5.2.3, $f_{\mathbf{m}_1}(X) - f_{\mathbf{m}_2}(X)$ has at most $k-1$ roots. Thus, the weight of $RS(\mathbf{m}_2) - RS(\mathbf{m}_1)$ is at least $n-(k-1) = n-k+1$. Therefore $d \geq n-k+1$, and since the Singleton bound (Theorem 4.3.1) implies that $d \leq n-k+1$, we have $d = n-k+1$.[6] The argument above also shows that distinct polynomials $f_{\mathbf{m}_1}(X), f_{\mathbf{m}_2}(X) \in \mathbb{F}_q[X]$ are mapped to distinct codewords. (This is because the Hamming distance between any two codewords is at least $n-k+1 \geq 1$, where the last inequality follows as $k \leq n$.) Therefore, the code contains q^k codewords and has dimension k. The claim on linearity of the code follows from Claim 5.2.1. □

Recall that the Plotkin bound (Corollary 4.4.2) implies that to achieve the Singleton bound, the alphabet size cannot be a constant. Thus, some dependence of q on n in Reed-Solomon codes is unavoidable.

Let us now find a generator matrix for RS codes (which exists by Claim 5.2.1). By Definition 5.2.1, any basis $f_{\mathbf{m}_1}, ..., f_{\mathbf{m}_k}$ of polynomial of degree at most $k-1$ gives rise to a basis $RS(\mathbf{m}_1), ..., RS(\mathbf{m}_k)$ of the code. A particularly nice polynomial basis is the set of monomials $1, X, ..., X^i, ..., X^{k-1}$. The corresponding generator matrix, whose ith row (numbering rows from 0 to $k-1$) is

$$(\alpha_1^i, \alpha_2^i, ..., \alpha_j^i, ..., \alpha_n^i)$$

and this generator matrix is called the *Vandermonde* matrix of size $k \times n$:

[6] See Exercise 5.2 for an alternate direct argument.

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_j & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_j^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \alpha_1^i & \alpha_2^i & \cdots & \alpha_j^i & \cdots & \alpha_n^i \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & \cdots & \alpha_j^{k-1} & \cdots & \alpha_n^{k-1} \end{pmatrix}$$

The class of codes that match the Singleton bound have their own name, which we define and study next.

5.3 A Property of MDS Codes

Definition 5.3.1 (MDS codes)**.** An $(n,k,d)_q$ code is called *Maximum Distance Separable (MDS)* if $d = n-k+1$.

Thus, Reed-Solomon codes are MDS codes.

Next, we prove an interesting property of an MDS code $C \subseteq \Sigma^n$ with integral dimension k. We begin with the following notation.

Definition 5.3.2. For any subset of indices $S \subseteq [n]$ of size exactly k and a code $C \subseteq \Sigma^n$, C_S is the set of all codewords in C projected onto the indices in S.

MDS codes have the following nice property that we shall prove for the special case of Reed-Solomon codes first and subsequently for the general case as well.

Proposition 5.3.1. *Let $C \subseteq \Sigma^n$ of integral dimension k be an MDS code, then for all $S \subseteq [n]$ such that $|S| = k$, we have $|C_S| = \Sigma^k$.*

Before proving Proposition 5.3.1 in its full generality, we present its proof for the special case of Reed-Solomon codes.
Consider any $S \subseteq [n]$ of size k and fix an arbitrary $\mathbf{v} = (v_1, \ldots, v_k) \in \mathbb{F}_q^k$, we need to show that there exists a codeword $\mathbf{c} \in RS$ (assume that the RS code evaluates polynomials of degree at most $k-1$ over $\alpha_1, \ldots, \alpha_n \subseteq \mathbb{F}_q$) such that $\mathbf{c}_S = \mathbf{v}$. Consider a generic degree $k-1$ polynomial $F(X) = \sum_{i=0}^{k-1} f_i X^i$. Thus, we need to show that there exists $F(X)$ such that $F(\alpha_i) = v_i$ for all $i \in S$, where $|S| = k$.

For notational simplicity, assume that $S = [k]$. We think of f_i's as unknowns in the equations that arise out of the relations $F(\alpha_i) = v_i$. Thus, we need to show that there is a solution to the following system of linear equations:

$$\begin{pmatrix} p_0 & p_1 & \cdots & p_{k-1} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ \alpha_1 & \alpha_i & \alpha_k \\ \alpha_1^2 & \alpha_i^2 & \alpha_k^2 \\ \vdots & \vdots & \vdots \\ \alpha_1^{k-1} & \alpha_i^{k-1} & \alpha_k^{k-1} \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_k \end{pmatrix}$$

The above constraint matrix is a Vandermonde matrix and is known to have full rank (see Exercise 5.6). Hence, by Exercise 2.5, there always exists a unique solution for $(p_0, \ldots, p_{k-1})$. This completes the proof for Reed-Solomon codes.

Next, we prove the property for the general case which is presented below

Proof of Proposition 5.3.1. Consider a $|C| \times n$ matrix where each row represents a codeword in C. Hence, there are $|C| = |\Sigma|^k$ rows in the matrix. The number of columns is equal to the block length n of the code. Since C is Maximum Distance Separable, its distance $d = n - k + 1$.

Let $S \subseteq [n]$ be of size exactly k. It is easy to see that for any $\mathbf{c}^i \neq \mathbf{c}^j \in C$, the corresponding projections $\mathbf{c}^i_S$ and $\mathbf{c}^j_S \in C_S$ are not the same. As otherwise $\triangle(\mathbf{c}^i, \mathbf{c}^j) \leq d-1$, which is not possible as the minimum distance of the code C is d. Therefore, every codeword in C gets mapped to a distinct codeword in C_S. As a result, $|C_S| = |C| = |\Sigma|^k$. As $C_S \subseteq \Sigma^k$, this implies that $C_S = \Sigma^k$, as desired. □

Proposition 5.3.1 implies an important property in pseudorandomness: see Exercise 5.7 for more.

5.4 Exercises

Exercise 5.1. Prove that $X^2 + X + 1$ is the unique irreducible polynomial of degree two over $\mathbb{F}_2$.

Exercise 5.2. For any $[n,k]_q$ Reed-Solomon code, exhibit two codewords that are at Hamming distance exactly $n - k + 1$.

Exercise 5.3. Let $\mathrm{RS}_{\mathbb{F}_q^*}[n,k]$ denote the Reed-Solomon code over $\mathbb{F}_q$ where the evaluation points is $\mathbb{F}_q$ (i.e. $n = q$). Prove that

$$\left(\mathrm{RS}_{\mathbb{F}_q}[n,k]\right)^{\perp} = \mathrm{RS}_{\mathbb{F}_q}[n, n-k],$$

that is, the dual of these Reed-Solomon codes are Reed-Solomon codes themselves. Conclude that Reed-Solomon codes contain self-dual codes (see Exercise 2.29 for a definition).

Hint: Exercise 2.2 might be useful.

Exercise 5.4. Since Reed-Solomon codes are linear codes, by Proposition 2.3.3, one can do error detection for Reed-Solomon codes in quadratic time. In this problem, we will see that one can design even more efficient error detection algorithm for Reed-Solomon codes. In particular, we

will consider *data streaming algorithms* (see Section 20.5 for more motivation on this class of algorithms). A data stream algorithm makes a sequential pass on the input, uses poly-logarithmic space and spend only poly-logarithmic time on each location in the input. In this problem we show that there exists a randomized data stream algorithm to solve the error detection problem for Reed-Solomon codes.

1. Give a randomized data stream algorithm that given as input $\mathbf{y} \in \mathbb{F}_q^m$ decides whether $\mathbf{y} = \mathbf{0}$ with probability at least 2/3. Your algorithm should use $O(\log qm)$ space and polylog(qm) time per position of $\mathbf{y}$. For simplicity, you can assume that given an integer $t \geq 1$ and prime power q, the algorithm has oracle access to an irreducible polynomial of degree t over $\mathbb{F}_q$.
   Hint: Use Reed-Solomon codes.

2. Given $[q,k]_q$ Reed-Solomon code C (i.e. with the evaluation points being $\mathbb{F}_q$), present a data stream algorithm for error detection of C with $O(\log q)$ space and polylog$q$ time per position of the received word. The algorithm should work correctly with probability at least 2/3. You should assume that the data stream algorithm has access to the values of k and q (and knows that C has $\mathbb{F}_q$ as its evaluation points).
   Hint: Part 1 and Exercise 5.3 should be helpful.

Exercise 5.5. We have defined Reed-Solomon in this chapter and Hadamard codes in Section 2.7. In this problem we will prove that certain alternate definitions also suffice.

1. Consider the Reed-Solomon code over a field $\mathbb{F}_q$ and block length $n = q-1$ defined as

$$\mathrm{RS}_{\mathbb{F}_q^*}[n,k,n-k+1] = \{(p(1),p(\alpha),\ldots,p(\alpha^{n-1})) \mid p(X) \in \mathbb{F}[X] \text{ has degree} \leq k-1\}$$

   where α is the generator of the multiplicative group $\mathbb{F}^*$ of $\mathbb{F}$.[7]

   Prove that

$$\begin{aligned}\mathrm{RS}_{\mathbb{F}_q^*}[n,k,n-k+1] &= \{(c_0,c_1,\ldots,c_{n-1}) \in \mathbb{F}^n \mid c(\alpha^\ell) = 0 \text{ for } 1 \leq \ell \leq n-k\,, \\ &\qquad \text{where } c(X) = c_0 + c_1X + \cdots + c_{n-1}X^{n-1}\}\,. \end{aligned} \tag{5.2}$$

   Hint: Exercise 2.2 might be useful.

2. Recall that the $[2^r, r, 2^{r-1}]_2$ Hadamard code is generated by the $r \times 2^r$ matrix whose ith (for $0 \leq i \leq 2^r-1$) column is the binary representation of i. Briefly argue that the Hadamard codeword for the message $(m_1,m_2,\ldots,m_r) \in \{0,1\}^r$ is the evaluation of the (multivariate) polynomial $m_1X_1 + m_2X_2 + \cdots + m_rX_r$ (where $X_1,\ldots,X_r$ are the r variables) over all the possible assignments to the variables $(X_1,\ldots,X_r)$ from $\{0,1\}^r$.

   Using the definition of Hadamard codes above (re)prove the fact that the code has distance 2^{r-1}.

[7] This means that $\mathbb{F}_q^* = \{1,\alpha,\ldots,\alpha^{n-1}\}$. Further, $\alpha^n = 1$.

Exercise 5.6. Prove that the $k \times k$ Vandermonde matrix (where the (i,j)th entry is α_j^i) has full rank (where $\alpha_1, \ldots, \alpha_k$ are distinct).

Exercise 5.7. A set $S \subseteq \mathbb{F}_q^n$ is said to be a t-wise independent source (for some $1 \leq t \leq n$) if given a uniformly random sample $(X_1, \ldots, X_n)$ from S, the n random variables are t-wise independent: i.e. any subset of t variables are uniformly independent random variables over $\mathbb{F}_q$. We will explore properties of these objects in this exercise.

1. Argue that the definition of t-wise independent source is equivalent to the definition in Exercise 2.12.

2. Argue that any $[n,k]_q$ code C is an 1-wise independent source.

3. Prove that any $[n,k]_q$ MDS code is a k-wise independent source.

4. Using part 3 or otherwise prove that there exists a k-wise independent source over $\mathbb{F}_2$ of size at most $(2n)^k$. Conclude that $k(\log_2 n + 1)$ uniformly and independent random bits are enough to compute n random bits that are k-wise independent.

5. For $0 < p \leq 1/2$, we say the n binary random variables $X_1, \ldots, X_n$ are p-biased and t-wise independent if any of the t random variables are independent and $\Pr[X_i = 1] = p$ for every $i \in [n]$. For the rest of the problem, let p be a power of $1/2$. Then show that any $t \cdot \log_2(1/p)$-wise independent random variables can be converted into t-wise independent p-biased random variables. Conclude that one can construct such sources with $k \log_2(1/p)(1 + \log_2 n)$ uniformly random bits. Then improve this bound to $k(1 + \max(\log_2(1/p), \log_2 n))$ uniformly random bits.

Exercise 5.8. In many applications, errors occur in "bursts"– i.e. all the error locations are contained in a contiguous region (think of a scratch on a DVD or disk). In this problem we will use how one can use Reed-Solomon codes to correct bursty errors.

An error vector $\mathbf{e} \in \{0,1\}^n$ is called a *t-single burst* error pattern if all the non-zero bits in $\mathbf{e}$ occur in the range $[i, i+t-1]$ for some $1 \leq i \leq n = t+1$. Further, a vector $\mathbf{e} \in \{0,1\}^n$ is called a (s,t)-burst error pattern if it is the union of at most s t-single burst error pattern (i.e. all non-zero bits in $\mathbf{e}$ are contained in one of at most s contiguous ranges in $[n]$).

We call a binary code $C \subseteq \{0,1\}^n$ to be (s,t)-*burst error correcting* if one can uniquely decode from any (s,t)-burst error pattern. More precisely, given an (s,t)-burst error pattern $\mathbf{e}$ and any codeword $\mathbf{c} \in C$, the only codeword $\mathbf{c}' \in C$ such that $(\mathbf{c} + \mathbf{e}) - \mathbf{c}'$ is an (s,t)-burst error pattern satisfies $\mathbf{c}' = \mathbf{c}$.

1. Argue that if C is (st)-error correcting (in the sense of Definition 1.3.3), then it is also (s,t)-burst error correcting. Conclude that for any $\varepsilon > 0$, there exists code with rate $\Omega(\varepsilon^2)$ and block length n that is (s,t)-burst error correcting for any s,t such that $s \cdot t \leq \left(\frac{1}{4} - \varepsilon\right) \cdot n$.

2. Argue that for any rate $R > 0$ and for large enough n, there exist (s,t)-burst error correcting as long as $s \cdot t \leq \left(\frac{1-R-\varepsilon}{2}\right) \cdot n$ and $t \geq \Omega\left(\frac{\log n}{\varepsilon}\right)$. In particular, one can correct from $\frac{1}{2} - \varepsilon$ fraction of burst-errors (as long as each burst is "long enough") with rate $\Omega(\varepsilon)$ (compare this with

item 1).
Hint: Use Reed-Solomon codes.

Exercise 5.9. In this problem we will look at a very important class of codes called BCH codes[8].
Let $\mathbb{F} = \mathbb{F}_{2^m}$. Consider the binary code C_{BCH} defined as $\text{RS}_{\mathbb{F}}[n, k, n-k+1] \cap \mathbb{F}_2^n$.

1. Prove that C_{BCH} is a binary linear code of distance at least $d = n-k+1$ and dimension at least $n - (d-1)\log_2(n+1)$.

   Hint: Use the characterization (5.2) of the Reed-Solomon code from Exercise 5.5.

2. Prove a better lower bound of $n - \left\lceil \frac{d-1}{2} \right\rceil \log_2(n+1)$ on the dimension of C_{BCH}.

   Hint: Try to find redundant checks amongst the "natural" parity checks defining C_{BCH}).

3. For $d = 3$, C_{BCH} is the same as another code we have seen. What is that code?

4. For constant d (and growing n), prove that C_{BCH} have nearly optimal dimension for distance d, in that the dimension cannot be $n - t\log_2(n+1)$ for $t < \frac{d-1}{2}$.

Exercise 5.10. In this exercise, we continue in the theme of Exercise 5.9 and look at the intersection of a Reed-Solomon code with $\mathbb{F}_2^n$ to get a binary code. Let $\mathbb{F} = \mathbb{F}_{2^m}$. Fix positive integers d, n with $(d-1)m < n < 2^m$, and a set $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ of n distinct nonzero elements of $\mathbb{F}$. For a vector $\mathbf{v} = (v_1, \ldots, v_n) \in (\mathbb{F}^*)^n$ of n not necessarily distinct nonzero elements from $\mathbb{F}$, define the *Generalized Reed-Solomon code* $\text{GRS}_{S,\mathbf{v},d}$ as follows:

$$\text{GRS}_{S,\mathbf{v},d} = \{(v_1 p(\alpha_1), v_2 p(\alpha_2), \ldots, v_n p(\alpha_n)) \mid p(X) \in \mathbb{F}[X] \text{ has degree} \leq n-d\} .$$

1. Prove that $\text{GRS}_{S,\mathbf{v},d}$ is an $[n, n-d+1, d]_{\mathbb{F}}$ linear code.

2. Argue that $\text{GRS}_{S,\mathbf{v},d} \cap \mathbb{F}_2^n$ is a binary linear code of rate at least $1 - \frac{(d-1)m}{n}$.

3. Let $\mathbf{c} \in \mathbb{F}_2^n$ be a nonzero binary vector. Prove that (for every choice of d, S) there are at most $(2^m - 1)^{n-d+1}$ choices of the vector $\mathbf{v}$ for which $\mathbf{c} \in \text{GRS}_{S,\mathbf{v},d}$.

4. Using the above, prove that if the integer D satisfies $\text{Vol}_2(n, D-1) < (2^m - 1)^{d-1}$ (where $\text{Vol}_2(n, D-1) = \sum_{i=0}^{D-1} \binom{n}{i}$), then there exists a vector $\mathbf{v} \in (\mathbb{F}^*)^n$ such that the minimum distance of the binary code $\text{GRS}_{S,\mathbf{v},d} \cap \mathbb{F}_2^n$ is at least D.

5. Using parts 2 and 4 above (or otherwise), argue that the family of codes $\text{GRS}_{S,\mathbf{v},d} \cap \mathbb{F}_2^n$ contains binary linear codes that meet the Gilbert-Varshamov bound.

[8]The acronym BCH stands for Bose-Chaudhuri-Hocquenghem, the discoverers of this family of codes.

Exercise 5.11. In this exercise we will show that the dual of a GRS code is a GRS itself with different parameters. First, we state the obvious definition of GRS codes over a general finite field $\mathbb{F}_q$ (as opposed to the definition over fields of characteristic two in Exercise 5.10). In particular, define the code $\mathrm{GRS}_{S,\mathbf{v},d,q}$ as follows:

$$\mathrm{GRS}_{S,\mathbf{v},d,q} = \{(v_1 p(\alpha_1), v_2 p(\alpha_2), \ldots, v_n p(\alpha_n)) \mid p(X) \in \mathbb{F}_q[X] \text{ has degree} \leq n-d\} .$$

Then show that

$$\left(\mathrm{GRS}_{S,\mathbf{v},d,q}\right)^{\perp} = \mathrm{GRS}_{S,\mathbf{v}',n-d+2,q},$$

where $\mathbf{v}' \in \mathbb{F}_q^n$ is a vector with all non-zero components.

Exercise 5.12. In Exercise 2.15, we saw that any linear code can be converted in to a systematic code. In other words, there is a map to convert Reed-Solomon codes into a systematic one. In this exercise the goal is to come up with an explicit encoding function that results in a systematic Reed-Solomon code.

In particular, given the set of evaluation points $\alpha_1, \ldots, \alpha_n$, design an explicit map f from $\mathbb{F}_q^k$ to a polynomial of degree at most $k-1$ such that the following holds. For every message $\mathbf{m} \in \mathbb{F}_q^k$, if the corresponding polynomial is $f_{\mathbf{m}}(X)$, then the vector $\left(f_{\mathbf{m}}(\alpha_i)\right)_{i\in[n]}$ has the message $\mathbf{m}$ appear in the corresponding codeword (say in its first k positions). Further, argue that this map results in an $[n,k,n-k+1]_q$ code.

Exercise 5.13. In this problem, we will consider the number-theoretic counterpart of Reed-Solomon codes. Let $1 \leq k < n$ be integers and let $p_1 < p_2 < \cdots < p_n$ be n distinct primes. Denote $K = \prod_{i=1}^{k} p_i$ and $N = \prod_{i=1}^{n} p_i$. The notation $\mathbb{Z}_M$ stands for integers modulo M, i.e., the set $\{0,1,\ldots,M-1\}$. Consider the *Chinese Remainder code* defined by the encoding map $E : \mathbb{Z}_K \to \mathbb{Z}_{p_1} \times \mathbb{Z}_{p_2} \times \cdots \times \mathbb{Z}_{p_n}$ defined by:

$$E(m) = (m \mod p_1,\ m \mod p_2,\ \cdots,\ m \mod p_n) .$$

(Note that this is not a code in the usual sense we have been studying since the symbols at different positions belong to different alphabets. Still notions such as distance of this code make sense and are studied in the question below.)

Suppose that $m_1 \neq m_2$. For $1 \leq i \leq n$, define the indicator variable $b_i = 1$ if $E(m_1)_i \neq E(m_2)_i$ and $b_i = 0$ otherwise. Prove that $\prod_{i=1}^{n} p_i^{b_i} > N/K$.

Use the above to deduce that when $m_1 \neq m_2$, the encodings $E(m_1)$ and $E(m_2)$ differ in at least $n-k+1$ locations.

Exercise 5.14. In this problem, we will consider derivatives over a finite field $\mathbb{F}_q$. Unlike the case of derivatives over reals, derivatives over finite fields do not have any physical interpretation but as we shall see shortly, the notion of derivatives over finite fields is still a useful concept. In particular, given a polynomial $f(X) = \sum_{i=0}^{t} f_i X^i$ over $\mathbb{F}_q$, we define its derivative as

$$f'(X) = \sum_{i=0}^{t-1} (i+1) \cdot f_{i+1} \cdot X^i.$$

Further, we will denote by $f^{(i)}(X)$, the result of applying the derivative on f i times. In this problem, we record some useful facts about derivatives.

1. Define $R(X,Z) = f(X+Z) = \sum_{i=0}^{t} r_i(X) \cdot Z^i$. Then for any $j \geq 1$,

$$f^{(j)}(X) = j! \cdot r_j(X).$$

2. Using part 1 or otherwise, show that for any $j \geq \text{char}(\mathbb{F}_q)$,[9] $f^{(j)}(X) \equiv 0$.

3. Let $j \leq \text{char}(\mathbb{F}_q)$. Further, assume that for every $0 \leq i < j$, $f^{(i)}(\alpha) = 0$ for some $\alpha \in \mathbb{F}_q$. Then prove that $(X-\alpha)^j$ divides $f(X)$.

4. Finally, we will prove the following generalization of the degree mantra (Proposition 5.2.3). Let $f(X)$ be a non-zero polynomial of degree t and $m \leq \text{char}(\mathbb{F}_q)$. Then there exists at most $\left\lfloor \frac{t}{m} \right\rfloor$ distinct elements $\alpha \in \mathbb{F}_q$ such that $f^{(j)}(\alpha) = 0$ for every $0 \leq j < m$.

Exercise 5.15. In this exercise, we will consider a code that is related to Reed-Solomon codes and uses derivatives from Exercise 5.14. These codes are called *derivative codes*.

Let $m \geq 1$ be an integer parameter and consider parameters $k > \text{char}(\mathbb{F}_q)$ and n such that $m < k < nm$. Then the derivative code with parameters (n,k,m) is defined as follow. Consider any message $\mathbf{m} \in \mathbb{F}_q^k$ and let $f_{\mathbf{m}}(X)$ be the message polynomial as defined for the Reed-Solomon code. Let $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$ be distinct elements. Then the codeword for $\mathbf{m}$ is given by

$$\begin{pmatrix} f_{\mathbf{m}}(\alpha_1) & f_{\mathbf{m}}(\alpha_2) & \cdots & f_{\mathbf{m}}(\alpha_n) \\ f_{\mathbf{m}}^{(1)}(\alpha_1) & f_{\mathbf{m}}^{(1)}(\alpha_2) & \cdots & f_{\mathbf{m}}^{(1)}(\alpha_n) \\ \vdots & \vdots & \vdots & \vdots \\ f_{\mathbf{m}}^{(m-1)}(\alpha_1) & f_{\mathbf{m}}^{(m-1)}(\alpha_2) & \cdots & f_{\mathbf{m}}^{(m-1)}(\alpha_n) \end{pmatrix}.$$

Prove that the above code is an $\left[n, \frac{k}{m}, n - \left\lfloor \frac{k-1}{m} \right\rfloor\right]_{q^m}$-code (and is thus MDS).

Exercise 5.16. In this exercise, we will consider another code related to Reed-Solomon codes that are called *Folded Reed-Solomon* codes. We will see a lot more of these codes in Chapter 17.

Let $m \geq 1$ be an integer parameter and let $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$ are distinct elements such that for some element $\gamma \in \mathbb{F}_q^*$, the sets

$$\{\alpha_i, \alpha_i\gamma, \alpha_i\gamma^2, \ldots, \alpha_i\gamma^{m-1}\}, \tag{5.3}$$

are pair-wise disjoint for different $i \in [n]$. Then the folded Reed-Solomon code with parameters $(m,k,n,\gamma,\alpha_1,\ldots,\alpha_n)$ is defined as follows. Consider any message $\mathbf{m} \in \mathbb{F}_q^k$ and let $f_{\mathbf{m}}(X)$ be the message polynomial as defined for the Reed-Solomon code. Then the codeword for $\mathbf{m}$ is given by:

$$\begin{pmatrix} f_{\mathbf{m}}(\alpha_1) & f_{\mathbf{m}}(\alpha_2) & \cdots & f_{\mathbf{m}}(\alpha_n) \\ f_{\mathbf{m}}(\alpha_1 \cdot \gamma) & f_{\mathbf{m}}(\alpha_2 \cdot \gamma) & \cdots & f_{\mathbf{m}}(\alpha_n \cdot \gamma) \\ \vdots & \vdots & \vdots & \vdots \\ f_{\mathbf{m}}(\alpha_1 \cdot \gamma^{m-1}) & f_{\mathbf{m}}(\alpha_2 \cdot \gamma^{m-1}) & \cdots & f_{\mathbf{m}}(\alpha_n \cdot \gamma^{m-1}) \end{pmatrix}.$$

Prove that the above code is an $\left[n, \frac{k}{m}, n - \left\lfloor \frac{k-1}{m} \right\rfloor\right]_{q^m}$-code (and is thus, MDS).

[9] $\text{char}(\mathbb{F}_q)$ denotes the *characteristic* of $\mathbb{F}_q$. That is, if $q = p^s$ for some prime p, then $\text{char}(\mathbb{F}_q) = p$. Any natural number i in $\mathbb{F}_q$ is equivalent to $i \mod \text{char}(\mathbb{F}_q)$.

Exercise 5.17. In this problem we will see that Reed-Solomon codes, derivative codes (Exercise 5.15) and folded Reed-Solomon codes (Exercise 5.16) are all essentially special cases of a large family of codes that are based on polynomials. We begin with the definition of these codes.

Let $m \geq 1$ be an integer parameter and define $m < k \leq n$. Further, let $E_1(X), \ldots, E_n(X)$ be n polynomials over $\mathbb{F}_q$, each of degree m. Further, these polynomials pair-wise do not have any non-trivial factors (i.e. $\gcd(E_i(X), E_j(X))$ has degree 0 for every $i \neq j \in [n]$.) Consider any message $\mathbf{m} \in \mathbb{F}_q^k$ and let $f_{\mathbf{m}}(X)$ be the message polynomial as defined for the Reed-Solomon code. Then the codeword for $\mathbf{m}$ is given by:

$$\left(f_{\mathbf{m}}(X) \mod E_1(X), f_{\mathbf{m}}(X) \mod E_2(X), \ldots, f_{\mathbf{m}}(X) \mod E_n(X)\right).$$

In the above we think of $f_{\mathbf{m}}(X) \mod E_i(X)$ as an element of $\mathbb{F}_{q^m}$. In particular, given given a polynomial of degree at most $m-1$, we will consider any bijection between the q^m such polynomials and $\mathbb{F}_{q^m}$. We will first see that this code is MDS and then we will see why it contains Reed-Solomon and related codes as special cases.

1. Prove that the above code is an $\left[n, \frac{k}{m}, n - \left\lfloor \frac{k-1}{m} \right\rfloor\right]_{q^m}$-code (and is thus, MDS).

2. Let $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$ be distinct elements. Define $E_i(X) = X - \alpha_i$. Argue that for this special case the above code (with $m = 1$) is the Reed-Solomon code.

3. Let $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$ be distinct elements. Define $E_i(X) = (X - \alpha_i)^m$. Argue that for this special case the above code is the derivative code (with an appropriate mapping from polynomials of degree at most $m-1$ and $\mathbb{F}_q^m$, where the mapping could be different for each $i \in [n]$ and can depend on $E_i(X)$).

4. Let $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$ be distinct elements and $\gamma \in \mathbb{F}_q^*$ such that (5.3) is satisfied. Define $E_i(X) = \prod_{j=0}^{m-1}(X - \alpha_i \cdot \gamma^j)$. Argue that for this special case the above code is the folded Reed-Solomon code (with an appropriate mapping from polynomials of degree at most $m-1$ and $\mathbb{F}_q^m$, where the mapping could be different for each $i \in [n]$ and can depend on $E_i(X)$).

Exercise 5.18. In this exercise we will develop a sufficient condition to determine the irreducibility of certain polynomials called the *Eisenstein's criterion*.

Let $F(X, Y)$ be a polynomial of $\mathbb{F}_q$. Think of this polynomial as over X with coefficients as polynomials in Y over $\mathbb{F}_q$. Technically, we think of the coefficients as coming from the ring of polynomials in Y over $\mathbb{F}_q$. We will denote the ring of polynomials in Y over $\mathbb{F}_q$ as $\mathbb{F}_q(Y)$ and we will denote the polynomials in X with coefficients from $\mathbb{F}_q(Y)$ as $\mathbb{F}_q(Y)[X]$.

In particular, let

$$F(X, Y) = X^t + f_{t-1}(Y) \cdot X^{t-1} + \cdots + f_0(Y),$$

where each $f_i(Y) \in \mathbb{F}_q(Y)$. Let $P(Y)$ be a *prime* for $\mathbb{F}_q(Y)$ (i.e. $P(Y)$ has degree at least one and if $P(Y)$ divides $A(Y) \cdot B(Y)$ then $P(Y)$ divides at least one of $A(Y)$ or $B(Y)$). If the following conditions hold:

(i) $P(Y)$ divides $f_i(Y)$ for every $0 \le i < t$; but

(ii) $P^2(Y)$ does not divide $f_0(Y)$

then $F(X,Y)$ does not have any non-trivial factors over $\mathbb{F}_q(Y)[X]$ (i.e. all factors have either degree t or 0 in X).

In the rest of the problem, we will prove this result in a sequence of steps:

1. For the sake of contradiction assume that $F(X,Y) = G(X,Y) \cdot H(X,Y)$ where
$$G(X,Y) = \sum_{i=0}^{t_1} g_i(Y) \cdot X^I \text{ and } H(X,Y) = \sum_{i=0}^{t_2} h_i(Y) \cdot X^i,$$
where $0 < t_1, t_2 < t$. Then argue that $P(Y)$ does not divide both of $g_0(Y)$ and $h_0(Y)$.
For the rest of the problem WLOG assume that $P(Y)$ divides $g_0(Y)$ (and hence does not divide $h_0(Y)$).

2. Argue that there exists an i^* such that $P(Y)$ divide $g_i(Y)$ for every $0 \le i < i^*$ but $P(Y)$ does not divide $g_{i^*}(Y)$ (define $g_t(Y) = 1$).

3. Argue that $P(Y)$ does not divide $f_i(Y)$. Conclude that $F(X,Y)$ does not have any non-trivial factors, as desired.

Exercise 5.19. We have mentioned objects called algebraic-geometric (AG) codes, that generalize Reed-Solomon codes and have some amazing properties: see for example, Section 4.6. The objective of this exercise is to construct one such AG code, and establish its rate vs distance trade-off.

Let p be a prime and $q = p^2$. Consider the equation
$$Y^p + Y = X^{p+1} \tag{5.4}$$
over $\mathbb{F}_q$.

1. Prove that there are exactly p^3 solutions in $\mathbb{F}_q \times \mathbb{F}_q$ to (5.4). That is, if $S \subseteq \mathbb{F}_q^2$ is defined as
$$S = \left\{(\alpha,\beta) \in \mathbb{F}_q^2 \mid \beta^p + \beta = \alpha^{p+1}\right\}$$
then $|S| = p^3$.

2. Prove that the polynomial $F(X,Y) = Y^p + Y - X^{p+1}$ is irreducible over $\mathbb{F}_q$.
Hint: Exercise 5.18 could be useful.

3. Let $n = p^3$. Consider the evaluation map $\mathrm{ev} : \mathbb{F}_q[X,Y] \to \mathbb{F}_q^n$ defined by
$$\mathrm{ev}(f) = (f(\alpha,\beta) : (\alpha,\beta) \in S)\ .$$
Argue that if $f \neq 0$ and is not divisible by $Y^p + Y - X^{p+1}$, then $\mathrm{ev}(f)$ has Hamming weight at least $n - \deg(f)(p+1)$, where $\deg(f)$ denotes the *total* degree of f.
Hint: You are allowed to make use of *Bézout's theorem*, which states that if $f, g \in \mathbb{F}_q[X,Y]$ are nonzero polynomials *with no common factors*, then they have at most $\deg(f)\deg(g)$ common zeroes.

4. For an integer parameter $\ell \geq 1$, consider the set $\mathscr{F}_\ell$ of bivariate polynomials

$$\mathscr{F}_\ell = \left\{ f \in \mathbb{F}_q[X, Y] \mid \deg(f) \leq \ell, \deg_X(f) \leq p \right\}$$

where $\deg_X(f)$ denotes the degree of f in X.

Argue that $\mathscr{F}_\ell$ is an $\mathbb{F}_q$-linear space of dimension $(\ell+1)(p+1) - \frac{p(p+1)}{2}$.

5. Consider the code $C \subseteq \mathbb{F}_q^n$ for $n = p^3$ defined by

$$C = \left\{ \mathrm{ev}(f) \mid f \in \mathscr{F}_\ell \right\} .$$

Prove that C is a linear code with minimum distance at least $n - \ell(p+1)$.

6. Deduce a construction of an $[n,k]_q$ code with distance $d \geq n - k + 1 - p(p-1)/2$.

(Note that Reed-Solomon codes have $d = n - k + 1$, whereas these codes are off by $p(p-1)/2$ from the Singleton bound. However they are much longer than Reed-Solomon codes, with a block length of $n = q^{3/2}$, and the deficiency from the Singleton bound is only $o(n)$.)

5.5 Bibliographic Notes

Reed-Solomon codes were invented by Irving Reed and Gus Solomon [60]. Even though Reed-Solomon codes need $q \geq n$, they are used widely in practice. For example, Reed-Solomon codes are used in storage of information in CDs and DVDs. This is because they are robust against burst-errors that come in contiguous manner. In this scenario, a large alphabet is then a good thing since bursty errors will tend to corrupt the entire symbol in $\mathbb{F}_q$ unlike partial errors, e.g. errors over bits. (See Exercise 5.8.)

It is a big open question to present a deterministic algorithm to compute an irreducible polynomial of a given degree with the same time complexity as in Corollary 5.1.3. Such results are known in general if one is happy with polynomial dependence on q instead of $\log q$. See the book by Shoup [64] for more details.