# ALGEBRAIC GEOMETRIC CODES FROM CURVES AND HIGHER DIMENSIONAL VARIETIES

JOSHUA CRUZ

Abstract. Linear codes are important tools in error correction. Although it is extraordinarily easy to construct a linear code, constructing an efficient linear code proves to be a more involved task. Linear codes can be constructed utilizing tools from algebraic geometry. This paper first outlines the construction of Reed-Solomon codes as a motivating example of a good linear code. After some discussion of algebraic geometry, we construct the Goppa codes, which are generalizations of Reed-Solomon codes. Finally, by viewing our developments in algebraic geometry in the language of sheaves and line bundles, we are able to further generalize the notion of a Goppa code to higher dimensional algebraic varieties.

Contents



## 1. Introduction

Error-correcting codes are important objects that facilitate accurate transmission of messages over unreliable or noisy channels. Errors often occur during the transmission of data, causing the received data to contain different information than the original data. Error-correcting codes can detect and correct errors during transmission to ensure that the receiver receives the intended data.

Error-correcting codes (from here on, referred to as codes) are a set of codewords. If a sender wishes to send a message of length $k$, he can first embed his original message into the set of codewords, which can be seen as $n$-tuples over an alphabet $A$. For example, digital data is often transmitted as a sequence of bits such as '0011010′. This can be seen as a codeword of length 7 over the alphabet $A = \mathbb{F}_2$.

As the message is transmitted across a communications channel, random noise may cause the codeword to mutate into a message that is not a codeword. However, when mutated data is received, we can find the original data sent by imposing a metric on the set of codewords and transforming the mutated data into the nearest codeword.

When constructing codes, there are two important parameters worth investigating. First, we must consider how much extraneous data is generated when encoding a message as a codeword. If we wanted to send a message of length 1 but had to send it as a codeword of length 100, we are sending a lot of extraneous information. Such a code would be quite inefficient. Another form of efficiency that is important to consider is the number of errors a code can correct.

This paper details the construction of codes using algebraic varieties. In particular, we will emphasize the construction of codes using algebraic curves and use many of their properties to motivate us towards more general constructions that work for higher dimensional algebraic varieties.

In Section 2, will discuss the general theory of linear codes and generate define the Reed-Solomon codes as a motivating example. Section 3 and 4 analyze general code parameters to view guaranteed bounds on a code's efficiency or inefficiency.

Sections 5 and 6 outline topics in algebraic geometry necessary to construct Goppa codes, which are generalizations of Reed-Solomon codes. These topics include basic notions of algebraic curves, the projective plane, and divisors. Divisors are the main tools that will need to understand our construction of algebraic geometric codes which appear in Section 7.

Section 8 provides a basic overview of sheaves on algebraic varieties to prepare the reader for the discussion of line bundles and divisors in Section 9. This discussion of line bundles allows us to further generalize the discussion of algebraic geometric codes by showing that we can generate a line bundle from a divisor. Section 10 illustrates how we can use this concept of line bundle to construct more general families of higher dimensional algebraic geometric codes.

## 2. Linear Codes

**Definition 2.1.** If $A$ is a field, a subset $C \subseteq A^n$ is called a *code*. Elements of the code are called *codewords*. If $C$ is a vector subspace then it is called a *linear code*.

Throughout this paper, we will set the alphabet $A := \mathbb{F}_q$ where $q$ is a prime power $p^m$.

**Definition 2.2.** If $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in A^n$, the *Hamming distance* from $x$ to $y$ is the number of coordinates of $x$ that differ from $y$. More precisely, the Hamming distance is defined as

$$d(x, y) = |\{i | x_i \neq y_i\}|.$$

We similarly define the *Hamming weight* of a code as $\mathrm{wt}(x) = d(x, (0, 0, \ldots, 0))$.

**Definition 2.3.** Given a linear code $C \subset A^n$, the *parameters* of $C$ are

- $n$ is the *length* of $C$
- $k := dim(C)$ is the *dimension* of $C$
- $d := \min\{d(x, y) | x, y \in C \text{ and } x \neq y\}$ is the *minimum Hamming distance* of $C$

If $C$ is a code over $\mathbb{F}_q$ with parameters $n, k, d$, we say that $C$ is a $[n, k, d]_q$ code.

*Remark* 2.4. If we say that $C$ is an $[n, k, d]_q$ code, this implicitly assumes that $C$ is linear.

**Proposition 2.5.** *For any code $C$, the minimum Hamming distance $d$ is the same as the minimum Hamming weight.*

*Proof.* Suppose $x, y \in C$ satisfy $d(x, y) = d$, where $d(x, y)$ is the Hamming distance between $x$ and $y$. Let $w := \min\{w(z)|z \in C, z \neq 0\}$ Observe that $\mathrm{wt}(x - y) = d(x, y) = d$. Therefore, we have that $d \leq w$. Since each $\mathrm{wt}(x) = d(x, 0)$, we also have that $w \leq d$. Therefore, $d = w$. □

Suppose we have a $[n, k, d]_q$ code $C$. We can correct errors in the following manner: First, fix an embedding $E : \mathbb{F}_q^k \to \mathbb{F}_q^n$ such that $E(\mathbb{F}_q^k) = C$. Given a message $a$ of length $k$, transmit the message $E(a)$ of length $n$. After transmission, it is likely that the received message $E(a)'$ is not the same as $E(a)$ and perhaps is not even in $C$. However, we can take $E(a)'' \in C$ such that the Hamming distance in between $E(a)'$ and $E(a)''$ is minimized and consider $E(a)''$ to be the decoded message.

How good is this method of error correction? We can analyze a code's performance through its parameters. Firstly, we can consider how efficiently the code encodes information. The parameters $n$ and $k$ induce another parameter $n - k$ called the *redundancy* of a code. The sender wants to transmit a message of length $k$ but must instead transmit a message of length $n$. The length of the message is increased from $k$ to $n$, so in the interest of not sending extraneous amounts of information, an efficient code is one with a small redundancy. This condition is equivalent to having a large value of $k$ relative to $n$.

Another way to analyze a code's performance is by the number of errors a code can correct. Let $a$ be a codeword in $C$. Let $B_{\lfloor \frac{d-1}{2} \rfloor}(a)$ be the ball of radius $\lfloor \frac{d-1}{2} \rfloor$ around $a$ where the utilized metric is the Hamming distance. We observe that $B_{\lfloor \frac{d-1}{2} \rfloor}(a) \cap C = \{a\}$. That is, the only element in $C$ within Hamming distance $\lfloor \frac{d-1}{2} \rfloor$ from $a$ is itself. Therefore, we if the communication channel makes at most $\lfloor \frac{d-1}{2} \rfloor$ errors in transmission, we can always decode the original codeword.

Through this, we see that a $[n, k, d]_q$ code is considered good if both $k$ and $d$ are large relative to $n$. However, the Singleton Bound gives a bound on how good $k$ and $d$ can be.

**Theorem 2.6.** *(Singleton Bound) Let $C$ be an $[n, k, d]$ code over $\mathbb{F}_q$. Then*

$$d \leq n - k + 1.$$

*Proof.* Consider the following subspace $S$.

$$S = \{x = (x_1, \ldots, x_n)|x_d = x_{d+1} = \cdots = x_n = 0\} \subseteq \mathbb{F}_q^n$$

The subspace $S$ has dimension $d - 1$. Additionally if $x \in S$ then $\mathrm{wt}(x) \leq d - 1$. Since $d$ is the minimum Hamming weight of an element in $C$, then $C \cap S = \{0\}$. Therefore, $\dim(C \cup S) = \dim(C) + \dim(S) = k + d - 1$. Since $\dim(C \cup S) \leq n$, we conclude that

$$k + d - 1 \leq n \implies d \leq n - k + 1.$$

□

Now we will construct a linear code that satisfies equality in the singleton bound. First, we will define the space $L_r := \{f \in \mathbb{F}_q[x] : \deg(f) \leq r\} \cup \{0\}$. It is easy to verify that $L_r$ is actually a vector space with dimension $r+1$.

**Definition 2.7.** If we label the nonzero elements of $\mathbb{F}_q$ as $\alpha_1, \ldots, \alpha_{q-1}$ and fix some $k \in \mathbb{Z}$ with $1 \leq k \leq q-1$, we define the *Reed-Solomon Code* $RS(k,q)$ to be

$$RS(k,q) := \{(f(\alpha_1), \ldots, f(\alpha_{q-1}) | f \in L_{k-1}\}.$$

**Proposition 2.8.** *The Reed-Solomon Code $RS(k,q)$ satisfies equality in the Singleton Bound.*

*Proof.* We can define the evaluation mapping as

$$\begin{aligned} \mathrm{ev}_{\mathbb{F}_q} : L_{k-1} &\to \mathbb{F}_q^{q-1} \\ f &\to (f(\alpha_1), \ldots, f(\alpha_{q-1})). \end{aligned}$$

By defining this evaluation mapping, we see that $RS(k,q) = \mathrm{ev}_{\mathbb{F}_q}(L_{k-1})$. Now, let us evaluate the parameters of $RS(k,q)$. Clearly, $n = q-1$. To calculate $k$, first we note that $\dim(RS(k,q)) \leq \dim(L_{k-1}) = k$. Now, we prove that $\mathrm{ev}_{\mathbb{F}_q}$ is injective. If $\mathrm{ev}_{\mathbb{F}_q}(f) = \mathrm{ev}_{\mathbb{F}_q}(g)$, then the polynomial $h = f - g$ has roots at $\alpha_1, \ldots, a_{q-1}$. Therefore, $h$ has at least $q-1$ roots. However, since $\deg(f), \deg(g) \leq k-1$, we have that $\deg(h) \leq k-1$. Since we set $k \leq q-1$, if $h$ is not the zero polynomial then $h$ has at most $q-2$ roots. Since $h$ has at least $q-1$ roots, we must have that $h = 0$ and thus $f = g$. Since $\mathrm{ev}_{\mathbb{F}_q}$ is injective, we get that $\dim(RS(k,q)) = \dim(L_{k-1}) = k$.

To find bounds on $d$, suppose that $f \in L_{k-1}$ is such that its evaluation mapping has minimum Hamming weight in $RS(k,q)$. That is $\mathrm{wt}(\mathrm{ev}_{k,q}(f)) = d$. Therefore, $f$ has at least $q-1-d = n-d$ zeros. So $f$ has degree at least $n-d$. Since $\deg(f) \leq k-1$, we get that $k-1 \geq n-d$ which is equivalent to $d \geq n-k+1$. However, by the Singleton Bound, $d \leq n-k+1$. Therefore, $d = n-k+1$. □

Proposition 2.8 shows that for linear codes of length $n = q-1$ and dimension $k$, the Reed-Solomon Code is the best we can do. However, a limitation of the Reed-Solomon codes is that their length is small with respect to the size of our alphabet $\mathbb{F}_q$. If we have an alphabet with $q$ letters, we can only transmit messages of length $q-1$. Therefore, we would like to construct codes that may not be optimal with respect to the Singleton Bound but enable the transmission of long messages while also efficiently encoding information and correcting errors.

## 3. Bounds on Code Parameters

**Definition 3.1.** Let $q$ be a prime power and let $n, d$ be positive integers with $d \leq n$. Then we define $A_q(n,d)$ to be the maximum value of $M$ such that there is a code $C$ over $\mathbb{F}_q$ containing $M$ codewords with length $n$ and minimum distance $d$.

*Remark* 3.2. Observe that $A_q(n,d)$ neither accounts for the dimension of $C$ nor imposes that the code $C$ with $|C| = A_q(n,d)$ even be linear. So although we may get a code with many codewords, it may have a large amount of redundancy or might not be linear.

**Theorem 3.3.** *(Plotkin Bound) Let $n, d \in \mathbb{Z}$ and $q$ be a prime power. If $d < \left(1 - \frac{1}{q}\right) n$ then $A_q(n, d) = 0$ If $d > \left(1 - \frac{1}{q}\right) n$ then*

$$A_q(n, d) \leq \frac{d}{d - \left(1 - \frac{1}{q}\right) n}.$$

*Proof.* Let $C$ be a code with $M$ codewords, length $n$, and minimum Hamming distance $d$. Define the quantity $S = \sum d(x, y)$ where the sum is taken over ordered pairs of distinct $x, y \in C$. Since $d(x, y) \geq d$ and since there are $M(M-1)$ ordered pairs, we have that $S \geq M(M-1)d$.

Now we attempt to derive an upper bound for $S$. Consider the $M \times n$ matrix that has a codeword of $C$ in each row. Consider any one column, and define $m_\alpha$ to be the number of times that $\alpha$ appears in this column. Observe that $\sum_{\alpha \in \mathbb{F}_q} m_\alpha = M$. Since there are $M - m_\alpha$ codewords without $\alpha$ in the $i$th position, if we assume that we are looking at the column where codewords differ the most, we get that

$$\begin{aligned} S &\leq n \sum_{\alpha \in \mathbb{F}_q} m_\alpha (M - m_\alpha) \\ &= n(M^2 - \sum_{\alpha \in \mathbb{F}_q} m_\alpha^2). \end{aligned}$$

Using Cauchy-Schwarz inequality:

$$\begin{aligned} \sum_{\alpha \in \mathbb{F}_q} m_\alpha &\leq \left( \sum_{\alpha \in \mathbb{F}_q} m_\alpha^2 \right)^{\frac{1}{2}} \sqrt{q} \\ \frac{1}{q} \left( \sum_{\alpha \in \mathbb{F}_q} m_\alpha \right)^2 &\leq \sum_{\alpha \in \mathbb{F}_q} m_\alpha^2. \end{aligned}$$

Therefore

$$\begin{aligned} S &\leq n \left( M^2 - \frac{1}{q} \left( \sum_{\alpha \in \mathbb{F}_q} m_\alpha \right)^2 \right) \\ &= nM^2 \left(1 - \frac{1}{q}\right), \\ dM(M-1) &\leq nM^2 \left(1 - \frac{1}{q}\right), \\ M \left( d - n \left(1 - \frac{1}{q}\right) \right) &\leq d, \text{ and} \\ M &\leq \frac{d}{d - \left(1 - \frac{1}{q}\right) n} \end{aligned}$$

Therefore, $A_q(n, d) \leq \frac{d}{d - \left(1 - \frac{1}{q}\right) n}$ as required. □

Now we look for a lower bound on $A_q(n, d)$. First we observe that for any $x \in \mathbb{F}_q^n$, the number of elements in $B_r(x)$ (using the Hamming metric) is completely

independent of $x$. Given some $x = (x_1, \ldots, x_n)$, we have that $y \in B_r(x)$ if the entries of $y$ differ from the entries of $x$ by at most $r$ elements. We can choose up to $\binom{n}{r}$ entries to change. For each entry that we choose to change, we have $q-1$ different choices. Therefore, the number of elements at exactly distance $i$ from $x$ is $(q-1)^i\binom{n}{i}$. Therefore, we have that

$$|B_r(x)| = \sum_{i=0}^{r} \binom{n}{i}(q-1)^i.$$

We define the number $V_q(n,r)$ as the number of elements in $B_r(x)$ if $x \in \mathbb{F}_q^n$. That is,

$$V_q(n,r) := |B_r(x)| = \sum_{i=0}^{r} \binom{n}{i}(q-1)^i.$$

We can use this to find a lower bound of $A_q(n,d)$.

**Theorem 3.4.** *(Gilbert-Varshamov Bound) Let $n, d \in \mathbb{Z}$ and $q$ be a prime power. Then*

$$A_q(n,d) \geq \frac{q^n}{V_q(n,d-1)}.$$

*Proof.* Suppose $C$ is a code with $|C| = A_q(n,d)$. Let $y \in \mathbb{F}_q^n$. Suppose that for all $x \in C$, $d(x,y) \geq d$. Then $C \cup \{y\}$ is a code of length $n$ and minimum distance $d$. But $|C \cup \{y\}| = A_q(n,d) + 1$, a contradiction since $A_q(n,d)$ is the maximum number of codewords in a code of length $d$ and minimum distance $d$. Therefore for all $y \in \mathbb{F}_q^n$ there exists some $x \in C$ such that $y \in B_{d-1}(x)$.

Therefore, $\bigcup_{x \in C} B_{d-1}(x) = \mathbb{F}_q^n$. Since $|B_{d-1}| = V_q(n, d-1)$ and $|\mathbb{F}_q^n| = q^n$, we have that

$$q^n \leq A_q(n,d)V_q(n,d-1).$$

Rearranging, we get

$$A_q(n,d) \geq \frac{q^n}{V_q(n,d-1)},$$

as required. □

*Remark* 3.5. Note that a $k$-dimensional subspace of $\mathbb{F}_q^n$ contains $q^k$ elements. Therefore, $A_q(n,d)$ implicitly gives an upper bound on the dimension of a linear code. However, it does not necessarily give a lower bound on the dimension of a linear code, since $A_q(n,d)$ ranges over all codes of length $n$ and minimum distance $d$, regardless of whether the code is linear or not.

## 4. Asymptotic Bounds on Code Parameters

As discussed in Section 2, we would like our codes to be "large" with respect to $n$. This corresponds to the number of codewords in $C$, and in the case of a linear code, this also corresponds to the dimension of the code. Since our notion of "largeness" is with respect to the length of the code, we normalize our parameters with respect to $n$.

**Definition 4.1.** Let $C$ be a code over $\mathbb{F}_q$ with length $n$, minimum distance $d$, and $q^k$ codewords (if $C$ is non-linear, then $k$ is not necessarily an integer). We define

- The *information rate* of $C$ is $R := \frac{k}{n}$.
- The *relative minimum distance* of $C$ is $\delta := \frac{d}{n}$.

Suppose that we want to send messages of length $k$ utilizing a code of length $n$. The information rate quantifies how fast we can send multiple messages of length $k$ when our corresponding code is of length $n$. If our code has zero redundancy ($k = n$), then our information rate is 1 because we can send out information at the same rate as we send out codes. This idea provides further justification as to why we wish $k$ is large with respect to $n$. With these new normalized parameters, we can find some asymptotic bounds.

**Definition 4.2.** Suppose $q$ is a prime power and $0 \leq \delta \leq 1$. We define

$$\alpha_q(\delta) := \limsup_{n\to\infty} \frac{1}{n} \log_q A_q(n, \lfloor \delta n \rfloor).$$

Let us explore what this quantity gives us. Suppose that we define $\tilde{A}_q(n, d)$ to be equal to the maximum number value of $M$ such that there is a *linear* code $C$ over $\mathbb{F}_q$ containing $M$ codewords with length $n$ and minimum distance $d$. Similarly, let us define

$$\tilde{\alpha}_q(\delta) := \limsup_{n\to\infty} \frac{1}{n} \log_q \tilde{A}_q(n, \lfloor \delta n \rfloor).$$

If $\delta = \frac{d}{n}$ for $d \in \mathbb{Z}$, for some $k_n \in \mathbb{Z}$, we have that

$$\begin{aligned}
\tilde{\alpha}_q(\delta) &= \limsup_{n\to\infty} \frac{1}{n} \log_q \tilde{A}_q(n, d) \\
&= \limsup_{n\to\infty} \frac{\log_q q^{k_n}}{n} \\
&= \limsup_{n\to\infty} \frac{k_n}{n}.
\end{aligned}$$

Thus, $\tilde{\alpha}_q(\delta)$ gives us a limit of information rates for linear codes with length $n$ and minimum distance $d$. In other words, there exists a sequence of codes $\{C_n\}$ with length $n$ and minimum distance $d_n$ such that their information rates converge to $\tilde{\alpha}_q(\delta)$. If we lift the constraint that the codes are linear, $\alpha_q(\delta)$ gives us the largest value of $R$ such that there exist codes of length $n$ with relative minimum distance converging to $\delta$ and information rates converging to $R$.

We now will state and prove an asymptotic version of the Plotkin bound.

**Theorem 4.3.** *(Asymptotic Plotkin Bound) Let $q$ be a prime power. Then we have*

$$\begin{cases} \alpha_q(\delta) \leq 1 - \frac{\delta}{1-\frac{1}{q}} & \text{if } 0 \leq \delta \leq 1 - \frac{1}{q} \\ \alpha_q(\delta) = 0 & \text{if } 1 - \frac{1}{q} < \delta \leq 1 \end{cases}.$$

*Proof.* First let $1 - \frac{1}{q} < \delta \leq 1$. We apply the Plotkin bound on $A_q(n, \lfloor \delta n \rfloor)$.

$$\begin{aligned}
\alpha_q(\delta) &= \limsup_{n\to\infty} \frac{1}{n} \log_q A_q(n, \lfloor \delta n \rfloor) \\
&\leq \limsup_{n\to\infty} \frac{1}{n} \log_q \left( \frac{\lfloor \delta n \rfloor}{\lfloor \delta n \rfloor - \left(1 - \frac{1}{q}\right) n} \right) \\
&= \limsup_{n\to\infty} \frac{1}{n} \log_q \left( \frac{\delta}{\delta - \left(1 - \frac{1}{q}\right)} \right) \\
&= 0.
\end{aligned}$$

Now let $0 \leq \delta \leq 1 - \frac{1}{q}$. We observe the following: Suppose $C$ is a code over $\mathbb{F}_q$ of length $n$, $M$ codewords, and minimum distance $d$. Using the Pigeonhole Principle, we get that there exists a subset $S \subseteq C$ with $\frac{M}{q}$ codewords ending with the same element of $\mathbb{F}_q$. We can find a new code $C'$ in the following way by taking the elements of $S$ and deleting the final entry. $C'$ is a code of length $n-1$ and at least $\frac{M}{q}$ codewords. Additionally, since all elements of $S$ had the same last coordinate, $C'$ does not decrease the minimum distance and so the minimum distance of $C'$ is at least $d$.

Let $C$ be a code of length $n$ and minimum distance $\lfloor \delta n \rfloor$. Define $n' := \lfloor \frac{d-1}{1-\frac{1}{q}} \rfloor$. Use this shortening algorithm on $C$ $n - n'$ times to get a code of length $n'$ with $M' \geq \frac{M}{q^{n-n'}}$ codewords. We use the Plotkin bound to obtain the following bound:

$$\frac{M}{q^{n-n'}} \leq M' \leq A_q(n - n', \lfloor \delta n \rfloor) \leq \frac{\lfloor \delta n \rfloor}{\lfloor \delta n \rfloor - \left(1 - \frac{1}{q}\right) n'} \leq \lfloor \delta n \rfloor.$$

Therefore,

$$M \leq \lfloor \delta n \rfloor q^{n-n'}.$$

Thus, we get

$$\begin{aligned}
\alpha_q(\delta) &= \limsup_{n\to\infty} \frac{1}{n} \log_q A_q(n, \lfloor \delta n \rfloor) \\
&\leq \limsup_{n\to\infty} \frac{1}{n} \log_q \left( \lfloor \delta n \rfloor q^{n-n'} \right) \\
&= \limsup_{n\to\infty} \frac{\log_q \lfloor \delta n \rfloor}{n} + 1 - \frac{n'}{n} \\
&= \limsup_{n\to\infty} \left( 1 - \frac{1}{n} \left\lfloor \frac{\lfloor \delta n \rfloor - 1}{1 - \frac{1}{q}} \right\rfloor \right) \\
&= 1 - \frac{\delta}{1 - \frac{1}{q}}.
\end{aligned}$$

as required. □

The Asymptotic Plotkin Bound shows that if we have a sequence of codes with large relative minimum distances, the information rate tends to 0. If the relative minimum distance is bounded above by $1 - \frac{1}{q}$, we get a bound on the information rate dependent on $\delta$ and $q$.

## 5. A Brief Introduction to Algebraic Geometry

In the following sections, we will discuss some basic constructions in algebraic geometry necessary to build algebraic geometric codes. Basic properties of these objects will be presented without proof for the sake of keeping this paper as brief as possible. Examples will be provided to give basic intuition. The goal of this section is to provide the basic framework needed to construct algebraic geometric codes and the proper setting within which our relevant theorems and results are true.

**Notation 5.1.** In the following sections on algebraic geometry, we will use $k$ to denote a field. Although we used $k$ as a parameter in our discussion about codes, $k$ is quite commonly used to denote fields in literature on algebraic geometry.

**Definition 5.2.** Let $k$ be a field and $f_1, \ldots, f_m \in k[x_1, \ldots, x_n]$ be polynomials. An *affine algebraic set* is one of the form

$$V(f_1, \ldots, f_m) = \{(a_1, \ldots, a_n) \in k^n | f_i(a_1, \ldots, a_n) = 0 \text{ for } 1 \leq i \leq m\}.$$

*Remark* 5.3. Notice that $V(f_1, \ldots, f_m) = \bigcap_{i=1}^{m} V(f_i)$.

**Definition 5.4.** An affine algebraic set $X$ is *reducible* if there exist distinct nonempty affine algebraic sets $X_1 \neq X \neq X_2$ and $X = X_1 \cup X_2$.

**Definition 5.5.** An irreducible affine algebraic set $X$ is called an *affine algebraic variety*.

From now on, we will omit the term 'algebraic' and simply say 'affine set' or 'affine variety' when understood.

**Example 5.6.** $V(xy)$ is an affine set over $\mathbb{R}$ that is not an affine variety because we can write $V(xy) = V(x) \cup V(y)$.

**Example 5.7.** $V(x^2 + y^2 - 1)$ is an affine variety over $\mathbb{R}$.

The variety in Example 5.7 is the unit circle in $\mathbb{R}^2$ and one can intuitively see that it is irreducible. However, it is important to formalize a more general way to understand irreducibility, even in cases where we cannot visualize the variety. As expected, we will have to connect the variety to the polynomials that generate it. Given an algebraic variety $X$, we can generate an ideal of $k[x_1, \ldots, x_n]$ defined as

$$I(X) := \{f \in k[x_1, \ldots, x_n] : f(x) = 0 \text{ for all } x \in X\}$$

By checking addition of polynomials in $I(X)$ and multiplication by any other polynomial in $k[x_1, \ldots, x_n]$, one can verify that $I(X)$ is an ideal of $k[x_1, \ldots, x_n]$. Therefore, we have a method by which we can go from varieties to ideals. However, our method of going from ideals to varieties is not as clear. For example, consider the varieties $V(x)$ and $V(x^2)$. We have that $V(x) = V(x^2) = \{0\}$. So when we try to translate the variety $\{0\}$ into an ideal, we do not have a unique representation. This problem is fixed using radical ideals.

**Definition 5.8.** Let $J$ be an ideal of a ring $R$. Its *radical* $\sqrt{J}$ is the ideal such that $x \in \sqrt{J}$ if and only if $x^m \in J$ for some integer $m$.

**Theorem 5.9.** *(Hilbert's Nullstellensatz) Let $k$ be an algebraically closed field. If $p \in k[x_1, \ldots, x_n]$ that vanishes on $V(J)$ for some ideal $J \subseteq k[x_1, \ldots, x_n]$, then there exists some $m$ such that $p^m \in J$.*

In other words, if $k$ is an algebraically closed field, we have that $I(V(J)) = \sqrt{J}$. Now, our correspondence between ideals and varieties over algebraically closed fields is unambiguous. In the following sections on algebraic geometry, we will assume that $k$ is algebraically closed. When we return to our discussion on $\mathbb{F}_q$, we will take its algebraic closure $\overline{\mathbb{F}}_q$ (in fact, we will take its *projective closure*

**Example 5.10.** $I(V(x^2)) = (x)$

Radical ideals give us a method of distinguishing between algebraic sets and algebraic varieties.

**Proposition 5.11.** *An algebraic set $V(J)$ is an algebraic variety if and only if $\sqrt{J}$ is a prime ideal.*

*Proof.* $\Rightarrow$ Suppose that $V(J)$ is an irreducible variety. Suppose for contradiction that $ab \in \sqrt{J}$ but $a, b \notin \sqrt{J}$. Consider the following varieties $V_1 = V(\sqrt{J} + \langle a \rangle)$, $V_2 = V(\sqrt{J} + \langle b \rangle)$. Since $J \subseteq \sqrt{J} \subseteq \sqrt{J} + \langle a \rangle$, we have that $V(J) \supseteq V(\sqrt{J} + \langle a \rangle) = V_1$. Similarly, we have $V_2 \subseteq V(J)$. However, since $ab \in \sqrt{J}$, we have that if $x \in V(J)$ then $x \in V_1$ or $x \in V_2$. Therefore $V(J) = V_1 \cup V_2$, contradicting that $V(J)$ is irreducible.

$\Leftarrow$ Suppose that $\sqrt{J}$ is a prime ideal. Suppose for contradiction that we can write $V(J) = V_1 \cup V_2$. We can take some $a \in I(V_1) \setminus I(V_2)$ and $b \in I(V_2) \setminus I(V_1)$. We have that $ab \in \sqrt{J}$, but $a, b \notin \sqrt{J}$, contradicting that $\sqrt{J}$ is prime. □

Let us assume that all our algebraic sets are algebraic varieties. Now, let us focus on a particular type of variety called an affine algebraic curve.

**Definition 5.12.** An *affine algebraic curve* $C_f$ is a variety of the form $C_f = V(f)$ where $f \in k[x, y]$.

In other words, an algebraic curve is the zero set of one polynomial in two variables. Given two algebraic curves, we may be interested in where these curves intersect. For example, consider the polynomials $f(x, y) = y - x^2$ and $g(x, y) = y - c$. If $c < 0$, over $\mathbb{R}$ we get no points of intersection. By taking $\mathbb{C}$, the algebraic closure of $\mathbb{R}$, we get two points of intersection at $(\sqrt{|c|}i, c)$ and $(-\sqrt{|c|}i, c)$. If $c = 0$, it seems that we have one point of intersection $(0, 0)$, but since $C_f$ and $C_g$ lie tangent to one another, it is as if this point of intersection has multiplicity 2.

However, now consider $f(x, y) = y - x^2$ and $h(x, y) = x - c$. Regardless of whether we take these polynomials to be over $\mathbb{R}$ or to be over $\mathbb{C}$, we still get only one point of intersection with multiplicity 1. By considering the intersection of $C_f$ and $C_g$, we should expect that $C_f$ and $C_h$ intersect at 2 points or at one point of multiplicity 1. The algebraic closure does not tell the whole story, but the projective plane does.

**Definition 5.13.** Let $k$ be a field. The *projective plane* $\mathbb{P}^2(k)$ is defined as

$$\mathbb{P}^2(k) = \{k^3 \setminus \{(0, 0, 0)\}\} / \sim .$$

where $(X_0 : Y_0 : Z_0) \sim (X_1 : Y_1 : Z_1)$ if there exists some $\lambda \in k^\times$ such that $X_0 = \lambda X_1, Y_0 = \lambda Y_1$, and $Z_0 = \lambda Z_1$.

**Definition 5.14.** Let $P = (X : Y : Z) \in \mathbb{P}^2(k)$. If $Z \neq 0$, then $P$ is called an *affine point*. If $Z = 0$, then $P$ is called a *point at infinity*.

To gain some intuition about the projective plane, we will look at $\mathbb{P}^2(\mathbb{R})$. Suppose that $(X : Y : Z)$ is an affine point. For every affine point, we can associate the point $(\frac{X}{Z} : \frac{Y}{Z} : 1)$ since $(X : Y : Z) \sim (\frac{X}{Z} : \frac{Y}{Z} : 1)$. Therefore, equivalence classes of affine points in $\mathbb{P}^2(\mathbb{R})$ are lines in $\mathbb{R}^3$ that go through the origin and intersect different points on the hyperplane $S = \{(x, y, 1)\}$. Therefore, affine points are considered equivalent if their projection onto $S$ by a line through the origin is the same.

Now let us consider points at infinity. First suppose that we have an affine point $(\frac{X}{Z} : \frac{Y}{Z} : 1)$. As $Z$ approaches 0, the quantities $\frac{X}{Z}$ and $\frac{Y}{Z}$ approach infinity. Suppose $L$ is the line goes through the origin and passes through $(\frac{X}{Z}, \frac{Y}{Z}, 1)$. Notice that as $Z$ approaches 0, the angle that $L$ makes with the $x$-$y$ plane also approaches 0. Thus, points at infinity $(X : Y : 0)$ correspond to lines that go through the origin that lie completely in the $x$-$y$ plane.

**Definition 5.15.** Suppose $f \in k[x, y]$ is of degree $d$. The *homogenization* of $f$ is the polynomial $F \in k[X, Y, Z]$ defined as

$$F(X, Y, Z) := Z^d f\left(\frac{X}{Z}, \frac{Y}{Z}\right).$$

**Definition 5.16.** Suppose $f \in k[x, y]$ of degree $d$, with homogenization $F \in k[X, Y, Z]$. Let $C_f$ be the affine algebraic curve associated with $f$. The *projective closure* $\widehat{C_f}$ is defined as

$$\widehat{C_f} := \{(X_0 : Y_0 : Z_0) \in \mathbb{P}^2 | F(X_0 : Y_0 : Z_0) = 0\}$$

We make the following observations:

- Every term in the homogenization $F$ has degree $d$.
- $f(x, y) = 0$ if and only if $F(x : y : 1) = 0$.
- For $\alpha \in k^\times$, we have that $F(\alpha X, \alpha Y, \alpha Z) = (\alpha Z)^d f\left(\frac{\alpha X}{\alpha Z}, \frac{\alpha Y}{\alpha Z}\right) = \alpha^d F(X, Y, Z)$. Therefore, if $F(X, Y, Z) = 0$, $F(X_0, Y_0, Z_0) = 0$ for all $(X_0, Y_0, Z_0) \in (X : Y : Z)$

*Remark* 5.17. Now that we have figured out how to define polynomials in the projective plane, we can define projective sets, projective varieties, and projective curves in the same way we defined affine sets, varieties, and curves.

Let us reconsider our example $f(x, y) = y - x^2$ and $h(x, y) = x - c$. We know that $(c, c^2)$ is a point of intersection between $C_f$ and $C_h$, but we would like to find another point at infinity. We start by homogenizing to $F(X, Y, Z) = YZ - X^2$ and $H(X, Y, Z) = X - CZ$. We have already found one affine point of intersection between $\widehat{C_f}$ and $\widehat{C_h}$, namely $(c : c^2 : 1)$. Now we search for another point of intersection at infinity. We have that $F(X, Y, 0) = -X^2$ and $H(X, Y, 0) = X$. Setting $X = 0$, we have that the projective point $(0 : 1 : 0)$ is a point of intersection between $\widehat{C_f}$ and $\widehat{C_h}$.

Our discussion on the projective plane leads us to the following important theorem.

**Theorem 5.18.** *(Bezout's Theorem) Let $f, g \in k[x, y]$ be polynomials with degrees $d_f$ and $d_g$ and no non-constant common factors. Then $\widehat{C_f}$ and $\widehat{C_g}$ intersect at exactly $d_f d_g$ points in $\mathbb{P}^2(k)$ counting multiplicity.*

## 6. Functions and Divisors on a Curve

In this section, I will discuss the field of rational functions on a curve to develop the notion of a divisors on curves, which are central objects in constructing algebraic geometric codes.

**Definition 6.1.** Let $k$ be a field and $C$ be a projective plane curve with $C = V(F)$ for some homogeneous polynomial $F(X, Y, Z) \in k[X, Y, Z]$. If $K$ is a field containing $k$, we call $(X : Y : Z) \in \mathbb{P}^2(K)$ satisfying $F(X, Y, Z) = 0$ a *$K$-rational point* on $C$. The set of $K$ rational points on $C$ is denoted $C(K)$.

**Example 6.2.** Consider the curve defined by $C = V(X^2+Y^2-Z^2)$. We have that $(3 : 4 : 5) \in C(\mathbb{R}) \subset C(\mathbb{C})$. However, $(3 : 2i : 5) \in C(\mathbb{C})$ but $(3 : 2i : 5) \notin C(\mathbb{R})$.

**Definition 6.3.** Let $C$ be a projective plane curve defined by some homogeneous polynomial $F(X, Y, Z) \in k[X, Y, Z]$. The *field of rational functions on $C$* is defined as

$$k(C) := \left( \left\{ \frac{g(X,Y,Z)}{h(X,Y,Z)} \middle| g, h \in k[X, Y, Z] \text{ homogeneous of the same degree} \right\} \cup \{0\} \right) / \sim$$

where $\frac{g}{h} \sim \frac{g'}{h'}$ if $gh' - g'h \in (F)$.

**Example 6.4.** Let $f = x - y$ We have that $\frac{X^2}{YZ}$ and $\frac{X}{Z}$ are in the field of rational functions on $\widehat{C_f}$. Additionally, we have that $\frac{X^2}{YZ} \sim \frac{X}{Z}$ over $k(\widehat{C_f})$ because $X^2Z - XYZ = XZ(X - Y) \in (X - Y)$. Intuitively, this means that $X^2Z - XYZ = 0$ when restricted to the curve $X - Y = 0$.

**Definition 6.5.** Let $C$ be a projective plane curve defined over $k$. A *divisor* on $D$ is an element of the free abelian group of points on the curve $C$. That is, every divisor $D$ can be written as the formal sum of finitely many points on $C$ as $D = \sum_{P \in C} n_P P$ where $n_P \in \mathbb{Z}$. The *support* of a divisor $D$ are points $P$ in the formal sum representation of $D$ such that $n_P \neq 0$. Finally, a divisor is *effective* if $n_P \geq 0$ for all $P$ and is written as $D \geq 0$.

This formal sum construction is just one way to understand divisors. We will explore some other perspectives on divisors in later sections, but this construction will suffice for now.

We can also construct divisors out of rational functions.

**Definition 6.6.** Let $C$ be a projective plane curve defined over $k$ and let $f = \frac{g}{h} \in k(C)$. The *divisor of $f$* is defined to be $\operatorname{div}(f) = \sum P - \sum Q$ where each $P$ is a point on the intersection $C \cap V(g)$ and each $Q$ is a point on the intersection $C \cap V(h)$. Divisors of the form $\operatorname{div}(f)$ are called *principal divisors*.

Intuitively, $\operatorname{div}(f)$ "adds the zeros of $f$" and "subtracts the poles of $f$," counted with multiplicity. Since $g$ and $h$ have the same degree, $f$ has the same number of zeros as poles. Thus, the sum of the coefficients of $\operatorname{div}(f)$ is 0. Additionally, if $C$ is the zero set of some homogeneous polynomial $F \in k[X, Y, Z]$ with degree $d$ and $g, h \in k(C)$ each have degree $e$, by Bezout's theorem we have that $|C \cap V(g)| = |C \cap V(h)| = de$.

**Definition 6.7.** Let $D$ be a divisor over a projective plane curve $C$ over $\mathbb{F}_q$. The *space of rational functions associated to $D$* is defined as

$$L(D) := \{f \in \mathbb{F}_q(C) | \operatorname{div}(f) + D \geq 0\}$$

If $D$ is a divsior and $f \in L(D)$, then intuitively, we have that $f$ has "enough" zeros and "not too many" poles.

## 7. Goppa Codes from Algebraic Geometry

Let us recall how we defined Reed-Solomon codes $RS(k,q)$. If $q$ is a prime power, $k$ is some integer $1 \leq k \leq q-1$, and $L_{k-1}$ is the set of polynomials of degree at most $k-1$, we define the evaluation mapping $\text{ev}_{\mathbb{F}_q}$

$$\begin{aligned}\text{ev}_{\mathbb{F}_q} : L_{k-1} &\to \mathbb{F}_q^{q-1} \\ f &\to (f(\alpha_1), \ldots, f(\alpha_{q-1}))\end{aligned}$$

We defined Reed-Solomon codes as $RS(k,q) = \text{ev}_{\mathbb{F}_q}(L_{k-1})$.

**Definition 7.1.** The *projective line* $\mathbb{P}^1(k)$ is defined as

$$\mathbb{P}^1(k) := \{k^2 \setminus \{(0,0)\}\}/\sim$$

where $(X_0 : Y_0) \sim (X_1 : Y_1)$ if there exists some $\lambda \in k^\times$ such that $X_0 = \lambda X_1$ and $Y_0 = \lambda Y_1$.

**Proposition 7.2.** *If we denote $P_\infty = (1:0) \in \mathbb{P}^2(k)$, then the divisor $D = (k-1)P_\infty$ has that $L(D) = L_{k-1}$.*

*Proof.* We will prove this by showing both inclusions.

First let $f \in L_{k-1}$. Suppose $f$ has degree $d \leq k-1$. Let $f(x) = \sum_{i=0}^{d} a_k x^k$. Homogenizing $f(x)$ to $F(X,Y)$, we get that

$$F(X,Y) = \frac{\sum_{i=0}^{d} a_k X^k Y^{d-k}}{Y^d}.$$

We see that $F$ has a degree $d$ pole at $Y = P_\infty$. If $P_i$ are the zeros of $F$, we have that $\text{div}(F) = \sum_{i=1}^{d} P_i - dP_\infty$. Therefore, $D + \text{div}(f) = \sum_{i=1}^{d} P_i + (k-1-d)P_\infty \geq 0$ since $d \leq k-1$ and thus $F \in L(D)$.

Now suppose $F \in L(D)$. We have that $(k-1)P_\infty + \text{div}(F) \geq 0$. Since $\text{div}(F) = \sum P - \sum Q$, we must have that $F$ only has poles at $P_\infty$. Additionally, $P_\infty$ must be a pole of degree at most $k-1$. Therefore, $F$ must be a polynomial of degree at most $k-1$ and thus $F \in L_{k-1}$. □

This proposition provides us a new perspective on the Reed-Solomon codes. The Reed-Solomon code $RS(k,q)$ is the image of an evaluation mapping where the set of points evaluated are the points in $\mathbb{F}_q$ and the set of functions acted on in the evaluation mapping is $L((k-1)P_\infty)$. The Goppa codes are a generalization of this idea and is constructed by using an arbitrary divisor and an arbitrary set of points.

**Definition 7.3.** Let $X$ be a projective plane curve over $\mathbb{F}_q$. Let $D$ be a divisor on $X$ and $\mathcal{P} = \{P_1, \ldots, P_n\}$ be a set of points on $X$ such that $\mathcal{P} \cap \text{supp}\, D = \varnothing$. The *Goppa code* $C(X, \mathcal{P}, D)$ is defined as

$$C(X, \mathcal{P}, D) := \{(f(P_1), \ldots, f(P_n)) | f \in L(D)\}.$$

Equivalently, we can define the evaluation map

$$\begin{aligned} \mathrm{ev}_{\mathcal{P}} : L(D) &\to \mathbb{F}_q^n \\ f &\to (f(P_1), \ldots, f(P_n)) \end{aligned}$$

Then define

$$C(X, \mathcal{P}, D) := \mathrm{ev}_{\mathcal{P}}(L(D))$$

Let us discuss the parameters of $C(X, \mathcal{P}, D)$. We have that $n = |\mathcal{P}|$. Since $C(X, \mathcal{P}, D)$ is the image of a linear transformation of $L(D)$, we have that $k \leq \dim L(D)$. We will have that $k = \dim L(D)$ if $\ker \mathrm{ev}_{\mathcal{P}} = \{0\}$. Suppose that $\mathrm{ev}_{\mathcal{P}}(f) = 0$. So $f$ has roots at $P_1, \ldots, P_n$. Therefore, we have that $\mathrm{div}(f) + D - \sum_{i=1}^{n} P_i \geq 0$ so $f \in L(D - \sum_{i=1}^{n} P_i)$. If we impose that $\deg D < n$, then we get that $L(D - \sum_{i=1}^{n} P_i) = \{0\}$ and therefore $f = 0$. Therefore, if $\deg D < n$, then $k = \dim L(D)$.

In fact, under certain conditions, we can even get a guaranteed bound on the minimum distance parameter. We will use two notions that I have not formally defined, in the sake of brevity. The first is that of nonsingularity. Intuitively, a nonsingular curve is one with no self-intersection points. The next is the notion of genus, which is intuitively the number of "holes" in a variety. The following theorem requires too much background to prove, but its bound provides an important bound for algebraic geometric curves.

**Theorem 7.4.** *(Riemann-Roch Theorem) Let $C$ be a nonsingular projective plane curve over $\mathbb{F}_q$. Suppose $C$ has genus $g$. If $D$ is a divisor over $C$, then* $\dim L(D) \geq \deg D + 1 - g$. *If* $\deg D > 2g - 2$, *then* $\dim L(D) = \deg D + 1 - g$.

With the Riemann-Roch Theorem, we obtain the following bound on the parameters of a Goppa code.

**Theorem 7.5.** *Let $X$ be a nonsingular projective plane curve over $\mathbb{F}_q$. Suppose $X$ has genus $g$ and $\mathcal{P} \subset X(\mathbb{F}_q)$ is a set of $\mathbb{F}_q$-rational points on $X$. Let $D$ be a divisor on $C$ such that $2g - 2 < \deg D < n$. Then the Goppa code $C(X, \mathcal{P}, D)$ has parameters*

$$\begin{aligned} n &= |\mathcal{P}| \\ k &= \deg D + 1 - g \\ d &\geq n - \deg D \end{aligned}$$

*Proof.* Our previous discussion yields that $n = |\mathcal{P}|$ and $k = \dim L(D)$. By the Riemann-Roch Theorem, $k = \deg D + 1 - g$.

Let us now bound the minimum distance. Suppose $f \in L(D)$ such that $\mathrm{ev}_{\mathcal{P}}(f)$ has minimum nonzero Hamming weight. We have that $\mathrm{ev}_{\mathcal{P}}(f)$ has $d$ nonzero coordinates. WLOG, suppose that $f(P_{d+1}) = \cdots = f(P_n) = 0$. Since $P_{d+1}, \ldots, P_n$ are all zeros of $f$, we have that $\mathrm{div}(f) + D - P_{d+1} - \cdots - P_n \geq 0$. Since $f \neq 0$, we must have that $D - P_{d+1} - \cdots - P_n$ has nonnegative degree. Thus, $\deg(D) - (n - d) \geq 0$ and so $d \geq n - \deg D$ as required. □

## 8. Sheaves on Algebraic Varieties

The Goppa construction provides us with a family of good linear codes built off of an algebraic curve $X$, a set of points $\mathcal{P}$, and a divisor $D$ on $X$. However, given our definition of a divisor as a formal sum, there is no clear indication as to why our construction of the Goppa code works or what it does. In this section, we will build up the theory of sheaves to be able to discuss local data on a variety. In our previous definition of a divisor, we had finitely many points in the support of the divisor. When studying the properties of the divisor, studying the entire variety would provide us with a lot of extraneous data. Instead, we should study local properties on Zariski open neighborhoods around the points in the support of the divisor. The study of sheaves is well-suited to studying local data.

**Definition 8.1.** Given a variety $X$, for every open set $U \subseteq X$ we associate an abelian group $\mathcal{F}(U)$ such that $\mathcal{F}(\varnothing) = \{0\}$. An element $s \in \mathcal{F}(U)$ is called *section*. If $V \subseteq U$, we should have a restriction map $\text{res}_{V,U} : \mathcal{F}(U) \to \mathcal{F}(V)$ such that

1. $\text{res}_{U,U}$ is the identity on $U$.
2. If $W \subseteq V \subseteq U$, then $\text{res}_{W,V} \circ \text{res}_{V,U} = \text{res}_{W,U}$

This system of sets and mappings is called a *presheaf*.

**Notation 8.2.** If $s \in U$, we often denote $\text{res}_{V,U}(s) := s|_V$.

Presheaves provide local data to open sets in a variety. The restriction morphisms give us a way of understanding the local data of a subset of a set. Note that presheaves can be more generally defined for topological spaces, not only varieties. The following example will use the real topology to build the intuition behind presheaves and highlight important properties we would like to our presheaves to possess.

**Example 8.3.** Let $X = \mathbb{R}$ and let $\mathcal{F}(U)$ be the set of bounded continuous functions over an open set $U$. If $f \in \mathcal{F}(U)$ and $V \subseteq U$, we can define $\text{res}_{V,U}(f) = f|_V$.

However, this presheaf has the limitation that the local data cannot effectively capture the global data. For example, suppose that $U_n = (-n, n)$. The identity function $f(x) = x$ has that $f \in \mathcal{F}(U_n)$ for every $n$. However, if we take the infinite union $\bigcup_{n=0}^{\infty} U_n = \mathbb{R}$, then $f \notin \mathcal{F}(\mathbb{R})$ since $f$ is not bounded on the real line. Therefore, despite the fact that $f \in \mathcal{F}(U_n)$ for every $n$, the fact that $f$ is not in their infinite union raises an issue. If we take an open cover of a set, we would like that the local data of the covering sets is usable to make claims on the global data of a set. This notions leads us into the definition of a sheaf.

**Definition 8.4.** A *sheaf* is a presheaf with the following additional properties:

1. Let $\{U_i\}$ be an open cover for some set $U$. If $s|_{U_i} = t|_{U_i}$ for all $U_i$ then $s = t$.
2. Let $\{U_i\}$ be an open cover for some open set $U$. If for all $i$, there is some $s_i \in \mathcal{F}(U_i)$ such that $s_i|_{U_i \cap U_j} = s_j|_{U_i \cap U_j}$, then there exists some unique $s \in \mathcal{F}(U)$ such that $s_i = s|_{U_i}$.

**Notation 8.5.** The set of global sections of a sheaf $\mathcal{F}$ over $X$ is often denoted $\Gamma(X, \mathcal{F})$.

The additional properties of a sheaf create a notion of global compatibility of local data.

**Example 8.6.** Let $X = \mathbb{R}$ and let $\mathcal{F}(U)$ be the set of continuous functions over an open set $U$. Again, we take the restriction map to be the restriction of a function on a set. Let us verify that this presheaf is a sheaf. Suppose that $\{U_i\}$ is an open cover. If $f|_{U_i} = g|_{U_i}$ for all $i$, then $f$ and $g$ agree on the entire real line and thus $f = g$. Without loss of generality, suppose that each $U_i$ is an open interval since the collection of all open intervals form a basis of the topology of $\mathbb{R}$. If $f_i|_{U_i \cap U_j} = f_j|_{U_i \cap U_j}$, then by defining $f = f_i$ on $U_i$ and $f = f_j$ on $U_j$, we get that $f$ is still continuous on each $U_i$ and that $f_{U_i} = f_i$.

Here are some important examples of sheaves on varieties that we will use:

1. $K$ is the constant sheaf of *rational functions* on $X$. If $U \subseteq X$, then $K(U) = K(X)$ which is the field of rational functions of $X$.
2. $\mathcal{O}_X$ is the sheaf of *regular functions* on $X$. A *regular function* $f$ on $X$ is one that can be expressed as a polynomial on $X$. More precisely, $f$ is a regular function if and only if $f \in k[x_0, \ldots, x_n]/I(X)$
3. $\mathcal{O}_X^*$ is the sheaf of *regular invertible functions* on $X$.

By taking the natural choice of restriction morphism, we can see that each of these examples is a sheaf.

## 9. Line Bundles and Divisors

Now that we have a notion of local data over a variety, we can use this information to see how local properties near the points in the support of a divisor allow us to generate codes. We will see that divisors generate line bundles, and we will be able to use the ideas of line bundles to generate different linear codes.

**Definition 9.1.** Let $X, Y$ be a projective varieties over a field $k$ and $\pi : Y \to X$ a surjective morphism of varieties. Suppose that we have the following properties:

1. For all $x \in X$, the *fiber* $\pi^{-1}(\{x\})$ is a finite-dimensional vector space.
2. For every point $x \in X$, there is an open neighborhood $U \subseteq X$ around $x$ such that $\pi^{-1}(U) \simeq U \times k^r$.

This structure is called a *vector bundle of rank* $r$ and the homeomorphism $\varphi : U \times k^r \to \pi^{-1}(U)$ is called the *local trivialization of* $U$.

**Definition 9.2.** A vector bundle of rank 1 is called a *line bundle*.

**Example 9.3.** (Trivial line bundle) Let $X$ be a variety and consider the line bundle $X \times k$ where $\pi(x, a) = x$.

**Example 9.4.** (Tautological line bundle) Let $\mathbb{P}^n$ be $n$-dimensional projective space. Let $x \in \mathbb{P}^n$, and let $V_x \subseteq k^{n+1}$ be the 1-dimensional vector space consisting of points contained in the equivalence class of $x \in \mathbb{P}^n$. We define

$$E := \{(x, v) | x \in \mathbb{P}^n, v \in V_x\} \subset \mathbb{P}^n \times k^{n+1}.$$

We let $\pi : E \to \mathbb{P}^n$ be defined as $\pi(x, v) = x$. To show that $E$ forms a line bundle, we need to find a local trivialization. Suppose that points $(x, v) \in E$ take the form

$$(x, v) = ((x_0 : \cdots : x_n), (v_0, \ldots, v_n)).$$

Over the open cover chart $U_i = \{x \in \mathbb{P}^n | x_i \neq 0\} \subset \mathbb{P}^n$, we can construct our local trivialization by defining the following map:

$$\begin{aligned} \varphi_i : \pi^{-1}(U_i) &\to U_i \times k \\ (x, v) &\to (x, v_i) \\ \varphi_i^{-1} : U_i \times k &\to \pi^{-1}(U_i) \\ (x, v_i) &\to \left(x, v_i \frac{x}{x_i}\right). \end{aligned}$$

The map $\varphi_i^{-1}$ is well-defined since $x_i \neq 0$ over $U_i$ and thus $\frac{x}{x_i}$ if a function on $U_i$.

A line bundle induces a sheaf of sections $\mathcal{L}$. The sheaf of sections over an open neighborhood $U$ is the set of homomorphisms $s \in \mathrm{Hom}(U, \pi^{-1}(U))$ which satisfy $\pi \circ s = Id_U$. Because $\pi^{-1}(U) \simeq U \times k^1$, it will be easier to look at homomorphisms in $\mathrm{Hom}(U, U \times k^1)$.

**Proposition 9.5.** *$\mathcal{L}$ is a sheaf over $X$.*

*Proof.* By taking the usual restriction mapping, $\mathcal{L}$ is a presheaf. If $\{U_i\}$ is an open cover of $X$ and $s|_{U_i} = t|_{U_i}$ for every $U_i$ in the cover, then in particular $s(x) = t(x)$ for all $x \in X$ and so $s = t$. Finally, if there are sections such that $s_i|_{U_i \cap U_j} = s_j|_{U_i \cap U_j}$ then by simply taking $s = s_i|_{U_i}$, the local sections $\mathrm{Hom}_{U_i}(U_i, \pi^{-1}(U_i))$ glue together to give a global section $\mathrm{Hom}_X(X, \pi^{-1}(X))$. □

**Example 9.6.** (Sheaf of sections of the trivial line bundle) Given the trivial line bundle, sections $s \in \mathcal{L}(U)$ takes the form $s(u) = (u, f(u))$ for some $f : U \to k$.

Given two sections $s_1, s_2 \in \mathcal{L}(U)$, we can define pointwise addition by $(s_1 + s_2)(x) = s_1(x) + s_2(x)$. Additionally, if $f \in \mathcal{O}_X(U)$ is a regular function, then we can define $(fs)(x) = f(x)s(x)$. Note that $s(x) \in \pi^{-1}(U)$, where $f(x)$ is a scalar quantity in the affine space. Therefore, $\mathcal{L}(U)$ is an $\mathcal{O}_X(U)$-module. In fact, we have the following proposition that makes it simpler to classify line bundles by using $\mathcal{O}_x$-modules.

**Proposition 9.7.** *A map $f : Y \to X$ makes $Y$ into a line bundle over $X$ if and only if its associated sheaf of sections $\mathcal{L}$ is a locally free $\mathcal{O}_X$-module of rank* 1

*Proof.* I will only prove the 'only if' direction, to show an intuitive connection between line bundles and sheaves that is accessible based off of the provided discussion.

By saying $\mathcal{L}$ is a locally free $\mathcal{O}_X$-module of rank 1, we mean that for any open set $U \subseteq X$, $\mathcal{L}(U)$ is a free $\mathcal{O}_X(U)$-module of rank 1.

$\Rightarrow$ Suppose that $f : Y \to X$ makes $Y$ into a line bundle over $X$. For any open neighborhood $U \subset X$, we have that $f^{-1}(U) \simeq U \times \mathbb{A}^1$. We have that $\mathcal{O}(f^{-1}(U)) \simeq \mathcal{O}(U \times \mathbb{A}^1)$. Therefore, we note that $f$ induces a map on $\mathcal{O}(U)$ to $\mathcal{O}(U \times k^1)$ defined by pullback.

$$\begin{aligned} \varphi : \mathcal{O}(U) &\to \mathcal{O}(U \times k^1) \\ g &\to g \circ f \end{aligned}$$

Similarly, given a section $s \in \mathcal{L}(U)$, we get another map

$$\begin{aligned} \psi : \mathcal{O}(U \times k^1) &\to \mathcal{O}(U) \\ h &\to h \circ s \end{aligned}$$

Now we observe that $\mathcal{O}(U \times \mathbb{A}^1) \simeq \mathcal{O}(U)[t]$. Since $\mathcal{O}(U \times \mathbb{A}^1)$ is the ring of polynomial functions $f(x_0, \ldots, x_n, t)$, we can write each $f \in \mathcal{O}(U \times k^1)$ as $f = \sum_{i=0}^{d} f_i t^i$ where $d$ is the maximum degree $t$ term and $f_i$ are regular functions over $U$.

Suppose that $s \in \mathcal{O}(U)$. We get a map $\varphi_S : \mathcal{O}(U) \to \mathcal{O}(U)$ that is generated by the following composition:

$$\begin{aligned} \varphi_S : \mathcal{O}(U) &\to \mathcal{O}(U)[t] \to \mathcal{O}(U) \\ g &\to g \circ f \to g \circ f \circ s \end{aligned}$$

Recall that for an open neighborhood $U \subseteq X$, we have that $s \in \mathcal{L}(U)$ if $f \circ s = Id_U$. Therefore, if $s \in \mathcal{L}(U)$ then $\varphi_S = Id_{\mathcal{O}(U)}$. Therefore, the map $\varphi_S$ is associated with the section $s' \in \mathcal{O}(U)$ such that from $\mathcal{O}(U)[t]$, $t$ gets mapped to $s'$. So each section in $\mathcal{L}(U)$ is associated with some section in $\mathcal{O}(U)$, namely the section where $t$ goes to in the mapping $\mathcal{O}(U)[t] \to \mathcal{O}(U)$. Thereofre, $\mathcal{L}(U) \simeq \mathcal{O}(U)$ and thus $\mathcal{L}(U)$ is a free $\mathcal{O}_U$-module of rank 1 and thus is a locally free $\mathcal{O}$-module of rank 1. □

Using this proposition, we can connect the idea of a divisor to line bundles by using divisors to construct a sheaf of sections that is an $\mathcal{O}_X$-module of rank 1. Suppose that we have a curve $X$ and a divisor $D$ on $X$. For each $p \in \operatorname{supp} D$, there exists a Zariski neighborhood $U_1^p$ of $p$ in $X$ such that $U_1^p \cap \operatorname{supp} D = p$. Intuitively, we find a Zariski neighborhood around $p$ that is small enough so that it does not contain any other points in the support.

Recall the sheaf $K(X)$ defines the field of rational functions of $X$. We can find a nonconstant rational function $g_p \in K(X)$ which has either a zero or pole at $p$. By further shrinking the neighborhood $U_1^p$ to another neighborhood $U_2^p$, we can ensure that the only zero or pole of $g_p$ restricted to $U_2^p$ is at $p$ itself.

Recall that the sheaf $\mathcal{O}_X$ is the sheaf of regular functions on $X$ and $\mathcal{O}_X^*$ is the sheaf of regular invertible functions on $X$. Let us consider the neighborhood $U_2^p - p$. Because of the way that we defined $U_2^p$, we have that $g_p$ has no zeroes or poles in $U_2^p - p$ and thus $\frac{1}{g_p}$ also has no zeroes or poles in $U_2^p - p$. Thus, $g_p \in \mathcal{O}_X^*(U_2^p - p)$.

Through this process, we have found for each $p \in \operatorname{supp} D$ some Zariski neighborhood $U_2^p$, which we will now denote as $U^p$, and some $g_p \in K(X)$ such that $g_p \in \mathcal{O}_X^*(U^p - p)$. Together, the $U_p$ and $g_p$ provide a sense of local data which suggests that the set of $g_p$ form a sheaf of sections $\mathcal{O}_X(D)$. We will show that this sheaf of sections is a locally free $\mathcal{O}_X$-module of rank 1.

**Proposition 9.8.** *To each divisor $D$ on $X$ we can associate $\mathcal{O}_X(D)$, a locally free sheaf of $\mathcal{O}_X$-modules of rank* 1.

*Proof.* We will show that over each neighborhood $U^p$, we have that $\mathcal{O}_X(D)$ is generated as a sheaf of $\mathcal{O}_X$-modules by $g_p^{-1}$. Consider two $p, q \in \operatorname{supp} D$. By how we defined $U^p$ and $U^q$, $g_p$ and $g_q$ have no zeroes or poles on the overlap $U^p \cap U^q$. Thus $g_p, g_q \in \mathcal{O}_X^*(U^p \cap U^q)$. So over the neighborhood $U^P \cap U^q$, $g_p$ and $g_q$ are both invertible. Therefore, we have that

$$\mathcal{O}_X(U^p \cap U^q) \cdot g_p = \mathcal{O}_X(U^p \cap U^q) \cdot g_q.$$

By defining $\mathcal{O}_{U^p}(D) := \mathcal{O}_X(U^p) \cdot g_p^{-1}$ we can glue together the local sheaves $\mathcal{O}_{U^p}(D)$ to form $\mathcal{O}_X(D) \subset K(X)$. Since the $g_p^{-1}$ are the generators of $\mathcal{O}_{U^p}(D)$, we have that $\mathcal{O}_X(D)$ is locally free of rank 1. □

Now let us connect divisors with line bundles in one more way. Recall that in Definition 6.7, if we have a projective plane curve $X$ over $\mathbb{F}_q$ we defined the space of rational functions associated to a divisor $D$ as

$$L(D) := \{f \in \mathbb{F}_q(X) | \text{div}(f) + D \geq 0\}.$$

**Proposition 9.9.** *If $D$ is a divisor of a projective plane curve $X$ over $\mathbb{F}_q$, we have that*

$$\Gamma(X, \mathcal{O}_X(D)) = L(D)$$

*Proof.* In our construction of $\mathcal{O}_X(D)$, we chose local functions $g_p \in K(X)$ that had zeros or poles at $p \in \text{supp}\, D$ and additionally had that $p$ is the only zero or pole in a neighborhood of $p$. Since we have that $\mathcal{O}_X(D)$ is a sheaf, we can glue together these $g_p$ over the different neighborhoods $U^p$ to generate some function $g$ such that $\text{div}(g) = D$. Therefore, $g|_{U_p} = g_p$ and so $\text{div}(g_p) = D$.

First suppose that $f \in L(D)$. Over a neighborhood $U^p$ we would like to show that $f|_{U^p} \in \mathcal{O}_X(U^p) \cdot g_p^{-1}$. I claim that $f|_{U^p} g_p \in \mathcal{O}_X(U^p)$. Indeed,

$$\text{div}(f|_{U^p} g_p) = \text{div}(f) + D \geq 0.$$

Therefore, any poles $f|_{U^p} g_p$ are cancelled out by zeros of equal or higher multiplicity. Therefore, $f|_{U^p} \in \mathcal{O}_X(U^p) \cdot g_p^{-1}$ meaning that $f \in \Gamma(X, \mathcal{O}_X(D))$.

Now let $f \in \Gamma(X, \mathcal{O}_X(D)) \subset \Gamma(X, K) = K(X)$. We have that $K(X)$ lies over $U^p$ in $\mathcal{O}_X(U^p) \cdot g_p^{-1}$. Suppose that $f = f_p g_p^{-1}$ for some $f_p \in \mathcal{O}_X(U_p)$. Since $f_p \in \mathcal{O}_X(U_p)$, $f_p$ has no poles in $\mathcal{O}_X(U_p)$ and thus $\text{div}(f_p) \geq 0$. Therefore, we have that

$$\begin{aligned} \text{div}(f) &= \text{div}(f_p) + \text{div}(g_p^{-1}) \\ \text{div}(f) &= \text{div}(f_p) - D \\ \text{div}(f) + D &= \text{div}(f_p) \geq 0 \end{aligned}$$

Therefore, $f \in L(D)$. □

## 10. Higher Dimensional Algebraic Geometric Codes

In the previous section, we drew a connection between line bundles and divisors. Since we used divisors to construct linear codes, it is natural to consider if we can get more code constructions utilizing line bundles. Indeed, we can do this by considering line bundles on an algebraic variety that is not necessarily an algebraic curve. To begin, let us view another way that presheaves can be used to capture local data.

**Definition 10.1.** Let $\mathcal{F}$ be a presheaf on a variety $X$ and $x \in X$. The *stalk* of $\mathcal{F}$ at $x$ is defined as

$$\mathcal{F}_x := \varinjlim_{U \ni x} \mathcal{F}(U).$$

The limit is taken over open sets $U$ containing $x$. A stalk captures the local data of variety around a point by looking at the data of open sets containing the point. In the context of a variety $X$, let us look at the specific case when $\mathcal{F}$ is some subsheaf of the sheaf of rational functions $K(X)$. Given some point $p \in X$, we have that $\mathcal{F}_p$ is the limit of $\mathbb{F}_q$-rational functions on decreasing neighborhoods around

$p$. As the neighborhood $U$ decreases, we have fewer points to worry about being poles.

With this idea in mind, we can take the stalk $\mathcal{F}_p$ modulo functions that vanish at $p$. More precisely, $\overline{\mathcal{F}}_p$ is the image of $p$ under the maps in $\mathcal{F}_p$. Suppose that $\mathcal{L}$ is the associated sheaf of sections over some line bundle of $X$. If $X$ is a variety over $\mathbb{F}_q$ and $p \in X$ is an $\mathbb{F}_q$-rational point, we get that $\overline{\mathcal{L}}_p \simeq \mathbb{F}_q$.

Equipped with these ideas, and drawing from previous notions of codes as images of evaluation maps, we construct the germ map.

**Definition 10.2.** Let $X$ be a smooth projective variety over $\mathbb{F}_q$. Let $\mathcal{P} = \{P_1, \ldots, P_n\}$ be a set of $\mathbb{F}_q$-rational points on $X$ and let $\mathcal{L}$ be the sheaf of sections associated to a line bundle over $X$. Suppose we have chosen an appropriate trivialization of $\mathcal{L}$ such that the sheaf of sections so that $\mathcal{L}_p \simeq \mathbb{F}_q$. We can define the *germ map* as

$$\alpha : \Gamma(X, \mathcal{L}) \to \bigoplus_{i=1}^{n} \overline{L}_{P_i} \simeq \mathbb{F}_q^n$$

The image of $\alpha$ is a linear code which we denote $\alpha\left(\Gamma(X, \mathcal{L})\right) := C\left(X, \mathcal{L}, \mathcal{P}\right)$

**Proposition 10.3.** *If $X$ is a smooth projective curve, $\mathcal{L} = \mathcal{O}_X(D)$ for some divisor $D$ over $X$, and $\mathcal{P} \cap \operatorname{supp} D = \varnothing$, then the code generated by the germ map is the same as the corresponding Goppa code.*

*Proof.* Let us recall that the Goppa codes were constructed as the image of the evaluation mapping

$$\begin{aligned} \operatorname{ev}_{\mathcal{P}} : L(D) &\to \mathbb{F}_q^n \\ f &\to (f(P_1), \ldots, f(P_n)) \end{aligned}$$

where

$$L(D) := \{f \in \mathbb{F}_q(C) | \operatorname{div}(f) + D \geq 0\}$$

Since our definition using the germ map is also the image of a similar evaluation mapping, this is an immediate corollary of Proposition 9.9, in which we proved $\Gamma(X, \mathcal{O}_X(D)) = L(D)$. □

Our notion of Goppa codes are consistent with the notions introducted with the germ map. The germ map thus provides a generalization of Goppa codes that does not require our variety to be an algebraic curve.

Although this paper is unable to explain the usefulness of these codes, it shows algebraic geometry can be utilized to provide a new perspective of our well-established notion of codes, as well as opening up new methods of code construction that can be efficiently encoded and decoded.

**Acknowledgments.** Firstly, I would like to thank Professor Peter May for organizing yet another incredible REU. Secondly, I would like to thank all the lecturers for providing such great learning opportunities. In particular, I would like to mention Professor Antoni Rangachev, for piquing my interest in algebraic geometry. Last, but certainly not least, I would like to express my gratitude towards my mentor Owen Barrett for his incredible patience, mentorship, and contagious passion for algebraic geometry. Without him, this project would not have been possible.

## References

[1] William Fulton. Algebraic Curves: An Introduction to Algebraic Geometry. Addison-Wesley Publishing Company. 1989.
[2] John B. Little. Algebraic Geometry Codes From higher Dimensional Varieties. arXiv:0802.2349v1 [cs.IT] 2008.
[3] Igor R. Shafarevich. Basic Algebraic Geometry. Springer-Verlag. 1997.
[4] Judy L. Walker. Codes and Curves. University of Nebraska Press. 1991.