Projective Geometry Lecture Notes

Thomas Baird

March 26, 2012

Contents

1 Introduction ................................................................. 2

2 Vector Spaces and Projective Spaces ................................. 3
  2.1 Vector spaces and their duals ......................................... 3
  2.2 Projective spaces and homogeneous coordinates .................. 5
    2.2.1 Visualizing projective space ...................................... 5
    2.2.2 Homogeneous coordinates ........................................ 6
  2.3 Linear subspaces ..................................................... 7
    2.3.1 Two points determine a line ..................................... 7
    2.3.2 Two planar lines intersect at a point ........................... 8
  2.4 Projective transformations and the Erlangen Program .......... 8
    2.4.1 Erlangen Program ................................................ 9
    2.4.2 Projective versus linear ......................................... 10
    2.4.3 Examples of projective transformations ......................... 11
    2.4.4 Direct sums ..................................................... 12
    2.4.5 General position ................................................. 13
  2.5 Classical Theorems .................................................. 15
    2.5.1 Desargues’ Theorem ............................................... 15
    2.5.2 Pappus’ Theorem ................................................ 16
  2.6 Duality .............................................................. 17

3 Quadrics and Conics ..................................................... 20
  3.1 Affine algebraic sets ................................................. 20
  3.2 Projective algebraic sets ............................................. 21
  3.3 Bilinear and quadratic forms ....................................... 22
    3.3.1 Quadratic forms ................................................ 24
    3.3.2 Change of basis ................................................ 25
    3.3.3 Digression on the Hessian ...................................... 27
  3.4 Quadrics and Conics ................................................ 28
  3.5 The rational parametrization of the circle ....................... 30
  3.6 Polars ............................................................ 33
  3.7 Linear subspaces of quadrics and ruled surfaces ................. 35
1 Introduction

Projective geometry has its origins in Renaissance Italy, in the development of perspective in painting: the problem of capturing a 3-dimensional image on a 2-dimensional canvas. It is a familiar fact that objects appear smaller as the they get farther away, and that the apparent angle between straight lines depends on the vantage point of the observer. The more familiar Euclidean geometry is not well equipped to make sense of this, because in Euclidean geometry length and angle are well-defined, measurable quantities independent of the observer. Projective geometry provides a better framework for understanding how shapes change as perspective shifts.

The projective geometry most relevant to painting is called the real projective plane, and is denoted $\mathbb{R}P^2$ or $P(\mathbb{R}^3)$.

**Definition 1.** The real projective plane, $\mathbb{R}P^2 = P(\mathbb{R}^3)$ is the set of 1-dimensional subspaces of $\mathbb{R}^3$.

This definition is best motivated by a picture. Imagine an observer sitting at the origin in $\mathbb{R}^3$ looking out into 3-dimensional space. The 1-dimensional subspaces of $\mathbb{R}^3$ can be understood as lines of sight. If we now situate a (Euclidean) plane $P$ that doesn’t contain the origin, then each point in $P$ determines unique sight line. Objects in (subsets of) $\mathbb{R}^3$ can now be “projected” onto the plane $P$, enabling us to translate a 3-dimensional scene onto a 2-dimensional scene. Since 1-dimensional lines translate into points on the plane, we call the 1-dimensional lines *projective points*.

Of course, not every projective point corresponds to a point in $P$, because some 1-dimensional subspaces are parallel to $P$. Such points are called *points at infinity*. To motivate this terminology, consider a family of projective points that rotate from projective points that intersect $P$ to one that is parallel. The projection onto $P$ becomes a family of points that diverges to infinity and then disappears. But as projective points they converge to a point at infinity. It is important to note that a projective point is only at infinity with respect to some Euclidean plane $P$.

One of the characteristic features of projective geometry is that every distinct pair of *projective lines* in the projective plane intersect. This runs contrary to the *parallel postulate* in Euclidean geometry, which says that lines in the plane intersect except when they are parallel. We will see that two lines that appear parallel in a Euclidean plane will intersect at a point at infinity when they are considered as projective lines. As a general rule, theorems about intersections between geometric sets are easier to prove, and require fewer exceptions when considered in projective geometry.
2 Vector Spaces and Projective Spaces

2.1 Vector spaces and their duals

Let $F$ denote a field (for our purposes $F$ is either the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$). A vector space $V$ over $F$ is set equipped with two operations:

- **Addition**: $V \times V \to V$, $(v, w) \mapsto v + w$.
- **Scalar multiplication**: $F \times V$, $(\lambda, v) \mapsto \lambda v$

satisfying the following list of axioms for all $u, v, w \in V$ and $\lambda, \mu \in F$.

- $(u + v) + w = u + (v + w)$ (additive associativity)
- $u + v = v + u$ (additive commutativity)
- There exists a vector $0 \in V$ called the zero vector such that, $0 + v = v$ for all $v \in V$ (additive identity)
- For every $v \in V$ there exists $-v \in V$ such that $v + -v = 0$ (additive inverses)
- $(\lambda + \mu)v = \lambda v + \mu v$ (distributivity)
- $\lambda(u + v) = \lambda u + \lambda v$ (distributivity)
- $\lambda(\mu v) = (\lambda \mu)v$ (associativity of scalar multiplication)
- $1v = v$ (identity of scalar multiplication)

I will leave it as an exercise to prove that $0v = 0$ and that $-1v = -v$.

**Example 1.** The set $V = \mathbb{R}^n$ of $n$-tuples of real numbers is a vector space over $F = \mathbb{R}$ in the usual way. Similarly, $\mathbb{C}^n$ is a vector space over $\mathbb{C}$. The vector spaces we consider in this course will always be isomorphic to one of these examples.

Given two vector spaces $V, W$ over $F$, a map $\phi : V \to W$ is called linear if for all $v_1, v_2 \in V$ and $\lambda_1, \lambda_2 \in F$, we have

$$\phi(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 \phi(v_1) + \lambda_2 \phi(v_2).$$

**Example 2.** A linear map $\phi$ from $\mathbb{R}^m$ to $\mathbb{R}^n$ can be uniquely represented by an $n \times m$ matrix of real numbers. In this representation, vectors in $\mathbb{R}^m$ and $\mathbb{R}^n$ are represented by column vectors, and $\phi$ is defined using matrix multiplication. Similarly, a linear map from $\mathbb{C}^m$ to $\mathbb{C}^n$ can be uniquely represented by an $n \times m$ matrix of complex numbers.

We say that a linear map $\phi : V \to W$ is an isomorphism if $\phi$ is a one-to-one and onto. In this case, the inverse map $\phi^{-1}$ is also linear, hence is also an isomorphism. Two vector spaces are called isomorphic if there exists an isomorphism between them. A vector space $V$ over $F$ is called $n$-dimensional if it is isomorphic to $F^n$. $V$ is called finite dimensional.
if it is \( n \)-dimensional for some \( n \in \{0, 1, 2, \ldots\} \). Otherwise we say that \( V \) is infinite dimensional.

A **vector subspace** of a vector space \( V \) is a subset \( U \subset V \) which is closed under addition and scalar multiplication. This makes \( U \) into a vector space in its own right, and the inclusion map \( U \hookrightarrow V \) is linear. Given any linear map \( \phi : V \to W \), we can define two vector subspaces: the **image**

\[
\phi(V) := \{ \phi(v) | v \in V \} \subseteq W
\]

and the **kernel**

\[
\ker(\phi) := \{ v \in V | \phi(v) = 0 \} \subseteq V
\]

The dimension of \( \phi(V) \) is called the **rank** of \( \phi \), and the dimension of \( \ker(\phi) \) is called the **nullity** of \( \phi \). The **rank-nullity theorem** states that the rank plus the nullity equals the dimension of \( V \). I.e., given a linear map \( \phi : V \to W \)

\[
\dim(V) = \dim(\ker(\phi)) + \dim(\phi(V)).
\]

Given a set of vectors \( S := \{ v_1, \ldots, v_k \} \subset V \) a linear combination is an expression of the form

\[
\lambda_1 v_1 + \ldots + \lambda_k v_k
\]

where \( \lambda_1, \ldots, \lambda_k \in F \) are scalars. The set of vectors that can be written as a linear combination of vectors in \( S \) is called the **span** of \( S \). The set \( S \) is called **linearly independent** if the only linear combination satisfying

\[
\lambda_1 v_1 + \ldots + \lambda_k v_k = 0
\]

is when \( \lambda_1 = \lambda_2 = \ldots = \lambda_k = 0 \). If some non-trivial linear combination equals zero, then we say \( S \) is **linearly dependent**.

A **basis** of \( V \) is linearly independent set \( S \subset V \) that spans \( V \). Every vector space has a basis, and the dimension of a vector space is equal to the size of any basis of that space. If \( \{ v_1, \ldots, v_n \} \) is a basis of \( V \), then this defines an isomorphism \( \Phi : F^n \to V \) by

\[
\Phi((x_1, \ldots, x_n)) = x_1 v_1 + \ldots + x_n v_n.
\]

Conversely, given an isomorphism \( \Phi : F^n \to V \), we obtain a basis \( \{ v_1, \ldots, v_n \} \subset V \) by setting \( v_i = \Phi(e_i) \) where \( \{ e_1, \ldots, e_n \} \in F^n \) is the standard basis. So choosing a basis for \( V \) amounts to the same thing as defining an isomorphism from \( F^n \) to \( V \).

**Proposition 2.1.** Suppose that \( W \subseteq V \) is a vector subspace. Then \( \dim(W) \leq \dim(V) \). Furthermore, we have equality \( \dim(W) = \dim(V) \) if and only if \( W = V \).

**Proof.** Exercise. \( \square \)

Given any vector space \( V \), the **dual** vector space \( V^* \) is the set of linear maps from \( V \) to \( F \)

\[
V^* := \{ \phi : V \to F \}.
\]
The dual $V^*$ is naturally a vector space under the operations

$$(\phi_1 + \phi_2)(v) = \phi_1(v) + \phi_2(v)$$

and

$$(\lambda \phi)(v) = \lambda(\phi(v)).$$

**Example 3.** Suppose $V \cong F^n$ is the set of $1 \times n$ column vectors. Then $V^*$ is identified with the set of $n \times 1$ row vectors under matrix multiplication.

Suppose that $v_1, ..., v_n \in V$ is a basis. Then the *dual basis* $v_1^*, ..., v_n^* \in V^*$ is the unique set of linear maps satisfying

$$v_i^*(v_j) = \delta^i_j,$$

where $\delta^i_j$ is the *Dirac delta* defined by $\delta^i_i = 1$ and $\delta^i_j = 0$ if $i \neq j$.

### 2.2 Projective spaces and homogeneous coordinates

**Definition 2.** Let $V$ be a vector space. The *projective space* $P(V)$ of $V$ is the set of 1-dimensional subspaces of $V$.

**Definition 3.** If $V$ is a vector space of dimension $n + 1$, then we say that $P(V)$ is a projective space of dimension $n$. A 0-dimensional projective space is called a *projective point*, a 1-dimensional vector space is called a *projective line*, and a 2-dimensional vector space is called a *projective plane*. If $V$ is one of the standard vector spaces $\mathbb{R}^{n+1}$ or $\mathbb{C}^{n+1}$ we also use notation $\mathbb{R}P^n := P(\mathbb{R}^{n+1})$ and $\mathbb{C}P^n := P(\mathbb{C}^{n+1})$.

#### 2.2.1 Visualizing projective space

To develop some intuition about projective spaces, we consider some low dimensional examples.

If $V$ is one-dimensional, then the only one dimensional subspace is $V$ itself, and consequently $P(V)$ is simply a point (a one element set).

Now consider the case of the real projective line. Let $V = \mathbb{R}^2$ with standard coordinates $x, y$ and let $L \subseteq \mathbb{R}^2$ be the line defined $x = 1$. With the exception of the $y$-axis, each 1-dimensional subspace of $\mathbb{R}^2$ intersects $L$ in exactly one point. Thus $\mathbb{R}P^1 = P(\mathbb{R}^2)$ is isomorphic as a set with the union of $L$ with a single point. Traditionally, we identify $L$ with $\mathbb{R}$ and call the extra point $\{\infty\}$, the *point at infinity*. Thus we have a bijection of sets

$$\mathbb{R}P^1 \cong \mathbb{R} \cup \{\infty\}.$$

Another approach is to observe that every 1-dimensional subspace intersects the unit circle $S^1$ in a pair of *antipodal points*. Thus a point in $\mathbb{R}P^1$ corresponds to a pair of antipodal points in $S^1$. Topologically, we can construct $\mathbb{R}P^1$ by taking a closed half circle and identifying the endpoints. Thus $\mathbb{R}P^1$ is itself a circle.

By similar reasoning we can see that the complex projective line $\mathbb{C}P^1$ is in bijection with $\mathbb{C} \cup \{\infty\}$

$$\mathbb{C}P^1 \cong \mathbb{C} \cup \{\infty\}.$$
Topologically, \( \mathbb{C}P^1 \) is isomorphic to the 2-dimensional sphere \( S^2 \) and is sometimes called the *Riemann Sphere*.

We turn next to the real projective plane \( \mathbb{R}P^2 = P(\mathbb{R}^3) \). Let \( x, y, z \) be the standard coordinates on \( \mathbb{R}^3 \). Then every 1-dimensional subspace of \( \mathbb{R}^3 \) either intersects the plane \( z = 1 \) or lies in the plane \( z = 0 \). The first class of 1-dimensional subspaces is in bijection with \( \mathbb{R}^2 \), and the second class is in bijection with \( \mathbb{R}P^1 \). Thus we get a bijection

\[
\mathbb{R}P^2 \cong \mathbb{R}^2 \cup \mathbb{R}P^1 \cong \mathbb{R}^2 \cup \mathbb{R} \cup \{\infty\}.
\]

It is not hard to see that this reasoning works in every dimension \( n \) to determine the following bijection

\[
\mathbb{R}P^n \cong \mathbb{R}^n \cup \mathbb{R}P^{n-1} \cong \mathbb{R}^n \cup \mathbb{R}^{n-1} \cup \ldots \cup \mathbb{R}^0 \tag{1}
\]

and similarly

\[
\mathbb{C}P^n \cong \mathbb{C}^n \cup \mathbb{C}P^{n-1} \cong \mathbb{C}^n \cup \mathbb{C}^{n-1} \cup \ldots \cup \mathbb{C}^0. \tag{2}
\]

Returning now to the projective plane, observe that every 1-dimensional subspace of \( \mathbb{R}^3 \) intersects the unit sphere \( S^2 \) in a pair of antipodal points. Thus we can construct \( \mathbb{R}P^2 \) by taking (say) the northern hemisphere and identifying antipodal points along the equator.

### 2.2.2 Homogeneous coordinates

To go further it is convenient to introduce some notation. Given a non-zero vector \( v \in V \setminus \{0\} \), define \([v] \in P(V)\) to be the 1-dimensional spanned by \( v \). If \( \lambda \in F \setminus \{0\} \) is a non-zero scalar, then

\[ [\lambda v] = [v]. \]

Now suppose we choose a basis \( \{v_0, \ldots, v_n\} \) for \( V \) (recall this is basically the same as defining an isomorphism \( V \cong F^n \)). Then every vector in \( v \in V \) can be written

\[ v = \sum_{i=0}^{n} x_i v_i \]

for some set of scalars \( x_i \) and we can write

\[ [v] = [x_0, \ldots, x_n] \]

with respect to this basis. These are known as *homogeneous coordinates*. Once again, for \( \lambda \neq 0 \) we have

\[ [\lambda x_0, \ldots, \lambda x_n] = [x_0, \ldots, x_n] \]
Consider now the subset $U_0 \subset P(V)$ consisting of points $[x_0, ..., x_n]$ with $x_0 \neq 0$ (Observe that this subset is well-defined independent of the representative $(x_0, ..., x_n)$ because for $\lambda \neq 0$ we have $x_0 \neq 0$ if and only if $\lambda x_0 \neq 0$). The for points in $U_0$

$$[x_0, ..., x_n] = [x_0, x_0(x_1/x_0), ..., x_0(x_n/x_0)] = [1, x_1/x_0, ..., x_n/x_0]$$

Thus we can uniquely represent any point in $U_0$ by a representative vector of the form $(1, y_1, ..., y_n)$. It follows that $U_0 = F^n$. This result has already been demonstrated geometrically in equations (1) and (2) because $U_0$ is the set of 1-dimensional subspaces in $V$ that intersect the “hyperplane” $x_0 = 1$. The complement of $U_0$ is the set of points with $x_0 = 0$, which form a copy of $P(F^n)$.

### 2.3 Linear subspaces

**Definition 4.** A linear subspace of a projective space $P(V)$ is a subset

$$P(W) \subseteq P(V)$$

consisting of all 1-dimensional subspaces lying in some vector subspace $W \subseteq V$. Observe that any linear subspace $P(W)$ is a projective space in its own right.

For example, each point in $P(V)$ is a linear subspace because it corresponds to a 1-dimensional subspace. If $W \subseteq V$ is 2-dimensional, we call $P(W) \subseteq P(V)$ a projective line in $P(V)$. Similarly, if $W$ is 3-dimensional then $P(W) \subseteq P(V)$ is a projective plane in $P(V)$. Observe that when each of these linear subspaces are intersected with the subset $U_0 = \{x_0 = 1\} \cong F^n$ then they become points, lines and planes in the traditional Euclidean sense.

#### 2.3.1 Two points determine a line

**Theorem 2.2.** Given two distinct points $[v_1], [v_2] \in P(V)$, there is a unique projective line containing them both.

**Proof.** Since $[v_1]$ and $[v_2]$ represent distinct points in $P(V)$, then the representative vectors $v_1, v_2 \in V$ must be linearly independent, and thus span a 2-dimensional subspace $W \subseteq V$. Thus $P(W)$ is a projective line containing both $[v_1]$ and $[v_2]$.

Now suppose that there exists another projective line $P(W') \subseteq P(V)$ containing $[v_1]$ and $[v_2]$. Then $v_1, v_2 \in W$, so $W \subseteq W'$. Since both $W$ and $W'$ are 2-dimensional it follows that $W = W'$ and $P(W) = P(W')$.

**Example 4.** Consider the model of $\mathbb{R}P^2$ as pairs of antipodal points on the sphere $S^2$. A projective line in this model is represented by a great circle. Proposition 2.2 amounts to the result that any pair of distinct and non-antipodal points in $S^2$ lie on a unique great circle (by the way, this great circle also describes the optimal flight route for airplanes flying between two points on the planet).
2.3.2 Two planar lines intersect at a point

Theorem 2.3. In a projective plane, two distinct projective lines intersect in a unique point.

Proof. Let \( P(V) \) be a projective plane where \( \dim(V) = 3 \). Two distinct projective lines \( P(U) \) and \( P(W) \) correspond to two distinct subspaces \( U,W \subseteq V \) each of dimension 2. Then we have an equality

\[
P(U) \cap P(W) = P(U \cap W)
\]

so proposition amounts to showing that \( U \cap W \subset V \) is a vector subspace of dimension 1. Certainly

\[
\dim(U \cap W) \leq \dim(U) = 2.
\]

Indeed we see that \( \dim(U \cap W) \leq 1 \) for otherwise \( \dim(U \cap W) = 2 = \dim(U) = \dim(W) \)
so \( U \cap W = U = W \) which contradicts the condition that \( U,W \) are distinct.

Thus to prove \( \dim(U \cap W) = 1 \) it is enough to show that \( U \cap W \) contains a non-zero vector. Choose bases \( u_1,u_2 \in U \) and \( w_1,w_2 \in W \). Then the subset \( \{ u_1,u_2,w_1,w_2 \} \subset V \) must be linearly dependent because \( V \) only has dimension 3. Therefore, there exist scalars \( \lambda_1,\ldots,\lambda_4 \) not all zero such that

\[
\lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 w_1 + \lambda_4 w_2 = 0
\]

from which it follows that \( \lambda_1 u_1 + \lambda_2 u_2 = -\lambda_3 w_1 - \lambda_4 w_2 \) is a non-zero element in \( U \cap W \).

Example 5. Using the model of \( \mathbb{R}P^2 \) as pairs of antipodal points in \( S^2 \), then Proposition 2.3 is the geometric fact that pairs of great circles always intersect in a pair of antipodal points.

2.4 Projective transformations and the Erlangen Program

So far we have explored how a vectors space \( V \) determines a projective space \( P(V) \) and how a vector subspace \( W \subseteq V \) determines a linear subspace \( P(W) \). What about linear maps?

It is not true in general that a linear map \( T : V \to V' \) induces a map between the projective spaces \( P(V) \) and \( P(V') \). This is because if \( T \) has a non-zero kernel, then it sends some 1-dimensional subspaces of \( V \) to zero in \( V' \). However, if \( \ker(T) = 0 \), or equivalently if \( T \) is injective, then this works fine.

Definition 5. Let \( T : V \to V' \) be an injective linear map. Then the map

\[
\tau : P(V) \to P(V')
\]

defined by

\[
\tau([v]) = [T(v)]
\]

is called the projective morphism induced by \( T \).

If \( T \) is an isomorphism, then \( \tau \) is bijective and we call \( \tau \) a projective transformation.

Projective transformations from a projective space to itself \( \tau : P(V) \to P(V) \) are particularly interesting. These are considered to be group of symmetries of projective space.
2.4.1 Erlangen Program

To better understand the meaning of this phrase, we make a brief digression to consider the more familiar Euclidean geometry in the plane. This kind of geometry deals with certain subsets of \( \mathbb{R}^2 \) and studies their properties. Especially important are points, lines, lengths and angles. For example, a triangle is an object consisting of three points and three line segments joining them; the line segments have definite lengths and their intersections have definite angles that add up to 180 degrees. A circle is a subset of \( \mathbb{R}^2 \) of points lying a fixed distance \( r \) from a point \( p \in \mathbb{R}^2 \). The length \( r \) is called the radius and circumference of the circle is \( 2\pi r \).

The *group of symmetries* of the Euclidean plane (or Euclidean transformations) consists of translations, rotations, and combinations of these. All of the objects and quantities studied in Euclidean geometry are preserved by these symmetries: lengths of line segments, angles between lines, circles are sent to circles and triangles are sent to triangles. Two objects in the plane are called *congruent* if they are related by Euclidean transformations, and this is the concept of “isomorphic” in this context.

In 1872, Felix Klein initiated a highly influential re-examination of geometry called the Erlangen Program. At the time, several different types of “non-Euclidean” geometries had been introduced, and it was unclear how they related to each other. Klein’s philosophy was that a fundamental role was played by the group of symmetries in the geometry; that the meaningful concepts in a geometry are those that preserved by the symmetries and that geometries can be related in a hierarchy according to how much symmetry they possess.

With this philosophy in mind, let us explore the relationship between projective and Euclidean geometry in more depth. Recall from \( \S 2.2.2 \) that we have a subset \( U_0 \subset \mathbb{R}P^n \) consisting of those points with homogeneous coordinates \([1, y_1, \ldots, y_n]\). There is an obvious bijection

\[ U_0 \cong \mathbb{R}^n, \quad [1, y_1, \ldots, y_n] \leftrightarrow (y_1, \ldots, y_n) \]

**Proposition 2.4.** Any Euclidean transformation of \( U_0 \cong \mathbb{R}^n \) extends uniquely to a projective transformation of \( \mathbb{R}P^n \).

**Proof.** We focus on the case \( n = 2 \) for notational simplicity. It is enough to show that translations and rotations of \( U_0 \) extend uniquely.

Consider first translation in \( \mathbb{R}^2 \) by a vector \((a_1, a_2)\). That is the map \((y_1, y_2) \mapsto (y_1 + a_1, y_2 + a_2)\). The corresponding projective transformation is associated to the linear transformation

\[
\begin{pmatrix}
1 \\
y_1 \\
y_2
\end{pmatrix}
\mapsto
\begin{pmatrix}
1 & 0 & 0 \\
a_1 & 1 & 0 \\
a_2 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
1 \\
y_1 \\
y_2
\end{pmatrix}
= \begin{pmatrix}
1 \\
y_1 + a_1 \\
y_2 + a_2
\end{pmatrix}
\]

A rotation by angle \( \theta \) of \( U_0 = \mathbb{R}^2 \) extends to the projective transformation induced by the linear transformation
\[
\begin{pmatrix}
1 \\
y_1 \\
y_2
\end{pmatrix}
\mapsto
\begin{pmatrix}
1 & 0 & 0 \\
0 & \cos(\theta) & -\sin(\theta) \\
0 & \sin(\theta) & \cos(\theta)
\end{pmatrix}
\begin{pmatrix}
1 \\
y_1 \\
y_2
\end{pmatrix}
= 
\begin{pmatrix}
1 \\
\cos(\theta)y_1 + \sin(\theta)y_2 \\
\cos(\theta)y_2 - \sin(\theta)y_1
\end{pmatrix}
\]

Uniqueness will be clear after we learn about general position. \hfill \Box

From the point of view of the Erlangen Program, Proposition ?? tells us that projective geometry more basic than Euclidean geometry because it has more symmetry. This means that theorems and concepts in projective geometry translate into theorems and concepts in Euclidean geometry, but not vice versa. For example, we have lines and points and triangles in projective geometry, but no angles or lengths. Furthermore, objects in Euclidean geometry that are not isomorphic (or congruent) to each other can be isomorphic to each other in projective geometry. We will show for example that all projective triangles are isomorphic to one another. We will also learn that circles, ellipses, parabolas and hyperbolas all look the same in projective geometry, where they are called conics.

### 2.4.2 Projective versus linear

Next, we should clarify the relationship between linear maps and projective morphisms.

**Proposition 2.5.** Two injective linear maps \( T : V \to W \) and \( T' : V \to W \) determine the same projective morphism if and only if \( T = \lambda T' \) for some non-zero scalar \( \lambda \).

**Proof.** Suppose that \( T = \lambda T' \). Then for \([v] \in P(V)\),

\[
[T(v)] = [\lambda T'(v)] = [T'(v)],
\]

so \( T \) and \( T' \) define the same projective morphism.

Conversely, suppose that for all \([v] \in P(V)\) that \([T(v)] = [T'(v)]\). Then for any basis \(v_0, ..., v_n \in V\), there exist non-zero scalars \(\lambda_0, ..., \lambda_n\) such that

\[
T(v_i) = \lambda_i T'(v_i).
\]

However it is also true that \([T(v_0 + ... + v_n)] = [T'(v_0 + ... + v_n)]\) so for some non-zero scalar \(\lambda\) we must have

\[
T(v_0 + ... + v_n) = \lambda T'(v_0 + ... + v_n).
\]

Combining these equations and using linearity, we get

\[
\sum_{i=0}^{n} \lambda T'(v_i) = \lambda T'(\sum_{i=0}^{n} v_i) = T(\sum_{i=0}^{n} v_i) = \sum_{i=0}^{n} T(v_i) = \sum_{i=0}^{n} \lambda_i T'(v_i)
\]

Subtracting gives \(\sum_{i=0}^{n} (\lambda - \lambda_i) T'(v_i) = 0\). Since \(T'\) is an injective, \(\{T'(v_i)\}\) is a linearly independent set, so \(\lambda = \lambda_i\) for all \(i\) and \(T = \lambda T'\). \hfill \Box
2.4.3 Examples of projective transformations

To develop some intuition about projective transformations, let’s consider transformations of the projective line $FP^1 = P(F^2)$ for $F = \mathbb{R}, \mathbb{C}$. A linear transformation of $F^2$ is uniquely described by a $2 \times 2$ matrix,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}$$

so for a projective point of the form $[1 : y]$ we have

$$\tau([1 : y]) = [a + by : c + dy] = [1 : \frac{c + dy}{a + by}].$$

In the case $F = \mathbb{C}$, the function $f : \mathbb{C} \to \mathbb{C} \cup \{\infty\}$ of the complex numbers

$$f(y) = \frac{c + dy}{a + by}$$

is called a Mobius transformation and plays an important role in complex analysis.

**Proposition 2.6.** Expressed in a Euclidean coordinate $y \leftrightarrow [1 : y]$, every projective transformation of the line $FP^1$ is equal to a combination of the following transformations:

- A translation: $y \mapsto y + a$
- A scaling: $y \mapsto \lambda y, \quad \lambda \neq 0$
- An inversion: $y \mapsto \frac{1}{y}$

**Proof.** Recall from linear algebra that any invertible matrix can be obtained from the identity matrix using elementary row operations: multiplying a row by a non-zero scalar, permuting rows, and adding one row to another. Equivalently, every invertible matrix can be written as a product of matrices of the form

$$\begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

and these correspond to scaling, inversion and translation respectively. For instance,

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ y \end{bmatrix} = \begin{bmatrix} y \\ 1 \end{bmatrix} = y \begin{bmatrix} 1 \\ 1/y \end{bmatrix}$$

so corresponds to the inversion $y \mapsto \frac{1}{y}$.

It is also informative to try and picture these transformations using our model where $\mathbb{R}P^1$ is a circle and $\mathbb{C}P^1$ is a sphere (leave this to class discussion).
Remark 1. More generally, a projective transformation of $\mathbb{R}P^n$ takes the form

$$\bar{y} \mapsto \frac{A\bar{y} + \bar{c}}{\lambda + b \cdot \bar{y}}$$

where $A$ is an $n \times n$ matrix, $\bar{b}, \bar{c}$ are vectors in $\mathbb{R}^n$ and $\lambda$ is scalar and $\cdot$ is the dot product. Proving this is an exercise.

A more geometrically defined transformation is obtained as follows. Let $P(V)$ be a projective plane, let $P(U), P(U')$ be two projective lines in $P(V)$, and let $[w] \in P(V)$ be a projective points not lying on $P(U)$ or $P(U')$. We define a map

$$\tau : P(U) \rightarrow P(U')$$

as follows: Given $[u] \in P(U)$, there is a unique line $L \subset P(V)$ containing both $[w]$ and $[u]$. Define $\tau([u])$ to be the unique intersection point $L \cap P(U')$. Observe that $\tau$ is well defined by Theorem 2.2 and 2.3.

**Proposition 2.7.** The map $\tau : P(U) \rightarrow P(U')$ is a projective transformation.

Before proving this proposition, it is convenient to introduce a concept from linear algebra: the **direct sum**.

### 2.4.4 Direct sums

Let $V$ and $W$ be vector spaces over the same field $F$. The direct sum $V \oplus W$ is the vector space with vector set equal to the cartesian product

$$V \oplus W := V \times W := \{(v, w) | v \in V, w \in W\}$$

with addition and scalar multiplication defined by $(v, w) + (v', w') = (v + v', w + w')$ and $\lambda(v, w) = (\lambda v, \lambda w)$. Exercise: show that

$$\dim(V \oplus W) = \dim(V) + \dim(W).$$

There are natural linear maps

$$i : V \rightarrow V \oplus W, \quad i(v) = (v, 0)$$

and

$$p : V \oplus W \rightarrow V, \quad p(v, w) = v$$

called inclusion and projection maps respectively. Similar maps exist for $W$ as well.

**Proposition 2.8.** Let $V$ be a vector space and let $U, W$ be vector subspaces such that $\text{span}\{U, W\} = V$. Then the natural map

$$\Phi : U \oplus W \rightarrow V, \quad (u, w) \mapsto u + w$$

is surjective with kernel

$$\ker(\Phi) \cong U \cap W.$$

In particular, if $U \cap W = \{0\}$, then $\Phi$ defines an isomorphism $U \oplus W \cong V$ and we say that $V$ is the internal direct sum of $U$ and $W$. 

12
Proof. Because $\text{span}\{U, W\} = V$ every vector in $V$ is a sum of vectors in $U$ and $W$, so it follows that $\Phi$ is surjective. To find the kernel:

$$\ker(\Phi) = \{(u, w) \in U \times W \mid u + w = 0\} = \{(u, -u) \in U \times W\} \cong U \cap W.$$ 

Finally, if $\ker(\Phi) = U \cap W = \{0\}$ then $\Phi$ is both injective and surjective, so it is an isomorphism. \hfill \Box

Proof of Proposition 2.7. Let $W = \text{span}\{w\}$ be the one-dimensional subspace of $V$ corresponding to $[w]$. Then since $[w] \notin \text{P}(U')$, it follows that $W \cap U' = \{0\}$. Counting dimensions, we deduce that

$$V \cong W \oplus U'$$

is an internal direct sum. Denote by $p : V \to U'$ the resulting projection map.

Now let $[u] \in \text{P}(U)$ and take $[u'] = \tau([u]) \in \text{P}(U')$. By definition of $\tau$, $[u']$ lies in the line determined by $[w]$ and $[u]$, so $u'$ is a linear combination

$$u' = \lambda w + \mu u$$

for some scalars $\lambda, \mu \in F$. Moreover, $[w] \in \text{P}(U')$ so $[w] \neq [u']$ and we deduce that $\mu \neq 0$, so

$$u = \frac{\lambda}{\mu} w - \frac{1}{\mu} u'$$

so $p(u) = -\frac{1}{\mu} u'$ and $\tau([u]) = [p(u)]$. It follows that $\tau$ is the projective transformation induced by the composition of linear maps $i : U \hookrightarrow V$ and $p : V \to U'$.

\hfill \Box

2.4.5 General position

We gain a firmer grasp of the freedom afforded by projective transformations using the following concept.

**Definition 6.** Let $P(V)$ be a projective space of dimension $n$. A set of $n + 2$ points in $P(V)$ are said to lie in general position if the representative vectors of any proper subset are linearly independent in $V$ (A subset $S' \subset S$ is called a proper subset if $S'$ is non-empty and $S'$ does not equal $S$).

**Example 6.** Any pair of distinct points in a projective line are represented by linearly independent vectors, so any three distinct points on a projective line lie in general position.

**Example 7.** Four points in a projective plane lie in general position if and only if no three of them is collinear (prove this).

**Theorem 2.9.** If $P(V)$ and $P(W)$ are projective spaces of the same dimension $n$ and both $p_1, \ldots, p_{n+2} \in P(V)$ and $q_1, \ldots, q_{n+2} \in P(W)$ lie in general position, then there is a unique projective transformation

$$\tau : P(V) \to P(W)$$

sending $\tau(p_i) = q_i$ for all $i = 1, 2, \ldots, n + 2$. 

13
Proof. Choose representative vectors $v_1, \ldots, v_{n+2} \in V$ such that $[v_i] = p_i$ for all $i$. Because the points $p_i$ lie in general position, it follows that the first $n + 1$ vectors $v_1, \ldots, v_{n+1}$ are linearly independent and thus form a basis of $V$. It follows that $v_{n+2}$ can be written as a linear combination

$$\sum_{i=1}^{n+1} \lambda_i v_i = v_{n+2}$$

Moreover, all of the scalars $\lambda_i$ are non-zero, for otherwise the non-trivial linear expression

$$\sum_{i=1}^{n+1} \lambda_i v_i - v_{n+2} = 0$$

would imply that some proper subset of $\{v_1, \ldots, v_{n+2}\}$ is linearly dependent, in contradiction with the points lying in general position. By replacing our representative vectors $v_i$ with the alternative representing vectors $\lambda_i v_i$, we may assume without loss of generality that

$$\sum_{i=1}^{n+1} v_i = v_{n+2}.$$

Similarly, we may choose representing vectors $w_i \in W$ for $q_i \in P(W)$ such that

$$\sum_{i=1}^{n+1} w_i = w_{n+2}.$$

Since $v_1, \ldots, v_{n+1} \in V$ are linearly independent, they form a basis for $V$ and we can define a linear map $T : V \to W$ by sending $T(v_i) = w_i$ for $i = 1, \ldots, n+1$ and extending linearly. By linearity we have

$$T(v_{n+2}) = T(\sum_{i=1}^{n+1} v_i) = \sum_{i=1}^{n+1} T(v_i) = \sum_{i=1}^{n+1} w_i = w_{n+2}.$$

So it induces a projective transformation satisfying $\tau(p_i) = q_i$ for all $i$.

To prove uniqueness, assume that there is another linear map $T' : V \to W$ satisfying the conditions of the theorem. Then for each $i$ there exists a non-zero scalar $\mu_i$ such that

$$T'(v_i) = \mu_i w_i.$$

But then by linearity

$$\mu_{n+2} w_{n+2} = T'(v_{n+2}) = T'(\sum_{i=1}^{n+1} v_i) = \sum_{i=1}^{n+1} \mu_i w_i$$

so

$$\sum_{i=1}^{n+1} w_i = w_{n+2} = \sum_{i=1}^{n+1} \frac{\mu_i}{\mu_{n+2}} w_i$$

and we deduce that $\frac{\mu_i}{\mu_{n+2}} = 1$ for all $i$ and thus that $T' = \mu T$ for some non-zero constant scalar $\mu$. \qed
Example 8. There is a unique projective transformation $\tau : FP^1 \to FP^1$ sending any ordered set of three points to any other ordered set of three points. In the context of complex analysis, we traditionally say that a mobius transformation is determined by where it sends $0, 1, \infty$.

Remark 2. In the proof of Theorem 2.9, we proved that for points $p_1, \ldots, p_{n+2} \in P(V)$ in general position, and given a representative vector $v_{n+2} \in V$ such that $[v_{n+2}] = p_{n+2}$, we may choose vectors $v_1, \ldots, v_{n+1} \in V$ with $[v_i] = p_i$ for $i = 1, \ldots, n+1$ and satisfying

$$\sum_{i=1}^{n+1} v_i = v_{n+2}.$$

This is a useful fact that we will apply repeatedly in the rest of the course.

2.5 Classical Theorems

2.5.1 Desargues’ Theorem

Desargues (1591-1661) a french mathematician, architect and engineer, is considered one of the founders of projective geometry. His work was forgotten for many years, until it was rediscovered in the 19th century during the golden age of projective geometry. The following result is named in his honour.

First, we adopt some notation commonly used in Euclidean geometry. Given distinct points $A, B$ in a projective space, let $AB$ denote the unique line passing through them both.

Theorem 2.10 (Desargues’ Theorem). Let $A, B, C, A', B', C'$ be distinct points in a projective space $P(V)$ such that the lines $AA', BB', CC'$ are distinct and concurrent (meaning the three lines intersect at a common point). Then the three points of intersection $X := AB \cap A'B'$, $Y := BC \cap B'C'$ and $Z := AC \cap A'C'$ are collinear (meaning they lie on a common line).

Proof. Let $P$ be the common point of intersection of the three lines $AA', BB'$ and $CC'$. First suppose that $P$ coincides with one of the points $A, B, C, A', B', C'$. Without loss of generality, let $P = A$. Then we have equality of the lines $AB = BB'$ and $AC = CC'$, so it follows that $X = AB \cap A'B' = B'$, $Z = AC \cap A'C' = C'$, so $B'C' = XZ$ must be collinear with $Y = BC \cap B'C'$ and the result is proven.

Now suppose that $P$ is distinct from $A, B, C, A', B', C'$. Then $P, A$ and $A'$ are three distinct points lying on a projective line, so they are in general position. By Remark 2, given a representative $p \in V$ of $P$, we can find representative vectors $a, a'$ of $A, A'$ such that

$$p = a + a'.$$

Similarly, we have representative vectors $b, b'$ for $B, B'$ and $c, c'$ for $C, C'$ such that

$$p = b + b', \quad p = c + c'$$
It follows that $a + a' = b + b'$ so

$$x := a - b = b' - a'$$

must represent the intersection point $X = AB \cap A'B'$, because $x$ is a linear combination of both $a, b$ and of $a', b'$. Similarly we get

$$y := b - c = c' - b' \quad \text{and} \quad z := c - a = a' - c'$$

representing the intersection points $Y$ and $Z$ respectively.

To see that $X, Y$ and $Z$ are collinear, observe that

$$x + y + z = (a - b) + (b - c) + (c - a) = 0$$

so $x, y, z$ are linearly dependent and thus must lie in a 2-dimensional subspace of $V$ corresponding to a projective line.

Desargues’ Theorem is our first example of a result that is much easier to prove using projective geometry than using Euclidean geometry. Our next result, named after Pappus of Alexandria (290-350) is similar.

### 2.5.2 Pappus’ Theorem

**Theorem 2.11.** Let $A, B, C$ and $A', B', C'$ be two pairs of collinear triples of distinct points in a projective plane. Then the three points $X = BC' \cap B'C$, $Y := CA' \cap C'A$ and $Z := AB' \cap A'B$ are collinear.

**Proof.** Without loss of generality, we can assume that $A, B, C'$, $B'$ lie in general position. If not, then two of the three required points coincide, so the conclusion is trivial. By Theorem 2.9, we can assume that

$$A = [1 : 0 : 0], \quad B = [0 : 1 : 0], \quad C' = [0 : 0 : 1], \quad B' = [1 : 1 : 1]$$

The line $AB$ corresponds to the two dimensional vector space spanned be $(1, 0, 0)$ and $(0, 1, 0)$ in $F^3$, so the point $C \in AB$ must have the form

$$C = [1 : c : 0]$$

(because it is not equal to $A$) for some $c \neq 0$. Similarly, the line $B'C'$ corresponds to the vector space spanned by $(0, 0, 1)$ and $(1, 1, 1)$ so

$$A' = [1, 1, a]$$

for some $a \neq 1$.

Next compute the intersection points of lines

$$BC' = \text{span}\{(0, 1, 0), (0, 0, 1)\} = \{(x_0, x_1, x_2) \mid x_0 = 0\}$$

$$B'C = \text{span}\{(1, 1, 1), (1, c, 0)\}$$
so we deduce that $BC' \cap B'C$ is represented by $(1, 1, 1) - (1, c, 0) = (0, 1 - c, 1)$.

$$AC' = \text{span}\{(1, 0, 0), (0, 0, 1)\} = \{x_1 = 0\}$$

$$A'C = \text{span}\{(1, 1, a), (1, c, 0)\}$$

so we deduce that $A'C \cap AC'$ is represented by $(1, c, 0) - c(1, 1, a) = (1 - c, 0, -ca)$.

$$AB' = \text{span}\{(1, 0, 0), (1, 1, 1)\} = \{x_1 = x_2\}$$

$$A'B = \text{span}\{(1, 1, a), (0, 1, 0)\}$$

so we deduce that $AB' \cap A'B$ is represented by $(a - 1)(0, 1, 0) + (1, 1, a) = (1, a, a)$.

Finally, we must check that the intersection points $[0 : 1 - c : 1], [1 - c : 0 : -ca]$ and $[1 : a : a]$ are collinear, which is equivalent to showing that $(0, 1 - c, 1), (1 - c, 0, -ca)$ and $(1, a, a)$ are linearly dependent. This can be accomplished by row reducing the matrix,

$$\begin{pmatrix}
1 & a & a \\
1 - c & 0 & -ca \\
0 & 1 - c & 1
\end{pmatrix} \Rightarrow \begin{pmatrix}
1 & a & a \\
0 & (c - 1)a & -a \\
0 & 1 - c & 1
\end{pmatrix} \Rightarrow \begin{pmatrix}
1 & a & a \\
0 & 0 & 0 \\
0 & 1 - c & 1
\end{pmatrix}$$

which has rank two, so $\text{span}\{(0, 1 - c, 1), (1 - c, 0, -ca), (1, a, a)\}$ has dimension two and the vectors are linearly dependent.

\[\square\]

### 2.6 Duality

In this section we explore the relationship between the geometry of $P(V)$ and that of $P(V^*)$ where $V^*$ denotes the dual vector space. Recall that given any vector space $V$, the dual vector space $V^*$ is the set of linear maps from $V$ to $F$

$$V^*: = \{\phi : V \rightarrow F\} = \text{Hom}(V, F).$$

- The dual $V^*$ is naturally a vector space under the operations

  $$(\phi_1 + \phi_2)(v) = \phi_1(v) + \phi_2(v)$$

  $$(\lambda \phi)(v) = \lambda(\phi(v)).$$

- Suppose that $v_1, ..., v_n \in V$ is a basis. Then the dual basis $\phi_1, ..., \phi_n \in V^*$ is the unique set of linear maps satisfying $\phi_i(v_j) = 1$, if $i = j$ and $\phi_i(v_j) = 0$ otherwise.

- If $T: V \rightarrow W$ is a linear map, then there is a natural linear map $T^*: W^* \rightarrow V^*$ (called the transpose) defined by

  $$T^*(f)(v) = f(T(v)).$$

Although the vector spaces $V$ and $V^*$ have the same dimension, there is no natural isomorphism between them. However there is a natural correspondence between subspaces.
Definition 7. Let $U \subseteq V$ be a vector subspace. The annihilator $U^0 \subseteq V^*$ is defined by

$$U^0 := \{ \phi \in V^* \mid \phi(u) = 0, \text{ for all } u \in U \}. $$

To check that $U^0$ is a subspace of $V^*$, observe that if $\phi_1, \phi_2 \in U^0$ and $u \in U$ then

$$(\lambda_1 \phi_1 + \lambda_2 \phi_2)(u) = \lambda_1 \phi_1(u) + \lambda_2 \phi_2(u) = 0 + 0 = 0$$

so $(\lambda_1 \phi_1 + \lambda_2 \phi_2) \in U^0$, so $U_0$ is closed under linear combinations.

Proposition 2.12. If $U_1 \subset U_2 \subset V$ then $U_1^0 \supset U_2^0$.

Proof. This is simply because if $\phi(v) = 0$ for all $v \in U_2$ then necessarily $\phi(v) = 0$ for all $v \in U_1$.

We also have the following result.

Proposition 2.13. $\dim(U) + \dim(U^0) = \dim(V)$

Proof. Choose a basis $u_1, \ldots, u_m \in U$ and extend to a basis $u_1, \ldots, u_m, v_{m+1}, \ldots, v_n$ of $V$. Let $\phi_1, \ldots, \phi_n$ be the dual basis. For a general element

$$\phi = \sum_{i=1}^n \lambda_i \phi_i \in V^*$$

we have $\phi(u_i) = \lambda_i$. Thus $\phi \in U^0$ if and only if $\lambda_i = 0$ for $i = 1, \ldots, m$ so $U^0 \subseteq V^*$ is the subspace with basis $\phi_{m+1}, \ldots, \phi_n$ which has dimension $n - m$.

Next, we explain the use of the word duality.

Proposition 2.14. There is a natural isomorphism $V \cong (V^*)^*$, defined by $v \mapsto S(v)$ where $S(v)(\phi) = \phi(v)$.

Proof. We already know that $\dim(V) = \dim(V^*) = \dim((V^*)^*)$, so (by ) it only remains to prove that $S : V \to (V^*)^*$ is injective. But if $v \in V$ is non-zero, we may extend $v_1 := v$ to a basis and then in the dual basis $\phi_1(v_1) = S(v_1)(\phi_1) = 1 \neq 0$ so $S(v) \neq 0$. Thus $\ker(S) = \{0\}$ and $S$ is injective.

Remark 3. Proposition 2.14 depends on $V$ being finite dimensional. If $V$ is infinite dimensional, then $S$ is injective but not surjective.

In practice, we tend to replace the phrase “naturally isomorphic” with “equals” and write $V = (V^*)^*$. Using this abuse of terminology, one may easily check that $(U^0)^0 = U$, so the operation of taking a vector space to its dual is reversible.

Proposition 2.15. Let $P(V)$ be a projective space of dimension $n$. Then there is a one-to-one correspondence between linear subspaces $P(U) \subset P(V)$ of dimension $m$ and linear subspaces of $P(V^*)$ of dimension $n - m - 1$ by the rule

$$P(U) \leftrightarrow P(U^0).$$

We use notation $P(U^0) = P(U)^0$. 18
Proof. By preceding remarks and Proposition there is a one-to-one correspondence

\[ U \leftrightarrow U^0 \]

between subspaces of dimension \( m + 1 \) in \( V \) and subspaces of dimension \( n - m \) in \( V^* \). The result follows immediately. \( \square \)

It follows from Proposition 2.15 that geometric statements about linear subspaces of \( P(V) \) translate into geometric statements about linear subspaces of \( P(V^*) \).

**Example 9.** Let \( P(V) \) be a projective plane, so \( P(V^*) \) is also a projective plane. There is a one-to-one correspondence between points in \( P(V) \) and lines in \( P(V^*) \). Similarly, there is one-to-one correspondence between lines in \( P(V) \) and points in \( P(V^*) \).

**Proposition 2.16.** Let \( A, B, C \) be collinear points in a projective plane lying on the line \( L \). Then \( A^0, B^0, C^0 \) are concurrent lines intersecting at the point \( L^0 \).

**Proof.** If \( A, B, C \subset L \) then \( L^0 \subset A^0, L^0 \subset B^0 \) and \( L^0 \subset C^0 \), so \( L^0 = A^0 \cap B^0 \cap C^0 \). \( \square \)

Duality can be used to establish “dual” versions of theorems. For example, simply by exchanging points with lines and collinear with concurrent, we get the following results.

**Theorem 2.17** (Dual version of Desargues’ Theorem). Let \( \alpha, \beta, \gamma, \alpha', \beta', \gamma' \) be distinct lines in a projective plane such that the points \( \alpha \cap \alpha', \beta \cap \beta' \) and \( \gamma \cap \gamma' \) are distinct and collinear. Then the lines joining \( \alpha \cap \beta \) to \( \alpha' \cap \beta' \), \( \alpha \cap \gamma \) to \( \alpha' \cap \gamma' \) and \( \beta \cap \gamma \) to \( \beta' \cap \gamma' \) are concurrent.

**Theorem 2.18** (Dual version of Pappus’ Theorem). Let \( \alpha, \beta, \gamma \) and \( \alpha', \beta', \gamma' \) be two sets of concurrent triples of distinct lines in a projective plane. Then the three lines determined by pairs of points \( \beta \cap \gamma' \) and \( \beta' \cap \gamma \), \( \gamma \cap \alpha' \) and \( \gamma' \cap \alpha \), \( \alpha \cap \beta' \) and \( \alpha' \cap \beta \) are concurrent.

In fact, with a little effort one can show that the statement of Pappus’ Theorem is equivalent to the Dual of Pappus’ Theorem (see Hitchin’s notes for details). Thus Pappus’ Theorem may be considered “self-dual”.

As a final result, we describe the set of lines in \( \mathbb{R}^2 \).

**Proposition 2.19.** The set of lines in \( \mathbb{R}^2 \) is in one-to-one correspondence with the Mobius strip.

**Proof.** The set of lines in \( \mathbb{R}^2 \) is the set of lines in \( \mathbb{R}P^2 \) minus the line at infinity. By duality, this is identified with \( \mathbb{R}P^2 \) minus a single point, which is the same as a mobius strip. Alternatively,
3 Quadrics and Conics

3.1 Affine algebraic sets

Given a set \( \{x_1, \ldots, x_n\} \) of variables, a polynomial \( f(x_1, \ldots, x_n) \) is an expression of the form

\[
f = f(x_1, \ldots, x_n) := \sum_I a_I x^I
\]

where

- the sum is over \((n+1)\)-tuples of non-negative integers \( I = (i_1, \ldots, i_n) \),
- \( x^I = x_1^{i_1} \cdots x_n^{i_n} \) are called monomials, and
- \( a_I \in F \) are scalars called coefficients, of which all but finitely many are zero.

The degree of a monomial \( x^I \) is the sum \(|I| = i_1 + \ldots + i_n\). The degree of a polynomial \( f \) equals the largest degree of a monomial occurring with non-zero coefficient.

Example 10.

- \( x_1 - x_2 \) is a polynomial of degree one.
- \( x^2 + y + 7 \) is a polynomial of degree two.
- \( xy^2 + 2x^3 - xyz + yz - 11 \) is a polynomial of degree three.

A polynomial \( f(x_1, \ldots, x_n) \) in \( n \) variables may be interpreted as a function

\[
f : F^n \to F
\]

by simply plugging in the entries \( n \)-tuples in \( F^n \) in place of the variables \( x_1, \ldots, x_n \).

Definition 8. Given a \( n \) variable polynomial \( f \), the set

\[
Z_{aff}(f) := \{ (x_1, \ldots, x_n) \in F^n \mid f(x_1, \ldots, x_n) = 0 \}
\]

is called the (affine) zero set of \( f \). More generally, if \( f_1, \ldots, f_k \) are polynomials then

\[
Z_{aff}(f_1, \ldots, f_k) := Z_{aff}(f_1) \cap \ldots \cap Z(f_k)
\]

is called the zero set of \( f_1, \ldots, f_k \). An affine algebraic set \( X \subseteq F^n \) is a set that equals \( Z_{aff}(f_1, \ldots, f_k) \) for some set of polynomials \( f_1, \ldots, f_k \).

Example 11. An affine line in \( F^2 \) is defined by a single equation \( ax + by + c = 0 \), where \( a \) and \( b \) are not both zero. Thus a line in \( F^2 \) is the zero set of a single degree one polynomial. More generally, an affine linear subset of \( F^n \) (points, lines, planes, etc.) is the zero set of a collection of degree one polynomials.

Example 12. The graph of a one variable polynomial \( f(x) \) is defined by the equation \( y = f(x) \), thus is equal to the algebraic set \( Z_{aff}(y - f(x)) \).
Example 13. Any finite set of points \( \{r_1, ..., r_d\} \subset F \) is an algebraic set \( Z(f) \) where \( f(x) = (x - r_1)(x - r_2)...(x - r_d) \).

Definition 9. An affine quadric \( X \subset F^n \) is a subset of the form \( X = Z_{aff}(f) \) where \( f \) is a single polynomial of degree 2. In the special case that \( X \subset F^2 \), we call \( X \) a (affine) conic.

Example 14. Some examples of affine in \( \mathbb{R}^2 \) conics include: circles, ellipses, hyperbolas and parabolas. The unit sphere \( Z_{aff}(x_1^2 + ... + x_n^2 - 1) \) is a quadric in \( \mathbb{R}^n \).

3.2 Projective algebraic sets

What is the right notion of an algebraic set in projective geometry? We begin with a definition.

Definition 10. A polynomial is called homogeneous if all non-zero monomials have the same degree.

Example 15. • \( x - y \) is a homogeneous polynomial of degree one.
• \( x^2 + y + 7 \) is a non-homogeneous polynomial.
• \( yz^2 + 2x^3 - xyz \) is a homogeneous polynomial of degree three.

Projective algebraic sets in \( FP^n \) are defined using homogeneous polynomials in \( n + 1 \) variables, which we normally denote \( \{x_0, x_1, ..., x_n\} \). The crucial property that makes homogenous polynomials fit into projective geometry is the following.

Proposition 3.1. Let \( f(x_0, ..., x_n) \) be a homogeneous polynomial of degree \( d \). Then for any scalar \( \lambda \in F \), we have
\[
 f(\lambda x_0, ..., \lambda x_n) = \lambda^d f(x_0, ..., x_n).
\]

It follows from Proposition 3.1 that for \( \lambda \) a non-zero scalar, \( f(x_0, ..., x_n) = 0 \) if and only if \( f(\lambda x_0, ..., \lambda x_n) = 0 \). This makes the following definition well-defined.

Definition 11. Let \( f_1, ..., f_k \) be a set of homogeneous polynomials in the variables \( x_0, ..., x_n \). The zero set
\[
 Z(f_1, ..., f_k) := \{[x_0 : ... : x_n] \in FP^n \mid f_i(x_0, ..., x_n) = 0 \text{ for all } i=1,...,k\}
\]
is a called a projective algebraic set.

The main example we have encountered so far are linear subsets of projective space. These are always of the form \( Z(l_1, ..., l_k) \) for some set of homogeneous linear polynomials \( l_1, ..., l_k \).

To understand the correspondence between projective and algebraic sets we recall that within \( FP^n \), we have a copy of the affine plane \( U_0 := \{[1 : y_1, ..., : y_n] \in FP^n \} \cong F^n \). Under this correspondence, we have an equality
$Z(f_1, ..., f_k) \cap U_0 = Z_{aff}(g_1, ..., g_k)$

where $g_i(x_1, ..., x_n) := f(1, x_1, ..., x_n)$.

Conversely, given an affine algebraic set we construct a projective algebraic set as follows. Let $f(x_1, ..., x_n)$ be a non-homogeneous polynomial of degree $d$. By multiplying the monomials of $f$ by appropriate powers of $x_0$, we can construct a homogeneous, degree $d$ polynomial $h(x_0, ..., x_n)$ such that $h(1, x_1, ..., x_n) = f(x_1, ..., x_n)$. We call $h$ the homogenization of $f$.

**Example 16.** Here are some examples of homogenizing polynomials

- $g(x) = 1 - x + 2x^2 \Rightarrow h(x, y) = y^2 - xy + 2x^2$
- $g(x, y) = y - x^2 + 1 \Rightarrow h(x, y, z) = yz - x^2 + z^2$
- $g(x_1, x_2) = x_1 + x_2^3 + 7 \Rightarrow h(x_0, x_1, x_2) = x_0^2 x_1 + x_2^3 + 7x_0^3$

Now suppose that $g_1, ..., g_k$ are non-homogeneous polynomials with homogenizations $h_1, ..., h_k$. Then we have equality

$$Z(h_1, ..., h_k) = U_0 \cap Z_{aff}(g_1, ..., g_k).$$

The study of algebraic sets is called algebraic geometry. It is a very active field of research today, of which projective geometry is a small part. So far we have only considered linear algebraic sets, where the polynomials all have order one. Our next big topic are algebraic sets of the form $Z(f)$ where $f$ is a degree two polynomial. Such algebraic sets are called quadrics.

### 3.3 Bilinear and quadratic forms

**Definition 12.** A symmetric bilinear form $B$ on a vector space $V$ is a map

$$B : V \times V \to F$$

such that

- $B(u, v) = B(v, u)$, (Symmetric)
- $B(\lambda_1 u_1 + \lambda_2 u_2, v) = \lambda_1 B(u_1, v) + \lambda_2 B(u_2, v)$, (Linear)

Observe that by combining symmetry and linearity, a bilinear form also satisfies

$$B(u, \lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 B(u, v_1) + \lambda_2 B(u, v_2).$$

We say that $B$ is linear in both entries, or bilinear. We say that $B$ is nondegenerate if for each non-zero $v \in V$, there is a $w \in V$ such that $B(v, w) \neq 0$. 

22
Example 17. The standard example of a symmetric bilinear form is the dot product

\[ B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \quad B((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \sum_{i=1}^n x_i y_i. \]

The dot product is nondegenerate because \( B(v, v) = |v|^2 > 0 \) if \( v \neq 0 \).

Example 18. The bilinear form \( B : \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R} \) defined by

\[ B((x_0, x_1, x_2, x_3), (y_0, y_1, y_2, y_3)) = -x_0 y_0 + x_1 y_1 + x_2 y_2 + x_3 y_3 \]

is called the Minkowski metric that measures lengths in space-time according to special relativity. In contrast with the dot product, the equation \( B(v, v) = 0 \) holds for some non-trivial vectors called null-vectors. The null-vectors form a subset of \( \mathbb{R}^4 \) called the lightcone dividing those vectors satisfying \( B(v, v) < 0 \) (called time-like vectors) from those satisfying \( B(v, v) > 0 \) (called space-like vectors).

Now let \( v_1, \ldots, v_n \in V \) be a basis. For any pair of vectors \( u = x_1 v_1 + \ldots x_n v_n \) and \( w = y_1 v_1 + \ldots + y_n v_n \) we have

**Lemma 3.2.** Let \( v_1, \ldots, v_n \in V \) be a basis. For any pair of vectors \( u = x_1 v_1 + \ldots x_n v_n \) and \( w = y_1 v_1 + \ldots + y_n v_n \) we have

\[ B(u, w) = \sum_{i,j=1}^n x_i y_j B(v_i, v_j). \]

This means that the bilinear form is completely determined by the \( n \times n \)-matrix \( \beta \) with entries \( \beta_{i,j} = B(v_i, v_j) \).

**Proof.**

\[
B(u, w) = B(x_1 v_1 + \ldots x_n v_n, w) \\
= \sum_{i=1}^n x_i B(v_i, w) = \sum_{i=1}^n x_i B(v_i, y_1 v_1 + \ldots + y_n v_n) \\
= \sum_{i=1}^n \sum_{j=1}^n x_i y_j B(v_i, v_j) = \sum_{i,j=1}^n x_i y_j \beta_{i,j}
\]

\( \square \)

Lemma 3.2 provides a nice classification of bilinear forms: Given a vector space \( V \) equipped with a basis \( v_1, \ldots, v_n \), then there is a one-to-one correspondence between symmetric bilinear forms \( B \) on \( V \) and symmetric \( n \times n \) matrices \( \beta \), according to the rule

\[ B(v, w) = \bar{v}^T \beta \bar{w} \]

where \( \bar{v}, \bar{w} \) denote the vectors \( v, w \) expressed as column vectors and \( T \) denotes transpose, so that \( \bar{v}^T \) is \( v \) as a row vector.
Proposition 3.3. The bilinear form $B$ is non-degenerate if and only if the matrix $\beta$ has non-zero determinant.

Proof. Suppose that $\det(\beta)$ is zero if and only if there exists a nonzero row vector $\bar{v}^T$ such that $\bar{v}^T \beta = 0$ if and only if $\bar{v}^T \beta \bar{w} = 0$ for all column vectors $\bar{w}$ if and only if there exists $v \in V \setminus \{0\}$ such that $B(v, w) = 0$ for all $w \in V$ if and only if $B$ is degenerate. \(\square\)

3.3.1 Quadratic forms

Given a symmetric bilinear form $B : V \times V \to F$, we can associate a function called a quadratic form $Q : V \to F$ by the rule $Q(v) = B(v, v)$.

Proposition 3.4. Let $V$ be a vector space over $F = \mathbb{R}$ or $\mathbb{C}$, let $B$ be a symmetric bilinear form on $V$ and let $Q : V \to F$ be the associated quadratic form. Then for any $u, v \in V$ we have

$$B(u, v) = (Q(u + v) - Q(u) - Q(v))/2.$$  

In particular, $B$ is completely determined by its quadratic form $Q$.

Proof. Using bilinearity and symmetry, we have

$$Q(u + v) = B(u + v, u + v) = B(u, u) + B(u, v) + B(v, u) + B(v, v) = Q(u) + Q(v) + 2B(u, v)$$

from which the result follows. \(\square\)

Remark 4. Proposition ... does not hold when $F$ is the field of Boolean numbers because we are not allowed to divide Boolean numbers by two!

We can understand these relationship better using a basis. Let $v_1, ..., v_n$ for $V$ be a basis so $B$ can be expressed in terms of the $n \times n$ matrix $\beta$. According to ...

$$Q(x_1v_1 + ... + x_nv_n) = B(x_1v_1 + ... + x_nv_n, x_1v_1 + ... + x_nv_n) = \sum_{i,j} \beta_{i,j}x_ix_j$$

is a homogeneous quadratic polynomial.

Some examples:

$$\begin{bmatrix} a \end{bmatrix} \leftrightarrow ax^2$$

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} \leftrightarrow ax^2 + 2bxy + cy^2$$

$$\begin{bmatrix} a & b & d \\ b & c & e \\ d & e & f \end{bmatrix} \leftrightarrow ax^2 + 2bxy + cy^2 + 2dxz + 2exz + fz^2$$

24
3.3.2 Change of basis

We’ve seen so far that a symmetric bilinear form $B$ can be described using a symmetric matrix with respect to a basis. What happens if we change the basis?

Let $v_1, \ldots, v_n$ and $w_1, \ldots, w_n$ be two different bases of $V$. Then we kind find scalars $\{P_{i,j}\}_{i,j=1}^n$ such that

$$w_j = \sum_i P_{i,j} v_i \quad (3)$$

The matrix $P$ with entries $P_{i,j}$ is called the change of basis matrix. Observe that $P$ is invertible, with $P^{-1}$ providing the change of matrix from $w_1, \ldots, w_n$ back to $v_1, \ldots, v_n$. In fact, if $P'$ is an invertible matrix then it is the change of basis matrix from $v_1, \ldots, v_n$ to the new basis $w_1, \ldots, w_n$ satisfying (3).

**Proposition 3.5.** Let $B$ be a bilinear form on $V$, and let $\beta$, $\beta'$ be symmetric matrices representing $B$ with respect to bases $v_1, \ldots, v_n$ and $w_1, \ldots, w_n$ respectively. Then

$$\beta' = P^T \beta P$$

where $P$ is the change of basis matrix.

**Proof.** There are two approaches. Simply calculate:

$$\beta'_{i,j} = B(w_i, w_j) = B\left(\sum_k P_{k,i} v_k, \sum_l P_{l,j} v_l\right)$$

$$= \sum_{k,l} P_{k,i} P_{l,j} \beta_{k,l} = \sum_{k,l} P_{i,k}^T \beta_{k,l} P_{l,j}$$

$$= (P^T \beta P)_{i,j}$$

Alternatively, we can argue more conceptually as follows. Let $e_i$ denote the $i$th standard basis column vector. Then in terms of the basis $v_1, \ldots, v_n$, $w_i$ corresponds to the column vector $Pe_i$. Then

$$\beta'_{i,j} = B(w_i, w_j) = (Pe_i)^T \beta(Pe_j) = e_i^T P^T \beta P e_j = (P^T \beta P)_{i,j}.$$ 

Our next task is to classify bilinear forms up to change of basis. First we need a lemma.

**Lemma 3.6.** Let $V$ be a vector space of dimension $n$ equipped with a symmetric bilinear form $B$. For $w \in V$ define

$$w^\perp := \{v \in V | B(v, w) = 0\}.$$ 

Then $w^\perp \subset V$ is a vector space of dimension $\geq n - 1$.  

25
Proof. Consider the map \( \phi : V \to F \), defined by \( \phi(v) = B(v, w) \). Then \( \phi \) is linear because
\[
\phi(\lambda_1 v_1 + \lambda_2 v_2) = B(\lambda_1 v_1 + \lambda_2 v_2, w) = \lambda_1 B(v_1, w) + \lambda_2 B(v_2, w) = \lambda_1 \phi(v_1) + \lambda_2 \phi(v_2)
\]
so \( w^\perp \) is equal to the kernel of \( \phi \), hence it is a vector subspace. By the rank nullity theorem,
\[
\dim(w^\perp) = \dim(V) - \dim(\text{im}(\phi)) = \begin{cases} n & \text{im}(\phi) = \{0\} \\ n - 1 & \text{im}(\phi) = F. \end{cases}
\]

\[ \square \]

**Theorem 3.7.** Let \( B \) be a symmetric, bilinear form over a vector space \( V \). Then there exists a basis \( v_1, \ldots, v_n \) such that
\[
B(v_i, v_j) = 0 \tag{4}
\]
for \( i \neq j \) and such that for all \( i = 1, \ldots, n \)
\[
B(v_i, v_i) \in \begin{cases} 0, 1 & F = \mathbb{C} \\ 0, \pm 1 & F = \mathbb{R}. \end{cases} \tag{5}
\]

**Proof.** As a first step construct a basis satisfying (4) using induction on the dimension of \( V \). In the base case \( \dim(V) = 1 \), (4) holds vacuously.

Now suppose by induction that all \( n - 1 \) dimensional symmetric bilinear forms can be diagonalized. Let \( V \) be \( n \) dimensional. If \( B \) is trivial, then \( \beta \) is the zero matrix with respect to any basis (hence diagonal).

If \( B \) is non-trivial then by Proposition 3.4, there exists a vector \( w_n \in V \) such that \( B(w_n, w_n) \neq 0 \). Clearly \( w_n \notin v^\perp_n \), so \( w^\perp_n \) is \( n - 1 \) dimensional. By induction we may choose a basis \( w_1, \ldots, w_{n-1} \in w^\perp_n \) satisfying (4) for all \( i, j < n \). Furthermore, for all \( i < n \), \( B(w_i, w_n) = 0 \) because \( w_i \in w^\perp_n \). Thus \( w_1, \ldots, w_n \) satisfy (4).

Finally, we want to modify \( w_1, \ldots, w_n \) to satisfy (5). Suppose that \( B(w_i, w_i) = c \in F \).
If \( c = 0 \), then let \( v_i = w_i \). If \( c \neq 0 \) and \( F = \mathbb{C} \), let \( v_i = \frac{1}{\sqrt{c}} w_i \) so that
\[
B(v_i, v_i) = B(\frac{1}{\sqrt{c}} w_i, \frac{1}{\sqrt{c}} w_i) = \left( \frac{1}{\sqrt{c}} \right)^2 B(w_i, w_i) = \frac{c}{c} = 1.
\]
If \( c \neq 0 \) and \( F = \mathbb{R} \) then let \( v_i := \frac{1}{\sqrt{|c|}} w_i \) so
\[
B(v_i, v_i) = \frac{c}{|c|} = \pm 1.
\]

\[ \square \]

**Corollary 3.8.** Let \( V \) and \( B \) be as above and let \( Q(v) = B(v, v) \). Then

- If \( F = \mathbb{C} \) then for some \( m \leq n \) there exists a basis such that if \( v = \sum_{i=1}^{n} z_i v_i \)
\[
Q(v) = \sum_{i=1}^{m} z_i^2
\]
• If \( F = \mathbb{R} \) then for some \( p, q \) with \( p + q \leq n \) there exists a basis such that

\[
Q(v) = \sum_{i=1}^{p} z_i^2 - \sum_{i=p+1}^{p+q} z_i^2
\]

Proof. Choose the basis \( v_1, ..., v_n \) so that \( B(v_i, v_j) \) satisfy the equations (4) and (5). Then

\[
Q(z_1v_1 + ... + z_nv_n) = B(z_1v_1 + ... + z_nv_n, z_1v_1 + ... + z_nv_n) = \sum_{i,j=1}^{n} z_iz_j B(v_i, v_j) = \sum_{i=1}^{n} z_i^2 B(v_i, v_i),
\]

where the \( B(v_i, v_i) \) are 0, 1 or \(-1\) as appropriate.

Exercise: Show that the numbers \( m, p, q \) occurring above are independent of the choice of basis.

It follows that \( B \) and \( Q \) are non-degenerate if and only if \( m = n \) or \( m = p + q \). In the case of a real vector space, we say that \( B \) is positive definite if \( p = n \), negative definite if \( q = n \) and indefinite otherwise.

### 3.3.3 Digression on the Hessian

This discussion will not be on the test or problem sets.

An important place where symmetric bilinear forms occur in mathematics is the study
of critical points of differentiable functions. Let \( f(x_1, ..., x_n) \) be a differentiable function
and let \((a_1, ..., a_n)\) be a point at which all partial derivatives vanish:

\[
f_{x_i}(a_1, ..., a_n) = \frac{\partial f}{\partial x_i}|_{(a_1, ..., a_n)} = 0, \quad \text{for all } i = 1, ..., n
\]

If the the second order partials are defined and continuous then the \( n \times n \)-matrix \( H \) with
entries

\[
H_{i,j} = f_{x_ix_j}(a_1, ..., a_n) = \frac{\partial^2 f}{\partial x_j \partial x_i}|_{(a_1, ..., a_n)}
\]

is a symmetric matrix (\( H_{i,j} = H_{j,i} \)) defining a symmetric bilinear form on the set of
tangent vectors called the Hessian. The corresponding quadratic form is equal to the
second order Taylor polynomial of the function.

The Hessian can be used to determine if \( f \) has a local maximum or minimum at
\((a_1, ..., a_n)\) (using the Second Derivative Test). Specifically, if \( H \) is non-degenerate then
\( f \) has a local maximum at \((a_1, ..., a_n)\) if \( H \) is negative definite, has a local minimum if \( H \)
is positive definite, and has neither a max or min if \( H \) is neither positive nor negative
definite.

Similarly, there is a notion of a complex valued function of complex variables being
differentiable (also called analytic). Reasoning as above one can show that for a complex-
valued function \( f \), \(|f|\) has no local maximums where \( f \) is differentiable. This is a version
of the maximum principle.
3.4 Quadrics and Conics

We now can state a coordinate independent definition of a quadric.

**Definition 13.** A *quadric* $X \subset P(V)$ is a subset defined by

$$X = \{ [v] \in V | B(v, v) = 0 \}.$$ 

By convention, we exclude the trivial case $B = 0$.

Is the case $P(V) = FP^n$, this definition is equivalent to $X = Z(f)$ for some homogeneous degree two polynomial $f$.

Consider now the implications of Corollary 3.8 for quadrics in a projective line. Over the reals we have such a quadric has the form $Z(f) = Z(-f)$ where (up to a change of coordinates) $f$ equals

- $x^2$
- $x^2 + y^2$
- $x^2 - y^2$

In the first case, $Z(x^2)$ consists of the single point $[0 : 1]$ which is counted with “multiplicity” two. The case $Z(x^2 + y^2)$ is empty set because there are no solutions with both $x$ and $y$ non zero. Finally $Z(x^2 - y^2) = Z((x - y)(x + y))$ has two solutions: $[1 : 1]$ and $[1 : -1]$. These three possibilities correspond to a familiar fact about roots of quadratic polynomials $0 = ax^2 + bx + c$: such a polynomial can have two roots, one root, or no roots. A consequence of Corollary 3.8 is that up to projective transformations, the number of roots completely characterizes the quadric.

Over the complex numbers, the classification is even simpler: a quadric has the form $Z(f)$ where $f$ is one of (up to coordinate change)

- $x^2$
- $x^2 + y^2$

The first case $Z(x^2)$ is a single point of multiplicity 2, while $Z(x^2 + y^2) = Z((x - iy)(x + iy))$ consists of the two points $[i : 1]$ and $[i : -1]$. This corresponds to the fact that over the complex numbers $0 = ax^2 + bx + c$ with $a \neq 0$ has either two distinct roots or a single root of order two (complex polynomials in one variable can always be factored!).

We now consider the case $F = \mathbb{R}$ and $P(V)$ is a projective plane (in this case a quadric is called a conic). We may choose coordinates so that $X = Z(f)$ where $f$ is one of:

- $x$,
- $x^2 + y^2$,
- $x^2 + y^2 + z^2$, 

28
- $x^2 - y^2$,
- $x^2 + y^2 - z^2$,

up to multiplication by $-1$ which doesn’t change $Z(f) = Z(-f)$.

The first example

$$x^2 = xx = 0$$

has the same solution set as $x = 0$, which is projective line $\{(0 : a : b)\} \cong \mathbb{RP}^1$, however this line is counted twice in a sense we won’t make precise (similar two a double root of single variable polynomial). The second example

$$x^2 + y^2 = 0$$

has a single solution $[0 : 0 : 1]$. The third example $x^2 + y^2 + z^2 = 0$, has no projective solutions (remember $(0, 0, 0)$ does not represent a projective point!). The example

$$x^2 - y^2 = (x - y)(x + y)$$

factors as a product of linear polynomials. It follows that $Z(x^2 - y^2)$ is a union of two distinct lines intersecting at one point (a degenerate conic).

The last example $Z(x^2 + y^2 - z^2)$ is (up to coordinate change) the only non-empty and non-degenerate conic, up to projective transformation. The apparent difference between the affine conics: ellipses, parabolas, hyperbolas, can be understood as how the conic relates to the line at infinity. A conic looks like an ellipse if it is disjoint form the line at infinity, looks like parabola if it is tangent to the line at infinity, and looks like a hyperbola if it intersects the line at infinity at two points.

To illustrate, consider what happens when we set $z = 1$ and treat $x$ and $y$ like affine coordinates (so the equation $z = 0$ is the “line at infinity”). Then the polynomial equation $x^2 + y^2 - z^2 = 0$ becomes $x^2 + y^2 - 1 = 0$ or

$$x^2 + y^2 = 1$$

which defines the unit circle in the x-y plane.

Now instead, set $y = 1$ and consider $x, z$ as affine coordinates (so $y = 0$ defines the line at infinity). Then we get the equation

$$x^2 + 1 - z^2 = 0,$$

or

$$x^2 - z^2 = (x - z)(x + z) = 1$$

which is the equation of a hyperbola asymptotic to the lines $x = z$ and $x = -z$.

Finally, consider performing the linear change of variables $z' = y + z$. In these new coordinates, the equation becomes

$$x^2 + y^2 - (z' - y)^2 = x^2 + 2z'y - z'^2 = 0.$$
Passing to affine coordinates $x, y$ by setting $z' = 1$ we get

$$x^2 + 2y - 1 = 0$$

or

$$y = -x^2/2 + 1/2$$

which defines a parabola.

In case $F = \mathbb{C}$ then the possible conics is even more restricted. Up to change of coordinates, the possible quadratics in the complex projective plane are defined by the quadratic polynomial equations

- $x^2 = 0$,
- $x^2 + y^2 = 0$,
- $x^2 + y^2 + z^2 = 0$.

The equation $x^2 = 0$ defines a projective line that is counted twice or has multiplicity two (in a sense we won’t make precise). The equation

$$x^2 + y^2 = (x + iy)(x - iy) = 0$$

defines the union of two distinct lines determined by the equations $x + iy = 0$ and $x - iy = 0$. The last equation $x^2 + y^2 + z^2 = 0$ defines the non-degenerate quadric curve. We study what this curve looks like in the next section.

### 3.5 The rational parametrization of the circle

**Theorem 3.9.** Let $C$ be a non-degenerate conic in a projective plane $P(V)$ over the field $F$, and let $A$ be a point on $C$. Let $P(U) \subset P(V)$ be a projective line not containing $A$. Then there is a bijection

$$\alpha : P(U) \to C$$

such that, for $X \in P(U)$, the points $A, X, \alpha(X)$ are collinear.

**Proof.** Suppose the conic is defined by the symmetric bilinear form $B$. Let $a \in V$ represent the fixed point $A$. Because $A \in X$, we know that $B(a, a) = 0$. Let $x \in V$ represent a point $X \in P(U)$. Because $A \not\in P(U)$, it follows that $X \neq A$ so the vectors $a, x$ are linearly independent, so we can extend them to a basis $a, x, y \in V$.

The restriction of $B$ to the span of $a, x$ is not identically zero. If it were then the matrix for $B$ associated to the basis $a, x, y$ would be of the form

$$
\begin{pmatrix}
0 & 0 & * \\
0 & 0 & * \\
* & * & *
\end{pmatrix}
$$

which has determinant zero, which is impossible because $B$ is non-degenerate. It follows that either $B(a, x)$ or $B(x, x)$ is non-zero.
An arbitrary point on the line $AX$ has the form $[\lambda a + \mu x]$ for some scalars $\lambda$ and $\mu$. Such a point lies on the conic $C$ if and only if

$$B(\lambda a + \mu x, \lambda a + \mu x) = \lambda^2 B(a, a) + 2\lambda\mu B(a, x) + \mu^2 B(x, x)$$

$$= \mu(2\lambda B(a, x) + \mu B(x, x)) = 0$$

for which there are two solutions in the homogeneous coordinates $\lambda$ and $\mu$: The solution $\mu = 0$ corresponding the point $A$ and solution to $2\lambda B(a, x) + \mu B(x, x) = 0$ corresponding to the point represented by the vector

$$w := 2B(a, x)x - B(x, x)a$$

which is non-zero because one of the coefficients $2B(a, x)$ or $B(x, x)$ is non-zero.

We define the map $\alpha : P(U) \to C$ by

$$\alpha(X) = [w].$$

It remains to prove that $\alpha$ is a bijection. If $Y$ is a point on $C$ distinct from $A$, then by Theorem 2.3 there is a unique point of intersection between $AY$ and $P(U)$ so

$$\alpha^{-1}(Y) = AY \cap P(U)$$

so the function is invertible on the set $C \setminus \{A\}$. To determine $\alpha^{-1}(A)$, observe that

$$\alpha(X) = [2B(a, x)x - B(x, x)a] = A$$

if and only if $B(a, x) = 0$, if and only $x \in a^\perp \cap U$. Since $a^\perp$ is 2-dimensional (Lemma 3.6), and $a^\perp \neq U$ ( $a \in a^\perp$ but not in $U$), then $a^\perp \cap U$ is an intersection of distinct 2-dimensional subspaces of a 3-dimensional space and hence is 1-dimensional and represents a single point $X = P(a^\perp \cap U)$.

A consequence of Theorem 3.9 is that a real projective conic is “homeomorphic” to a real projective line, which in turn is homeomorphic to a circle - a fact that we’ve already seen. A second consequence is that the complex projective conic is homeomorphic to $\mathbb{C}P^1$ which in turn is homeomorphic to the two sphere $S^2$.

Let us now consider what the map $\alpha$ looks like in coordinates in a special case. Let $C = Z(x^2 + y^2 - z^2)$ corresponding to the matrix

$$\beta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$ 

Choose $A = [1 : 0 : 1]$ and let $P(U) = Z(x)$. Then points $X \in P(U)$ have the form $[0 : 1 : t]$ or $[0 : 0 : 1]$ and the map $\alpha : P(U) \to C$ satisfies

$$\alpha([0 : 1 : t]) = [2B((1, 0, 1), (0, 1, t))(0, 1, t) - B((0, 1, t), (0, 1, t))(1, 0, 1)]$$

$$= [-2t(0, 1, t) - (1 - t^2)(1, 0, 1)]$$

$$= [(t^2 - 1, -2t, -1 - t^2)]$$

$$= [1 - t^2 : 2t : 1 + t^2]$$

31
and

$$\alpha([0 : 0 : 1]) = [2B((1, 0, 1), (0, 0, 1))(0, 0, 1) - B((0, 0, 1), (0, 0, 1))(1, 0, 1)]$$

$$= [-2(0, 0, 1) + 1(1, 0, 1)]$$

$$= [1 : 0 : -1]$$

These formulas can be used to parametrization of the circle using rational functions. For $t$ a real number $1 + t^2 > 0$ is non-zero so we can divide

$$[1 - t^2 : 2t : 1 + t^2] = [\frac{1 - t^2}{1 + t^2} : \frac{2t}{1 + t^2} : 1]$$

and we gain a parametrization of the unit circle in affine co-ordinates $\alpha' : \mathbb{R} \to S^1 = \mathbb{Z}_{aff}(x^2 + y^2 - 1)$

$$\alpha'(t) = \left(1 - \frac{t^2}{1 + t^2}, \frac{2t}{1 + t^2}\right)$$

That hits all points except $(-1, 0)$.

These formulas have an interesting application in the case $F = \mathbb{Q}$ is the field of rational numbers (those of the form $a/b$ where $a$ and $b$ are integers).

Since we have focused on the cases $F = \mathbb{R}$ and $F = \mathbb{C}$ so far, we take a moment to discuss informally linear algebra and projective geometry over $F = \mathbb{Q}$. To keep matters simple, we concentrate on the standard dimension $n$ vector space over $\mathbb{Q}$, $\mathbb{Q}^n$ consisting of $n$-tuples of rational numbers. It is helpful geometrically to consider $\mathbb{Q}^n$ as a subset of $\mathbb{R}^n$ in the standard way. Elements in $\mathbb{Q}^n$ can be added in the usual way and can be multiplied by scalars from $\mathbb{Q}$, so it makes sense to form linear combinations

$$\lambda_1 v_1 + \ldots + \lambda_k v_k \in \mathbb{Q}^n,$$

where $\lambda_1, \ldots, \lambda_k \in \mathbb{Q}$ and $v_1, \ldots, v_k \in \mathbb{Q}^n$. The span of a set of vectors $span(v_1, \ldots, v_k) \subset \mathbb{Q}^n$ is the set of all linear combinations of $v_1, \ldots, v_k$. In particular, a one-dimensional subspace of $\mathbb{Q}^n$ is the span of a single non-zero vector. We define the projective space

$$\mathbb{Q}P^n = P(\mathbb{Q}^{n+1})$$

to be the set of one dimensional subspaces in $\mathbb{Q}^{n+1}$. Geometrically, we can regard

$$\mathbb{Q}P^n \subset \mathbb{R}P^n$$

consisting of those projective points that can be represented by $[v] \in \mathbb{R}P^n$ where $v \in \mathbb{Q}^{n+1}$.

Many of the results about projective geometry we have developed so far generalize immediately to geometry in $\mathbb{Q}P^n$ (one notable exception is the classification of symmetric bilinear forms because the proof depends on taking square roots but the square root of a rational number is not necessarily rational!).

In particular the formulas parametrizing the circle remain valid. That is, solutions to the equation

$$x^2 + y^2 = z^2$$
for \( x, y, z \in \mathbb{Q} \) must be of the form
\[
[x : y : z] = [1 - t^2 : 2t : 1 + t^2]
\]
for some rational number \( t \in \mathbb{Q} \). If we let \( t = a/b \) where \( a \) and \( b \) are integers sharing no common factors, then
\[
[x : y : z] = [1 - (a/b)^2 : 2(a/b) : 1 + (a/b)^2] = [b^2 - a^2 : 2ab : b^2 + a^2].
\]
It follows in particular that every integer solution to the equation \( x^2 + y^2 = z^2 \) must have the form
\[
\begin{align*}
x &= c(b^2 - a^2), \\
y &= c(2ab), \\
z &= c(b^2 + a^2)
\end{align*}
\]
for some integers \( a, b, c \). Such solutions are called Pythagorean triples because they define right triangles with integer side lengths. For example, \( a = c = 1, b = 2 \) determines \( x = 3, y = 4, z = 5 \), while \( a = 2, b = 3, c = 1 \) determines \( x = 5, y = 8, z = 13 \).

### 3.6 Polars

Let \( V \) be a vector space and \( B \) a symmetric, bilinear form on \( V \). Given a subset \( S \subset V \), define
\[
S^\perp := \{ v \in V | B(s, v) = 0, \text{ for all } v \in V \}.
\]
Observe that \( S^\perp \) must be a vector subspaces, because if \( v_1 \in S^\perp \) and \( v_2 \in S^\perp \) then for any scalars \( \lambda_1, \lambda_2 \in F \) and \( s \in S \) we have
\[
B(s, \lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 B(s, v_1) + \lambda_2 B(s, v_2) = 0
\]
so \( \lambda_1 v_1 + \lambda_2 v_2 \in S \). Indeed, by a similar linearity argument it can be shown that
\[
S^\perp = \text{span}(S)^\perp
\]
so the operation of applying \( \perp \) can be thought of as assigning to a vector space \( U \subset V \) its orthogonal or perpendicular subspace \( U^\perp \) (relative to \( B \) of course).

**Proposition 3.10.** If \( B \) is a non-degenerate bilinear form, then for any subspace \( U \subset V \)

- \( (U^\perp)^\perp = U \)
- \( U_1 \subset U_2 \Rightarrow U_2^\perp \subset U_1^\perp \)
- \( \dim(U) + \dim(U^\perp) = \dim(V) \)
Proof. Let $V^* = \text{Hom}(V, F)$ denote the dual vector space and consider the map
\[ \phi : V \rightarrow V^* \]
defined by
\[ \phi(v)(w) = B(v, w). \]
This map is linear because
\[
\phi(\lambda v_1 + \lambda v_2)(w) = B(\lambda v_1 + \lambda v_2, w) \\
= \lambda_1 B(v_1, w) + \lambda_2 B(v_2, w) \\
= \lambda_1 \phi(v_1)(w) + \lambda_2 \phi(v_2)(w)
\]
so
\[ \phi(\lambda v_1 + \lambda v_2) = \lambda_1 \phi(v_1) + \lambda_2 \phi(v_2). \]
Because $B$ is non-degenerate, if $v \neq 0$ then there exists $w \in V$ such that $\phi(v)(w) = B(v, w) \neq 0$. Thus $\ker(\phi) = 0$ and $\phi$ is injective. By ... $\dim(V) = \dim(V^*)$ so $\phi$ is also surjective and hence an isomorphism between $V$ and $V^*$.

Finally
\[ \phi(U^\perp) = \{ \xi \in V^*| \xi(u) = 0 \text{ for all } u \in U \} = U^0. \]
Thus $U^\perp$ is simply the bijective image of the annihilator, so all the conclusions follow directly from Section 2.6.

**Definition 14.** Let $B$ be a non-degenerate symmetric bilinear form on $V$ and let $P(U) \subset P(V)$ be a linear subspace. Then the polar of $P(U)$ is the linear subspace $P(U^\perp)$.

It follows from Section 2.6 that if $P(V)$ has dimension $m$ and $P(U) \subset P(V)$ has dimension $m$ then $P(U^\perp)$ has dimension $n - m - 1$. If $P(V)$ is a projective plane, then the polar of point is a line and the polar of a line is a point.

**Proposition 3.11.** Let $C$ be a non-degenerate conic in a complex projective plane $P(V)$. Then

(i) Each line in the plane meets the conic in one or two points

(ii) If $P \in C$, its polar line is the unique tangent to $C$ passing through $P$

(iii) If $P \notin C$, the polar line of $P$ meets $C$ in two points, and the tangents to $C$ at these points intersect at $P$.

Proof. Let $U \subset V$ be the 2-dimensional subspace defining the projective line $P(U)$ and let $u, v \in U$ be a basis for $U$. A general point in $P(U)$ has the form $[\lambda u + \mu v]$ for scalars $\lambda, \mu \in \mathbb{C}$ not both zero. A point $[\lambda u + \mu v]$ lies on the conic $C$ if

\[ 0 = B(\lambda u + \mu v, \lambda u + \mu v) = \lambda^2 B(u, u) + 2\lambda \mu B(u, v) + \mu^2 B(v, v). \]  \hfill (6)
We think of \( \lambda \) and \( \mu \) as homogeneous coordinates on the projective line, so that (6) defines the roots of a quadratic polynomial. Over the complex numbers, (6) can be factored

\[
0 = (a\lambda - b\mu)(a'\lambda + b'\mu)
\]
determining two (possibly coincident) points of intersection between the line and the conic

\[
[bu + av], \quad [b'u + a'v],
\]
establishing (i).

To simplify equation (6), we choose \( u \) so that \([u]\) is one of the intersection points in \( C \cap P(U) \). Equivalently, \( B(u, u) = 0 \), so (6) becomes

\[
0 = 2\lambda \mu B(u, v) + \mu^2 B(v, v) = \mu(2\lambda B(u, v) + \mu B(v, v)).
\]

This equation has two (possibly coincident) solutions: \( \mu = 0 \) corresponding to \([u]\) and \( 2\lambda B(u, v) + \mu B(v, v) \) corresponding to \([w] = [B(v, v)u + 2B(u, v)v] \).

Observe that \([u] = [w]\) if and only if:

\[
[u] = [w] \iff B(u, v) = 0 \\
\iff B(u, \lambda u + \mu v) = 0 \text{ for all points } \lambda u + \mu v \in U \\
\iff P(U) \text{ is the polar of } [u].
\]
establishing (ii).

Now let \([a]\) denote a point not lying on \( C \), and let \( P(U) \) denote the polar of \([a]\). From the calculation above, it follows that \( P(U) \) meets \( C \) in two distinct points \([u]\) and \([w]\) with \([u] \neq [w]\). Because \([u]\) lies on \( C \) we know \( B(u, u) = 0 \). Because \([u]\) lies on the polar of \([a]\) we know \( B(u, a) = 0 \). It follows that for all points on the line joining \([u]\) and \([a]\) that

\[
B(u, \lambda u + \mu a) = \lambda B(u, u) + \mu B(u, a) = 0
\]
so the line joining \([u]\) and \([a]\) is the polar of \([u]\) hence also the tangent to \( C \) at \([u]\). The same argument applies to \([w]\) and this establishes (iii).

\[
\square
\]

### 3.7 Linear subspaces of quadrics and ruled surfaces

Quadrics in projective spaces \( P(V) \) of dimension greater than 2 often contain linear spaces. However, such a linear space can only be so big.

**Proposition 3.12.** Suppose that \( Q \subset P(V) \) is a non-degenerate quadric. Then if some subspace \( P(U) \) is a subset of \( Q \), then

\[
\dim(P(U)) \leq (\dim(P(V)) - 1)/2
\]
Proof. Suppose that \( P(U) \subseteq Q \). Then for all \( u \in U \), \( B(u, u) = 0 \). Furthermore, for any pair \( u, u' \in U \) we have by Proposition 3.4 that
\[
B(u, u') = \frac{1}{2}(B(u + u', u + u') - B(u, u) - B(u', u')) = 0
\]
from which it follows that \( U \subseteq U^\perp \). According to Proposition 2.13 this means that
\[
2 \dim(U) \leq \dim(U^\perp) = \dim(V)
\]
from which the conclusion follows.

Indeed, over the complex numbers the maximum value of (7) is always realized. If \( n \) is odd then we can factor
\[
x_0^2 + \ldots + x_n^2 = (x_0 + ix_1)(x_0 - ix_1) + \ldots + (x_{n-1} + ix_n)(x_{n-1} - ix_n)
\]
so the non-degenerate quadric \( Z(x_0^2 + \ldots + x_n^2) \) in \( \mathbb{C}P^n \) contains the linear space
\[
Z((x_0 - ix_1), \ldots, (x_{n-1} - ix_n))
\]
of dimension \( (n - 1)/2 \).

In the real case, the maximum value of (7) is achieved when \( n \) is odd and \( n + 1 = 2p = 2q \) (we say \( B \) has split signature) . In this case,
\[
x_0^2 + \ldots + x_{p-1}^2 - x_p^2 - \ldots - x_{2p-1}^2 = (x_0 + x_p)(x_0 - x_p) + \ldots + (x_{p-1} + x_{2p-1})(x_{p-1} - x_{2p-1})
\]
so the quadric contains the linear subspace \( Z(x_0 - x_p, \ldots, x_{p-1} - x_n) \).

This is best pictured in the case \( n = 3 \):
\[
x^2 + y^2 - z^2 - w^2 = (x - z)(x + z) + (y - w)(y + w) = 0
\]
passing to affine coordinates by setting \( w = 1 \) gives
\[
(x - z)(x + z) + (y - 1)(y + 1) = 0
\]
which defines a surface called a hyperboloid or “cooling tower”.

Indeed, the hyperboloid can be filled or “ruled” by a family of lines satisfying the equations
\[
\lambda(x - z) = \mu(y - 1)
\]
\[
\mu(x + z) = \lambda(y + 1)
\]
for some homogeneous coordinates \([\lambda : \mu]\). It can also by ruled by the family of lines satisfying
\[
\lambda(x - z) = \mu(y + 1)
\]
\[
\mu(x + z) = \lambda(y - 1)
\]
for some homogeneous coordinates \([\lambda : \mu]\).
4 Ideas for Projects

- Projective spaces over finite fields and the Fano plane
- Cramer’s paradox
- Cross-ratio
- Cubic curves
- Hopf Fibration
- Incidence geometries and non-Desarguesian planes
- Elliptic curves and group laws
- Symmetric bilinear forms in calculus and the Hessian
- The Grassmanian