Real projective manifolds

Stephan Tillmann

(Version 12.7.2012)

Course description

These notes accompany the course “Real projective manifolds” given by the author at Tsinghua University in July 2012. The purpose of the course is to give a flavour of the study of real projective manifolds in particular, and of the interplay between algebra, geometry and topology in general.

The course begins with an introduction to real projective geometry, including the notions of projective transformation, cross ratio, duality, convexity and Hilbert metric. The special rôle of elliptic, euclidean and hyperbolic geometry will be discussed, as well as affine geometry. An emphasis will be placed on key examples of Hilbert geometries (such as the Klein model for hyperbolic space and the Hilbert geometry modelled on the simplex) and of their quotients.

Key notions for the study of convex projective manifolds will then be described in detail. This includes a look at the large and small scale properties of a Hilbert geometry as well as a study of its isometries, especially with view towards their fixed point sets and dynamics. The notes end with a proof of the Margulis Lemma for properly convex projective manifolds and a brief description of one of its applications: the thick-thin decomposition of strictly convex projective manifolds.

About these notes

In preparing the notes, much use was made of the main references and survey articles listed in the bibliography. In particular, about a quarter of the contents is taken straight from [2]. The notes in their current form lack many of the proofs and pictures, which were given in lectures. An appendix gives a self-contained account of the notions from algebraic topology that are used or alluded to.

Contents

§1 Elements of projective geometry 2
§2 Hilbert geometries 10
§3 Projective manifolds 16
§4 Hilbert geometries on large and small scales 20
§5 Isometries of the Hilbert metric 24
§6 The Margulis lemma 27
A Manifolds, fundamental group and covering spaces 29
B Bibliography 35
1 Elements of projective geometry

This first section introduces aspects of projective geometry that are fundamental to the study of projective manifolds. Examples given include euclidean, affine and elliptic geometry. The last section discusses a direction not taken in this course: the axiomatic study of projective geometry.

References: Busemann and Kelly [1], Goldman [4], Cooper, Long and Tillmann [2], de la Harpe [47]

1.1 What is geometry?

According to Felix Klein’s influential Erlanger program of 1872, geometry is the study of the properties of a space, which are invariant under a group of transformations. In Klein’s framework, the familiar euclidean geometry consists of \( \mathbb{R}^n \) and its group of isometries. In general, a geometry is a pair \((G,X)\), where \( X \) is a (sufficiently nice) space and \( G \) is a (sufficiently nice) group acting on the space. Geometric properties are precisely those that are preserved by the group. A geometry in Klein’s sense may not allow the concepts of distance or angle that are the familiar starting points of euclidean geometry. Examples of such geometries include affine geometry and projective geometry.

If \( H \leq G \) is a subgroup, then \( H \) corresponds to the group of transformations of a more rigid structure on \( X \). This is called a stiffening of \((G,X)\), since an \((H,X)\) structure inherits the geometric properties of \((G,X)\), but may have more. Examples of stiffenings are:

- euclidean geometry is a stiffening of affine geometry
- affine geometry is a stiffening of projective geometry.

The contents of projective geometry arises from the structure of the underlying vector space; affine geometry allows the additional notions of parallel lines and ratios of distances along parallel lines; and euclidean geometry allows the additional notions of distance and angle.

A subtle point is that the euclidean and affine transformations leave a subspace of projective space invariant, and this invariant subspace plays a rôle in the definition of the additional geometric properties. In a similar way, hyperbolic geometry and more general Hilbert geometries are encountered as stiffenings of projective geometry at the end of this lecture.

1.2 Real projective space

I’ll first define the basic objects without coordinates. Let \( V \) be a real vector space of dimension \( n + 1 \). Then the set of all 1–dimensional vector subspaces of \( V \) is denoted \( \mathbb{P}(V) \) and called a projective space of dimension \( n \).

If \( \dim V = 0 \), then \( \mathbb{P}(V) = \emptyset \). A useful convention is to set \( \dim \emptyset = -1 \).

For \( 0 \neq v \in V \), denote \( [v] \in \mathbb{P}(V) \) the 1–dimensional subspace spanned by \( v \).

It will be convenient to have names for the projective spaces of small dimensions. A 2–dimensional projective space is called a projective plane, a 1–dimensional projective space is a projective line, and a 0–dimensional projective space is simply a point.

1.3 Projective linear subspaces

If \( U \subseteq V \) is a vector subspace, then \( \mathbb{P}(U) \) is a projective space of dimension \( \dim(U) - 1 \), and we have a natural inclusion \( \mathbb{P}(U) \subseteq \mathbb{P}(V) \). We call \( \mathbb{P}(U) \) a projective linear subspace of \( \mathbb{P}(V) \). The intersection of projective linear subspaces is again a projective linear subspace; it is the projectivisation of the intersection of the corresponding subspaces of \( V \).

If \( \dim(U) = n \), \( \mathbb{P}(U) \) is called a hyperplane in \( \mathbb{P}(V) \).

For a set \( S \subseteq \mathbb{P}(V) \) denote \( \langle S \rangle \subseteq \mathbb{P}(V) \) the smallest subspace containing \( S \). This is easily seen to be the projectivisation of the linear span of \( S \). We have the following useful fact:
Lemma 1.1 Let \( A, B \subseteq \mathbb{P}(V) \) be projective linear subspaces. Then
\[
\dim(A \cap B) = \dim A + \dim B - \dim(A \cup B).
\]

1.4 The dual space

The dual vector space of \( V \) is \( V^* = \text{Hom}(V, \mathbb{R}) \), the vector space of all linear transformations \( V \to \mathbb{R} \).

A codimension-1 vector subspace \( U \leq V \) determines a 1-dimensional subspace of \( V^* \), namely the set of all \( \varphi \in V^* \) with \( U \subset \ker \varphi \). This is generalised as follows.

Given any subspace \( U \leq V \), one has the annihilator
\[
\text{Ann}_V(U) = \{ \varphi \in V^* \mid \varphi(U) = 0 \}.
\]

One also has the notation \( \text{Ann}_V(U) = \text{Ann}(U) = U_\perp \). I’ll settle for the latter. Then \( V_\perp = \{0\} \), and we have the following relationship between dimensions:
\[
\dim U + \dim U_\perp = \dim V.
\]

Moreover, if \( U \leq W \leq V \), then \( W_\perp \leq U_\perp \), since any linear functional that vanishes on \( W \) must also vanish on \( U \). You should also verify that
\[
\langle U \cap W \rangle_\perp = \langle U_\perp, W_\perp \rangle
\]
and
\[
\langle U, W \rangle_\perp = U_\perp \cap W_\perp.
\]

So the annihilator gives us a bijection between \( k + 1 \)-dimensional subspaces of \( V \) and \( (n-k) \)-dimensional subspaces of \( V^* \), taking subspaces to subspaces, intersections to spans and spans to intersections.

Stated in term of the projective spaces, there is a natural bijection called duality between \( k \)-dimensional subspaces of \( \mathbb{P}(V) \) and \( (n-k-1) \)-dimensional subspaces of \( \mathbb{P}(V^*) \), which has the same properties. Denote \( W_\perp \leq \mathbb{P}(V^*) \) the dual subspace of the projective subspace \( W \leq \mathbb{P}(V) \).

One can go one step further and consider the dual of the dual: \( (V^*)^* = V^{**} \).

Exercise 1.2 There is a natural isomorphism \( V \to V^{**} \).

1.5 Projective maps

Suppose \( T : V \to W \) is an injective linear transformation, then it takes the 1-dimensional vector subspaces of \( V \) to 1-dimensional vector subspaces of \( W \). Hence it gives a well-defined projective map \([T] : \mathbb{P}(V) \to \mathbb{P}(W)\). An important class of such examples arises when \( V = W \). In this case, \( T \) is injective if and only if \( T \in \text{GL}(V) \), and \([T]\) acts non-trivially on \( \mathbb{P}(V) \) if and only if \( T \notin \{ \lambda E \mid \lambda \in \mathbb{R} \setminus \{0\} \} \). Hence the group \( \text{PGL}(V) = \text{GL}(V)/\{\lambda E\} \) acts effectively on \( \mathbb{P}(V) \); this means that the only element that acts trivially is the identity. If \( V = \mathbb{R}^{n+1} \), we also write \( \text{GL}(V) = \text{GL}(n+1) \).

Going back to Klein’s definition of geometry:

projective geometry is the study of the properties of \( \mathbb{P}(V) \), which are invariant under \( \text{PGL}(V) \).

The contents of projective geometry thus arises (at first) from the linear algebra of \( V \).

Given the subsets \( S_1, \ldots, S_m \subseteq \mathbb{P}(V) \), the following defines a subgroup of \( \text{PGL}(V) \):
\[
\text{PGL}(S_1, \ldots, S_m) = \{ T \in \text{PGL}(V) \mid T(S_k) \subseteq S_k \text{ for all } k \}.
\]

We will sometimes consider the following more general projective maps:
If $T : V \to W$ is an arbitrary linear transformation, then there is an induced projective map
\[ [T] : \mathbb{P}(V) \setminus \mathbb{P}(\ker T) \to \mathbb{P}(W). \]
An example of such a map is the following *radial projection*. Let $0 \neq v \in V$, and denote $P : V \to V / \langle v \rangle$ the natural quotient map. Then
\[ D_v = [P] : \mathbb{P}(V) \setminus \{[v]\} \to \mathbb{P}(V / \langle v \rangle). \]
Another construction of projective maps arises from embeddings in higher dimensions:

Let $H_0$ and $H_1$ be hyperplanes in $\mathbb{P}^n$ and let $o$ be a point not in their union. The *perspectivity* $T : H_0 \to H_1$ from $o$ is obtained by mapping $p \in H_0$ to the point of intersection $Tp$ of $H_1$ with the projective line through $p$ and $o$. This is well defined by the intersection formula for subspaces.

**Exercise 1.3** Show that the perspectivity is a projective map between the $(n - 1)$-dimensional projective spaces $H_0$ and $H_1$.

### 1.6 Dual projective maps

If $\mathcal{B} = \{b_1, \ldots, b_{n+1}\}$ is a basis of $V$, then there is a dual basis $\mathcal{B}^* = \{f_1, \ldots, f_{n+1}\}$ of $V^*$ characterised by $f_j(b_k) = \delta_{jk}$ for all $1 \leq j, k \leq n + 1$.

An element $T \in \text{GL}(V)$ acts on $V^*$ by:
\[ T^*(f) = f \circ T^{-1}. \]
In particular, the matrices with respect to the bases are related by:
\[ [T^*]_{\mathcal{B}^*} = \text{transpose}([T^{-1}])_{\mathcal{B}}. \]

**Exercise 1.4** Why is the action defined by $T^*(f) = f \circ T^{-1}$ and not (as found in most linear algebra texts) by $T^*(f) = f \circ T$?

### 1.7 Projective coordinates

Choosing a basis $\{b_1, \ldots, b_{n+1}\}$ for $V$ gives projective coordinates
\[ [t_1 b_1 + \cdots + t_{n+1} b_{n+1}] = [t_1 : \cdots : t_{n+1}] = [\lambda t_1 : \cdots : \lambda t_{n+1}], \]
where $\lambda \neq 0$. Given a hyperplane $H \subset \mathbb{P}(V)$, one can choose projective coordinates so that
\[ H = \{[x_1 : \cdots : x_n : 0] \mid x_k \in \mathbb{R}, \text{ not all } x_k \text{ equal to 0}\}. \]

An *affine patch* is the complement of a hyperplane; the complement of the above $H$ is given by:
\[ \mathbb{P}(V) \setminus H = \{[x_1 : \cdots : x_n : 1] \mid x_k \in \mathbb{R}\}. \]

Dehomogenising identifies the affine patch $\mathbb{P}(V) \setminus H$ with $\mathbb{R}^n$:
\[ [x_1 : \cdots : x_n : 1] \leftrightarrow (x_1, \ldots, x_n). \]

This also gives an identification
\[ \text{PGL}(\mathbb{P}(V) \setminus H) \leftrightarrow \text{Aff}(n) = \{x \mapsto Ax + b \mid A \in \text{GL}(n), b \in \mathbb{R}^n\} = \text{GL}(n) \ltimes \mathbb{R}^n, \]
where the embedding of $\text{Aff}(n)$ into $\text{PGL}(n + 1)$ is given via the following map to $\text{GL}(n + 1)$:
\[ (x \mapsto Ax + b) \mapsto \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}. \]
It follows that $(\mathbb{P}(V) \setminus H, \text{PGL}(\mathbb{P}(V) \setminus H))$ is identified with $n$-dimensional affine space $(\mathbb{R}^n, \text{Aff}(n))$, and hence affine geometry is a stiffening of projective geometry.

A further stiffening is *euclidean geometry*, which again has underlying space $\mathbb{R}^n$, but with the smaller group
\[ \text{Euc}(n) = \{x \mapsto Ax + b \mid A \in \text{O}(n), b \in \mathbb{R}^n\} = \text{O}(n) \ltimes \mathbb{R}^n, \]
where $\text{O}(n) = \{A \in \text{GL}(n) \mid A^T A = E\}$ is the orthogonal group. If $\mathbb{R}^n$ is imbued with the geometry of the standard inner product, turning it into $\mathbb{E}^n$, we have $\text{Isom}(\mathbb{E}^n) = \text{Euc}(n)$.
1.8 More on euclidean geometry

The euclidean metric on \( \mathbb{R}^n \) is given via

\[ d_{\text{euc}}(x,y) = ||x - y|| = \sqrt{(x - y) \cdot (x - y)} = \sqrt{\sum (x_k - y_k)^2}, \]

and for each \( T \in \text{Euc}(n) \) and all \( x,y \in \mathbb{R}^n \), one has \( d_{\text{euc}}(Tx,Ty) = d_{\text{euc}}(x,y) \).

Denote \( n \)-dimensional Euclidean space \( E^n = (\mathbb{R}^n,d_{\text{euc}}) \). You may have seen the proof that \( \text{Isom}E^n = \text{Euc}(n) \) and that the latter has the claimed structure; here is a sketch. To begin with, one shows that every isometry \( \alpha \in \text{Isom}E^n \) can be written in the form \( \alpha(x) = Ax + b \), where \( A \in O(n) \) and \( b \in \mathbb{R}^n \). To do this, one uses linear algebra to first show that an isometry fixing the origin must be of the form \( \alpha(x) = Ax \), where \( A \in O(n) \). A general isometry will take the origin to some \( b \in \mathbb{R}^n \), so composing this with the translation by \(-b\) gives an isometry fixing the origin.

This decomposition of \( \alpha \) into a “rotational” part and a translational part now leads to the group structure. Denote the translation \( x \mapsto x + b \) by \( T_b \). The subgroup of translations is normal in \( \text{Isom}E^n \) and isomorphic to \( \mathbb{R}^n \). The isometries fixing the origin also form a subgroup, and this is isomorphic with \( O(n) \). Now every isometry in the above form is written uniquely as an isometry fixing the origin followed by a translation (it can also be written uniquely as a translation followed by an isometry fixing the origin, but this is irrelevant). So with the representation \( \rho : O(n) \to \text{Aut}(\mathbb{R}^n) \) taking \( A \in O(n) \) to the automorphism \( T_b \to AT_bA^{-1} = T_{Ab} \), one has

\[ \text{Isom}E^n \cong \mathbb{R}^n \rtimes \rho O(n). \]

1.9 More on affine geometry

In affine geometry, there is no interesting function on pairs of points that is invariant under the action of \( \text{Aff}(n) \), but there is an interesting such function on triples of points. This will be shown in the following exercise.

**Exercise 1.5** Let \( L_0 \) and \( L_1 \) be (affine) lines in \( \mathbb{R}^n \) and \( x,y \in L_0 \) and \( a,b \in L_1 \) be pairs of distinct points. Show that there is \( T \in \text{Aff}(n) \) such that \( Tx = a \) and \( Ty = b \). Moreover, show that any two such affine maps agree on \( L_0 \).

Hence deduce that any function \( f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \) which is invariant under \( \text{Aff}(n) \) is constant.

Next, identify \( \mathbb{R} \) with the affine line \( L_1 = \{(x,0,\ldots,0) \mid x \in \mathbb{R} \} \), let \( L_0 \) be any line and \( x,y,z \in L_0 \) be pairwise distinct points. Define \([x,y,z] = Tz\), where \( T \) is an affine transformation mapping \( Tx = 0 \) and \( Ty = 1 \). Show that if \( L_0 = \mathbb{R} \), then

\[ [x,y,z] = \frac{z - x}{y - x}. \]

This last line can be paraphrased as: affine geometry allows the concept of “ratios of distances along parallel lines.” Also note that the ratio can be extended to triples of points of which at least two are distinct.

1.10 Elliptic space

Spherical geometry is usually described on the round \( n \)-sphere, but the treatment becomes more succinct when one identifies antipodal points; for instance two points on the sphere lie on a unique (spherical) line if and only if they are not antipodal.

The \( n \)-sphere \( S^n \) has a natural Riemannian geometry of constant curvature +1, which co-incides with the path metric that it inherits as the level set \( x_1^2 + \ldots + x_{n+1}^2 = 1 \) in \( \mathbb{R}^{n+1} \). The antipodal map is an isometry of this metric, and hence the quotient metric space, called *elliptic space* has a Riemannian geometry of constant curvature +1. In the quotient, any two points determine a unique (elliptic) line and are connected by a unique shortest
geodesic. The elliptic metric can be given explicitly as follows. View $\mathbb{P}^n$ as the quotient of the closed unit ball in $\mathbb{R}^n$ by the antipodal map. Then the line element for the metric is given by:

$$ds^2 = \frac{4}{1 + ||x||^2}d^2\sqrt{1 + ||x||^2}.$$ 

It follows from its description that as a set, elliptic space can be identified with $\mathbb{P}^n = \mathbb{P}(\mathbb{R}^{n+1})$. The group of isometries of elliptic space is naturally identified with PO$(n + 1)$, and it follows that

elliptic geometry $(\mathbb{P}^n, \text{PO}(n + 1))$ is a stiffening of projective geometry $(\mathbb{P}^n, \text{PGL}(n + 1))$.

### 1.11 Desargues and Pappus

The following two results are classics in projective geometry, and highlight how results that were first proved on the euclidean plane using geometric techniques generalise to projective geometry, often with more elegant statements and simpler proofs.

**Theorem 1.6** (Desargues) If two triangles $\Delta(x, y, z)$ and $\Delta(a, b, c)$ are in perspective from $o \in \mathbb{P}^n$, then the corresponding sides meet in 3 collinear points; i.e. there is a projective line containing the points $\langle x, y \rangle \cap \langle a, b \rangle$, $\langle y, z \rangle \cap \langle b, c \rangle$ and $\langle z, x \rangle \cap \langle c, a \rangle$.

**Theorem 1.7** (Pappus) Let $L_0$ and $L_1$ be lines in $\mathbb{P}^2$, and let $x, y, z \in L_0$ and $x, y, z \in L_1$ be triples of pairwise distinct points and not equal to $L_0 \cap L_1$. Then the three points $\langle x, b \rangle \cap \langle y, a \rangle$, $\langle y, c \rangle \cap \langle z, b \rangle$ and $\langle z, a \rangle \cap \langle x, c \rangle$ are collinear.

### 1.12 Topologies on $\mathbb{P}(V)$ and $\text{PGL}(V)$

Using coordinates, $\mathbb{P}(V)$ is identified with $\mathbb{P}(\mathbb{R}^{n+1}) = \mathbb{P}^n$. One can view $\mathbb{P}^n$ as the quotient topological space of $\mathbb{R}^{n+1} \setminus \{0\}$ (with the usual euclidean topology) by the equivalence relation $x \sim y$ if and only if $x = \lambda y$ for some $\lambda \in \mathbb{R} \setminus \{0\}$. The quotient topology from this description coincides with the metric topology that arises from viewing $\mathbb{P}^n$ as elliptic space since the metric on $\mathbb{S}^n$ is the path metric from the Euclidean metric.

The elements of $\text{GL}(n + 1) = \text{GL}(\mathbb{R}^{n+1})$ are continuous maps $\mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$, and descend to continuous maps $\mathbb{P}^n \to \mathbb{P}^n$. Last, $\text{GL}(n + 1)$ has a natural subspace topology from $\mathbb{R}^{(n+1)^2}$, which makes it into a Lie group, and $\text{PGL}(n + 1)$ is the quotient topology.

Using the identification of $V$ with $\mathbb{R}^{n+1}$ and $\mathbb{P}(V)$ with $\mathbb{P}^n$, and hence of $\text{GL}(V)$ with $\text{GL}(n + 1)$ and of $\text{PGL}(V)$ with $\text{PGL}(n + 1)$, we have natural topologies on these spaces, and they are independent of the choice of coordinates.

We often identify a projective line with $\mathbb{P}^1$, which (via its projective coordinates) is identified with $\mathbb{R} \cup \{\infty\}$ using the map $[x : y] \to \frac{x}{y}$.

### 1.13 Projective basis and fundamental theorems

A set of $n + 2$ points in $\mathbb{P}(V)$ is in general position if each subset of $n + 1$ points has representative vectors in $V$ which are linearly independent. Such a set of $n + 2$ points is called a projective basis or projective frame of reference or simplex of reference of $\mathbb{P}(V)$.

**Lemma 1.8** Given two sets of points that are in general position in $\mathbb{P}(V)$, $\{p_1, \ldots, p_{n+2}\}$ and $\{q_1, \ldots, q_{n+2}\}$, there is a unique projective map $\varphi \in \text{PGL}(V)$ such that $\varphi(p_k) = q_k$ for each $k \in \{1, \ldots, n+2\}$. 
Here are three useful facts that make use of the topologies described in the previous section; they can be viewed as different guises of a fundamental theorem of projective geometry.

**First Fundamental Theorem 1.9** Suppose $n \geq 2$. Let $A$ and $B$ be open, connected subsets of $\mathbb{P}(V)$ and suppose $f: A \to B$ is a continuous map that sends the intersection of any projective line with $A$ to the intersection with $B$ and a line, i.e. for each line $l$ there exists a line $l'$ such that $f(A \cap l) = B \cap l'$. Then $f$ is the restriction to $A$ of an element in $\text{PGL}(V)$.

**Second Fundamental Theorem 1.10** Fix a projective basis $B_0$ of $\mathbb{P}(V)$, and define $h: \text{PGL}(V) \to \{ B \mid B \text{ is a projective basis of } \mathbb{P}(V) \} \subset \mathbb{P}(V)^{n+2}$ by $h(\varphi) = \varphi(B_0)$. Then $h$ is a homeomorphism (where $\mathbb{P}(V)^{n+2}$ is given the product topology).

**Corollary 1.11** (Third Fundamental Theorem) There exists a unique projective map of a projective line taking any three pairwise distinct points on the line to any other three pairwise distinct points.

### 1.14 Cross ratio

As a result of the above corollary, there can be no non-trivial function on pairs or triples of points on $\mathbb{P}^1$ that is invariant under all projective transformations of $\mathbb{P}^1$. However, in analogy with the ratio in affine geometry, one can define a cross ratio as follows.

Let $L \subset \mathbb{P}(V)$ be a line. Given $w,x,y,z \in L$, with $x,y,z$ pairwise distinct, denote $T: L \to \mathbb{P}^1$ the unique projective map such that $Tx = 0$, $Ty = 1$ and $Tz = \infty$. Then define

$$[x,y,w,z] = Tw.$$

**Exercise 1.12** If $L = \mathbb{P}^1$, this is given by the formula:

$$[x,y,w,z] = \frac{(x-w)(z-y)}{(x-y)(z-w)}.$$

The cross ratio can be extended to quadruples of collinear points in $\mathbb{P}(V)$ of which at least three are pairwise distinct.

A pair $(y,w)$ is harmonic with respect to $(x,z)$, in which case $(x,y,w,z)$ is a harmonic quadruple, if and only if

$$[x,y,w,z] = -1.$$

**Exercise 1.13** Let $\sigma$ be a permutation on four symbols. Show that there exists a linear fractional transformation $\Phi_\sigma$ such that

$$[x_{\sigma(1)}, x_{\sigma(3)}, x_{\sigma(3)}, x_{\sigma(4)}] = \Phi_\sigma([x_1, x_2, x_3, x_4]).$$

In particular, determine which permutations leave cross ratio invariant.

**Exercise 1.14** Let $f: \mathbb{P}^1 \to \mathbb{P}^1$ be a map. Show that the following are equivalent:

1. $f$ is a projective map;
2. $f$ is a bijection and takes harmonic quadruples to harmonic quadruples;
3. $f$ preserves cross ratio; i.e. for all quadruples $(w,x,y,z)$ the cross ratios satisfy:

$$[f(x), f(y), f(w), f(z)] = [x,y,w,z].$$
Exercise 1.15  Let \( \mathbb{R}^n \) be an affine patch, imbued with the Euclidean metric. Suppose \( w,x,y,z \) lie on an affine line such that \( x \) and \( z \) are the endpoints of a segment on this line containing \( y \) and \( w \), and such that \( y \) lies between \( x \) and \( w \) on the line segment. Show that

\[
[x,y,w,z] = [x-w] \cdot [z-y] /[x-y] \cdot [z-w].
\]

Notice that the cross ratio in this set-up satisfies \([x,y,w,z] \geq 1\).

Exercise 1.16  Given \([x],[z] \in \mathbb{P}^n\), let \([w] = [\lambda x + \mu z] \) and \([v] = [\lambda' x + \mu' z] \). Show that

\[
[x,y,w,z] = \frac{\lambda \mu'}{\lambda' \mu}
\]

Exercise 1.17 (Duality)  Show that there is a well-defined cross ratio of four lines in \( \mathbb{P}^2 \) meeting in a single point. How does this statement (and its proof) generalise to \( \mathbb{P}^n \)?

1.15 Convexity

The set \( C \subseteq \mathbb{P}(V) \) is convex if the intersection of every line with \( C \) is connected.

A convex subset \( C \subseteq \mathbb{P}(V) \) is properly convex if its closure is contained in an affine patch.

We will distinguish three classes of boundary points. Let \( p \in \partial \overline{C} \), then \( p \) is a

1. strictly convex point if it is not contained in a line segment of positive length in \( \partial \overline{C} \).
2. \( C^1 \) point if it is contained in a unique supporting hyperplane to \( \partial \overline{C} \).
3. round point if it is both a strictly convex point and a \( C^1 \) point.

For instance, a triangle has strictly convex points that are not \( C^1 \), and \( C^1 \) points that are not strictly convex.

The terminology “\( C^1 \) point” has the following justification. At a point \( p \in \partial \Omega \), the boundary \( \partial \Omega \) is locally the graph of a function defined on a neighborhood of \( p \) in a supporting hyperplane \( H \). By (2.7 of [46]) this function is \( C^1 \) at \( p \) if and only if \( H \) is the unique supporting hyperplane at \( p \).

The set \( C \) is strictly convex if it is properly convex and strictly convex at every point in \( \partial \overline{C} \).

Proposition 1.18 (convex decomposition)  If \( \Omega \) is an open convex subset of \( \mathbb{P}^n \) which contains no projective line, then it is a subset \( \mathbb{A}^k \times C \) of some affine patch \( \mathbb{A}^k \times \mathbb{A}^{n-k} \subset \mathbb{P}^n \), where \( k \geq 0 \) and \( C \subset \mathbb{A}^{n-k} \) is a properly convex set. One factor might be a single point. The set \( C \) is unique up to projective isomorphism.

1.16 Using the double cover

It will often be convenient to refer to eigenvalues rather than ratios of eigenvalues; for instance in the discussion of isometries. This is achieved by working with the double cover of projective space, the sphere

\[
S(V) = (V \setminus \{0\})/\mathbb{R}^+.
\]

with automorphism group \( \text{SL}_\pm(V) = \{ A \in \text{GL}(V) \mid \det A = \pm 1 \} \). This is a double cover of geometries since also the homomorphism \( \text{SL}_\pm(V) \to \text{PGL}(V) \) is 2–to–1. Notice that spherical geometry is a stiffening of \((S(V),\text{SL}_\pm(V))\). In analogy with \( \mathbb{P}^n = \mathbb{P}(\mathbb{R}^{n+1}) \), we write \( S^n = S(\mathbb{R}^{n+1}) \) and \( \text{SL}_\pm(n+1) = \text{SL}_\pm(\mathbb{R}^{n+1}) \).

Let \( \pi : S(V) \to \mathbb{P}(V) \) denote the double cover. The preimage in \( S(V) \) of a projective line in \( \mathbb{P}(V) \) is a great circle and the preimage of an affine patch is an open hemisphere. A subset of \( S(V) \) is convex if its intersection with each great circle is connected. A subset of the sphere \( S(V) \) is properly convex if it is convex and its closure is contained in an open hemisphere.

If \( C \) is a properly convex subset of \( \mathbb{P}(V) \), then \( \pi^{-1} C = C_0 \cup C_1 \) has two components, each with closure contained in an open hemisphere, and the restriction of \( \pi \) to each component gives a homeomorphism. We therefore choose one component as a lift (say \( C_0 \)) and define \( \text{SL}_\pm(C) = \text{SL}_\pm(C_0) \) for the subgroup of \( \text{SL}_\pm(V) \) which preserves \( C_0 \). This economical notation should not cause any confusion.
Exercise 1.19  The group $\text{SL}_\pm(C)$ is naturally isomorphic to the subgroup $\text{PGL}(C) \leq \text{PGL}(V)$. 

The conveniency of this set-up lies in the easy of switching back and forth between projective space and its double cover, and between $\text{PGL}(C)$ and $\text{SL}_\pm(C)$.

The sphere also provides a good framework to define the dual of a convex domain. A subset $C \subseteq V$ is a cone if $\lambda \cdot C = C$ for all $\lambda > 0$. The cone $C$ is convex if $a, b > 0$ and $x, y \in C$ always implies $ax + by \in C$. The convex cone $C$ is sharp if it contains no affine line.

The properly convex set $C \subset S(V)$ determines the sharp convex cone $\mathcal{C}(C) = \mathbb{R}_+ \cdot C \subset V$.

If $C$ is a sharp convex cone, then its dual cone is defined by:

$$
\mathcal{C}^* = \{ \phi \in V^* \mid \phi(x) > 0 \text{ for all } x \in \overline{C} \}.
$$

Lemma 1.20  [4] The dual cone of a sharp convex cone is a sharp convex cone.

Given a properly convex domain $\Omega \subset \mathbb{P}(V)$, we take a lift $\Omega_0 \subset S(V)$, and hence obtain a sharp convex cone $\mathcal{C}(\Omega_0)$. Intersecting the dual convex cone $(\mathcal{C}(\Omega_0))^*$ with the sphere $S(V^*)$ and projecting to $\mathbb{P}(V^*)$ gives the dual domain $\Omega^*$. The dual domain is independent of the chosen lift. It follows from the lemma, that $\Omega^*$ is also properly convex.

From now on, we will write $\Omega = \Omega_0$, and it should be clear from context whether $\Omega \subset \mathbb{P}(V)$ or $S(V)$.

Exercise 1.21  Letting $\text{SL}_\pm(\mathcal{C}(\Omega)) \leq \text{SL}_\pm(V)$ denote the subgroup that preserves $\mathcal{C}(\Omega)$, we have $\text{SL}_\pm(\mathcal{C}(\Omega)) = \text{SL}_\pm(\Omega)$.

Exercise 1.22  The canonical isomorphism $V \rightarrow V^{**}$ maps $\Omega$ onto $\Omega^{**}$.

Exercise 1.23  The following are equivalent:

1. Every point on $\partial \Omega$ is a strictly convex point.
2. Every point on $\partial \Omega^*$ is a $C^1$ point.

Exercise 1.24  Show that if $\Delta \subset \mathbb{P}^2$ is a triangle, then its dual domain $\Delta^*$ is also a triangle.

1.17 Axiomatic projective geometry

These lectures have a bias towards real projective geometry. David Hilbert formulated an axiomatic approach to projective geometry, consisting of the abstract sets $V$ (for vertices) and $L$ (for lines) as well as a subset of $V \times L$ termed incidence. The requirements on the incidence relation are:

1. Every line passes through at least 3 points.
2. Every point is on at least 3 lines.
3. Given any 2 points, there is a unique line passing through both of them.
4. Given any 2 lines, there is a unique point lying on both of them.

Hilbert showed that coordinates can be found for each projective space that satisfies his axioms. This again takes the form $\mathbb{P}(k^{n+1})$, but in this case $k$ is a division ring (not necessarily a field), so all that is required is that the equation $ax = b$ has a solution for any $a \neq 0$, but $k$ may not be commutative. Many of the above facts may still hold, taking care of the fact that $x \sim y$ if and only if $y = kx$ uses left-multiplication. For instance:

$k$ is an associative ring $\iff$ Desargues’ theorem holds in $\mathbb{P}(k^{n+1})$,

$k$ is a commutative ring $\iff$ Pappus’ theorem holds in $\mathbb{P}(k^{n+1})$.

Projective geometry in this setting has many applications, for instance in coding theory and cryptography.
2 Hilbert geometries

After the preliminary notions from the last section, we now turn to the study of the Hilbert metric on a properly convex domain in projective space. After looking at some key examples in detail, we focus on basic properties of geodesics and the distance function before turning to general properties on the large and the small scale.

References: Busemann and Kelly [1], de la Harpe [47], Socié-Méthou [3], Cooper, Long and Tillmann [2]

2.1 The Hilbert metric

Let $\Omega$ be a properly convex open set contained in an affine patch. The Hilbert metric $d_\Omega$ on $\Omega$ is

$$d_\Omega(a, b) = \log \left( \frac{\|b - x\| \cdot \|a - y\|}{\|b - y\| \cdot \|a - x\|} \right),$$

where $x, y \in \partial \Omega$ are the endpoints of a line segment in $\Omega$ containing $a$ and $b$ such that $a$ lies between $x$ and $b$ on the line segment. The definition is worded to include the case $a = b$.

By definition and Exercise 1.15, we have positivity, i.e. $d_\Omega(a, b) \geq 0$ for all $a$ and $b$ and $d_\Omega(a, b) = 0$ if and only if $a = b$. Symmetry follows from Exercise 1.13. Proofs of the triangle inequality using cross ratios can be found in [1] and [47] (I gave the latter in the lecture), and a proof using the Kobayashi metric is given in [4].

It is easy to verify that for any $c$ lying on the segment between $a$ and $b$, we have $d_\Omega(a, b) = d_\Omega(a, c) + d_\Omega(c, b)$. Whence every segment of a projective line in $\Omega$ is length minimising, and in the strictly convex case these are the only geodesics (see the discussion of geodesics in §2.6).

An interesting comparison principle was observed in the proof of the triangle inequality: If $\Omega \subseteq \Omega'$ and $a, b \in \Omega$, then

$$d_\Omega(a, b) \geq d_{\Omega'}(a, b).$$

Since projective transformations preserve cross ratio, $\text{PSL}(\Omega)$ is a group of isometries of the Hilbert metric. The inclusion $\text{PSL}(\Omega) \leq \text{Isom}(\Omega, d_\Omega)$ may be strict (as will be seen in §2.3). However, it will be shown in §2.7 that if $\Omega$ is strictly convex, then $\text{PSL}(\Omega) = \text{Isom}(\Omega, d_\Omega)$. The triple

$$\left(\Omega, d_\Omega, \text{PGL}(\Omega)\right)$$

is called a Hilbert geometry. It is both a non-trivial stiffening of $(\mathbb{P}^n, \text{PGL}(n + 1))$ and a (possibly trivial) stiffening of $(\Omega, \text{Isom}(\Omega, d_\Omega))$. It is time to look at interesting examples of Hilbert geometries.

2.2 The Klein model of hyperbolic space

The Klein model or projective model of hyperbolic space $\mathbb{H}^n$ is the Hilbert geometry with domain the standard open unit ball $D^n \subseteq \mathbb{P}^n$. We have:

$$\text{PSL}(D^n) \cong \text{PO}(1, n) = \{A \in \text{GL}(n + 1) \mid A'JA = J\}/\{\pm E\},$$

where $J = \text{diago}(-1, 1, \ldots, 1)$. The Hilbert metric gives twice the usual hyperbolic metric (information on adjusting the curvature is given in §2.10). More generally, when $\Omega$ is the interior of an ellipsoid, one obtains a Hilbert geometry isometric with this, since one can take any ellipsoid to a ball by an affine map.
2.3 The Hilbert metric on a simplex

The Hex plane is obtained by taking $\Omega = \Delta$, the interior of an open 2-simplex. Then $\text{PSL}(\Delta)$ is the semi-direct product of positive diagonal matrices of determinant 1 and permutations of the three vertices:

$$\text{SL}_+ (\Delta) = \{ \text{diag}(a,b,c) \mid a,b,c > 0, abc = 1 \} \times \text{Sym}(3).$$

An excellent reference for this geometry is de la Harpe [47]. This Hilbert geometry is isometric to a normed vector space, where the unit ball is a regular hexagon. To make this explicit, let $(\mathbb{R}^2, || \cdot ||_{\text{hex}})$ be the vector space with norm:

$$||(x,y)||_{\text{hex}} = \frac{1}{2\sqrt{3}}(|2y| + |y + \sqrt{3}x| + |y - \sqrt{3}x|).$$

Choosing the vertices $a_1 = (1,0), a_2 = (-1/2, \sqrt{3}/2)$ and $a_3 = (-1/2, -\sqrt{3}/2)$ of the hexagon, an isometry between $(\Delta, d_\Delta)$ and $(\mathbb{R}^2, || \cdot ||_{\text{hex}})$ is given by

$$[e^0] e_1 + [e^2] e_2 + [e^3] e_3 = [e^0] : e^{0'} \mapsto t_1 a_1 + t_2 a_2 + t_3 a_3.$$

Since the unit ball is not strictly convex geodesics are not even locally unique. The isometry $x \mapsto -x$ of $(\mathbb{R}^2, || \cdot ||_{\text{hex}})$ does not correspond to a projective transformation of $(\Delta, d_\Delta)$ since it is given by

$$[x : y : z] \mapsto \left[ \frac{1}{x} : \frac{1}{y} : \frac{1}{z} \right].$$

It turns out that

$$\text{Isom}(\Delta, d_\Delta) = \{ \text{diag}(a,b,c) \mid a,b,c > 0, abc = 1 \} \rtimes D_6,$$

where $D_6$ is the symmetry group of a regular hexagon (this dihedral group is also often denoted $D_{12}$ to indicate its order rather than the object it acts on). Hence $\text{PSL}(\Delta)$ has index 2 in $\text{Isom}(\Delta, d_\Delta)$.

More generally, the Hilbert geometry on the $n$–simplex has with similar features, as it includes many hex planes. In dimension 3, the unit ball of this geometry is a rhombic dodecahedron, which may be better known from beehives or crystals. The Hilbert metric on a simplex plays the following special rôle:

**Theorem 2.1** (Förtsch–Karlsson [43]) The Hilbert geometry $(\Omega, d_\Omega)$ is isometric to a normed vector space if and only if $\Omega$ is the interior of an $n$–simplex.

Slicing a tetrahedron by a plane produces either a triangle or a quadrilateral. The following looks at the latter:

**Exercise 2.2** If $\square$ is a quadrilateral in $\mathbb{P}^2$, then $\text{PGL}(\square) = \text{Isom}(\square, d_{\square}) \cong D_4$, where $D_4$ is the symmetry group of the square.

More generally, de la Harpe [47] shows that for any convex $n$–gon $P$ in $\mathbb{P}^2$ with $n \geq 4$, both groups $\text{PGL}(P)$ and $\text{Isom}(P, d_P)$ are finite.

2.4 An interesting cone

Let $\Omega = D^2 \ast \{ p \} \subset \mathbb{R}P^1$ be the open cone on a round disc $D^2$. The restriction of the Hilbert metric to $D^2 \times \{ x \} \subset \Omega$ is hyperbolic geometry. Restricted to the cone on a line in $D^2$ gives a hex plane. We have

$$\text{SL}_\pm (D^2 \ast \{ p \}) = \left\{ \begin{pmatrix} tA & 0 \\ 0 & t^{-1} \end{pmatrix} \mid A \in O(2,1), t > 0 \right\}.$$
2.5 Transitive actions

In addition to marrying hyperbolic and hex geometry, the cone over the disc has the following remarkable property. Rogers announced in 1975 [54] (page 20) and in 1981 [69] a result with Larman and Mani that completely claims to classify all properly convex open sets \( \Omega \) with the property that \( \text{PGL}(\Omega) \) acts transitively on \( \Omega \). In dimension two, these are just the disc and the 2–simplex, in dimension 3 the ellipsoid, the 3–simplex and the above cone, and in higher dimensions the list of possibilities grows. An explicit description for \( n \leq 4 \) can be found in [47] and for \( n \leq 6 \) in [69].

2.6 Geodesics

A path \( \gamma \) connecting two points \( p \) and \( q \) of \( \Omega \) is a geodesic if for each point \( r \) in the image of \( \gamma \) we have \( d(p,q) = d(p,r) + d(r,q) \). Straight line segments are geodesics by definition of the Hilbert metric, but there may be more as was seen in the Hex plane. A proof of below fact can be found in [47].

**Lemma 2.3** Suppose \( r \notin [p,q] \), but \( d(p,q) = d(p,r) + d(r,q) \). Then the projective plane containing \( p, q \) and \( r \) meets \( \Omega \) in a disc \( D \), which contains two distinct maximal line segments in its boundary.

**Corollary 2.4** Any two points in \( (\Omega, d_\Omega) \) are connected by a unique geodesic if and only if the intersection of \( \Omega \) with any projective plane contains at most one maximal line segment in its boundary.

The hypothesis of the corollary in particular holds if \( \Omega \) is strictly convex.

2.7 The isometry group

We now make a first examination of isometries following [47]. A classification of the isometries in \( \text{PGL}(\Omega) \) will be undertaken later.

**Lemma 2.5** Let \( \Omega \) be a strictly convex domain and \( \gamma \in \text{Isom}(\Omega, d_\Omega) \). Then:

1. The image under \( \gamma \) of the intersection of a line with \( \Omega \) is again the intersection of a line with \( \Omega \).
2. The map \( \gamma \) extends to a homeomorphism \( \overline{\Omega} \to \overline{\Omega} \).
3. Cross-ratios of collinear points in \( \Omega \) are preserved by \( \gamma \).

Theorem 1.9 now implies that:

**Corollary 2.6** Let \( \Omega \) be a strictly convex domain. Then \( \text{PGL}(\Omega) = \text{Isom}(\Omega, d_\Omega) \).

2.8 Rigidity of domains

Since projective transformations preserve cross-ratios, convexity and take affine patches to affine patches, any \( T \in \text{PGL}(V) \) restricts to an isometry between \( (\Omega, d_\Omega) \) and \( (T(\Omega), d_{T(\Omega)}) \). Socié-Méthou has observed that this statement has a very strong converse:

**Lemma 2.7** (Local rigidity [3]) Suppose \( \Omega \) and \( \Omega' \) are contained in an affine patch of \( \mathbb{P}(V) \). If \( d_\Omega \) and \( d_{\Omega'} \) agree on an open, non-empty subset of \( \Omega \cap \Omega' \), then \( \Omega = \Omega' \).
2.9 The geometry of the distance function

The following material is taken from [2], §1.

**Lemma 2.8** If \( \Omega \) is properly (resp. strictly) convex, then metric balls of the Hilbert metric are convex (resp. strictly convex).

**Proof** Refer to Figure 1(a). Suppose \( R = d(x,y) = d(x,z) \). We need to show that for every \( p \in [y,z] \), we have \( d(x,p) \leq R \). The extreme case is obtained by taking the quadrilateral \( Q \subset \Omega \) which is the convex hull of the four points on \( \partial \Omega \), where the extensions of the segments \([x,z]\) and \([x,y]\) meet \( \partial \Omega \). Then \( d_\Omega \leq d_\Omega \) and the ball of radius \( R \) in \( Q \) centered \( x \) is a convex quadrilateral.

A function defined on a convex set is **convex** if the restriction to every line segment is convex. The statement that metric balls centered at the point \( p \) are convex is equivalent to the statement that the function on \( \Omega \) defined by \( f(x) = d_\Omega(p,x) \) is convex. Socié-Méthou [67] showed that \( d_\Omega(x,y) \) is not a geodesically convex function, in contrast to the situation in hyperbolic and Euclidean space. However, the following lemma leads to a maximum principle for the distance function.

**Lemma 2.9** (4 points) Suppose \( a,b,c,d \) are points in a properly convex set \( \Omega \) and that \( R = d_\Omega(a,b) = d_\Omega(c,d) \). Then every point on \([a,c]\) is within distance \( R \) of \([b,d]\).

**Proof** Refer to Figure 1(b). Let \( A,B \) be the points in \( \partial \Omega \) such that the line \([A,B]\) contains \([a,b]\). Define \([C,D]\) similarly. Let \( \sigma \) be the interior of the convex hull of \( A,B,C,D \). Then \( \sigma \subset \Omega \), so \( d_\sigma \geq d_\Omega \). The formula for the Hilbert metric on \( \sigma \) makes sense for pairs of points on the same edge in the 1–skeleton of \( \sigma \). Then, by construction \( d_\sigma(a,b) = d_\Omega(a,b) \) and \( d_\sigma(c,d) = d_\Omega(c,d) \). Thus it suffices to prove the result when \( \Omega = \sigma \).

We may therefore assume that \( \Omega = \sigma \) is a possibly degenerate 3-simplex. The degenerate case follows from the non-degenerate case by a continuity argument.

The identity component \( H \) of \( \text{SL}_+(\sigma) \) fixes the vertices of \( \sigma \) and acts simply transitively on \( \sigma \). If we choose coordinates so that the vertices of \( \sigma \) are represented by basis vectors, then \( H \) is the group of positive diagonal matrices with determinant 1. A point \( x \) in the interior of \( \sigma \) lies on a unique line segment, \( \ell = [a,c] \), in \( \sigma \) with one endpoint \( a \in (A,B) \) and the other \( c \in (C,D) \). It follows that the subgroup of \( H \) that preserves \( \ell \) is a one-parameter group which acts simply-transitively on \( \ell \).

The point \( x \) also lies on a unique segment \([X,Y]\) with \( X \in (A,C) \) and \( Y \in (B,D) \). Let \( G = G_1 \cdot G_2 \) be the two parameter subgroup of \( H \) that is the product of the stabilizers, \( G_1 \) of \([a,c]\) and \( G_2 \) of \([X,Y]\). The \( G_2 \)-orbit of \( x \) is a doubly ruled surface: a hyperbolic paraboloid. The \( G_1 \)-orbit of the line \( G_2 \cdot x = (X,Y) \) gives one ruling. The \( G_2 \)-orbit of the line \( G_1 \cdot x = (a,c) \) gives the other ruling. This surface is the interior of a twisted square with corners \( A,B,C,D \). Since \( G \) acts by isometries and \( d_\sigma(a,b) = d_\sigma(c,d) \), it follows that \([a,c]\) is sent to \([b,d]\) by an element of \( G \). Thus \([b,d]\) intersects \([X,Y]\) at a point \( y \). The segment \([x,y]\) can be moved by elements of \( G \) arbitrarily close to both \([a,b]\) and to \([c,d]\). Furthermore, \( d_\sigma(g \cdot x, g \cdot y) \) is independent of \( G \). It follows by continuity of cross-ratio that this constant is \( d_\sigma(a,b) \).

![Diagram](image.png)
A point \( x \) in a set \( K \) in Euclidean space is an extreme point if it is not contained in the interior of a line segment in \( K \). It is clear that the extreme points of a compact set \( K \) must lie on its frontier and that \( K \) is the convex hull of its extreme points, [51]. If \( \Omega \) is properly convex, a function \( f : \Omega \to \mathbb{R} \) satisfies the maximum principle if for every compact subset \( K \subset \Omega \) the restriction \( f|K \) attains its maximum at an extreme point of \( K \).

**Corollary 2.10** (Maximum principle) If \( C \) is a closed convex set in a properly convex domain \( \Omega \), then the distance of a point in \( C \) from \( \Omega \) satisfies the maximum principle.

**Proof** The function \( f(x) = d_\Omega(x,C) \) is 1-Lipschitz, therefore continuous. Let \( K \subset \Omega \) be a compact set then \( f|K \) attains its maximum at some point \( y \). There is a finite minimal set, \( S \), of extreme points of \( K \) such that \( y \) is in their convex hull. Choose \( y \) to minimise \( |S| \). If \( S \) contains more than one point then \( y \) is in the interior of a segment \( [a,b] \subset K \) with \( a \in S \) and \( b \in \) the convex hull of \( S' = S \setminus y \). Since \( C \) is closed and \( f \) is continuous there are \( c,d \in C \) with \( f(a) = d_\Omega(a,c) \) and \( f(b) = d_\Omega(b,d) \). Since \( C \) is convex \( [c,d] \subset C \).

Assume for purposes of contradiction that \( f(y) > f(a) = d_\Omega(a,[c,d]) \) and \( f(y) > f(b) = d_\Omega(b,[c,d]) \). Then we may find \( a',b' \) on \([a,b]\) such that \( y \in [a',b'] \) and \( f(a') = f(b') < f(y) \). By the 4-points lemma \( d_\Omega(y,[c,d]) \leq f(a') \). However, \([c,d] \subset C \) and so \( f(y) \leq d_\Omega(y,[c,d]) \), giving the contradiction \( f(y) \leq f(a') \).

**Corollary 2.11** (convexity of \( r \)-neighborhoods) If \( C \) is a closed convex set in a properly convex domain \( \Omega \) and \( r > 0 \), then the \( r \)-neighborhood of \( C \) is convex. In particular, an \( r \)-neighborhood of a line segment is convex.

**Lemma 2.12** (diverging lines) Suppose \( L \) and \( L' \) are two distinct line segments in a strictly convex domain \( \Omega \) which start at \( p \in \partial \Omega \). Let \( x(t) \) and \( x'(t) \) be parameterizations of \( L \) and \( L' \) by arc length so that increasing the parameter moves away from \( p \).

Then \( f(s) = d_\Omega(x(s),L') \) is a monotonic increasing homeomorphism \( f : \mathbb{R} \to (\alpha, \infty) \) for some \( \alpha \geq 0 \). Furthermore \( \alpha = 0 \) if \( p \) is a \( C^1 \) point.

**Proof** Refer to figure 2. We may reduce to two dimensions by intersecting with a plane containing the two lines. The function is 1-Lipschitz, thus continuous. Let \( x'(s') \) be some point on \( L' \) closest to \( x(s) \), and let \( \Omega_s \) be the subdomain of \( \Omega \) which is the triangle with vertices \( p,q(s),r(s) \) shown dotted. The following facts are evident. The distance between \( x(s) \) and \( x'(s') \) is the same in both \( \Omega \) and \( \Omega_s \). For \( t > 0 \) we have \( f(s-t) \leq d_{\Omega_{s-t}}(x(s-t),x'(s'-t)) \). Finally \( d_{\Omega_s}(x(s-t),x'(s'-t)) \) is constant for \( t > 0 \). The obvious comparison applied to triangular domains \( \Omega_s \) and \( \Omega_{s-t} \) gives the monotonicity statement.

If now \( p \) is a \( C^1 \) point, then there is an unique tangent line to \( \partial \Omega \) at \( p \) and the triangular domains have the angle at \( p \) increasingly close to \( \pi \). This implies that the distance tends to zero.

It only remains to show \( f \) is not bounded above. Let \( a(s) = |q(s) - x(s)| \) and \( b(s) = |r(s) - x'(s')| \). If \( f(s) = d_\Omega(x(s),x'(s')) \) is bounded above as \( s \to \infty \) then, using the cross ratio formula for distance and the fact \( x(s) - x'(s') \) is bounded away from zero, \( a(s) \) and \( b(s) \) are bounded away from \( 0 \). Using the fact that \( \Omega \) is convex, the limit as \( s \to \infty \) of the segment with endpoints \( q(s) \) and \( r(s) \) is a line segment in \( \partial \Omega \). 

---

Figure 2: Diverging Lines

---
\section*{2.10 The Hilbert metric as a Finsler metric or Riemannian metric}

Busemann and Kelly \cite{BusemannKelly} give a slightly more general definition of the Hilbert metric, which inserts a constant factor in the previous definition:

\[ d_\Omega(a,b) = \frac{k}{2} \log[x,a,b,y] = \frac{k}{2} \log \frac{\|b-x\| \cdot \|a-y\|}{\|b-y\| \cdot \|a-x\|} \]

where \( k > 0 \). For any properly convex domain \( \Omega \), this gives a complete Finsler metric with the Finsler norm of \( u \in \mathbb{R}^n \) at \( a \in \Omega \) given by:

\[ F_\Omega(a,u) = \frac{k\|u\|}{2} \left( \frac{1}{\|x-u^-\|} + \frac{1}{\|x-u^+\|} \right), \]

where \( u^- \) and \( u^+ \) are the intersection points of \( \partial \Omega \) with the line based at \( a \) in direction \( u \), such that \( u^+ \) is in the forwards direction. The length of a rectifiable path \( \gamma: I \to \Omega \) is then given by:

\[ L_\Omega(\gamma) = \int_I F_\Omega(\gamma(t), \gamma'(t)) \, dt. \]

(See Vernicos \cite{Vernicos} for a proof of this fact.) Some authors specialise to the case \( k = 1 \) (giving the hyperbolic metric in the Klein model) and others to the case \( k = 2 \) (eliminating constants in many expressions).

The differentiability class of the Finsler norm is precisely the differentiability class of \( \partial \Omega \). In general, convexity implies that the boundary of \( \Omega \) is \( C^1 \) almost everywhere with respect to Lebesgue measure on \( \mathbb{P}^{n-1} \). In order to apply techniques from Finsler geometry, one often needs to make additional hypotheses on the regularity of the boundary of \( \Omega \), generally making it at least of class \( C^2 \). See Socié-Méthou \cite{SociéMéthou} for a complete treatment of these and other aspects of the Finsler geometry of Hilbert geometries. In particular, the above metric (under the right regularity conditions) guarantees that the usual Finsler curvature is defined, making \( (\Omega, d_\Omega) \) into a Finsler manifold of constant negative curvature \(-\frac{1}{k^2}\).

The relationship between Hilbert metric and Riemannian geometry has been known for quite some time. The only proof in print appears to be due to Colbois and Verovic:

\begin{theorem}
A Hilbert geometry \( (\Omega, d_\Omega) \) is Riemannian if and only if \( \Omega \) is an ellipsoid. In this case, \( (\Omega, d_\Omega) \) is isometric with hyperbolic space.
\end{theorem}

The up-shot of this discussion is that if one imposes extra regularity conditions on \( \partial \Omega \), then many technical obstacles will either not occur or may become easier to navigate as there are additional techniques from Finsler geometry at one’s disposal.
3 Projective manifolds

The only Hilbert geometries that are isometric to either a normed vector space or a Riemannian manifold are those associated with a simplex and an ellipsoid respectively. On the other hand, every Hilbert geometry is both a stiffening of projective geometry as well as a Finsler geometry. This special place of Hilbert geometry should justify the development of intrinsic tools, whilst providing important examples for the more general settings.

But does it?

The first section stated that geometry is the study of properties invariant under a group of transformations. The second section introduced Hilbert geometries and derived some general facts about their groups of transformations. However, for many properly convex domains, the group of projective automorphisms fixing them will just contain the trivial element, and domains having a transitive group action are rare. The goal of this section is to show that there are indeed examples that motivate and justify the development of a rich theory. In doing so, the concepts of developing map and holonomy representation, which are fundamental in many areas of geometry, are introduced.

References: Goldman [4, 44], Thurston [5], Cooper, Long and Tillmann [2], Sullivan and Thurston [74]

3.1 Gluing constructions

I constructed geometric structures on surfaces by gluing opposite sides of a square or an octagon, giving the following examples:

1. a Euclidean structure on the torus;
2. a Hex structure on the torus;
3. a hyperbolic structure on the once-punctured torus.

The ideas developed in these examples are now formalised.

Moreover, it is not difficult to show that there is a hyperbolic structure on any closed surface of negative Euler characteristic, by considering $4n$-gons in the Poincaré disc. More generally, the uniformisation theorem of complex analysis implies that every surface admits a spherical, euclidean or hyperbolic structure, and hence a real projective structure.

3.2 Real projective structure

A projective atlas of the $n$–manifold $M$ is a collection of maps

$$\{\Phi: U \rightarrow V \subseteq \mathbb{P}^n\},$$

where the sets $U$ are open and their union covers $M$, and the maps $\Phi$ are homeomorphisms. A real projective structure on $M$ is a projective atlas with the property that all transition maps

$$\Phi_j \circ \Phi_i^{-1}: \Phi_i(U_i \cap U_j) \rightarrow \Phi_j(U_i \cap U_j)$$

are the restrictions of projective maps. It follows from Theorem 1.9 that this condition only needs to be verified on $\Phi_i(U_i \cap U_j)$. A given manifold may have more than one projective structure; a manifold with a fixed projective structure is called a real projective manifold.
3.3 Developing map and holonomy

By “unrolling” a real projective manifold \( M \) in \( \mathbb{P}^n \), one obtains the important notions of a \textit{developing map} 

\[
dev: \tilde{M} \to \mathbb{P}^n,
\]

where \( \tilde{M} \) is the universal cover of \( M \), and a \textit{holonomy representation} 

\[
\text{hol}: \pi_1(M) \to \text{PGL}(n+1).
\]

I will not defined these carefully (see Thurston [5] for definitions and discussion). They were introduced informally by the examples in §3.1. I will now give more examples that help clarify the concepts:

1. an affine structure on the torus that is not euclidean;
2. a projective structure on the torus that is not affine.

A proof of the following fact due to Benzecri can be found in Goldman’s notes [4]:

**Theorem 3.1** (Benzecri 1955) Let \( S \) be a closed surface with an affine structure. Then \( \chi(S) = 0 \).

The affine and projective structures that I constructed on the torus had the same holonomy. To sum up, we have now seen examples of the following situations:

1. The developing map is not a covering map.
2. The holonomy representation is not injective.
3. The holonomy representation does not uniquely determine the geometric structure.

The next section shows under which conditions we can think of a geometric structure as a quotient \( \Omega/\Gamma \).

3.4 Properly discontinuous actions

In the examples, we have obtained geometric structures on surfaces by gluing sides of polygons, and obtained a tiling group generated by the face pairings. In general, we would like to start with a group action and have a criterion that tells us when the quotient space is a manifold or orbifold. This will now be given.

A group, \( G \), of homeomorphisms of a locally compact Hausdorff space \( X \) acts \textit{properly discontinuously} if for every compact \( K \subset X \) the set \( K \cap gK \) is non-empty for at most finitely many \( g \in G \).

Now suppose that \((X,d)\) is a metric space and \( G \) is a group of isometries of \( X \). Then the quotient semi-metric \( \overline{d} \) on \( X/G \) is given by:

\[
\overline{d}(\overline{x},\overline{y}) = \inf\{d(x,\gamma \cdot y) \mid \gamma \in G\},
\]

where \( x \mapsto \overline{x} \) is the quotient map \( X \to X/G \). If the action of \( G \) on \( X \) is properly discontinuous, then \( \overline{d} \) is a metric. If the action is properly discontinuous and free (\( gx = x \iff g = 1 \)), then \( X/G \) is locally isometric to \( X \). Details can be found in [5], [28] and [29].

**Proposition 3.2** [2] Let \( \Omega \) be a properly convex domain and \( \Gamma \subset \text{PGL}(\Omega) \). Then \( \Gamma \) is a discrete subgroup of \( \text{PGL}(n+1) \) if and only if \( \Gamma \) acts properly discontinuously on \( \Omega \).

**Proof** Suppose there is a sequence of distinct elements \( \gamma_i \in \Gamma \) converging to the identity in \( \text{PGL}(n+1) \). Let \( K \subset \Omega \) be a compact set containing \([v]\) in its interior. Then \( \gamma_i[v] \in K \) for all sufficiently large \( i \) so \( \Gamma \) does not act properly discontinuously. Conversely, suppose \( K \subset \Omega \) is compact and there is a sequence of distinct elements \( \gamma_i \in \Gamma \) with \( K \cap \gamma_i K \neq \emptyset \). Choose a projective basis \( B = (x_0, \ldots, x_n) \subset \Omega \) with \( x_0 \in K \). After taking a subsequence we may assume \( \gamma_i B \) converges to a subset of \( \Omega \). The sequence \( \delta_i = \gamma_i^{-1} \gamma_i \in \Gamma \) has the property \( \delta_i B \to B \) because \( \delta_i \) is an \textit{isometry}. By Theorem 1.10, this implies \( \delta_i \) converges to the identity. \( \blacksquare \)
The manifolds of interest will come to us as quotients of properly convex domains \( \Omega \subset \mathbb{P}^n \). Recall that the triple

\[
X_\Omega = (\Omega, d_\Omega, \text{PGL}(\Omega))
\]

is called a Hilbert geometry. The manifold \( M \) is a \textit{properly convex projective manifold} or (more specifically) has an \( X_\Omega \)-\textit{structure} if \( M = \Omega/\Gamma \), where \( \Gamma \leq \text{PGL}(\Omega) \) is discrete and acts freely. In particular, \( M \) is locally isometric to \((\Omega, d_\Omega)\). If \( \Omega \) is strictly convex, we say that \( M \) is strictly convex.

More generally, a \textit{properly convex projective orbifold} is \( Q = \Omega/\Gamma \), where \( \Omega \) is an open properly convex set and \( \Gamma \leq \text{PSL}(\Omega) \) is a discrete group. Similarly for \textit{strictly convex}. This orbifold is a manifold if and only if \( \Gamma \) is torsion free.

Since points in \( \Omega^* \) are the duals of hyperplanes disjoint from \( \Omega \) it follows that under the dual action \( \text{PSL}(\Omega) \) preserves \( \Omega^* \). Thus given a properly convex projective orbifold \( Q \), there is a dual orbifold \( Q^* = \Omega^*/\Gamma^* \). Two projective orbifolds \( \Omega/\Gamma \) and \( \Omega'/\Gamma' \) are \textit{projectively equivalent} if there is a homeomorphism between them, which is covered by the restriction of a projective transformation mapping \( \Omega \) to \( \Omega' \). Cooper and Delp [36] have shown that in general \( Q \) is not projectively equivalent to \( Q^* \).

### 3.5 Deformation spaces

We have seen surfaces that admit inequivalent projective structures. Even if one restricts to convex projective structures, there turns out to be an abundance of structures, as shown in the following result.

Given a convex projective structure on a surface \( S \), the holonomy representation \( \text{hol}: \pi_1(S) \rightarrow \text{PGL}(3) \) is well defined up to conjugacy in \( \text{PGL}(3) \), and the set of projective equivalence classes of convex projective structures \( \mathcal{D}(S) \) can be identified with an open subset of \( \text{Hom}(\pi_1(S), \text{PGL}(3))/\text{PGL}(3) \). We also call \( \mathcal{D}(S) \) the deformation space of convex projective structures.

**Theorem 3.3** (Goldman [44]) Let \( S \) be a closed surface with \( \chi(S) < 0 \). Then \( \mathcal{D}(S) \) is diffeomorphic to \( \mathbb{R}^{8\chi(S)} \).

In contrast, the deformation space of a torus is neither Hausdorff nor a non-Hausdorff manifold. The reader is strongly encouraged to study Goldman’s paper. To illustrate the idea of a deformation space of geometric structures, we will first look at euclidean structures on the torus. The analysis here is simplified by a number of facts: the domain \( \Omega \) does not vary, but can be taken as a fixed affine patch; the groups we consider are subgroups of the smaller group \( \text{Euc}(n) \); the fundamental group of the torus is the free abelian group of rank 2.

### 3.6 Teichmüller space and moduli space of the torus

On considers two euclidean structures on the torus equivalent, if they are \textit{similar}, meaning that they are isometric after possibly scaling one of them. So what matters is \textit{shape} and not \textit{size}.

We identify the euclidean plane with the vector space \( \mathbb{R}^2 \) with the geometry of the euclidean inner product. A euclidean torus is of the form \( \mathbb{R}^2/\Gamma \), where \( \Gamma = \langle a_1, a_2 \rangle \cong \mathbb{Z}^2 \) is a lattice in \( \mathbb{R}^2 \). Now \( \mathbb{R}^2 \) is oriented using the standard basis \( \{e_1, e_2\} \), and the torus is given the induced orientation. We can therefore assume that the vectors \( a_1 \) and \( a_2 \) are such that we can rotate the parallelogram spanned by them so that \( a_1 \) is a positive multiple of \( e_1 \) and \( a_2 \) is in the upper half plane.

Up to scaling we may assume \( a_1 = e_1 \), and we identify \( a_2 = (x, y) \) with the corresponding complex number \( z = x + iy \in \mathbb{H}^2 = \{z \in \mathbb{C} \mid \Im(z) > 0\} \). Whence the open upper half plane parameterises all euclidean structures on the torus (up to similarity) given a fixed basis for the torus. This is called the Teichmüller space of the torus. A point not only describes a euclidean structure on the torus, but also how it is worn. A common analogy for
this is the situation, where one puts a stripy jumpsuit on a toddler; if the toddler wiggles a lot whilst putting it on, a leg may be twisted, and this is difficult to untwist once the heel has locked into place.

Let us now ignore the way a metric is worn; the space parametrising euclidean structures up to similarity is called the moduli space. An orientation preserving change of basis corresponds to an element of \( \text{SL}(2, \mathbb{Z}) \). If \( z \in \mathbb{H}^2 \) parameterises a given euclidean structure on the torus, then there is an isometry between the structure determined by \( z \) and the structure determined by

\[
A \cdot z = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az + b}{cz + d}.
\]

It turns out that the converse is also true: given an isometry between two euclidean structures on the torus, there is an associated element of \( \text{SL}(2, \mathbb{Z}) \) relating them as above. The action of \( A \) on \( \mathbb{H}^2 \) is the same as the action of \( -A \), and one obtains the following result:

**Theorem 3.4** The moduli space of all euclidean structures of the torus is \( \mathbb{H}^2 / \text{PSL}(2, \mathbb{Z}) \).

The moduli space is a hyperbolic orbifold with total space a once-punctured disc and one cone point of order 2 and one cone point of order 3.

### 3.7 Teichmüller space and moduli space of hyperbolic surfaces

For a closed orientable surface \( S \) of negative Euler characteristic, one also has a Teichmüller space and a moduli space. Since these spaces were introduced informally for the torus, here is a precise definition:

Two hyperbolic metrics \( d_0 \) and \( d_1 \) on the surface \( S \) are Teichmüller equivalent if there is an isometry

\[
f: (S, d_0) \to (S, d_1),
\]

such that \( f \) is homotopic to the identity map. The metrics are moduli equivalent if there is an isometry \( f: (S, d_0) \to (S, d_1) \). The set of all Teichmüller equivalence classes is called the Teichmüller space, denoted \( \text{Teich}(S) \), and the set of all moduli equivalence classes is called the moduli space of \( S \), denoted \( \text{Mod}(S) \).

The relationship between these spaces is straightforward: \( \text{Mod}(S) \) is the quotient of \( \text{Teich}(S) \) under the action of the group of homotopy classes of homeomorphisms \( S \to S \) (called the mapping class group of \( S \)).

**Theorem 3.5** (Fricke–Klein 1912) Teichmüller space \( \text{Teich}(S) \) is naturally homeomorphic with \( \mathbb{R}^{-3\chi(S)} \).

Every hyperbolic structure on \( S \) is a strictly convex projective structure on \( S \), and two hyperbolic structures are equivalent as hyperbolic structures if and only if they are equivalent as projective structures. One therefore obtains an embedding \( \text{Teich}(S) \to \mathcal{D}(S) \). Being able to deform hyperbolic metrics to more general projective metrics (and the much greater dimension of the deformation space) is one indication of the scope of projective geometry.

### 3.8 Mostow-Prasad rigidity

The following result implies that the analogue of Teichmüller space for hyperbolic manifolds of higher dimensions consists of just a point:

**Theorem 3.6** (Mostow-Prasad) The hyperbolic structure on a closed, connected, orientable, hyperbolic \( n \)-manifold with \( n \ge 2 \) is unique up to isometry homotopic to the identity.

However, projective geometry again brings additional aspects to the study of these manifolds as some closed hyperbolic manifolds admit projective deformations and others do not; see [38, 37]. It is not known what properties make a projective manifold rigid.

**Remark 3.7** Examples are known, in all dimensions \( \ge 2 \), of strictly convex projective manifolds that do not admit a hyperbolic structure but have Gromov hyperbolic fundamental group. See Kapovich [49] and Marquis [58] for two different constructions.
4 Hilbert geometries on large and small scales

It is natural to ask what else can be said about it a Hilbert geometry in general (both on the large scale and on the small scale). Two notions that capture this information are Gromov hyperbolicity and the CAT($k$) property. It turns out that the only Hilbert geometry that satisfies the CAT($k$) property is hyperbolic geometry. Then a compactness result due to Benzécri is used to show that any Hilbert geometry is $K$–quasi-homogeneous even though its isometry group does not act transitively in general. The last section contains a construction due to Vinberg, which shows that there is a bilipschitz map between a convex projective orbifold and its dual.

References: Goldman [4, 44], Socié-Méthou [3], Vernicos [7], Cooper, Long and Tillmann [2]

4.1 Gromov hyperbolicity

A metric space $(X,d)$ is a geodesic metric space if every pair of points is connected by a geodesic. A triangle in a geodesic metric space consists of three geodesics arranged in the usual way.

A triangle is $\delta$–thin if every point on each side of the triangle is within distance $\delta$ of the union of the other two sides. A triangle that is not $\delta$–thin is called $\delta$-fat. If $(X,d)$ is a locally compact, complete geodesic metric space and every triangle in $(X,d)$ is $\delta$–thin then $(X,d)$ is called $\delta$-hyperbolic or Gromov hyperbolic.

These ideas can be applied to a properly convex domain with the Hilbert metric. Some care is required with terminology in view of the fact that if $\Omega$ is strictly convex, then geodesics are precisely projective line segments, otherwise if $\Omega$ is only properly convex, there may be geodesics which are not segments of projective lines, and triangles with geodesic sides which are not planar. These complications are in fact mythological in the presence of $\delta$–hyperbolicity:

**Theorem 4.1** ($\delta$–hyperbolic implies strictly convex) If $(\Omega,d_\Omega)$ is $\delta$–hyperbolic, then $\Omega$ is strictly convex.

Necessary and sufficient conditions for $\delta$–hyperbolicity in terms of orbits of domains were given by Benoist [20]. Sufficient conditions for $\delta$–hyperbolicity in terms of group actions can be found in [21] and [2].

The notion of $\delta$–hyperbolicity includes hyperbolic manifolds, but applies to a larger class of manifolds. Benoist [24] has shown that there are $\delta$–hyperbolic domains which have quotient manifolds that do not admit a complete hyperbolic structure.

4.2 CAT(0)

The notion of $\delta$–hyperbolicity only captures the large scale geometry of a space; it does not see small deviations. Fine details are seen by a property termed CAT($k$), which compares a metric space with the simply connected Riemannian manifold of constant curvature $k$. I will only discuss the case $k = 0$, where the metric space is compared with Euclidean geometry. Generally, if a space is CAT($k$), then it is CAT($k'$) for any $k' \leq k$.

The geodesic metric space $(X,d)$ is CAT(0) if it has the following property. Given any geodesic triangle $\Delta(abc)$, construct a comparison triangle $\Delta(d'b'c')$ in $\mathbb{R}^2$ which has corresponding sides of the same lengths, e.g. $d(a,b) = ||b - a||$. Choose any point $C$ on the side $ab$ and any point $B$ on the side $ac$, and take the corresponding points $C'$ and $B'$ in the comparison triangle that have the same distances to the respective vertices. Then:

$$d(B,C) \leq ||C' - B'||.$$

**Theorem 4.2** (Vernicos [7]) A Hilbert geometry $(\Omega,d_\Omega)$ is CAT($k$) for some $k \in \mathbb{R}$ if and only if $\Omega$ is an ellipsoid.
4.3 Volume

Volume in a Hilbert geometry is defined using Hausdorff measure; an excellent treatment of this can be found in Stein and Shakarchi [73]. Here is a summary. For a set $S$ in a metric space $(X,d)$, its diameter is

$$\text{diam} S = \sup \{ d(x,y) \mid x,y \in S \}.$$ 

Then the $\alpha$–dimensional Hausdorff measure of a subset $Y \subset X$, is defined by first putting

$$\mathcal{H}_\alpha(Y) = \inf \left\{ \sum_k (\text{diam} F_k)^\alpha \mid Y \subset \bigcup_{k=1}^\infty F_k, \text{diam} F_k \leq \delta \text{ for all } k \right\},$$

and then letting

$$\mu_\alpha(Y) = \lim_{\delta \to 0} \mathcal{H}_\alpha(Y).$$

This exists (allowing the value $\infty$) because the quantity $\mathcal{H}_\alpha(Y)$ increases as one decreases $\delta$. For any measurable set $Y$, there is a unique $\alpha$ such that

$$\mu_\beta(Y) = \begin{cases} \infty & \text{if } \beta > \alpha, \\ 0 & \text{if } \beta < \alpha. \end{cases}$$

This number $\alpha$ (which is not necessarily an integer) is called the Hausdorff dimension of $Y$.

Specialising to a properly convex domain $\Omega \subset \mathbb{P}^n$ with the Hilbert metric, it turns out that the Hausdorff dimension of any open set is $n$. We have the $n$–dimensional Hausdorff measure $\mu_\Omega = \mu_n$ on $\Omega$, which depends only on $d_\Omega$.

It is not difficult to see that $\mu_\Omega(\Omega) = \infty$. Corollary 4.8 below shows that $\mu_\Omega$ is absolutely continuous with respect to Lebesgue measure $\mu_\mathcal{X}$ (i.e. $\mu_\mathcal{X}(Y) = 0$ implies $\mu_\Omega(Y) = 0$ for any set $Y$ that is measurable with respect to both measures) and vice versa.

Since the measure is defined in terms of the metric and $\text{PGL}(\Omega)$ acts by isometries, it preserves the measure, and one therefore obtains a well-defined measure on any projective manifold modelled on a Hilbert metric. We can therefore distinguish between finite volume projective manifolds and infinite volume projective manifolds.

4.4 Benzécri’s compactness result

In studying the deformation of one properly convex projective structure on a manifold into another, one faces the difficulty that not only the holonomy varies, but also the image of the developing map. A fundamental tool in overcoming this obstacle is a compactness result due to Benzécri [25]. A simplified proof of his result can be found in Goldman [4] pages 49–63. Most of the following summary can be found in [2], §6.

The set of all compact subsets of the metric space $(X,d)$ is made into a metric space as follows. Denote the $\delta$–neighbourhood of a subset $A \subset X$ by:

$$N_\delta(A) = \{ x \in X \mid d(x,A) \leq \delta \}.$$ 

For compact subsets $A$ and $B$ of $X$ define

$$d_H(A,B) = \inf \{ \delta \mid A \subset N_\delta(B) \text{ and } B \subset N_\delta(A) \}.$$ 

This defines the Hausdorff metric (and hence the Hausdorff topology) on the set of all compact subsets of $X$.

Let $\mathcal{C}$ be the set of all properly convex compact subsets in $\mathbb{P}^n$ with non-empty interior and equip this with the Hausdorff topology (using the elliptic metric on $\mathbb{P}^n$). Since we are interested in studying Hilbert metrics up to equivalence, one may ask what properties the quotient space $\mathcal{C}/\text{PGL}(n+1)$ has. The group $\text{PGL}(n+1)$ acts continuously on $\mathcal{C}$.

Suppose $\Omega \subset \mathbb{P}^2$ is a properly convex domain whose boundary contains a point which is not $\mathcal{C}^1$. Up to projective isometries, one may assume that this point is $[e_1]$ and that $\Omega$ is contained in the triangle $\Delta = \Delta([e_1],[e_2],[e_3])$. Letting $\gamma_n = [\text{diag}(n^{-2},n,n)]$, one sees that the sequence of domains $\gamma_n \Omega$ converges to $\Delta$. 

21
Exercise 4.3  Show that if $\Omega \subset P^2$ is not strictly convex, then there is a sequence $\gamma_t \in \text{PGL}(3)$ such that $\gamma_t \overline{\Omega}$ converges to $\overline{\Delta}$.

Exercise 4.4  Show that if $\Omega \subset P^2$ has $C^1$ boundary, but there is a point which is $C^2$, then there is a sequence $\gamma_t \in \text{PGL}(3)$ such that $\gamma_t \overline{\Omega}$ converges to the (orbit of) the domain $C = \{ [x : y : z] \mid xy - z^2 = 0 \}$.

The examples show that the orbits of proper domains under $\text{PGL}(n+1)$ are generally not closed, and hence the quotient space $\mathcal{C}/\text{PGL}(n+1)$ will not be a Hausdorff space. The important thing to notice in the examples is that under the sequence $\gamma_t$ any fixed point in the interior of $\Omega$ will converge to the boundary of $\Delta$. This observation leads to the following set-up.

Let $\mathcal{C}_s$ be the space of all $(C, p) \in \mathcal{C} \times P^n$ with $p$ a point in the interior of $C$ and equipped with the product topology.

Theorem 4.5  (Benzécri compactness)  The quotient of $\mathcal{C}_s$ by the natural action of $\text{PGL}(n+1)$ is a compact Hausdorff space.

Given a metric space $X$ with metric $d$ the closed ball in $X$ center $p$ radius $r$ is

$$B_r(p; X, d) = \{ x \in X : d(x, p) \leq r \}.$$  

In what follows $B(r)$ denotes the closed ball of Euclidean radius $r$ centered on the origin in Euclidean space.

Corollary 4.6  (Benzécri charts, [4] page 61 C.24)  For every $n \geq 2$ there is a constant $R_\emptyset = R_\emptyset(n) > 1$ with the following property: If $\Omega \subset P^n$ is a properly convex open set and $p \in \Omega$ then there is a projective automorphism $\tau$ called a Benzécri chart such that $B(1) \subset \tau(\Omega) \subset B(R_\emptyset) \subset P^n$ and $\tau(p) = 0$.

An open convex set $\Omega$ is called a Benzécri domain if $B(1) \subset \overline{\Omega} \subset B(R_\emptyset(n))$.

Proposition 4.7  Let $\mathcal{B}$ be the set of all Benzécri domains in $P^n$. Then $\mathcal{B}$ is compact with the Hausdorff metric induced by the Euclidean metric on $P^n$.

Corollary 4.8  (Hilbert balls are uniformly bilipschitz)  For every dimension $n \geq 2$ and every $r > 0$:

- There is $K = K(n, r) > 0$ such that for every properly convex domain $\Omega \subset P^n$ and $p \in \Omega$ there is a $K$-bilipschitz homeomorphism from $B_r(p; \Omega, d_\Omega)$ to $B(r)$.
- There is $K_\mu = K_\mu(n, r) > 0$ such that if $\Omega$ is a Benzécri domain and $\mu_\Omega$ is the Hausdorff measure on $\Omega$ induced by the Hilbert metric and $\mu_\mu$ is Lebesgue measure on $P^n$ then for every open set $U \subset B_r(0; \Omega, d_\Omega)$

$$K_\mu^{-1} \cdot \mu_\mu(U) \leq \mu_\Omega(U) \leq K_\mu \cdot \mu_\mu(U).$$

As a result of the corollary, $(\Omega, d_\Omega)$ is $K$–quasi-homogeneous for every properly convex domain $\Omega$ even though $\text{Isom}(\Omega, d_\Omega)$ does not act transitively in general.

4.5  The Vinberg hypersurface

As the name suggests, the hypersurface described in this section is due to Vinberg. Simplified proofs of below summary of results (which is again lifted from [2], §6) can be found in Goldman [4] pages 49–63.

Suppose $\mathcal{C} = \mathcal{C}(\Omega) \subset V$ is a sharp convex cone and $\mathcal{C}^* \subset V^*$ is the dual cone. Let $d\psi$ be a volume form on $V^*$. The characteristic function $f : \mathcal{C} \to \mathbb{R}$ defined by

$$f(x) = \int_{\mathcal{C}} e^{-\psi(x)} \, d\psi$$

is real analytic and $f(tx) = t^{-1} f(x)$ for $t > 0$. For each $t > 0$ the level set $S_t = f^{-1}(t)$ is called a Vinberg hypersurface. It is the boundary of the sublevel set $\mathcal{C}_t = f^{-1}(0, t] \subset \mathcal{C}$. For example, the hyperboloids $z^2 = x^2 + y^2 + t$ are Vinberg hypersurfaces in the cone $z^2 > x^2 + y^2$.  

22
Theorem 4.9 (Vinberg [76], see also [4] (C1), (C6) pages 51–52) The Vinberg hypersurfaces are an analytic foliation of $\mathcal{C}$.

- The radial projection $\pi : S_t \rightarrow \Omega$ is a diffeomorphism.
- $\mathcal{C}_t$ has smooth strictly convex boundary.
- $S_t$ is preserved by $SL(\mathcal{C})$.

At each point $p$ on a Vinberg surface there is a unique supporting tangent hyperplane $\ker df_p$. This gives a duality map $\Phi_\Omega : \Omega \rightarrow \Omega^*$. Another description of this map is that $\Phi_\Omega(x)$ is the centroid of the intersection of $\mathcal{C}^*$ with the hyperplane $\{\psi \in V^* \mid \psi(x) = n\} \subset V^*$. Benzécri’s compactness theorem has the following consequences:

Theorem 4.10 $\Phi_\Omega$ is $K$-bilipschitz with respect to the Hilbert metrics where $K = K(n)$ only depends on $n = \dim \Omega$.

Corollary 4.11 The duality map descends to a $K$-bilipschitz map between a properly convex orbifold $M$ and its dual $M^*$. In particular, $M$ has finite volume if and only if $M^*$ has finite volume.
5 Isometries of the Hilbert metric

This section gives a very brief introduction to the study of isometries of the Hilbert metric.

References: Cooper, Long and Tillmann [2], Goldman [4]

5.1 Elliptic, parabolic and hyperbolic

Let \( \Omega \subset \mathbb{P}^n \) be a properly convex open set and \( \varphi \in \text{PGL}(\Omega) \) be a projective isometry of its Hilbert metric. Since \( \overline{\Omega} \) is a closed ball, the Brouwer fixed point theorem implies that \( \varphi \) has a fixed point in \( \overline{\Omega} \). We say that \( \varphi \) is elliptic if \( \varphi \) fixes a point in \( \Omega \). If \( \varphi \) does not fix a point in \( \Omega \) (i.e. it acts freely on \( \Omega \)) we turn to its eigenvalues to make further distinctions. To facilitate this, we replace \( \Omega \) by a lift to \( S^n \) and \( \text{PGL}(\Omega) \) by \( \text{SL}_\pm(\Omega) \).

We then say that a non-elliptic isometry \( \varphi \in \text{SL}_\pm(\Omega) \) is a parabolic if every (possibly complex) eigenvalue has modulus 1, and else it is hyperbolic.

Recall that \( \text{SL}_\pm(\Omega, p) \subseteq \text{SL}_\pm(\Omega) \) is defined as the subgroup which stabilises both \( \Omega \) and \( p \). It is easy to see that if \( p \in \Omega \), then this group is compact. This gives the following, reassuring fact:

**Lemma 5.1** (Elliptics are standard) If \( \Omega \) is a properly convex domain, then \( A \in \text{SL}_\pm(\Omega) \) is elliptic if and only if it is conjugate in \( \text{SL}(n+1) \) into \( \text{O}(n+1) \). Furthermore, if \( p \in \Omega \), then \( \text{SL}_\pm(\Omega, p) \) is conjugate in \( \text{SL}(n+1) \) into \( \text{O}(n+1) \).

5.2 Fixed point sets

A fixed point of \( \varphi \in \text{SL}_\pm(\Omega) \) (elliptic or not) corresponds to the span of an eigenvector \( [v] \in \mathbb{P}^n \). As we have passed to the double cover, notice that the associated eigenvalue \( \lambda \) is necessarily positive, so \( \varphi v = \lambda v \) with \( \lambda > 0 \). Denote the \( \lambda \)–eigenspace \( E_\lambda(\varphi) \), let

\[
\text{Fix}_\Omega(\varphi) = E_\lambda(\varphi) \cap \overline{\Omega},
\]

and

\[
\text{Fix}_\Omega(\varphi) = \bigcup_{\lambda > 0} \text{Fix}_\Omega(\varphi) \subseteq \overline{\Omega},
\]

where the union is taken over all positive eigenvalues of \( \varphi \). For a given positive eigenvalue, the set \( \text{Fix}_\Omega(\varphi) \) may be empty; if it is not empty, then it is convex. Moreover, as observed above, \( \text{Fix}_\Omega(\varphi) \) is non-empty by the Brouwer fixed point theorem.

**Exercise 5.2** Determine projective isometries of the Hex plane \( (\Delta, d_\Delta) \) with the following properties:

1. A hyperbolic isometry that fixes exactly the 3 vertices.
2. A hyperbolic isometry that fixes one vertex and the opposite edge.
3. An elliptic isometry without fixed points on the boundary

Is there a parabolic isometry?

5.3 Translation length and minset

Let \( \Omega \subseteq \mathbb{P}^n \) be an open properly convex domain. The translation length of \( A \in \text{SL}_\pm(\Omega) \) is

\[
t(A) = \inf_{x \in X} d_\Omega(x, Ax).
\]

The subset of \( \Omega \) for which this infimum is attained is called the minset of \( A \). It might be empty. For an elliptic element, we have \( t(A) = 0 \) and the minset is precisely the set of fixed points in \( \Omega \).

Later we derive the following algebraic formula for translation length which implies parabolics also have translation length zero, and hyperbolics have positive translation length. The following result is proved in [2]:
Proposition 5.3 Let $\Omega \subseteq \mathbb{P}^n$ be an open properly convex domain and $A \in \text{SL}_\pm(\Omega)$. Then

$$t(A) = \log \frac{\lambda}{\mu},$$

where $\lambda$ and $\mu$ are eigenvalues of $A$ of maximum and minimum modulus respectively.

Example 5.4 The element $A = [\text{diag}(e^a, e^b, e^c)] \in \text{PGL}(\Delta)$ is a hyperbolic. Suppose $a \geq b \geq c$. Then the translation length of $A$ is $t(A) = a - c$ according to the above formula. Determine this directly by showing that $A$ corresponds to the translation on $(\mathbb{R}^2, \| \cdot \|_{\text{hex}})$ given by

$$x \mapsto x + (aa_1 + ba_2 + ca_3).$$

Determine both the minset and the translation length for the latter.

We have seen that making a domain smaller increases distances between points. This can be used to exhibit hyperbolic elements with empty minset:

Exercise 5.5 Show that the domain $\Omega = \{(x, y) \mid xy > 1\}$ is projectively equivalent to a subset of the hex plane $\Delta$. Show that $A \in \text{SL}_\pm(\Omega)$ given by $A(x, y) = (2x, y/2)$ is a hyperbolic (and hence has translation length $\log 4$), but that is has has empty minset.

5.4 Fixed points

The next step is to describe the fixed points in $\partial \overline{\Omega}$ and the dynamics of a projective isometry.

Lemma 5.6 Suppose $\Omega$ is a strictly convex domain. A non-elliptic element of $\text{SL}_\pm(\Omega)$ fixes precisely one point in $\partial \Omega$ if it is parabolic, and fixes precisely two points in $\partial \Omega$ if it is hyperbolic.

Proof Let $\gamma \in \text{SL}_\pm(\Omega)$ be a non-elliptic element, and suppose that $\phi$ fixes two points in $\partial \Omega$. By hypothesis, the projective line through these points meets $\Omega$ in an open interval. Whence $\gamma$ must act as a translation along this line. This implies that the two points correspond to eigenvectors with distinct positive real eigenvalues, and hence at least one of them is not 1. This shows that $\gamma$ is hyperbolic.

Now suppose that $\gamma$ also fixes a third point on $\partial \Omega$. Let $\Delta = \Delta(a, b, c)$ be the triangle in $\overline{\Omega}$ with vertices the three fixed points. Since $\Omega$ is strictly convex, the interiors of the three sides of $\Delta$ are in $\Omega$. Moreover, $\gamma$ maps each side of $\Delta$ to itself by a translation. Pick a point $p$ on the side $[a, b]$, and choose the point $q$ on $[a, c]$ closest to $p$. This point is unique since metric balls in $\Omega$ are strictly convex. Now $\gamma(p) \in [a, b]$ and $\gamma(q) \in [a, c]$ and $d_{\Omega}(\gamma(p), \gamma(q)) = d_{\Omega}(p, q)$. Since $\gamma$ is not elliptic, the segments $[\gamma^{-n}(p), \gamma^n(p)]$ cover $[a, b]$. It follows from Lemma 2.9 that $[a, b]$ and $[a, c]$ are within bounded distance of each other. This contradicts Lemma 2.12.

It follows that if the non-elliptic element $\gamma$ has two fixed points on $\partial \Omega$, then it has precisely two fixed points and is hyperbolic. Whence it is parabolic if it has exactly one fixed point.

The following characterisation of parabolics in general is given in [2]:

Proposition 5.7 Suppose $\Omega$ is a properly convex domain and $A \in \text{SL}(\Omega, p)$ is not elliptic. The following are equivalent:

1. $A$ is parabolic,
2. every eigenvalue has modulus 1,
3. the translation length $t(A) = 0$,
4. the subset of $\partial \overline{\Omega}$ fixed by $A$ is non-empty, convex and connected,
5.5 Dynamics

The next results from [2] gives a picture of the dynamics of a projective isometry. If \( W \) is a codimension-2 projective subspace then the set of codimension-1 projective hyperplanes containing \( W \) is called a pencil of hyperplanes and \( W \) is the center of the pencil. The hyperplanes in the pencil are dual to a line \( W_\perp \) in the dual projective space.

**Proposition 5.8** (isometry permutes pencil) *Suppose that \( \Omega \) is a properly convex domain and \( A \in \text{SL}_\pm(\Omega) \) is a parabolic or hyperbolic. Then there is a pencil of hyperplanes that is preserved by \( A \). The intersection of this pencil with \( \Omega \) is a foliation and no leaf is stabilized by \( A \). Thus \( M = \Omega / \langle A \rangle \) is a bundle over the circle with fibers subsets of hyperplanes.*

![Figure 3: Pencils of hyperplanes](image)

This and Lemma 2.12 imply:

**Corollary 5.9** *If \( \Omega \) is strictly convex and \( A \in \text{SL}_\pm(\Omega) \) is not elliptic, then \( f(x) = d_{\Omega}(x, Ax) \) is not bounded above.*
6 The Margulis lemma

This section gives the proof of the Margulis lemma for properly convex projective manifolds (straight from §7 of [2]) and briefly describes one of its main applications, the thick-thin decomposition of strictly convex projective manifolds.

References: Cooper, Long and Tillmann [2], Thurston [5]

6.1 The Margulis lemma

Theorem 6.1 (properly convex Margulis) For each dimension \( n \geq 2 \) there is a Margulis constant \( \mu_n > 0 \) with the following property. If \( M \) is a properly convex projective \( n \)-manifold and \( x \) is a point in \( M \), then the subgroup of \( \pi_1(M,x) \) generated by loops based at \( x \) of length less than \( \mu_n \) is virtually nilpotent. In fact, there is a nilpotent subgroup of index bounded above by \( m = m(n) \).

Furthermore, if \( M \) is strictly convex and of finite volume, this nilpotent subgroup is abelian. If \( M \) is strictly convex and closed, this nilpotent subgroup is trivial or infinite cyclic.

The remainder of this subsection is devoted to the proof of the first part of this theorem.

Theorem 6.2 (Isometry Bound) For every \( d > 0 \) there is a compact subset \( K \subset \text{SL}_\pm(n+1) \) with the following property. Suppose that \( \Omega \) is a Benzecri domain and \( A \in \text{SL}_\pm(\Omega) \) moves the origin a distance at most \( d \) in the Hilbert metric on \( \Omega \). Then \( A \in K \).

Proof Let \( p \) denote the origin. Suppose we have a sequence \( (\Omega_k,A_k) \) where each \( \Omega_k \) is a Benzecri domain and \( A_k \in \text{SL}(\Omega_k) \) moves a Hilbert distance at most \( d \). It suffices to show \( A_k \) has a convergent subsequence in \( \text{SL}_\pm(n+1) \).

By 4.7 we can pass to a subsequence so that \( \Omega_k \) converges to a Benzecri domain \( \Omega_\infty \). Choose a projective basis \( \mathcal{B} = (p_0,p_1,p_2,\ldots,p_{n+1}) \) in \( B(1/10) \). This ensures that \( \mathcal{B} \subset B_1(p;\Omega,d_\Omega) \) for every Benzecri domain \( \Omega \). We can choose a subsequence so that the projective bases \( \mathcal{B}_k = A_k(\mathcal{B}) \) converge to an \((n+2)\)-tuple \( \mathcal{B}_\infty = (q_0,\ldots,q_{n+1}) \subset \Omega_\infty \). We need to show this set is a projective basis.

Since every \( A_k \) moves \( p \) a distance at most \( d \), it follows that \( \mathcal{B}_\infty \subset B_{d+1}(p;\Omega_\infty,d_\Omega) \). Let \( \sigma_i \) be the \( n \)-simplex with vertices \( \mathcal{B} \setminus \{p_i\} \). Since metric balls are convex by Lemma 2.8, it follows that \( \sigma_i \subset B_{d+1}(p;\Omega_\infty,d_\Omega) \). Note that each \( A_i \) has determinant 1, so preserves Lebesgue measure.

Let \( V = (K_{\mu}(n,d+1))^{-1} \min_i \mu_L(\sigma_i) \). It follows from Corollary 4.8 that \( \mu_{\Omega_k}(\sigma_i) \geq V \). Let \( \sigma_i^m \) be the possibly degenerate \( n \)-simplex with vertices the \((n+2)\)-tuple \( \mathcal{B}_\infty \) with \( q_i \) deleted. Then \( \sigma_i^m = \lim_k A_k(\sigma_i) \). It is easy to see that \( \mu_{\Omega_\infty}(\sigma_i^m) = \lim_k \mu_{\Omega_k}(A_k\sigma_i) \geq V > 0 \). In particular \( \sigma_i^m \) is not degenerate therefore \( \mathcal{B}_\infty \) is a projective basis. There is a unique element \( A_\infty \in \text{SL}_\pm(n+1) \) sending \( \mathcal{B} \) to \( \mathcal{B}_\infty \). It is easy to check that \( A_\infty = \lim A_k \). \( \blacksquare \)

Theorem 6.2 and Theorem 4.6 have the following consequence:

Corollary 6.3 For every \( d > 0 \) there is a compact subset \( K \subset \text{SL}_\pm(n+1) \) so that if \( \Omega \) is any properly convex domain and \( p \) is a point in \( \Omega \) and \( S = S(\Omega,p,d) \) is the subset of \( \text{SL}_\pm(\Omega) \) consisting of all maps that move \( p \in \Omega \) a distance at most \( d \) in the Hilbert metric on \( \Omega \), then \( S \) is conjugate into \( K \), i.e. there is \( B \in \text{SL}_\pm(n+1) \) such that \( B \cdot S \cdot B^{-1} \subset K \).

We will use the following fact from (6.2.3) in Eberlein [42]:

Proposition 6.4 (Zassenhaus neighborhood) There is a neighborhood \( U \) of the identity in \( \text{SL}(n+1) \) such that if \( \Gamma \) is a discrete subgroup of \( \text{SL}(n+1) \) then the subgroup generated by \( \Gamma \cap U \) is nilpotent.
The following statement and proof is essentially (4.1.16) in Thurston [5]. However the hypotheses are different.

**Proposition 6.5** (short motion almost nilpotent) For every dimension \( n \geq 2 \) there is an integer \( m > 0 \) and a Margulis constant \( \mu > 0 \) with the following property:

Suppose that \( \Omega \) is a properly convex domain and \( p \) is a point in \( \Omega \) and \( \Gamma \subset \text{SL}_+ \Omega \) is a discrete subgroup generated by isometries that move \( p \) a distance less than \( \mu \) in the Hilbert metric on \( \Omega \). Then

1. There is a normal nilpotent subgroup of index at most \( m \) in \( \Gamma \).
2. \( \Gamma \) is contained in a closed subgroup of \( \text{SL}_+ \left( n+1 \right) \) with no more than \( m \) components and with a nilpotent identity component.

**Proof** By Theorem 4.6 we may assume \( \Omega \) is a Benzecri domain and \( p \) is the origin. Let \( K \subset \text{SL}_+ \left( n+1 \right) \) be a compact subset as provided by Theorem 6.2 when \( d = 1 \) (for example). Since \( K \) is compact, it is covered by some finite number, \( m \), of left translates of the Zassenhaus neighborhood \( U \) given by Proposition 6.4. Define \( \mu = d/m \).

Let \( W \subset \text{SL}_+ \Omega \) be the subset of all \( A \) such that \( A \) moves \( p \) a distance less than \( \mu \). Then \( W = W^{-1} \) and \( W^m \subset K \). By hypothesis the group \( \Gamma \) is generated by \( \Gamma \cap W \). Define \( \Gamma_U \) to be the nilpotent subgroup generated by \( \Gamma \cap U \). We claim there are at most \( m \) left cosets of \( \Gamma_U \) in \( \Gamma \).

Otherwise there are \( m + 1 \) distinct left cosets of \( \Gamma_U \) which have representatives each of which is the product of at most \( m \) elements of an arbitrary symmetric generating set of \( \Gamma \) (see [5], 4.1.15). Choose the symmetric generating set \( \Gamma \cap U \subset W \). Hence these representatives are in \( W^m \subset K \). But \( K \) is covered by \( m \) left cosets of \( U \). Thus there are two representatives \( g, g' \in \Gamma \cap W \) such that \( g, g' \) are in the same left translate of \( U \). Thus \( g^{-1}g' \in \Gamma \cap U \subset \Gamma_U \), hence \( g\Gamma_U = g'\Gamma_U \) which contradicts the existence of \( m + 1 \) distinct cosets of \( \Gamma_U \) in \( \Gamma \).

It follows that \( \Gamma_U \) has index at most \( m \) in \( \Gamma \).

It remains to prove there is a normal subgroup of index at most \( m \) and the statement concerning the closed subgroup. We follow the last three paragraphs of Thurston’s proof (4.1.16) [5] verbatim, subject only to the change that he uses \( \varepsilon \) in place of our \( \mu \). During the course of that proof, \( m \) is replaced by another constant. ■

The first part of the Margulis Lemma (Theorem 6.1) follows from this. The second part, starting with “Furthermore…” follows from below thick-thin decomposition.

### 6.2 The thick-thin decomposition

The Margulis lemma is the basis for the thick-thin decomposition, which will now be described. Each component of the thin part (where the injectivity radius is less than \( \mu/2 \)) consists of Margulis tubes (tubular neighborhoods of short geodesics) and cusps. The proof of below theorem requires an understanding of isometries beyond the material of §5, as well as a definition and analysis of cusps. The details are in [2].

**Theorem 6.6** (strictly convex thick-thin) Suppose that \( M \) is a strictly convex projective \( n \)-manifold. Then \( M = A \cup B \), where \( A \) and \( B \) are smooth submanifolds and \( \overline{A} \cap \overline{B} = \partial A = \partial B \), and \( B \) is nonempty, and \( A \) is a possibly empty submanifold with the following properties:

1. If \( \text{inj}(x) \leq t_n \), then \( x \in A \), where \( t_n = 3^{-(n+1)}\mu_n \).
2. If \( x \in A \), then \( \text{inj}(x) \leq \mu_n/2 \).
3. Each component of \( A \) is a Margulis tube or a cusp.

The manifold \( B \) is referred to as the thickish part and \( A \) as the thinnish part. The terminology reflects the fact that the injectivity radius on \( \partial A \) is between \( t_n \) and \( \mu_n/2 \). It follows from the description of the thinnish part that the thickish part is connected in dimension greater than \( 2 \). Strict convexity is necessary since given any properly convex manifold \( M \), there is a properly convex structure on \( M \times S^1 \), where the circle factor is arbitrarily short. In this case the whole manifold is thinnish.

28
A Manifolds, fundamental group and covering spaces

This section gives a brief introduction to the topics in the title, which is adapted to this course.

A.1 Manifold

The notion of a manifold is introduced informally in the lectures using gluing constructions. This section provides an alternative approach that is adapted to this course, and may help clarify both the concept and terminology.

**Definition A.1** (Manifold) *An n–dimensional manifold is a metric space with the property that every point has a neighbourhood that is homeomorphic to \( \mathbb{R}^n \).*

The above definition is not the standard definition of a manifold. Usually one does not encounter a “metric space” in the definition, but a “paracompact Hausdorff space” or a “second countable Hausdorff space.” All of these notions are equivalent if the space has at most a countable number of connected components, and there is no need to go into the relevant point-set topology. We will mainly be interested in the question of whether the quotient of a manifold under the action of a group of isometries is also a manifold.

**Remark A.2** Thurstons gives a more general definition of manifolds in [5] §3.1, and defines different structures on manifolds. Thurstons’s definition includes non-Hausdorff spaces that are locally homeomorphic to \( \mathbb{R}^n \). Such spaces often arise naturally as quotient spaces of certain group actions or flows on a (Hausdorff) manifold.

Using methods from algebraic topology, one can show that \( \mathbb{R}^n \) is homeomorphic to \( \mathbb{R}^m \) if and only if \( n = m \). It follows from this *invariance of domain* that the dimension of a manifold is indeed well-defined. To save space, “n–dimensional manifold” is abbreviated to “n–manifold.” It turns out that any n–manifold can be embedded in \( \mathbb{R}^m \) for large enough \( m \) (depending on the manifold). For example, an open interval and the circle are 1–manifolds. The former can be embedded in \( \mathbb{R}^1 \) and the latter in \( \mathbb{R}^2 \). A class of examples generalising the standard embedding of the circle in the plane is given in Proposition A.4.

**Example A.3** (1–manifolds) Up to homeomorphism, the line and the circle are the only connected 1–manifolds. The circle is therefore the unique compact, connected 1–manifold.

The following is a useful tool to obtain many examples of manifolds:

**Proposition A.4** (Regular level set is manifold) Let \( f: \mathbb{R}^{n+1} \to \mathbb{R} \) be a \( C^1 \) function and suppose \( a \in \mathbb{R} \) is a regular value. Then the level set \( f^{-1}(a) \) is an n–dimensional manifold.

**Proof** Since \( f^{-1}(a) \) is a (possibly not connected) subset of \( \mathbb{R}^{n+1} \), we can turn it into a metric space by putting the induced metric on it. It remains to show that every point has a neighbourhood that is homeomorphic to \( \mathbb{R}^n \). This follows from the Implicit Function Theorem.

**Examples A.5** (Two n–manifolds)

1. The \( n–\)sphere \( S^n \) is the level set \( x_0^2 + \ldots + x_n^2 = 1 \).
2. The hyperboloid \( H^n \) is the level set \( -x_0^2 + x_1^2 + \ldots + x_n^2 = -1 \).

**Example A.6** \( \text{SL}_n(\mathbb{R}) \) is an \((n^2 − 1)–\)dimensional manifold. This can be seen by viewing the set of all \( n \times n \) matrices as \( \mathbb{R}^{n^2} \). The determinant is a polynomial map, hence \( C^\infty \), with 1 as a regular value. So \( \text{det}^{-1}(1) = \text{SL}_n(\mathbb{R}) \) is a manifold of the claimed dimension. Moreover, matrix multiplication \( (M,N) \to MN \) is a continuous map with respect to the product metric, and the inverse \( M \to M^{-1} \) is also a continuous map. This is an example of a Lie group.
The following standard terminology (that may be confusing to start with) makes sense once one starts to consider manifolds with boundary. A compact manifold is called a closed manifold. An \(n\)-manifold-with-boundary is a metric space with the property that every point has a neighbourhood that is homeomorphic to either \(\mathbb{R}^n\) or the half-space

\[
\mathbb{R}^n_\geq = \{x \in \mathbb{R}^n \mid x_n \geq 0\}.
\]

The circle is a closed 1–manifold, whilst the interval \([0,1]\) is a compact 1–manifold-with-boundary, and \((0,1]\) is a non-compact 1–manifold-with-boundary. In terms of their topology, each of these spaces is both open and closed, since it is regarded as a metric space in its own right.

The set of points whose neighbourhood is homeomorphic with the half-space is termed the boundary of the manifold-with-boundary, and the symbol \(\partial\) is (again confusingly) used for the boundary. Whence \(\partial [0,1] = \{0,1\}\) and \(\partial (0,1] = \{1\}\), when both are regarded as manifolds-with-boundary.

If the boundary is empty, a manifold-with-boundary is simply a manifold as defined earlier. The hyphens will usually be dropped to allow statements such as “Let \(M\) be a manifold with (possibly empty) boundary…”

Last, a topologically finite manifold is either a closed manifold or the interior of a compact manifold-with-boundary.

### A.2 Homotopy equivalence

Let \(X\) and \(Y\) be topological spaces and \(I = [0,1]\) be the unit interval. Two maps \(f, g : X \to Y\) are homotopic if there is a continuous map \(F : X \times I \to Y\) such that

\[
F(x,0) = f(x) \text{ and } F(x,1) = g(x) \text{ for all } x \in X.
\]

One often writes \(f_t(x) = F(x,t)\), so a homotopy gives a continuous family of maps between \(f_0 = f\) and \(f_1 = g\).

**Exercise A.7** Homotopy is an equivalence relation on the set \(C(X,Y)\) of all continuous maps from \(X\) to \(Y\).

We therefore write \(f \sim g\) if \(f\) and \(g\) are homotopic. We will mostly be concerned with homotopies of a path that keep its end-points fixed, so it is useful to have the following terminology.

The paths \(f, g : I \to X\) (where \(I = [0,1]\) is the unit interval) are path homotopic if they have the same initial point \(f(0) = x_0 = g(0)\) and the same terminal point \(f(1) = x_1 = g(1)\) and there is a homotopy \(F : I \times I \to Y\), satisfying

\[
F(0,t) = x_0 \text{ and } F(1,t) = x_1 \text{ for all } t \in I.
\]

(The first \(I\) in the homotopy is the domain of \(f\) and \(g\) and the second is the \(I\)–factor from the homotopy)

### A.3 Fundamental group

A loop in the topological space \(X\) is a path \(f : I \to X\) with \(f(0) = f(1)\). We’ll fix a basepoint \(x_0 \in X\), and study the loops based at \(x_0\), i.e. the loops \(f : I \to X\) with

\[
f(0) = x_0 = f(1).
\]

Given two loops based at \(x_0\), \(f, g : I \to X\), we define the loop \(g \ast f : I \to X\) based at \(x_0\) by:

\[
(g \ast f)(x) = \begin{cases} 
  f(2x) & \text{if } x \in [0, \frac{1}{2}], \\
  g(2x - 1) & \text{if } x \in [\frac{1}{2}, 1].
\end{cases}
\]

This is well-defined since \(f(1) = x_0 = g(0)\). The loop \(g \ast f\) runs first along \(f\) and then along \(g\), but each loop is traversed twice as fast to keep the domain \(I\). This concatenation looks as though it should give a nice multiplication of loops. However, in general we have

\[
h \ast (g \ast f) \neq (h \ast g) \ast f,
\]

30
since the domain of $h$ on the left hand side is $[0, 1]$, whilst on the right hand side it is $[0, 1]$. 

This is fixed by considering loops up to an appropriate equivalence. We say that loops $f, g: I \rightarrow X$ based at $x_0$, are equivalent, if there is a path homotopy $F: I \times I \rightarrow X$ between the loops. As above, this gives an equivalence relation on the set of all loops based at $x_0$. The equivalence class of $f: I \rightarrow X$ will be denoted $[f]$ and called the homotopy class of $f$ (based at $x_0$).

We can now define the fundamental group of $X$ at $x_0$, denoted $\pi_1(X, x_0)$. Its elements are precisely the homotopy classes of loops based at $x_0$. The multiplication $\circ: \pi_1(X, x_0) \times \pi_1(X, x_0) \rightarrow \pi_1(X, x_0)$ is defined by:

$$[g] \circ [f] = [g \ast f].$$

One now needs to check that the multiplication

1. is well-defined, i.e. it does not depend on the representatives $f$ and $g$; and
2. satisfies the axioms of a group.

The first part is elementary, but tedious, and you can work out the details as an exercise or read them in [52]. Once this is done, the associative law is not too terrible, and then you can check that the identity is given as

The converse of the proposition is not true. For instance, the annulus and the Möbius band both have fundamental group isomorphic to $\mathbb{Z} \mathbb{Z}$. Why do they have fundamental group isomorphic to $\mathbb{Z} \mathbb{Z}$? For the purpose of this course, the most natural way to compute fundamental group is via the universal covering space. The annulus is obtained, for instance, as the orbit space of the action of the euclidean (or hyperbolic) plane by a translation, and the Möbius band is obtained as the orbit space of a glide reflection. Both of these groups are infinite cyclic, and the remainder of these notes explains why this implies that their respective fundamental groups are also infinite cyclic.

By the proposition, the fundamental group is an invariant of a topological space, and by the example, it is not a complete invariant.
A.4 Simply connected and contractible

**Definition A.9** (Simply connected) The path connected topological space $X$ is simply connected if $\pi_1(X, x_0)$ is trivial for some (and hence every) $x_0 \in X$.

An example of a simply connected space is $\mathbb{R}^n$ for any $n \geq 1$: By compactness, any loop is contained in some ball, and hence is homotopic to a constant path. The spheres $S^n$ for $n \geq 2$ are also simply connected, but this is more difficult to see. One can give a similar argument to the one given for $\mathbb{R}^n$ if there is a point on the sphere which is not in the image of a given loop. However, one also has to worry about space-filling loops.

Simply connected spaces have some nice properties; here is one of them:

**Lemma A.10** In a simply connected space, any two paths having the same initial and terminal points are path homotopic.

**Proof** Let $f, g : I \to X$ be paths with $f(0) = x_0 = g(0)$ and $f(1) = x_1 = g(1)$. Denote $\overline{g}(x) = g(1-x)$ the path $g$ traversed in the opposite direction. Then $f \star \overline{g} : I \to X$ is well-defined and a loop based at $x_0$. Since $X$ is simply connected, it the loop $f \star \overline{g}$ is path homotopic to the constant map $e(t) = x_0$. But then the path $(f \star \overline{g}) \star g$ is homotopic to $e \star g$. It is also clear that $(f \star \overline{g}) \star g$ is path homotopic to $f$, and $e \star g$ is path homotopic to $g$.

**Definition A.11** (Contractible) A topological space $X$ is contractible if the identity map $X \to X$ is homotopic to the constant map.

A contractible space is automatically path connected and simply connected. The spaces $\mathbb{R}^n$ are contractible. However, the spheres $S^n$ for $n \geq 2$ are not contractible, though this is a story for a different course.

A.5 Covering spaces

Let $Y$ and $X$ be path connected topological spaces and $p : Y \to X$ be a continuous map. A path connected open set $U \subseteq X$ is evenly covered if $p^{-1}(U)$ consists of a countable union of pairwise disjoint sets $\tilde{U}_k$ such that each restriction $p|_{\tilde{U}_k} : \tilde{U}_k \to U$ is a homeomorphism. We say that $p : Y \to X$ is a covering map (and $Y$ is a cover of $X$) if every point in $X$ has a path connected open neighbourhood that is evenly covered.

**Example A.12** The standard example is the map $\mathbb{R} \to S^1$ given by $x \to \exp(2\pi i x)$. We have already seen much more interesting examples!

Since $X$ is path connected, it is not difficult to see that for each evenly covered open set $U \subseteq X$, the cardinality of $p^{-1}(U)$ is identical, and its number is called the degree of the cover. Once the degree is infinite, there are different distinctions that will be made based on group theory.

We will now see that a covering "unwraps" parts of the space $X$, and that the way a loop $\alpha$ in $X$ is unwrapped corresponds to an action of its homotopy class $[\alpha]$ on the covering space.

A.6 The deck group

Let $p : Y \to X$ be a covering map. A homeomorphism $h : Y \to Y$ with the property that $p \circ h = p$ is called a deck transformation. The set of deck transformations form a group under composition.

**Example A.13** The group of deck transformations for the cover $\mathbb{R} \to S^1$ is isomorphic to $\mathbb{Z}$, with $h_n(x) = x + n$.

For any $x \in X$, the components of a fibre $p^{-1}(x) \subseteq Y$ are permuted by the deck transformation $h$. If the action of the deck group is transitive on some fibre, then it is transitive on every fibre and the covering is called regular.
A.7 Lifting properties

One often runs into the question of whether a path or a homotopy in the base space can be lifted to a covering space, and if so, what conclusions one can expect. Here are some useful results.

Definition A.14 Let \( p : Y \to X \) be a covering map, and suppose \( f : Z \to X \) is continuous. A lift of \( f \) to \( Y \) is a map \( \tilde{f} : Z \to Y \) such that \( p \circ \tilde{f} = f \).

Lemma A.15 Let \( p : Y \to X \) be a covering map, and \( p(y_0) = x_0 \). Let \( \alpha : [0,1] \to X \) be a path with \( \alpha(0) = x_0 \). Then \( \alpha \) has a unique lift \( \tilde{\alpha} : [0,1] \to Y \) satisfying \( \tilde{\alpha}(0) = y_0 \).

The lemma is proved by covering the image of \( \alpha \) by finitely many open sets that are evenly covered and then defining the lift step by step over pre-images of these open sets.

Lemma A.16 Let \( p : Y \to X \) be a covering map, and \( p(y_0) = x_0 \). Suppose the map \( F : I \times I \to X \) is continuous with \( F(0,0) = x_0 \). Then there is a unique lift \( \tilde{F} : I \times I \to Y \) such that \( \tilde{F}(0,0) = y_0 \). Moreover, if \( F \) is a path homotopy, then \( \tilde{F} \) is a path homotopy.

This second lemma is a 2-dimensional version of the first. In this case, one subdivides the square \( I \times I \) into finitely many smaller squares (or triangles) so that each piece in the subdivision is mapped by \( F \) into an evenly covered open set. One now lifts the map \( F \) by starting with the bottom left square (or triangle) and extending it step by step over all of \( I \times I \).

Combining the above two lemmata gives:

Corollary A.17 Let \( p : Y \to X \) be a covering map, and \( p(y_0) = x_0 \). Let \( \alpha \) and \( \beta \) be paths in \( X \) from \( x_0 \) to \( x_1 \), and denote their respective lifts to \( Y \) starting at \( y_0 \) by \( \tilde{\alpha} \) and \( \tilde{\beta} \). If \( \alpha \) and \( \beta \) are path homotopic, then \( \tilde{\alpha} \) and \( \tilde{\beta} \) end at the same point of \( Y \) and are also path homotopic.

A.8 The universal cover

Recall that a path connected topological space \( X \) is simply connected if \( \pi_1(X) \) is the trivial group.

Let \( p : \tilde{X} \to X \) be a covering map, and suppose that \( \pi_1(\tilde{X}) \) is trivial. Then \( \tilde{X} \) is called a universal cover of \( X \). The universal property of this cover is that any other covering map \( Y \to X \) factors as \( \tilde{X} \to Y \to X \). So, in particular, a universal cover is unique up to homeomorphism and therefore called the universal cover.

Not all topological spaces have universal covers, the fancy properties needed are that \( X \) is path connected (fine), locally path connected (every neighbourhood of every point contains a neighbourhood that is path connected) and semi-locally simply connected (every point has a neighbourhood with the property that every loop in that neighbourhood is contractible in \( X \)). These properties are clearly necessary for the existence of a universal cover, and they turn out to be sufficient also. In particular, every manifold has a universal cover.

Theorem A.18 If \( X \) is path connected, locally path connected and semi-locally simply connected, then \( X \) has a universal cover.

Proof Here is a sketch. The points of \( \tilde{X} \) are equivalence classes of paths in \( X \) starting at a fixed point \( x_0 \in X \).

Two paths \( \alpha \) and \( \beta \) are termed equivalent if and only if their concatenation \( \alpha \beta^{-1} \) is homotopic to the constant path. The map \( \tilde{X} \to X \) is then defined by taking an equivalence class of paths \( [\alpha] \) to its end-point \( \alpha(1) \). So far, we have only used the assumption that \( X \) is path connected.

The other two properties are now used to define a topology on \( \tilde{X} \). A basic open set is defined as follows. Take a path connected open set \( V \) in \( X \) with the property that every loop in \( V \) is contractible in \( X \), and take a point
\[\{\alpha\} \in \tilde{X} \text{ such that } \alpha(1) \in V. \]

Now define the basic open set \(\tilde{V} = \tilde{V}(V, [\alpha])\) as the subset of \(\tilde{X}\) represented by paths of the form \(\beta \ast \alpha\), where \(\beta\) is a path in \(V\) starting at \(\alpha(1)\). The topology on \(\tilde{X}\) is defined to be the topology generated by these basic open sets.

An important property of the basic open sets is that the map \(\tilde{X} \to X\) restricts to a bijection \(\tilde{V} = \tilde{V}(V, [\alpha]) \to V\). One now needs to check that this map is indeed well-defined, that it is a homomorphism, and that it is both injective and surjective. The trickiest part is to prove surjectivity: for each \(\gamma \in \pi_1(X, x_0)\) by \(\varphi(h) = [\alpha]\).

One now needs to check that this map is indeed well-defined, that it is a homomorphism, and that it is both injective and surjective. The trickiest part is to prove surjectivity: for each \([\alpha] \in \pi_1(X, x_0)\), one has to construct a homeomorphism \(h: \tilde{X} \to \tilde{X}\). To define \(h(\tilde{y})\) for any given \(\tilde{y} \in \tilde{X}\), choose a path \(\tilde{\beta}\) from \(\tilde{x}_0\) to \(\tilde{y}\), giving a path \(\beta = p \circ \tilde{\beta}\) from \(x_0\) to \(p(\tilde{y})\). Now \(\gamma = \alpha \ast \tilde{\beta}\) is another path in \(\tilde{X}\) starting at \(\tilde{x}_0\). Define the end-point of this path to be \(h(\tilde{y})\). One needs to check that this is well-defined and continuous. Since the construction applied to \([\alpha]^{-1}\) will produce the map \(h^{-1}\), it follows that \(h\) is invertible and has a continuous inverse; whence \(h\) is a homeomorphism.

\section{A.9 Covers for all subgroups}

The main result stated above was that if \(\tilde{X} \to X\) is the universal cover of \(X\), then the deck group is isomorphic to \(\pi_1(X)\). Now if \(p: Y \to X\) is a covering map, by the universal property, \(\tilde{X}\) is also the universal cover of \(Y\), and hence the deck group of \(\tilde{X} \to Y\) is isomorphic to \(\pi_1(Y)\). How are these groups related?

Choose a base point \(y_0 \in Y\), and denote \(x_0\) its image in \(X\). By continuity of the covering map, we get a homomorphism \(p_*: \pi_1(Y, y_0) \to \pi_1(X, x_0)\). One can show that this is in fact a monomorphism, so \(\pi_1(Y)\) can be identified with a subgroup of \(\pi_1(X)\). Moreover, the degree of the cover equals the index

\[\left[\pi_1(X) : p_*\pi_1(Y)\right].\]

Conversely, for every subgroup of \(\pi_1(X)\), there is a cover corresponding to that subgroup, and any two such covers are equivalent. The cover is the orbit space of the action of the subgroup on the universal cover.

One can also show that the cover \(p: Y \to X\) is regular if and only if \(p_*\pi_1(Y)\) is a normal subgroup of \(\pi_1(X)\), and in this case, the group of deck transformations is isomorphic to the quotient group

\[\pi_1(X) / p_*\pi_1(Y).\]

One often qualifies covers by properties of the deck group. So a cyclic cover has a cyclic deck group, an infinite cyclic cover has deck group isomorphic with \(\mathbb{Z}\), an abelian cover has an abelian deck group, and so forth.

\section{A.10 Branched covering maps}

The map \(p: Y \to X\) is a \textit{branched covering map} if there is a nowhere dense set \(S \subset X\) such that the restriction \(p|_{X \setminus S}: p^{-1}(X \setminus S) \to X \setminus S\) is a covering map. A minimal such set consists of all points which do not have an evenly covered neighbourhood (and this minimal set is necessarily closed).

We call \(Y\) a branched cover of \(X\), \(X \setminus S\) the set of regular points, \(S\) the set of singular points, and \(p^{-1}(S)\) the set of branch points. By definition, a branched covering map has a well-defined degree, namely the degree of the covering map \(p|_{X \setminus S}\). A branched cover is a cover if the set of singular points can be taken to be empty.
Bibliography

Main references

Survey articles

Bed-time reading

Geometry games and software

Sources


School of Mathematics and Statistics
The University of Sydney
NSW 2006 Australia

Email: tillmann@math.usyd.edu.au