Lecture 2
What is Geometry?

2.1 Groups of Transformations

Let X be some set. In the following we will refer to X as "space". By a *transformation* or *bijection* of X we will mean a mapping $f : X \to X$ that is "one-to-one and onto". This means that:

1) if $f(x_1) = f(x_2)$ then $x_1 = x_2$,
2) every y in X is of the form $f(x)$ for some x in X—which by 1) is clearly unique,

and then the inverse of f, denoted by f^{-1}, is defined using 2) to be the mapping of X to itself $f^{-1}(y) := x$.

▷ **2.1—Exercise 1.** Check that f^{-1} is also a bijection of X.

Recall that if f and g are any two self-mappings of X, then their *composition* $f \circ g$ is the mapping of X to X defined by $(f \circ g)(x) := f(g(x))$.

▷ **2.1—Exercise 2.** Show that the composition of two bijections of X is again a bijection of X. Show by an example that composition of bijections is **not** necessarily commutative, i.e., it is not true in general that $f \circ g = g \circ f$. (Hint: Take for X the three element set $\{1, 2, 3\}$.) On the other hand, show that $f \circ f^{-1}$ and $f^{-1} \circ f$ are always both equal to the *identity transformation* of X, i.e., the transformtion of X that maps each elemet of X to itself. Show that the inverse of $f \circ g$ is $g^{-1} \circ f^{-1}$.

2.1.1 Definition. A set G of transformations of X is called a *group* of transformations of X if:

1) it contains the identity transformation,
2) it is closed under composition, and
3) if a transformation f is in G then so also is f^{-1}.

Clearly the set $\mathrm{Biject}(X)$ of all transformations of X is the largest group of transformations of X, and the set consisting of only the identity map of X is the smallest.

▷ **2.1—Exercise 3.** Show that the intersection of any collection of groups of transformations of X is itself a group of transformations of X. Deduce that if S is any set of transformations of X, then there is is a smallest group of transformations of X that includes S. This group is called the group generated by S.

2.1.2 Felix Klein's Erlanger Program. In 1872, the German mathematician Felix Klein, recently appointed to a chair at Erlanger, proposed a new viewpoint towards Geometry. For centuries after Euclid, there was only Euclidean geometry, but then, in the 1800s many new and different geometries were introduced and studied; for example spherical geometry, hyperbolic geometry, projective geometry, affine geometry, and more. Eventually mathematicians felt the need to elucidate just what a geometric theory was and how to classify the various different geometries. Klein observed that each geometry had associated to it a group of transformations, the symmetry group of the geometry, and two objects of the geometry could be considered equivalent if there was an element of that group that carried one object to the other. For example, the symmetry group of Euclidean geometry is the so-called Euclidean Group—generated by the rotations and translations of Euclidean space—and two triangles are considered to be "equivalent" in Euclidean geometry if they are congruent, meaning that there is some element of the Euclidean group that carries one triangle into another. To quote Klein:

> "...geometrical properties are characterized by their invariance under a group of transformations."

That last sentence epitomizes what has come to be known as Klein's Erlanger Program, and just what it means should gradually become more clear as we proceed.

2.2 Euclidean Spaces and Their Symmetries

We shall as usual denote by $\mathbf{R}^n$ the space of all ordered n-tuples $x = (x_1, \ldots, x_n)$ of real numbers. The space $\mathbf{R}^n$ has two important structures. The first is algebraic, namely $\mathbf{R}^n$ is a real vector space (of dimension n). If α is a "scalar", i.e., a real number, then the product αx is the n-tuple $(\alpha x_1, \ldots, \alpha x_n)$, and if $y = (y_1, \ldots, y_n)$ is a second element of $\mathbf{R}^n$, then the vector sum $x + y$ is the n-tuple $(x_1 + y_1, \ldots, x_n + y_n)$. The second important structure is the so-called *inner product*, which is a real-valued function on $\mathbf{R}^n \times \mathbf{R}^n$, namely $(x, y) \mapsto \langle x, y \rangle := x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$. As we shall see, it is this inner product that allows us to define geometric concepts like length and angle in $\mathbf{R}^n$. Sometimes the inner product is referred to as the "dot-product" and written as $x \cdot y$.

The inner product has three characteristic properties that give it its importance, namely:

1) Symmetry: $\langle x, y \rangle = \langle y, x \rangle$ for all $x, y \in \mathbf{R}^n$
2) Positive Definiteness: $\langle x, x \rangle \geq 0$, with equality if and only if $x = 0$.
3) Bilinearity: $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$, for all $x, y, z \in \mathbf{R}^n$ and all $\alpha, \beta \in \mathbf{R}$.

▹ **2.2—Exercise 1.** Verify that $\langle x, y \rangle$ has these three properties.

More generally:

2.2.1 Definition. . If V is a real vector space, then a real-valued function on $V \times V$, $(v_1, v_2) \mapsto \langle v_1, v_2 \rangle$ is called an *inner product for V* if it is symmetric, positive definite, and bilinear. An *inner-product space* is a real vector space V together with a fixed choice of inner product for V.

We recall that that for any $x \in \mathbf{R}^n$, we define its norm, $\|x\|$, by $\|x\| := \sqrt{\langle x, x \rangle}$, and of course, we make this same definition in any inner product space.

▷ **2.2—Exercise 2.** Show that if $x, y \in \mathbf{R}^n$ and $t \in \mathbf{R}$, then $\|tx + y\|^2$ is a quadratic polynomial function of t, namely:

$$\|tx + y\|^2 = \langle tx + y, tx + y \rangle = \|x\|^2 t^2 + 2 \langle x, y \rangle t + \|y\|^2 .$$

and note the important special case

$$\|x + y\|^2 = \|x\|^2 + 2 \langle x, y \rangle + \|y\|^2 .$$

Finally, for reasons we shall see a little later, the two vectors x an y are called *orthogonal* if $\langle x, y \rangle = 0$, so in this case we have:

Pythagorean Identity. *If x and y are orthogonal vectors in an inner product space then* $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

Now let me remind you of some basic facts from high-school mathematics concerning an arbitrary real polynomial $P(t) = at^2 + bt + c$ in a variable t. (For simplicity, we will assume tha $a > 0$.) Recall that the *discriminant* of $P(t)$ is the quantity $b^2 - 4ac$, and it distinguishes what kind of roots the polynomial has. In fact, the so-called "Quadratic Formula" says that the two (possible complex) roots of $P(t)$ are $(-b \pm \sqrt{b^2 - 4ac})/2a$. Thus there are three cases:

Case 1: $b^2 - 4ac > 0$. Then $P(t)$ has two real roots. Between these roots $P(t)$ is negative and outside of the interval between the roots it is positive.

Case 2: $b^2 - 4ac = 0$. Then $P(t)$ has only the single real root $-b/2a$, and elsewhere $P(t) > 0$.

Case 3: $b^2 - 4ac < 0$. Then $P(t)$ has no real roots and $P(t)$ is positive for all real t.

In the case of the polynomial $\|tx + y\|^2$, we see that $a = \|x\|^2$, $c = \|y\|^2$, and $b = 2 \langle x, y \rangle$, so the discriminant is $4(|\langle x, y \rangle|^2 - \|x\|^2 \|y\|^2)$. Now Case 1 is ruled out by positive definiteness. In Case 2, we have $|\langle x, y \rangle| = \|x\| \|y\|$, so if t is the root of the polynomial then $\|x + ty\| = 0$, so x = -ty, and we see that in this case x and y are linearly dependent. Finally, in Case 3, $|\langle x, y \rangle| < \|x\| \|y\|$, and since $x + ty$ is never zero, x and y are linearly independent. This proves one of the most important inequalities in all of mathematics,

Schwartz Inequality. *For all $x, y \in \mathbf{R}^n$, $|\langle x, y \rangle| \leq \|x\| \|y\|$, with equality if and only if x and y are linearly dependent.*

Of course, since our proof only used the three properties that define an inner product, the Schwartz Inequality is valid in any inner-product space.

▷ **2.2—Exercise 3.** Use the Schwartz Inequality to deduce the Triangle Inequality:

$$\|x + y\| \leq \|x\| + \|y\| .$$

(Hint: Square both sides.)

2.2—Example 1. Let $C([a, b])$ denote the vector space of continuous real-valued functions on the interval $[a, b]$ (with pointwise vector operations, as usual). For $f, g \in C([a, b])$ define $\langle f, g \rangle = \int_a^b f(x)g(x)\, dx$. it is easy to check that this satisfies our three conditions for an inner product. What does the Schwartz Inequality say in this case?