Chapter 1

Lecture 2: Isometries of $\mathbb{R}^2$

1.1 The synthetic vs. analytic approach; axiomatics

We discussed briefly what an axiomatic system is: a collection of *undefined terms* related by *axioms*. Further terms are defined in terms of the undefined terms and previously defined terms, and theorems as derived as logical consequences of the axioms. We discussed which logical rules are acceptable for proofs. If our axiomatic system is:

Ax. (i): All blorks are glebes.

Ax. (ii): All glebes are frimes.

then the undefined terms are blorks, glebes, frimes. Rosenberg's Thm.: *All blorks are frimes* follows from the accepted logical rule (modus ponens): $A \Rightarrow B, B \Rightarrow C$ implies $A \Rightarrow C$. Another accepted rule (modus tollens) is $A \Rightarrow B$ iff $\sim B \Rightarrow \sim A$ (proof by contradiction).

Often difficult questions about axiomatic systems:

1. Consistency: Show that your axiomatic system has no contradictory consequences. For example, the system

   Ax. (i) All blorks are glebes.

   Ax. (ii) All blorks are not glebes.

   Ax. (iii) There exists a blork.

   has as a logical consequence that there is a blork which is both a glebe and not a glebe. Note that the system would be consistent if we deleted Ax. (iii)

2. Completeness: Given a theory (e.g. number theory of $\mathbb{Z}$) show that a given set of axioms is strong enough to derive all true statements in the theory.

   As a dumb example, you cannot prove all the theorems of group theory if you leave out the axiom of associativity. Our handout "All triangles are isocseles" shows that Euclid's 5 axioms are not complete, as there is no notion of betweeness/sidedness (at least in the Playfair version of Ax. (v)).

3. Miminality/indenpendence: Show that your set of axioms contains no redundancies.

   For example, if we add Rosenberg's Thm. to the first set of axioms, the axioms are not minimal.

   A question from perhaps Euclid's time until the work of Gauss and Lobachevsky: Is the parallel postulate independent of the previous axioms?

4. Uniqueness of models/examples: For a given set of axioms, does there exist more than one example?

   For e.g. the axioms of a group, we certainly don't have or want uniqueness of examples. For Euclidean geometry, we certainly do want uniqueness. It's reasonable to say that this question never occurred to Euclid – he had the only model he cared about in front of him. This single minded focus obscured the relatively easy answer of the independence of the parallel postulate for 2000 years!

We looked at Hilbert's axioms for Euclidean geometry, and noted that the difficult questions above are indeed difficult for this axiomatic system.

To move ahead, we'll assume both that Hilbert's axioms are consistent and that we "understand" $\mathbb{R}$ well enough.

The connection between the synthetic and analytic approach is the infamous number line. The connection is given by some sort of statement like:

"A line ℓ is a set of points. There is a bijection $f : \ell \to \mathbb{R}$ such that f is between-ness preserving: if B is between A and C for one set of points on ℓ, then either $f(A) < f(B) < f(C)$ or $f(A) > f(B) > f(C)$, and this same relation then holds for all triples of points."

Is this statement sufficient to translate every result of synthetic geometry a la Hilbert into every result in analytic geometry?

1.2 Isometries

We have to decide if a symmetry of $\mathbb{R}^2$ means a function that preserves lines or preserves distances. Here the distance between points is given by the usual distance formula:

$$d((x_1, y_1), (x_2, y_2)) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

We choose distance preserving maps and call such symmetries *isometries*:

Definition. An isometry of R^2 is a function $f : \mathbb{R}^2 \to \mathbb{R}^2$ such that $d(P, Q) = d(f(P), f(Q))$ for all $P, Q \in \mathbb{R}^2$.

Example. Is $f(x, y) = (x^2, y^2 + 3x^2y + e^{xy})$ an isometry?

Note that we can now define congruence (which is an undefined term for Hilbert) by: sets $S_1, S_2 \subset \mathbb{R}^2$ are congruent if there exists an isometry f with $f(S_1) = S_2$. (This assume isometries are bijections; see below). This is certainly much cleaner than "two figures are congruent if we can pick one up and drop it on the other."

Certainly translations by fixed vectors, rotations around a fixed point, reflections across a fixed line, and compositions of such are isometries. (Proof?) Is that it?

Theorem. *The set of isometries of* $\mathbb{R}^2$ *is a group.*

Closure is easy, the identity function is the identity element, associativity is automatic (!) since isometries are functions on a set. The existence of inverses: we only need that isometries are bijections. Injectivity is easy: $f(P) = f(Q) \Rightarrow d(f(P), f(Q)) = 0 \Rightarrow d(P, Q) = 0 \Rightarrow P = Q$. Surjectivity seems hard. So there is a gap is our proof.

Warmup on Isom($\mathbb{R}$): Here $d(x, y) = |x - y|$. Say f is an isometry of $\mathbb{R}$. Set $f(0) = c$ and let T_{-c} be translation by $-c$. Then $g = T_{-c} \circ f$ has $g(0) = 0$. Therefore

$$|g(x)| = |g(x) - g(0)| = |x - 0| = |x|,$$

so $g(x)/x = \pm 1$ for all $x \neq 0$. You check that if the sign is different for some x_0, x_1, then $|g(x_0) - g(x_1)| \neq |x_0 - x_1|$, a contradiction. Thus $g(x) = \pm x$ for all x, and composing on the left with T_c given $f(x) = \pm x + c$.

Let $\mathscr{F}$ denote either of the functions $h(x) = \pm x$. Then we've shown:

Proposition. *Every isometry of* $\mathbb{R}$ *has a unique decomposition of the form* $f = T_c \circ \mathscr{F}$, *where* $c = f(0)$ *and* $\mathscr{F}$ *is one of* $h(x) = \pm x$.

The uniqueness of the decomposition is shown below for isometries of $\mathbb{R}^2$.

1.3 Back to groups

Example. Which of the following are groups? $(\mathbb{Z}, +), (\mathbb{Z}, \cdot)$, $(\mathbb{Z}_n, +)$, $(\mathbb{Z}_n, \cdot)$, $(\mathbb{R}, +)$, $(\mathbb{R}, \cdot)$, $(\mathbb{R}^*, \cdot)$, $(\mathbb{Z}_n^*, \cdot)$, $(\mathbb{Q}^*, \cdot)$, $(\mathbb{C}^*, \cdot)$, the symmetries of a triangle and of a square.

Let G and H be groups. The *product group* $G \times H$ is defined by

$$(g_1, h_1) \cdot_{G\times H} (g_2, h_2) = (g_1 \cdot_G g_2, h_1 \cdot_H h_2).$$

It's easy to check that this is indeed a group (e.g. the identity is (e_G, e_H)). Check that $G \times H$ is abelian iff G and H are abelian.

There is certainly a bijection between Isom($\mathbb{R}$) and $\mathbb{Z}_2 \times \mathbb{R}$, where we send $f(x) = x+c$ to $(0, c)$ and $f(x) = -x+c$ to $(1, c)$. However, this bijection doesn't "take multiplication table to multiplication table:" for $f_1(x) = x+3, f_2(x) = -x+2$, then

$$(f_2 \circ f_1)(x) = -x-1 \mapsto (1,-1);\ f_2 \mapsto (1,2),\ f_1 \mapsto (0,3),\ (1,2)+(0,3) = (1,5) \neq (1,-1).$$

(Note that there is no problem for $(f_1 \circ f_2)(x) = -x+5$.) Intuitively, we can't have a multiplication table preserving bijection, as Isom($\mathbb{R}$) is nonabelian and $\mathbb{Z}_2 \times \mathbb{R}$ is abelian.

Say $f : G \to H$ is a multiplication table preserving bijection. Then if $f(g_1) = h_1, f(g_2) = h_2$, we must have

$$f(g_1 \cdot_G g_2) = h_1 \cdot_H h_2 = f(g_1) \cdot_H f(g_2).$$

Definition. A bijection $f : G \to H$ is an isomorphism if $f(g_1 \cdot_G g_2) = f(g_1) \cdot_H f(g_2)$ for all $g_1, g_2 \in G$. G and H are called isomorphic, denoted $G \simeq H$.

Definition. A subset $K \subset G$ of a group $(G, \cdot_G)$ is a subgroup if $(K, \cdot_G)$ is a group.

Note that we have to check that K is closed under $\cdot_G$ and that inverses of elements in K are again in K. It is not necessary to check that the identity element of G is also in K (Why?), but it's a good idea. Associativity is automatic, since the larger set G is associative.

Question. What are all the subgroups of $(\mathbb{Z}, +)$? This uses some interesting number theory.

1.4 Isometries of $\mathbb{R}^2$

Denote the isometries of $\mathbb{R}^2$ by Isom($\mathbb{R}^2$) or just $\mathcal{I}$.

Facts about $\mathcal{I}$.

1. $f \in \mathcal{I}$ implies f is injective.

   The proof is as for Isom($\mathbb{R}$).

2. $f \in \mathcal{I}$ implies f is continuous.

   An intuitive understanding of continuity is good enough for us. A function $g : \mathbb{R} \to \mathbb{R}$ is continuous if its graph has no jumps or "lack of limit' like $\sin(1/x)$ at $x = 0$. This translates into: f is continuous if $x \to x_0$ implies $f(x) \to f(x_0)$, where the arrow denotes "gets closer and closer to." We could take a sequence of points x_i getting closer and closer to x_0 and write $x_i \to x_0$ implies $f(x_i) \to f(x_0)$. The same definition works for $f : \mathbb{R}^2 \to R^2$. Then $f \in \mathcal{I}$ gives

   $$P_i \to P \Rightarrow d(P_i, P) \to 0 \Rightarrow d(f(P_i), f(P)) \to 0 \Rightarrow f(P_i) \to f(P).$$

3. Isometries take lines to lines.

   We use the *Triangle Inequality*: for $P, Q, R \in \mathbb{R}^2$,

   $$d(P,Q) + d(Q,R) \leq d(P,R)$$

   with equality iff P, Q, R are collinear with Q between P and R.

   The proof is a (analytic) homework problem, which should be done carefully: what is the analytic definition of betweenness? Why is it ok to assume that one vertex of the triangle is at the origin and one side is on the positive x-axis?

   Anyway, if P, Q, R are on a line, then they satisfy triangle equality. Then $f(P), f(Q), f(R)$ satisfy triangle equality (why?), so they are also collinear. Now finish the argument.

4. Isometries preserve the dot product of vectors.

   Here $\vec{P} = (p^1, p^2), \vec{Q} = (q^1, q^2)$ has $\vec{P} \cdot \vec{Q} = p^1q^1 + p^2q^2$. Note that $|\vec{P}| \overset{\text{def}}{=} (\vec{P} \cdot \vec{P})^{1/2}$ is the length of $\vec{P}$. If $\vec{P}$ has its tail at the origin, $|\vec{P}| = d((0,0), P)$. (Note that the last P is a point, the head of $\vec{P}$.) Properly distinguishing between vectors with tails at the origin and their heads, we get

   $$|P-Q|^2 = d((0,0), P-Q)^2 = d(P,Q)^2 = d(f(P), f(Q))^2 = |f(P)-f(Q)|^2,$$

   and in particular $|P|^2 = |f(P)|^2$. Expanding the last equation using $|P - Q|^2 = (P-Q) \cdot (P-Q)$ etc. and canceling, we get $P \cdot Q = f(P) \cdot f(Q)$.

   Note the lack of a short but careful discussion of the difference between points and vectors.

5. Isometries preserve angles

   Let θ be the angle between $\vec{v}, \vec{w}$ (say with tails at the origin), and let ψ be the angle between $f(\vec{v}), f(\vec{w})$. Recall the Law of Cosines (which has an analytic proof based on the Pythagorean Theorem, which has an analytic proof):

   $$\begin{aligned} |v|^2 + |w|^2 - 2|v||w|\cos(\theta) &= |v-w|^2 \\ v \cdot v + w \cdot w - 2|v||w|\cos(\theta) &= (v-w) \cdot (v-w) = v \cdot v - 2v \cdot w - w \cdot w. \end{aligned}$$

   This gives by 4.

   $$\cos(\theta) = \frac{v \cdot w}{|v||w|} = \frac{f(v) \cdot f(w)}{|f(v)||f(w)|} = \cos(\psi)$$

   which is enough to conclude $\theta = \psi$.

   As a bonus, we get

   Cauchy-Schwarz Inequality: $|v \cdot w| \leq |v||w|$ with equality iff v and w are collinear.

   Also, $v \perp w$ iff $v \cdot w = 0$.

Now we get to work on determining the structure of an isometry f of $\mathbb{R}^2$.

Just as with Isom($\mathbb{R}$), say $f(\vec{0}) = \vec{v}$ and let T_{-v} be translation by the vector $-v$. (I hate typing arrows.) Then $T_{-v} \circ f = g$ is an isometry with $g(\vec{0}) = \vec{0}$. Let $\vec{i} = e_1 = (1,0)$ and $\vec{j} = e_2 = (0,1)$ be the standard axis vectors in their various notations. Dropping parentheses, we know $|g\vec{i}| = |\vec{i}| = 1, |g\vec{j}| = 1$, and $g\vec{i} \cdot g\vec{j} = \vec{i} \cdot \vec{j} = 0$. Therefore $g\vec{i}, g\vec{j}$ lie on the unit circle S^1 and are perpendicular. Looking at rotations, we see that $g\vec{i}$ can be any element of S^1. Note that once we know where $g\vec{i}$ is on S^1, there are only two choices for $g\vec{j}$.

Take an arbitrary point/vector (e, f) with tail at the origin and draw the standard coordinate box with vertices at the origin, $(e,0)$, (e,f), $(0,f)$. Say $e, f > 1$ for simplicity. By the triangle equality, $d(0,i)+d(i,(e,0)) = d(0,(e,0)) = e$, so $d(0,gi) + d(gi, g(e,0)) = d(0, g(e,0)) = e$. This implies that $g(e,0)$ is e units from the origin along the line joining 0 and $g(e,0)$. In particular, we have $g(e,0) = eg(1,0)$. The same holds for $g(0,f)$, namely $g(0,f) = fg(0,1)$. Now the coordinate box must be congruent to the new coordinate box with vertices the origin, $g(e,0), g(e,f), g(0,f)$ (since angles and distances are preserved), so the diagonal of the first coordinate box must go to the diagonal of the new coordinate box. (Make this precise.) This says

$$g(e,f) = g(e,0) + g(0,f) = eg(1,0) + fg(0,1).$$

If we write $g(1,0) = (a,c), g(0,1) = (b,d)$, this becomes

$$g(e,f) = e(a,c) + f(b,d) = (ae+bf, ce+bd), \text{ or } g\begin{pmatrix} e \\ f \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} e \\ f \end{pmatrix}$$

Note that the columns of the matrix are $g(1,0) = gi, g(0,1) = gj$.

It is interesting that the strange rules of matrix multiplication appear naturally in this geometric context.

Slight Digression: You can check directly that

(i) $g(v+w) = gv + gw$

(ii) $\lambda gv = g(\lambda v)$ for all $\lambda \in \mathbb{R}$.

Definition. A function $g : \mathbb{R}^2 \to \mathbb{R}^2$ satisfying (i), (ii) above is a linear transformation.

So a 2×2 matrix gives a linear transformation. Conversely, if a linear transformation h has $h(1,0) = (a,c), h(0,1) = (b,d)$, then

$$\begin{aligned} \begin{pmatrix} e \\ f \end{pmatrix} &= g\left[e\begin{pmatrix} 1 \\ 0 \end{pmatrix} + f\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right] \overset{(i)}{=} g\left(e\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) + g\left(f\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) \\ &\overset{(ii)}{=} eg\begin{pmatrix} 1 \\ 0 \end{pmatrix} + fg\begin{pmatrix} 0 \\ 1 \end{pmatrix} = e\begin{pmatrix} a \\ c \end{pmatrix} + f\begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} ae+bf \\ ce+bd \end{pmatrix} \\ &= \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} e \\ f \end{pmatrix} \end{aligned}$$

This shows:

Theorem. *There is a bijection between the set of linear transformations of* $\mathbb{R}^2$ *and the set of* 2×2 *matrices (with real numbers as entries).*

Back to isometries: So far, for $f \in \mathcal{I}$, we know that $T_{-v} \circ f = g$ fixes the origin and is a linear transformation.

Let $F_x : \mathbb{R}^2 \to \mathbb{R}^2$ be the flip on the x-axis; $F_x \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ -b \end{pmatrix}$. You check that F_x is a linear transformation, and its matrix is given by $F_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Let R_θ be counterclockwise rotation around the origin by angle θ.

Claim. There exists a θ such that either $g = R_\theta$ or $g = R_\theta \circ F_x$.

Proof. Define an isometry h as follows: if the counterclockwise angle between gi and gj is $\pi/2$, set $h = g$; if the counterclockwise angle is $3\pi/2$, set $h = g \circ F_x$. Check that in both cases the counterclockwise angle between hi and hj is $\pi/2$.

Since hi in on S^1, we have $hi = (\cos\theta, \sin\theta)$ for some θ. This implies $hj = (-\sin\theta, \cos\theta)$. Since h is linear transformation, it is given by the matrix

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Let $v = (e, f)$ make angle ψ with the origin. Then the coordinates of v satisfy $(e, f) = (|v|\cos\psi, |v|\sin\psi)$. (Why?) Therefore

$$\begin{aligned} hv &= h\begin{pmatrix} e \\ f \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} e \\ f \end{pmatrix} \\ &= \begin{pmatrix} e\cos\theta - f\sin\theta \\ e\sin\theta + f\cos\theta \end{pmatrix} = \begin{pmatrix} (|v|\cos\psi\cos\theta - |v|\sin\psi\cos\theta \\ |v|\cos\psi\sin\theta + |v|\sin\psi\cos\theta \end{pmatrix} \\ &= \begin{pmatrix} |v|\cos(\psi+\theta) \\ |v|\sin(\psi+\theta) \end{pmatrix} \\ &= R_\theta \begin{pmatrix} e \\ f \end{pmatrix}. \end{aligned}$$

Thus $h = R_\theta$. □

We now know $g \circ \mathscr{F} = h = R_\theta$, where $\mathscr{F}$ is either the identity or F_x. $g = T_{-v} \circ f$, so we get

$$T_{-v} \circ f \circ \mathscr{F} = R_\theta$$

for some θ. If we compose this equation on the right by $\mathscr{F}$, we get

$$T_{-v} \circ f = R_\theta \circ \mathscr{F},$$

since $\mathscr{F} \circ \mathscr{F} = Id$. If we now compose by T_v on the left, we get

$$f = T_v \circ R_\theta \circ \mathscr{F}.$$

This is the famous TRF decomposition of an isometry of $\mathbb{R}^2$:

Theorem. *Every isometry $f : \mathbb{R}^2 \to \mathbb{R}^2$ can be uniquely written in the form $f = T_v \circ R_\theta \circ \mathscr{F}$, where $v = f(\vec{0})$, θ is the counterclockwise angle from a horizontal axis through $f(\vec{0})$ to $f\vec{i}$, and $\mathscr{F}$ is the identity or the reflection F_x depending on whether the counterclockwise angle from $f\vec{i}$ to $f\vec{j}$ is $\pi/2$ or $3\pi/2$.*

Proof. We have done everything except the uniqueness. If $f = T_1R_1F_1 = T_2R_2F_2$ has two decompositions, then

$$\begin{aligned} T_1R_1F_1F_1 &= T_2R_2F_2F_1 \\ T_2^{-1}T_1R_1 &= R_2F_2F_1 \\ R_2^{-1}T_2^{-1}T_1R_1 &= F_2F_1 \end{aligned}$$

Since the left hand side is orientation-preserving (i.e. the counterclockwise angle from $f\vec{i}$ to $f\vec{j}$ is $\pi/2$), F_2F_1 is orientation-preserving, which means that either both are flips or both are Id. In either case, $F_2 = F_1$. So far, $T_1R_1F_1 = T_2R_2F_2 \Rightarrow T_1R_1 = T_2R_2$, so

$$T_2^{-1}T_1 = R_2R_1^{-1}$$

Since $R_2R_1^{-1}$ fixes $\vec{0}$, we get $T_2^{-1}T_1 = \text{Id} \Rightarrow T_1 = T_2$. So $R_2R_1^{-1} = \text{Id}$, which implies $R_1 = R_2$. □

Remark. (i) We can now finish the proof that $\mathcal{I}$ is a group. Since every isometry f has a TRF decomposition, and since translations, rotations and flips are surjective, f must be surjective.

(ii) $\mathcal{I}$ has many interesting subgroups. The group of rotations around the origin is in bijection with points of S^1, and is called SO(2), the special orthogonal group of the plane. The set $\{R_\theta \circ \mathscr{F} : \mathscr{F} = Id \text{ or } F_x\}$ is also a group, called the orthogonal group O(2). (To check closure: the only nontrivial case is $R_\theta \circ F_x \circ R_\psi \circ \mathscr{F} = R_\theta \circ R_{-\psi} \circ F_x \circ \mathscr{F} = R_{\theta-\psi} \circ \mathscr{F}$ for some possibly new value of $\mathscr{F}$. Check this!) The set of all translations is also a subgroup, and is in bijection with $\mathbb{R}^2$. For every natural number n, there is a subgroup of order n, namely the set of rotations by angles $2\pi k/n, k = 1, ..., n$.