Introduction to Algebraic topology, fall 2013

Exercises 2 - Solutions

1. Suppose V is a finite dimensional vector space, $A \subset V$ and $m \in \mathbb{N}$. Prove that the affine dimension of A is m if and only if the following conditions are true.

   (a) For any affinely independent subset $\{\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_k\}$ of A we have that $k \leq m$.

   (b) There exists an affinely independent subset $\{\mathbf{w}_0, \mathbf{w}_1, \ldots, \mathbf{w}_m\} \subset A$ with precisely $m+1$ vectors in it.

   Solution: Since V is finite dimensional, by Linear Algebra we know that V cannot contain arbitrary large linearly independent sequences. In fact Proposition 1.5 ii) states that if $\dim V = n$ and $(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k)$ is a linearly independent sequence of vectors of V, then $k \leq n$.

   Let $\{\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be affinely independent set of vectors belonging to A. Then, by Lemma 2.10, the sequence $(\mathbf{v}_1 - \mathbf{v}_0, \ldots, \mathbf{v}_k - \mathbf{v}_0)$ is linearly independent, so, by previous paragraph, $k \leq \dim V$. In particular A cannot contain arbitrary long affinely independent sequences, so there must exist a **maximal** affinely independent subset of A consisting of vectors $\mathbf{w}_0, \mathbf{w}_1, \ldots, \mathbf{w}_m$. "Maximal" here means that it has a maximal number of vectors an affinely independent subset of A can have. In other words it is exactly a set that satisfies conditions 1) and 2) of the exercise.

   Let $W = \operatorname{aff}\{\mathbf{w}_0, \mathbf{w}_1, \ldots, \mathbf{w}_m\}$. It is enough to show that
   a) affine dimension of W is exactly m,
   b) W is the affine hull of A.

   To prove a) let $\mathbf{v}_i = \mathbf{w}_i - \mathbf{v}_0$, for all $i = 1, \ldots, m$. Clearly (or "easy to see", if you prefer that notorious expression)

   $$W = \mathbf{v}_0 + U,$$

   where U is the vector space spanned by vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$. By Lemma 2.10 the sequence $(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ is linearly independent, hence it is a basis for U. Thus U is m-dimensional and since $W = \mathbf{v}_0 + U$, the affine dimension of W is exactly m. Claim a) is proved.

Next we show that W is precisely the affine hull of A. Let $W' = \text{aff}(A)$. Since

$$W = \text{aff}\{\mathbf{w}_0, \mathbf{w}_1, \ldots, \mathbf{w}_m\}$$

and $\{\mathbf{w}_0, \mathbf{w}_1, \ldots, \mathbf{w}_m\} \subset A \subset W'$, we have that $W \subset W'$ (the bigger set, the bigger its affine hull). Indeed, this inclusion shows that W' is *some* affine set that includes the points $\mathbf{w}_0, \mathbf{w}_1, \ldots, \mathbf{w}_m$ and by definition W is *the smallest* such a set.

To prove the converse inclusion $W' \subset W$ it is enough to show that $A \subset W$, by the same reasoning as in the previous paragraph. Let $a \in A$ be arbitrary. If $a \in \{\mathbf{w}_0, \mathbf{w}_1, \ldots, \mathbf{w}_m\}$, then $a \in W$ by definition. Otherwise the set $\{\mathbf{w}_0, \ldots, \mathbf{w}_m, a\}$ has more elements then the set $\{\mathbf{w}_0, \mathbf{w}_1, \ldots, \mathbf{w}_m\}$, so, by maximality of the former set, this bigger set cannot be affinely independent. By Lemma 2.10, (2), there exists scalars $t_0, \ldots, t_m, t_{m+1}$ such that

$$t_0\mathbf{w}_0 + \ldots + t_m\mathbf{w}_n + t_{m+1}a = 0$$

and $t_0 + \ldots + t_m + t_{m+1} = 0$, but not all scalars $t_0, \ldots, t_m, t_{m+1}$ equal to zero. Here we must have $t_{m+1} \neq 0$, because otherwise we would have a representation

$$t_0\mathbf{w}_0 + \ldots + t_m\mathbf{w}_m = 0$$

and $t_0 + \ldots + t_m = 0$, and at least one of the scalars $t_0, \ldots, t_m$, is not a zero. This would contradict (by the same Lemma 2.10) the fact that the sequence $\mathbf{w}_0, \mathbf{w}_1, \ldots, \mathbf{w}_m$ is affinely independent. Thus $t_{m+1} = 0$, so, solving for a, we have the equation of the form

$$a = \frac{-t_0}{t_{m+1}}\mathbf{w}_0 + \frac{-t_1}{t_{m+1}}\mathbf{w}_1 + \ldots + \frac{-t_m}{t_{m+1}}\mathbf{w}_m.$$

Here the sum of scalars is

$$\sum_{i=0}^{m} \frac{-t_i}{t_{m+1}} = \frac{1}{t_{m+1}}(-t_0 - t_1 - \ldots - t_m) = \frac{1}{t_{m+1}}t_{m+1} = 1,$$

because $t_0 + t_1 + \ldots + t_m + t_{m+1} = 0$. By Lemma 2.7

$$a \in \text{aff}\{\mathbf{w}_0, \mathbf{w}_1, \ldots, \mathbf{w}_m\} = W,$$

which is what we had to prove.

2. a) Suppose V is a vector space and $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are three different elements of V that lie on the same line. Prove that one of these points lies on the closed interval between the other two (i.e. on the 1-simplex the other two points span).
b) Suppose $C, D \subset V$ are convex sets and $f\colon C \to D$ is an affine mapping. Prove that

$$f((1-t)\mathbf{x} + t\mathbf{y}) = (1-t)f(\mathbf{x}) + tf(\mathbf{y})$$

whenever $\mathbf{x}, \mathbf{y} \in C$ and $t \in \mathbb{R}$ is such that

$$(1-t)\mathbf{x} + t\mathbf{y} \in C.$$

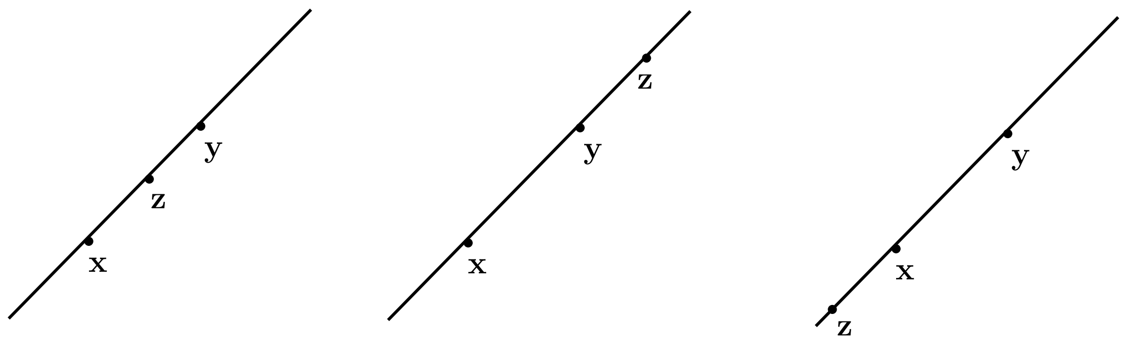
Solution: a) By Exercise 1.1 the unique line ℓ that contains both $\mathbf{x}$ and $\mathbf{y}$ is the set

$$\ell = \{(1-t)\mathbf{x} + t\mathbf{y} \mid t \in \mathbb{R}\}.$$

Assumptions imply that $z \in \ell$ as well, so there exists $t \in \mathbb{R}$ such that

$$z = (1-t)\mathbf{x} + t\mathbf{y}.$$

There are three possibilities -
1) $t[0,1]$,
2) $t > 1$ or
3) $t < 0$.



0<t<1 t>1 t<0

In case $t \in [0,1]$ the point $\mathbf{z}$ lies on the interval $[\mathbf{x}, \mathbf{y}]$ by definition and we are done.

In case $t > 1$ the picture suggests that $\mathbf{y}$ should lie on the interval $[\mathbf{x}, \mathbf{z}]$, so we solve the equation

$$\mathbf{z} = (1-t)\mathbf{x} + t\mathbf{y}$$

for $\mathbf{y}$, obtaining

$$\mathbf{y} = \frac{1}{t}(\mathbf{z} - (1-t)\mathbf{x}) = \frac{1}{t}\mathbf{z} + (1 - \frac{1}{t})\mathbf{x}.$$

Since $t > 1$, $0 < 1/t < 1$, so $\mathbf{y}$ indeed lies on the interval $[\mathbf{x}, \mathbf{z}]$.

In case $t < 0$ we solve for $\mathbf{x}$, obtaining the combination

$$\mathbf{x} = \frac{-t}{1-t}\mathbf{y} + \frac{1}{1-t}\mathbf{z}.$$

Here both scalars $\frac{-t}{1-t}$ and $\frac{1}{1-t}$ are positive and

$$\frac{-t}{1-t} + \frac{1}{1-t} = \frac{1-t}{1-t} = 1,$$

so the combination is convex. In other words $\mathbf{x}$ lies on the interval $[\mathbf{z}, \mathbf{y}]$.

b) Suppose $C, D \subset V$ are convex sets and $f \colon C \to D$ is an affine mapping. Suppose $\mathbf{x}, \mathbf{y}$

$$\mathbf{z} = (1-t)\mathbf{x} + t\mathbf{y}$$

are elements of C, where $t \in \mathbb{R}$. This means that $\mathbf{x}, \mathbf{y}, \mathbf{z}$ lie on the same line.

If $t \in [0, 1]$ then

$$f(\mathbf{z}) = (1-t)f(\mathbf{x}) + tf(\mathbf{y})$$

by the definition of affine mapping.

Suppose $t > 1$. By the proof of a) there is a simplicial combination

$$\mathbf{y} = \frac{1}{t}\mathbf{z} + (1 - \frac{1}{t})\mathbf{x},$$

so, since f is affine, we have that

$$f(\mathbf{y}) = \frac{1}{t}f(\mathbf{z}) + (1 - \frac{1}{t})f(\mathbf{x}).$$

Solving this equation back for $\mathbf{z}$ we obtain

$$f(\mathbf{z}) = (1-t)f(\mathbf{x}) + tf(\mathbf{y}).$$

The proof for the case $t < 0$ is similar - in that case we have the simplicial combination

$$\mathbf{x} = \frac{-t}{1-t}\mathbf{y} + \frac{1}{1-t}\mathbf{z},$$

so

$$f(\mathbf{x}) = \frac{-t}{1-t}f(\mathbf{y}) + \frac{1}{1-t}f(\mathbf{z})$$

and solving back for $\mathbf{z}$ yields the equation

$$f(\mathbf{z}) = (1-t)f(\mathbf{x}) + tf(\mathbf{y})$$

once more.

Remark: The claim of b) can be generalized to arbitrary long affine combinations. In other words suppose $f\colon C \to D$, where C is convex, is an affine mapping, $\mathbf{v}_0, \ldots, \mathbf{v}_n \in C$ and $t_0, \ldots, t_n \in \mathbb{R}$ are such that $t_0 + t_1 + \ldots + t_n = 1$ and

$$\mathbf{w} = t_0\mathbf{v}_0 + \ldots + t_n\mathbf{v}_n \in C.$$

We claim that then

$$f(\mathbf{w}) = t_0 f(\mathbf{v}_0) + \ldots + t_n f(\mathbf{v}_n).$$

It is tempting to prove this claim by induction on n. The part b) of this exercise is precisely the claim for $n = 2$. Case $n = 1$ is trivial.

Let us try to continue with the inductive step. Suppose the claim is true for some $n \geq 2$ and suppose $\mathbf{v}_0, \ldots, \mathbf{v}_{n+1} \in C$ and $t_0, \ldots, t_{+1}n \in \mathbb{R}$ are such that $t_0 + t_1 + \ldots + t_{n+1} = 1$ and

$$\mathbf{w} = t_0\mathbf{v}_0 + \ldots + t_{n+1}\mathbf{v}_{n+1} \in C.$$

The natural way to use inductive assumption is now to choose t_i such that $t_i \neq 1$ (such must exist), for instance we may assume that $t_{n+1} \neq 1$, and let

$$\mathbf{w}' = \frac{t_0}{1-t_{n+1}}\mathbf{v}_0 + \ldots + \frac{t_n}{1-t_{n+1}}\mathbf{v}_n.$$

Then $\mathbf{w} = (1-t_{n+1})\mathbf{w}' + t_{n+1}\mathbf{w}$, so it seems that we can apply the case $n = 2$ proved above separately. But the huge problem with this proof is that nothing guarantees that $\mathbf{w}'$ is a point of C! The point $\mathbf{w}'$ **is** an affine combination of the points $\mathbf{v}_0, \ldots, \mathbf{v}_n$ belonging to C, since the coefficients add up to 1,

$$\frac{t_0}{1-t_{n+1}}+\frac{t_1}{1-t_{n+1}}+\ldots+\frac{t_n}{1-t_{n+1}}=\frac{1}{1-t_{n+1}}(t_0+t_1+\ldots+t_n)=\frac{1}{1-t_{n+1}}(1-t_{n+1})=1.$$

But that does not help, since we are only assuming that C is convex, not necessarily affine. However this attempt is not entirely waste of time, because if the original combination

$$\mathbf{w}=t_0\mathbf{v}_0+\ldots+t_{n+1}\mathbf{v}_{n+1}$$

were simplicial, i.e. all scalars are also non-negative, the point

$$\mathbf{w}'=\frac{t_0}{1-t_{n+1}}\mathbf{v}_0+\ldots+\frac{t_n}{1-t_{n+1}}\mathbf{v}_n$$

will also be a simplicial combination of the points from C, so will belong to C. In this case we would have that $\mathbf{w}=(1-t_{n+1})\mathbf{w}'+t_{n+1}\mathbf{w}$, so, since f is affine and we assume $0 \leq t_{n+1} \leq 1$, directly from the definition we obtain

$$f(\mathbf{w})=(1-t_{n+1})f(\mathbf{w}')+t_{n+1}f(\mathbf{w}).$$

Next we use the inductive assumption, which is the claim that

$$f(\mathbf{w}')=\frac{t_0}{1-t_{n+1}}f(\mathbf{v}_0)+\ldots+\frac{t_n}{1-t_{n+1}}f(\mathbf{v}_n),$$

whenever

$$\mathbf{w}'=\frac{t_0}{1-t_{n+1}}\mathbf{v}_0+\ldots+\frac{t_n}{1-t_{n+1}}\mathbf{v}_n$$

is a **simplicial** combination. This gives us the claim for all $n \geq 1$ for **simplicial** combinations (the initial case $n=1$ is trivial and the case $n=2$ is a definition of an affine mapping).

Now that we have the claim for simplicial combinations, we prove the claim for affine combinations as following. Suppose

$$\mathbf{c}=t_0\mathbf{v}_0+t_1\mathbf{v}_1+\ldots+t_n\mathbf{v}_n \in C,$$

where $\mathbf{v}_0, \ldots, \mathbf{v}_n \in C$ and $t_0, \ldots, t_n \in \mathbb{R}$ are such that $t_0+t_1+\ldots+t_n=1$. We may assume that $t_0, t_1, \ldots, t_k \geq 0$ and $t_{k+1}, \ldots, t_n < 0$ for some $k \geq 0$. Then we have that

$$t_0\mathbf{v}_0+\ldots+t_k\mathbf{v}_k=1\cdot\mathbf{c}-t_{k+1}\mathbf{v}_{k+1}-\ldots-t_n\mathbf{v}_n. \tag{0.1}$$

Scalars on the left and on the right side are all non-negative and sum up to the same number

$$t_0 + \ldots + t_k = t = 1 - t_{k+1} - \ldots - t_n.$$

Moreover $t > 0$. Hence if we devide the equation 0.1 by t, we obtain the equation

$$\frac{t_0}{t}\mathbf{v}_0 + \ldots + \frac{t_k}{t}\mathbf{v}_k = \frac{1}{t}\mathbf{c} + \frac{-t_{k+1}}{t}\mathbf{v}_{k+1} + \ldots + \frac{-t_n}{t}\mathbf{v}_n,$$

where both left and right side are simplicial combinations. Since we already have proved the claim for simplicial combinations above, we have that

$$\frac{t_0}{t}f(\mathbf{v}_0) + \ldots + \frac{t_k}{t}f(\mathbf{v}_k) = \frac{1}{t}f(\mathbf{c}) + \frac{-t_{k+1}}{t}f(\mathbf{v}_{k+1}) + \ldots + \frac{-t_n}{t}f(\mathbf{v}_n).$$

Multiplicating by t and rearranging produces equation

$$f(\mathbf{c}) = t_0 f(\mathbf{v}_0) + t_1 f(\mathbf{v}_1) + \ldots + t_n f(\mathbf{v}_n) \in C,$$

which is precisely what we needed to show.

3. Consider the set

$$A = \{(2,1,-3), (6,3,-4), (5,2,-8), (9,4,-9)\} \subset \mathbb{R}^3$$

from the exercise 1.3. Construct an affine isomorphism $f\colon I^2 \to \mathrm{conv}(A)$. Is such an affine isomorphism between I^2 and $\mathrm{conv}(A)$ unique? If not, can you guess (no exact proof required) how many there are?

Solution: Let

$$\begin{aligned}
\mathbf{v}_0 &= (2,1,-3),\\
\mathbf{v}_1 &= (6,3,-4),\\
\mathbf{v}_2 &= (5,2,-8),\\
\mathbf{v}_2 &= (9,4,-9).
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{v}_1 - \mathbf{v}_0 = (4,2,-1),\\
\mathbf{w}_2 &= \mathbf{v}_2 - \mathbf{v}_0 = (3,1,-5),\\
\mathbf{w}_3 &= \mathbf{v}_3 - \mathbf{v}_0 = (7,3,-6).
\end{aligned}$$

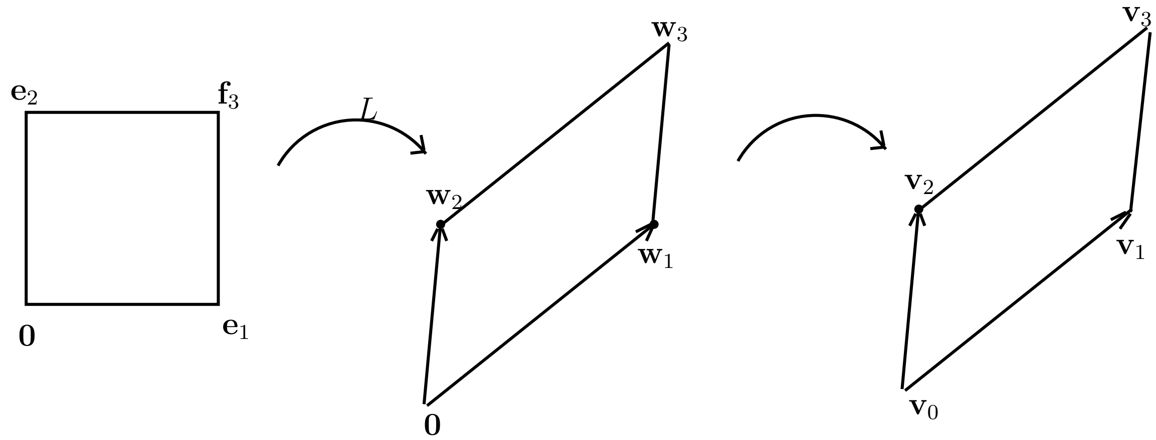
In the exercise 1.3 we have noticed that A is not affinely independent and

$$\mathbf{w}_1 + \mathbf{w}_2 = \mathbf{w}_3. \tag{0.2}$$

The vertices of the square I^2 have the same property. Precisely, let us denote them, as usual, $\mathbf{e}_0 = (0,0)$, $\mathbf{e}_1 = (1,0)$, $\mathbf{e}_2 = (0,1)$, $\mathbf{f}_3 = (1,1)$ (here we don't use $\mathbf{e}_3$ since this notation is already taken, $\mathbf{e}_3$ means the vector $(0,0,1)$ of $\mathbb{R}^3$). Then substituting $\mathbf{e}_0$ changes nothing and

$$\mathbf{e}_1 + \mathbf{e}_2 = \mathbf{f}_3,$$

the property analogous to 0.2. Hence the natural approach one can take is to first map square affinely to the parallelogram $\operatorname{conv}(B)$ spanned by the sequence $B = \{(0,0), \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$, in a natural way, mapping vertices $\mathbf{e}_0, \mathbf{e}_1 = (1,0), \mathbf{e}_2, \mathbf{f}_3$ to the vertices $(0,0), \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$, in that order. It is easy to realise that the simplest choice (and in fact the only) is a (restriction of) the linear mapping L that maps the basis $(\mathbf{e}_1, \mathbf{e}_2)$ onto $(\mathbf{w}_1, \mathbf{w}_2)$. This mapping will be an affine isomorphism between I^2 and the convex hull of $\{(0,0), \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$. Combining this mapping with the translation $\mathbf{x} \to \mathbf{x} + \mathbf{v}_0$ will produce affine isomorphism $I^2 \to \operatorname{conv}(A)$ required (see the picture).



It remains to check the details. Mapping L is the restriction of the unique linear mapping $\mathbb{R}^2 \to \mathbb{R}^3$ that maps standard basis $(\mathbf{e}_1, \mathbf{e}_2)$ onto $(\mathbf{w}_1, \mathbf{w}_2)$. Writing this down will give a formula

$$L(x_1, x_2) = x_1\mathbf{w}_1 + x_2\mathbf{w}_2 = x_1(4,2,-1) + x_2(3,1,-5) = (4x_1+3x_2, 2x_1+x_2, -x_1-5x_2).$$

This formula defines a well-defined mapping $L\colon I^2 \to \operatorname{conv}(B)$. Moreover it is affine, since it is even linear (linear mappings clearly satisfy the definition of an affine mapping). Since $L(I^2)$ is convex and clearly, by construction, contains B, $L(I^2) = \operatorname{conv}(B)$ and hence L is a surjection. Using the formula for L one can easily verify that it is an injection. Hence $L\colon I^2 \to \operatorname{conv}(B)$ is indeed affine bijection. The translation $\mathbf{x} \to \mathbf{x}+\mathbf{v}_0$ is clearly a well-defined affine bijection $\operatorname{conv}(B) \to \operatorname{conv}(A)$. Since the composition of affine bijections is an affine bijection (check!), this composition is an affine bijection $I^2 \to \operatorname{conv}(A)$.

How many affine isomorphisms $f\colon I^2 \to \operatorname{conv}(A)$ exists?

Fix any affine isomorphism $f\colon I^2 \to \operatorname{conv}(A)$, for instance the one we have constructed before. Then for any other affine isomorphism $g\colon I^2 \to \operatorname{conv}(A)$ the composite mapping $g^{-1} \circ f\colon I^2 \to I^2$ is an **automorphism** of the square. "Automorphism" is an isomorphism of the set to itself. Conversely any automorphism $\alpha\colon I^2 \to I^2$ of the square defines an affine isomorphism $g = f \circ \alpha\colon I^2 \to \operatorname{conv}(A)$. It is easy to see that the correspondence $\alpha \to f \circ \alpha$ is bijective, so it is enough to count how many affine isomorphisms $\alpha\colon I^2 \to I^2$ of the square to itself there is.

The first observation is that any affine isomorphism $\alpha\colon I^2 \to I^2$ preserves **extreme points** (for the concept of extreme points see exercise 1). Since extreme points of I^2 are exactly the vertices $\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{f}_3$ (see exercise 1.4), we see that an affine isomorphism $\alpha\colon I^2 \to I^2$ permutes these 4 extreme points. There are $4! = 24$ permutations of the set of four elements, however not all of these can be extended to an affine mapping.

Next we observe that any affine isomorphism $\alpha\colon I^2 \to I^2$ maps the "middle" point $(\frac{1}{2}, \frac{1}{2})$ to itself i.e. leaves it fixed. The reason for that is that this point is the **the only** point $\mathbf{x}$ of the square that can be written in the form

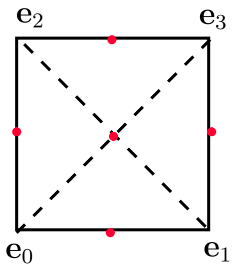
$$\mathbf{x} = \frac{1}{2}\mathbf{v}_i + \frac{1}{2}\mathbf{v}_j$$

where $\mathbf{v}_i$ and $\mathbf{v}_j$ are two different vertices of the square, **in two different ways**. Indeed,

$$(\frac{1}{2}, \frac{1}{2}) = \frac{1}{2}\mathbf{e}_0 + \frac{1}{2}\mathbf{e}_3 \text{ and}$$

$$(\frac{1}{2}, \frac{1}{2}) = \frac{1}{2}\mathbf{e}_1 + \frac{1}{2}\mathbf{e}_2.$$

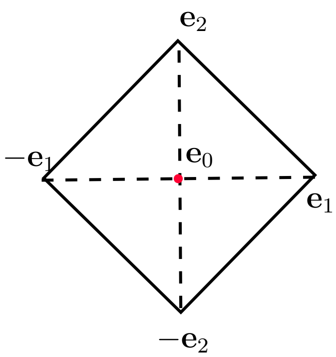
All the other middle points of the intervals, which end points are vertices, are the middle points of only one such interval - see the picture, in which all such points are marked with red colour.



Now since affine isomorphism $\alpha\colon I^2 \to I^2$ maps the point of the form

$$\mathbf{x} = \frac{1}{2}\mathbf{v}_i + \frac{1}{2}\mathbf{v}_j$$

to the point of the same form, it follows that the middle points $(1/2, 1/2)$ is mapped to itself. Translating this point to the origin and rotating the square we can reduce the problem to the problem of counting how many affine automorphism of the rhombus C with the vertices $(\mathbf{e}_1, \mathbf{e}_2, -\mathbf{e}_1, -\mathbf{e}_2)$, that map origin to itself, there is.



Let us show that an affine mapping $\alpha\colon C \to C$ that maps origin to itself is (restriction of) the linear mapping. Suppose $\mathbf{x} \in C$ and $t \in \mathbb{R}$ are such that $t\mathbf{x} =\in C$. Then

$$t\mathbf{x} = (1-t)\mathbf{0} + t\mathbf{x}$$

belongs to C and is an affine combination of two points of C. By Exercise 2b) above, affine mapping α preserves such combinations, so

$$\alpha(t\mathbf{x}) = (1-t)\alpha(\mathbf{0}) + t\alpha(\mathbf{x}) = t\alpha(\mathbf{x}),$$

since we are assuming that $\alpha(\mathbf{0}) = \mathbf{0}$.

Next suppose $\mathbf{x}, \mathbf{y} \in C$ are such that

$$\mathbf{x} + \mathbf{y} \in C.$$

By convexity of C we also have that

$$\mathbf{u} = \frac{1}{2}\mathbf{x} + \frac{1}{2}\mathbf{y} \in C.$$

Now $\mathbf{u}$ and

$$2\mathbf{u} = \mathbf{x} + \mathbf{y} \in C.$$

By the property $\alpha(t\mathbf{x}) = t\alpha(\mathbf{x})$ proved above and applied for $t = 2$, we have, using also the fact that α is an affine mapping, that

$$\alpha(\mathbf{x} + \mathbf{y}) = \alpha(2\mathbf{u}) = 2\alpha(\mathbf{u}) = 2(\alpha(\frac{1}{2}\mathbf{x} + \frac{1}{2}\mathbf{y})) =$$

$$= 2(\frac{1}{2}\alpha(\mathbf{x}) + \alpha(\frac{1}{2}\mathbf{y})) = \alpha(\mathbf{x}) + \alpha(\mathbf{y}).$$

We have thus shown that α satisfies the conditions of a linear mapping i.e.

$$\alpha(t\mathbf{x}) = t\alpha(\mathbf{x}) \text{ and}$$

$$\alpha(\mathbf{x} + \mathbf{y}) = \alpha(\mathbf{x}) + \alpha(\mathbf{y})$$

whenever they are well-defined, i.e. when all arguments of α belong to C. Using this we proceed as following. Suppose $\mathbf{x} = (x_1, x_2) \in C$. Then

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2,$$

where $x_1\mathbf{e}_1$ and $x_2\mathbf{e}_2$ both belong to C. Hence

$$\alpha(\mathbf{x}) = \alpha(x_1\mathbf{e}_1) + \alpha(x_2\mathbf{e}_2) = x_1\alpha(\mathbf{e}_1) + x_2\alpha(\mathbf{e}_2).$$

This implies that α is just a restriction of a linear mapping

$$L\colon \mathbb{R}^2 \to \mathbb{R}^2$$

defined by

$$L(x_1, x_2) = x_1\mathbf{a}_1 + x_2\mathbf{a}_2,$$

where $\mathbf{a}_1 = \alpha(\mathbf{e}_1)$, $\mathbf{a}_2 = \alpha(\mathbf{e}_2)$.

Now all we have to do is to calculate how many there are linear isomorphisms $L\colon \mathbb{R}^2 \to \mathbb{R}^2$, that map rhombus C to itself. The linear mapping is completely determined by the image of the basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$, which belong to the rhombus and are in fact its extreme points. Since extreme points must map to extreme points, $\mathbf{e}_1$ must map onto one of the vertices $\mathbf{e}_1, \mathbf{e}_2, -\mathbf{e}_1, -\mathbf{e}_2$. Thus there are 4 choices for the image $L(\mathbf{e}_1)$. Once this choice is made, there are only 2 choices left for the image $L(\mathbf{e}_2)$. Indeed, suppose $L(\mathbf{e}_1) \in \{\mathbf{e}_1, -\mathbf{e}_1\}$. Then $L(\mathbf{e}_2)$ cannot map to the same points, since then, by linearity, L will not be isomorphism. Thus $L(\mathbf{e}_2)$ is $\mathbf{e}_2$ or $-\mathbf{e}_2$. The similar conclusion holds when $L(\mathbf{e}_1) \in \{\mathbf{e}_2, -\mathbf{e}_2\}$. Thus amount of different choices is $4 \cdot 2 = 8$.

Answer: There are exactly 8 different affine automorphisms of the square and consequently 8 different affine isomorphisms $I^2 \to \mathrm{conv}(A)$.

Remark: Similarly as above one can show in general that an affine mapping is always a composition of some linear mapping and a translation mapping.

4. a) Recall (or google) the precise definitions (and be ready to present them) of the following concepts - inner product in a vector space, norm defined by an inner product.
b) Recall (or google) the proof of the **Schwartz inequality**

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq |\mathbf{v}||\mathbf{w}|.$$

Here $\langle , \rangle$ is an inner product in a vector space and $|\cdot|$ is a norm defined by this inner product. Also prove that the equality

$$\langle \mathbf{v}, \mathbf{w} \rangle = |\mathbf{v}||\mathbf{w}|.$$

holds if and only if $\mathbf{v} = \mathbf{0}$ or there exists $t \geq 0$ such that $\mathbf{w} = t\mathbf{w}$.
c) Recall (or google) the proof of the **triangle inequality**

$$|\mathbf{v} + \mathbf{w}| \leq |\mathbf{v}| + |\mathbf{w}|$$

(using Schwartz inequality). Here $\langle , \rangle$ is an inner product in a vector space and $|\cdot|$ is a norm defined by this inner product. Also prove that the equality

$$|\mathbf{v} + \mathbf{w}| = |\mathbf{v}| + |\mathbf{w}|.$$

holds if and only if $\mathbf{v} = \mathbf{0}$ or there exists $t \geq 0$ such that $\mathbf{w} = t\mathbf{v}$.
d) Apply the previous result to show that every point of the sphere S^{n-1} is an extreme point of the convex set $\overline{B}^n$. The concept of the

extreme point was defined in Exercises 1.

Solution: a) Suppose V is vector space. A function $\langle , \rangle : V \times V \to \mathbb{R}$ is called *an inner product* in V if it satisfies the equations

$$\langle \mathbf{v} + \mathbf{v}', \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{v}', \mathbf{w} \rangle,$$

$$\langle \mathbf{v}, \mathbf{w} + \mathbf{w}' \rangle = \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w}' \rangle,$$

$$\langle t\mathbf{v}, \mathbf{w} \rangle = t\langle \mathbf{v}, \mathbf{w} \rangle,$$

$$\langle \mathbf{v}, t\mathbf{w} \rangle = \langle \mathbf{v}, t\mathbf{w} \rangle,$$

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle,$$

$$\langle \mathbf{v}, \mathbf{v} \rangle > 0 \text{ for all } \mathbf{v} \in V, \mathbf{v} \neq 0.$$

In other words inner product is a symmetric real-valued function of two arguments (both are vectors from V) which is linear with respect to every argument i.e. so-called *bilinear mapping.* The last condition says that the inner product of every non-zero vector with itself is a strictly positive real number. The inner product of zero vector with itself, or any vector, is always zero, this follows from bilinearity.

The norm $|\cdot|$ defined by the inner product $\langle , \rangle$ in a vector space V is defined by the formula

$$|\mathbf{v}| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

The standard norm of $\mathbb{R}^n$ is induced by the standard inner product $\cdot$ in $\mathbb{R}^n$ defined by

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i.$$

b) The Schwartz inequality

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq |\mathbf{v}||\mathbf{w}|$$

can be proved as following. In case $\mathbf{v} = 0$ the inequality is clear, and in fact even the equality holds.

Assume $\mathbf{v} \neq 0$. Let $t \in \mathbb{R}$ be arbitrary. Then

$$0 \leq \langle \mathbf{w} - t\mathbf{v}, \mathbf{w} - t\mathbf{v} \rangle = t^2 \langle \mathbf{v}, \mathbf{v} \rangle - 2t\langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle =$$

$$= t^2 |\mathbf{v}|^2 - 2t\langle \mathbf{v}, \mathbf{w} \rangle + |\mathbf{w}|^2.$$

The expression on the right is the polynomial of second degree in terms of t, which is now non-negative for all $t \in \mathbb{R}$. In this s This is equivalent to the assertion that the discriminant D of this polynomial is non-negative, i.e.

$$D = 4|\langle \mathbf{v}, \mathbf{w} \rangle|^2 - 4|\mathbf{v}|^2|\mathbf{w}|^2 \geq 0.$$

This is equivalent to

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq |\mathbf{v}||\mathbf{w}|$$

which is exactly the Schwartz inequality.

Let us investigate when the Schwartz inequality is the strict equality

$$|\langle \mathbf{v}, \mathbf{w} \rangle| = |\mathbf{v}||\mathbf{w}|.$$

The case $\mathbf{v} = 0$ is already taken care of, so we assume again that $\mathbf{v} \neq 0$.

By the proof above the equality holds if and only if $D = 0$, which is equivalent to the existence of $t \in \mathbb{R}$ for which the quadratic function above is zero, i.e. the existence of $t \in \mathbb{R}$ for which

$$0 = \langle \mathbf{w} - t\mathbf{v}, \mathbf{w} - t\mathbf{v} \rangle = |\langle \mathbf{w} - t\mathbf{v}, \mathbf{w} - t\mathbf{v} \rangle|.$$

The norm of the vector is, by definition of the inner product, can be negative if and only if the vector is zero. Hence the Schwartz inequality is the equality precisely when $\mathbf{v} = 0$ or when there exists $t \in \mathbb{R}$ such that $\mathbf{w} - t\mathbf{v} = 0$ i.e.

$$\mathbf{w} = t\mathbf{v}.$$

When $t < 0$ and $\mathbf{v} \neq 0$, the equation holds in the form

$$-\langle \mathbf{v}, \mathbf{w} \rangle = |\mathbf{v}||\mathbf{w}|.$$

When $t \geq 0$, the equation holds in the form

$$\langle \mathbf{v}, \mathbf{w} \rangle = |\mathbf{v}||\mathbf{w}|.$$

Thus the equation

$$\langle \mathbf{v}, \mathbf{w} \rangle = |\mathbf{v}||\mathbf{w}|.$$

holds if and only if $\mathbf{v} = \mathbf{0}$ or when there exists $t \geq 0$ such that $\mathbf{w} = t\mathbf{v}$.

c) The standard proof of the triangle inequality uses Schwartz inequality,

$$|\mathbf{v}+\mathbf{w}|^2 = \langle \mathbf{v}+\mathbf{w} \rangle = |\mathbf{v}|^2 + 2\langle \mathbf{v}, \mathbf{w} \rangle + |\mathbf{w}|^2 \leq |\mathbf{v}|^2 + 2|\mathbf{v}||\mathbf{w}| + |\mathbf{w}|^2 = \Big(|\mathbf{v}| + |\mathbf{w}|\Big)^2.$$

Taking square root from both sides yields the triangle inequality. Also, the prove above shows that the equality holds if and only

$$\langle \mathbf{v}, \mathbf{w} \rangle = |\mathbf{v}||\mathbf{w}|.$$

By b) this is possible if and only if $\mathbf{v} = \mathbf{0}$ or there exists $t \geq 0$ such that $\mathbf{w} = t\mathbf{v}$.

d) Suppose $\mathbf{x} \in S^{n-1}$. We prove that $\mathbf{x}$ is an extreme point of the ball $\overline{B}^n$. Let us make a counter-assumption, suppose

$$\mathbf{x} = (1-t)\mathbf{y} + t\mathbf{z}$$

for some $\mathbf{y}, \mathbf{z} \in \overline{B}^n$, $t \in]0,1[$ and $\mathbf{y} \neq \mathbf{z}$. Then, by the triangle inequality

$$1 = |\mathbf{x}| \leq (1-t)|\mathbf{y}| + t|\mathbf{z}|.$$

Here $|\mathbf{y}|, |\mathbf{z}| \leq 1$. Moreover, if at least one of the inequalities is strict, for instance $|\mathbf{y}| < 1$, then also

$$1 \leq (1-t)|\mathbf{y}| + t|\mathbf{z}| < (1-t) + t = 1,$$

which is impossible. Hence $\mathbf{y}, \mathbf{z} \in S^{n-1}$ and in the application of the triangle inequality above, the equality must hold. By c), there must exist $s \geq 0$ such that

$$(1-t)\mathbf{y} = st\mathbf{z}.$$

Taking norms from both sides of this equality gives us

$$(1-t) = st,$$

hence

$$(1-t)\mathbf{y} = st\mathbf{z} = (1-t)\mathbf{z},$$

which implies that $\mathbf{y} = \mathbf{z}$. This contradicts the assumption.

5. Define the topology τ in $\mathbb{R}$ as following. Suppose $U \subset \mathbb{R}$. Then U is open if and only if for every $x \in U$ there exists half-open interval $[a, b[$ such that $x \in [a, b[$ and $[a, b[\subset U$.
a) Show that τ is indeed a topology in $\mathbb{R}$. Show that every subset of $\mathbb{R}$ which is open with respect to the standard topology of $\mathbb{R}$ is also open with respect to τ, but the converse statement is not true.
b) Is open interval $]a, b[$ open or closed with respect to τ? What about intervals of the form $[a, b[$, $]a, b]$, $[a, b]$?

c) Show that the closed interval $[a, b]$ is not compact with respect to τ.
d) The topological space is called totally disconnected if the only connected subsets of the space are empty set and singletons. Show that all intervals of $\mathbb{R}$ are totally disconnected with respect to τ.

Solution:
a) Emptyset and the whole real line are trivially open. Suppose $(U_i)_{i \in I}$ is a collection of open sets and let

$$U = \bigcup_{i \in I} U_i.$$

Suppose $x \in U$. Then $x \in U_i$ for some $i \in I$ and by definition of τ there exists a, b such that

$$x \in [a, b[\subset U_i \subset U.$$

Since this is true for every $x \in U$, $U \in \tau$.

Suppose $U_1, U_2, \ldots, U_n \in \tau$. Let

$$x \in \bigcap_{i=1}^{n} U_i.$$

For every $i = 1, \ldots, n$ there exists a_i, b_i such that $x \in [a_i, b_i[\subset U_i$. Let

$$a = \max\{a_1, \ldots, a_n\},$$

$$b = \min\{b_1, \ldots, b_n\}.$$

Then

$$x \in [a, b[\subset \bigcap_{i=1}^{n} U_i.$$

Hence also finite intersection of open sets is open.

Suppose U is open in standard topology of $\mathbb{R}$ and let $x \in U$. By definition of standard topology there exist $a, b \in \mathbb{R}$ such that

$$x \in]a, b[\subset U.$$

But then

$$x \in [x, b[\subset U.$$

This implies that U is also open with respect to τ.

The converse is not true - interval $[a, b[$ is open w.r.t. τ, but not with respect to standard topology.

b) Every open interval is open w.r.t to τ by a), since it is open even w.r.t. standard topology. It is not closed since its compliment

$$\mathbb{R}\backslash]a, b[=]-\infty, a] \cup [b, +\infty[= B$$

is not open. This is seen by observing that for $a \in B$ it is impossible to find an interval of the form $[c, d[$ such that $a \in [c, d[$ and $[c, d[\subset B$.

Closed interval $[a, b]$ is closed, since its compliment is open in standard topology, so also open w.r.t. to τ. Closed interval is not open, since for a point $b \in [a, b]$ we cannot find an interval of the form $[c, d[$ such that $b \in [c, d[$ and $[c, d[\subset [a, b]$.

Half-open interval $[a, b[$ is open in τ by the very definition of τ - every point $x \in [a, b[$ has interval $[a, b[$ such that

$$x \in [a, b[\subset [a, b[.$$

Half-open interval $[a, b[$ is also closed, since its compliment

$$\mathbb{R} \setminus [a, b[=]-\infty, a[\cup [b, +\infty[$$

is open, as a union of two open sets first is open since it is open in standard topology and the second is the union

$$[b, \infty[= \bigcup_{n \geq 1} [b, b+n[$$

of half-open intervals that we already proved to be open in this topology.

Half-open interval $]a, b]$ is not open. This is seen as before observing that b cannot have a suitable neighbourhood $[c, d[$ containing in the set. For the similar reason the compliment

$$\mathbb{R}\backslash]a, b] =]-\infty, a] \cup]b, +\infty[$$

is not open (problem in a), so the half-open interval $]a, b]$ is not closed either.

c) The covering

$$\{[a, b - \frac{1}{n}[\mid n \geq 1\} \cup \{[b, b+1[\}$$

is an open covering of $[a, b]$ that has no finite subcover - any finite union of sets from this cover do not contain points $b - \varepsilon$ for small $\varepsilon > 0$. Thus $[a, b]$ is not compact.

d) It is enough to prove that the only non-empty connected subsets of $\mathbb{R}$ w.r.t. τ are singletons. Suppose $A \subset \mathbb{R}$ is a set that contains at least two different points $\mathbf{x}, \mathbf{y} \in A$. Choose a point $\mathbf{a}$ between $\mathbf{x}$ and $\mathbf{y}$. Then intervals $B_1 =]-\infty[, a[$ and $B_2 = [a, \infty[$ are open, disjoint and their union is the whole $\mathbb{R}$. This implies that $A \cap B_1$ and $A \cap B_2$ is a separation of A into two non-empty, open, disjoint sets, so A is not connected.

Remark: Notice that the definition of τ is very similar to the standard topology of $\mathbb{R}$, the only difference is that the interval $[a, b[$ instead of the interval $]a, b[$ is used. Observe however dramatic difference that adding of one end point has made. The closed intervals which are compact in standard topology are not compact any more. Moreover intervals, which are connected in standard topology, are even totally disconnected w.r.t. to τ.

6. a) Suppose $\mathbf{y} \in S^n, \mathbf{y} \neq \mathbf{e}_{n+1}$. Show that the unique line ℓ that goes through both $\mathbf{y}$ and $\mathbf{e}_{n+1}$ intersects the set

$$\mathbb{R}^n = \{\mathbf{x} \in \mathbb{R}^{n+1} \mid \mathbf{x}_{n+1} = 0\} \subset \mathbb{R}^{n+1}.$$

in exactly one point, which we denote $p(\mathbf{y})$.
b) Show that $p \colon S^n \setminus \{\mathbf{e}_{n+1}\} \to \mathbb{R}^n$ is given by the formula

$$p(\mathbf{y}) = \frac{1}{1 - \mathbf{y}_{n+1}}(\mathbf{y}_1, \ldots, \mathbf{y}_n).$$

Show that p is a continuous bijection and its inverse is given by the formula

$$p^{-1}(\mathbf{y}) = \frac{1}{|\mathbf{y}|^2 + 1}(2\mathbf{y} + (|\mathbf{y}|^2 - 1)\mathbf{e}_{n+1}).$$

Solution: a) By Exercise 1.1 the unique line ℓ that contains $\mathbf{y}$ and $\mathbf{e}_{n+1}$ is the set

$$\{(1-t)\mathbf{y} + t\mathbf{e}_{n+1} \mid t \in \mathbb{R}\}.$$

The point of the line is in the subspace $\mathbb{R}^n$, if its last $(n+1)$-component is exactly zero i.e. if

$$(1-t)y_{n+1} + t = 0.$$

This is true if and only

$$t = \frac{y_{n+1}}{y_{n+1} - 1}.$$

Thus the unique intersection point of ℓ and $\mathbb{R}^n$ is the point

$$(1 - \frac{y_{n+1}}{y_{n+1}-1})\mathbf{y} + \frac{y_{n+1}}{y_{n+1}-1}\mathbf{e}_{n+1} = \frac{1}{1-\mathbf{y}_{n+1}}(\mathbf{y}_1, \ldots, \mathbf{y}_n),$$

which gives the formula for p.

b) The mapping $p\colon S^n \setminus \{\mathbf{e}_{n+1}\} \to \mathbb{R}^n$ defined by

$$p(\mathbf{y}) = \frac{1}{1-\mathbf{y}_{n+1}}(\mathbf{y}_1, \ldots, \mathbf{y}_n)$$

is clearly well-defined and continuous. Since the formula for the inverse mapping $q\colon \mathbb{R}^n \to S^n \setminus \{\mathbf{e}_{n+1}\}$ is given,

$$q(\mathbf{y}) = \frac{1}{|\mathbf{y}|^2 + 1}(2\mathbf{y} + (|\mathbf{y}|^2 - 1)\mathbf{e}_{n+1}),$$

it is enough to verify that this mapping is well-defined i.e. maps $\mathbb{R}^n$ onto $S^n \setminus \{\mathbf{e}_{n+1}\}$ and satisfies equations $p \circ q = \mathrm{id}$, $q \circ p$ is identity. The continuity of q is clear.

For the first assertion one calculates

$$|q(\mathbf{y})|^2 = \frac{1}{(|\mathbf{y}|^2 + 1)^2}(4|\mathbf{y}|^2 + (|\mathbf{y}|^2 - 1)^2 = 1.$$

Also n first coordinates of $q(\mathbf{y})$ can all be zeros if and only if $\mathbf{y} = 0$, in which case the last coordinate is -1, not 1. Hence q indeed map $\mathbb{R}^n$ into $S^n \setminus \{\mathbf{e}_{n+1}\}$. The verification of equations $p \circ q = \mathrm{id}$ and $q \circ p = \mathrm{id}$ is a straightforward calculation.

If the formula for the inverse q of p would not be given in the exercise, it can be easily constructed as following. Let $\mathbf{y} \in \mathbb{R}^n$ be arbitrary. The definition of p implies that $q(\mathbf{y})$ must be unique point of the set $S^n \setminus \{\mathbf{e}_{n+1}\}$ that lies on the line

$$\{(1-t)\mathbf{y} + t\mathbf{e}_{n+1} \mid t \in \mathbb{R}\}$$

(the unique line passing through $\mathbf{y}$ and $\mathbf{e}_{n+1}$).

Thus we must find $t \in \mathbb{R}$ for which the point of the form $(1-t)\mathbf{y}+t\mathbf{e}_{n+1}$ lies on the sphere but is not its north pole. This implies the equation

$$(1-t)^2|\mathbf{y}|^2 + t^2 = 1.$$

Simplifying yields equation

$$(|\mathbf{y}|^2 + 1)t^2 - 2|\mathbf{y}|^2 t + (|\mathbf{y}^2| - 1) = 0.$$

Solving this equation of the second degree will give

$$t = \frac{|y|^2 \pm 1}{|y|^2 + 1}.$$

When the sigh in the numerator is $+$, we obtain the north pole. When the sign is $-$, the point

$$(1-t)\mathbf{y} + t\mathbf{e}_{n+1}$$

will be precisely $q(\mathbf{y})$. This also provides with a geometric way to prove that p is bijection and its inverse is q.