

22.1 Affine Spaces and Affine Maps

Every space in this book so far has had a distinguished point: the zero vector. Geometry does not. A plane drawn on paper has no privileged spot, and none of the statements a geometer cares about — three points are collinear, two lines are parallel, this point is the midpoint of that segment — mentions one. This section builds the object that has all the structure of a vector space except the origin, and the maps between such objects. Everything in the rest of the chapter is written in this language.

22.1.1 A set with no zero in it

Chapter 3 already produced such an object without naming it. Let $\mathbf{A} \in M_{m \times n}(F)$ and $\mathbf{b} \in F^m$, and suppose $\mathbf{Ax} = \mathbf{b}$ is consistent. By Theorem 3.1.3 the solution set is $\mathbf{p} + N$, where $N = \text{nul}(\mathbf{A})$ and $\mathbf{p}$ is any one solution. When $\mathbf{b} \neq \mathbf{0}$ this set is **not** a subspace: it misses $\mathbf{0}$, and the sum of two solutions is not a solution. Yet it is not shapeless either. The difference of two solutions is a vector of N , and adding a vector of N to a solution gives a solution. So the solution set carries an action of N , and N records exactly how any two solutions differ.

That is the whole structure, and it deserves a name. The only operation is: *subtract two points and get a vector; add a vector to a point and get a point*. Adding two points is meaningless, and so is multiplying a point by a scalar.

An affine space is a set of points in which differences of points are vectors, and every vector moves every point to exactly one other.

Definition 22.1.1. Affine Space

Let V be a vector space over a field F . An **affine space with direction space V** is a **nonempty** set $\mathcal{A}$, whose elements are called **points**, together with a map

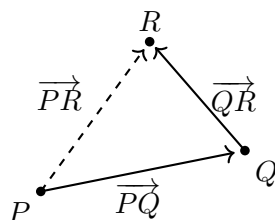
$$\mathcal{A} \times \mathcal{A} \rightarrow V, \quad (P, Q) \mapsto \overrightarrow{PQ},$$

such that:

- (AS1) (**Chasles's relation**) $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$ for all $P, Q, R \in \mathcal{A}$;
- (AS2) **for every** $P \in \mathcal{A}$, the map $\mathcal{A} \rightarrow V, Q \mapsto \overrightarrow{PQ}$, is a **bijection**.

The **dimension** of $\mathcal{A}$ is $\dim \mathcal{A} := \dim V$. For $P \in \mathcal{A}$ and $v \in V$ we write $P + v$ for the unique point Q with $\overrightarrow{PQ} = v$, which exists and is unique by (AS2).

Read the two clauses. (AS1) says that traveling from P to Q and then from Q to R is traveling from P to R ; it is the one law that makes the arrows behave like displacements. (AS2) says that from any fixed point P , every vector of V is the displacement to exactly one point: **existence** makes $\mathcal{A}$ big enough (the action is *transitive*), **uniqueness** makes it no bigger (the action is *free*). The set $\mathcal{A}$ is not a vector space and has no zero; V is a vector space and has one.



Two consequences are immediate and will be used constantly.

Lemma 22.1.1. First consequences of Chasles

Let $\mathcal{A}$ be an affine space with direction space V . For all $P, Q \in \mathcal{A}$ and $u, w \in V$:

- (a) $\overrightarrow{PP} = 0$, and $\overrightarrow{QP} = -\overrightarrow{PQ}$;
- (b) $P + 0 = P$ and $(P + u) + w = P + (u + w)$;
- (c) $\overrightarrow{PQ} = 0$ if and only if $P = Q$.

Proof.

- (a) Taking $P = Q = R$ in (AS1) gives $\overrightarrow{PP} + \overrightarrow{PP} = \overrightarrow{PP}$, so $\overrightarrow{PP} = 0$ by cancellation in the group $(V, +)$ (Theorem 0.10.1). Taking $R = P$ in (AS1) gives $\overrightarrow{PQ} + \overrightarrow{QP} = \overrightarrow{PP} = 0$.
- (b) $P + 0 = P$ restates $\overrightarrow{PP} = 0$. For the second identity, put $Q = P + u$ and $R = Q + w$, so that $\overrightarrow{PQ} = u$ and $\overrightarrow{QR} = w$. By (AS1), $\overrightarrow{PR} = u + w$, which says $R = P + (u + w)$.
- (c) $(\Leftarrow)$ is (a). $(\Rightarrow)$ If $\overrightarrow{PQ} = 0 = \overrightarrow{PP}$, then $Q = P$ by the injectivity in (AS2).

■

The clauses in part (b) are the axioms of an action of the additive group of V on $\mathcal{A}$, and the definition is often given that way instead. The two presentations carry the same information.

Proposition 22.1.2. The two forms of the definition agree

Let V be a vector space over F and $\mathcal{A}$ a nonempty set. Giving a map $\mathcal{A} \times \mathcal{A} \rightarrow V$ satisfying (AS1) and (AS2) is the same as giving a map $\mathcal{A} \times V \rightarrow \mathcal{A}$, written $(P, v) \mapsto P + v$, such that

- (AC1) $P + 0 = P$ for every $P \in \mathcal{A}$;
- (AC2) $(P + u) + w = P + (u + w)$ for all $P \in \mathcal{A}$ and $u, w \in V$;
- (AC3) for all $P, Q \in \mathcal{A}$ there is **exactly one** $v \in V$ with $P + v = Q$.

The two constructions are inverse to each other.

Idea. There is nothing to find: in one direction the addition was already defined after [Definition 22.1.1](#), in the other the vector $\overrightarrow{PQ}$ is the one whose uniqueness (AC3) asserts. The work is checking the axioms, and each check is one line.

Proof. Suppose first that $(P, Q) \mapsto \overrightarrow{PQ}$ satisfies (AS1) and (AS2), and define $P + v$ as after [Definition 22.1.1](#). Then (AC1) and (AC2) are [Lemma 22.1.1](#) (b), and (AC3) is (AS2), since $P + v = Q$ says $\overrightarrow{PQ} = v$.

Conversely, suppose $(P, v) \mapsto P + v$ satisfies (AC1)–(AC3), and let $\overrightarrow{PQ}$ be the unique v with $P + v = Q$. For (AS1), put $u = \overrightarrow{PQ}$ and $w = \overrightarrow{QR}$; then by (AC2),

$$P + (u + w) = (P + u) + w = Q + w = R,$$

so $\overrightarrow{PR} = u + w$ by the uniqueness in (AC3). For (AS2), fix P . The map $Q \mapsto \overrightarrow{PQ}$ is surjective because v is the image of $P + v$, and injective because $\overrightarrow{PQ} = \overrightarrow{PQ'} = v$ gives $Q = P + v = Q'$.

The two passages are inverse: starting from $\overrightarrow{}$, the constructed addition has $P + v = Q$ exactly when $\overrightarrow{PQ} = v$, so the difference map it returns is the original one; and starting from an addition, the same equivalence returns it unchanged. This proves the proposition. ■

Remark. The converse half used only (AC2) and (AC3), and it could not have used more: (AC1) follows from the other two. Indeed (AC3) with $Q = P$ gives exactly one v with $P + v = P$, and that v has $P + (v + v) = (P + v) + v = P + v = P$ by (AC2), so $v + v = v$ by the same uniqueness and $v = 0$. We list (AC1) anyway, because that is how the axioms of a group action are usually written.

Now the examples. The first is the one every other example is modeled on.

Example. Four affine spaces

Check each of the following against [Definition 22.1.1](#), and say what its direction space and dimension are.

- (a) A vector space V over F , with $\overrightarrow{xy} := y - x$.
- (b) A coset $p + U$ of a subspace U of V (Definition 4.4.2), with the same difference map.
- (c) The solution set of a **consistent** system $\mathbf{Ax} = \mathbf{b}$, $\mathbf{A} \in M_{m \times n}(F)$.
- (d) The set $\mathbb{F}_2^2$ of four points, with $\overrightarrow{\mathbf{x}\mathbf{y}} = \mathbf{y} - \mathbf{x}$.

Solution.

- (a) Chasles reads $(y - x) + (z - y) = z - x$, which is true in any vector space. For fixed x , the map $y \mapsto y - x$ is a bijection $V \rightarrow V$ with inverse $v \mapsto x + v$. Direction space V ; dimension $\dim V$. Here the point x and the vector x are the same object, and only the language distinguishes them.
- (b) The set $p + U$ is nonempty (it contains p) and the difference of two of its elements is $(p + u) - (p + w) = u - w \in U$, so the map does land in U . Chasles is again the identity in (a). For fixed $x \in p + U$, the map $y \mapsto y - x$ is a bijection $p + U \rightarrow U$: it is injective, and $u \in U$ is the image of $x + u$, which lies in $p + U$ because U is closed under addition. Direction space U ; dimension $\dim U$. Over $\mathbb{R}$ these are exactly Chapter 18's affine subspaces of V (Definition 18.1.5 (b)), with the same dimension.
- (c) By Theorem 3.1.3 the solution set is $\mathbf{p} + \text{nul}(\mathbf{A})$ for any one solution $\mathbf{p}$, so this is a special case of (b). Direction space $\text{nul}(\mathbf{A})$; dimension $\text{nullity}(\mathbf{A}) = n - \text{rank } \mathbf{A}$ by Theorem 3.5.7. Consistency is what makes the set nonempty, and Theorem 3.5.8 says when it holds.
- (d) A special case of (a) with $V = \mathbb{F}_2^2$: four points, direction space $\mathbb{F}_2^2$, dimension 2. It is worth seeing because it is so small. Every pair of distinct points $P \neq Q$ has $\overrightarrow{PQ} = \overrightarrow{QP}$, since $-1 = 1$ in $\mathbb{F}_2$; the next section will call a coset of a one-dimensional subspace a **line**, and here each such coset has exactly two points. So every one of the $\binom{4}{2} = 6$ pairs

of distinct points is itself a line, and the six lines fall into three pairs with no point in common.

A nonexample, by minimal change. Delete one point from the plane: let $\mathcal{A} = \mathbb{R}^2 \setminus \{\mathbf{0}\}$, with $\overrightarrow{\mathbf{x}\mathbf{y}} = \mathbf{y} - \mathbf{x}$. Chasles still holds, because it is an identity about subtraction. But (AS2) fails: for $P = (1, 0)$, no point Q of $\mathcal{A}$ has $\overrightarrow{PQ} = (-1, 0)$, since the only candidate is $\mathbf{0}$, which we removed. The map $Q \mapsto \overrightarrow{PQ}$ is injective but not surjective. It is exactly the transitivity of the action that fails: the set has a hole.

Warning. A point is not a vector. In an affine space there is no zero point; and over a field in which $2 \neq 0$ and $3 \neq 0$ there is no sum $P + Q$ of two points and no multiple $3P$. The only legal expressions are $\overrightarrow{PQ} \in V$, $P + v \in \mathcal{A}$, and the affine combinations defined below. If you find yourself writing $P + Q$, you have silently chosen an origin; the next subsection says what that costs.

22.1.2 Choosing an origin

An affine space looks like its direction space as soon as one point is nominated as the zero. This is the chapter's first named move, and it is used in almost every proof that follows: *to compute, pick an origin; to state a result, do not.*

Proposition 22.1.3. Choice of Origin

Let $\mathcal{A}$ be an affine space with direction space V , and let $O \in \mathcal{A}$. Then

$$\theta_O: \mathcal{A} \rightarrow V, \quad \theta_O(P) = \overrightarrow{OP},$$

is a bijection with $\theta_O(O) = 0$, and it carries the affine structure of $\mathcal{A}$ to that of V in [Example \(Four affine spaces\)](#) (a):

$$\overrightarrow{PQ} = \theta_O(Q) - \theta_O(P) \quad (P, Q \in \mathcal{A}).$$

If $O' \in \mathcal{A}$ is another point, then $\theta_{O'}(P) = \theta_O(P) - \theta_O(O')$ for every P : the two identifications differ by a translation of V .

Proof. θ_O is a bijection by (AS2), and $\theta_O(O) = \overrightarrow{OO} = 0$ by [Lemma 22.1.1](#) (a). By Chasles and the same lemma,

$$\theta_O(Q) - \theta_O(P) = \overrightarrow{OQ} - \overrightarrow{OP} = \overrightarrow{OQ} + \overrightarrow{PO} = \overrightarrow{PQ}.$$

For the last statement, apply this with P and O' in place of Q and P : $\theta_{O'}(P) = \overrightarrow{O'P} = \theta_O(P) - \theta_O(O')$. ■

So an n -dimensional affine space over F is a copy of F^n with the labels rubbed off. Choosing O writes the labels back on, and any two choices differ by a translation. We write $\mathbb{A}^n(F)$ for F^n regarded as an affine space over itself, and $\mathbb{A}^n$ when $F = \mathbb{R}$; its points are written

$P = (p_1, \dots, p_n)$, in plain type, even though the same list of numbers written $\mathbf{p}$ would be a vector.

Warning. **The identification is never canonical.** Proposition 22.1.3 does not say that $\mathcal{A}$ “is” V ; it says that each point of $\mathcal{A}$ gives one isomorphism, and there are as many as there are points. A statement proved after choosing O is a statement about $\mathcal{A}$ only if it does not depend on which O was chosen. Checking that independence is a real obligation, and the next subsection is nothing but that check, done once for all.

22.1.3 Affine combinations

Sums of points are forbidden, and yet the centroid $\frac{1}{3}(P_1 + P_2 + P_3)$ of a triangle is a genuine point of the plane, independent of any coordinate system. Something with weights adding up to 1 survives the loss of the origin. Here is the computation that says which expressions survive.

Fix $O \in \mathcal{A}$, points $P_1, \dots, P_k \in \mathcal{A}$ and scalars $t_1, \dots, t_k \in F$ with $s := t_1 + \dots + t_k$. Let $O' \in \mathcal{A}$ and $w = \overrightarrow{OO'}$. Since $\overrightarrow{O'P_i} = \overrightarrow{O'O} + \overrightarrow{OP_i} = \overrightarrow{OP_i} - w$,

$$\sum_{i=1}^k t_i \overrightarrow{O'P_i} = \sum_{i=1}^k t_i \overrightarrow{OP_i} - sw, \quad (22.1.1)$$

and therefore, adding this vector to the point $O' = O + w$ and using Lemma 22.1.1 (b), we get the companion identity for points.

$$O' + \sum_{i=1}^k t_i \overrightarrow{O'P_i} = \left(O + \sum_{i=1}^k t_i \overrightarrow{OP_i} \right) + (1-s)w. \quad (22.1.2)$$

The vector on the left of Eq. 22.1.1 does not depend on the origin exactly when $s = 0$; the point on the left of Eq. 22.1.2 does not depend on it exactly when $s = 1$. Both cases are worth a name.

Weights adding to one make a point; weights adding to zero make a vector; nothing else means anything.

Definition 22.1.2. Affine Combination

Let $\mathcal{A}$ be an affine space with direction space V , let $P_1, \dots, P_k \in \mathcal{A}$ with $k \geq 1$, and let $t_1, \dots, t_k \in F$.

- (a) If $t_1 + \dots + t_k = 1$, the **affine combination** $\sum_i t_i P_i$ is the **point** $O + \sum_i t_i \overrightarrow{OP_i}$, for any $O \in \mathcal{A}$.
- (b) If $t_1 + \dots + t_k = 0$, the expression $\sum_i t_i P_i$ denotes the **vector** $\sum_i t_i \overrightarrow{OP_i} \in V$, for any $O \in \mathcal{A}$.

For no other value of $t_1 + \dots + t_k$ is $\sum_i t_i P_i$ defined.

In words: the same symbol means a point when the weights add up to 1 and a vector when they add up to 0, and the type of the answer is read off from the weights. The simplest instance of

(b) is $k = 2$ with weights $(-1, 1)$: the vector $Q - P$ is $\overrightarrow{PQ}$. The simplest instance of (a) is $k = 1$ with $t_1 = 1$, which is P_1 .

Proposition 22.1.4. Affine combinations are exactly the origin-free ones

With the notation of Definition 22.1.2, put $s = t_1 + \cdots + t_k$.

- (a) If $s = 1$, the point $O + \sum_i t_i \overrightarrow{OP_i}$ is the same for every $O \in \mathcal{A}$.
- (b) If $s = 0$, the vector $\sum_i t_i \overrightarrow{OP_i}$ is the same for every $O \in \mathcal{A}$.
- (c) Suppose $\dim \mathcal{A} \geq 1$. If $s \neq 1$, there are points $P_1, \dots, P_k$ and origins O, O' for which $O + \sum_i t_i \overrightarrow{OP_i}$ and $O' + \sum_i t_i \overrightarrow{O'P_i}$ differ; and if $s \neq 0$, the same is true of the vectors $\sum_i t_i \overrightarrow{OP_i}$.

Proof.

- (a) and (b) are Eq. 22.1.2 and Eq. 22.1.1 with $1 - s = 0$ and $s = 0$ respectively, since O' was an arbitrary point of $\mathcal{A}$.
- (b) Since $\dim \mathcal{A} \geq 1$, the direction space V contains some $w \neq 0$. Take any $O \in \mathcal{A}$, put $O' = O + w$, and take $P_1 = \cdots = P_k = O$, so that $\overrightarrow{OP_i} = 0$ for every i . Then Eq. 22.1.2 reads

$$O' + \sum_i t_i \overrightarrow{O'P_i} = O + (1 - s)w,$$

which differs from $O + \sum_i t_i \overrightarrow{OP_i} = O$ when $(1 - s)w \neq 0$, that is, when $s \neq 1$, by Lemma 22.1.1 (c). Likewise Eq. 22.1.1 reads $\sum_i t_i \overrightarrow{O'P_i} = -sw$, which differs from $\sum_i t_i \overrightarrow{OP_i} = 0$ when $s \neq 0$. ■

Taking $\mathcal{A} = V$ over $\mathbb{R}$ and $O = 0$, part (a) becomes the ordinary vector expression $\sum_i t_i x_i$ with $\sum_i t_i = 1$, which is Chapter 18's affine combination (Definition 18.1.5 (a)). So Definition 22.1.2 extends that notion from a real vector space to an affine space over any field, and adds the companion vector-valued case. Chapter 18's affine hull $\text{aff } S$ and its description as a coset (Proposition 18.1.4) are the real case of the flats of the next section.

Quick check. Let $f: \mathbb{R} \rightarrow \mathbb{R}$ be $f(x) = 2x + 3$, which is not linear. Compute $f(3 \cdot 1 - 2 \cdot 2)$ and $3f(1) - 2f(2)$; then compute $f(1 + 2)$ and $f(1) + f(2)$. What do the two pairs show?

Solution. First pair: $3 \cdot 1 - 2 \cdot 2 = -1$ and $f(-1) = 1$, while $3f(1) - 2f(2) = 3 \cdot 5 - 2 \cdot 7 = 1$. They agree, and the weights 3 and -2 add up to 1. Second pair: $f(3) = 9$, while $f(1) + f(2) = 5 + 7 = 12$. They differ, and the weights 1 and 1 add up to 2. So f respects combinations whose weights add up to 1 and destroys the others: it is affine but not linear. That is the content of the next subsection.

22.1.4 Affine maps

Linear maps are the functions that respect the structure of a vector space. Affine maps should be the functions that respect the structure of an affine space, and by [Definition 22.1.2](#) that structure is: which affine combinations equal which points. So there is only one reasonable definition.

An affine map is one that can be applied inside an affine combination.

Definition 22.1.3. Affine Map

Let $\mathcal{A}$ and $\mathcal{B}$ be affine spaces over the same field F , with direction spaces V and W . A map $f: \mathcal{A} \rightarrow \mathcal{B}$ is **affine** if

$$f\left(\sum_{i=1}^k t_i P_i\right) = \sum_{i=1}^k t_i f(P_i)$$

for every $k \geq 1$, all $P_1, \dots, P_k \in \mathcal{A}$ and all $t_1, \dots, t_k \in F$ with $t_1 + \dots + t_k = 1$.

Both sides are points, by [Definition 22.1.2](#) (a) applied in $\mathcal{A}$ and in $\mathcal{B}$. Nothing is asked about weights summing to anything else, and by the Quick check above nothing could be.

The definition quantifies over all k , which is awkward to verify. The next theorem replaces it by a single linear map, and is the result the rest of the chapter uses.

Theorem 22.1.5. An Affine Map Is Linear Plus a Translation

Let $\mathcal{A}, \mathcal{B}$ be affine spaces over F with direction spaces V, W , and let $f: \mathcal{A} \rightarrow \mathcal{B}$. The following are equivalent.

- (a) f is affine.
- (b) There is a linear map $T: V \rightarrow W$ with $\overrightarrow{f(P)f(Q)} = T(\overrightarrow{PQ})$ for all $P, Q \in \mathcal{A}$.
- (c) For **some** $O \in \mathcal{A}$ there is a linear map $T: V \rightarrow W$ with $f(P) = f(O) + T(\overrightarrow{OP})$ for all $P \in \mathcal{A}$.

The map T in (b) and (c) is unique, is the same in both, and does not depend on O . It is called the **linear part** of f and written $\vec{f}$.

Idea. The cycle is (a) $\Rightarrow$ (c) $\Rightarrow$ (b) $\Rightarrow$ (a). For (a) $\Rightarrow$ (c) we must manufacture a linear map out of a condition about weights, so we need the two vector-space axioms written as affine combinations. They are

$$O + (u + w) = (O + u) + (O + w) - O, \quad O + \lambda u = \lambda(O + u) + (1 - \lambda)O,$$

whose weights add up to $1 - 1 + 1 = 1$ and $\lambda + (1 - \lambda) = 1$. Applying f to each is additivity and homogeneity of T .

Proof.

- (a) $\Rightarrow$ (c). Fix $O \in \mathcal{A}$ and define $T: V \rightarrow W$ by $T(v) = \overrightarrow{f(O)f(O+v)}$; this is a map because $O + v$ is a well-defined point. Then $f(O + v) = f(O) + T(v)$ by the definition

of $+$ in $\mathcal{B}$, so once T is shown to be linear, taking $v = \overrightarrow{OP}$ gives (c).

Additivity. Let $u, w \in V$. The three points $O + u$, $O + w$, O with weights $1, 1, -1$ have weight sum 1, and their affine combination, computed with origin O , is $O + (u + w - 0) = O + (u + w)$. Applying (a) and computing the right-hand side with origin $f(O)$,

$$\begin{aligned} f(O + u + w) &= f(O + u) + f(O + w) - f(O) \\ &= f(O) + (T(u) + T(w)), \end{aligned}$$

so $T(u + w) = T(u) + T(w)$.

Homogeneity. Let $\lambda \in F$. The two points $O + u$ and O with weights λ and $1 - \lambda$ have weight sum 1, and their affine combination with origin O is $O + \lambda u$. Applying (a) and computing with origin $f(O)$,

$$f(O + \lambda u) = \lambda f(O + u) + (1 - \lambda)f(O) = f(O) + \lambda T(u),$$

so $T(\lambda u) = \lambda T(u)$. Hence T is linear.

(c) $\Rightarrow$ (b). Let O and T be as in (c) and let $P, Q \in \mathcal{A}$. Then $f(P) = f(O) + T(\overrightarrow{OP})$ and $f(Q) = f(O) + T(\overrightarrow{OQ})$, so by [Proposition 22.1.3](#) applied in $\mathcal{B}$ with origin $f(O)$, and then linearity of T ,

$$\overrightarrow{f(P)f(Q)} = T(\overrightarrow{OQ}) - T(\overrightarrow{OP}) = T(\overrightarrow{OQ} - \overrightarrow{OP}) = T(\overrightarrow{PQ}),$$

the last equality by [Proposition 22.1.3](#) in $\mathcal{A}$.

(d) $\Rightarrow$ (a). Let T be as in (b), let $\sum_i t_i = 1$, and put $X = \sum_i t_i P_i$. Choose any $O \in \mathcal{A}$, so that $\overrightarrow{OX} = \sum_i t_i \overrightarrow{OP_i}$ by [Definition 22.1.2](#). Applying (b) to the pairs (O, X) and (O, P_i) , and then linearity,

$$\overrightarrow{f(O)f(X)} = T(\overrightarrow{OX}) = \sum_i t_i T(\overrightarrow{OP_i}) = \sum_i t_i \overrightarrow{f(O)f(P_i)}.$$

The right-hand side is, by [Definition 22.1.2](#) computed in $\mathcal{B}$ with origin $f(O)$, the displacement from $f(O)$ to $\sum_i t_i f(P_i)$. Hence $f(X) = \sum_i t_i f(P_i)$, which is (a).

Uniqueness. Suppose T and T' both satisfy (b). Given $v \in V$, fix O and put $Q = O + v$; then $T(v) = \overrightarrow{f(O)f(Q)} = T'(v)$. So T is determined by f , independently of any origin; and a T satisfying (c) for one O satisfies (b) by the implication above, hence is this same map. This proves the theorem. ■

So an affine map is *one linear map plus one displacement*, no more. In $\mathbb{A}^n(F) \rightarrow \mathbb{A}^m(F)$ it is

$$f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}, \quad \mathbf{A} \in M_{m \times n}(F), \quad \mathbf{b} \in F^m,$$

with $\mathbf{A}$ the matrix of $\vec{f}$ in the standard bases and $\mathbf{b} = f(\mathbf{0})$.

Warning. Affine is weaker than linear, and the difference is exactly the origin. An affine map $f: V \rightarrow W$ between *vector spaces* is linear if and only if $f(0) = 0$: if $f(x) = T(x) + b$ then $f(0) = b$, and f is linear exactly when $b = 0$. The translation $\tau(x) = x + b$ with $b \neq 0$ is affine with linear part id_V and is **not** linear — Chapter 2 recorded it as Example

(Translation), the standard nonexample of a linear map. Conversely $x \mapsto x^2$ on $\mathbb{R}$ fixes 0 and is not affine: it sends the affine combination $2 \cdot 1 + (-1) \cdot 0 = 1$ to 1, while $2f(1) - f(0) = 2$.

An affine map is determined by very little, but not by too little. Here is the negative half; the positive half is the subject of the next section.

Example. Two points are not enough in the plane

Let $f, g: \mathbb{A}^2 \rightarrow \mathbb{A}^2$ be $g(x, y) = (x, y)$ and $f(x, y) = (x + y, y)$. Both are affine, being linear. They agree at the two distinct points $(0, 0)$ and $(1, 0)$ — indeed at every point of the line $y = 0$ — yet $f(0, 1) = (1, 1) \neq (0, 1) = g(0, 1)$. So two points, even distinct ones, do not determine an affine map of the plane. The count to remember is $n + 1$: the next section proves that 3 suitably chosen points do determine an affine map of the plane, and that 2 never will.

22.1.5 The affine group, and the matrix that computes it

Affine maps compose, and the linear part is functorial. That single fact organizes every computation in the chapter.

Theorem 22.1.6. Composition, Inverses, and the Affine Group

Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be affine spaces over F , and write V and W for the direction spaces of $\mathcal{A}$ and $\mathcal{B}$.

- (a) If $f: \mathcal{A} \rightarrow \mathcal{B}$ and $g: \mathcal{B} \rightarrow \mathcal{C}$ are affine, then $g \circ f$ is affine and $\overrightarrow{g \circ f} = \vec{g} \circ \vec{f}$.
- (b) An affine $f: \mathcal{A} \rightarrow \mathcal{B}$ is bijective if and only if $\vec{f}$ is bijective, and then f^{-1} is affine with $\overrightarrow{f^{-1}} = (\vec{f})^{-1}$.
- (c) The bijective affine maps $\mathcal{A} \rightarrow \mathcal{A}$ form a group $\text{Aff}(\mathcal{A})$ under composition, and $f \mapsto \vec{f}$ is a surjective group homomorphism $\text{Aff}(\mathcal{A}) \rightarrow \text{GL}(V)$ whose kernel is the set of **translations** $\tau_v: P \mapsto P + v$, $v \in V$.

Idea. Everything is read off from the identity $\overrightarrow{f(P)f(Q)} = \vec{f}(\overrightarrow{PQ})$ of [Theorem 22.1.5](#) (b), which turns statements about points into statements about vectors. For (c), the kernel consists of the affine maps whose linear part is the identity, and such a map moves every point by the same vector.

Proof.

- (a) For $P, Q \in \mathcal{A}$, applying [Theorem 22.1.5](#) (b) twice,

$$\overrightarrow{g(f(P))g(f(Q))} = \vec{g}(\overrightarrow{f(P)f(Q)}) = \vec{g}(\vec{f}(\overrightarrow{PQ})).$$

Since $\vec{g} \circ \vec{f}$ is linear (Theorem 2.4.2), condition (b) of that theorem holds for $g \circ f$ with this map, so $g \circ f$ is affine and, by the uniqueness clause, $\overrightarrow{g \circ f} = \vec{g} \circ \vec{f}$.

- (b) ($\Leftarrow$) Suppose $\vec{f}$ is bijective. If $f(P) = f(Q)$ then $\vec{f}(\overrightarrow{PQ}) = \overrightarrow{f(P)f(Q)} = 0$, so $\overrightarrow{PQ} = 0$ and $P = Q$ by Lemma 22.1.1 (c); thus f is injective. For surjectivity, fix $O \in \mathcal{A}$ and let $R \in \mathcal{B}$. As $\vec{f}$ is surjective, $\overrightarrow{f(O)R} = \vec{f}(v)$ for some $v \in V$, and then $f(O + v) = f(O) + \vec{f}(v) = R$ by part (c) of Theorem 22.1.5.
- ($\Rightarrow$) Suppose f is bijective and fix $O \in \mathcal{A}$. If $\vec{f}(v) = 0$, then $f(O + v) = f(O) + 0 = f(O)$, so $O + v = O$ and $v = 0$; thus $\vec{f}$ is injective. Given $w \in W$, the point $f(O) + w$ is $f(P)$ for some P , and then $w = \overrightarrow{f(O)f(P)} = \vec{f}(\overrightarrow{OP})$; thus $\vec{f}$ is surjective.
- For the inverse, let $R, S \in \mathcal{B}$ and put $P = f^{-1}(R)$, $Q = f^{-1}(S)$. Then $\vec{f}(\overrightarrow{PQ}) = \overrightarrow{RS}$, so $f^{-1}(R)f^{-1}(S) = (\vec{f})^{-1}(\overrightarrow{RS})$. Since $(\vec{f})^{-1}$ is linear (Theorem 2.5.1), f^{-1} is affine with the stated linear part.
- (c) By (a) and (b), $\text{Aff}(\mathcal{A})$ is closed under composition and inverses and contains $\text{id}_{\mathcal{A}}$, whose linear part is id_V ; composition of functions is associative, so it is a group (Definition 0.10.1). By (a) the map $f \mapsto \vec{f}$ preserves composition, and by (b) it lands in $\text{GL}(V)$; so it is a homomorphism (Definition 0.10.3). It is surjective: given $T \in \text{GL}(V)$, fix O and set $f(P) = O + T(\overrightarrow{OP})$, which is affine with $\vec{f} = T$ by Theorem 22.1.5 (c) and bijective by (b).

Its kernel is $\{f : \vec{f} = \text{id}_V\}$. Each translation τ_v lies there, since $\overrightarrow{\tau_v(P)\tau_v(Q)} = \overrightarrow{PQ}$ by Chasles. Conversely suppose $\vec{f} = \text{id}_V$, fix O and put $v = \overrightarrow{Of(O)}$. For any P , Chasles gives

$$\overrightarrow{Pf(P)} = \overrightarrow{PO} + \overrightarrow{Of(O)} + \overrightarrow{f(O)f(P)} = \overrightarrow{PO} + v + \overrightarrow{OP} = v,$$

so $f(P) = P + v = \tau_v(P)$. This proves the theorem. ■

In coordinates the homomorphism of (c) is visible in the shape of a matrix. Identify $\mathbf{x} \in F^n$ with the column $\begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} \in F^{n+1}$ — the points of $\mathbb{A}^n(F)$ become the vectors of F^{n+1} whose last coordinate is 1.

Proposition 22.1.7. The Block Matrix of an Affine Map

Let $f: \mathbb{A}^n(F) \rightarrow \mathbb{A}^m(F)$ be affine, $f(\mathbf{x}) = \mathbf{Ax} + \mathbf{b}$ with $\mathbf{A} \in M_{m \times n}(F)$ the matrix of $\vec{f}$ and $\mathbf{b} = f(\mathbf{0})$. Put

$$\widehat{f} = \begin{pmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{0}^\top & 1 \end{pmatrix} \in M_{(m+1) \times (n+1)}(F).$$

Then $\widehat{f}\begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} = \begin{pmatrix} f(\mathbf{x}) \\ 1 \end{pmatrix}$ for every $\mathbf{x} \in F^n$; and for affine $f: \mathbb{A}^n(F) \rightarrow \mathbb{A}^m(F)$ and $g: \mathbb{A}^m(F) \rightarrow \mathbb{A}^p(F)$ one has $\widehat{g \circ f} = \widehat{g} \widehat{f}$. For $m = n$, f is bijective if and only if $\mathbf{A}$ is invertible, and then

$$\widehat{f^{-1}} = (\widehat{f})^{-1} = \begin{pmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{b} \\ \mathbf{0}^\top & 1 \end{pmatrix}.$$

Proof. Block multiplication (Theorem 8.1.1) gives

$$\begin{pmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{0}^\top & 1 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{Ax} + \mathbf{b} \\ 1 \end{pmatrix},$$

which is the first claim. Writing $g(\mathbf{y}) = \mathbf{C}\mathbf{y} + \mathbf{d}$, the same rule gives

$$\begin{pmatrix} \mathbf{C} & \mathbf{d} \\ \mathbf{0}^\top & 1 \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{0}^\top & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{CA} & \mathbf{Cb} + \mathbf{d} \\ \mathbf{0}^\top & 1 \end{pmatrix},$$

and $(g \circ f)(\mathbf{x}) = \mathbf{C}(\mathbf{Ax} + \mathbf{b}) + \mathbf{d} = (\mathbf{CA})\mathbf{x} + (\mathbf{Cb} + \mathbf{d})$, so the two agree. For $m = n$: f is bijective exactly when $\vec{f}$ is, by Theorem 22.1.6 (b), and $\vec{f}$ is bijective exactly when its matrix $\mathbf{A}$ is invertible, by Theorem 4.1.5 (b). In that case $\mathbf{x} \mapsto \mathbf{A}^{-1}\mathbf{x} - \mathbf{A}^{-1}\mathbf{b}$ is affine and is a two-sided inverse of f , and its block matrix is the one displayed, which is therefore $(\hat{f})^{-1}$ by the composition rule. ■

So composing affine maps is multiplying $(n+1) \times (n+1)$ matrices, and $\text{Aff}(\mathbb{A}^n(F))$ is a group of such matrices. The last row $(\mathbf{0}^\top \mid 1)$ is what records “these are points, not vectors”; the projective sections later in this chapter read that row again, as the equation of the hyperplane at infinity, after the extra coordinate is moved from last place to first.

Example. Composing a shear and a quarter turn

In $\mathbb{A}^2$ let $f(\mathbf{x}) = \mathbf{Ax} + \mathbf{b}$ and $g(\mathbf{y}) = \mathbf{Cy} + \mathbf{d}$, where

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{d} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

Compute $g \circ f$ in block form, and find f^{-1} .

Solution. The block matrices multiply as

$$\begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 3 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

so $(g \circ f)(x, y) = (-y + 3, x + 2y + 1)$. **Check directly:** $f(x, y) = (x + 2y + 1, y - 1)$, and applying g gives $-(y - 1) + 2, x + 2y + 1 = (3 - y, x + 2y + 1)$, the same.

For the inverse, $\mathbf{A}^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$ and $-\mathbf{A}^{-1}\mathbf{b} = -\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$, so $f^{-1}(x, y) = (x - 2y - 3, y + 1)$. **Check:** $f^{-1}(f(x, y)) = ((x + 2y + 1) - 2(y - 1) - 3, (y - 1) + 1) = (x, y)$.

Two things have been set up. Affine maps are linear maps that have forgotten where the origin is, and they are computed by one matrix of size one larger. The next section uses both to describe the subsets of an affine space that are themselves affine spaces.

22.1.6 Exercises

22.1.6.1 A. Check your understanding

Exercise. A1

- (a) State Chasles's relation, and deduce $\overrightarrow{QP} = -\overrightarrow{PQ}$ from it.
- (b) In an affine space over a field in which $2 \neq 0$ and $3 \neq 0$, which of $P + Q$, $3P$, $\overrightarrow{PQ} + \overrightarrow{RS}$, $P + \overrightarrow{QR}$, $\frac{1}{2}P + \frac{1}{2}Q$ are defined? For each defined one, say whether it is a point or a vector.
- (c) True or false, with a reason: a map $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ with $f(\mathbf{0}) = \mathbf{0}$ is affine if and only if it is linear.
- (d) An affine map $f: \mathbb{A}^3 \rightarrow \mathbb{A}^3$ has $\vec{f}$ of rank 2. Is f injective? Justify your answer.

Solution.

- (a) $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$ for all points P, Q, R . Taking $R = P$ gives $\overrightarrow{PQ} + \overrightarrow{QP} = \overrightarrow{PP}$, and $\overrightarrow{PP} = \mathbf{0}$ by Lemma 22.1.1 (a); hence $\overrightarrow{QP} = -\overrightarrow{PQ}$.
- (b) $P + Q$ and $3P$ are **not** defined: the weights add up to 2 and to 3, and under the hypotheses on F neither sum is 0 or 1, so Definition 22.1.2 gives these expressions no meaning; Proposition 22.1.4 (c) says why nothing could, since for $\dim \mathcal{A} \geq 1$ no such expression is origin-free. $\overrightarrow{PQ} + \overrightarrow{RS}$ is a vector (a sum in V). $P + \overrightarrow{QR}$ is a point. $\frac{1}{2}P + \frac{1}{2}Q$ is a point, the midpoint, and $\frac{1}{2}$ exists because $2 \neq 0$. (The hypotheses on F are needed: over $\mathbb{F}_2$ the weights of $P + Q$ add up to $2 = 0$, and Definition 22.1.2 (b) makes $P + Q$ the vector $\overrightarrow{OP} + \overrightarrow{OQ} = \overrightarrow{PQ}$.)
- (c) True. An affine map is $f(\mathbf{x}) = T(\mathbf{x}) + \mathbf{b}$ with T linear and $\mathbf{b} = f(\mathbf{0})$ (Theorem 22.1.5), so $f(\mathbf{0}) = \mathbf{0}$ forces $\mathbf{b} = \mathbf{0}$ and $f = T$. Conversely every linear map is affine, with $\vec{f} = f$, and sends $\mathbf{0}$ to $\mathbf{0}$.
- (d) No. By Theorem 22.1.6 (b), f is bijective exactly when $\vec{f}$ is, and for an operator on a finite-dimensional space injective, surjective, and bijective agree (Theorem 2.5.7). Here $\text{rank } \vec{f} = 2 < 3$, so $\vec{f}$ is not injective; taking $\mathbf{v} \neq \mathbf{0}$ in $\ker \vec{f}$ and any P , the points P and $P + \mathbf{v}$ have the same image, since $\overrightarrow{f(P)f(P+\mathbf{v})} = \vec{f}(\mathbf{v}) = \mathbf{0}$.

22.1.6.2 B. Practice
Exercise. B1: Determine which are affine

Determine which of the following maps are affine. Justify your answer; for each affine one give its linear part, and for each other one name a failing affine combination.

- (a) $f: \mathbb{A}^2 \rightarrow \mathbb{A}^2$, $f(x, y) = (2x - y + 1, 5)$.
- (b) $f: \mathbb{A}^2 \rightarrow \mathbb{A}^1$, $f(x, y) = xy$.
- (c) $f: \mathbb{A}^1 \rightarrow \mathbb{A}^2$, $f(t) = (1 - t, 2 + 3t)$.
- (d) $f: \mathbb{A}^2 \rightarrow \mathbb{A}^2$, $f(x, y) = (x + 1, |y|)$.

Solution.

- (a) Affine: $f(\mathbf{x}) = \mathbf{Ax} + \mathbf{b}$ with $\mathbf{A} = \begin{pmatrix} 2 & -1 \\ 0 & 0 \end{pmatrix}$ and $\mathbf{b} = (1, 5)$. Its linear part is $\mathbf{x} \mapsto \mathbf{Ax}$, of rank 1.

- (b) Not affine. Take the affine combination $2(1, 1) - 1 \cdot (0, 0) = (2, 2)$, with weights $2 + (-1) = 1$. Then $f(2, 2) = 4$, while $2f(1, 1) - f(0, 0) = 2$.
- (c) Affine: $f(t) = \mathbf{A}t + \mathbf{b}$ with $\mathbf{A} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$ and $\mathbf{b} = (1, 2)$. Linear part $t \mapsto (-t, 3t)$, of rank 1.
- (d) Not affine. Take the affine combination $2(0, 1) - 1 \cdot (0, 3) = (0, -1)$, with weights $2 + (-1) = 1$. Then $f(0, -1) = (1, 1)$, while $2f(0, 1) - f(0, 3) = 2(1, 1) - (1, 3) = (1, -1)$. The two differ, so f is not affine. (The first coordinate is affine in x ; it is $|y|$ that fails.)

Exercise. B2: Block matrices

In $\mathbb{A}^2$, let r be the affine map with $\hat{r} = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and let τ be the translation by $(2, 3)$.

- (a) Write r and τ in the form $\mathbf{x} \mapsto \mathbf{A}\mathbf{x} + \mathbf{b}$.
- (b) Compute $\widehat{\tau \circ r}$ and $\widehat{r \circ \tau}$, and say whether the two maps are equal.
- (c) Find all fixed points of r .

Solution.

- (a) $r(x, y) = (-y + 1, x)$, that is, $\mathbf{A}_r = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $\mathbf{b}_r = (1, 0)$. And $\tau(x, y) = (x + 2, y + 3)$, with $\mathbf{A}_\tau = \mathbf{I}_2$, $\mathbf{b}_\tau = (2, 3)$.
- (b) By [Proposition 22.1.7](#), $\widehat{\tau \circ r} = \hat{\tau} \hat{r}$ and $\widehat{r \circ \tau} = \hat{r} \hat{\tau}$:

$$\hat{\tau} \hat{r} = \begin{pmatrix} 0 & -1 & 3 \\ 1 & 0 & 3 \\ 0 & 0 & 1 \end{pmatrix}, \quad \hat{r} \hat{\tau} = \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

They differ, so $\tau \circ r \neq r \circ \tau$. (The linear parts agree; the translation vectors do not, because $\vec{r}$ moves $\mathbf{b}_\tau$.)

- (c) $r(x, y) = (x, y)$ reads $-y + 1 = x$ and $x = y$, so $x = y = \frac{1}{2}$. The unique fixed point is $(\frac{1}{2}, \frac{1}{2})$. Equivalently, $(\mathbf{I} - \mathbf{A}_r)\mathbf{x} = \mathbf{b}_r$ with $\mathbf{I} - \mathbf{A}_r = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ invertible.

Exercise. B3: The four-point plane

Let $\mathcal{A} = \mathbb{F}_2^2$ as in [Example \(Four affine spaces\)](#) (d), with points $O = (0, 0)$, $P = (1, 0)$, $Q = (0, 1)$, $R = (1, 1)$.

- (a) Compute $\overrightarrow{PQ}$ and $\overrightarrow{QP}$, and explain why they are equal.
- (b) How many affine combinations of the three points O, P, Q are there, and which points do they give? (A combination is counted by its list of weights.)
- (c) How many bijective affine maps $\mathcal{A} \rightarrow \mathcal{A}$ are there?

Solution.

- (a) $\overrightarrow{PQ} = Q - P = (0, 1) - (1, 0) = (-1, 1)$ and $\overrightarrow{QP} = P - Q = (1, 0) - (0, 1) = (1, -1)$. They are equal because $\overrightarrow{QP} = -\overrightarrow{PQ}$ (Lemma 22.1.1 (a)) and $-\mathbf{v} = \mathbf{v}$ for every $\mathbf{v}$ over $\mathbb{F}_2$.
- (b) A weight list $(t_0, t_1, t_2) \in \mathbb{F}_2^3$ with $t_0 + t_1 + t_2 = 1$ has an odd number of ones, so it is one of $(1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 1)$: four lists. The first three give O, P, Q , and $(1, 1, 1)$ gives $O + P + Q = (1, 1) = R$. So the four affine combinations give the four points, each exactly once.
- (c) By Theorem 22.1.6 (c) an affine bijection is a translation composed with a linear one, and the number of pairs $(T, \mathbf{v}) \in \text{GL}(\mathbb{F}_2^2) \times \mathbb{F}_2^2$ is $|\text{GL}_2(\mathbb{F}_2)| \cdot 4$. A matrix in $\text{GL}_2(\mathbb{F}_2)$ has a nonzero first column (3 choices) and a second column outside the span of the first ($4 - 2 = 2$ choices), so $|\text{GL}_2(\mathbb{F}_2)| = 6$ and there are 24 affine bijections. Since $|\mathcal{A}| = 4$, these are **all** $4! = 24$ permutations of the four points.

22.1.6.3 C. Going deeper

Exercise. C1: Fixed points

Let $f: \mathbb{A}^n \rightarrow \mathbb{A}^n$ be affine, $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$.

- (a) Prove that f has a fixed point if and only if $\mathbf{b} \in \text{col}(\mathbf{I} - \mathbf{A})$, and that the set of fixed points, when nonempty, is a coset of $\text{nul}(\mathbf{I} - \mathbf{A})$.
- (b) Deduce that if 1 is not an eigenvalue of $\mathbf{A}$, then f has exactly one fixed point.
- (c) Give an affine map of $\mathbb{A}^2$ with no fixed point and linear part $\neq \text{id}$.

Solution.

- (a) $f(\mathbf{x}) = \mathbf{x}$ says $\mathbf{A}\mathbf{x} + \mathbf{b} = \mathbf{x}$, that is, $(\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{b}$. This system is consistent exactly when $\mathbf{b} \in \text{col}(\mathbf{I} - \mathbf{A})$ (Theorem 3.1.1), and then its solution set is $\mathbf{p} + \text{nul}(\mathbf{I} - \mathbf{A})$ for any one solution $\mathbf{p}$, by Theorem 3.1.3.
- (b) If 1 is not an eigenvalue of $\mathbf{A}$, then $\text{nul}(\mathbf{I} - \mathbf{A}) = \{\mathbf{0}\}$, so $\mathbf{I} - \mathbf{A}$ is invertible (Theorem 3.4.4). Hence $(\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}$.
- (c) Take $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ and $\mathbf{b} = (1, 0)$, so $f(x, y) = (x + 1, 2y)$. Then $\mathbf{I} - \mathbf{A} = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$ and $\text{col}(\mathbf{I} - \mathbf{A}) = \text{span}(\mathbf{e}_2)$, which does not contain $\mathbf{b} = \mathbf{e}_1$; so there is no fixed point, by (a). And $\mathbf{A} \neq \mathbf{I}$.

Exercise. C2: Affine maps on two points, over two fields

Let $f: \mathcal{A} \rightarrow \mathcal{B}$ be a map of affine spaces over F that preserves affine combinations of **two** points, that is, $f((1-t)P + tQ) = (1-t)f(P) + tf(Q)$ for all $P, Q \in \mathcal{A}$ and $t \in F$.

- (a) Prove that if $\text{char } F \neq 2$, then f is affine. *Hint: run the proof of Theorem 22.1.5, replacing the three-point combination by two two-point ones.*
- (b) Show that over $\mathbb{F}_2$ the hypothesis is empty, and give a nonaffine map $\mathbb{F}_2^2 \rightarrow \mathbb{F}_2^2$ satisfying it.

Solution.

- (a) Fix $O \in \mathcal{A}$ and define $T(v) = \overrightarrow{f(O)f(O+v)}$ as in [Theorem 22.1.5](#). Homogeneity is the hypothesis with $P = O + v$, $Q = O$: $O + \lambda v = \lambda(O + v) + (1 - \lambda)O$, so $f(O + \lambda v) = \lambda f(O + v) + (1 - \lambda)f(O) = f(O) + \lambda T(v)$, giving $T(\lambda v) = \lambda T(v)$. For additivity, let $u, w \in V$ and let $M = \frac{1}{2}(O + u) + \frac{1}{2}(O + w)$, which is legal since $2 \neq 0$ in F . With origin O , $M = O + \frac{1}{2}(u + w)$. By the hypothesis, $f(M) = \frac{1}{2}f(O + u) + \frac{1}{2}f(O + w)$, which with origin $f(O)$ reads $f(M) = f(O) + \frac{1}{2}(T(u) + T(w))$. So $T(\frac{1}{2}(u + w)) = \frac{1}{2}(T(u) + T(w))$. Applying homogeneity with $\lambda = 2$ to the left-hand side gives $T(u + w) = T(u) + T(w)$. Hence T is linear, and $f(P) = f(O) + T(\overrightarrow{OP})$ by construction, so f is affine by [Theorem 22.1.5](#) (c).
- (b) Over $\mathbb{F}_2$ the only scalars are $t = 0$ and $t = 1$, for which $(1 - t)P + tQ$ is P or Q ; the required identity then reads $f(P) = f(P)$ or $f(Q) = f(Q)$. So every map satisfies the hypothesis. Take $f: \mathbb{F}_2^2 \rightarrow \mathbb{F}_2$ with $f(0, 0) = f(1, 0) = f(0, 1) = 0$ and $f(1, 1) = 1$. It is not affine: by [Exercise \(B3: The four-point plane\)](#) (b), $(1, 1)$ is the affine combination $O + P + Q$ of the other three points with weights $(1, 1, 1)$, so an affine f would have $f(1, 1) = f(O) + f(P) + f(Q) = 0$.

Exercise. C3: Conjugating a translation

Let $\mathcal{A}$ be an affine space with direction space V , let $f \in \text{Aff}(\mathcal{A})$ and let $v \in V$.

- (a) Prove that $f \circ \tau_v \circ f^{-1} = \tau_{\vec{f}(v)}$.
- (b) Hence prove that the translations form a subgroup of $\text{Aff}(\mathcal{A})$ that is carried to itself by every conjugation, and that $v \mapsto \tau_v$ is an isomorphism from $(V, +)$ onto it.
- (c) Deduce that if $\dim \mathcal{A} \geq 2$, then $\text{Aff}(\mathcal{A})$ is not abelian.

Solution.

- (a) Both sides are affine maps $\mathcal{A} \rightarrow \mathcal{A}$. By [Theorem 22.1.6](#) (a) the linear part of the left-hand side is $\vec{f} \circ \text{id}_V \circ (\vec{f})^{-1} = \text{id}_V$, so by the kernel description in [Theorem 22.1.6](#) (c) it is a translation τ_u for some u . To identify u , fix $O \in \mathcal{A}$ and put $P = f^{-1}(O)$. Then

$$u = \overrightarrow{O(f \circ \tau_v \circ f^{-1})(O)} = \overrightarrow{f(P)f(P+v)} = \vec{f}(v),$$

using [Theorem 22.1.5](#) (b) and $\overrightarrow{P(P+v)} = v$.

- (b) The translations are the kernel of $f \mapsto \vec{f}$ ([Theorem 22.1.6](#) (c)), hence a subgroup by [Theorem 0.10.3](#), and (a) says every conjugate of a translation is a translation. The map $v \mapsto \tau_v$ satisfies $\tau_u \circ \tau_w = \tau_{u+w}$ by [Lemma 22.1.1](#) (b), so it is a homomorphism $(V, +) \rightarrow \text{Aff}(\mathcal{A})$; it is surjective onto the translations by definition, and injective because $\tau_v = \text{id}$ forces $P + v = P$, hence $v = 0$.
- (c) Since $\dim V \geq 2$, pick independent $v, w \in V$ and $T \in \text{GL}(V)$ with $T(v) = w$ (extend $\{v, w\}$ to a basis of V by [Theorem 1.8.2](#), and map that basis to itself with v and w swapped, using [Theorem 2.1.6](#); this T is invertible because $T \circ T$ fixes every basis vector and so is id_V , by the uniqueness in that theorem). By [Theorem 22.1.6](#) (c) there is $f \in \text{Aff}(\mathcal{A})$ with $\vec{f} = T$. If $\text{Aff}(\mathcal{A})$ were abelian then $f \circ \tau_v \circ f^{-1} = \tau_v$, while (a) gives $\tau_{T(v)} = \tau_w$; by the injectivity in (b) that forces $v = w$, contradicting independence.