CHAPTER 1

MATRICES & TRANSFORMATIONS

LUKE COLLINS
maths.com.mt

FEBRUARY MMXXI

1 Linear Transformations

Recall that all pairs (x, y) of real numbers are regarded as points in the xy-plane, where the set of all such points is denoted by

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}.$$

Here we will interpret the pair (x, y) in two ways: sometimes as the point (x, y) in the plane just as before, which we will call the *position* (x, y); other times as the *directed line segment* taking us from the *origin* $(0, 0)$ to the point (x, y), which we call the *vector* (x, y).

The distinction between the two interpretations is rarely important, and whenever the distinction is important, it is often clear from the context.

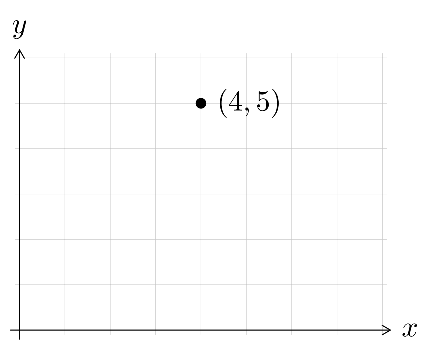

FIGURE 1: The position $(4, 5)$

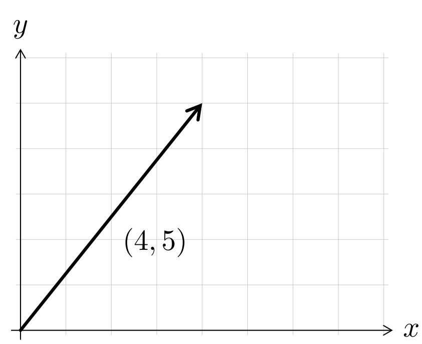

FIGURE 2: The vector $(4, 5)$

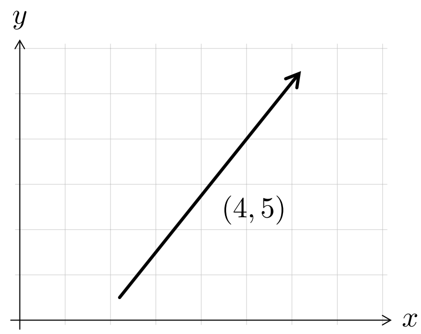

FIGURE 3: Still the vector $(4, 5)$

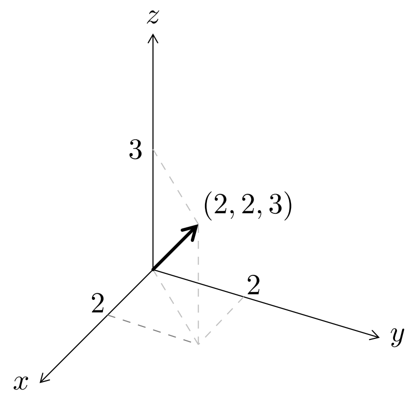

FIGURE 4: A vector in $\mathbb{R}^3$

When vectors are translated in the plane (that is, when their tails do not sit at the origin $(0, 0)$), they still correspond to the same pair of coordinates (x, y), since what the pair of numbers represent in this case is the *displacement* or *movement* from the tip of the arrow to its head. Thus if a vector is translated, we treat the tip as the "new origin", and read off the coordinates at the head of the arrow, thus obtaining the same pair (x, y). This pair of numbers is not telling you the coordinates of the position of the arrowhead, but rather, by how much you need to move in the x- and y-directions to go from the tail of the vector to its tip.

These ideas easily extend to the ordered triples (x, y, z) of real numbers, corresponding to points or vectors in three dimensional space

$$\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$$

and beyond, but for the scope of the MATSEC intermediate, we will focus mainly on two dimensional space.

1.1 Vector Operations

Let us start by introducing two important operations on vectors. We will be denoting vectors using single letters in bold typeface, such as $\boldsymbol{v} = (x, y)$ for example. In writing, you are encouraged to underline vectors to distinguish them from numbers, e.g., writing $\underline{v}$, for $\boldsymbol{v}$.

Definition 1.1 (Vector Addition)**.** Let $\boldsymbol{u} = (x_1, y_1)$ and $\boldsymbol{v} = (x_2, y_2)$ be two vectors in $\mathbb{R}^2$. Then the *sum* $\boldsymbol{u} + \boldsymbol{v}$ is defined by

$$\boldsymbol{u} + \boldsymbol{v} \stackrel{\text{def}}{=} (x_1 + x_2, y_1 + y_2).$$

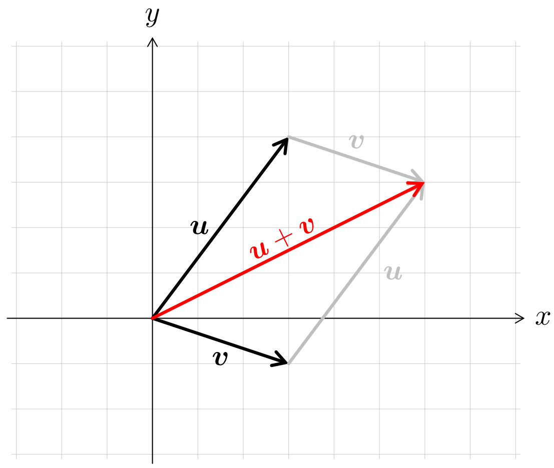

FIGURE 5: Illustration of the parallelogram law in $\mathbb{R}^2$

Example 1.2. If $\boldsymbol{u} = (3, 4)$ and $\boldsymbol{v} = (3, -1)$, then

$$\boldsymbol{u} + \boldsymbol{v} = (3 + 3, 4 + (-1)) = (6, 3).$$

Remark 1.3. Observe that the vector sum $\boldsymbol{u} + \boldsymbol{v}$ corresponds to the position obtained when translating the vector $\boldsymbol{v}$ such that its tail is at the head of the vector $\boldsymbol{u}$, or vice-versa; as shown in figure 5. This is a consequence of the fact that, as we've already mentioned, we think of vectors away from the origin as representing *movement*, and not position. This way, $\boldsymbol{u} + \boldsymbol{v}$ is the vector taking us to where we end up if we move along $\boldsymbol{u}$, and then carry out the displacement represented by the vector $\boldsymbol{v}$. Because of this behaviour, addition is sometimes referred to as *the parallelogram law*.

Notice also that $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$, i.e., vector addition is *commutative*. This is obvious from the definition since $x_1 + x_2 = x_2 + x_1$ and similarly for y_1 and y_2, but it is good to ponder the geometric meaning of $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$: it doesn't matter if we start from $\boldsymbol{u}$ or from $\boldsymbol{v}$ (see figure 5).

Definition 1.4 (Scalar Multiplication)**.** Let $\lambda \in \mathbb{R}$, and let $\boldsymbol{v} = (x, y)$ be a vector in $\mathbb{R}^2$. Then the *scalar multiplication* of $\boldsymbol{v}$ by λ, denoted $\lambda\boldsymbol{v}$, is the vector given by

$$\lambda\boldsymbol{v} \stackrel{\text{def}}{=} (\lambda x, \lambda y).$$

Example 1.5. If $\boldsymbol{u} = (1, 2)$, then $5\boldsymbol{u} = (5, 10)$ and $(-3)\boldsymbol{u} = (-3, -6)$.

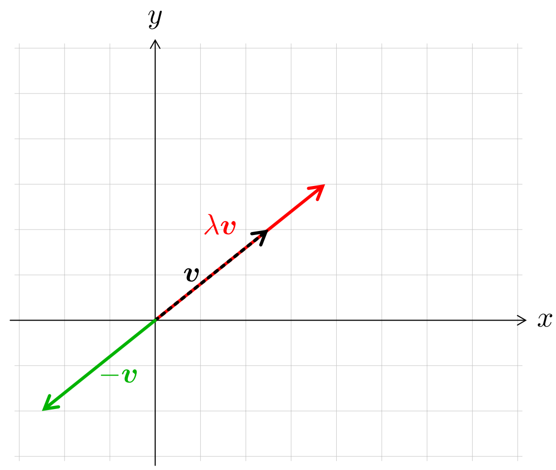

FIGURE 6: Illustration of scaling in $\mathbb{R}^2$

Remark 1.6. The reason we call this operation *scalar* multiplication is that the result of $\lambda\boldsymbol{v}$ is a *scaled* version of $\boldsymbol{v}$ by a factor of λ (see figure 6). When $\lambda < 0$, then the direction of $\boldsymbol{v}$ is reversed. In particular, $-1\boldsymbol{v}$, which we denote by $-\boldsymbol{v}$, corresponds to the vector with the arrow head and tail interchanged.

As a consequence of this scaling behaviour, we call single real numbers *scalars* instead of numbers throughout this chapter. Thus the entries in a vector are scalars, for example.

Notation. As mentioned in remark 1.6, we denote $-1\boldsymbol{v}$ by $-\boldsymbol{v}$, and we also introduce the *difference* between two vectors, denoted $\boldsymbol{u} - \boldsymbol{v}$, defined by

$$\boldsymbol{u} - \boldsymbol{v} = \boldsymbol{u} + (-\boldsymbol{v}).$$

This ends up being the same as subtraction componentwise.

Example 1.7. If $\boldsymbol{u} = (1, 2)$ and $\boldsymbol{v} = (3, -5)$, then

$$\boldsymbol{u} - \boldsymbol{v} = (1 - 3, 2 - (-5)) = (-2, 3).$$

Definition 1.8 (Zero vector)**.** We denote the vector $(0, 0)$ by $\mathbf{0}$ and call it the *zero vector* or the *origin*.

Note. $\mathbf{0} \neq 0$. One is a vector, the other is a scalar.

The two operations of addition and scalar multiplication turn $\mathbb{R}^2$ into a structure we call a *vector space*. Vector spaces are characterised by having the following properties.

Theorem 1.9 (Vector space properties in $\mathbb{R}^2$)**.** *Let* $\boldsymbol{u}$*,* $\boldsymbol{v}$*,* $\boldsymbol{w}$ *be three vectors in* $\mathbb{R}^2$*, and let* $\lambda, \mu \in \mathbb{R}$ *be scalars. Then the following properties hold:*

I) $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$

II) $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$

III) $\boldsymbol{u} + \mathbf{0} = \boldsymbol{u}$

IV) $\boldsymbol{v} + (-\boldsymbol{v}) = \mathbf{0}$

V) $\lambda(\mu\boldsymbol{v}) = (\lambda\mu)\boldsymbol{v}$

VI) $1\boldsymbol{v} = \boldsymbol{v}$

VII) $\lambda(\boldsymbol{u} + \boldsymbol{v}) = \lambda\boldsymbol{u} + \lambda\boldsymbol{v}$

VIII) $0\boldsymbol{v} = \mathbf{0}$

IX) $(\lambda + \mu)\boldsymbol{v} = \lambda\boldsymbol{v} + \mu\boldsymbol{v}$

Proof. These results all easily follow from the definitions, and properties inherited from real numbers, e.g. for I, if we write $\boldsymbol{u} = (x_1, y_1)$, $\boldsymbol{v} = (x_2, y_2)$ and $\boldsymbol{w} = (x_3, y_3)$, then

$$\begin{aligned}
\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) &= (x_1, y_1) + ((x_2, y_2) + (x_3, y_3)) \\
&= (x_1, y_1) + (x_2 + x_3, y_2 + y_3) && \text{(by definition 1.1)} \\
&= (x_1 + (x_2 + x_3), y_1 + (y_2 + y_3)) && \text{(by definition 1.1)} \\
&= ((x_1 + x_2) + x_3, (y_1 + y_2) + y_3) && \text{(ordinary + in } \mathbb{R}) \\
&= (x_1 + x_2, y_1 + y_2) + (x_3, y_3) && \text{(by definition 1.1)} \\
&= ((x_1, y_1) + (x_2, y_2)) + (x_3, y_3) && \text{(by definition 1.1)} \\
&= (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w},
\end{aligned}$$

as required. Similarly for VIII, we have

$$0\boldsymbol{v} = (0x_2, 0y_2) = (0, 0) = \mathbf{0}.$$

The proofs of the remaining properties are left as an exercise. □

Remark 1.10. Notice that a lot of these facts seem obvious, but this is because we are using the same conventions (namely the symbol + and juxtaposition) which we usually use for addition and multiplication of numbers. If we denote vector addition using a different symbol, say $\boldsymbol{u} \oplus \boldsymbol{v}$, and scalar multiplication using $\lambda \odot \boldsymbol{v}$, then it is clearer that what we are saying requires proof; e.g., V becomes

$$\lambda \odot (\mu \odot \boldsymbol{v}) = (\lambda\mu) \odot \boldsymbol{v}$$

and IX is

$$(\lambda + \mu) \odot \boldsymbol{v} = (\lambda \odot \boldsymbol{v}) \oplus (\mu \odot \boldsymbol{v}).$$

Mathematicians like to use the same symbol for various different things (this is usually called *overloading* the symbol). We tend to do this when the two ideas that the symbols represent share a number of features. For instance, even though $+$ for numbers and $+$ for vectors are technically distinct operations, they do have various things in common (theorem 1.9 is basically a list of those things), so it makes sense to use it for vectors. Also, if we write $1+3$ and $(3,2)+(1,6)$, you can't really mistake one for the other, it's obvious which of the two operations we need to use in either case.

Exercise 1.11. 1. Watch: https://youtu.be/fNk_zzaMoSs.

2. Let $\boldsymbol{u} = (1,2)$, $\boldsymbol{v} = (3,-5)$ and $\boldsymbol{w} = (-1,-1)$.
   (a) Work out $\boldsymbol{u}+\boldsymbol{v}$ and $3\boldsymbol{u} - 5\boldsymbol{v} + \frac{1}{8}\boldsymbol{w}$.
   (b) Draw a diagram which illustrates that $\boldsymbol{u}+\boldsymbol{v} = \boldsymbol{v}+\boldsymbol{u}$.
   (c) Determine constants a and b such that $a\boldsymbol{u} + b\boldsymbol{v} = \boldsymbol{w}$.
   (d) Determine constants x and y such that $3(x,y) - 5\boldsymbol{u} = 2\boldsymbol{w}$.
3. Complete the proof of theorem 1.9.
4. Let $\boldsymbol{u} = (x_1, y_1)$ and $\boldsymbol{v} = (x_2, y_2)$, and define their *dot product* $\boldsymbol{u} \cdot \boldsymbol{v}$ by $\boldsymbol{u} \cdot \boldsymbol{v} \stackrel{\text{def}}{=} x_1 x_2 + y_1 y_2$.
   (a) Prove that $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$.
   (b) Prove that $\boldsymbol{u}\cdot(\boldsymbol{v}+\boldsymbol{w}) = \boldsymbol{u}\cdot\boldsymbol{v}+\boldsymbol{u}\cdot\boldsymbol{w}$ and $\boldsymbol{u}\cdot(\lambda\boldsymbol{v}) = \lambda(\boldsymbol{u}\cdot\boldsymbol{v})$.
   (c) (Cauchy–Schwarz inequality). Prove that
   $$(\boldsymbol{u}\cdot\boldsymbol{v})^2 \leqslant (\boldsymbol{u}\cdot\boldsymbol{u})(\boldsymbol{v}\cdot\boldsymbol{v}).$$
   (d) By choosing appropriate vectors in the inequality in (c), prove that for any $a, b \in \mathbb{R}$, we have $a + b \leqslant \sqrt{2}\sqrt{a^2+b^2}$. Hence or otherwise, deduce that for any $x, y > 0$,
   $$\sqrt{\frac{x}{x+y}} + \sqrt{\frac{y}{x+y}} \leqslant \sqrt{2}.$$
   Is this bound sharp? (i.e., is there an assignment of x and y for which we have equality instead of $\leqslant$?)

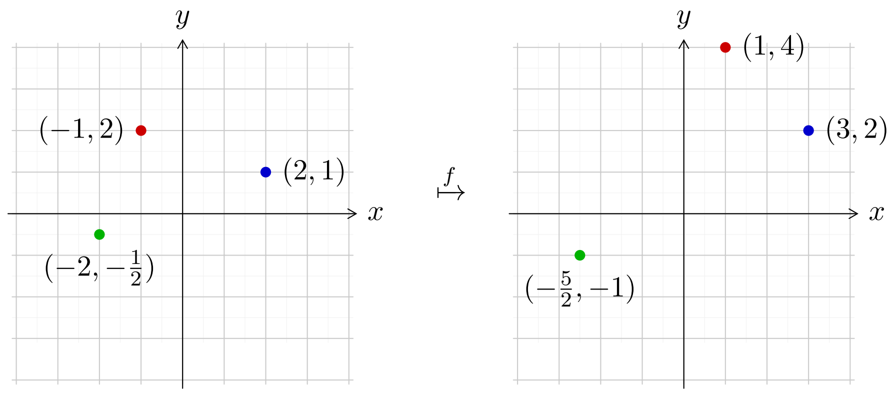

FIGURE 7: The function $f\colon (x,y) \mapsto (x+y, 2y)$ applied to some points in $\mathbb{R}^2$

1.2 Linear Transformations

A transformation is simply another word for a function. In this chapter, we will deal with functions from $\mathbb{R}^2$ to $\mathbb{R}^2$, i.e., functions whose inputs are points in $\mathbb{R}^2$, and whose outputs are also points in $\mathbb{R}^2$. For example,

$$f(x,y) = (x+y, 2y)$$

is such a transformation. See how it is applied to some points in the plane in figure 7. If we use vector notation, writing $\boldsymbol{u} = (x,y)$, then we also write $f(\boldsymbol{u})$ for $f(x,y)$.

The use of the word *transformation* is indicative of the fact that we like to think of f "moving" points from their starting position to their destination (see figure 8).

Another nice way to think of a transformations is as a process which does something to the plane as a whole, as opposed to individual points. We can try to visualise this by applying the transformation to the grid lines in our diagrams. In other words, if we transform *each point* which makes up the grid lines, we get new grid lines which should give us a better picture of what f "does" to the whole plane. See figure 9, as well as and this interactive web page:

`https://www.desmos.com/calculator/8etkl7jm0r`.

Notice that the transformed points retain their old coordinates with respect to these new grid lines. In other words, if we denote by $[x,y]$ the point obtained by travelling along the new grid lines rather than the usual ones, then the image of $(2,1)$ under f is $[2,1]$ (for instance), see figure 10.

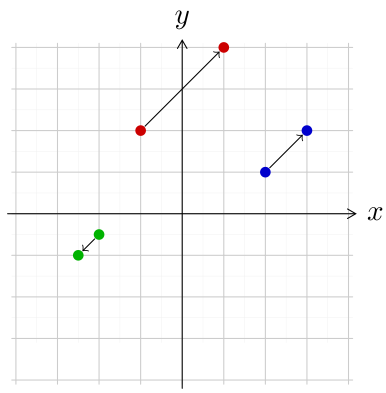

FIGURE 8: The transformation f visualised as “moving” points in space

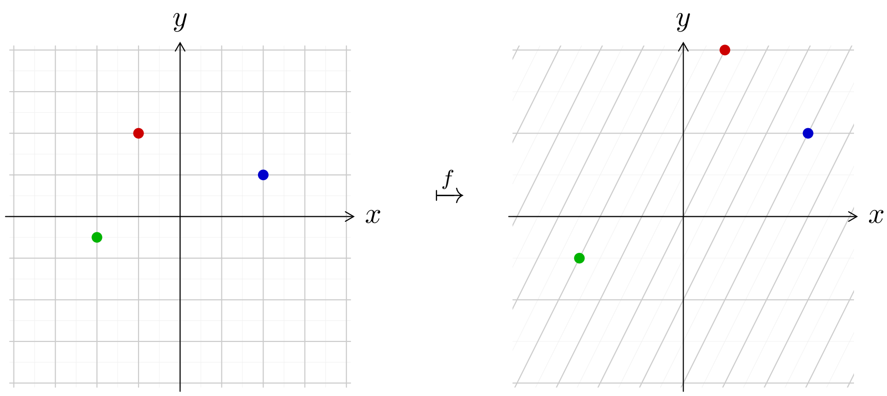

FIGURE 9: The function $f\colon (x, y) \mapsto (x + y, 2y)$ applied to the grid lines

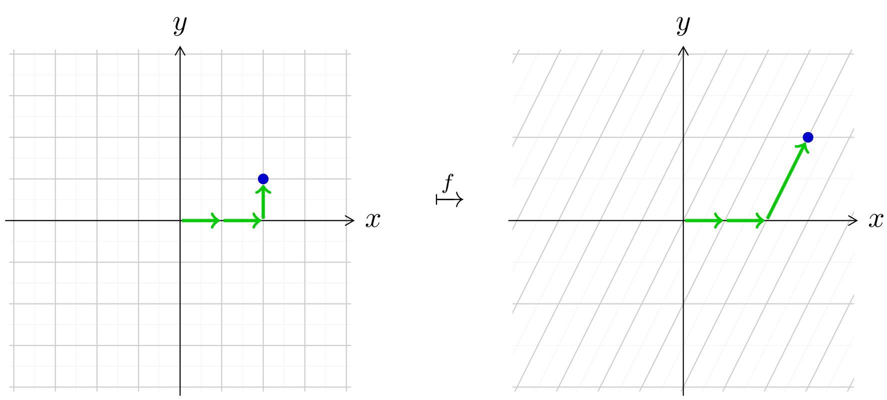

FIGURE 10: $(2, 1)$ and its image $[2, 1]$

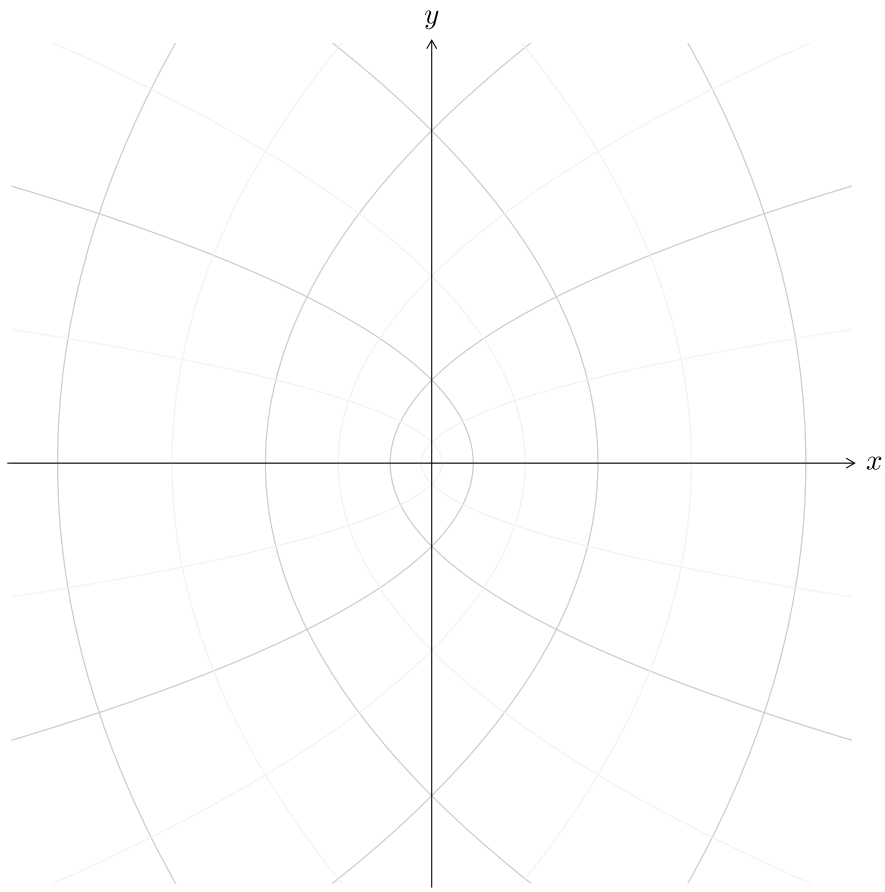

FIGURE 11: The image of the plane under the z^2 transformation

It is not necessarily the case that grid lines remain straight. For instance, a well known transformation is

$$g(x, y) = (x^2 - y^2, 2xy)$$

(this is called the z^2 transformation). When this is applied, the plane ends up looking like what's in figure 11. To get a better idea, it would be helpful to look at the online version to see the transformation animated:

https://www.desmos.com/calculator/bd9c1z6bll.

Notice that it is still true that the image of $(1, 2)$ (say) ends up at $[1, 2]$, where we instead have to travel along the curvy grid lines rather than the old ones. These sorts of transformations are complicated to study generally:

we will focus our attention on transformations where the grid lines remain parallel, evenly spaced lines. These are called *linear transformations*.

We will not give this as *the* definition of "linear" though, we will instead give a definition which it is easier to do maths directly with (it will be set to you as an exercise to check that the definition we give is equivalent to "lines remain lines" at a later stage). Here is our (and any standard textbook's) definition of linear:

Definition 1.12 (Linear)**.** Let $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ be a transformation. Then f is said to be *linear* if for all vectors $\boldsymbol{u}, \boldsymbol{v}$ and scalars λ, we have

$$f(\boldsymbol{u}+\boldsymbol{v}) = f(\boldsymbol{u}) + f(\boldsymbol{v}) \qquad \text{and} \qquad f(\lambda \boldsymbol{u}) = \lambda\, f(\boldsymbol{u}).$$

Our initial example $f(x,y) = (x+y, 2y)$ is an example of a linear transformation (indeed, lines remained lines when we applied it to the plane as a whole in figure 9). Let us verify that it satisfies the definition. If $\boldsymbol{u} = (x_1, y_1)$ and $\boldsymbol{v} = (x_2, y_2)$, then

$$\begin{aligned} f(\boldsymbol{u}+\boldsymbol{v}) &= f((x_1,y_1)+(x_2,y_2)) \\ &= f(x_1+x_2, y_1+y_2) \\ &= \big((x_1+x_2)+(y_1+y_2), 2(y_1+y_2)\big) \\ &= (x_1+y_1+x_2+y_2, 2y_1+2y_2) \\ &= (x_1+y_1, 2y_1) + (x_2+y_2, 2y_2) \\ &= f(\boldsymbol{u}) + f(\boldsymbol{v}), \end{aligned}$$

and

$$f(\lambda \boldsymbol{u}) = f(\lambda x_1, \lambda y_1) = (\lambda x_1 + \lambda y_1, 2(\lambda y_1)) = \lambda(x_1+y_1, 2y_1) = \lambda\, f(\boldsymbol{u}),$$

which proves that f complies with the definition.

By contrast, the z^2 transformation has

$$g(1,1) = (0,2) \qquad \text{and} \qquad g(2,2) = (0,8).$$

Since $(2,2) = 2(1,1)$, it should be the case (if g is linear) that

$$g(2,2) = g(2(1,1)) \overset{\text{linearity}}{=\!=\!=} 2\,g(1,1) = 2(0,2) = (0,4),$$

which is false. This shows that g is not linear.

Before we give some more examples, let us prove a very important fact about linear transformations.

Theorem 1.13. *Linear transformations do not move the origin.*

Proof. Observe that $0\mathbf{0} = 0(0,0) = \mathbf{0}$. Moreover, $0\boldsymbol{v} = \mathbf{0}$ for any vector $\boldsymbol{v}$ (theorem 1.9, VIII). Thus, for any linear transformation f, we have

$$f(\mathbf{0}) = f(0\mathbf{0}) = 0\, f(\mathbf{0}) = \mathbf{0},$$

i.e., $f(\mathbf{0}) = \mathbf{0}$. □

In particular, this implies that translations, i.e., transformations of the form

$$f(\boldsymbol{v}) = \boldsymbol{v} + \boldsymbol{t}$$

where $\boldsymbol{t}$ is some non-zero vector, are not linear transformations.

Example 1.14. For instance, the transformation $f(\boldsymbol{v}) = \boldsymbol{v} + (1,2)$ or written differently, $f(x,y) = (x+1, y+2)$, is not linear.

On the other hand, rotations (about the origin) and reflections in straight lines (which go through the origin) are linear transformations.[1] We will see that this is the case in a moment. But before we continue, let us introduce some notation.

Notation. For reasons which will become clear later, we will also write points/vectors as columns, i.e., we will write $\boldsymbol{v} = (x,y)$ as $\binom{x}{y}$. This means we have three different ways to write $f(\boldsymbol{v})$, namely, as

$$f(\boldsymbol{v}), \qquad f(x,y) \qquad \text{or} \qquad f\begin{pmatrix} x \\ y \end{pmatrix}.$$

Make sure you feel comfortable with all three.

Exercise 1.15.

1. (a) Prove that scaling, i.e., $f(\boldsymbol{v}) = \lambda\boldsymbol{v}$, is a linear transformation for all $\lambda \in \mathbb{R}$.
   (b) When is a constant function $f(\boldsymbol{v}) = \boldsymbol{u}$ linear? ($\boldsymbol{u}$ is a fixed vector here.)
2. (a) Prove that $f(x,y) = (3x+2y, 7x-8y)$ is linear.
   (b) Prove that $f(x,y) = (3x+2, 7x-8y)$ is not linear.

[1] Notice we need the conditions in brackets, otherwise the origin wouldn't stay fixed by the corresponding transformations.

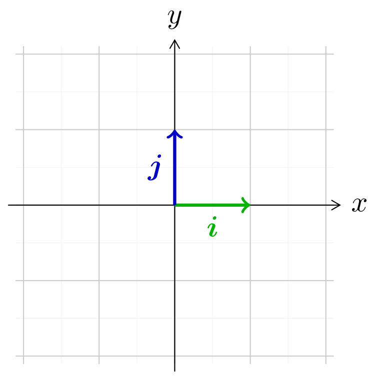

FIGURE 12: Basis vectors $\boldsymbol{i}$ and $\boldsymbol{j}$ for $\mathbb{R}^2$

1.3 Basis Vectors and Matrices

Let $\boldsymbol{i} = (1, 0)$ and $\boldsymbol{j} = (0, 1)$, we call these the *basis vectors* for $\mathbb{R}^2$. Notice that any point (x, y) can be written as

$$(x, y) = x\boldsymbol{i} + y\boldsymbol{j}$$

since $x\boldsymbol{i} + y\boldsymbol{j} = x(1, 0) + y(0, 1) = (x, 0) + (0, y) = (x, y)$. This seems like a trivial observation, essentially $\boldsymbol{i}$ represents a single unit of motion in the x-direction, and $\boldsymbol{j}$ represents a single unit of motion in the y-direction. Thus, all we're saying is that

$$(x, y) = x \times (1 \text{ step in the } x\text{-direction}) + y \times (1 \text{ step in the } y\text{-direction}),$$

which is obvious almost by definition of what we understand by the coordinate pair (x, y). But then combining this with the definition of linearity, we get that for any linear transformation f, we must have

$$\begin{aligned} f(x, y) &= f(x\boldsymbol{i} + y\boldsymbol{j}) \\ &= f(x\boldsymbol{i}) + f(y\boldsymbol{j}) \\ &= x\, f(\boldsymbol{i}) + y\, f(\boldsymbol{j}). \end{aligned}$$

In other words, in order to know what f does to *any point* (x, y), all we need to know is what f does to $\boldsymbol{i}$ and $\boldsymbol{j}$. This is quite remarkable. It should make sense intuitively if we remember what linear means visually: that the grid lines remain parallel and evenly spaced. If we have $f(\boldsymbol{i})$ and $f(\boldsymbol{j})$, there is only one way to continue the grid lines in such a way that they are parallel and evenly spaced, and so we are able to infer what f does to the whole plane (see figure 13).

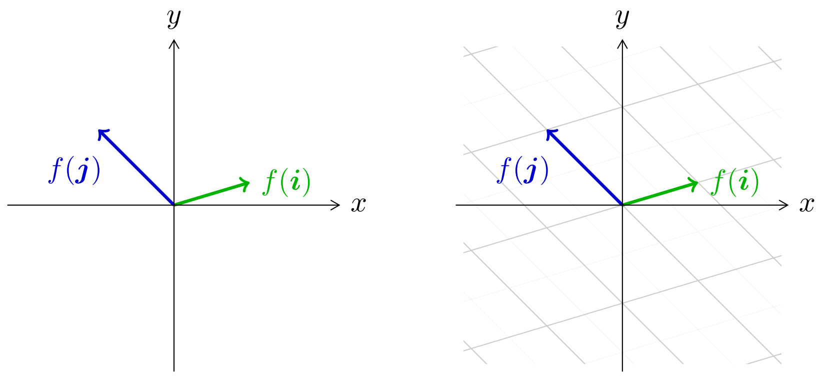

FIGURE 13: Given $f(\boldsymbol{i})$ and $f(\boldsymbol{j})$, there is only one way to continue the grid lines, this gives us an image of what happens to every point in space, effectively determining what $f(\boldsymbol{v})$ is for any $\boldsymbol{v}$

Example 1.16. Given that f is linear,

$$f\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}2\\3\end{pmatrix}\qquad\text{and}\qquad f\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}1\\-2\end{pmatrix},\qquad\text{what is}\qquad f\begin{pmatrix}5\\-8\end{pmatrix}?$$

We can find this since by linearity,

$$f\begin{pmatrix}5\\-8\end{pmatrix}=5f\begin{pmatrix}1\\0\end{pmatrix}+(-8)f\begin{pmatrix}0\\1\end{pmatrix}=5\begin{pmatrix}2\\3\end{pmatrix}-8\begin{pmatrix}1\\-2\end{pmatrix}=\begin{pmatrix}2\\31\end{pmatrix}.$$

Example 1.17. Given that f is linear, $f(1,0)=(2,2)$ and $f(0,1)=(2,2)$, determine a general formula for $f(x,y)$.

We have $f(x,y)=x\,f(1,0)+y\,f(0,1)=x(2,2)+y(2,2)=(2x+2y,2x+2y)$.

To summarise our observations, we have the following theorem.

Theorem 1.18. *Let $f\colon\mathbb{R}^2\to\mathbb{R}^2$ be a linear transformation. If $f(\boldsymbol{i})=\binom{a}{c}$ and $f(\boldsymbol{j})=\binom{b}{d}$, then for any vector $(x,y)\in\mathbb{R}^2$, we have*

$$f\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}ax+by\\cx+dy\end{pmatrix}.$$

Proof. We have

$$f\begin{pmatrix}x\\y\end{pmatrix}=f(x\boldsymbol{i}+y\boldsymbol{j})=x\,f(\boldsymbol{i})+y\,f(\boldsymbol{j})=x\begin{pmatrix}a\\c\end{pmatrix}+y\begin{pmatrix}b\\d\end{pmatrix}=\begin{pmatrix}ax+by\\cx+dy\end{pmatrix},$$

as required. □

Since $f(\boldsymbol{i})$ and $f(\boldsymbol{j})$ determine f completely, we introduce the following notation which "summarises" the behaviour f:

$$f = \begin{pmatrix} f(\boldsymbol{i}) & f(\boldsymbol{j}) \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

This 2×2 grid of numbers is called a *matrix*. Acting on the left, we interpret it as being the same as f. In other words,

$$\underbrace{\begin{pmatrix} a & b \\ c & d \end{pmatrix}}_{=f} \begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} a \\ c \end{pmatrix} + y \begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

by theorem 1.18. We call the act of evaluating the result of $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ *matrix-vector multiplication.*

Example 1.19. We evaluate the matrix-vector product:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 7 \\ -5 \end{pmatrix} = \begin{pmatrix} 1 \cdot 7 + 2 \cdot (-5) \\ 3 \cdot 7 + 4 \cdot (-5) \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \end{pmatrix}.$$

Understand what we've done here: we've found what the linear transformation for which $f(\boldsymbol{i}) = \binom{1}{3}$ and $f(\boldsymbol{j}) = \binom{2}{4}$ does to the point $\binom{7}{-5}$.

Matrices are of interest in computer graphics, since they allow us to translate the manipulation of a picture on a screen into the process of evaluating a matrix vector product.

Exercise 1.20. 1. Suppose $f(\boldsymbol{i}) = (-2, 3)$ and $f(\boldsymbol{j}) = (5, -7)$.

   (a) Determine $f(2, 3)$, $f(7, 3)$ and $f(5, 2)$.

   (b) Determine a general formula for $f(x, y)$.

   (c) Write f as a matrix.

2. The z^2 transformation does $g(\boldsymbol{i}) = (1, 0)$ and $g(\boldsymbol{j}) = (-1, 0)$.

   (a) What is the linear transformation f which agrees with it at these points?

   (b) Does it agree with it anywhere else? (i.e., are there other points $\boldsymbol{v}$ such that $f(\boldsymbol{v}) = g(\boldsymbol{v})$ other than $\boldsymbol{i}$ and $\boldsymbol{j}$?)

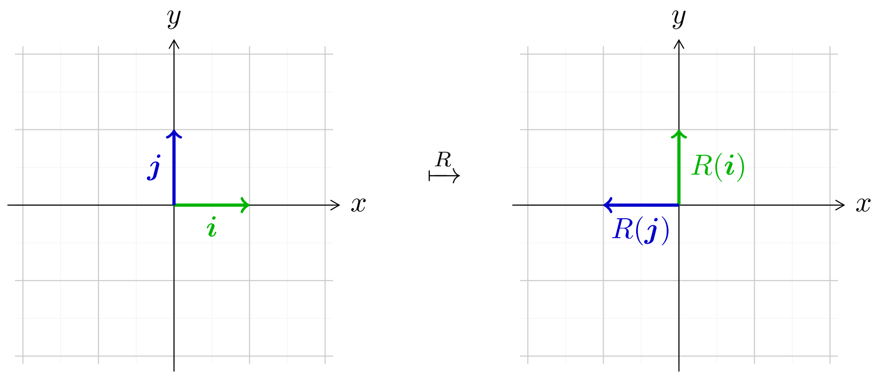

FIGURE 14: Rotation by 90° anti-clockwise

1.4 Finding and Interpreting 2×2 Matrices

Intuitively, we see that rotations, reflections and enlargements all keep the grid lines of the plane parallel and evenly spaced, so these are linear transformations. How can we find their matrices so that we can easily apply them to things?

All we need to do is think about what happens to the L-shape corresponding to the vectors $\boldsymbol{i}$ and $\boldsymbol{j}$. For instance, say we want to find the matrix R corresponding to a rotation by 90° (anti-clockwise). If we draw a rough sketch and think about where $\boldsymbol{i}$ and $\boldsymbol{j}$ end up (figure 14), we get that $R(\boldsymbol{i}) = \binom{0}{1}$ and $R(\boldsymbol{j}) = \binom{-1}{0}$, so the matrix is

$$R = \begin{pmatrix} R(\boldsymbol{i}) & R(\boldsymbol{j}) \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Here is a tool which carries out transformations represented by 2×2 matrices:

https://maths.mt/matrices.[2]

Try out this matrix that we found and verify that it does indeed rotate the plane by 90°.

Example 1.21. As another example, we compute the matrix T corresponding to a reflection in the line $y = -x$. With reference to figure 15, the desired matrix is easily seen to be

$$T = \begin{pmatrix} T(\boldsymbol{i}) & T(\boldsymbol{j}) \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}.$$

[2]Warning: the red/blue vectors which appear in the tool are *not* where $\boldsymbol{i}$ and $\boldsymbol{j}$ end up, they are what's called *eigenvectors* of the matrix. Just ignore them.

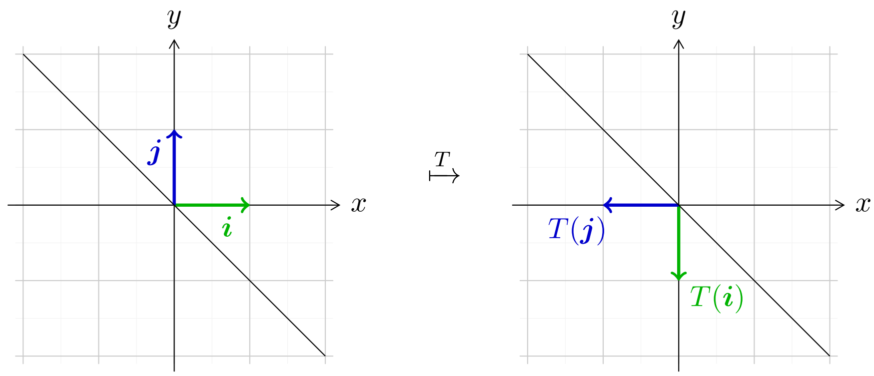

FIGURE 15: Reflection in the line $y = -x$

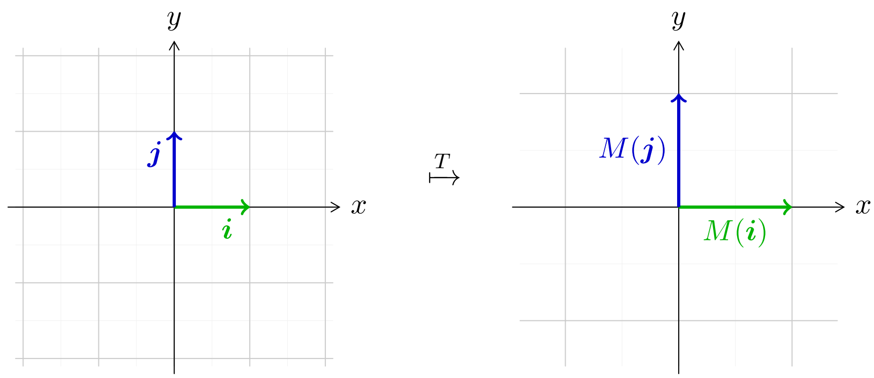

FIGURE 16: What M does to the plane

Example 1.22. This time we interpret the geometric meaning of a given matrix. Let

$$M = \begin{pmatrix} \frac{3}{2} & 0 \\ 0 & \frac{3}{2} \end{pmatrix}.$$

This tells us that $M(\boldsymbol{i}) = \binom{3/2}{0}$ and $M(\boldsymbol{j}) = \binom{0}{3/2}$. With reference to figure 16, we see that after applying the transformation, the plane is uniformly enlarged by a scale factor of $\frac{3}{2}$. Double check this using the visualiser tool.

Exercise 1.23. 1. Determine the matrices corresponding to the following transformations.

(a) *Clockwise* rotation of $90°$ about the origin.