Linear Transformations

Thus far we have studied vector spaces individually. In this chapter we will discuss transformations between vector spaces. They may very well have different vector additions and scalar multiplications, but we will require that they have the same field of scalars: $\mathbb{F}$.

Definition 1

Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces. A function $T : \mathbb{V} \to \mathbb{W}$ is called a **linear transformation** if for all $\vec{u}$, $\vec{v} \in \mathbb{V}$ and all $r \in \mathbb{F}$

(a) $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$

(b) $T(r \cdot \vec{u}) = r \cdot T(\vec{u})$

Strictly speaking the two vector spaces might be quite different, with different operations:

Example 1

(1) For $M_{2\times 2}(\mathbb{R})$ "$+$" is the addition of matrices, and "$\cdot$" the product of a scalar with a matrix.

(2) For $P_2(\mathbb{R})$ "$+$" is the addition of polynomials, and "$\cdot$" the product of a scalar with a polynomial.

(3) For $C([0,1],\mathbb{R})$ "$+$" is the addition of continuous functions, and "$\cdot$" the product of a scalar with a continuous function.

(4) For $S(\mathbb{R})$ "$+$" is the addition of sequences, and "$\cdot$" the product of a scalar with a sequence.

Note 1

If $T : A \to B$, then T is a function "from" A "to" B. We call A the **domain** of the function and B the **codomain** of the function. The terms "transformation" and "map" are just different names for functions

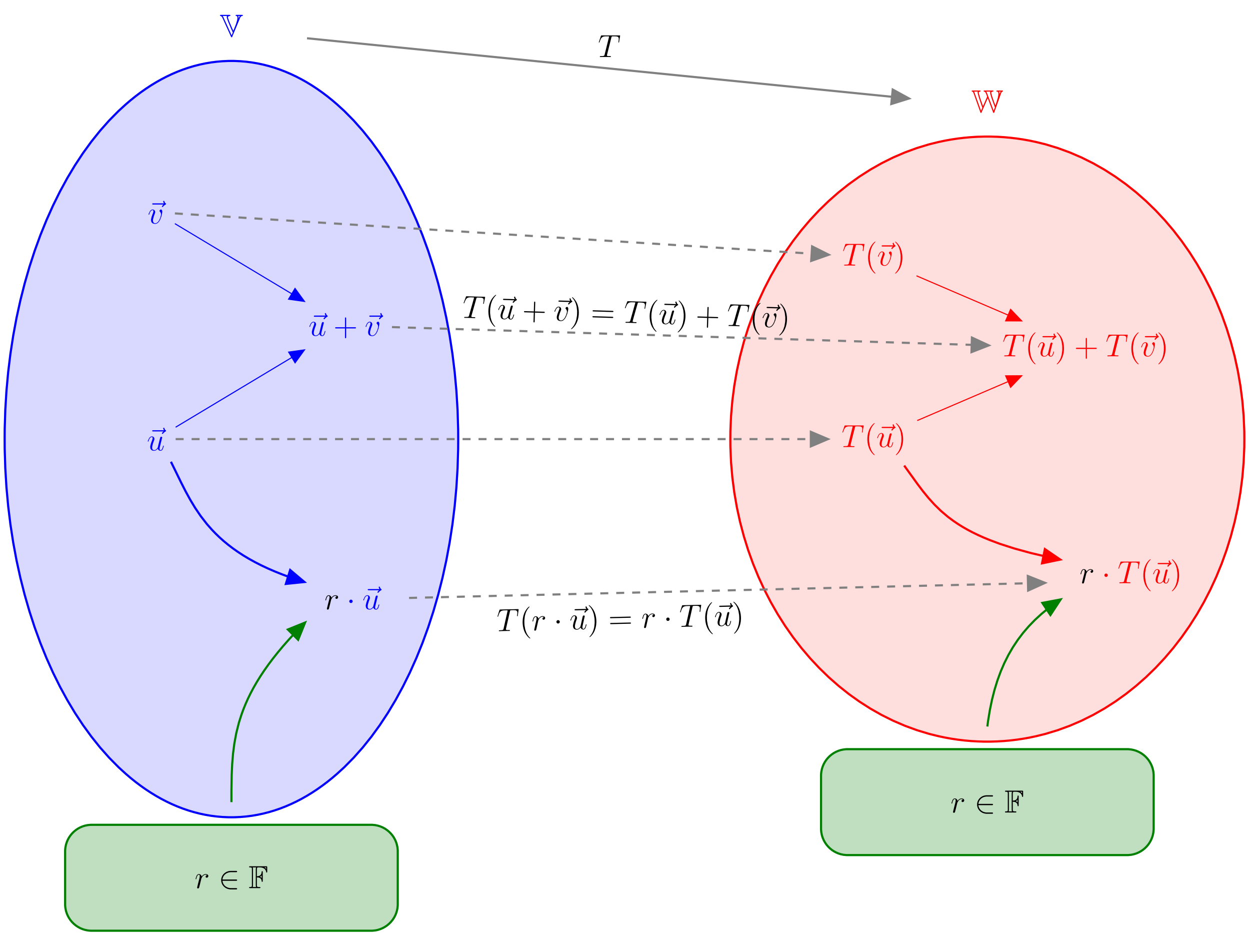
Example 2

If $T : M_{2\times 2}(\mathbb{R}) \to P_2(\mathbb{R})$ then the vector addition in the domain space $M_{2\times 2}(\mathbb{R})$ is the addition of matrices, while the vector addition in the codomain $P_2(\mathbb{R})$ is the addition of polynomials. Similarly for the scalar multiplication

However the operations in the two spaces are in sync!

$T:\ \mathbb{V} \to \mathbb{W}$ a Linear Transformation

($\mathbb{V}$ and $\mathbb{W}$ both over $\mathbb{F}$)



The vector additions in $\mathbb{V}$ and $\mathbb{W}$ are in sync

$$T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$$

and the scalar multiplications in $\mathbb{V}$ and $\mathbb{W}$ are in sync

$$r \cdot T(\vec{u}) = T(r \cdot \vec{u})$$

So even though the transformation T maps one space to another, it does it in such a way that the operations are synchronized. The operations in the domain correspond with the operations in the codomain, the world of the images

It would be better to denote the operations with different symbols in the definition to emphasize their differences: e.g.

$$\langle \mathbb{V}, \mathbb{F}, \oplus, \odot \rangle \qquad\qquad \langle \mathbb{W}, \mathbb{F}, \boxplus, \boxdot \rangle$$

Here $\oplus$ is the vector edition in $\mathbb{V}$, while $\boxplus$ is the vector edition in $\mathbb{W}$,
and $\odot$ is the scalar multiplication in $\mathbb{V}$, while $\boxdot$ is the scalar multiplication in $\mathbb{W}$.

We could use the following definition instead:

Definition 2

Let $\langle \mathbb{V}, \mathbb{F}, \oplus, \odot \rangle$ and $\langle \mathbb{W}, \mathbb{F}, \boxplus, \boxdot \rangle$ be two vector spaces. A function $T: \mathbb{V} \to \mathbb{W}$ is called a **linear transformation** if for all $\vec{u}$, $\vec{v} \in \mathbb{V}$ and all $r \in \mathbb{F}$

(a) $T(\vec{u} \oplus \vec{v}) = T(\vec{u}) \boxplus T(\vec{v})$

(b) $T(r \odot \vec{u}) = r \boxdot T(\vec{u})$

Of course for most people this looks rather cryptic. It certainly is unnecessary when the operations of vector addition and scalar multiplication in the vector spaces are well understood. In that case it is much nicer to use the first definition, which looks much more digestible, but then you have to keep in mind that the "+" and "·" symbols that are being used pertain to the corresponding vector spaces, and thus may be very different operations.

Example 3 $T: \mathbb{R}^3 \to M_{2\times 2}(\mathbb{R})$ defined by $T\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a+b & b-c \\ b & a+c \end{bmatrix}$.

Let's verify $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$ for $\vec{u} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} 1 \\ -4 \\ 5 \end{bmatrix}$. Of course the two $+$ signs are different:

$$T\left(\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ -4 \\ 5 \end{bmatrix}\right) = T\left(\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}\right) + T\left(\begin{bmatrix} 1 \\ -4 \\ 5 \end{bmatrix}\right)$$

vector addition in $\mathbb{R}^3$ matrix addition in $M_{2\times 2}(\mathbb{R})$

$$T\left(\begin{bmatrix} 4 \\ -3 \\ 7 \end{bmatrix}\right) = \begin{bmatrix} 4 & -1 \\ 1 & 5 \end{bmatrix} + \begin{bmatrix} -3 & -9 \\ -4 & 6 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -10 \\ -3 & 11 \end{bmatrix} = \begin{bmatrix} 1 & -10 \\ -3 & 11 \end{bmatrix} \qquad \checkmark$$

To check that this is a linear transformation we need to check this for **all** $\vec{u},\ \vec{v} \in \mathbb{V}$ and all $r \in \mathbb{F}$, not only the one we just picked, we have to prove

$$\textbf{(a)} \quad T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$$

$$\textbf{(b)} \quad T(r\cdot\vec{u}) = r\cdot T(\vec{u})$$

in general. This is not hard, but takes some algebra.

Let $\vec{u} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} A \\ B \\ C \end{bmatrix}$ be arbitrary vectors in $\mathbb{R}^3$ and $r \in \mathbb{F}$ an arbitrary scalar.

$$\begin{aligned}
\textbf{(a)} \quad T(\vec{u}+\vec{v}) = T\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix} + \begin{bmatrix} A \\ B \\ C \end{bmatrix}\right) &= T\left(\begin{bmatrix} a+A \\ b+B \\ c+C \end{bmatrix}\right) \\
&= \begin{bmatrix} (a+A)+(b+B) & (b+B)-(c+C) \\ b+B & (a+A)+(c+C) \end{bmatrix} \\
&= \begin{bmatrix} a+b & b-c \\ b & a+c \end{bmatrix} + \begin{bmatrix} A+B & B-C \\ B & A+C \end{bmatrix} \\
&= T\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) + \left(\begin{bmatrix} A \\ B \\ C \end{bmatrix}\right) = T\vec{u}) + T(\vec{v}) \quad \checkmark
\end{aligned}$$

$$\begin{aligned}
\textbf{(b)} \quad T(r\cdot\vec{u}) = T\left(r\cdot\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) &= T\left(\begin{bmatrix} r\cdot a \\ r\cdot b \\ r\cdot c \end{bmatrix}\right) \\
&= \begin{bmatrix} r\cdot a + r\cdot b & r\cdot b - r\cdot c \\ r\cdot b & r\cdot a + r\cdot c \end{bmatrix} \\
&= r\cdot\begin{bmatrix} a+b & b-c \\ b & a+c \end{bmatrix} \\
&= r\cdot T\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) = r\cdot T(\vec{u}) \quad \checkmark
\end{aligned}$$

This proves that T is a linear transformation. $\square$

Note 2 Sometimes the two properties of linearity are combined in one as follows

$$T(r\cdot\vec{u}+\vec{v}) = r\cdot T(\vec{u}) + T(\vec{v})$$

This is one of the great properties of linear transformations

$$\boxed{T(\, a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_m\vec{v}_m) \;=\; a_1T(\,\vec{v}_1\,) + a_2T(\,\vec{v}_2\,) + \cdots + a_mT(\,\vec{v}_m\,)}$$

i.e.

Theorem 1

If T is a linear transformation then

$$T\left(\sum_{i=1}^{m} a_i\vec{v}_i\right) \;=\; \sum_{i=1}^{m} a_i\, T(\,\vec{v}_i\,)$$

Example 4 Here are some examples of linear transformations.

(1) $T : P_2(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ defined by $T(\,a\,t^2 + b\,t + c\,) = \begin{bmatrix} a+6b & 2b-c \\ c-2b & a+3c \end{bmatrix}$

(2) $T : M_{2\times 2}(\mathbb{R}) \to \mathbb{R}^3$ defined by $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a-3b \\ b+c \\ 3c-a \end{bmatrix}$

(3) $T : M_{2\times 2}(\mathbb{R}) \to \mathbb{R}$ defined by $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \int_0^1 (a\,t^2 + (b-c)\,t + d)\,dt$

(4) $T : P_3(\mathbb{R}) \to P_2(\mathbb{R})$ defined by $T(a\,t^3 + b\,t^2 + c\,t + d) = 3a\,t^2 + 2b\,t + c$ or in calculus terms $T(p(t)) = \dfrac{dp(t)}{dt}$

(5) $T : \mathbb{R}^3 \to \mathbb{R}^2$ defined by $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = (4x+y)\begin{bmatrix} -1 \\ 2 \end{bmatrix} + (x-2y+3z)\begin{bmatrix} 3 \\ -5 \end{bmatrix}$.

(5) $T : \mathbb{R}^3 \to \mathbb{R}$ defined by $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix} \times \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, where $\times$ is the cross product.

(6) $T_{\vec{v},\theta} : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T_{\vec{v},\theta}(\vec{x}) = (1-\cos\theta)\dfrac{\vec{v}\,\vec{v}^{\top}}{\vec{v}\cdot\vec{v}}\vec{x} + \cos\theta\cdot I + \dfrac{\sin\theta}{||\,\vec{v}\,||}\cdot\vec{v}\times\vec{x}$

(7) $\varphi_\alpha : \mathbb{V} \to \mathbb{F}^n$ defined by $\varphi_\alpha(\,\vec{v}\,) = [\,\vec{v}\,]_\alpha$ where α is a basis of the vector space $\mathbb{V}$ over the field $\mathbb{F}$, and $\dim(\mathbb{V}) = n$. This is the transformation that takes a vector of a vector space and maps it to its coordinate vector with respect to the basis α.

Let's show this last map, $\varphi_\alpha : \mathbb{V} \to \mathbb{F}^n$ defined by $\varphi_\alpha(\vec{v}) = [\vec{v}]_\alpha$, is indeed linear.

Let $\vec{u}$ and $\vec{v}$ be arbitrary vectors of $\mathbb{V}$, and let $\alpha = \{\vec{a}_1, \vec{a}_2, \vec{a}_3, \cdots, \vec{a}_n\}$.

Since α is a basis, there is only one way to write any vector as a linear combination of the vectors in α. Suppose the unique coefficients for the vectors $\vec{u}$ and $\vec{v}$ are

$$\vec{u} = u_1 \cdot \vec{a}_1 + u_2 \cdot \vec{a}_2 + u_3 \cdot \vec{a}_3 + \cdots + u_n \cdot \vec{a}_n$$

$$\vec{v} = v_1 \cdot \vec{a}_1 + v_2 \cdot \vec{a}_2 + v_3 \cdot \vec{a}_3 + \cdots + v_n \cdot \vec{a}_n$$

Or equivalently $[\vec{u}]_\alpha = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $[\vec{v}]_\alpha = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$

Clearly $\vec{u} + \vec{v} = (u_1 + v_1) \cdot \vec{a}_1 + (u_2 + v_2) \cdot \vec{a}_2 + (u_3 + v_3) \cdot \vec{a}_3 + \cdots + (u_n + v_n) \cdot \vec{a}_n$, and since this expression with respect to the basis α is unique:

$$[\vec{v} + \vec{u}]_\alpha = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}$$

Hence

(1) $\varphi_\alpha(\vec{u} + \vec{v}) = [\vec{v} + \vec{u}]_\alpha = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = [\vec{u}]_\alpha + [\vec{v}]_\alpha = \varphi_\alpha(\vec{u}) + \varphi_\alpha(\vec{v})$

Similarly, $r \cdot \vec{u} = r\, u_1 \cdot \vec{a}_1 + r\, u_2 \cdot \vec{a}_2 + r\, u_3 \cdot \vec{a}_3 + \cdots + r\, u_n \cdot \vec{a}_n$, is the unique representation of $r \cdot \vec{u}$ with respect to u: i.e.

$$[r \cdot \vec{u}]_\alpha = \begin{bmatrix} r\, u_1 \\ r\, u_2 \\ \vdots \\ r\, u_n \end{bmatrix}$$

Hence

(2) $\varphi_\alpha(r \cdot \vec{u}) = [r \cdot \vec{u}]_\alpha = \begin{bmatrix} r\, u_1 \\ r\, u_2 \\ \vdots \\ r\, u_n \end{bmatrix} = r \cdot \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix} = r \cdot \varphi_\alpha(\vec{u})$

So that φ_α is a linear transformation.

Here are some facts that are good to know.

Theorem 2

If $T: \mathbb{V} \to \mathbb{W}$ is a linear transformation.

then **(a)** $T(\vec{0}_{\mathbb{V}}) = \vec{0}_{\mathbb{W}}$

(b) $T(-\vec{u}) = -T(\vec{u})$

(c) $T(\vec{u} - \vec{v}) = T(\vec{u}) - T(\vec{v})$

Proof: Let $\vec{u},\ \vec{v} \in \mathbb{V}$

(a) $T(\vec{0}_{\mathbb{V}}) = T(0 \cdot \vec{u}) = 0 \cdot T(\vec{u}) = \vec{0}_{\mathbb{W}}$

(b) $T(-\vec{u}) = T\Big((-1) \cdot \vec{u}\Big) = (-1) \cdot T(\vec{u}) = -T(\vec{u})$

(c) $T(\vec{u} - \vec{v}) = T\Big(\vec{u} + (-1) \cdot \vec{v}\Big) = T(\vec{u}) + (-1) \cdot T(\vec{v}) = T(\vec{u}) - T(\vec{v})$ ◻

Theorem 3

A linear transformation $T: \mathbb{V} \to \mathbb{W}$ is completely determined by what happens to a basis of $\mathbb{V}$.

Proof: Suppose $\alpha = \{\vec{a}_1,\ \vec{a}_2,\ \vec{a}_3, \cdots,\ \vec{a}_n\}$ ia a basis, and that

$$T(\vec{a}_1) = \vec{b}_1, \qquad T(\vec{a}_2) = \vec{b}_2, \qquad T(\vec{a}_3) = \vec{b}_3, \qquad \cdots \quad , \ T(\vec{a}_n) = \vec{b}_n$$

Hence we can now compute the image of **any** vector $\vec{v} \in \mathbb{V}$ as follows:

$$\begin{aligned} T(\vec{v}) &= T\Big(\, v_1 \cdot \vec{a}_1 + v_2 \cdot \vec{a}_2 + v_3 \cdot \vec{a}_3 + \cdots + v_n \cdot \vec{a}_n \Big) \\ &= v_1 \cdot T(\vec{a}_1) + v_2 \cdot T(\vec{a}_2) + v_3 \cdot T(\vec{a}_3) + \cdots + v_n \cdot T(\vec{a}_n) \\ &= v_1 \cdot \vec{b}_1 + v_2 \cdot \vec{b}_2 + v_3 \cdot \vec{b}_3 + \cdots + v_n \cdot \vec{b}_n \end{aligned}$$ ◻

Example 5 Let $T: \mathbb{R}^3 \to P_2(\mathbb{R})$ be a linear transformation with

$$T\begin{bmatrix}1\\0\\0\end{bmatrix} = t^2 + 3t - 1, \quad T\begin{bmatrix}0\\1\\0\end{bmatrix} = 5t^2 - 4t + 6, \quad T\begin{bmatrix}0\\0\\1\end{bmatrix} = 7t + 4$$

Compute the image of $\begin{bmatrix}2\\3\\-1\end{bmatrix}$:

$$
\begin{aligned}
T\begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix} &= T\left(2\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) \\
&= 2\cdot T\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 3\cdot T\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - T\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\
&= 2\cdot\left(t^2+3t-1\right) + 3\cdot\left(5t^2-4t+6\right) - \left(7t+4\right) \\
&= 17t^2 - 13t + 12
\end{aligned}
$$

Of course knowing what happens to the standard basis makes things easier. Let's do a harder example.

Example 6 Let $T : \mathbb{R}^3 \to P_2(\mathbb{R})$ be a linear transformation with

$$
T\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = t^2+3t-1, \quad T\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = 5t^2-4t+6, \quad T\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = 7t+4
$$

Now if we want to compute the image of $\begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$ we have to do a rref first to express the vector with respect to the new basis:

$$
\operatorname{rref}\begin{bmatrix} 1 & 2 & 0 & 2 \\ 1 & 1 & 1 & 3 \\ 1 & 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -6 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 5 \end{bmatrix}
$$

so that

$$
\begin{aligned}
T\begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix} &= T\left(-6\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + 4\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + 5\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right) \\
&= -6\cdot T\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + 4\cdot T\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + 5\cdot T\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \\
&= -6\cdot\left(t^2+3t-1\right) + 4\cdot\left(5t^2-4t+6\right) + 5\cdot\left(7t+4\right) \\
&= 14t^2 + t + 50
\end{aligned}
$$

Let's define some terms related to functions: domain, codomain, image, preimage, range, one-to-one and onto.

Definition 3

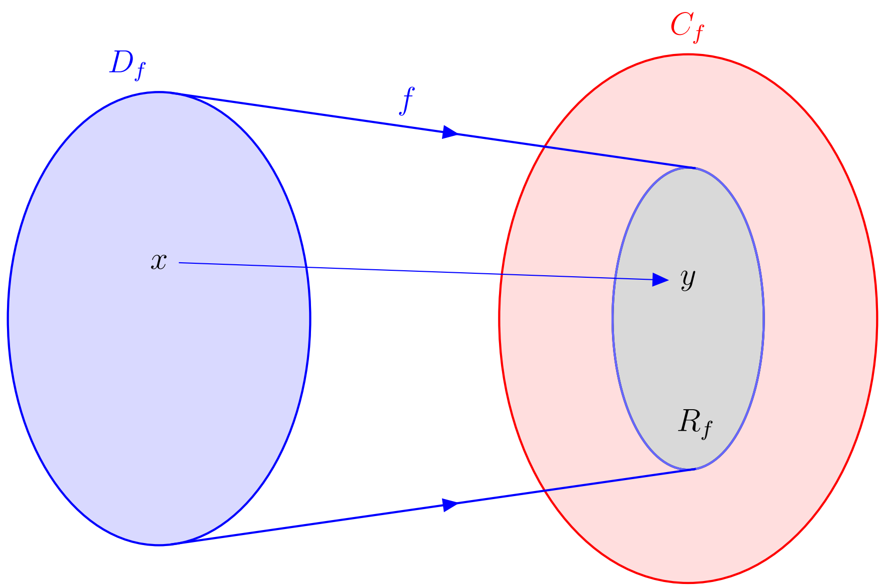
Let $f : A \to B$ be a function from the set A to the set B.

(a) A is called the **domain** of f. [Notation: D_f]

(b) B is called the **codomain** of f. [Notation: C_f]

(c) If $f : x \mapsto y$ then y is called the **image** of x, and x is called a **pre-image** of y.

(d) The set of all images is called the **range** of f. [R_f]



Note 3

If $f : x \mapsto y$ then y is called **the** image of x, since f is a function. By definition a function can have only one image for every element x in the domain.

Note 4

If $f : x \mapsto y$ then x is called **a** pre-image of y. An image can have multiple pre-images. Many elements of A can be mapped to the same image.

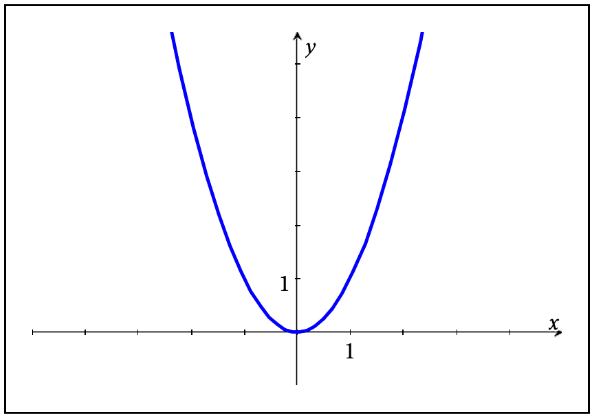
Note 5

The range is a subset of the codomain: $R_f \subseteq C_f$. It may very well be that there are elements in C_f that are not images of any element of the domain: nothing gets mapped to these elements by f. But it could also be that every element of the codomain is the image of some element in the domain.

Example 7

Let $f : \mathbb{R} \to \mathbb{R}$ is defined by $f(x) = x^2$. In this case we have

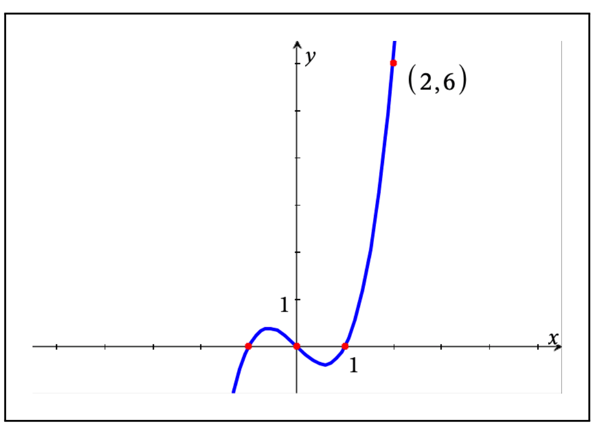
- $D_f = \mathbb{R}$, the domain is the set of all real numbers.
- $C_f = \mathbb{R}$, the codomain is also the set of all real numbers.
- $R_f = [\,0, \infty\,)$, the range is the set of all non-negative real numbers.
- $f(2) = 4$ hence 4 is **the** image of 2, but 2 is only one of the pre-images of 4, in fact both 2 and -2 are mapped to 4.



Example 8

Let $f : \mathbb{R} \to \mathbb{R}$ is defined by $f(x) = x^3 - x$. In this case we have

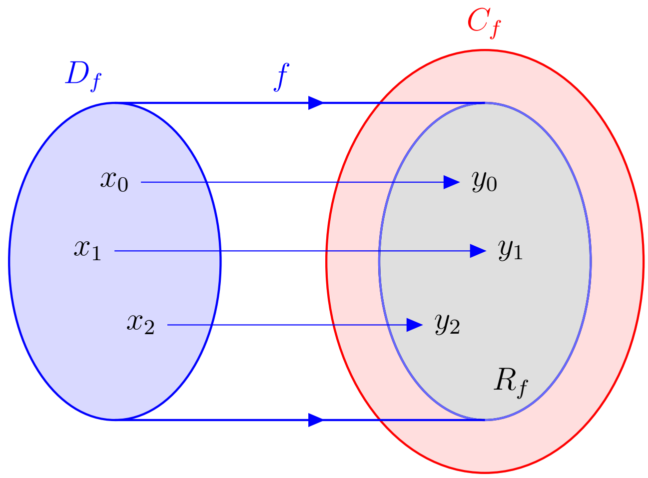
- $D_f = \mathbb{R}$, the domain is the set of all real numbers.
- $C_f = \mathbb{R}$, the codomain is also the set of all real numbers.
- $R_f = \mathbb{R}$, the range is also the set of real numbers.
- $f(1) = 0$ hence 0 is **the** image of 1, but 1 is only one of the pre-images of 0, in fact -1, 0 and -2 are all mapped to 0.
- $f(2) = 6$ hence 6 is **the** image of 2, and in this case 2 is the only pre-image of 6.



Here are two more concepts: one-to-one, and onto.

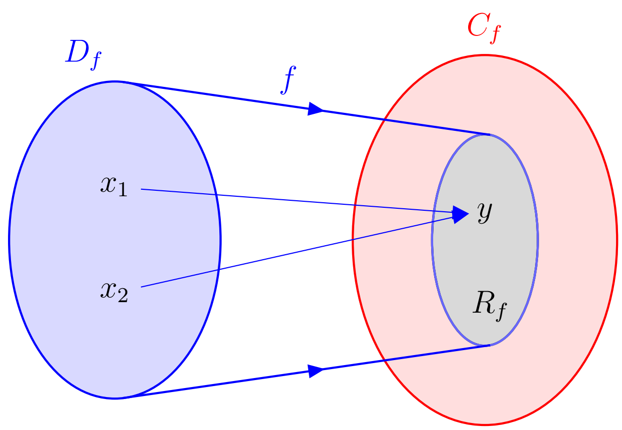
Definition 4

A function $f : A \to B$ is called **one-to-one** if every element of the range has only one pre-image.



Note 6

A function is **not** one-to-one if there is one image with at least two pre-images



Example 9

The functions in example 7 and 8:

$f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ is not one-to-one

$f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3 - x$ is not one-to-one

But $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = e^x$ is one-to-one.

Theorem 4

A function $f : A \to B$ is one-to-one, when

$$f(x_1) = f(x_2) \quad \Rightarrow \quad x_1 = x_2$$

Proof: If there exist x_1 and x_2 that have the same image y

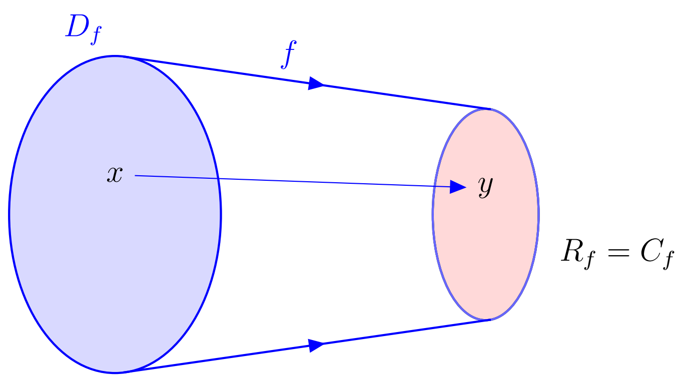
$$f(x_1) = y = f(x_2)$$

then y would have two pre-images, hence f would not be one-to-one. $\square$.

This is often the 'acid' test that is used. Does $f(x_1) = f(x_2)$ imply $x_1 = x_2$; or are there $x_1 \neq x_2$ with $f(x_1) = f(x_2)$?

Definition 5

A function $f : A \to B$ is called **onto** if $R_f = C_f$, i.e. if every element of the codomain has a pre-image.



Example 10

The function in example 7

$f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ is not onto

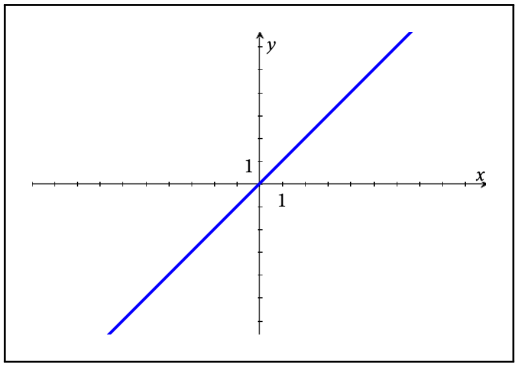
whereas the function in example 8

$f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3 - x$ is onto

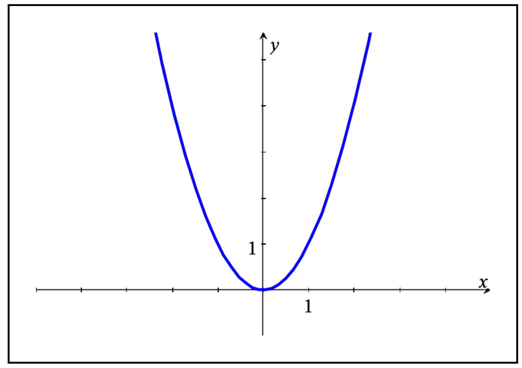
Note 7

The notions of **onto** and **one-to-one** are not related, in the sense that the one doesn't imply the other. There are functions that are both one-to-one and onto. There are functions that are neither one-to-one nor onto, and there are functions that or one or the other.

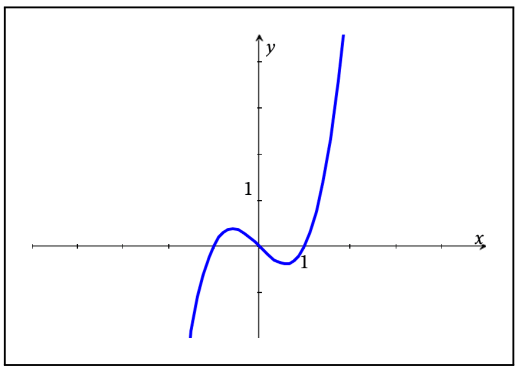
Example 11 Here are examples of all four cases. In each case $f : \mathbb{R} \to \mathbb{R}$



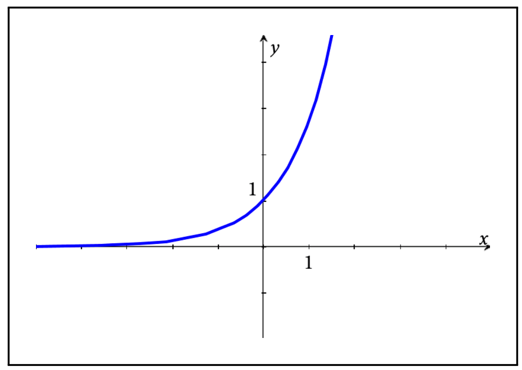
$f(x) = x$ one-to-one **and** onto



$f(x) = x^2$ **neither** one-to-one **nor** onto



$f(x) = x^3 - x$ **not** one-to-one **but** onto



$f(x) = e^x$ one-to-one but **not** onto

Some books call a function that is one-to-one "injective", and a function that is onto "surjective". If a function is both an injection and a surjection it is called a bijection. In our terminology:

Definition 6 A function $f : A \to B$ that is **both** **one-to-one and onto** is called a **bijection**.

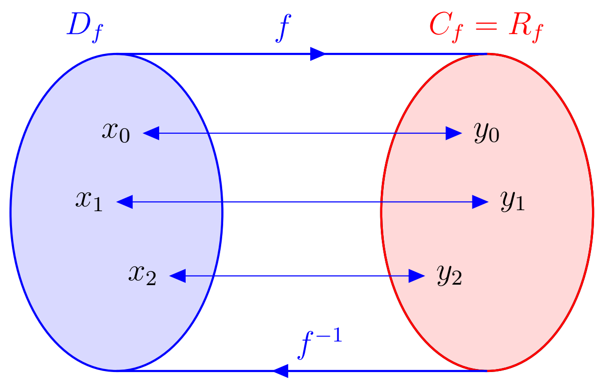
If the function $f : A \to B$ is a bijection, then every element of B has one and only one pre-image. There is a one-to-one match up between all the elements of A and B. The function pairs every element of A with one element of B and every element of B with one element of A, its unique pre-image. Such a function is called invertible.

[Some authors call a function $f : A \to B$ invertible when it is one-to-one. But of course in that case the inverse function f^{-1} has to be defined as a map from R_f to A. We will link invertibility to bijections, functions that are both one-to-one and onto.]

Definition 7

If $f : A \to B$ is a bijection it is called **invertible**.

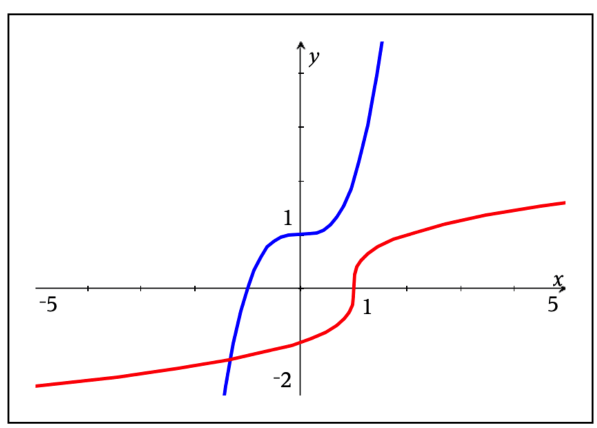
The inverse function, which is denoted by f^{-1}, maps each element of B to its unique pre-image in A.



Example 12

$f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3 + 1$ is a bijection, hence invertible.

Its inverse is the function $f^{-1} : \mathbb{R} \to \mathbb{R}$ defined by $f^{-1}(x) = \sqrt[3]{x-1}$



The graph of f is colored blue, while the graph of f^{-1} is colored red. Note that

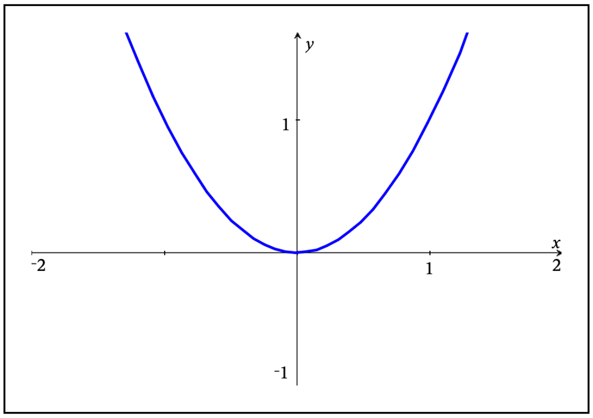
$$x \underset{f^{-1}}{\overset{f}{\rightleftarrows}} y$$

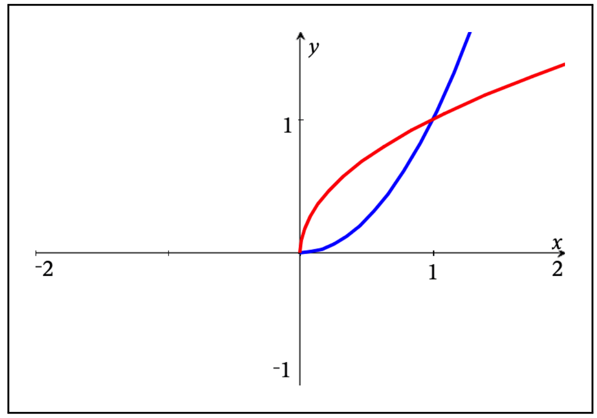
Basically the roles of x and y are flipped. Input and output are switched. Hence a point (x_1, y_1) on the graph of f corresponds to the point (y_1, x_1) on the graph of f^{-1} and vice versa. So that the graphs of f and f^{-1} are symmetric over the line $x = y$.

Example 13

$f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ is a not a bijection, hence not invertible. It is not one-to-one nor onto. But if we change the domain and codomain we can can obtain an invertible (part of the) function.

The function $f : [0, \infty] \to [0, \infty]$ defined by $f(x) = x^2$ **is** a bijection, hence invertible. Its inverse is $f^{-1} : [0, \infty] \to [0, \infty]$ is $f^{-1}(x) = \sqrt{x}$. Both graphs are shown on the right.





Note that these concepts, domain, codomain, image, pre-image, range, one-to-one, onto, bijection, inverse pertain to functions in general. We will certainly use them in the context of linear algebra as well. In linear algebra the only functions we discuss are linear functions. Still a linear function has a domain, codomain, range etc. There are some special definitions like:

Definition 8

A transformation $T : \mathbb{V} \to \mathbb{W}$ that is linear and invertible is called an **isomorpism**.

If there exists an isomorpism $T : \mathbb{V} \to \mathbb{W}$ between two vector spaces $\mathbb{V}$ and $\mathbb{W}$ they are called **isomorphic**, which is denoted by $\mathbb{V} \cong \mathbb{W}$

Theorem 5

Let $\mathbb{V}$ be an n-dimensional vector space over the field $\mathbb{F}$. If α is a basis of $\mathbb{V}$, then the linear transformation $\varphi_\alpha : \mathbb{V} \to \mathbb{F}^n$ defined by $\varphi_\alpha(\vec{v}) = [\,\vec{v}\,]_\alpha$ is an isomorphism, i.e.

$$\mathbb{V} \text{ and } \mathbb{F}^n \text{ are isomorphic}$$

Notation: $\boxed{\mathbb{V} \cong \mathbb{F}^n}$

Proof: We proved earlier that this map is linear. The only thing that needs to be shown, is that this map is one-to-one and onto. This we will leave to the reader as an exercise. □

Example 14

(1) $M_{2\times 2}(\mathbb{R}) \cong \mathbb{R}^4$

(2) $P_2(\mathbb{R}) \cong \mathbb{R}^3$

(3) $M_{3\times 2}(\mathbb{F}_7) \cong \mathbb{F}_7^6$

(4) $P_3(\mathbb{C}) \cong \mathbb{C}^4$

(5) $P_3(\mathbb{R}) \cong M_{2\times 2}(\mathbb{R})$

To prove that $P_3(\mathbb{R}) \cong M_{2\times 2}(\mathbb{R})$ we need to find an isomorphism $T : P_3(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$. The following map will do:

$$T(\, a\,t^3 + b\,t^2 + c\,t + d\,) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Exercise: Show that this T is linear, one-to-one and onto.

We'll conclude with some useful theorems:

Theorem 6 The composition of linear functions is linear

Proof: Let $T : A \to B$ and $U : B \to C$ then

(a) $$\begin{aligned}(U \circ T)(\vec{u} + \vec{v}) &= U\Big(T(\vec{u} + \vec{v})\Big) \\ &= U\Big(T(\vec{u}) + T(\vec{v})\Big) \\ &= U\Big(T(\vec{u})\Big) + U\Big(T(\vec{v})\Big) \\ &= (U \circ T)(\vec{u}) + (U \circ T)(\vec{v})\end{aligned}$$

(b) $$\begin{aligned}(U \circ T)(r \cdot \vec{u}) &= U\Big(T(r \cdot \vec{u})\Big) \\ &= U\Big(r \cdot T(\vec{u})\Big) \\ &= r \cdot \Big(U\Big(T(\vec{u})\Big)\Big) \\ &= r \cdot \Big((U \circ T)(\vec{u})\Big)\end{aligned}$$ □

Theorem 7

(a) The composition of one-to-one functions is one-to-one

(b) The composition of onto functions is onto.

(c) The composition of bijections is a bijection.

(d) The composition of isomorphisms is an isomorphism.

The proofs are easy, and left as an exercise.