3.3 Composition of Functions

Definition 1. *Let $f : X \to Y$ and $g : Y \to Z$ be two functions. Then the* ***composition*** *of f and g is the function*

$$(g \circ f) : X \to Z$$

defined by

$$(g \circ f)(x) = g(f(x)).$$

(See Figure 3.5.)

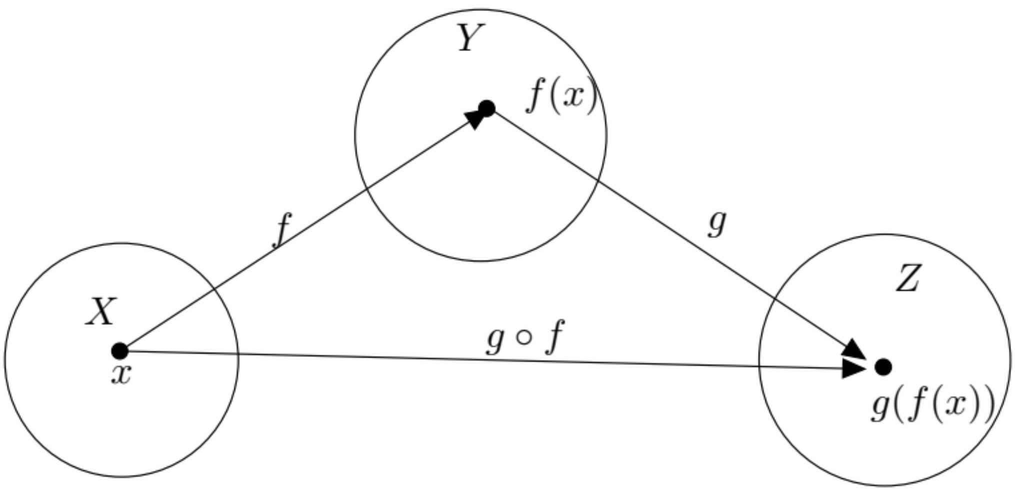

Figure 3.5: Composition of functions

Note that $(g \circ f)(x)$ is defined for every $x \in X$ because the image set of f is contained in the domain of g.

Suppose functions $f : X \to Y$ and $g : W \to Z$ are given. If $Y \neq W$, can you define $g \circ f$? Yes, if $\mathrm{R}(f) \subseteq W$. In that case, $g(f(x))$ is defined for every $x \in X$, and therefore $g \circ f : X \to Z$ is defined.

Exercise 0.1. Let $f, g : \mathbb{R} \to \mathbb{R}$ be defined by

$$f(x) = x^2, \qquad g(x) = 3x + 2.$$

Find $g \circ f$ and $f \circ g$. Are they the same?

Solution:

$$(g \circ f)(x) = g(f(x)) = g(x^2) = 3x^2 + 2.$$

$$(f \circ g)(x) = f(g(x)) = f(3x + 2) = (3x + 2)^2 = 9x^2 + 12x + 4.$$

Hence,

$$(g \circ f)(x) = 3x^2 + 2, \qquad (f \circ g)(x) = 9x^2 + 12x + 4,$$

so $(g \circ f)(x) \neq (f \circ g)(x)$.

Example 2. *Consider*

$$f : \mathbb{R} \to [0, 1), \quad f(x) = x^2,$$

and

$$g : [0, 1) \to \mathbb{R}, \quad g(x) = \sqrt{x}, \text{ the nonnegative square root.}$$

Then

$$(g \circ f)(x) = g(x^2) = \sqrt{x^2} = |x|, \quad x \in \mathbb{R},$$

and

$$(f \circ g)(x) = (\sqrt{x})^2 = x, \quad x \in [0, 1).$$

Thus the two compositions have different domains and are not the same as functions on $\mathbb{R}$*, but they agree on the common domain* $[0, 1)$*:*

$$(g \circ f)(x) = (f \circ g)(x) \quad \text{for all } x \in [0, 1).$$

Exercise 0.2. Let $f, g : \mathbb{R} \to \mathbb{R}$ be defined by

$$f(x) = x^3 + 2, \qquad g(x) = \sqrt[3]{x} \text{ (the real cube root).}$$

Find $f \circ g$ and $g \circ f$. Is $f \circ g = g \circ f$?

Solution. For all $x \in \mathbb{R}$,

$$(f \circ g)(x) = f(g(x)) = f(\sqrt[3]{x}) = (\sqrt[3]{x})^3 + 2 = x + 2.$$

Next,

$$(g \circ f)(x) = g(f(x)) = g(x^3 + 2) = \sqrt[3]{x^3 + 2}.$$

Therefore

$$(f \circ g)(x) = x + 2, \qquad (g \circ f)(x) = \sqrt[3]{x^3 + 2}.$$

These two functions are not equal for general $x \in \mathbb{R}$ (for example, at $x = 0$ we have $(f \circ g)(0) = 2$ but $(g \circ f)(0) = \sqrt[3]{2}$). Hence $f \circ g \neq g \circ f$.

Exercise 3. Let $f : M(2, \mathbb{R}) \to M(2, \mathbb{R})$ be defined by $f(A) = A^T A$, and $g : M(2, \mathbb{R}) \to \mathbb{R}$ be defined by $g(A) = \mathrm{tr}(A)$. Find $(g \circ f)(A)$ for $A \in M(2, \mathbb{R})$ in terms of the entries a_{ij} of A.

Solution. Write

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Then

$$A^T = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix},$$

and

$$A^T A = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11}^2 + a_{21}^2 & a_{11}a_{12} + a_{21}a_{22} \\ a_{12}a_{11} + a_{22}a_{21} & a_{12}^2 + a_{22}^2 \end{pmatrix}.$$

Taking trace,

$$(g \circ f)(A) = \mathrm{tr}\,(A^T A) = (a_{11}^2 + a_{21}^2) + (a_{12}^2 + a_{22}^2) = \sum_{i=1}^{2} \sum_{j=1}^{2} a_{ij}^2.$$

Thus $(g \circ f)(A)$ equals the sum of squares of all entries of A:

$$(g \circ f)(A) = a_{11}^2 + a_{12}^2 + a_{21}^2 + a_{22}^2.$$

Theorem 3. *Let $f : X \to Y$ and $g : Y \to Z$ be functions.*

(a) If f and g are injective (one-one), then $g \circ f$ is injective.

(b) If f and g are surjective (onto), then $g \circ f$ is surjective.

(c) If f and g are bijections, then $g \circ f$ is a bijection.

Proof. **(a) Injective case:**
Assume f and g are injective. Suppose $(g \circ f)(x_1) = (g \circ f)(x_2)$ for some $x_1, x_2 \in X$. Then

$$g(f(x_1)) = g(f(x_2)).$$

Since g is injective, we have

$$f(x_1) = f(x_2).$$

Since f is injective, it follows that

$$x_1 = x_2.$$

Hence, $g \circ f$ is injective.

(b) Surjective case:

Assume f and g are surjective. Let $z \in Z$ be arbitrary. Since g is surjective, there exists $y \in Y$ such that $g(y) = z$. Since f is surjective, there exists $x \in X$ such that $f(x) = y$. Then

$$(g \circ f)(x) = g(f(x)) = g(y) = z.$$

Thus, $g \circ f$ is surjective.

(c) Bijection case:

If f and g are bijections, then by (a) $g \circ f$ is injective, and by (b) $g \circ f$ is surjective. Therefore, $g \circ f$ is bijective. ■

If we suppose $g \circ f$ is surjective, g must necessarily be surjective, but f need not be.

Theorem 4. *Let $f : X \to Y$ and $g : Y \to Z$ be functions.*

(a) If $g \circ f$ is injective, then f is injective.

(b) If $g \circ f$ is surjective, then g is surjective.

Proof. **(a) Injective case:**

Suppose $g \circ f$ is injective. Let $x_1, x_2 \in X$ be such that $f(x_1) = f(x_2)$. Then

$$(g \circ f)(x_1) = g(f(x_1)) = g(f(x_2)) = (g \circ f)(x_2).$$

Since $g \circ f$ is injective, it follows that $x_1 = x_2$. Hence, f is injective.

(b) Surjective case:

Suppose $g \circ f$ is surjective. Let $z \in Z$ be arbitrary. Since $g \circ f$ is surjective, there exists $x \in X$ such that

$$(g \circ f)(x) = g(f(x)) = z.$$

Let $y = f(x) \in Y$. Then $g(y) = z$. Since z was arbitrary, g is surjective. ■

Exercise 0.3. Give examples of functions $f, g : \mathbb{Z} \to \mathbb{Z}$ such that

(a) g is not injective but $g \circ f$ is injective.

(b) f is not surjective but $g \circ f$ is surjective.

Solution:

(a) Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = e^x$, and $g : \mathbb{R} \to \mathbb{R}$ be defined by $g(x) = x^2$. Then the composition $(g \circ f)(x)$ is given by

$$(g \circ f)(x) = g(f(x)) = (e^x)^2 = e^{2x}.$$

Here, $g \circ f$ is injective (one-one) on $\mathbb{R}$ but g is not injective on its entire domain $\mathbb{R}$.

(b) Let $f, g : \mathbb{Z} \to \mathbb{Z}$ be defined by $f(n) = 2n$ and $g(n) = n/2$. Then f is not surjective (odd integers not in range), but $(g \circ f)(n) = g(f(n)) = g(2n) = n$, which is surjective on $\mathbb{Z}$.

—

Exercise 0.4. Let $f : X \to Y$ and $g : Y \to Z$ be such that $g \circ f$ is injective and f is surjective. Is g injective?

Solution: Yes, g must be injective.

Proof. Suppose $g(y_1) = g(y_2)$ for some $y_1, y_2 \in Y$. Since f is surjective, there exist $x_1, x_2 \in X$ such that $f(x_1) = y_1$ and $f(x_2) = y_2$. Then

$$(g \circ f)(x_1) = g(f(x_1)) = g(y_1) = g(y_2) = g(f(x_2)) = (g \circ f)(x_2).$$

Since $g \circ f$ is injective, $x_1 = x_2$, so $y_1 = y_2$. Hence, g is injective. ∎

—

Exercise 0.5. Let $f : X \to X$ be a function such that $f \circ f$ is bijective. Should f be bijective?

Solution: Yes, f must be bijective.

Proof. Since $f \circ f$ is injective, if $f(x_1) = f(x_2)$, then $(f \circ f)(x_1) = (f \circ f)(x_2)$ implies $x_1 = x_2$. Hence, f is injective. Since $f \circ f$ is surjective, for any $y \in X$, there exists $x \in X$ such that $(f \circ f)(x) = y$. Let $z = f(x)$. Then $f(z) = y$, so f is surjective. Being both injective and surjective, f is bijective. ■

—

Exercise 0.6. Suppose $f : X \to Y$ and $g : Y \to X$ are such that $g \circ f = \mathrm{Id}_X$. Show that f is one-one and g is onto.

Proof. **(i)** f **is injective:** Suppose $f(x_1) = f(x_2)$. Applying g to both sides gives

$$g(f(x_1)) = g(f(x_2)) \implies \mathrm{Id}_X(x_1) = \mathrm{Id}_X(x_2) \implies x_1 = x_2.$$

Hence, f is injective.

(ii) g **is surjective:** Let $x \in X$ be arbitrary. Then $x = \mathrm{Id}_X(x) = (g \circ f)(x) = g(f(x))$. Let $y = f(x) \in Y$. Then $g(y) = x$. Since x was arbitrary, g is surjective. ■

Exercise 0.7. Suppose $f : X \to Y$ and $g : Y \to X$ are functions such that

$$g \circ f = \mathrm{Id}_X \quad \text{and} \quad f \circ g = \mathrm{Id}_Y.$$

Show that f and g are bijections.

Proof. **(i)** f **is injective:** Suppose $f(x_1) = f(x_2)$ for $x_1, x_2 \in X$. Applying g gives

$$g(f(x_1)) = g(f(x_2)) \implies \mathrm{Id}_X(x_1) = \mathrm{Id}_X(x_2) \implies x_1 = x_2.$$

Hence, f is injective.

(ii) f **is surjective:** Let $y \in Y$. Then

$$(f \circ g)(y) = \mathrm{Id}_Y(y) = y.$$

Let $x = g(y) \in X$. Then $f(x) = f(g(y)) = y$. Hence, f is surjective. Being both injective and surjective, f is bijective.

(iii) g **is bijective:** Since f has an inverse g, g is also bijective. ■

—

Proposition 5. *For functions $f : X \to Y$, $g : Y \to Z$, and $h : Z \to W$,*

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Proof: Since $g \circ f : X \to Z$ and $h : Z \to W$, we have $h \circ (g \circ f) : X \to W$. Similarly, $(h \circ g) \circ f : X \to W$. Therefore, the two functions have the same domain and codomain.

Let $x \in X$. We need to show that

$$h \circ (g \circ f)(x) = (h \circ g) \circ f(x).$$

Assume $y = f(x) \in Y$, $z = g(y) \in Z$, and $w = h(z) \in W$. Then

$$(g \circ f)(x) = g(f(x)) = g(y) = z, \quad \text{and therefore} \quad h \circ (g \circ f)(x) = h(z) = w.$$

Again,

$$(h \circ g)(y) = h(g(y)) = h(z) = w, \quad \text{and therefore} \quad (h \circ g) \circ f(x) = (h \circ g)(f(x)) = (h \circ g)(y) = w.$$

Thus,

$$h \circ (g \circ f)(x) = w = (h \circ g) \circ f(x).$$

This completes the proof.