Currying

Claim: $\mathrm{Fun}(X \times Y, Z) \cong \mathrm{Fun}(X, \mathrm{Fun}(Y, Z))$

For any three sets X, Y & Z.

proof: Our goal will be to construct a bijection together with an inverse then show that the composite functions (both ways around) result in the identity functions on the appropriate sets.

We first introduce some more notation in order to limit the amount of writing this will take.

Set $A := \mathrm{Fun}(X \times Y, Z)$, $B := \mathrm{Fun}(X, \mathrm{Fun}(Y, Z))$

and $C := \mathrm{Fun}(Y, Z)$. Our goal can be rephrased as saying, we wish to find $\phi : A \longrightarrow B$ and $\psi : B \longrightarrow A$ such that $\psi \circ \phi = 1_A$ and $\phi \circ \psi = 1_B$

That is, we want

$$\begin{tikzcd} a \in A \arrow[rr, "\phi"] \arrow[dr, dashed, "1_A"'] & & B \ni \phi(a) \arrow[dl, "\psi"] \\ & A \ni (\psi \circ \phi)(a) = a & \end{tikzcd} \quad \& \quad \begin{tikzcd} b \in B \arrow[rr, "\psi"] \arrow[dr, dashed, "1_B"'] & & A \ni \psi(b) \arrow[dl, "\phi"] \\ & B \ni (\phi \circ \psi)(b) = b & \end{tikzcd}$$

First, we must <u>define</u> the action of Φ on elements of its domain, and similarly, we will need to decide how the function Ψ is going to work. Only once these two functions have been fully specified, this is to say, we have provided a complete understanding of how each converts an arbitrary input into a specific output value. No function has been defined until this information is given. What confuses this situation is simply the fact that there will be several moments in time where a function we would like to define "takes values in a set of functions," in other words, the codomain of a function we are attempting to specify is itself a collection of functions. When that happens, our original function's definition is incomplete until we are able to describe the action of its output functions on their respective inputs.

For example, Ψ has $A = \mathrm{Fun}(X \times Y, Z)$ as its codomain. Since this is a collection of functions, the complication just mentioned above becomes relevant. Before diving into the math, I will explain how Ψ is supposed to work using English.

$$\Psi : \mathrm{Fun}(X, \mathrm{Fun}(Y, Z)) \longrightarrow \mathrm{Fun}(X \times Y, Z)$$

Ψ begins by taking, as input, a function whose codomain is a set of functions, and it packages all of that information, or "data," into a function whose codomain is simply a set (of abstract elements) as opposed to functions (between two sets of abstract elements) - this is more or less the difference between a function taking values in $\mathrm{Fun}(Y, Z)$ versus just Z itself. Strictly speaking, Z may be a set of functions because we said all three of the sets involved here (X, Y, and Z) were arbitrary sets, but this will not change our analysis of the situation because the interesting part of this bijection deals with the way Ψ manipulates functions relating the elements of X, Y and Z which is independent of the specific nature of those elements. In this sense Ψ simplifies the codomain of functions it receives as input without changing the overall effect of this function. How does it do that? Well, you might say Ψ converts inputs into "Lazy" or "Smart" versions of themselves.

Here is what I mean by that — the point of view taken by a function in the codomain of Ψ is that any potential user of one of the functions in Ψ's domain has the ultimate goal of producing elements of Z. Therefore, these functions refuse to do anything until exactly enough information needed to obtain an element of Z has been given. Consequently, one must specify two inputs at once (one that is an element of X, and another that is an element of Y) before anything happens, whereas before inputting an element of X would have resulted in the creation of a function from Y to Z so that inputting an element of y next would cause an element of Z to "pop out."

To summarize,

Functions in ⇒	Ψ's Domain	Ψ's Codomain
# of times input may be taken	2	1
# of inputs needed each time	1	2

Below is the precise definition of ψ

$$\begin{array}{ccc} \mathrm{Fun}(X, \mathrm{Fun}(Y, Z)) & \xrightarrow{\quad\psi\quad} & \mathrm{Fun}(X \times Y, Z) \\ \rotatebox{90}{$\in$} & & \\ \sigma & \longmapsto & \Big((x,y) \xmapsto{\psi(\sigma)} \sigma(x)(y) \Big) \end{array}$$

A function from X to $\mathrm{Fun}(Y, Z)$. If $x \in X$ then

$$\begin{array}{ccc} \sigma(x) : Y & \longrightarrow & Z \\ \rotatebox{90}{$\in$} & & \rotatebox{90}{$\in$} \\ y & \longmapsto & \sigma(x)(y) \end{array}$$

an ordered pair of elements where $x \in X$ and $y \in Y$

the output of $\sigma(x)$ when it is evaluated at the element $y \in Y$

the name of the function that converts elements $(x,y) \in X \times Y$ into elements

$$\sigma(x)(y) \in Z$$

This completes our description of $\psi : B \longrightarrow A$ in both English & in Math. We now move on to a similar style of describing $\phi : A \longrightarrow B$. The bijection in this direction is slightly more complicated because we run into the issue of sets of functions appearing in the codomain of functions twice as opposed to just once like in the example above. As you can see, the codomain of ϕ is $B := \mathrm{Fun}(X, \mathrm{Fun}(Y, Z))$ which is a set of functions in which every constituent function has a set of functions in its domain. Thus our understanding of Φ will be delayed twice rather than once like the case above where we defined Ψ on functions, but then didn't really understand Ψ's behavior until the output functions had been fully specified.

First, we describe how the bijection Φ works in plain English, then move on to a more precise mathematical description.

Φ converts a function requiring two inputs (one input from the set X and another that belongs to the set Y) to be plugged in AT THE SAME TIME, into a function that takes two inputs ONE AT A TIME. The function that Φ outputs is essentially the same function that was fed into it, the only difference is in the way inputs are accepted. Rather than requiring an ordered pair $(x, y) \in X \times Y$ like the functions one is allowed to input into Φ, the functions Φ "spits out" accept first an element $x \in X$ at which point, this information is converted into another function that accepts elements $y \in Y$ and finally, now that both an x-value and a y-value (in that exact order) have been entered in, an output value - which, in our case must be an element of the set Z, is produced & agrees with the element of Z that would have resulted from sticking (x, y) in to the original 2-variable function. Quite a mouth-full, but the idea is simple

$$\left\{ \begin{array}{c} \text{Functions of} \\ \text{Two variables} \end{array} \right\} \cong \underbrace{\left\{ \begin{array}{c} \text{Families of functions} \\ \text{of a single variable} \end{array} \right\}}$$

the "single variable" mentioned here refers to the "second" variable taken by the "function of two variables" on the other side of this bijection, and the "family of functions" is controlled by the "first" variable on the other side

Here is the formal, mathematical definition of ϕ.

$$\mathrm{Fun}(X \times Y, Z) \xrightarrow{\ \phi\ } \mathrm{Fun}(X, \mathrm{Fun}(Y, Z))$$

$$\begin{array}{ccc} \mathrm{Fun}(X\times Y, Z) & & \\ \rotatebox{90}{$\in$} & & \\ t & \longmapsto & \Big(x \overset{\phi(t)}{\longmapsto} \underbrace{\overset{\phi(t)(x)}{\big(y \longmapsto t(x,y)\big)}}\Big) \end{array}$$

this function takes (x,y) as input & outputs $t(x,y) \in Z$

first input

second input

produces a function of a single variable that "eats" y & "spits out" $t(x,y)$. This function is named $\psi(t)(x)$

When I say "family of functions" above, I mean that for every element $x \in X$, $\phi(t)$ produces an element of $\mathrm{Fun}(Y, Z)$, so every function in the so-called "family" has domain Y and codomain Z. Moreover, there is one member of this family for every element $x \in X$ and the function $\phi(t)$ mediates this correspondence.

Now that both ϕ and ψ have been given definitions, we are able to check that they are mutual inverses. This requires two calculations, we must show that

(i) $\psi \circ \phi = 1_A$ and (ii) $\phi \circ \psi = 1_B$

Remember, two functions are equal if and only if they always output the same thing as one another when the same inputs are entered into both. For instance, this buys us the ability to prove (i) by checking that $\forall\ a \in A$ we have $(\psi \circ \phi)(a) = 1_A(a) = a$ in other words, (I) $\psi(\phi(a)) = a$ for all $a \in A$

Similarly, we will have proven (ii) once we establish the equality $(\phi \circ \psi)(b) = 1_B(b) = b$ for all $b \in B$, or equivalently (II) $\forall b \in B$: $\phi(\psi(b)) = b$.

We begin by proving the first of these two statements

__claim:__ $\psi(\phi(a)) = a$ for all $a \in A$

__proof:__ The fact we are required to prove is an equality of two functions in $A := \mathrm{Fun}(X \times Y, Z)$ so by the remark underlined in red above, it suffices to show that the outputs of these two functions agree for all possible inputs from $X \times Y$. Thus, we may reduce this problem once more. The claim above is equivalent to the statement that $\forall (x,y) \in X \times Y$, we have $\psi(\phi(a))(x,y) = a(x,y)$ no matter what $a \in A$ we happened to have picked.

Notice: $a(x,y) \in Z$, so $\psi(\phi(a))(x,y)$ better be an element of Z also.

Let's see if we can simplify the Left-hand side of the equality $\psi(\phi(a))(x,y)$ in order to (ideally) end with the expression $a(x,y)$.

In order to have fewer symbols on the page to confuse matters, we temporarily set $\alpha := \phi(a)$.

Notice: $\phi: A \longrightarrow B$ so $\phi(a) \in B =$ Domain of ψ

So it makes sense to plug $\phi(a)$ in to ψ.

What is the result of plugging α in to ψ, then successively plugging in (x,y)? well, we've already done the heavy lifting, there is not really any more thought that needs to go in to this problem since we have a formula for the action of the outputs of ψ on (x,y). That formula used σ as the input function from $A := \mathrm{Fun}(X, \mathrm{Fun}(Y, Z))$

So all we need to do is replace all of the σ's from our definition of ψ with α's. In particular, we get $\psi(\alpha)(x,y) = \alpha(x)(y)$ because

$$\begin{array}{ccc} & B & \xrightarrow{\quad\psi\quad} & A \\ & \rotatebox{90}{$\in$} & & \\ \phi(a) =: & \alpha & \longmapsto & \left((x,y) \overset{\psi(\alpha)}{\longmapsto} \alpha(x)(y)\right) \end{array}$$

Now that we have sorted that out, we replace $\phi(a)$ back in for α and we have the expression $\psi(\phi(a))(x,y) = \phi(a)(x)(y)$

because
$$\begin{array}{ccc} B & \xrightarrow{\quad\psi\quad} & A \\ \rotatebox{90}{$\in$} & & \\ \phi(a) & \longmapsto & \left((x,y) \overset{\psi(\phi(a))}{\longmapsto} \phi(a)(x)(y)\right) \end{array}$$

So, if we are able to show that $\underline{\phi(a)(x)(y) = a(x,y)}$ then we are done with the proof of this claim. It is time now to go back to the definition of ϕ so that the equality underlined in purple above is true. Happily, this is exactly what the definition of ϕ was! If you're having trouble seeing exactly how this equation relates to the definition of ϕ

Simply replace t from that definition with a.

$$\begin{array}{l} A \xrightarrow{\ \ \phi\ \ } B \\ a \longmapsto \left(x \overset{\phi(a)}{\longmapsto} \left(y \longmapsto a(x,y)\right)\right) \end{array}$$

$\square$

Remark: the two statements in blue above are doing something called "type checking." However, I won't say more about this — I mention it in case you would like to look it up. But, know that type checking is a very good habit to practice — in fact most mathematical statements become immediately understandable once all of the symbols have been type-checked.

We just proved that $\psi \circ \phi = 1_A$, which implies ϕ is injective & ψ is surjective.

⚠ Why? See if you can prove this implication

We now move on to the second, and final statement that needs to be proven, that is $\phi \circ \psi = 1_B$. Recall that we said this was equivalent to showing

$$(\mathrm{II}) \quad \forall b \in B : \phi(\psi(b)) = b$$

Using the fact that the elements of $B := \mathrm{Fun}(X, \mathrm{Fun}(Y, Z))$ are functions, we can reduce this problem to demonstrating that

$$\forall x \in X \ \forall b \in B \quad \underline{\phi(\psi(b))(x) = b(x)}$$

Q: What kind of object is $b(x)$?

A: well, $b(x) \in \mathrm{Fun}(Y, Z)$

Thus, we are able to show that the equality underlined in green above is true by showing that

$$\forall y \in Y \qquad b(x)(y) = \phi(\psi(b))(x)(y) \in Z$$

no matter what $b \in B$ or $x \in X$ we pick.

Claim: $\forall b \in B \ \forall x \in X \ \forall y \in Y$

$$\phi(\psi(b))(x)(y) = b(x)(y) \quad \text{as elements of } Z.$$

proof: Temporarily set $\psi(b)$ equal to β so that the notation looks a little less scary. Our proof strategy here is to compute, using the definitions of ϕ and ψ, the value of $\phi(\beta)(x)(y)$ so as to arrive at the conclusion that it equals $b(x)(y)$.

By definition of ϕ with β used in place of t we learn

$$A \xrightarrow{\ \phi\ } B := \mathrm{Fun}(X, \mathrm{Fun}(Y,Z))$$

$$\psi(b) =: \underset{\beta}{\rotatebox{90}{$\in$}} \longmapsto \left(x \overset{\phi(\beta)}{\longmapsto} \left(y \overset{\phi(\beta)(x)}{\longmapsto} \beta(x,y) \right) \right)$$

Now, it remains to interpret β for what it really is. We have

$$\phi(\beta)(x) = \underbrace{\phi(\psi(b))(x)}_{\substack{\text{this is a function taking} \\ \text{elements } y \in Y \text{ to } \psi(b)(x,y)}}$$

thus $\phi(\psi(b))(x)(y) = \psi(b)(x,y)$. Now, by definition of ψ we see that

$$\psi(b)(x,y) = b(x)(y) \quad \text{as desired.}$$

But don't take my word for it – double check yourself, by replacing σ from the definition above with $b \in B$.

This completes the proof. $\square$

Consequently, $\phi \circ \psi = 1_B$ so ψ is injective & ϕ is surjective $\xRightarrow[\text{symmetric fact proven earlier}]{\text{combine with}}$ Both are Bijections!