Physics 610 Midterm Fall 2020

Jeremy Welsh-Kavan

Let X and Y be sets.

Lemma 1: Let $f : X \to Y$ and $g : Y \to X$ be mappings such that $f \circ g = \mathrm{id}_Y$. Then f is surjective and g is injective.

Proof of Lemma 1. Let $y \in Y$. Since g is a mapping, there is $x \in X$ such that $g(y) = x$ and, since f is a mapping, $f(x) \in Y$. But $f(x) = f(g(y)) = (f \circ g)(y)$ and, by our assumption, $f \circ g = \mathrm{id}_Y$. So $(f \circ g)(y) = y$. Therefore, for each $y \in Y$ there exists $x \in X$ such that $f(x) = y$, so f is surjective.
Let $y_1, y_2 \in Y$ and assume $g(y_1) = g(y_2)$. Then $f(g(y_1)) = f(g(y_2))$. But, by assumption $f \circ g = \mathrm{id}_Y$, so $y_1 = f(g(y_1)) = f(g(y_2)) = y_2$. Therefore, if $y_1, y_2 \in Y$ and $g(y_1) = g(y_2)$ then $y_1 = y_2$, so g is injective. ∎

Lemma 2: Let $f : X \to Y$ be surjective. Then there exists an injective mapping $g : Y \to X$ such that $f \circ g = \mathrm{id}_Y$.

Proof of Lemma 2. Define a relation $\sim$ on X such that, for $x_1, x_2 \in X$, $x_1 \sim x_2$ if $f(x_1) = f(x_2)$. We can check that this is an equivalence relation. If $x_1, x_2, x_3 \in X$ then $f(x_1) = f(x_1)$, and $f(x_1) = f(x_2)$ implies $f(x_2) = f(x_1)$, so $\sim$ is reflexive and symmetric. If $f(x_1) = f(x_2)$ and $f(x_2) = f(x_3)$ then $f(x_1) = f(x_2) = f(x_3)$ so $f(x_1) = f(x_3)$. Hence, $\sim$ is reflexive, symmetric, and transitive and therefore defines an equivalence relation on X.
For each $y \in Y$, form the set $[x]_y = \{x : f(x) = y\}$. If $x_1, x_2 \in [x]_y$ then $f(x_1) = y = f(x_2)$, so $[x]_y$ are the equivalence classes of $\sim$. Observe that $[x]_y$ is nonempty for each y since f is surjective, and that $[x]_{y_1} \cap [x]_{y_2} = \emptyset$ if $y_1 \neq y_2$ since equivalences classes form a partition of X.
From each $[x]_y$ select exactly one element and call this element x_y.[1] Define a mapping $g : Y \to X$ such that if $y \in Y$ then $g(y) = x_y$. Since $[x]_y$ is nonempty, g maps each $y \in Y$ to some element $x_y \in X$, and since $[x]_y$ are disjoint g maps each element of Y to exactly one element of X. Therefore, $g : Y \to X$ is a true mapping.
Let $y \in Y$. Then $g(y) = x_y$ and, since $x_y \in [x]_y$, $f(x_y) = y$ by the definition of $[x]_y$. So $f(x_y) = f(g(y)) = y$. Therefore, since y was arbitrary, $f \circ g = \mathrm{id}_Y$. By Lemma 1, since f and g are mappings and $f \circ g = \mathrm{id}_Y$, we also have that g is injective. Thus, if $f : X \to Y$ is a surjective mapping then there exists an injective mapping $g : Y \to X$ such that $f \circ g = \mathrm{id}_Y$. ∎

Lemma 3: Let $f : X \to Y$ be bijective. Then there exists a unique bijective mapping $g : Y \to X$ such that $f \circ g = \mathrm{id}_Y$ and $g \circ f = \mathrm{id}_X$.

Proof of Lemma 3. We can proceed as we did in Lemma 2 by defining precisely the same equivalence relation and the same equivalence classes, $[x]_y$. However, since f is bijective, $[x]_y$ contains only one element for every $y \in Y$, so there is only one choice of x_y to make. Therefore, we can define $g : Y \to X$ such that $g(y) = x$ if and only if $f(x) = y$. Since f is surjective g maps each element of Y to an element in X, and since f is injective g maps each element of Y to exactly one element of X. Therefore, $g : Y \to X$ is a true mapping.

---

[1] In the case where there are infinitely many $[x]_y$, I'm pretty sure this requires the axiom of choice.

Let $y \in Y$ such that $f(x) = y$ for some $x \in X$. Then we have $g(y) = x$ and $f(g(y)) = f(x) = y$. Therefore, $f \circ g = \mathrm{id}_Y$.
Now let $x \in X$ such that $g(y) = x$ for some $y \in Y$. Then we have $f(x) = y$ and $g(f(x)) = g(y) = x$. Therefore, $g \circ f = \mathrm{id}_X$.
Let $y_1, y_2 \in Y$ and assume $g(y_1) = g(y_2)$. Then $y_1 = f(g(y_1)) = f(g(y_2)) = y_2$. Therefore, g is injective.
Let $x \in X$. Since f is a mapping there is $y \in Y$ such that $f(x) = y$. This implies that $g(f(x)) = g(y)$, and since $g \circ f = \mathrm{id}_X$, $x = g(y)$. Therefore, g is surjective.
We have shown that g is both injective and surjective. Therefore, g is bijective.
Clearly, g is unique. If it were not then we could find another mapping $\tilde{g} : Y \to X$ such that $\tilde{g}(y) = x$ if and only if $f(x) = y$. But then we would also have $f \circ \tilde{g} = \mathrm{id}_Y$ and $\tilde{g} \circ f = \mathrm{id}_X$. So for each $y \in Y$, we would have $f(\tilde{g}(y)) = y = f(g(y))$. In which case, $g(f(\tilde{g}(y))) = g(f(g(y)))$. But since $g \circ f = \mathrm{id}_X$, we have that $\tilde{g}(y) = g(y)$ for every $y \in Y$. Therefore, $g : Y \to X$ is unique.
Thus, we have shown that if $f : X \to Y$ is a bijective mapping then there exists a unique bijective mapping $g : Y \to X$ such that $f \circ g = \mathrm{id}_Y$ and $g \circ f = \mathrm{id}_X$. ■

Theorem: If $f : X \to Y$ is bijective, then there exists a unique mapping $f^{-1} : Y \to X$ that is also bijective and satisfies:

(i) $f \circ f^{-1} = \mathrm{id}_Y$

(ii) $f^{-1} \circ f = \mathrm{id}_X$

(iii) $(f^{-1})^{-1} = f$

Proof of Theorem. By Lemma 3, if $f : X \to Y$ is a bijective mapping then there exists a unique bijective mapping $f^{-1} : Y \to X$, which we called g in Lemma 3, such that

(i) $f \circ f^{-1} = \mathrm{id}_Y$

(ii) $f^{-1} \circ f = \mathrm{id}_X$

We will call f^{-1} "the inverse of f" and show that the inverse of f^{-1}, denoted $(f^{-1})^{-1}$, is f.
By Lemma 3, f^{-1} is bijective if f is bijective. Therefore, there exists a unique bijective mapping $g : X \to Y$ such that $f^{-1} \circ g = \mathrm{id}_X$ and $g \circ f^{-1} = \mathrm{id}_Y$.
Let $x \in X$. We have that $(f^{-1} \circ f)(x) = x$ since f^{-1} is the inverse of f and $(f^{-1} \circ g)(x) = x$ by Lemma 3. Therefore, $(f^{-1} \circ f)(x) = (f^{-1} \circ g)(x)$ and $f((f^{-1} \circ f)(x)) = f((f^{-1} \circ g)(x))$. But $f \circ f^{-1} = \mathrm{id}_Y$ so by the associativity of function composition, $\mathrm{id}_Y(f(x)) = \mathrm{id}_Y(g(x))$ which implies $f(x) = g(x)$, for all $x \in X$. Therefore, $g = f$. Thus, we have shown that if $f : X \to Y$ is a bijective mapping, the inverse of which is f^{-1}, then the inverse of f^{-1} is precisely f. This proves part (iii) of the theorem. ■