MATH 104-01, SPRING 2006, HANDOUT: ISOMORPHISMS

JOSHUA R. DAVIS

1. Sets

The precise definition of *set* is a foundational issue in mathematics, which we will happily ignore. For our purposes, a *set* is simply a "collection" of "things" (called *elements*), sometimes written in curly braces { and }. Examples include

- the set $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$ of *natural numbers*
- the set $\mathbb{R}$ of real numbers
- the set of people registered to vote in Durham, North Carolina
- the *empty set*, denoted $\emptyset$ or $\{\}$, which is the set containing no elements

The notation $x \in X$ means that x is an element of the set X.

Given two sets X and Y, we say that X is a *subset* of Y iff every element of X is also an element of Y. The empty set is a subset of every set, and any set is a subset of itself. We say that $X = Y$ iff X is a subset of Y and Y is a subset of X. We say that X is a *proper* subset of Y iff X is a subset of Y and $X \neq Y$.

Subset notation sometimes causes students confusion. The notation is roughly analogous to the $<$ notation for numbers, but common usage has $\subset$ being analogous to $\leq$, not $<$; this discrepancy causes there to be three symbols for only two concepts:

Notation	Meaning	Notation	Meaning
$X \subset Y$	X is a subset of Y	$x < y$	x is less than y
$X \subseteq Y$	X is a subset of Y	$x \leq y$	x is less than or equal to y
$X \subsetneq Y$	X is a proper subset of Y	$x \lneq y$	x is less than y

In practice, to prove that X is a subset of Y you typically begin by saying "Let $x \in X$." You then prove that $x \in Y$. Since x was an arbitrary element of X, it follows that every element of X is an element of Y, and thus that $X \subseteq Y$. To prove that $X = Y$, you typically prove both that $X \subseteq Y$ and that $Y \subseteq X$.

If X is a subset of Y, then $Y \setminus X$ denotes the set that results when you remove the elements of X from Y. It is called the *complement* of X in Y.

2. Functions

Given two sets X and Y, the *Cartesian product* $X \times Y$ is the set of all ordered pairs (x, y) where $x \in X$ and $y \in Y$. For example, $\mathbb{R}^2$ is the Cartesian product $\mathbb{R} \times \mathbb{R}$. The adjective "ordered" means that order matters: $(3, 5)$ and $(5, 3)$ are not the same element of $\mathbb{R}^2$.

A *relation* E between X and Y is simply a subset of $X \times Y$. The *inverse relation* E^{-1} is the subset of $Y \times X$ that contains the pair (y, x) if and only if (x, y) is in E.

A *function* f from X to Y is a relation such that for each $x \in X$ there is exactly one $y \in Y$ such that (x, y) is in f. It is common to conceptualize f as a "rule" or "machine" that produces one and only one element $y \in Y$ for each element $x \in X$. The y corresponding to a given x is denoted $f(x)$. Colloquially, we say that "f sends x to y" or that "y is hit by x through (or via, or under) f". The notations $f : X \to Y$ and $X \xrightarrow{f} Y$ denote a function f from X to Y.

When we graph a function $f : \mathbb{R} \to \mathbb{R}$ in calculus, the fact that f is a function means that every vertical line intersects the graph of f exactly once. There is significant abuse of terminology; for example, $y = 1/x$ is not, strictly speaking, a function $\mathbb{R} \to \mathbb{R}$ but rather a function $\mathbb{R} \setminus \{0\} \to \mathbb{R}$.

For any function $f : X \to Y$, we call X the *domain* of f and Y the *codomain* of f. The *image* or *range* of f is the subset of Y consisting of all y that are hit via f; that is, it is the set of $y \in Y$ such that there exists an $x \in X$ such that $f(x) = y$. The image is commonly denoted $\mathrm{im}(f)$ or $f(X)$.

The *identity function* on a set X is the function $\mathrm{id}_X : X \to X$ defined by $\mathrm{id}_X(x) = x$. More generally, if X is a subset of Y, then there is a function $i : X \to Y$ defined by $i(x) = x$; this is called the *inclusion* of X into Y. Inclusions are often written in their own special notation, $i : X \hookrightarrow Y$ or even $i : X \subseteq Y$.

Let $X \xrightarrow{f} Y$ and $Y \xrightarrow{g} Z$ be two functions. The *composition* $g \circ f$ is the function $X \to Z$ defined by $(g \circ f)(x) = g(f(x))$. Notice that $f \circ \mathrm{id}_X = f$ and $\mathrm{id}_Y \circ f = f$.

As defined above, any function $f : X \to Y$ is really a relation $f \subseteq X \times Y$. A calculus student would view the relation f as the graph of the function f. In any event, f has an inverse relation $f^{-1} \subseteq Y \times X$. However, the inverse will not be a function in general. We investigate this problem in the following subsections.

2.1. Injectivity.

A function $f : X \to Y$ is said to be *injective* (or *one-to-one*, or an *injection*) iff, for all $x_1, x_2 \in X$, if $f(x_1) = f(x_2)$ then it must be true that $x_1 = x_2$. That is, two distinct elements x_1 and x_2 cannot be sent to the same element of Y by f.

Any inclusion function $i : X \hookrightarrow Y$ is injective; in fact, the notation $f : X \hookrightarrow Y$ can be used for any injection, to emphasize that it is injective.

When we graph a function $f : \mathbb{R} \to \mathbb{R}$ in calculus, injectivity means that every horizontal line intersects the graph of $y = f(x)$ at most once. The functions x, x^3, $\arctan x$, e^x, and $\sqrt{x}$ are injective; functions 0, x^2, and $\sin x$ are not.

If f is injective, then there exists a function $F : Y \to X$ (not necessarily unique) such that $F(f(x)) = x$ for all $x \in X$. That is, $F \circ f = \mathrm{id}_X$; we say that F is a *left inverse* for f.

In practice, to prove that f is injective you typically begin by saying "Let $x_1, x_2 \in X$ be such that $f(x_1) = f(x_2)$." You then prove that $x_1 = x_2$. Since x_1 and x_2 were arbitrary, this implies that f is injective.

2.2. Surjectivity.

A function $f : X \to Y$ is said to be *surjective* (or *onto*, or a *surjection*) iff for every $y \in Y$ there exists an $x \in X$ such that $f(x) = y$. That is, every element in Y is hit by some element in X via f. In still other words, the image of f equals the codomain of f.

Surjections are sometimes written in the special notation $f : X \twoheadrightarrow Y$ to emphasize that they are surjective.

When we graph a function $f : \mathbb{R} \to \mathbb{R}$ in calculus, surjectivity means that every horizontal line intersects the graph of $y = f(x)$ at least once. The functions x, x^3, $\tan x$, and $\ln x$ are surjective; functions 0, x^2, $\sqrt{x}$ are not.

If f is surjective, then there exists a function $F : Y \to X$ (not necessarily unique) such that $f(F(y)) = y$ for all $y \in Y$. That is, $f \circ F = \mathrm{id}_Y$; we say that F is a *right inverse* for f.

In practice, to prove that f is surjective you typically begin by saying "Let $y \in Y$." You then prove that there is some x satisfying $f(x) = y$; often you do this by explicitly telling the reader how to figure out x from y. For example, $y = \tan x$ is surjective, since any y is hit by $\arctan y$: $\tan(\arctan y) = y$.

2.3. **Bijectivity.** A function $f : X \to Y$ is said to be *bijective* (or a *bijection*, or a *one-to-one correspondence*) iff it is both injective and surjective.

For functions $f : \mathbb{R} \to \mathbb{R}$, bijectivity means that every horizontal line crosses the graph $y = f(x)$ exactly once. The functions x and x^3 are bijective.

If f is bijective, then there is a (unique!) function $F : Y \to X$ such that $F \circ f = \mathrm{id}_X$ and $f \circ F = \mathrm{id}_Y$. This F is the *inverse* for f; it is usually denoted f^{-1}. It agrees with the inverse relation for f, regarded as a relation. In other words, the inverse of a function f is a function if and only if f is bijective.

Notice that if f is a bijection then so is f^{-1}. Also, if $f : X \to Y$ and $g : Y \to Z$ are both bijections, then so is $g \circ f$, and

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1} : Z \to X.$$

2.4. **Restriction.** Let $f : X \to Y$ and $U \subseteq X$. Then the *restriction* $f|_U$ of f to U is the function $f : U \to Y$ defined by $f|_U(x) = f(x)$. That is, $f|_U$ is essentially the same function as f, but defined on a smaller domain.

Suppose that $f : X \to Y$ has image $\mathrm{im}(f) \subseteq Y$. Then we may just as well regard f as a function $f : X \to \mathrm{im}(f)$. (Technically, the old f equals the composition of the inclusion $i : \mathrm{im}(f) \hookrightarrow Y$ with the new f.) This new f is then surjective. If f was already injective, then the new f is a bijection between X and $\mathrm{im}(f)$.

These two procedures — restricting the domain and "restricting the codomain" — can be used to build bijections from non-bijections. For example, consider $\sin : \mathbb{R} \to \mathbb{R}$. This function is neither injective nor surjective. However, if we restrict the domain to the interval $[-\pi/2, \pi/2]$, then it becomes injective. If we then restrict the codomain to $[-1, 1]$, then it becomes surjective. So we end up with a version of sin that is a bijection $[-\pi/2, \pi/2] \to [-1, 1]$. Since it is a bijection, it has an inverse $[-1, 1] \to [-\pi/2, \pi/2]$, which we call $\sin^{-1}$ or arcsin. This $\sin^{-1}$ is an inverse for the restricted sin function, but it is not a true inverse for the original sin function; $\sin \circ \sin^{-1} = \mathrm{id}_{[-1,1]}$, but $\sin^{-1} \circ \sin \neq \mathrm{id}_{\mathbb{R}}$.

3. Vector Space Isomorphisms

Let V and W be any two vector spaces. An *isomorphism* $f : V \to W$ is a bijection such that f and f^{-1} are both linear transformations. The following lemma shows that requiring f^{-1} to be linear is redundant.

Lemma 3.1. *If $f : V \to W$ is a bijective linear transformation, then its inverse $f^{-1} : W \to V$ is also a linear transformation.*

Proof. First let us note that $\vec{w} = f(f^{-1}(\vec{w}))$ for any $\vec{w} \in W$. Then the fact that f^{-1} respects addition follows from the fact that f respects addition:

$$\begin{aligned} f^{-1}(\vec{w_1} + \vec{w_2}) &= f^{-1}(f(f^{-1}(\vec{w_1})) + f(f^{-1}(\vec{w_2}))) \\ &= f^{-1}(f(f^{-1}(\vec{w_1}) + f^{-1}(\vec{w_2}))) \\ &= f^{-1}(\vec{w_1}) + f^{-1}(\vec{w_2}). \end{aligned}$$

Similarly, f^{-1} respects scalar multiplication because f does:

$$\begin{aligned} f^{-1}(c\vec{w}) &= f^{-1}(cf(f^{-1}(\vec{w}))) \\ &= f^{-1}(f(cf^{-1}(\vec{w}))) \\ &= cf^{-1}(\vec{w}). \end{aligned}$$

So we see that f^{-1} is a linear transformation. □

This means that any bijective linear transformation is an isomorphism of vector spaces. It is easy to see that the inverse of an isomorphism is also an isomorphism. The identity function $\mathrm{id}_V : V \to V$ is always an isomorphism. If $V \xrightarrow{f} W$ and $W \xrightarrow{g} U$ are isomorphisms, then so is $g \circ f : V \to U$.

We say that V is *isomorphic* to W iff there exists an isomorphism between them. Intuitively, two isomorphic vector spaces have equivalent structure; they are essentially identical, except that their elements happen to be written using different symbols. The isomorphism tells you how to translate the writing back and forth. An isomorphism between vector spaces is like a dictionary between two languages that differ only in vocabulary, not grammar.

Let V be an n-dimensional vector space and $\mathcal{V} = \{\vec{v_1}, \ldots, \vec{v_n}\}$ a basis for V. Then $\mathcal{V}$ induces a function $C_{\mathcal{V}} : V \to \mathbb{R}^n$ sending each vector $\vec{v}$ to its coordinates in the basis $\mathcal{V}$; that is,

$$C_{\mathcal{V}}(x_1\vec{v_1} + \cdots + x_n\vec{v_n}) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

It turns out that this function $C_{\mathcal{V}}$ is an isomorphism. So after choosing a basis for V we see that V is isomorphic to $\mathbb{R}^n$. It is important to note, however, that a different choice of basis for V induces a different isomorphism. Since there are infinitely many bases for V, there are infinitely many isomorphisms between V and $\mathbb{R}^n$, with no "standard" or "canonical" one that is any better than the others. Any n-dimensional vector space is isomorphic to $\mathbb{R}^n$, but not "canonically".

4. Exercises

1. Prove, for any function $f : X \to Y$, that f can be written as $h \circ g$, where g is a surjection and h is an injection. (First you need to figure out what the domains and codomains of g and h are.)

2. Prove, for any V with basis $\mathcal{V}$, that $C_{\mathcal{V}}$ is an isomorphism, as asserted above.

3. Explicitly describe isomorphisms $f : \mathcal{M}_{2\times 2} \to \mathbb{R}^4$ and $g : \mathcal{P}_3 \to \mathbb{R}^4$. Use these to construct an isomorphism from $\mathcal{M}_{2\times 2}$ to $\mathcal{P}_3$.