Chapter 1. The Real Number System.

1.1. Sets and Functions.

Note. It is impossible to define all objects in mathematics. This is because we can only define new objects in terms of old objects—at some point we must have foundational objects which are known to us through intuition. One such object is a *set* of *elements*. For more on the necessity of undefined terms, see my online notes for Introduction to Modern Geometry (MATH 4157/5157) on Section 1.3. Axiomatic Systems (notice Note 1.3.A).

Note. In this first chapter, we will deal with arbitrary sets, but later we will deal mostly with sets of real numbers. For more details on an axiomatic approach to set theory, see my online notes for Introduction to Set Theory (not a formal ETSU class). A more informal approach to set theory is given in my online notes on Naive Set Theory. These notes are "in preparation" at this time (fall 2023). You were also introduced to set theory in Mathematical Reasoning (MATH 3000), so you are probably familiar with some of these ideas. See my online notes for Mathematical Reasoning; notice the notes for Chapter 2 ("Sets") and Chapter 4 ("Finite and Infinite Sets"). Our approach in these notes are more intuitive than they are axiomatic.

Note 1.1.A. We now argue that there is no "largest set." So, there cannot be a set that contains "everything." In fact, sets can't be "too big." We can classify sets into two categories:

1. Those sets which have themselves as elements: For example, $A = \{1, 2, 3, A\}$, and
2. Those sets which do not have themselves as elements.

We now consider Ω = the set of all sets that fall into the second category. If we ask the question "In which of the two categories does Ω lie?" we are faced with a paradox. This is called *Russell's Paradox.* Russell's Paradox can be more easily explained in the following story. Suppose there is a town with a barber who cuts the hair of everyone who does not cut their own hair. Who cuts the barber's hair? For more details (and some history), see my online notes for Mathematical Reasoning (MATH 3000) on Section 2.2. Russell's Paradox.

Definition. The *null set* or the *empty set* is the set with no elements, denoted $\varnothing$. We denote the statement "x is an element of set A" as $x \in A$ and "x is not an element of set A" as $x \notin A$.

Definition. For a set A, the *complement of* A, denoted A^c, is the set of all elements not in A.

Note 1.1.B. A^c only makes sense when the background *universal set* is known (usually $\mathbb{R}$ for us). That is, we need to know what the set of "all elements" is so that we can distinguish between those elements in A and all those elements not in A.

Note. We can define a set using *set builder* notation: $A = \{x \mid P\}$. A is the set of all elements x (in the universal set) "such that" x satisfies property P. The existence of such a set is justified by the Axiom of Separation (see Section 2.1. Fundamentals in Mathematical Reasoning).

Definition. For sets A and B, A is a *subset* of B if every element of A is also an element of B. In this case, B is called a *superset* of A. This is denoted $A \subset B$ or $B \supset A$. For sets A and B, A *equals* B, denoted $A = B$, if A and B contain exactly the same elements.

Note. Some texts distinguish between *proper subsets* (when $A \subset B$ and $A \neq B$; also denoted $A \subsetneq B$) and *improper subsets* (when $A \subset B$ and possibly $A = B$, denoted $A \subseteq B$ or $B \supseteq A$). These notes follow the notation of Kirkwood and use "$A \subset B$" to indicate that A is a subset of B, whether it is a proper subset or not.

Note 1.1.C. If $A \subset B$ and $B \subset A$, then $A = B$. Therefore, to show the equality of sets A and B, it is sufficient to show (1) $A \subset B$ and (2) $B \subset A$. We will often use this to show the equality of sets.

Definition. We now define operations on the sets A and B.

1. The *union* of A and B, denoted $A \cup B$, is the set of all real elements that are in either A or B.

2. The *intersection* of A and B, denoted $A \cap B$, is the set of all elements in both A and B.

3. The *complement of B relative to A*, denoted $A \setminus B$, is the set of all elements in A but not in B.

Note 1.1.D. We can extend intersections and unions to a collection of *indexed sets*:

$$\cup_{i \in I} A_i = \{x \mid x \in A_i \text{ for some } i \in I\},$$

$$\cap_{i \in I} A_i = \{x \mid x \in A_i \text{ for every } i \in I\}.$$

Here, we are dealing with an arbitrary collection of sets. It is not assumed that the indexing set if finite, nor is it assumed that the indexing set if even countable (more on countable versus uncountable sets is given in Section 1.3. The Completeness Axiom; these topics are also covered in Mathematical Reasoning in Section 4.3. Countable and Uncountable Sets). We cannot say *how many* sets A_i there are, other than in terms of the cardinality of the indexing set: There are $|I|$ sets A_i.

Definition. For sets A and B, if $A \cap B = \varnothing$ then A and B are *disjoint*. A collection of sets $\{A_i \mid i \in I\}$ is *pairwise disjoint* if $A_i \cap A_j = \varnothing$ for all $i, j \in I$ such that $i \neq j$.

Definition. Let $A_1, A_2, \ldots, A_n$ be a finite collection of sets. The *Cartesian product* of these is

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \ldots, a_n) \mid a_i \in A_i \text{ for } i = 1, 2, \ldots, n\}.$$

Note. The idea of a Cartesian product can be extended to an arbitrary collection of sets $\{A_i \mid i \in I\}$, but we do not address this in Analysis 1 (MATH 4217/5217).

Note. We now state our first theorem. Throughout these notes, we use blue fonts to indicate a result for which a proof is given in the "Proofs of Theorems" supplements. Part (a) of our first theorem is given in the supplement and part (b) is to be proved in Exercise 1.1.7(a).

Theorem 1-1. If A, B, and C are sets then

(a) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$,

(b) $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$.

Note. As a corollary to Theorem 1-1, we can prove DeMorgan's Laws. These give an interaction between the operations of union and intersection with the operation of complement.

Corollary 1-1. DeMorgan's Laws.
If B and C are sets (with universal set A) then

(a) $(B \cup C)^c = B^c \cap C^c$,

(b) $(B \cap C)^c = B^c \cup C^c$.

Note. A proof of Corollary 1-1(b) is to be given in Exercise 1.1.7(b). In fact, DeMorgan's Law extends to an arbitrary collection of sets $\{A_i \mid i \in I\}$, as it to be shown in Exercise 1.1.8:

$$(\cap_{i\in I} A_i)^c = \cup_{i\in I} A_i^c \text{ and } (\cup_{i\in I} A_i)^c = \cap_{i\in I} A_i^c.$$

Definition. Let A and B be sets. A *function* f from A to B is a "rule" that associates with each element $x \in A$ a unique element of B, denoted $f(x)$. We write $f : A \to B$.

Note. We are primarily interested in functions $f : A \to \mathbb{R}$, where A is some subset of $\mathbb{R}$. For the sake of an example, we could have $A = \{1, 2, 3\}$ and $B = \{4, 5, 6\}$ with $f(1) = 4$, $f(2) = 4$ and $f(3) = 6$. In this example, two elements of A are both mapped to 4, and no element of A is mapped to $5 \in B$.

Definition. Suppose $f : A \to B$. The set of elements in B that have some point of A mapped into them by f is the *range* of f or the *image* of A under f. The set A is the *domain* of f, denoted $\mathcal{D}(f)$. Two functions f and g are *equal* if $\mathcal{D}(f) = \mathcal{D}(g)$ and $f(x) = g(x)$ for all $x \in \mathcal{D}(f) = \mathcal{D}(g)$.

Note. Symbolically, the range of $f : A \to B$ is

$$\mathcal{R}(f) = f(A) = \{y \mid y \in B \text{ and } y = f(x) \text{ for some } x \in A\}.$$

The domain of f is the set on which f is defined when it is defined.

Definition. $f : A \to B$ is *one-to-one* (sometimes denoted "1–1") if $f(x_1) = f(x_2)$ implies $x_1 = x_2$. f is *onto* B is $y \in B$ implies that there exists $x \in A$ such that $f(x) = y$.

Note 1.1.E. The contrapositive of the definition of one-to-one yields: $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$. Sometimes this is easier to use than the definition itself. As you see in high school math, a one-to-one function mapping a subset of the real numbers into the real numbers (so that it can be graphed in the Cartesian plane) has a graph that satisfies the "horizontal line test." See my online notes for Precalculus 1 (Algebra) (MATH 1710) on Section 5.2. One-to-One Functions; Inverse Functions and notice Theorem 5.2.A. One-to-one functions have an inverse, as we'll see below.

Note. For the function $f : A \to \mathbb{R}$, above where $A = \{1, 2, 3\}$ and $B = \{4, 5, 6\}$, with $f(1) = 4$, $f(2) = 4$ and $f(3) = 6$, we have that f is neither one-to-one (because $f(1) = f(2)$) nor onto (since $5 \in B$ is not the value of $f(x)$ for any $x \in A$). For $g : \mathbb{R} \to [-1, 1]$ defined as $g(x) = \sin x$, we have that g is not one-to-one (because, for example, $\sin x = 0$ for all $x = \pi n$ where $n \in \mathbb{N}$) but it is onto. For $h : \mathbb{R} \to \mathbb{R}$ defined as $h(x) = \tan^{-1} x$, we have that h is one-to-one but (because $\mathcal{R}(h) = (-\pi/2, \pi/2) \neq \mathbb{R}$) is not onto.

Note 1.1.F. Just as real numbers can be added, subtracted, multiplied, and divided (except for division by 0), these algebraic operations can also be applied to real valued functions as follows. We should quickly comment that there is really no such thing as "subtraction" or "division," but only addition and multiplication and their inverses! We'll elaborate on this in Section 1.2. Properties of the Real Numbers as an Ordered Field.

Definition. Suppose f and g are functions from the real numbers to the real numbers. Then

(i) $f \pm g$ is the function whose domain is $\mathcal{D}(f) \cap \mathcal{D}(g)$ and $(f \pm g)(x) = f(x) + g(x)$ for $x \in \mathcal{D}(f) \cap \mathcal{D}(g)$.

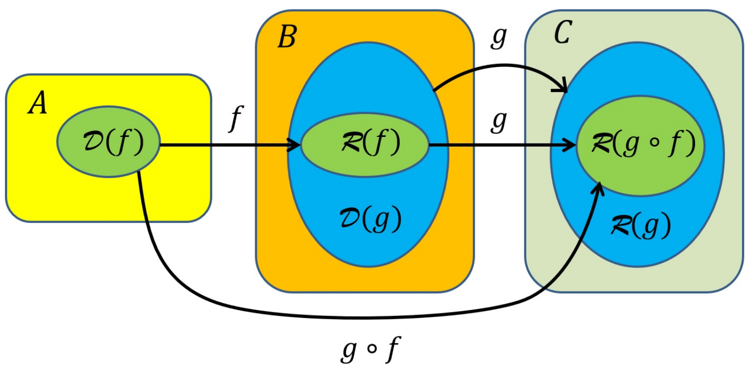
(ii) $f \cdot g$ is the function whose domain is $\mathcal{D}(f) \cap \mathcal{D}(g)$ and $(f \cdot g)(x) = f(x) \cdot g(x)$ for $x \in \mathcal{D}(f) \cap \mathcal{D}(g)$.

(iii) f/g is the function whose domain is $\mathcal{D}(f/g) = \mathcal{D}(f) \cap \mathcal{D}(g) \cap \{x \in \mathcal{D}(g) \mid g(x) \neq 0\}$ and $(f/g)(x) = f(x)/g(x)$ for $x \in \mathcal{D}(f/g)$.

Note. Though not stated in the book, we can also multiply a given function real valued f by a number (or "scalar") r to produce a new function rf with the same domain as f: $(rf)(x) = r(f(x))$ for all $x \in \mathcal{D}(f)$. In this way, we can take linear combinations of functions. For example, with $r, s \in \mathbb{R}$ and real valued functions f and g with common domain $\mathcal{D}(f) = \mathcal{D}(g)$ then the function $rf + sg$ is defined as $(rf + sg)(x) = r(f(x)) + s(g(x))$ for all $x \in \mathcal{D}(f) = \mathcal{D}(g)$. This allows us to make a *vector space* out of all real valued functions with a common domain. This is done in Linear Algebra (MATH 2010). See my online Linear Algebra notes on Section 3.1. Vector Spaces and notice Example 3.1.3.

Note. We can compose general functions (not just those involving real numbers), provided the domains and ranges are "properly related," as given in the next definition. Under appropriate conditions, we will use compositions to define inverse functions below.

Definition. Let A, B, and C be sets and f and g functions such that $\mathcal{D}(f) \subset A$, $\mathcal{F}(g) \subset B$, $\mathcal{R}(f) \subset B$, $\mathcal{R}(f) \subset \mathcal{D}(g)$, and $\mathcal{R}(g) \subset C$. The *composition* of g and f, denoted $g \circ f$, is the function from $\mathcal{D}(f)$ *into* C defined as $(g \circ f)(x) = g(f(x))$.



Example 1.1.5. Let $f(x) = x^2 + 3$ and $g(x) = \sqrt{x-2}$. We assume the domains are maximal subsets of the real numbers and that f and g are real valued. Then $\mathcal{D}(f) = \mathbb{R}$ and $\mathcal{D}(g) = \{x \mid x \in [2, \infty)\}$. Also, $\mathcal{R}(f) = \{x \mid x \in [3, \infty)\}$ and $\mathcal{R}(g) = \{x \mid x \in [2, \infty)\}$. Therefore $\mathcal{R}(f) \subset \mathcal{D}(g)$ so $g \circ f$ is defined with domain $\mathcal{D}(f) = \mathbb{R}$ and

$$(g \circ f)(x) = g(f(x)) = g(x^2 + 3) = \sqrt{(x^2 + 3) - 2} = \sqrt{x^2 + 1}.$$

Also, $\mathcal{R}(g) \subset \mathcal{D}(f)$ so $f \circ g$ is defined with domain $\mathcal{D}(g) = \{x \mid x \in [2, \infty)\}$ and

$$(f \circ g)(x) = f(g(x)) = f(\sqrt{x-2}) = (\sqrt{x-2})^2 + 3 = (x-2) + 3 = x + 1 \text{ for } x \in [2, \infty).$$

Notice that we have to explicitly state "for $x \in [2, \infty)$" as part of the formula for $f \circ g$. The function $h(x) = x+1$ is a function with a different domain from $\mathcal{D}(f \circ g)$ (namely, $h(x) = x+1$ has domain $\mathbb{R}$). Hence, as functions, $(f \circ g) \neq h$ because they have different domains (though, granted, they are equal on $[2, \infty)$, but functions are only defined to be "equal" when they have the same domain, as we saw above).

Note. To paraphrase, the next theorem shows that compositions of functions preserve the properties of one-to-one-ness and onto-ness. The proof of the second part of the theorem is to be given in Exercise 1.1.17.

Theorem 1-2. Suppose $f : A \to B$ and $g : B \to C$, and $g \circ f$ exists.

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

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

Note. We saw in Note 1.1.E that (in the real setting) a one-to-one function satisfies the horizontal line test. If we interchange the roles of x and y in the graph of $y = f(x)$, then the graph of $x = f(y)$ satisfies the vertical line test and so its graph defines a function. This is the reason for the one-to-one hypothesis in the next definition.

Definition. If $f : X \to Y$ is one-to-one, define $f^{-1} : \mathcal{R}(f) \to X$ as $f^{-1}(y) = x$ if $y = f(x)$. f^{-1} is called the *inverse* of f.

Note 1.1.G. With $f : X \to Y$ one-to-one, we have by the previous definition that $y = f(x)$ if and only if $x = f^{-1}(y)$. We have $f : \mathcal{D}(f) \to \mathcal{R}(f)$ and $f^{-1} : \mathcal{R}(f) \to \mathcal{D}(f)$, so that $\mathcal{D}(f^{-1}) = \mathcal{R}(f)$ and $\mathcal{R}(f^{-1}) = \mathcal{D}(f)$. Notice that:

(a) $f^{-1}(f(x)) = x$ for all $x \in \mathcal{D}(f)$, and

(b) $f(f^{-1}(y)) = y$ for all $y \in \mathcal{D}(f^{-1})$.

That is, $f^{-1} \circ f$ is the *identity function* on $\mathcal{D}(f)$, and $f \circ f^{-1}$ is the identity function on $\mathcal{D}(f^{-1})$. In different notation, $(f^{-1} \circ f)(x) = x$ for all $x \in \mathcal{D}(f) = \mathcal{D}(f^{-1} \circ f)$, and $(f \circ f^{-1})(y) = y$ for all $y \in \mathcal{D}(f^{-1}) = \mathcal{D}(f \circ f^{-1})$.

Example 1.1.6. Let $f(x) = x^3 + 1$. This is a one-to-one function (it is strictly increasing by the First Derivative Test, say), so it has an inverse. We claim the inverse is $g(x) = \sqrt[3]{x-1}$. Notice that $\mathcal{D}(f) = \mathcal{D}(g) = \mathcal{R}(f) = \mathcal{R}(g)$. We have

$$(f \circ g)(x) = f(g(x)) = f(\sqrt[3]{x-1}) = (\sqrt[3]{x-1})^3 + 1 = (x-1) + 1 = x$$

and

$$(g \circ f)(x) = g(f(x)) = g(x^3 + 1) = \sqrt[3]{(x^3+1) - 1} = \sqrt[3]{x^3} = x.$$

Since both compositions give the identity function, then $g(x) = f^{-1}(x)$, as claimed.

□

Note. Next we define a special subset of the domain of a function. This set will hold for any function, not just for a one-to-one function. BEWARE the fact that it refers to a set of values and not to a specific value. It does not assume the existence of the inverse function f^{-1}!

Definition. If $f : X \to Y$ and $B \subset Y$, define

$$f^{-1}(B) = \{x \mid x \in X \text{ and } f(x) \in B\}.$$

$f^{-1}(B)$ is called the *inverse image* of B.

Note 1.1.H. I routinely ask the students in my classes (from the pre-freshman level to the graduate level): "What is $\sqrt{9}$?" Given the choices of (a) 3, (b) -3, or (c) ± 3, the popular vote usually goes to ± 3 (at all levels of classes). But the correct answer is $\sqrt{9} = 3$ (as a calculator will tell you). This is because the square root function is a *function* and so can only have one output, and that output is 3. By convention, the square root function always gives the nonnegative square root (on the domain of the square root function, $[0, \infty)$). Sometimes this is called the "principal value" of the square root function (see my online notes for Precalculus 1 (Algebra) [MATH 1710] on Appendix A.1. Algebra Essentials; see Note A.1.E). A similar problem arises even in the complex setting, and a "principal branch" of the square root function (see my online notes for Complex Variables [MATH 4337/5337] on Section 3.33. Complex Exponents). Returning to the $\sqrt{9}$ question, notice that the question was *not*: "What numbers, when squared, give 9?" The answer to this different question *is* ± 3. In terms of functions, the squaring function is not one-to-one (it is two-to-one, except at 0), and so $f(x) = x^2$ had no inverse function. The square root function is the inverse of a different function (namely the squaring function with a restricted domain). That is, the function "$s(x) = x^2$ where $x \in [0, \infty)$" has as its inverse the function $s^{-1}(x) = \sqrt{x}$. Now the discussion of an inverse image does not require a one-to-one function, so we can consider inverse images of the function $f(x) = x^2$ (but this is an inverse image of *sets* and not of numbers). We have $f^{-1}(\{9\}) = \{-3, 3\}$. Had the original question been "what is the inverse image of set $\{9\}$ under the squaring function $f(x) = x^2$," then we could say "the set consisting of ± 3." We conclude this section with some properties images and inverse images under functions. These are given in Exercise 1.1.13.

Exercise 1.1.13. Let $f : X \to Y$ with $A_1, A_2 \subset X$ and $B_1, B_2 \subset Y$. Then

(a) $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$.

(b) $f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2)$.

(c) $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$.

(d) $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$.

(e) $f^{-1}(Y \setminus B_1) = X \setminus f^{-1}(B_1)$.

Exercise 1.1.13(d). Prove $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$.

Note. Notice that Exercise 1.1.13(f) requires an example that $f(A_1 \cap A_2) \neq f(A_1) \cap f(A_2)$. Also, each of parts (a) through (d) of Exercise 1.1.13 can be extended to arbitrary collections of sets (see Note 1.1.D). For example, $f^{-1}\left(\cap_{i \in I} B_i\right) = \cap_{i \in I} f^{-1}(B_i)$.

Revised: 12/4/2023