Foundations of Mathematics Course Notes

Spring 2011

David Lyons
Mathematical Sciences
Lebanon Valley College

Foundations of Mathematics
Course Notes

Spring 2011

David Lyons
Mathematical Sciences
Lebanon Valley College

Contents

The vocabulary of sets and functions is fundamental to all of mathematics, theoretical and applied. We present basic terminology in the first section of these notes.

1 Sets and Functions

1.1 Sets

A ***set*** is a collection of objects. The objects belonging to a set are called its ***elements*** or ***members***. To indicate that an object x is an element of the set A, we write $x \in A$, and pronounce those symbols "x is an element (or member) of A" or "x belongs to A". We write $x \notin A$ to denote that the object x is not an element of the set A.

Sets are specified by listing or describing the elements inside curly braces. For example, we write $A = \{x, y, z\}$ to specify that the set A consists of elements x, y, and z. We write {even whole numbers} or $\{\ldots, -4, -2, 0, 2, 4, \ldots\}$ to specify the set of even whole numbers. The ellipsis symbol "$\ldots$" indicates that the reader should infer an obvious pattern. Order and redundancy of the list inside curly braces are irrelevant. For example, for the set $A = \{x, y, z\}$, we have

$$A = \{x, y, z\} = \{y, z, x\} = \{x, x, y, z\}.$$

The colon symbol or vertical bar inside curly braces denotes the phrase "such that". For example, we may write $\{n^2 : n = 1, 2, 3\}$ or $\{n^2 \mid n = 1, 2, 3\}$ to describe the set $\{1, 4, 9\}$.

It is convenient to have a special set, called the ***empty set***, that contains no members. It plays a role among sets analogous to the role of zero among numbers. The empty set is denoted by an empty pair of curly braces { } or by the symbol $\emptyset$.

We write $A \subseteq B$ or $A \subset B$ to indicate that all the members of set A are also members of set B, and we express this by saying "A is a ***subset*** of B", "A is ***contained in*** B" or "B ***contains*** A". According to this definition, every set is a subset of itself. Less intuitive, but also a consequence of the defintion, is that the empty set is a subset of any other set. Some examples of subsets: the sets $\emptyset$, $\{y\}$, $\{z, x\}$, and $\{x, y, z\}$ are subsets of $\{x, y, z\}$. To indicate that A is a subset of B but not equal to B, we write $A \subsetneq B$ and say A is a ***proper subset*** of B, or A is ***properly contained*** in B.

We write $A \cap B$, pronounced "the ***intersection*** of A with B" or "A intersect B", to denote the set

$$A \cap B = \{x \colon x \in A \text{ and } x \in B\}$$

of all objects that are members both of set A and also of set B. We write $A \cup B$, pronounced "the ***union*** of A with B" or "A union B", to denote the set

$$A \cup B = \{x \colon x \in A \text{ or } x \in B\}$$

of all objects that are members of either set A or of set B or both. An important feature of this definition is that in mathematics, the word "or" is *always* used in the inclusive sense. That is, "or" means "one or the other or both". We

write $A \setminus B$ or $A - B$, pronounced "the ***complement*** of B ***relative to*** A" or "A minus B", to denote the set

$$A \setminus B = \{x \colon x \in A \text{ and } x \notin B\}$$

of all objects which are members of the set A *and are not* members of the set B. Sets A and B are ***disjoint*** or ***mutually exclusive*** if their intersection is the empty set. An example: let $A = \{x, y, z\}$ and let $B = \{a, b, y, z\}$. Then we have $A \cap B = \{y, z\}$, $A \cup B = \{a, b, x, y, z\}$, $A \setminus B = \{x\}$, and $B \setminus A = \{a, b\}$.

An ***ordered pair of elements from the set*** A is an ordered list (x, y) of two elements from A, where we allow the possibility that x equals y. It is important to not confuse ordered pairs with sets containing two elements. For example, in the set $A = \{x, y, z\}$, the symbols (y, x) denote the ordered list with y first and x second, which is different from the ordered pair (x, y). Both of these are different from the two-element set $\{x, y\}$. The set $A \times B$, called the ***(Cartesian) product*** of A and B, is the set of all ordered pairs (a, b) of elements from $A \cup B$ such that $a \in A$ and $b \in B$. Here is an example.

$$\begin{aligned} \{x, y, z\} \times \{a, b, y, z\} \quad = \quad & \{(x, a), (x, b), (x, y), (x, z), \\ & (y, a), (y, b), (y, y), (y, z), \\ & (z, a), (z, b), (z, y), (z, z)\} \end{aligned}$$

A technical consequence of this definition is that for any set A, we have $A \times \emptyset = \emptyset$ because there are *no* ordered pairs (a, b) with $a \in A$ and $b \in \emptyset$. We use the notation A^2 (pronounced "A squared") to denote the product $A \times A$ of a set A with itself. Given a finite collection of sets $A_1, A_2, \ldots, A_n$, the ***n-fold (Cartesian) product*** $A_1 \times A_2 \times \cdots \times A_n$ is the set of all ordered lists, also called ***n-tuples***, of the form $(a_1, a_2, \ldots, a_n)$, where $a_k \in A_k$ for every k in the range $1 \leq k \leq n$. We write A^n to denote the n-fold product of a set A with itself. For example,

$$\begin{aligned} \{a, b\}^3 \quad = \quad & \{(a, a, a), (a, a, b), (a, b, a), (a, b, b), \\ & (b, a, a), (b, a, b), (b, b, a), (b, b, b)\}. \end{aligned}$$

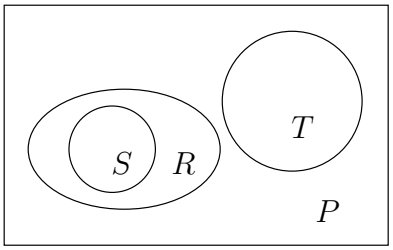

Figure **??**
Euler diagram example

It is often helpful to use pictures to visualize the relationships between sets. ***Euler diagrams*** depict sets as 2-dimensional regions in the plane. Figure **??** shows an Euler diagram illustrating the relationships between the set R of all rectangles, the set S of all squares, the set T of all triangles and the set P of all polygons. A special type of Euler diagram called a ***Venn diagram*** is used to visualize unions, intersections, and complements of sets. Figure **??** shows an example.

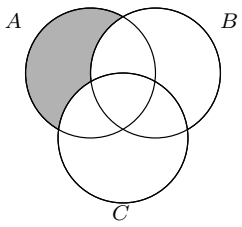

Figure **??**
Venn diagram showing
$A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$

Some important sets

Certain sets are so widely used in mathematics that they have standard names and symbols. One of the most important of these is the set $\mathbf{R}$ of real numbers, which is the set of points on a line. The name "real" indicates the notion that $\mathbf{R}$ is an appropriate set to represent quantities which can be measured in the "real" physical world, such as time, distance, temperature, etc. The set $\mathbf{R}^2 = \mathbf{R} \times \mathbf{R} = \{(x, y) : x, y \in \mathbf{R}\}$ is called the ***x, y-coordinate plane*** or the ***Euclidean plane***, named after Euclid (ca. 300 BC) because it is the

setting for classical plane geometry. The set of ***natural numbers*** is the set $\mathbf{N} = \{1, 2, 3, \ldots\}$ of counting numbers. The ***integers*** or ***whole numbers***, is the set $\mathbf{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$. The ***rational numbers*** or ***fractions***, denoted $\mathbf{Q}$, is the subset of all real numbers which can be written in the form m/n, where m and n are integers and $n \neq 0$.

Notes on terminology

The symbol $\mathbf{Z}$ for the integers comes from the German "Zahlennummern," which means "counting numbers." The symbol $\mathbf{Q}$ for the rationals comes from the word "quotient." The word "rational" comes from the root for *ratio*, meaning proportion. Beware that the notation for an open interval $(a, b) = \{x : a < x < b\}$ of the real line is identical to the notation for the point (a, b) in the x, y-plane; if context does not make clear which is meant, then some additional comment is appropriate on the part of the user.

1.2 Functions

A function is a mathematical model for a process or machine that takes "inputs," does something to them, then produces "outputs." While the idea is not complicated, it is difficult to give a precise definition in everyday language, so some formality is required. The collections of inputs and outputs are modeled by sets. The function itself is modeled by pairs of the form (input value, output value). The machine is not allowed to be ambiguous; for an input value a, there must be exactly one output value b. Here is the formal definition, using the language of sets, that captures this idea.

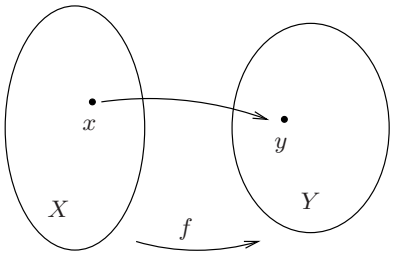

Figure **??**
Schematic diagram of a function

A ***function*** f ***from a set*** X ***to a set*** Y, denoted $f\colon X \to Y$, is a subset of $X \times Y$ in which each element $x \in X$ appears in exactly one ordered pair. That is, if $(x, y) \in f$ and $(x, y') \in f$, then it must be that $y = y'$. We write $f(x) = y$ or $x \mapsto y$ to mean $(x, y) \in f$. The set X is called the ***domain*** of f, and the set Y is called the ***codomain***. The arrows in the symbols $f\colon X \to Y$ and $x \mapsto y$ remind us that the machine takes input $x \in X$ and produces output $y = f(x) \in Y$.

Figure **??** shows a schematic representation of a function $f\colon X \to Y$. We use Euler diagrams for the domain and codomain sets with an arrow labeled f to indicate direction. An arrow from a point x in X to a point y in Y indicates that $f(x)$ is y. Figure **??** shows a version of a commonly used diagram that illustrates the conceptualization of a function as a machine.

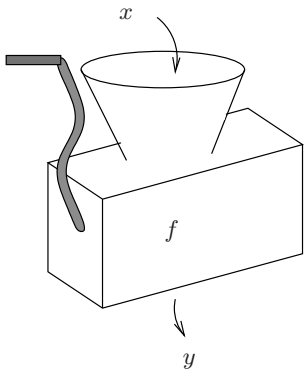

Figure **??**
A function "machine"

Functions are often specified by equations. For example, we write $f(x) = x^2$ or $g(x) = 2x + 3$ to define functions f and g whose domains are sets of real numbers. We often refer to "the function $f(x) = x^2$," or simply, "the function x^2," to mean the function f defined by the equation $f(x) = x^2$. This language carries the potential to confuse the function f with its value $f(x)$; care must be taken when the distinction makes a difference.

Usually, when a function is specified by an equation, the domain is not explicitly given. For example, one might say "the function $h(x) = \sqrt{x - 2}$". The convention in such cases is that the domain is all real x for which the equation specifying the function yields a meaningful real number. In this example, the domain for h is $\{x : 2 \leq x\}$.

The terms ***map*** and ***mapping*** are synonyms for the term *function*. We may also write $X \xrightarrow{f} Y$ to denote the function $f\colon X \to Y$. We write $x \mapsto y$ (pronounced "x goes to y" or "x maps to y") to denote that $f(x) = y$. We refer to $y = f(x)$ as the ***value of*** f ***at*** x or the ***image of*** x ***under*** f. The ***image of a function*** $f\colon X \to Y$ is the subset $\mathrm{Im}(f)$ of the codomain Y given by

$$\mathrm{Im}(f) = \{f(x) : x \in X\}.$$

Note: the term *range of a function* may refer to either (1) the image of the function or (2) the codomain; since range as two meanings, care must be taken to avoid ambiguity.

Given a subset A of the domain X of the function $f\colon X \to Y$, the ***image of*** A is defined to be the set

$$f(A) = \{f(x) : x \in A\}.$$

It is worth noting that the image of f is the same thing as $f(X)$. Given a subset B of the codomain Y, the ***preimage of*** B or the ***inverse image*** of B, is the set

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

For a point $y \in Y$, we write $f^{-1}(y)$ to denote the preimage $f^{-1}(\{y\})$. A function $f\colon X \to Y$ is called ***one-to-one*** or ***injective*** if $f^{-1}(y)$ has no more than 1 element for every $y \in Y$. The function f is called ***onto*** or ***surjective*** if $f^{-1}(y)$ has at least 1 element for every $y \in Y$. The function f is called a ***one-to-one correspondence*** or ***bijective*** if $f^{-1}(y)$ has exactly 1 element for every $y \in Y$, that is, if f is both one-to-one and onto. Examples: consider $f\colon \mathbf{Z} \to \mathbf{Z}$, $g\colon \mathbf{N} \to \mathbf{Z}$, and $h\colon \mathbf{Z} \to \mathbf{Z}$ given by $f(n) = g(n) = n^2$ and $h(n) = -n$. The function f is not one-to-one because the set $f^{-1}(4) = \{-2, 2\}$ has more than one element. The function g is one-to-one because the set $g^{-1}(n)$ is either empty (if n is not a perfect square) or is the 1-element set $\{\sqrt{n}\}$ (if n is a perfect square). Neither f nor g is onto because $f^{-1}(3) = g^{-1}(3) = \emptyset$. The function h is both one-to-one and onto because $h^{-1}(n) = \{-n\}$ for every n in $\mathbf{Z}$.

Given a set X, the ***identity function*** on X is the function $\mathrm{id}\colon X \to X$ given by $x \mapsto x$ for all x in X. A ***constant function*** is a function $f\colon X \to Y$ for which there is an element y_0 in Y such that $f(x) = y_0$ for all x in X.

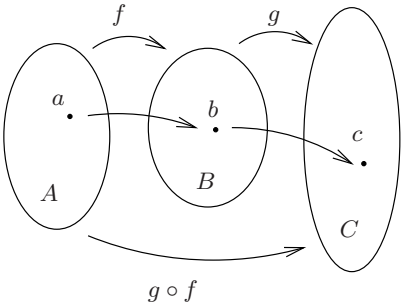

Figure **??**
Composition of functions

Composition

Given two functions $f\colon A \to B$ and $g\colon B \to C$, the ***composition*** $g \circ f$ is the function $g \circ f\colon A \to C$ defined by $(g \circ f)(a) = g(f(a))$. Note that the order matters; $g \circ f$ is not the same as $f \circ g$. Figure **??** shows a schematic diagram.

Inverse functions

Functions operate on input values to produce output values. It is often worthwhile to reverse this process. For example, suppose we have a function f which tells us the dollar value $A = f(t)$ of an investment at time t. A practical problem would be to determine the time value needed to realize a given investment value. In other words, we are seeking a "reversing function," say g, that operates on dollar values and produces the corresponding time values. To say that

g reverses the procedure f is to say $g(A) = t$ whenever $f(t) = A$. For any time value t or dollar value A, we would have the following.

$$g(f(t)) = t \qquad\qquad f(g(A)) = A$$

The above equations say that the composition $g \circ f$ is the identity function on the domain of time values of the function f, and $f \circ g$ is the identity function on the domain of dollar values of the function g. This example motivates the official definition of inverse functions.

(1.2.1) **Definition of Inverse Function.** Suppose $f\colon X \to Y$ and $g\colon Y \to X$ satisfy the equations

$$g \circ f = \mathrm{id}_X \qquad\qquad f \circ g = \mathrm{id}_Y$$

where id_X and id_Y denote the identity functions on X and Y, respectively. Then we say f and g are ***inverses*** of one another, and we write $g = f^{-1}$ and $f = g^{-1}$. The functions f and g are also called ***invertible***.

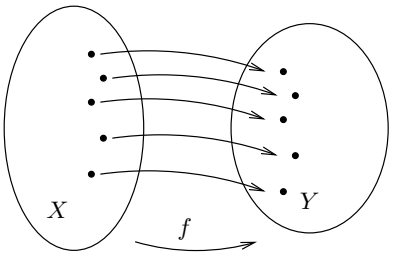

Figure **??**
An invertible function

(1.2.2) **Comment on notation clash, and an important fact about invertible functions.** The alert reader will have noticed that we have now given two different usages of the symbol f^{-1}. We write $f^{-1}(y)$ to denote the *preimage* set of an element y in the codomain, which is defined for *any* function f, and we write $f^{-1}(y)$ to denote the *image* of the element $y \in Y$ under the inverse function for f, which is defined only if f is invertible. Happily, the two meanings have a harmonious resolution when the latter is defined. If the preimage $f^{-1}(\{y\})$ is the 1-element set $\{x\}$, then the image $f^{-1}(y)$ of y under the inverse function f^{-1} is the element x, and vice-versa. It is an important fact that f is invertible if and only if f is one-to-one and onto.

A visual representation of an invertible function $f\colon X \to Y$ (see Figure **??**) shows the assignments made by f as arrows matching the elements of X and Y in a one-to-one manner. The picture of the inverse function f^{-1} is obtained by simply reversing the direction of all the arrows.

(1.2.3) **Examples of inverse functions.** Suppose X is a finite set, and $f\colon X \to Y$ is an invertible function. Since f matches the elements of X with the elements of Y in a one-to-one manner, Y must also be a finite set with the same number of elements as X.
Let $s\colon [0,\infty) \to [0,\infty)$ be the squaring function given by $x \mapsto x^2$. The inverse of s is the square root function $r\colon [0,\infty) \to [0,\infty)$ given by $x \mapsto \sqrt{x}$. Note that the domains are important here. Let $f\colon \mathbf{R} \to \mathbf{R}$ also be the squaring function $x \mapsto x^2$, but on the domain of all reals. The square root function is *not* an inverse for f because $r(f(-2)) = \sqrt{(-2)^2} = 2 \neq -2$. A lesson here is that a function given by an equation may be invertible with one domain, but not invertible with another domain.

Operations on real-valued functions

A ***real-valued function*** is a function whose codomain is a subset of the real numbers. Given two functions $f\colon A \to \mathbf{R}$, $g\colon A \to \mathbf{R}$ and a constant real number k, we define the functions kf, $f+g$, $f-g$, $f \cdot g$ and f/g by the following formulas, for all a in A (note that the last equation is defined only when $g(a) \neq 0$, so the

function f/g is defined to have domain $\{a \in A : g(a) \neq 0\}$).

$$\begin{aligned}(kf)(a) &= kf(a)\\ (f+g)(a) &= f(a)+g(a)\\ (f-g)(a) &= f(a)-g(a)\\ (f\cdot g)(a) &= f(a)g(a)\\ (f/g)(a) &= f(a)/g(a)\end{aligned}$$

Note: The notation fg is sometimes used to mean the product function $f\cdot g$, and sometimes to mean the composition $f \circ g$. Care should be taken when context does not make clear which is meant.

Summation, Intersection, and Union Notation

Let m, n be nonnegative whole numbers with $n \geq m$, and let $f\colon \{m, m+1, \ldots, n\} \to R$ be a function. We write $\sum_{i=m}^{n} f(i)$ to denote the ***sum***

$$\sum_{i=m}^{n} f(i) = f(m) + f(m+1) + \cdots + f(n).$$

The symbol $\sum$ is the capital Greek letter sigma, and denotes a sum. The variable i is called the ***index*** of the sum. More generally, we write $\sum_{i=m}^{n} x_i$ to denote the sum

$$x_m + x_{m+1} + x_{m+2} + \cdots + x_n$$

where m, n are integers with $m \leq n$.

Given a collection $A_1, A_2, \ldots, A_n$ of sets, we write $\bigcap_{i=1}^{n} A_i$, $\bigcup_{i=1}^{n} A_i$, to denote the ***intersection*** and ***union***, respectively, of the sets, defined as follows.

$$\bigcap_{i=1}^{n} A_i = \{x\colon x \in A_i \text{ for all } i, 1 \leq i \leq n\}$$
$$\bigcup_{i=1}^{n} A_i = \{x\colon x \in A_i \text{ for some } i, 1 \leq i \leq n\}$$

1.3 Exercises

1. List all the subsets of the set $X = \{a, b, c, d\}$.
2. Let $A = \{1, 3, 5, 7, 9\}$, let $B = \{2, 3, 4, 5, 6\}$ and let $D = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ be the set of all ten digits. Find $A \cup B$, $A \cap B$, $A \setminus B$, and the $D \setminus A$. Draw a single Euler diagram showing the relationships between A, B, D, and the set $C = \{0, 2, 6\}$.
3. Sketch a Venn diagram for the *symmetric difference* $(A \setminus B) \cup (B \setminus A)$ of two sets A and B for the sets A and B in the previous problem.

4. Use a Venn diagram to illustrate the following property of set operations (one of *DeMorgan's Laws*).

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

5. Let $A = \{a, b\}$ and $B = \{x, y, z\}$.
   (a) List all members of the set $A \times B$.
   (b) List all members of the set B^2.

6. Draw a single Euler diagram illustrating the relationships between the following sets: $\mathbf{R}$, $\mathbf{Q}$, $\mathbf{Z}$, and $[0, 1)$.

7. Let A and B be subsets of the real line $\mathbf{R}$ given by $A = \{x : -2 \leq x < 3\}$ and $B = \{x : 1 < x \leq 5\}$. Write in interval notation, set notation and sketch a picture of A, B, $A \cup B$, $A \cap B$, $A \setminus B$ and $\mathbf{R} \setminus A$. Example: in interval notation, set A is written $A = [-2, 3)$, in set notation $A = \{x : -2 \leq x < 3\}$, and the sketch of A is the shaded region of the real line marked with endpoint brackets at -2 on the left and 3 on the right.

8. Let $f(x) = x^2$ and $g(x) = x + 2$ define functions f and g from the reals to the reals.
   (a) Find $(f \circ g)(3)$
   (b) Find $(g \circ f)(3)$
   (c) Find $(g \cdot f)(3)$
   (d) Find $(f/g)(3)$
   (e) Find $(3f + g)(3)$
   (f) Write an equation for $(f \circ g)(x)$
   (g) Write an equation for $(g \circ f)(x)$

9. Let $X = \{a, b, c\}$ and $Y = \{1, 2, 3\}$. Describe all possible one-to-one correspondences between X and Y.

10. (a) Does there exist a function $f\colon \emptyset \to X$, where X is nonempty? If so, give an example. If not, explain.
    (b) Does there exist a function $f\colon X \to \emptyset$, where X is nonempty? If so, give an example. If not, explain.

11. (a) Evaluate $\displaystyle\sum_{i=1}^{10}(2i + 1)$.
    (b) Write the sum

$$(1^2 + 2 \cdot 1 + 3) + (2^2 + 2 \cdot 2 + 3) + (3^2 + 2 \cdot 3 + 3) + \cdots + (15^2 + 2 \cdot 15 + 3)$$

    using summation notation.

12. Let $A = \{x, y, z\}$ and let $f\colon A \to \mathbf{R}$ and $g\colon A \to \mathbf{R}$ be given by the following table of values.

element of A	value of f	value of g
x	2	5
y	0	1
z	-1	2

Let $s_1 = x, s_2 = y, s_3 = z$ be an ordered list of the elements in A. Find the following.

(a) $\displaystyle\sum_{i=1}^{3} f(s_i)$

(b) $\displaystyle\sum_{i=1}^{2} (f+g)(s_i)$

(c) $\displaystyle\sum_{i=2}^{3} (f \cdot g)(s_i)$

2 Logic

Logic was originally developed to make common sense reasoning more precise for analyzing and writing arguments. While logic is based on common sense, its demand for precision has led to certain usages which sometimes depart from everyday language (for example, see the discussion of the word "or" below). In this subsection we present the basics of logic that are routinely used in mathematics.

2.1 Statements

The fundamental objects in logic are declarative sentences called *statements*. A ***statement*** is an assertion that is either true or false. In the examples below, 1 and 2 are statements. We can agree that 1 is false and 2 is true, so 1 and 2 are examples of statements. While 3 is a valid sentence, it is not considered a statement since we cannot agree whether it is true or false.

1. $2 + 2 = 5$
2. The moon is a satellite of the earth.
3. How I wish I could fly.

We often represent statements by letters p, q etc.

2.2 Negation

Every statement p has an opposite statement, denoted $\neg p$ or $\sim p$, called its ***negation***. If p is true, its negation is false, and if p is false, its negation is true. The negations of statements 1 and 2 above are the following.

1. $2 + 2 \neq 5$
2. The moon is not a satellite of the earth.

Negation can be confusing when a statement is made about some or all members of a set. Here are some examples.

1. Every person in this room works for the fire department.
2. No person in this room works for the fire department.
3. Someone in this room does not work for the fire department.
4. Some people like ice climbing.
5. Everyone does not like ice climbing.
6. No one likes ice climbing.
7. No bird has flown to the moon.
8. All birds have flown to the moon.
9. Some bird has flown to the moon.

Statement 3 is the negation of 1, statement 6 is the negation of 4, and statement 9 is the negation of 7.

2.3 Connectives

Statements can be combined into compound statements using ***connectives***. The four essential connectives are *and*, *or*, *if-then*, and *if and only if*.

And

We write $p \wedge q$ to denote the statement "p and q." The statement $p \wedge q$ is true when both statements p and q are true, and false otherwise.

Or

In everyday English, we use the word "or" in two distinct ways. Sometimes we mean "one or the other but not both". This usage is called the ***exclusive or***. In logic and mathematics, we use the word "or" to mean "one or the other or both."

We write $p \vee q$ to denote the statement "p or q." The statement $p \vee q$ is true when either or both statements p and q are true, and false only when statements p and q are both false.

If-then

The statement "if it is raining, then I am carrying an umbrella" is a compound statement of the form "if p then q," where p is the statement "it is raining" and q is the statement "I am carrying an umbrella."

We write $p \Rightarrow q$ to denote the statement "if p then q" or "p implies q." The statement $p \Rightarrow q$ is true unless p is true and q is false. A statement of this form is called an ***if-then*** statement or an ***implication***.

The statement $q \Rightarrow p$ is called the ***converse*** of the statement $p \Rightarrow q$. It is important to note that an if-then statement is logically distinct from its converse. For example, if p is true and q is false, then the implication $p \Rightarrow q$ is false, but the converse $q \Rightarrow p$ is true.

The statement $(\neg q) \Rightarrow (\neg p)$ is called the ***contrapositive*** of the if-then statement $p \Rightarrow q$. The reader may check (see exercise **??** below) that an implication and its contrapositive have the same truth value for every possible combination of truth values of p and q.

If and only if

We write $p \Leftrightarrow q$ to denote the statement $(p \Rightarrow q) \wedge (q \Rightarrow p)$. In words, we say "$p$ if and only if q". The statement $p \Leftrightarrow q$ is true when p and q are either both true or both false.

We say that p and q are ***logically equivalent*** and write $p \equiv q$ when $p \Leftrightarrow q$ is true. For example, the statement $a \Rightarrow b$ is logically equivalent to the statement $(\neg a) \vee b$ for all statements a and b.

2.4 Order of precedence in notation

When we calculate $4+3\cdot 2$, we follow the convention that multiplication happens before addition so the result is $4+6=10$ and not $7\cdot 2=14$. There are similar conventions for logic notation. When we write $\neg p \wedge q \Rightarrow r$, we mean $((\neg p)\wedge q) \Rightarrow r$. Negation has the highest precedence, and/or connectives have the next level of precedence, if-then and if-and-only-if have the lowest precedence. This is analogous to the minus sign (highest precedence), multiplication (next highest), and addition (lowest precedence) in arithmetic.

2.5 Truth tables

Truth tables provide a convenient method for analyzing compound statements. For example, to describe $p \wedge q$ we make a table of 3 columns, one column for each of the statements p and q involved in the compound statement and one column for the compound statement.

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

There is one row for each of the *logical possibilities*, or possible combinations of truth values, for p and q. The third column gives the truth value for the compound statement for each logical possibility.

Here is how to use a truth table to show that $p \Rightarrow q$ is logically equivalent to $\neg p \vee q$. (Note the order of precedence, as explained above: $\neg p \vee q$ means $(\neg p)\vee q$. If we want to negate $p \vee q$ we must use parentheses and write $\neg(p \vee q)$.)

p	q	$p \Rightarrow q$	$\neg p \vee q$
T	T	T	T
T	F	F	F
F	T	T	T
F	F	T	T

Since the truth values in the last two columns are equal in every row, we conclude that the compound statements are logically equivalent.

2.6 Predicates

Given a set X, a map $p\colon X \to \mathcal{S}$ from X to the set $\mathcal{S}$ of statements is called a ***predicate***. For an example, let X be the set of integers and let $p(x)$ be the expression[1] "x is odd". The truth value of $p(x)$ depends on x. In this example, $p(2)$ is false and $p(3)$ is true.

Given a predicate $p\colon X \to \mathcal{S}$, statements of the form

$$\begin{aligned} A &= \text{"for every } x \text{ in } X,\ p(x) \text{ is true"}, \text{ and} \\ E &= \text{"there exists an } x \text{ in } X \text{ such that } p(x) \text{ is true"} \end{aligned}$$

[1] In orthodox logic, the expression "$p(x)$", where p is a predicate, is not considered to be a statement, on the grounds that a truth value cannot be assigned to $p(x)$ until x is specified.

are so common that there is a special shorthand notation. We write (when the set X is understood)

$$\begin{aligned} A &= \forall x\ p(x) \\ E &= \exists x\ p(x) \end{aligned}$$

where the symbol $\forall$, called the ***universal quantifier***, is pronounced "for all" and the symbol $\exists$, called the ***existential quantifier***, is pronounced "there exists". It is important to understand the following negations of the quantified statements above.

$$\begin{aligned} \neg(\forall x\ p(x)) &= \exists x \neg p(x) \\ \neg(\exists x\ p(x)) &= \forall x \neg p(x) \end{aligned}$$

Note on convention regarding implications involving predicates and a consequence for their negations: If p and q are predicates from a set X to the set of statements, the expression "$p(x) \Rightarrow q(x)$" is interpreted by convention to be the quantified statement "for every $x \in X$, $p(x) \Rightarrow q(x)$". For example, if $p(n)$ is "n is even" and $q(n)$ is "n^2 is even", and X is the set of integers, the expression "if n is even then n^2 is even" means "for every integer n, if n is even, then n^2 is even". The important consequence is that the negation of the statement "$p(x) \Rightarrow q(x)$" is

$$\neg(p(x) \Rightarrow q(x)) \equiv \neg(\forall x p(x) \Rightarrow q(x)) \equiv \exists x \neg(p(x) \Rightarrow q(x)) \equiv \exists x (p(x) \wedge \neg q(x)).$$

For example, the negation of "if n is even, then n^2 is even" is "there is an integer n such that n is even and n^2 is odd".

2.7 Comment on mathematical definitions

Mathematical definitions are commonly phrased using the word "if", as in the following example.

> An integer n is called ***even*** if there is an integer k such that $n = 2k$.

In definitions of this form, it is understood that we actually mean "if and only if".

2.8 Exercises

1. Form the negations of the following statements.
   (a) The cat is hungry.
   (b) All of the cats in the room are hungry.
   (c) Some of the cats in the room are hungry.
   (d) Some of the cats in the room are not hungry.
2. Use truth tables to establish *DeMorgan's Laws*.
   (a) $\neg(p \wedge q)$ is logically equivalent to $\neg p \vee \neg q$
   (b) $\neg(p \vee q)$ is logically equivalent to $\neg p \wedge \neg q$

3. Use a truth table to show that an implication is logically equivalent to its contrapositive.

4. Let p, q and r represent logical statements. Three compound statements are given below. Are any of them logically equivalent? Explain why or why not.

   (i) $(p \wedge \neg q) \Rightarrow \neg r$

   (ii) $\neg r \Rightarrow (p \wedge \neg q)$

   (iii) $r \Rightarrow (q \vee \neg p)$

3 More Core Material

3.1 More basic vocabulary of sets and functions

Power set

Given a set A, the ***power set of*** A, denoted $\mathscr{P}(A)$, is the set of all subsets of A. For example, $\mathscr{P}(\{x, y, z\})$ is given by

$$\mathscr{P}(\{x, y, z\}) = \{\emptyset, \{x\}, \{y\}, \{z\}, \{x, y\}, \{y, z\}, \{z, x\}, \{x, y, z\}\}.$$

Intersections, unions, sums, and products

Let Ω be a set and let $\mathcal{S}$ be a subset of the power set $\mathscr{P}(\Omega)$ of Ω. The intersection of the sets in $\mathcal{S}$ is defined to be

$$\bigcap_{A\in\mathcal{S}} A = \{x \in \Omega : x \in A \text{ for all } A \in \mathcal{S}\}$$

and the union of the sets in $\mathcal{S}$ is

$$\bigcup_{A\in\mathcal{S}} A = \{x \in \Omega : x \in A \text{ for some } A \in \mathcal{S}\}.$$

A consequence of these definitions is that if $\mathcal{S}$ is empty, then $\bigcap_{A\in\mathcal{S}} A = \Omega$ and $\bigcup_{A\in\mathcal{S}} A = \emptyset$.

Given a nonempty finite set A consisting of distinct elements $a_1, a_2, \ldots, a_n$ and a function $f\colon S \to \mathbf{R}$, we write $\sum_{x\in A} f(x)$ to denote the sum $f(a_1) + f(a_2) + \cdots + f(a_n)$, and we write $\prod_{x\in A} f(x)$ to denote the product $f(a_1)f(a_2)\cdots f(a_n)$. If A is empty, we define $\sum_{x\in A} f(x)$ to be zero, and we define $\prod_{x\in A} f(x)$ to be one. Given natural numbers m, n with $m \leq n$ and a function $f\colon I \to \mathbf{R}$, where $I = \{m, m+1, m+2, \ldots, n\}$, we write $\sum_{i=m}^{n} f(i)$ to denote $\sum_{i\in I} f(i)$ and we write $\prod_{i=m}^{n} f(i)$ to denote $\prod_{i\in I} f(i)$.

n-ary Operations on a set

A map $f\colon X \to X$ is called a ***unary operation*** on X. An example is $x \mapsto 1/x$, where X is the set of positive real numbers. A map $f\colon X \times X \to X$ is called a ***binary operation*** on X. An example is $(x, y) \mapsto xy$, where X is the set of real numbers. In general, a map $X^n \to X$ is called an ***n-ary operation*** on X. It is common to use an operation symbol, for example "$*$", to denote a value of a binary operation $f\colon X^2 \to X$ in the following way: we write $x * y$ to denote $f(x, y)$. Examples are $x + y, x \cdot y, x - y$, etc. A binary operation $*$ is

called ***associative*** if $x * (y * z) = (x * y) * z$ for all x, y, z in X, and is called ***commutative*** if $x * y = y * x$ for all x, y in X. For example, $+$ is associative and commutative on the real numbers, but $-$ is neither associative nor commutative.

Sequences

Given a set S, a ***sequence in*** S is a function $f\colon \mathbf{N} \to S$. Given a sequence f in S, we write x_n to denote $f(n)$, and we write (x_n) or $(x_1, x_2, x_3, \ldots)$ to denote the sequence f. It is often convenient to choose 0 instead of 1 for the starting index, so that a sequence is a function $f\colon \{0, 1, 2, \ldots\} \to \mathbf{R}$. Occasionally it is convenient to use a different integer other than 0 or 1 as a starting index.

A sequence in $\mathbf{R}$ of the form $(a, a+d, a+2d, \ldots)$, where a and d are constants, is called an ***arithmetic sequence*** (the stress is on the third syllable when "arithmetic" is used as an adjective). A sequence in $\mathbf{R}$ of the form $(a, ar, ar^2, \ldots)$, where a and r are constants with $a \neq 0$, $r > 0$ and $r \neq 1$, is called a ***geometric sequence***.

3.2 Basic vocabulary related to whole numbers

This section is a collection of vocabulary and facts that are basic and widely used in mathematics.

Given two integers a, b, we say a ***divides*** b, denoted $a|b$, to mean $b = ka$ for some integer k.

Given two natural numbers m, n, the ***greatest common divisor*** of m and n, denoted $\gcd(m, n)$, is the largest natural number d that divides both m and n. The ***least common multiple*** of m and n, denoted $\operatorname{lcm}(m, n)$, is the smallest natural number ℓ such that m and n both divide ℓ.

A natural number $p > 1$ is ***prime*** if $p|(ab)$ implies $p|a$ or $p|b$ for all natural numbers a, b. Here is an important fact that says primes are the fundamental building blocks of natural numbers.

(3.2.1) **Fundamental Theorem of Arithmetic.** Every natural number has a unique prime factorization, up to reordering the factors. This means the following. Let $n > 1$ be a natural number. There exist primes $p_1, p_2, \ldots, p_k$ such that $n = p_1 p_2 \cdots p_k$. If $n = q_1 q_2 \ldots q_r$ is another prime factorization of n, then $k = r$ and there is a relabeling of the indices of the p_j's so that $p_j = q_j$ for $1 \leq j \leq k$ for the reindexed p's.

(3.2.2) **Division Algorithm.** Let m be an integer and let n be a natural number. There exist unique integers q, r such that $m = qn + r$ and r is in the range $0 \leq r < n$. The number q is called the ***quotient*** of m by n, and r is called the ***remainder*** of m ***modulo*** n.

Given two integers m, n and a natural number $p > 1$, we say $m \equiv n \pmod{p}$, pronounced "m is congruent to n modulo p", to mean that $p|(m - n)$. For the special case $p = 2$, the division algorithm says that any integer m either has the form $m = 2k$ or $m = 2k + 1$ for some integer k. In the first case, we say m is an ***even*** number. In the latter, we say m is ***odd***. The evenness or oddness of a number is called its ***parity***.

3.3 Relations

Let X be a set. A ***relation on*** X is a subset S of $X^2 = X \times X$. We say x ***is related to*** y when (x, y) is in S. It is common to use a symbol, such as $\sim$, for the phrase "is related to", and to write $x \sim y$ to denote $(x, y) \in S$.

Examples:

1. Let X be the real numbers and let $\sim$ be $\leq$. Then S is the set $S = \{(x, y) : x \leq y\}$.
2. Let X be the natural numbers, let $p > 1$ be a fixed natural number, and define $x \sim y$ if $x \equiv y \pmod{p}$ for $x, y \in X$.
3. Let X be the set of LVC students, and let $x \sim y$ denote x is a Facebook friend of y for $x, y \in X$.

The relation S on the set X is ***reflexive*** if $x \sim x$ for all $x \in X$. We say S is ***symmetric*** if $x \sim y$ implies $y \sim x$ for all $x, y \in X$. We say S is ***antisymmetric*** if $(x \sim y \wedge y \sim x) \Rightarrow x = y$ for all $x, y \in X$. We say S is ***transitive*** if $(x \sim y \wedge y \sim z) \Rightarrow x \sim z$ for all $x, y, z \in X$. In the examples above, 1 and 2 are reflexive, but 3 is not. 2 and 3 are symmetric, but 1 is not. 1 is antisymmetric, but 2 and 3 are not. 1 and 2 are transitive, but 3 is not.

3.4 Equivalence relations and partitions

An ***equivalence relation*** on a set X is a relation that is reflexive, symmetric, and transitive. Here are two key examples.

(3.4.1) **Equivalence relation examples.**

1. Let $X = \mathbf{N} \times \mathbf{Z}$. We define a relation on X by $(m, n) \sim (p, q)$ if $mq = np$. The reader should check that this relation is reflexive, symmetric, and transitive.
2. Let $p > 1$ be a natural number. Let $X = \mathbf{Z}$ and define $n \sim m$ if $p|(n - m)$. This relation is called ***equivalence modulo*** p.

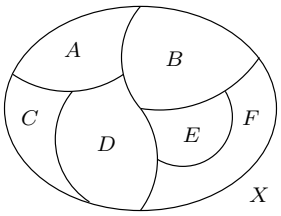

Figure **??**
{A,B,C,D,E,F} is a partition of X

A ***partition*** of a nonempty set X is a collection of nonempty subsets of X whose union is X and which are pairwise disjoint (that is, any two distinct sets in the collection have empty intersection). Figure **??** illustrates the intuitive idea of a partition with an Euler diagram. Here are two key examples.

(3.4.2) **Partition examples.**

1. Let $X = \mathbf{N} \times \mathbf{Z}$. For each rational number q, let L_q denote the graph of the line $y = qx$ in the plane $\mathbf{R}^2$, and let U_q denote the set $U_q = L_q \cap X$. The reader should check that the collection $\{U_q : q \in \mathbf{Q}\}$ is a partition of X. See Figure **??**.
2. Let $p > 1$ be a natural number, and let $X = \mathbf{Z}$. For each integer a in the range $0 \leq a < p$, let U_a be the set

$$U_a = \{\ldots, -3p + a, -2p + a, -p + a, a, p + a, 2p + a, 3p + a, \ldots\}.$$

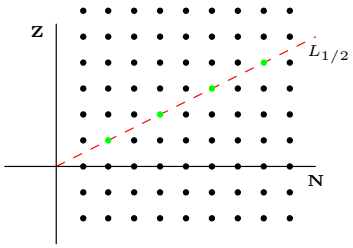

Figure **??**
Points along the dotted line constitute the set $U_{1/2}$ in a partition of $\mathbf{N} \times \mathbf{Z}$

The reader should check that the collection $\{U_a : 0 \leq a < p\}$ is a partition of X.

There is a natural correspondence between relations on a set X and collections of subsets of X, as follows. Given a relation S on X, define the set $U_x \subset X$ for each $x \in X$ by $U_x = \{y \in X : x \sim y\}$. Define a collection $\mathcal{P}(S)$ of subsets of X by $\mathcal{P}(S) = \{U_x : x \in X\}$. For example, for the relations given in (**??**), the corresponding families of sets are given in (**??**). Conversely, given a collection P of subsets of X, define a relation $\mathcal{S}(P)$ on X by $x \sim y$ if and only if x, y lie together in some element of P. For example, for the families of sets given in (**??**), the corresponding relations are given in (**??**). Here is the main fact about this correspondence.

(3.4.3) **Equivalence of Equivalence Relations and Partitions.** If S in an equivalence relation an a set X, then $\mathcal{P}(S)$ is an partition of X. Conversely, if P is a partition of X, then $\mathcal{S}(P)$ is an equivalence relation on X. Further, these correspondences are inverse to one another. That is, $\mathcal{S}(\mathcal{P}(S)) = S$ and $\mathcal{P}(\mathcal{S}(P)) = P$ for all equivalence relations S and partitions P.

Given an equivalence relation $\sim$ on X with associated partition P, we write $X/\!\sim$ to denote P. Given $x \in X$ we call the set $\{y \in X : x \sim y\}$ the ***equivalence class of*** x under the equivalence relation $\sim$, and we call the partition $P = X/\!\sim$ the ***set of equivalence classes***. The equivalence class of an element $x \in X$ is sometimes denoted $[x]$. There is a natural function $\pi\colon X \to X/\!\sim$ given by $x \mapsto [x]$. This map is often called the ***natural projection*** from X to its set of equivalence classes. For example, in (**??**) number 1, the set $X/\!\sim$ of equivalence classes is in one-to-one correspondence with the rational numbers, with the correspondence given by $[(m, n)] \leftrightarrow n/m$, which is the same as $U_q \leftrightarrow q$. In (**??**) number 2, the set $X/\!\sim$ of equivalence classes is in one-to-one correspondence with the set $\{0, 1, 2, \ldots, p-1\}$, where the correspondence is given by $U_a \leftrightarrow a$.

Here is an important theorem about functions on sets of equivalence classes. It's structure is called an ***existence and uniqueness*** theorem.

(3.4.4) **Functions on sets of equivalence classes.** Let X be a set with an equivalence relation $\sim$, and let $\pi\colon X \to X/\!\sim$ be the natural projection. Given a function $f\colon X \to Y$ from X to some set Y, there exists a unique function $g\colon X/\!\sim \to Y$ such that $g \circ \pi = f$ if and only if f is constant on equivalence classes, that is, $f(x) = f(y)$ whenever $x \sim y$.

Proof: First we prove the forward direction of the biconditional statement. Let $f\colon X \to Y$ be given, and suppose there exists a unique function $g\colon X/\!\sim \to Y$ such that $g \circ \pi = f$. If $x \sim y$ then we have $\pi(x) = \pi(y)$, so we have $f(x) = (g \circ \pi)(x) = (g \circ \pi)(y) = f(y)$. (Comment: this direction of the proof uses the existence, but not the uniqueness, of g.)

Next we prove the converse direction of the biconditional. Let $f\colon X \to Y$ be a function with the property that $x \sim y$ implies $f(x) = f(y)$. We define a function $g\colon X/\!\sim \to Y$ by $g(U) = f(x)$, where x is an arbitrary element of U. The assumption that f is constant on U guarantees that the choice of x does not matter. Observe that we have $(g \circ \pi)(x) = g([x]) = f(x)$, as desired. This shows the existence of g. To prove uniqueness, assume there is another function $g'\colon X/\!\sim \to Y$ such that $g' \circ \pi = g \circ \pi = f$. Let $[x]$ be an element of $X/\!\sim$. Then we have $g'([x]) = (g' \circ \pi)(x) = (g \circ \pi)(x) = g([x])$. Since this is true for all $[x]$,

we have $g' = g$. This concludes the proof. ■

Given a function $f\colon X \to Y$ on a set X with an equivalence relation $\sim$, we say that an expression of the form $g([x]) = f(x)$ is ***well-defined*** when f is constant on equivalence classes, and hence the expression $g([x]) = f(x)$ is indeed a legitimate definition for a function $g\colon X/\!\sim \to Y$.

3.5 Exercises

1. Show that parity is an equivalence relation on the integers. That is, define $m \sim n$ if m, n have the same parity. What are the equivalence classes of this relation?

2. Show that both relations in (**??**) are reflexive, symmetric, and transitive.

3. Show that both collections of subsets of X described in (**??**) satisfy the the definition of partition.

4. Give examples of two functions $f\colon \mathbf{N} \times \mathbf{Z} \to \mathbf{R}$, one of which leads to a well-defined function on $X/\!\sim$ (defined in (**??**) number 1) and the other which does not. Explain.

5. Let $f\colon X \to Y$ be a function that is onto. Show that $\{f^{-1}(y) : y \in Y\}$ is a partition of X.

6. Prove this surprisingly useful counting fact.

   (3.5.1) **The Pigeonhole Principle.** Let $f\colon X \to Y$ be a function from a set X with n elements to a set Y with $m < n$ elements. There is an element $y \in Y$ such that the preimage set $f^{-1}(y)$ contains at least two elements.

   Hint for the proof: apply the previous exercise to $\widehat{f}\colon X \to f(X)$, where $\widehat{f}$ is the same as f but with the codomain replaced by the image of f, so that $\widehat{f}$ is onto, then prove the contrapositive of the implication you wish to establish.

4 Proof

A ***proof*** is a logical argument that establishes the truth of a statement. Here is an example.

Proposition. Between every two rational numbers, there is another rational number.

PROOF: Let r and s be rational numbers, and write $r = m/n$ and $s = u/v$ for some integers m, n, u, v. The average $(r + s)/2$ of r and s is between r and s, and the average is

$$\frac{r+s}{2} = \frac{1}{2}\left(\frac{m}{n} + \frac{u}{v}\right) = \frac{vm + un}{2nv}$$

which is clearly rational. ■

The box "■" is a symbol that tells the reader the proof is concluded. It is also common to end a proof with a phrase such as "This concludes the proof" or with the initials "Q.E.D." which stands for the Latin phrase *quod erat demonstrandum* which means "that which was to be shown".

4.1 Direct Proof

Notice that the example proposition above can be rephrased "if r and s are two rational numbers, then there is another rational number between them". That is, the proposition can be rewritten in the form of an implication "if H then C" The proof begins "Let r and s be rational numbers", which can be rephrased as "suppose H is true", and concludes by establishing that C is true. The statement H is called the ***hypothesis*** and C is called the ***conclusion***. This structure of proposition and proof is very common. The proof method is called ***direct proof***. Here is a skeleton summary.

Proposition. $H \Rightarrow C$.

PROOF: **(Direct proof structure)**
Suppose the hypothesis H is true.
Show that C is true.
Conclude that the proposition$H \Rightarrow C$ is true. ■

A variation for proving "$H \Rightarrow C$" is to prove the logically equivalent contrapositive statement "$\neg C \Rightarrow \neg H$". This is sometimes called ***proof by contrapositive*** [2]. Here is an example.

Proposition. If $x + y \geq 100$, then $x \geq 50$ or $y \geq 50$.

PROOF: We prove the contrapositive statement. Suppose $x < 50$ and $y < 50$. Then $x + y < 50 + 50 = 100$. ■

To see why this is so useful, the reader should attempt to write a direct proof without using the contrapositive. It can be done, but the contrapositive proof is shorter and more elegant.

[2] Some authors consider proof by contrapositive to belong to a new category of proof types, called "indirect proof". Other authors consider proof by contrapositive to be a kind of direct proof. The author of these notes does not have a preference.

Direct Proof of Biconditional Statements

A biconditional statement has the form "$P \Leftrightarrow Q$", or in words, "P if and only if Q". A direct proof of a biconditional statement has two sections: first, a direct proof of the conditional "$P \Rightarrow Q$", and; second, a direct proof of the conditional statement "$Q \Rightarrow P$". Here is an example.

Proposition. A function $f\colon X \to Y$ is invertible if and only if it is one-to-one and onto.

PROOF: Suppose that $f\colon X \to Y$ is invertible, with inverse $g\colon Y \to X$. Let y be an element of Y, and let $g(y) = x$. Applying f to both sides, we have $y = f(x)$. This shows that $f^{-1}(\{y\})$ has at least one element, namely x. To show that $f^{-1}(\{y\})$ has at most one element, let $x, x' \in f^{-1}(\{y\})$, so that $f(x) = f(x') = y$. Applying g to both sides, we have $x' = x$. Thus $f^{-1}(\{y\})$ has exactly one element. This is true for all $y \in Y$, so f is a bijection.

Conversely, suppose that f is one-to-one and onto. Define $g\colon Y \to X$ as follows. Let y be a given element in Y. Because f is a bijection, the set $f^{-1}(\{y\})$ is a 1-element set, say $\{x\}$. Define $g(y)$ to be x. Now we check that $f \circ g$ is the identity on Y. Let $y \in Y$ and let $x = g(y)$. By the way we defined x, we have $f(x) = y$. Finally we check that $g \circ f$ is the identity on X. Let $x \in X$, and let $y = f(x)$. Again, by the way we defined g, we have $g(y) = x$. Thus f satisfies the definition of invertible, and the proof is complete. ■

4.2 Proof By Contradiction

Proof by contradiction, also called ***indirect proof*** [3], has the following structure. We seek to prove a proposition P. The proof argues that $\neg P \Rightarrow F$, where F is a false statement (called a ***contradiction***). The statement $\neg P \Rightarrow F$ is logically equivalent to its contrapositive $\neg F \Rightarrow P$. Since $\neg F$ is true, we conclude that P is true. Here is a famous example of proof by contradiction.

Proposition. The real number $\sqrt{2}$ is irrational.

PROOF: Suppose on the contrary that $\sqrt{2}$ is rational, so that there are integers m, n such that $\sqrt{2} = m/n$ and that m, n share no common factors other than 1. Rewriting our assumption, we have $\sqrt{2}n = m$. Squaring both sides yields $2n^2 = m^2$, which tells us that m^2 is an even number. It follows that m itself is an even number (because if m were odd, then m^2 would be odd), so we can write $m = 2k$ for some integer k. Substituting this into our last equation, we have $2n^2 = (2k)^2 = 4k^2$, which reduces to $n^2 = 2k^2$. This tells us that n^2 is an even number, and so n itself is an even number by the same reasoning as before. Since m, n are both even, they share 2 as a common factor. But this contradicts that m, n have no common factors. We conclude that the original assumption that $\sqrt{2}$ is rational is false, and so the proposition is proved. ■

Analysis of the proof. The proposition in this example is an assertion of impossibility. It says that no matter what integers m, n you pick, no fraction m/n can equal $\sqrt{2}$. Let X be the set $X = \mathbf{N} \times \mathbf{N}$, and let $p(x)$ be the assertion "$m/n = \sqrt{2}$" where $x = (m, n)$. In terms of these symbols, the proposition that

[3] Some authors use "proof by contradiction" and "indirect proof" as synonyms. Other authors consider proofs by contradiction to be a subset of indirect proofs. The author of these notes does not have a preference.

$\sqrt{2}$ is irrational is the quantified predicate statement

$$P = \forall x\ \neg p(x).$$

The negation of this statement is

$$\neg P = \exists x\ p(x).$$

The proof establishes that $\neg P \Rightarrow F$ is a true statement, where F is the false statement "integers m, n share no common factors other than 1 and m, n share the common factor 2". If we denote by A the statement "integers m, n share no common factors" then the contradiction F has the form $A \wedge \neg A$. This is common, but not universal, in proofs by contradiction. To summarize, this common proposition and proof structure is the following.

Proposition. $\forall x \neg P(x)$.

Proof: Suppose on the contrary that $\exists x P(x)$ is true.
Show $\exists P(x) \Rightarrow A$ is true.
Show that $\exists x P(x) \Rightarrow \neg A$ is true.
Since $P(x) \Rightarrow (A \wedge \neg A)$, we conclude that $\exists x P(x)$ is false, so the proposition is true. ■

4.3 Mathematical induction

Proof by ***mathematical induction*** refers to a proof that uses the following fact.

(4.3.1) **Principle of Mathematical Induction.**
Let $P(1), P(2), P(3), \ldots$ be an infinite sequence of statements. Suppose that

(i) $P(1)$ is true, and

(ii) for every $k \geq 1$, $P(k) \Rightarrow P(k+1)$ is true.

Then $P(n)$ is true for all $n \in \mathbf{N}$.

This is equivalent to the following.

(4.3.2) **Principle of Mathematical Induction, Version 2.**
Let $P(1), P(2), P(3), \ldots$ be an infinite sequence of statements. Suppose that

(i) $P(1)$ is true, and

(ii) for every $k \geq 1$, $P(1) \wedge P(2) \wedge \cdots \wedge P(k) \Rightarrow P(k+1)$ is true.

Then $P(n)$ is true for all $n \in \mathbf{N}$.

Here is an example of a proof that uses (**??**). First we give a definition: a natural number n is ***reducible*** if it can be factored as a product $n = pq$ of natural numbers p, q such that $p < n$ and $q < n$. A natural number is ***irreducible*** if it is not reducible. [Note: "irreducible" is **not** a synonym for "prime". Can you see why?]

Proposition. Let n be a natural number. Then either n is irreducible, or n is a product of irreducible numbers.

PROOF: The proof is by induction on the number n. For the base case, it is clear that $n = 1$ is irreducible. Now suppose that every natural number n is irreducible or a product of irreducibles for all n in the range $1 \leq n \leq k$, where k is some natural number. If $k+1$ is irreducible, we are done. If not, write $k+1 = pq$ for some natural numbers p, q with $p < k+1$ and $q < k+1$. The inductive hypothesis applies to p and q, so we can write each of them as a product of irreducibles

$$\begin{aligned} p &= p_1 p_2 \cdots p_r \\ q &= q_1 q_2 \cdots q_s \end{aligned}$$

where we take $r = 1$ and $p_1 = p$ if p itself is irreducible, and similarly for s and q_1 in regards to q. Thus we have $n = p_1 \cdots p_r q_1 \cdots q_s$ written as a product of irreducibles, as desired. ■

Analysis of the proof. The proposition has the form "$P(n)$ is true for $n \in \mathbf{N}$", where $P(n)$ is the statement "n is irreducible, or n is a product or irreducibles". The proof begins with the ***base case***, where $P(1)$ is shown to be true. This shows that part (i) of the principle of mathematical induction is satisfied. Next, the ***inductive hypothesis*** is stated. That is the statement "Suppose $P(1) \wedge P(2) \wedge \cdots \wedge P(k)$ is true". Next comes an argument that $P(k+1)$ follows from the inductive hypothesis. The proof is now complete.

4.4 Proof Writing Style

There is a wide range of latitude for style and personality in mathematical writing. There are, however, a few deep-seated elements of mathematical culture that should be observed.

- Personal pronouns: Use first person plural ("we"), **never** first person singular ("I") or the second person ("you").
- Voice: Use the active voice when possible. Use passive voice sparingly. For example, write "the equation has the following solutions ..." instead of "the solutions are found to be ...".
- Abbreviations: We commonly use certain abbreviations during oral discussion, such as WMS ("we must show"), WLOG ("without loss of generality"), and st ("such that"). In general, abbreviations should **not** be used in writing.

4.5 Exercises

1. Give a proof by contradiction that there is no smallest positive real number.
2. Prove DeMorgan's laws for sets: For all sets A, B, and C, we have
   (a) $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$, and
   (b) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$.

3. Given two sets A, B, their ***symmetric difference*** is defined to be

$$A \triangle B = (A \setminus B) \cup (B \setminus A).$$

Does symmetric difference distribute over union or intersection? That is, do we have

(a) $A \triangle (B \cap C) = (A \triangle B) \cap (A \triangle C)$, or
(b) $A \triangle (B \cup C) = (A \triangle B) \cup (A \triangle C)$?

Prove or give counterexamples.

4. The ***exclusive or*** logical connective, denoted "$\oplus$", is defined by $p \oplus q \equiv (p \vee q) \wedge \neg(p \wedge q)$ for statements p, q.

(a) Write a truth table for exclusive or.
(b) Prove or give a counterexample: $(p \oplus q) \Rightarrow r$ is logically equivalent to $p \vee q \vee r$.

5. Let X be a set. Given a subset A of X, define the function $\chi_A \colon X \to \{0, 1\}$ by

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

The function χ_A is called the ***characteristic function*** or the ***indicator function*** for A, since its value signals whether an input value is or is not a member of A.

(a) Show that $\chi_{A \cap B} = \chi_A \chi_B$ for all subsets A, B of X.
(b) Show that $\chi_{A \triangle B} = \chi_A \oplus \chi_B$ for all subsets A, B of X, where $A \triangle B$ is the symmetric difference of A and B (see exercise **??** above) and $\oplus$ denotes addition modulo 2.
(c) Show that $\chi_{A \cup B} = \chi_A \oplus \chi_B \oplus \chi_A \chi_B$ for all subsets A, B of X.
(d) Show that $\chi_{X \setminus A} = \chi_A \oplus 1$ for all subsets A of X.
(e) Show that $\chi_{A \setminus B} = \chi_A(\chi_B \oplus 1)$ for all subsets A, B of X.
(f) Use characteristic functions to show that symmetric difference is associative, that is, for any three sets A, B, C, we have

$$A \triangle (B \triangle C) = (A \triangle B) \triangle C.$$

using characteristic functions.

6. Recall that a *quadratic function* is a function $f \colon \mathbf{R} \to \mathbf{R}$ given by a formula $f(x) = ax^2 + bx + c$ for some $a \neq 0$. Prove that for every pair r, s of nonzero real numbers, there is a unique quadratic function that passes through $(-r, s)$, $(0, 0)$, and (r, s).

7. Let r, d be constants with $r \neq 1$, and let $(a_n) = (a_1, a_2, \ldots)$ be a sequence of real numbers defined by

$$\begin{aligned} a_0 &= 0 \\ a_n &= r a_{n-1} + d \end{aligned}$$

for $n = 1, 2, \ldots$. Use a proof by mathematical induction to show that

$$a_n = d \frac{r^n - 1}{r - 1}$$

for $n = 1, 2, \ldots$.

8. Use mathematical induction to prove the Leibniz Rule for nth order derivatives of products: For functions f, g such that their derivatives $f^{(k)}, g^{(k)}$ exist for $1 \leq k \leq n$, we have

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}.$$

[Recall that by convention the zero-th derivative $f^{(0)}$ is f, and the binomial coefficient $\binom{n}{k}$ is given by

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

for nonnegative integers n, k. You may use the identity

$$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$$

which can be established by algebra.]

9. Prove the equivalence of the two versions of the principle of mathematical induction.

Solutions to Exercises

Note: Most of the "solutions" posted here are not solutions at all, but are merely final answer keys, although some are complete. These are provided so that you can check your work; reading the answer keys is not a substitute for working the problems yourself.

1.3 Sets and Functions Solutions

1.

$$\emptyset, \{a\}, \{b\}, \{c\}, \{d\}, \{a,b\}, \{a,c\}, \{a,d\}, \{b,c\}, \{b,d\}, \{c,d\},$$
$$\{a,b,c\}, \{a,b,d\}, \{a,c,d\}, \{b,c,d\}, \{a,b,c,d\}$$

2.

$$\begin{aligned} A \cup B &= \{1,2,3,4,5,6,7,9\} \\ A \cap B &= \{3,5\} \\ A \setminus B &= \{1,7,9\} \\ D \setminus A &= \{0,2,4,6,8\} \end{aligned}$$

(also show an Euler diagram)

3. (Venn diagram)

4. (Venn diagram)

5. (a) $A \times B = \{(a,x),(a,y),(a,z),(b,x),(b,y),(b,z)\}$
   (b) $B^2 = \{(x,x),(x,y),(x,z),(y,x),(y,y),(y,z),(z,x),(z,y),(z,z)\}$

6. (Euler diagram)

7.

$$\begin{aligned} A &= \{x\colon -2 \le x < 3\} = [-2,3) \\ B &= \{x\colon 1 < x \le 5\} = (1,5] \\ A \cup B &= \{x\colon -2 \le x \le 5\} = [-2,5] \\ A \cap B &= \{x\colon 1 < x < 3\} = (1,3) \\ A \setminus B &= \{x\colon -2 \le x \le 1\} = (-2,1) \\ R \setminus A &= \{x\colon x < -2 \text{ or } 3 \le x\} = (-\infty,-2) \cup [3,\infty) \end{aligned}$$

(also sketch intervals)

8. Let $f(x) = x^2$ and $g(x) = x + 2$ define functions f and g from the reals to the reals.

   (a) $f(g(3)) = 25$
   (b) $g(f(3)) = 11$
   (c) $(g \cdot f)(3) = 45$
   (d) $(f/g)(3) = 9/5$

(e) $(3f + g)(3) = 32$

(f) $f(g(x)) = (x + 2)^2$

(g) $g(f(x)) = x^2 + 2$

9. There are six 1-1 correspondences, indicated below.

abc	abc	abc	abc	abc	abc
123	132	213	231	312	321

10. If either A or B is the empty set, then $A \times B$ is empty because there are no ordered pairs of the form (a, b) with a in A and b in B. Thus any subset of $A \times B$ is empty when one or both of A or B are empty. We check to see if the empty subset of $A \times B$ satisfies the definition of a function in two cases.

    (a) Suppose $A = \emptyset$ and $B = X \neq \emptyset$. We see that every element of A is the first coordinate of exactly one ordered pair of the empty set, so the empty set is indeed a function from A to B.

    (b) Suppose $A = X \neq \emptyset$ and $B = \emptyset$. Let a be an element of A. Since a is not the first coordinate of any ordered pair in the empty set, we conclude that the empty set is *not* a function from A to B, and hence that there are *no* functions from A to B.

11. (a) 120

    (b) $\displaystyle\sum_{i=1}^{15}(i^2 + 2i + 3)$

12. (a) 1

    (b) 8

    (c) -2

2.8 Logic Solutions

1. (a) The cat is not hungry.

    (b) Some of the cats in the room are not hungry.
    or
    There is at least one cat in the room that is not hungry.

    (c) None of the cats in the room are hungry.

    (d) All of the cats in the room are hungry.

2. (a) Here is the truth table.

p	q	$\neg(p \wedge q)$	$\neg p \vee \neg q$
T	T	F	F
T	F	T	T
F	T	T	T
F	F	T	T

    Since the two compound statements have the same truth values for all four logical possibilities, we are done.

(b) (similar to (a))

3. (similar to the previous problem)

4. Statements (i) and (iii) are logically equivalent and (ii) is not equivalent to these two. One way to see this is to observe that (i) and (iii) are contrapositives of one another, while (ii) is the converse of (i). A more basic method is to construct a truth table with eight rows and six columns. There is a row for each of the eight triples $TTT, TTF, TFT, \ldots, FFF$ of logical values for p, q, r, and there is a column for each of the three variables p, q, r and each of the statements (i)–(iii). One observes that the entries match in columns (i) and (iii), and column (ii) has no match.

3.5 More Core Material Solutions

1. We must show that parity is reflexive (R), symmetric (S), and transitive (T).

   (R) Let m be an integer. Since $2 \cdot 0 = 0$, we have $2|(m - m)$, so m is indeed related to itself.

   (S) Let m, n be integers such that $2|(m - n)$. Then there is some integer k such that $m - n = 2k$. It follows that $n - m = 2(-k)$, so $2|(n - m)$. This establishes symmetry.

   (T) Let m, n, p be integers such that $2|(m - n)$ and $2|(n - p)$. Let k, j be integers such that $2k = m - n$ and $2j = n - p$. Then $2(k + j) = (m - n) + (n - p) = m - p$, so $2|(m - p)$. This establishes transitivity.

   The equivalence classes of the parity relation are the following.

$$E = \{\ldots, -4, -2, 0, 2, 4, \ldots\}$$
$$O = \{\ldots, -5, -3, -1, 1, 3, 5, \ldots\}$$

2. We check reflexivity, symmetry, and transitivity for relation 1. The proof for 2 is the same as the proof for the previous problem, replacing 2 by p.

   (R) Let (m, n) be an element in $\mathbf{N} \times \mathbf{Z}$. Since $mn = mn$, we have that (m, n) is related to itself.

   (S) Let $(m, n), (p, q)$ be elements of $\mathbf{N} \times \mathbf{Z}$ such that $mq = np$. Then $np = mq$. This establishes symmetry.

   (T) Let $(m, n), (p, q), (r, s)$ be elements of $\mathbf{N} \times \mathbf{Z}$ such that $mq = np$ and $ps = qr$. Then we have (the product of the left hand sides equals the product of the right hand sides) $mpqs = npqr$. If q is nonzero, dividing by pq yields $ms = nr$. If q is zero, then $n = s = 0$, so $ms = nr$. In both cases we have that (m, n) is related to (r, s), thus establishing transitivity.

3. 1. To see that the union of the sets $\{U_q\}$ is all of X, observe that $(m, n) \in X$ lies in the set $U_{n/m}$. To see that the sets $\{U_q\}$ are pairwise disjoint, suppose that $U_q \cap U_r$ contains some point, say (m, n). Then we have $q = n/m = r$, so $U_q = U_r$. Thus (by contrapositive) if $U_q \neq U_r$, then $U_q \cap U_r = \emptyset$.

2. To see that the union of the sets $\{U_a\}$ is all of X, observe that $m \in X$ lies is the set U_a, where a is the remainder of m after dividing by p, that is, $m = pq + a$ for some integer q and an integer a in the range $0 \leq a < p$. To see that the sets $\{U_a\}$ are pairwise disjoint, suppose that $U_a \cap U_b$ contains some point, say m. Then we have that a, b are both the remainder of m after dividing by p, and hence must be equal by the division algorithm. Thus (by contrapositive) if $U_a \neq U_b$, then $U_a \cap U_b = \emptyset$.

4. Example of well-defined: $f(m, n) = n^2/m^2$. Example of not well-defined: $g(m, n) = \gcd(m, n)$. (Give explanation.)

5. To see that $\bigcup_{y \in Y} f^{-1}(y)$ is all of X, observe that the element $x \in X$ lies in the set $f^{-1}(f(x))$. To see that the inverse image sets are disjoint, suppose that $f^{-1}(y) \cap f^{-1}(y')$ contains some point x. Then we have $y = f(x) = y'$, so $f^{-1}(y) = f^{-1}(y')$. Thus (by contrapositive) if $f^{-1}(y) \neq f^{-1}(y')$, then the two sets are disjoint.

6. Suppose on the contrary that every preimage set $f^{-1}(y)$ has zero or one element. Since the collection $\{f^{-1}(y) \colon y \in f(X)\}$ is a partition of X (by the previous problem), we have

$$\begin{aligned} |X| &= \sum_{y \in f(X)} |f^{-1}(y)| \\ &\leq \sum_{y \in f(X)} 1 \\ &= |f(X)| \\ &\leq |Y|, \end{aligned}$$

but this contradicts the assumption that $|X| > |Y|$. This concludes the proof.

4.5 Proof Solutions

1. Suppose on the contrary that there is a smallest positive real number, say r. Observe that $r/2$ is a positive real number smaller than r. This contradicts the assumption that r is the smallest positive real number. We conclude that there is no smallest positive real number.

2. (a) We will show the desired equality of sets by showing

   ($\subset$) $A \setminus (B \cap C) \subset (A \setminus B) \cup (A \setminus C)$, and
   ($\supset$) $A \setminus (B \cap C) \supset (A \setminus B) \cup (A \setminus C)$.

   ($\subset$) Let x be an element of $A \setminus (B \cap C)$. Since x is not in $B \cap C$, it must be that x is not an element of B or x is not an element of C, or both. Since x is in A, it must be that x is in $A \setminus B$ or x is in $A \setminus C$, or both. Thus x is an element of $(A \setminus B) \cup (A \setminus C)$.

   ($\supset$) Let x be an element of $(A \setminus B) \cup (A \setminus C)$. Since x is in $A \setminus B$ or x is in $A \setminus C$ (or both), it must be that x is not in B or x is not in C (or both), so x is not in $B \cap C$. Since x is in A, we have that x is an element of $A \setminus (B \cap C)$.

   This concludes the proof of (a). A similar proof works for (b).

[Commentary: this proof illustrates a standard method for proving two sets are equal by proving that each one contains the other. The proof naturally divides into two sections. It is common to use subset and superset symbols to label the two sections.]

3. (a) A quick Venn diagram shows that the proposed equality is false. For a counterexample, let $A = \{1, 2\}$, let $B = \{2\}$, and let $C = \emptyset$. The left hand side is $A \triangle (B \cap C) = \{1, 2\}$, and the right hand side is $(A \triangle B) \cap (A \triangle C) = \{1\}$.

   A similar example can be found for (b).

4. (a)

p	q	$p \oplus q$
T	T	F
T	F	T
F	T	T
F	F	F

   (b) The two expressions are not logically equivalent. For example, when p, q, r all have truth value F, $p \oplus q$ is F, so $(p \oplus q) \Rightarrow r$ is T, but $p \vee q \vee r$ is F.

5. (a) We construct a table of values for the various functions involved.

$x \in A$?	$x \in B$?	$\chi_A(x)$	$\chi_B(x)$	$\chi_{A \cap B}(x)$	$\chi_A(x)\chi_B(x)$
T	T	1	1	1	1
T	F	1	0	0	0
F	T	0	1	0	0
F	F	0	0	0	0

   The entries in the two left columns comprise all possible cases. Since the two right columns contain the same entries for each row, the desired equality of functions is established.

   Similar tables establish parts (b)–(d). Part (e) follows from parts (a) and (d).

   (f) Let X be any set that contains A, B, and C. For example, we can let $X = A \cup B \cup C$. Observe that for any two subsets Y, Z of X, we have $Y = Z$ if and only if $\chi_Y = \chi_Z$. We will proof the proposition by applying this observation to $Y = A \triangle (B \triangle C)$ and $Z = (A \triangle B) \triangle C$. We have

$$\begin{aligned}
\chi_Y &= \chi_A \oplus \chi_{B \triangle C} && \text{(by part (b))} \\
&= \chi_A \oplus (\chi_B \oplus \chi_C) && \text{(by part (b))} \\
&= (\chi_A \oplus \chi_B) \oplus \chi_C && \text{(associativity of modular addition)} \\
&= \chi_{A \triangle B} \oplus \chi_C && \text{(by part (b))} \\
&= \chi_Z && \text{(by part (b)).}
\end{aligned}$$

[Commentary: This exercise illustrates the powerful principle that recasting a problem into new language can vastly simplify a proof. A direct proof that uses only the defintion of symmetric difference is lengthy, inelegant,

and painful to write. By contrast, the proof given here that translates the set theory proposition into a statement about characteristic functions is short and direct.]

6. The quadratic function $f(x) = s(x/r)^2$ passes through the three desired points. This establishes existence. If there is another quadratic function $g(x) = ax^2 + bx + c$ that also passes through the three desired points, then we have $0 = g(0) = c$. Further, we must have $ar^2 + br = g(r) = s = g(-r) = ar^2 - br$, which yields $b = 0$. Finally, we solve for a and find $a = s/r^2$, so $g = f$. This establishes uniqueness.

7. Let $P(n)$ be the statement "$a_n = d\frac{r^n-1}{r-1}$" for $n = 1, 2, \ldots$.

   (base case) For $n = 1$, we have $a_1 = r \cdot a_0 + d = d$ by defintion of the sequence (a_n), and we have $d\frac{r^1-1}{r-1} = d$, so $P(1)$ is true.

   (inductive step) Assume that $P(n)$ is true for $n \leq k$ for some $k \geq 1$. We have

$$\begin{aligned} a_{k+1} &= r \cdot a_k + d \quad \text{(by definition of } (a_n)\text{)} \\ &= r \cdot d\frac{r^k - 1}{r - 1} + d \quad \text{(inductive hypothesis)} \\ &= d\left(\frac{r^{k+1} - r}{r - 1} + 1\right) \quad \text{(factoring, distributing)} \\ &= d\left(\frac{r^{k+1} - r + r - 1}{r - 1}\right) \quad \text{(fraction arithmetic)} \\ &= d\frac{r^{k+1} - 1}{r - 1} \end{aligned}$$

   so we see that $P(k + 1)$ is true. QED.

8. Let n be a natural number and let $P(n)$ be the statement that says "given functions f, g whose derivatives exist up to and including order n, we have $(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}$".

   (base case) The statement $P(1)$ is the ordinary Leibniz Rule

$$(fg)' = f'g + fg'$$

   which we know to be true from calculus.

   (inductive step) Now suppose that $P(n)$ is true for $1 \leq n \leq k$, and let f, g be functions whose derivatives exist up to and including order $k + 1$.

Then we have

$$\begin{aligned}
(fg)^{k+1} &= ((fg)^k)' \\
&= \left(\sum_{j=0}^{k} \binom{k}{j} f^{(j)} g^{(k-j)} \right)' \quad \text{(inductive hypoth.)} \\
&= \sum_{j=0}^{k} \binom{k}{j} \left(f^{(j)} g^{(k-j)} \right)' \quad \text{(linearity of deriv.)} \\
&= \sum_{j=0}^{k} \binom{k}{j} \left(f^{(j+1)} g^{(k-j)} + f^{(j)} g^{(k-j+1)} \right) \quad \text{(Leibniz Rule)} \\
&= \sum_{j=0}^{k} \left(\binom{k}{j-1} + \binom{k}{j} \right) f^{(j)} g^{(k+1-j)} \quad \text{(rearranging terms)} \\
&= \sum_{j=0}^{k+1} \binom{k+1}{j} \left(f^{(j)} g^{(k+1-j)} \right) \quad \text{(property of bin. coeff.)}
\end{aligned}$$

so $P(k+1)$ is true. This completes the proof.

9. (Proof idea) To show 1 implies 2, let $Q(k) = P(1) \wedge \cdots \wedge P(k)$ and apply 1 to the sequence $((Q(n))$, and note that $Q(k)$ implies $P(k)$. To show 2 implies 1, simply note that $P(k) \Rightarrow P(k+1)$ implies $P(1) \wedge \cdots \wedge P(k) \Rightarrow P(k+1)$.