7261 1. Set Theory Fall 2012

Sets

I will assume basic set notation including: membership $\in$, subset $\subseteq$, equality $=$ ($A = B$ iff $A \subseteq B$ and $B \subseteq A$), union $\cup$, intersection $\cap$, set difference $A \setminus B = \{x \in A \mid x \notin B\}$, and power set $\mathcal{P}(A) = \{B \mid B \subseteq A\}$. Also the standard sets:

$$\begin{array}{lcllcl} \emptyset & = & \{\} & \mathbb{Q} & = & \{a/b \mid a, b \in \mathbb{Z}, b \neq 0\} \\ \mathbb{N} & = & \{0, 1, 2, \dots\} & \mathbb{R} & = & \{\text{real numbers}\} \\ \mathbb{Z} & = & \{\cdots -2, -1, 0, 1, 2, \dots\} & \mathbb{C} & = & \{\text{complex numbers}\} \end{array}$$

Care should be taken when defining sets since inconsistencies may arise if sets are defined in the form $\{x \mid \text{some property of } x \text{ holds}\}$. For example, if we define

$$S = \{x \mid x \notin x\},$$

then we get a contradiction since $S \in S \iff S \notin S$ (Russell's Paradox). To avoid this, sets should be defined in one of the following ways.

S1. By listing elements explicitly (finite sets only): $\{x, y, z\}$.

S2. As subsets of known sets: $\{x \in S \mid \text{some property of } x \text{ holds}\}$.

S3. By applying some construction to the elements of a known set: $\{F(x) \mid x \in S\}$.

S4. By using one of the operations $\mathcal{P}(S)$ or $\bigcup_{S \in \mathcal{S}} S$.

Ordered pairs

An **ordered pair** is an object (a, b) consisting of two elements a and b, in such a way that order matters, i.e., $(a, b) = (c, d)$ iff $a = c$ and $b = d$, so that in general $(a, b) \neq (b, a)$. There are a number of possible constructions of ordered pairs in terms of sets, the details of which are unimportant. Ordered triples (a, b, c) or n-tuples $(a_1, \dots, a_n)$ can also be defined.

Functions

A **function** (or **map**) $f \colon X \to Y$ assigns for each $x \in X$ a value $f(x) \in Y$. A function can be defined more precisely as an ordered triple (X, Y, Γ) consisting of the **domain** X, the **co-domain** Y, and the **graph** Γ of f. The graph is a set of ordered pairs (x, y) with the property that for each $x \in X$ there is precisely one value of $y \in Y$ such that $(x, y) \in \Gamma$. This y is then written as $f(x)$. Sometimes we also write $x \mapsto y$. The **image** or **range** of a map $f \colon X \to Y$ is the set of values taken: $\operatorname{Im} f = \{f(x) \mid x \in X\} \subseteq Y$. Note that the definition of a function includes the domain X and co-domain Y — for example $x \mapsto x^2$ is considered to be many different functions depending on whether we consider it as a function $\mathbb{R} \to \mathbb{R}$, or $\mathbb{R} \to [0, \infty)$, or $[0, \infty) \to [0, \infty)$, etc..

In order to define a function it is important to check that

F1. The function produces a value *in* Y for *every* $x \in X$.

F2. The function is **well defined**, in that each $x \in X$ gives only *one* value $y \in Y$.

If $f\colon X \to Y$ and $g\colon Y \to Z$ are two maps then the **composition** $g \circ f$ (or gf) is the map $c\colon X \to Z$ defined by $c(x) = g(f(x))$. Note that composition is always **associative** when defined: if in addition $h\colon Z \to W$ then $h \circ (g \circ f) = (h \circ g) \circ f$.

A map $f\colon X \to Y$ is **injective** (or **1-to-1**) if $f(x) = f(x')$ implies $x = x'$, i.e., each value $y \in Y$ occurs *at most* once as a value of f.

A map $f\colon X \to Y$ is **surjective** (or **onto**) if for every $y \in Y$ there is some x with $f(x) = y$, i.e., each value $y \in Y$ occurs *at least* once as a value of f.

A map $f\colon X \to Y$ is **bijective** if it is both injective and surjective, i.e., each value occurs *exactly* once, and so the function pairs up elements of X with elements of Y.

The map $f\colon X \to Y$ is a bijection if and only if it has a **two-sided inverse** $f^{-1}\colon Y \to X$, i.e., $f^{-1} \circ f = 1_X$, $f \circ f^{-1} = 1_Y$, where $1_X(x) = x$ and $1_Y(y) = y$ are the **identity maps** on X and Y.

If a bijection $f\colon X \to Y$ exists then X and Y have the same **cardinality**, $|X| = |Y|$ (this defines cardinality!). If there is a bijection from X to $\{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$ then X is **finite** and we write $|X| = n$. Otherwise X is **infinite**. Among infinite sets, sets that have a bijection to $\mathbb{N}$ are called **countable**, and those that do not are called **uncountable**.

A bijection from a set X to itself is called a **permutation** of X. We write S_X for the set of permutations of X. Note that if $|X| = |Y| < \infty$ and $f\colon X \to Y$ then

$$f \text{ is bijective } \iff f \text{ is injective } \iff f \text{ is surjective.}$$

If $A \subseteq Y$ then the **inclusion map** $i\colon A \to Y$ is the function defined by $i(x) = x$ for all $x \in A$. The identity map 1_X is just the inclusion $X \to X$. If $f\colon Y \to Z$ is any function then the **restriction** of f to A is $f_{|A} = f \circ i\colon A \to Z$, so $f_{|A}(x) = f(x)$ for all $x \in A$.

A function $a\colon \mathbb{N} \to X$ is also called a **sequence** and $a(i)$ is then often written as a_i.

Cartesian Products

The **Cartesian Product** $X \times Y$ is the set of all ordered pairs (x, y) with $x \in X$, $y \in Y$.

More generally if X_i are sets for all $i \in I$, then the Cartesian product $\prod_{i \in I} X_i$ can be defined as the set of functions $I \to \bigcup X_i$, $i \mapsto x_i$ with $x_i \in X_i$ for all i. This definition works for any set I including infinite or even uncountable sets. If $I = \{1, \ldots, n\}$ we can identify this with the set of all ordered n-tuples $(x_1, \ldots, x_n)$ with $x_i \in X_i$.

A function of two (or more) variables can be represented as a map from a Cartesian product. For example, if a function $f(x, y)$ takes values $x \in X$ and $y \in Y$ and produces $f(x, y) \in Z$, then f can be considered as a map $f\colon X \times Y \to Z$ with $f(x, y)$ a shorthand for $f((x, y))$.

The **Axiom of Choice** states that if $X_i \neq \emptyset$ for all $i \in I$ then $\prod_{i \in I} X_i \neq \emptyset$. In other words, there is a map $i \mapsto x_i$ with $x_i \in X_i$ for all $i \in I$. This map simultaneously 'chooses' an element x_i from each X_i.

7261 2. Equivalence relations Fall 2012

A **binary relation** on X is a subset $\Gamma \subseteq X \times X$. We write $x\Gamma y$ (or say that $x\Gamma y$ **holds**) iff $(x, y) \in \Gamma$. For example, we can define '$<$' on $X = \{1, 2, 3\}$ as the set $\{(1, 2), (1, 3), (2, 3)\}$.

(If we wish to be more formal, we could call Γ the **graph** of the relation, and define the relation as a triple (X, X, Γ) analogously to the definition of a function. More generally, we could define a relation on $X \times Y$, in which case a "function" would be a special case of a relation. However, functions and relations are used in very different ways, so perhaps this is not such a good way of thinking about them.)

An **equivalence relation** on X is a binary relation $\sim$ such that for all $x, y, z \in X$

E1. $x \sim x$ (reflexivity)

E2. $x \sim y \Rightarrow y \sim x$ (symmetry)

E3. ($x \sim y$ and $y \sim z$) $\Rightarrow x \sim z$ (transitivity)

Important example: If $f\colon X \to Z$ is any map then the relation defined by $x \sim y$ iff $f(x) = f(y)$ is an equivalence relation on X.

The **equivalence class** of x is $\bar{x} = \{y \mid y \sim x\}$.

Theorem 2.1 *The equivalence classes form a* **partition** *of X, i.e., X is the disjoint union of equivalence classes, each of which is non-empty. Conversely, any partition gives rise to an equivalence relation for which the partitions are the equivalence classes.*

Proof. Classes non-empty: $x \sim x$ (reflexivity), so $x \in \bar{x}$ and thus $\bar{x} \neq \emptyset$.

Union is X: Clearly $\{x\} \subseteq \bar{x} \subseteq X$. Taking unions over $x \in X$ gives $X \subseteq \bigcup_x \bar{x} \subseteq X$.

Classes disjoint: If $\bar{x} \cap \bar{y} \neq \emptyset$ then pick $z \in \bar{x} \cap \bar{y}$. Then $z \sim x$, $z \sim y$. If $t \in \bar{x}$ then $t \sim x$, $x \sim z$ (symmetry), so $t \sim z$ (transitivity), so $t \sim y$ (transitivity), so $t \in \bar{y}$. Thus $\bar{x} \subseteq \bar{y}$. Similarly $\bar{y} \subseteq \bar{x}$, so $\bar{x} = \bar{y}$ is just one equivalence class.

Conversely: Let $X = \bigcup_{i \in I} X_i$ be a partition. Define $x \sim y$ iff $x, y \in X_i$ for some $i \in I$. E1: every x lies in some X_i, and then $x, x \in X_i$. E2: if $x, y \in X_i$ then $y, x \in X_i$. E3: if $x, y \in X_i$ and $y, z \in X_j$, then $i = j$ (since the X_i are disjoint and both X_i and X_j contain y), and so $x, z \in X_i$. Finally, each X_i is non-empty. Pick $x \in X_i$. Now $x \notin X_j$ for $j \neq i$, and so $\bar{x} = \{y \mid y \in X_i\} = X_i$, so X_i is an equivalence class. □

Define the **quotient** $X/\!\sim$ as the set $\{\bar{x} \mid x \in X\}$ of equivalence classes, and the **quotient map** $\pi\colon X \to X/\!\sim$ by $\pi(x) = \bar{x}$. Note that π is always surjective and $x \sim y$ iff $\pi(x) = \pi(y)$ (so the important example above is in fact completely general).

Factoring maps

Lemma 2.2 *Suppose $g\colon X \to Z$ is surjective and $f\colon X \to Y$ is any map. Then there exists $h\colon Z \to Y$ with $f = h \circ g$ iff for all $x, x' \in X$, $g(x) = g(x')$ implies $f(x) = f(x')$. Morover h is unique.*

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle g}\downarrow & \nearrow_{h} & \\ Z & & \end{array}$$

Important special case: If you wish to define a map $X/\!\sim \to Y$ it is usually easier to construct a map $f\colon X \to Y$, check that $x \sim y$ implies $f(x) = f(y)$, and apply Lemma 2.2 with $g = \pi$.

Proof. $\Rightarrow$: If $g(x) = g(x')$ then $f(x) = h(g(x)) = h(g(x') = f(x')$.
$\Leftarrow$: For $z \in Z$, pick $x \in X$ with $g(x) = z$ (possible as g is surjective). Define $h(z) = f(x)$. This gives a value in Y for each $z \in Z$. To show h is well defined, suppose $x' \in X$ with $g(x') = z$. Then $g(x) = g(x')$ so $f(x) = f(x')$ and we obtain the same value for $h(z)$.
Uniqueness: If h, h' are two such maps, write $z = g(x)$, then $h(z) = f(x) = h'(z)$. □

Lemma 2.3 *Suppose $g\colon Z \to Y$ is injective and $f\colon X \to Y$ is any map. Then there exists $h\colon X \to Z$ with $f = g \circ h$ iff $\operatorname{Im} f \subseteq \operatorname{Im} g$. Morover h is unique.*

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ & {\scriptstyle h}\searrow & \uparrow{\scriptstyle g} \\ & & Z \end{array}$$

Proof. $\Rightarrow$: $f(x) = g(h(x)) \in \operatorname{Im} g$, so $\operatorname{Im} f \subseteq \operatorname{Im} g$.
$\Leftarrow$: Let $x \in X$ and define $h(x)$ to be y such that $g(y) = f(x)$. Since $f(x) \in \operatorname{Im} g$ such a y exists, and it is unique since g is injective. Uniqueness: If $f = g \circ h = g \circ h'$, then for any x, $g(h(x)) = g(h'(x))$ so $h(x) = h'(x)$ by injectivity of g. □

Corollary 2.4 *A function $g\colon X \to Y$ is bijective iff there exists an $h\colon Y \to X$ with $h \circ g = 1_X$ and $g \circ h = 1_Y$ where 1_X and 1_Y are the identity maps on X and Y.*

Proof. By Lemma 2.2, there is a $h\colon Y \to X$ with $1_X = h \circ g$ (using surjectivity of g *and* injectivity: $g(x) = g(x') \Rightarrow 1_X(x) = 1_X(x')$). Similarly Lemma 2.3, there is an $h'\colon Y \to X$ with $1_Y = g \circ h'$ (using injectivity of g *and* surjectivity: $\operatorname{Im} I_Y = Y \subseteq \operatorname{Im} g$). For all $y \in Y$, $h(y) = h(1_Y(y)) = h(g(h'(y))) = 1_X(h'(y)) = h'(y)$. so $h = h'$.
Conversely: $g(x) = g(x') \Rightarrow x = h(g(x)) = h(g(x')) = x'$, so g injective. If $y \in Y$, then $g(h(y)) = y$, so g surjective. □

Theorem 2.5 *Let $f\colon X \to Y$ be any map and define $x \sim y$ iff $f(x) = f(y)$. Then f can be "factored" as $f = i \circ \tilde{f} \circ \pi$ (i.e, $\tilde{f}(\bar{x}) = f(x)$), where*

- *$\pi\colon X \to X/\!\sim$ is the (surjective) quotient map,*
- *$\tilde{f}\colon X/\!\sim \to \operatorname{Im} f$ is a bijection,*
- *$i\colon \operatorname{Im} f \to Y$ is the (injective) inclusion map.*

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle \pi}\downarrow & & \uparrow{\scriptstyle i} \\ X/\!\sim & \xrightarrow{\tilde{f}} & \operatorname{Im} f \end{array}$$

Proof. Use Lemma 2.2 to write $f = h \circ \pi$, and then Lemma 2.3 to write $h = i \circ \tilde{f}$. The map $\tilde{f}$ is surjective, since if $y \in \operatorname{Im} f$, then $y = f(x)$ for some x and then $\tilde{f}(\pi(x)) = y$. The map $\tilde{f}$ is injective, since if $\tilde{f}(\bar{x}) = \tilde{f}(\bar{y})$ then $f(x) = f(y)$ and hence $x \sim y$ and $\bar{x} = \bar{y}$. □

Induction

Mathematical induction is normally applied in the following form:

Weak Induction: Let $P(n)$ be a property of the natural number n. If $P(0)$ holds and if $P(n)$ implies $P(n+1)$ for all n, then $P(n)$ holds for all n.

A slightly more general form is

Strong induction: Let $P(n)$ be a property of the natural number n. If for any n, the validity of $P(r)$ for all $r < n$ implies that $P(n)$ holds, then $P(n)$ holds for all n.

Note that the strong form implies the weak form, since for $n > 0$ we use the validity of $P(n-1)$ to deduce that $P(n)$ holds, and for $P(0)$ the validity of $P(r)$ for all $r < n$ is vacuous and must imply $P(0)$. Both forms of induction are simple consequences of

Well ordering: Any non-empty subset S of the natural numbers $\mathbb{N}$ has a smallest element.

To see this, let $S = \{n \mid P(n) \text{ does not hold}\}$. If $S = \emptyset$ then we are done. Otherwise S has a smallest element, n say. But then $P(r)$ holds for all $r < n$ and $P(n)$ does not hold, which contradicts our assumption on P.

Strong mathematical induction can be applied to any set with an ordering which satisfies the well ordering principle. For example, one can apply induction on ordered pairs (n, m) of natural numbers ordered lexicographically [$(n, m) < (n', m')$ iff either $n < n'$, or $n = n'$ and $m < m'$], which is in fact equivalent to double induction: an outer induction proving $P'(n) =$ "$\forall m\colon P(n, m)$", an an inner induction proving this by induction on m. Other well orderings on $\mathbb{N} \times \mathbb{N}$ are also possible.

Number theory

I will assume basic number theory. In particular, recall the

Division algorithm: If $a, b \in \mathbb{Z}$ with $b > 0$ then there exist unique integers q (the **quotient**) and r (the **remainder**) such that $a = qb + r$ and $0 \leq r < b$.

Proof. Apply well ordering to the set $\{a - qb \mid q \in \mathbb{Z}\} \cap \mathbb{N}$. □

For $a, b \in \mathbb{Z}$, write $a \mid b$ if there exists $c \in \mathbb{Z}$ such that $b = ca$. For example, $1 \mid a$ and $a \mid 0$ hold for all a, while $0 \mid a$ holds iff $a = 0$. Note that if $d \mid a, b$ then $d \mid ax + by$ for any $x, y \in \mathbb{Z}$, and if $d \mid d'$, $d' \mid d$, then $d = \pm d'$.

If $S \subseteq \mathbb{Z}$ then a **greatest common divisor** of S, $\gcd(S)$, is an integer d with the properties

G1. $d \mid a$ for each $a \in S$,

G2. if $c \mid a$ for each $a \in S$ then $c \mid d$.

Note: I use $c \mid d$ rather than $c \leq d$, as this allows a definition based purely on divisibility (which will later generalize) and not on ordering.

Lemma 3.1 *The gcd exists, is unique up to sign, and can always be written as a finite linear combination of elements of S:* $\gcd(S) = \sum_{i=1}^{r} c_i a_i$ *for some* $r, c_i \in \mathbb{Z}$ *and* $a_i \in S$.

Proof. The case when $S = \emptyset$ or $S = \{0\}$ is easy, so assume S contains some non-zero element. Let $I = \{\sum_{i=1}^{r} c_i a_i \mid r, c_i \in \mathbb{Z},\ a_i \in S\}$ be the set of all finite linear combinations of elements of S. Since S contains non-zero elements, I contains both positive and negative integers. Let d be the smallest positive element of I (well ordering). Now let $c \in I$ and write $c = qd + r$, $0 \leq r < d$. However, $c, d \in I$, so $r = c - qd$ is a finite linear combination of elements of S. Hence $r \in I$, which contradicts the choice of d unless $r = 0$. Thus every element of I is a multiple of d (and conversely it is clear that every multiple of d lies in I). In particular, every element of S lies in I, so is a multiple of d, so G1 holds. G2 is clear, since if c divides every element of S then it divides any linear combination of them, in particular d. For uniqueness, if d, d' are two gcd's, then G1 for d together with G2 for d' (with $c = d$) implies $d \mid d'$. Similarly $d' \mid d$. Thus $d' = \pm d$. □

Similarly, a **least common multiple** of S, $\mathrm{lcm}(S)$, is an integer d with the properties

L1. $a \mid d$ for each $a \in S$,

L2. if $a \mid c$ for each $a \in S$ then $d \mid c$.

The lcm also exists and is unique up to sign (exercise).

A **prime** is an integer $p \geq 2$ such that whenever $p \mid ab$ then either $p \mid a$ or $p \mid b$.

Unique factorization Any integer $n \geq 1$ can be written as a product of (not necessarily distinct) primes, and this product is unique up to order of the factors.

Proof. By induction on n. $n = 1$ is the empty product of primes. If $n > 1$ then either n is prime (so is a product of one prime), or $n \mid ab$ but $n \nmid a, b$ for some a, b. Let $m = \gcd(n, a)$, $m > 0$. Since $m \mid n$, $n = mt$ for some t. If $m = n$ then $n \mid a$, a contradiction. If $m = 1$ then $1 = ax + ny$, so $b = (ab)x + n(by)$ is divisible by n, a contradiction. Thus $1 < m < n$, so $m, t < n$ are products of primes and hence so is n. Uniqueness: If $n = p_1 \dots p_r = q_1 \dots q_s$ then $p_1 \mid q_1 \dots q_s$. Since p_1 is prime (and by induction on s), $p_1 \mid q_i$ for some i. But then $q_i = cp_1$ for some c. Since $q_i$is prime, $q_i \mid c$ or $q_i \mid p_1$. If $q_i \mid c$ then $|p_1| \leq 1$, a contradiction, so $q_i \mid p_1$ and $q_i = \pm p_1$. But $q_i, p_1 > 0$, so $q_i = p_1$. Dividing out by p_1 and using induction on n gives the result. □

In particular, if $n \geq 2$ has no positive factors other than 1 and n (n is **irreducible**), then n must be prime. Conversely any prime is irreducible.