ABOUT THE CLASS AND NOTES ON SET THEORY

About the Class

Evaluation. Final grade will be based 25%, 25%, 25%, 25%, on homework, midterm 1, midterm 2, final exam.

Exam dates. Midterm 1: February 19. Midterm 2: April 2. Final: May 14.

Homework. will be assigned on Courseworks each week. It is due before 1pm the following Wednesday in the Modern Algebra 1 homework box the 4th floor of the Math Building. Collaboration on homework is fine but the final write-up of homework solutions should be your own.

Extra Credit Problems. Homework problems labelled "extra credit" are optional and should be handed in separately directly to the instructor. The (rare) grade of $A+$ is for exceptional work, and cannot be earned without extra credit work.

Book. Dummit and Foote "Abstract Algebra" (DF for short) is strongly recommended. The course will cover roughly the content of chapters 1–6 (theory of groups), with some omissions, since DF is written as a graduate text.

We will start with two lectures on set theory, which is material not in DF.

The second semester of the course, Modern Algebra 2, will focus on rings, fields, Galois theory.

Notes. Class notes such as these will be posted on Courseworks when needed (material not in Bummit and Foote, etc.). The notes are not a replacement for class — they generally cover the bare essentials of what is described in class.

1. Basic set theory

(This material is mostly not in Dummit and Foote.)

We will start by explaining how "Russell's paradox" implies the necessity of being careful in the foundations of set theory.

A *set* is a collection of objects, but not every collection of objects can be called a set. For example the collection of *all* sets is not a set. The objects collected in a set are called the *elements* of the set.

1.1. Notations for sets.

- $a \in A$ means a is an element of the set A.
- $\{a_1, a_2, \dots, a_n\}$ is the set whose elements are $a_1, \dots, a_n$.
- $\emptyset$ is the set containing no elements.
- $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$ are standard notations for the sets of all integers, real numbers and complex numbers respectively.
- $\mathbb{N}$ is the set of *natural numbers* ("counting numbers"). Different authors use different conventions: it may mean $\{0, 1, 2, \dots\}$ or $\{1, 2, 3, \dots\}$. We will use the latter notation.
- If A is a set and P is a property which can be applied to elements of A then $\{a \in A : P(a)\}$ is a set which consists of the elements a of A for which property $P(a)$ is true. For example $\{n \in \mathbb{Z} : n \geq 1\}$ is the set $\mathbb{N}$.
- $A \subseteq B$ means A is a subset of B, i.e., every element of A is an element of B. $A = B$ means $A \subseteq B$ and $B \subseteq A$
- $A \subset B$ sometimes means the same as $A \subseteq B$ but some authors use it to mean A is a *proper* subset of B, i.e., $A \subseteq B$ and $A \neq B$.
- $|A|$ is the number of elements in A, called the size or *cardinality* of A. For example $|\emptyset| = 0$ and $|\{\emptyset\}| = 1$.
- If A and B are sets then $A \times B$ is the set of ordered pairs: $A \times B = \{(a, b) : a \in A \text{ and } b \in B\}$.
- If A is a set then its "power set" is the set whose elements are the subsets of A. It is usually denoted either by $\mathcal{P}(A)$ or sometimes 2^A.

1.2. **Functions or maps.** A *function* or *map* f from a set A to a set B is a rule which assigns an element of B to each element of A. We will generally use the word "map" rather than "function" but they are synonymous.

Notations:

- $f\colon A \to B$ means f is a map from A to B.
- If f is a map from A to B and $a \in A$ one writes $f(a)$ for the element of B assigned by f to a. It is called the *image* of a (by f).
- $f\colon a \mapsto b$ (read as "a mapsto b by f") is a synonym for $f(a) = b$.
- If $f\colon A \to B$ and $g\colon B \to C$ then $g \circ f$ denotes the *composition* of f and g defined by $g \circ f(a) = g(f(a))$.
- The function $id_A\colon A \to A$ is defined by $id_A(a) = a$ for all $a \in A$. It is also sometimes denoted 1_A. It is called the *identity function* of A.

Definition (Types of functions)**.**

- A function $f\colon A \to B$ is *surjective* (also called "onto"), if every element $b \in B$ is the image of some element of A. Notation: $f\colon A \twoheadrightarrow B$.
- A function $f\colon A \to B$ is *injective* (also called "one-one"), if no two elements of A have the same image by f, i.e., $f(a) = f(a') \Rightarrow a = a'$ for any two elements $a, a' \in A$. Notation: $f\colon A \rightarrowtail B$.
- A function $f\colon A \to B$ is *bijective* if it is both injective and surjective. Notation: $f\colon A \rightarrowtail\!\!\!\!\rightarrow B$.
  Equivalently, the function f has an *inverse function* (denoted f^{-1}), characterized by the property that $f \circ f^{-1} = id_B$ and $f^{-1} \circ f = id_A$.

1.3. Cardinality and sizes of infinity.

Definition (Cardinality)**.** We say two sets A and B have the same *cardinality* (a fancy name for "size"), written $|A| = |B|$ if there exists a bijective map $f\colon A \to B$. We write $|A| \leq |B|$ if there exists an injective map $f\colon A \to B$.

Theorem 1. *For two sets A and B either $|A| \leq |B|$ or $|B| \leq |A|$ (or both).*

Looking at the definition above, what this theorem says is that for any two set A and B there exists an injective map from one of the two sets to the other. This is a non-trivial fact, and to prove it one needs the "Axiom of Choice", one of the basic axioms of moderm set theory (which a few some logicians still reject).

Theorem 2 (Bernstein-Schroeder Theorem)**.** $|A| \leq |B|$ *and* $|B| \leq |A|$ *implies* $|A| = |B|$.

The proof of this will given later.

The following theorem implies that for any set there is a set with strictly greater cardinality. In particular, for every infinite "number" there is a larger one.

Theorem 3. *For any set A one has $|A| < |\mathcal{P}(A)|$ (i.e., $|A| \leq |\mathcal{P}(A)|$ and $|A| \neq |\mathcal{P}(A)|$).*

Recall $\mathcal{P}(A)$ is the "power set" (set of all subsets) of A. If $|A| = n$ is finite then the theorem just says that $n < 2^n$ for any non-negative integer n.

The following sets all have the same size: $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$. We also mentioned in class that the set of *algebraic numbers* (zeros of polynomials with integer coefficients), commonly called $\overline{\mathbb{Q}}$, also has the same size. The size of these sets is called $\aleph_0$ (spoken "aleph-zero"); it is the smallest infinite cardinal number and also goes by the name "countable infinity"[1]. The numbers which measure sizes of sets are called *cardinal numbers*, so they include 0, the natural numbers, $\aleph_0$, and larger sizes of infinity.

Theorem 3 implies that the cardinality of the power set of any one of the above countably infinite sets is strictly larger than $\aleph_0$. This larger cardinal number is called c (for "continuum"). It is also the size of $\mathbb{R}$ and $\mathbb{C}$.

A theorem of set theory says that for any cardinal there is a smallest cardinal larger than it. The second smallest infinite cardinal number is called $\aleph_1$, and after it come $\aleph_2$, etc. It was long an open question whether the *Continuum Hypothesis*, which says $c = \aleph_1$, is true. But it was then proved by Paul Cohen in 1963 that the Continuum Hypothesis is independent of the standard axioms of set theory, so one can choose to add it as an axiom or to add its negation as an axiom. Most logicians take the view that c is much larger than every $\aleph_n$, $n = 1, 2, \dots$.

We don't need to worry about this question. But the fact that there are different sizes of infinity has important consequences. For example, in the 19th century mathematicians worried whether *transcendental* numbers exist—numbers which are not algebraic, i.e., not a solution of a polynomial equation with integer coefficients. It took many years to prove that certain numbers, first e, then π and others, were transcendental. But the fact that there are only a countable infinity of algebraic numbers while the size of $\mathbb{R}$ and $\mathbb{C}$ is the larger cardinal c shows that, in fact, "almost every" number is transcendental.

1.4. **Some proofs.**

We start with the proofs of two of the theorems above. First some notation.

Notation. If $f\colon A \to B$ is a map and $A' \subseteq A$ one writes $f(A')$ for the set

$$f(A') := \{b \in B : b = f(a) \text{ for some } a \in A'\}\,.$$

It is called the *image of A' under f*. The set $f(A)$ is simply called the *image of f*.

If $B' \subseteq B$ one writes $f^{-1}(B')$ for the set

$$f^{-1}(B') := \{a \in A : f(a) \in B'\}$$

[1] Caution: If you google the word "countable" (and even in some recent books) you may see the statement that "countable" means "of size $\aleph_0$." This is WRONG. It means "either finite or of size $\aleph_0$"

(note that f^{-1} itself is not a map unless f is bijective—the notation $f^{-1}(B')$ should be seen as a whole). $f^{-1}(B')$ is called the *inverse image of B' by f.*

For $b \in B$ one abbreviates $f^{-1}(\{b\})$ as $f^{-1}(b)$. It is often called the *fiber of f over b.*
CAUTION: If f is bijective then f^{-1} exists, so the notation $f^{-1}(b)$ becomes ambiguous: if a is such that $f(a) = b$ then $f^{-1}(b)$ means "$\{a\}$" in the previous paragraph, but means "a" if interpreted as the function f^{-1} applied to b. Which is meant will usually be clear from context.

If f is just injective one still sometimes uses $f^{-1}(b)$ to mean the element a with $f(a) = b$, if this a exists, rather tnan the set $\{a\}$. Again, what is meant is usually clear from context. This occurs in the following proof, for example.

Proof of the Bernstein Schroeder Theorem 2. We are given that $|A| \leq |B|$ and $|B| \leq |A|$ and we are to show that $|A| = |B|$. Using the definitions of what this means, we have injective maps $\alpha\colon A \rightarrowtail B$ and $\beta\colon B \rightarrowtail A$, and we need to construct a bijective map $f\colon A \rightarrowtail\!\!\!\rightarrow B$.

We consider maximal chains of the form

$$\dots \overset{\beta}{\mapsto} a_i \overset{\alpha}{\mapsto} b_i \overset{\beta}{\mapsto} a_{i+1} \overset{\alpha}{\mapsto} b_{i+1} \overset{\beta}{\mapsto} \dots$$

Notice that such a chain always extends infinitely far to the right, since given an $a \in A$ we can always apply α to it to get $a \overset{\alpha}{\mapsto} \alpha(a) \in B$ and similarly for a $b \in B$ we have an arrow $b \overset{\beta}{\mapsto} \beta(b)$. On the other hand, the chain is not necessarily extendable to the left, since if some $a \in A$ is in the chain, there may or may not be a b with $\beta(b) = a$ (if there *is* such a b it is unique by injectivity of β, so we can then extend the chain one step to the left). Similarly for $b \in B$ in a chain there may or may not be an a with $\alpha(a) = b$ to let one extend the chain to the left. Thus each element of A and B is either in a bi-infinite chain (one that extends infinitely both left and right) or a chain that starts with an a which is not in the image of $\beta\colon B \to A$ (we call this an *A-chain*) or a chain which starts with a b which is not in the image of $\alpha\colon A \to B$ (a *B-chain*).

We define $f\colon A \to B$ and $g\colon B \to A$ as follows:

$$f(a) = \begin{cases} \alpha(a) & \text{if } a \text{ is in a bi-infinite chain or } A\text{-chain,} \\ \beta^{-1}(a) & \text{if } a \text{ is in a } B\text{-chain.} \end{cases}$$

$$g(b) = \begin{cases} \alpha^{-1}(b) & \text{if } b \text{ is in a bi-infinite chain or } A\text{-chain,} \\ \beta(b) & \text{if } b \text{ is in a } B\text{-chain.} \end{cases}$$

Note that f is well-defined, since $\alpha(a)$ is always defined, while $\beta^{-1}(a)$ is defined if a is in a B-chain (going to the left in such a chain stops at an element of B, not of A). Similarly g is well defined. If an element $a \in A$ is in a bi-infinite or A-chain we have $g \circ f(a) = \alpha^{-1} \circ \alpha(a) = a$ while if it is in a B-chain we have $g \circ f(a) = \beta \circ \beta^{-1}(a) = a$. Thus $g \circ f(a) = a$ for any $a \in A$, that is: $g \circ f = id_A$. Similarly one computes that $f \circ g = id_B$, so g is an inverse function for f, proving that f is bijective. □

(The little box □ is a standard sign in mathematics to signify the end of a proof. In old texts you sometimes see the box replaced by "*QED*" and the box is therefore called a QED-box. QED is short for *quod erat demonstrandum*, latin for "what was to be proved".)

Proof of Theorem 3. We want to show that if A is a set then $|A| < |\mathcal{P}(A)|$. Certainly $|A| \leq |\mathcal{P}(A)|$, since this means there is an injective map $A \to \mathcal{P}(A)$, and such a map can be given, for example, by $a \mapsto \{a\}$ for $a \in A$. To see that $|A| < |\mathcal{P}(A)|$ we must show that there is no bijective map between $f\colon A \to \mathcal{P}(A)$. Suppose there were such a map f. Then, in particular, f is surjective. We will show that this leads to a contradiction, showing f cannot exist.

So suppose $f\colon A \twoheadrightarrow \mathcal{P}(A)$ (recall the double-headed arrow means "surjective map"). Consider the following element of $\mathcal{P}(A)$ (i.e., subset of A):

$$B := \{a \in A : a \notin f(a)\}.$$

We will show that B is not in the image of the map f, contradicting surjectivity. Indeed, if we have an element a with $f(a) = B$, then we ask ourselves if the element a is in B or not. If $a \in B$ then the definition of B says $a \notin f(a)$, but $f(a) = B$, so $a \notin B$. Similarly, $a \notin B$ means $a \notin f(a)$, which by definition of B means $a \in B$. Thus, under the assumption that $f(a) = B$ we find that a is neither in nor not in B. So no such a can exist, and the proof is done. □

1.5. **Union, Intersection, Cartesian Product.** If $A_1, A_2, \ldots, A_n$ are sets we write

$$A_1 \cup \cdots \cup A_n \quad \text{and} \quad A_1 \cap \cdots \cap A_n$$

for the union and intersection of these sets. $A_1 \cup \cdots \cup A_n$ is the set of elements which occur in at least one of the A_i and $A_1 \cap \cdots \cap A_n$ is the set of elements which occur in all of them.

More generally, if I is some set (which need not be finite) and we have a set A_i for each $i \in I$ we write

$$\bigcup_{i\in I} A_i := \{a : \exists i, a \in A_i\} \quad \text{and} \quad \bigcap_{i\in I} A_i := \{a : \forall i, a \in A_i\}$$

for their union and intersection.

If $A_1, A_2, \ldots, A_n$ are sets one defines their *cartesian product* to be the set of ordered n-tuples

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

We also use the notation $A_1 \times \cdots \times A_n = \bigtimes_{i=1}^n A_i$.

One way of thinking of an element $(a_1, \ldots, a_n) \in A_1 \times \cdots \times A_n$ is that it is a choice of an element a_i in each set A_i. We can also think of this as a function with domain I, since choosing an element a_i for each $i \in \{1, \ldots, n\}$ is the same as considering the function defined on $\{1, \ldots, n\}$ given by $i \mapsto a_i$.

Now if I is again just some set (possibly infinite) and we have a set A_i for each $i \in I$ we would like to write the cartesian product of these sets, but a notation like $(\ldots, a_i, \ldots)$ does not work well, since there is no obvious order to put on the entries. But we can use the above idea of an element of the cartesian product being a function which chooses an element of A_i for each $i \in I$, and define:

$$\bigtimes_{i\in I} X_i := \{a \colon I \to \bigcup_{i\in I} A_i \mid a(i) \in A_i \text{ for each } i \in I\}.$$

A fundamental axiom of set theory is the *Axiom of Choice*, which says that if each X_i is non-empty then $\bigtimes_{i\in I} X_i$ is non-empty, i.e., there exists some choice of an element from each X_i. Although it might seem that this should be obviously true, it does not follow from the other basic axioms that are used to create modern set theory. As already mentioned earlier, one cannot prove Theorem 1 without the axiom of choice. Proving Theorem 1 would take us too far afield, so we won't do it here.

2. Binary Relations

A *binary relation* on A is a subset $R \subseteq A \times A$. The notation $a\,R\,b$ then means $(a, b) \in R$.

The binary relation R is an *equivalence relation* if it satisfies the properties:

(1) *Reflexivity*: $\forall a \in A \quad a\,R\,a$
(2) *Symmetry*: $\forall a, b \in A \quad a\,R\,b \Rightarrow b\,R\,a$
(3) *Transitivity*: $\forall a, b, c \in A \quad a\,R\,b$ and $b\,R\,c \Rightarrow a\,R\,c$

The binary relation R is a *partial order* if it is reflexive and transitive (properties (1) and (3)) and also satisfies

(2′) *Antisymmetry*$\forall a, b \in A \qquad a \, R \, b$ and $b \, R \, A \Rightarrow a = b$

A set with a partial order relation on it is called a *partially ordered set.*

Examples.

- *Equality.* If A is any set then the relation "=" $:= \{(a, a) : a \in A\}$ is the usual relation of equality. It is an equivalence relation.
- *Congruence.* The relation on the set $\mathbb{Z}$ of integers

$$\equiv_n := \{(a, b) \in \mathbb{Z} \times \mathbb{Z} : \exists k \in \mathbb{Z}, b = a + kn\}$$

is called *congruence mod n*. One writes "$a \equiv_n b$" or more commonly: "$a \equiv b \pmod{n}$". $\equiv_n$ is an equivalence relation.

- *Set inclusion* The relation "$\subseteq$" on the power set $\mathcal{P}(X)$ of a set X defined as "$\subseteq$" $:= \{(A, B) \in \mathcal{P}(X) \times \mathcal{P}(X) : A \subseteq B\}$ is a partial order.
- The usual relation of inequality $\leq$ defined on $\mathbb{N}$ or $\mathbb{R}$ or $\mathbb{Q}$ or $\mathbb{R}$ is a *total order*, i.e. a partial order with the additional property that for any a, b in the set at least one of $a \leq b$ or $b \leq a$ is true.

Theorem 4 (Zorn's Lemma)**.** *Suppose* $(A, \leq)$ *is a partially ordered set with the property that any infinite sequence in* A *of the form* $a_1 \leq a_2 \leq a_3 \leq \dots$ *has an upper bound (i.e., there is an* $a \in A$ *with* $a_i \leq a$ *for all* $i = 1, 2, 3, \dots$*). Then* A *has at least one* maximal element m *(i.e., an element with the property that* $m \leq a \Rightarrow m = a$*).*

This theorem is equivalent to the Axiom of Choice, and it turns out to be a powerful tool in proving existence of certain things (an example is in the homework as an extra credit problem).

2.1. Partitions.

A *partition* of a set is a decomposition of the set as a union of pairwise disjoint non-empty subsets. Precisely, if A is a set then a set $\{A_i : i \in I\}$ of subsets of A is a partition of A if it satisfies the properties

- Each A_i is non-empty;
- $\bigcup_{i \in I} A_i = X$;
- $A_i \cap A_j = \emptyset$ if $i \neq j$.

An example is the partition of the integers $\mathbb{Z}$ into the sets of even and odd integers.

A partition turns out to be just a different view of an equivalence relation.

2.2. Partition to equivalence relation.

If $\{A_i : i \in I\}$ is a partition of A then there is a corresponding equivalence relation $\sim$ on A given by $a \sim b$ if and only if a and b are in the same member of the partition (i.e., $a \sim b :\Leftrightarrow \exists i \in I : a, b \in A_i$).

2.3. **Equivalence relation to partition.** Given an equivalence relation $\sim$ on a set A, for each $a \in A$ the set

$$\{b \in A : a \sim b\}$$

is called the *equivalence class* of a. Common notations for this equivalence class are $\overline{a}$ or $[a]$. I will use $\overline{a}$ here, so

$$\overline{a} = \{b \in A : a \sim b\}\,.$$

We call a a *representative* for the equivalence class $\overline{a}$. Any two elements of $\overline{a}$ are representatives of $\overline{a}$:

Lemma 5. *If $\sim$ is an equivalence relation on A and $\overline{a}$ denotes the equivalence class of a then the following are equivalent:*

(1) $a \sim b$
(2) $\overline{a} = \overline{b}$
(3) $\overline{a} \cap \overline{b} \neq \emptyset$.

Proof. (1)⇒(2): Assume $a \sim b$. Then if $c \in \overline{b}$ we have $b \sim c$ so by transitivity of $\sim$ also $a \sim c$, whence $c \in \overline{a}$. Similarly $c \in \overline{b}$ implies $c \in \overline{a}$ so $\overline{a} = \overline{b}$.

(2)⇒(3) is trivial.

(3)⇒(1): Suppose c is a common element of $\overline{a}$ and $\overline{b}$. Then $c \in \overline{a}$ means $a \sim c$ and $c \in \overline{b}$ means $b \sim c$ and then $c \sim b$ by symmetry and then $a \sim b$ by transitivity. □

A "lemma" is a statement designed to help prove some other result. In this case the result we are proving is

Theorem 6. *If $\sim$ is an equivalence relation on A then the set of equivalence classes for $\sim$ is a partition of A.*

Proof. We need to prove the three defining properties of a partition. First, each $\overline{a}$ is nonempty, since $a \in \overline{a}$. Second, $\bigcup_{a\in A} \overline{a} = A$ since for any $b \in A$ we have $b \in \overline{b} \in \bigcup_{a\in A} \overline{a}$. Finally if two equivalence classes $\overline{a}$ and $\overline{b}$ are not equal then they are disjoint by the above lemma. □

Example. The partition of the integers $\mathbb{Z}$ into the sets of even and of odd integers corresponds to the equivalence relation $\equiv_2$ of congruence modulo 2.

More generally, congruence modulo n leads to the partition of the integers into the n sets:

$$\begin{aligned}
\overline{0} =& \{\dots, -2n, -n, 0, n, 2n, \dots\} \\
\overline{1} =& \{\dots, -2n+1, -n+1, 1, n+1, 2n+1, \dots\} \\
\overline{2} =& \{\dots, -2n+2, -n+2, 2, n+2, 2n+2, \dots\} \\
\vdots \quad & \quad \vdots \\
\overline{n-1} =& \{\dots, -n-1, -1, n-1, 2n-1, 3n-1, \dots\}
\end{aligned}$$