Math 113 Summer 2014. Some basic set theory stuff.

In this preliminary note we will introduce some of the fundamental language and notation that will be adopted in this course. It is intended to be an informal introduction to the language of sets and functions and logical quantifiers.

1.1 Basic Set Theory

For most mathematicians the notion of a *set* is fundamental and essential to their understanding of mathematics. In a sense, everything in sight is a set (even functions can be considered as sets![1]) .

Groups, rings, fields, vector spaces,... are all examples of a *set with structure*, so we need to ensure that we know what a set is and understand how to write down and describe sets using **set notation**.

Definition 1.1.1 (Informal Definition)**.** A *set* S is a collection of objects (or elements). We will denote the size of a set S by $|S|$; this will either be a natural number or infinite (we do discuss questions of cardinality of sets).

For example, we can consider the following sets:

- the set P of people in 71 Evans Hall at 4.20pm on 6/23/2014,
- the set B of all people in the city of Berkeley at 4.20pm on 6/23/2014,
- the set $\mathbb{R}$ of all real numbers,
- the set A of all real numbers that are greater than or equal to π,
- the set $M_{m\times n}(\mathbb{R})$ of all $m \times n$ matrices with real entries,
- the set $\mathrm{Aut}(\mathbb{Z})$ of automorphisms of the group $(\mathbb{Z}, +)$,
- the set $C(0, 1)$ of all real valued continuous functions with domain $(0, 1)$.

Don't worry if some of these words are new to you, we will define them shortly.

You will observe that there are some relations between these sets: for example,

- every person that is an object in the collection P is also an object in the collection B,
- every number that is an object of A is also an object of $\mathbb{R}$.

We say in this case that P (resp. A) is a *subset* of B (resp. $\mathbb{R}$), and write

$$P \subseteq B \text{ (resp. } A \subseteq \mathbb{R}).$$

Remark. In this class we will use the notations $\subseteq$ and $\subset$ interchangeably and make no distinction between them. On the blackboard I will write $\subseteq$ as this is a notational habit of mine whereas in these notes I shall usually write $\subset$ as it is a shorter command in LaTeX(the software I use to create these notes).

We can also write the following

$$P = \{x \in B \mid x \text{ is in 71 Evans Hall at 4.20pm on 6/23/2014}\},$$

or in words:

P is the set of those objects x in B such that x is in 71 Evans Hall at 4.20pm on 6/18/2014.

Here we have used

[1] A function $f : A \to B$; $x \mapsto f(x)$ is the same data as providing a subset $\Gamma_f \subset A \times B$, where $\Gamma_f = \{(x, f(x)) \mid x \in A\}$, the *graph of f*. Conversely, if $C \subset A \times B$ is a subset such that, $\forall a \in A, \exists b \in B$ such that $(a, b) \in C$, and $(a, b) = (a, b') \in C \implies b = b'$, then C is the graph of some function.

- the logical symbol '$\in$' which is to be translated as 'is a member of' or 'is an object in the collection',

- the vertical bar '$|$' which is to be translated as 'such that' or 'subject to the condition that'.

In general, we will write (sub)sets in the following way:

$$T = \{x \in S \,|\mathcal{P}\},$$

where $\mathcal{P}$ is some property or condition. In words, the above expression is translated as

T is the set of those objects x in the set S such that x satisfies the condition/property $\mathcal{P}$.

For example, we can write

$$A = \{x \in \mathbb{R} \mid x \geq \pi\}.$$

Definition 1.1.2. We will use the following symbols (or logical quantifiers) frequently:

- $\forall$ - translated as 'for all' or 'for every', (the *universal quantifier*)

- $\exists$ - translated as 'there exists' or 'there is, (the *existential quantifier*).

For example, the statement

'for every positive real number x, there exists some real number y such that $y^2 = x$',

can be written

$$\forall x \in \mathbb{R} \text{ with } x > 0, \exists y \in \mathbb{R} \text{ such that } y^2 = x.$$

Remark. Learning mathematics is difficult and can be made considerably more difficult if the basic language is not understood. If you ever encounter any notation that you do not understand please ask a fellow student or ask me and I will make sure to clear things up. I have spent many hours of my life staring blankly at a page due to misunderstood notation so I understand your pain in trying to get to grips with new notation and reading mathematics.

Notation. In this course we will adopt the following notational conventions:

- $\varnothing$, the *empty set* (ie the empty collection, or the collection of no objects),

- $[n] = \{1, 2, 3, \ldots, n\}$,

- $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$, the set of *natural numbers*,

- $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \ldots\}$, the set of *integers*,

- $\mathbb{Z}_{\geq a} = \{x \in \mathbb{Z} \mid x \geq a\}$, and similarly $\mathbb{Z}_{>a}, \mathbb{Z}_{\leq a}, \mathbb{Z}_{<a}$,

- $\mathbb{Q} = \{\frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0\}$, the set of *rational numbers*,

- $\mathbb{R}$, the set of *real numbers*,

- $\mathbb{C}$, the set of *complex numbers*.

Remark (Complex Numbers)**.** Complex numbers are poorly taught in most places so that most students have a fear and loathing of them. However, there is no need to be afraid! It really doesn't matter whether you consider imaginary numbers to be 'real' (or to exist in our domain of knowledge in this universe), all that matters is that you know their basic properties: a complex number $z \in \mathbb{C}$ is a'number' that can be expressed in the form

$$z = a + b\Delta, \; a, b \in \mathbb{R},$$

where, for now, Δ is just some symbol.

We can add and multiply the complex numbers $z = a + b\Delta, w = c + d\Delta \in \mathbb{C}$, as follows

$$z + w = (a + c) + (b + d)\Delta, \quad z.w = (ac - bd) + (bc + ad)\Delta.$$

If $z = a + b\Delta \in \mathbb{C}$ then the complex number $\tilde{z} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}\Delta$ satisfies

$$z.\tilde{z} = \tilde{z}.z = 1,$$

so that $\tilde{z}$ is the multiplicative inverse of z and we can therefore write $1/z = z^{-1} = \tilde{z}$. Hence, if $z = a + b\Delta, w = c + d\Delta \in \mathbb{C}$, then

$$z/w = z.w^{-1} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2}\Delta.$$

Of course, the number $i = 1.\Delta$ satisfies the property that $i^2 = -1$, so that $1.\Delta$ corresponds to the imaginary number i that you learned about in high school. However, as we will be using the letter i frequently for subscripts, we shall instead just write $\sqrt{-1}$ so that we will consider complex numbers to take the form

$$z = a + b\sqrt{-1}.$$

We have the following 'inclusions'

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C},$$

so that, in particular, *every real number is also a complex number* (if $a \in \mathbb{R}$ then we consider $a = a.1 + 0.\sqrt{-1} \in \mathbb{C}$).

Definition 1.1.3 (Operations on Sets)**.** • Suppose that S is a set and S_1, S_2 are subsets.

- the **union of** S_1 **and** S_2 is the set

$$S_1 \cup S_2 = \{x \in S \mid x \in S_1 \text{ or } x \in S_2\}.$$

- the **intersection of** S_1 **and** S_2 is the set

$$S_1 \cap S_2 = \{x \in S \mid x \in S_1 \text{ and } x \in S_2\}.$$

More generally, if $S_i \subset S, i \in J$, is a family of subsets of S, where J is some indexing set, then we can define

$$\bigcup_{i \in J} S_i = \{s \in S \mid s \in S_k, \text{ for some } k \in J\},$$

and

$$\bigcap_{i \in J} S_i = \{s \in S \mid s \in S_k, \forall k \in J\}.$$

• Let A, B be sets.

- the **Cartesian product of** A **and** B is the set

$$A \times B = \{(a, b) \mid a \in A, b \in B\},$$

so that the elements of $A \times B$ are **ordered pairs** (a, b), with $a \in A, b \in B$. In particular, it is not true that $A \times B = B \times A$.

Moreover, if $(a, b), (a', b') \in A \times B$ and $(a, b) = (a', b')$, then we must necessarily have $a = a'$ <u>and</u> $b = b'$.

For example, consider the following subsets of $\mathbb{R}$:

$$A = \{x \in \mathbb{R} \mid 0 < x < 2\},\ B = \{x \in \mathbb{R} \mid x > 1\},\ C = \{x \in \mathbb{R} \mid x < 0\}.$$

Then,

$$A \cup B = (0, \infty),\ A \cap B = (1, 2),\ A \cap C = \varnothing,\ A \cup B \cup C = \{x \in \mathbb{R} \mid x \neq 0\}.$$

Also, we have

$$A \times C = \{(x, y) \mid 0 < x < 2,\ y < 0\}.$$

1.2 Functions

Functions allow us to talk about certain relationships that exist between sets and allow us to formulate certain operations we may wish to apply to sets. You should already know what a function is but the notation to be introduced may not have been encountered before.

Definition 1.2.1. Let A, B be sets and suppose we have a function $f : A \to B$. We will write the information of the function f as follows:

$$f : A \to B,\ x \mapsto f(x),$$

where $x \mapsto f(x)$ is to be interpreted as providing the data of the function, ie, x is the input of the function and $f(x)$ is the output of the function. Moreover,

- A is called the **domain of** f,
- B is called the **codomain of** f.

For example, if we consider the function $|.| : \mathbb{R} \to [0, \infty)$, the 'absolute value' function, then we write

$$|.| : \mathbb{R} \to [0, \infty),\ x \mapsto \begin{cases} x, & \text{if } x \geq 0, \\ -x, & \text{if } x < 0. \end{cases}$$

The codomain of $|.|$ is $[0, \infty)$ and the domain of $|.|$ is $\mathbb{R}$.

Definition 1.2.2. Let $f : A \to B$ be a function.

- we say that f is **injective** if the following condition is satisfied:

$$\forall x, y \in A, \text{ if } f(x) = f(y) \text{ then } x = y,$$

- we say that f is **surjective** if the following condition is satisfied:

$$\forall y \in B, \exists x \in A \text{ such that } f(x) = y,$$

- we say that f is **bijective** if f is both injective and surjective.

It should be noted that the injectivity of f can also be expressed as the following (logically equivalent) condition:

$$\text{if } x, y \in A,\ x \neq y, \text{ then } f(x) \neq f(y).$$

Also, the notion of bijectivity can be expressed in the following way:

$$\forall y \in B, \text{ there is a unique } x \in A \text{ such that } f(x) = y.$$

Hence, if a function is bijective then there exists an inverse function $g : B \to A$ such that

$$\forall x \in A,\ g(f(x)) = x, \text{ and } \forall y \in B,\ f(g(y)) = y.$$

Remark. These properties of a function can be difficult to grasp at first. Students tend to find that injectivity is the hardest attribute of a function to comprehend. The next example is an attempt at providing a simple introduction to the concept of injectivity/surjectivity of functions.

Example 1.2.3. Consider the set P described above (so an object in P is a person in 71 Evans Hall, at 4.20pm on 6/23/2012) and let $\mathcal{C}$ denote the set of all possible cookie ice cream sandwiches available at C.R.E.A.M. on Telegraph Avenue (for example, vanilla ice cream on white chocolate chip cookies). Consider the following function

$$f : P \to \mathcal{C}\ ;\ x \mapsto f(x) = x\text{'s favourite cookie ice cream sandwich.}$$

In order for f to define a function we are assuming that nobody who is an element of P is indecisive so that they have precisely one favourite cookie ice cream sandwich.[2]

So, for example,

$$f(\textit{George}) = \text{ banana walnut ice cream on chocolate chip cookies.}$$

What does it mean for f to be

- injective? Let's go back to the definition: we require that for any two people $x, y \in P$, if $f(x) = f(y)$ then $x = y$, ie, if any two people in P have the same favourite cookie ice cream sandwich then those two people must be the same person. Or, what is the same, no two people in P have the same favourite cookie ice cream sandwich.

- surjective? Again, let's go back to the definition: we require that, if $y \in \mathcal{C}$ then there exists some $x \in P$ such that $f(x) = y$, ie, for any possible cookie ice cream sandwich y available at C.R.E.A.M. there must exist some person $x \in P$ for which y is x's favourite cookie ice cream sandwich.

There are a couple of things to notice here:

1. in order for f to be surjective, we must necessarily have at least as many objects in P as there are objects in $\mathcal{C}$. That is
$$f \text{ surjective} \implies |P| \geq |\mathcal{C}|.$$

2. in order for f to be injective, there must necessarily be more objects in $\mathcal{C}$ as there are in P. That is
$$f \text{ injective} \implies |P| \leq |\mathcal{C}|.$$

3. if P and $\mathcal{C}$ have the same number of objects then f is injective if and only if f is surjective.

You should understand and provide a short proof as to why these properties hold true.

The fact that these properties are true is dependent on the fact that both P and $\mathcal{C}$ are *finite sets*.

We will now include a basic lemma that will be useful throughout these notes. Its proof is left to the reader.

Lemma 1.2.4. *Let $f : R \to S$ and $g : S \to T$ be two functions.*

- *If f and g are both injective, then $g \circ f : R \to T$ is injective. Moreover, if $g \circ f$ is injective then f is injective.*

- *If f and g are both surjective, then $g \circ f : R \to T$ is surjective. Moreover, if $g \circ f$ is surjective then g is surjective.*

- *If f and g are bijective, then $g \circ f : R \to T$ is bijective.*

[2] Why are we making this assumption?