BACKGROUND

In this handout, we discuss some points of elementary logic that are apt to cause confusion, and also introduce ideas of set theory, and establish the basic terminology and notation. This is not examinable material, but read it carefully, as this forms the basic language of mathematics.

1. How to read mathematics

Don't just read it, fight it! Mathematics says a lot with a little. The reader must participate. After reading every sentence, stop, pause and think: do I really understand this sentence?

Don't read too fast. Reading mathematics too quickly results in frustration. A half hour of concentration while reading a novel perhaps buys you 20 pages with full comprehension. The same half hour in a math article buys you just a couple of lines. There is no substitute for work and time.

An easy way to progress in a mathematics course is to read the relevant section of the course notes or book before the lectures, and then once again on the same day after the lecture is over. Keep up, as Mathematics is different from other disciplines: you need to know yesterday's material to understand today's. Don't save it all for one long night of cramming, which simply won't work with Mathematics.

Before attempting the exercises, make sure you read the corresponding section from the lecture notes or book. After reading an exercise, stop and think if you know all the terms in the exercise, and if you understand what is being asked. Then think about what is given and what is required. You might then see a possible way of proceeding. You can do some rough work by writing down a few things in order to convince yourself that your strategy indeed works. Then write down your answer in a manner that a person can understand logically what your argument is. Justify each step. Writing proofs is an art, and one gets better at it only by practice. Every step in the proof is a (mathematical) statement, but it is a sentence in English! So make sure that each step in your argument reads like a simple sentence (so avoid the use of a chain of dangling '$\Rightarrow$' or '$\Leftrightarrow$', and pay attention to punctuation and grammar!).

2. Definitions, Lemmas, Theorems and all that

2.1. **Definitions.** A *definition* in Mathematics is a name given to a mathematical object by specifying what the mathematical object is. Just like in biology we define that

An animal is called a *fish* if it is a cold-blooded, water-dwelling vertebrate with gills.

Observe that in defining a fish, we have listed the characterizing properties that specify which animals are fish and which aren't. In the same way, in Mathematics, a mathematical object (a set or a function) is given a certain name if it satisfies certain properties.

Any definition of a mathematical term or a phrase will roughly have the form

$$\boxed{* * *} \text{ if } \boxed{\star \star \star},$$

where $* * *$ is the term which is being defined, and $\star\star\star$ is its defining property. For example:

Definition 1. An integer is called *even* if it is divisible by 2.

Definition 2. An integer is called *prime* if it is greater than 1 and it has no divisors other than 1 and itself.

Definition 3. A function $f : [a, b] \to \mathbb{R}$ is *continuous at* $c \in [a, b]$ if for every $\epsilon \in \mathbb{R}$ such that $\epsilon > 0$, there exists a $\delta \in \mathbb{R}$ such that $\delta > 0$ and for all $x \in [a, b]$ satisfying $|x - c| < \delta$, $|f(x) - f(c)| < \epsilon$.

In a definition the word 'if' really means 'if and only if'. So in the definition of odd integers, all integers which are not divisible by 2 are ruled out from being even.

Definitions are very important, as without knowing what the objects precisely mean, how can we prove things about them?

2.2. **Lemmas, Theorems, ...** Not all results proved in mathematics are called 'theorems'. Some are called 'lemmas' or 'corollaries'. Strictly speaking, there is no difference among them. However, the distinction rests on some aspects such as utility, depth and beauty. A *lemma* is useful in a limited context (often only as a preparatory step for some theorem) and is too technical to have an aesthetic appeal. A *theorem* carries with it some depth and a certain succinctness of form and often represents the culmination of some coherent piece of work, while a *corollary* is like an outgrowth of a theorem, and it is an easy consequence of the theorem.

3. Some remarks about logic

3.1. **Bivalued logic.** A mathematical statement is a sentence that allows only two possibilities: either it is true or it is false. Thus there is nothing between 'true' or 'false'. There is no such thing as 'very true', 'almost true', 'substantially true', 'partially true', or 'having an element of truth', although we commonly use such phrases in everyday life.

Of course there are many mathematical statements for which we do not know if they are true or false. An example of this is the *Goldbach's conjecture*, which states the following:

(1) Every even integer greater than 2 can be expressed as a sum of two prime numbers.

Although the above statement has been verified for an impressive range of cases, nobody has proved it for all cases. Nor has it been disproved, that is, nobody has so far discovered even a single even integer greater than 2 which cannot be expressed as a sum of two primes. Still, even today, the statement is either true or false, even though we do not know which way it is.

3.2. **Statements about a class.** The above remark about quantification of truth is important for statements about a class. The statement

(2) All rich men are happy

is about a class, namely that comprising rich men. Goldbach's conjecture above is a statement about the class of all even integers greater than 2.

A layman is apt to regard these statements as 'true' or 'nearly true' when they hold in a large number of cases. Even if there are a few exceptions, he is likely to ignore them as say 'The exception proves the rule!'. In mathematics, this is not so. Even a single exceptional case (a *counter-example*, as it is called), renders false a statement about a class. Thus even one unhappy rich man makes the statement (2) as false as millions of such men would do. In other words, in mathematics, we interpret the words 'all' and 'every' quite literally, not allowing even a single exception. If we want to make a true statement after taking the exceptional ones into account, we would have to make a different statement such as

All rich men other than Mr. X are happy.

But loose expressions such as ‘most’, ‘a great many’ or ‘almost all’ cannot be used in mathematical statements, unless, of course they have been precisely defined earlier.

There is another type of statements made about a class. These do not assert that something holds for *all* elements of the class, but instead that it holds for *at least one* element from that class. Take for example the statement

There exists a man whose height is 5 feet 7 inches

or

There exists a natural number k such that $1 < k < 4294967297$ that divides 4294967297.

These statements refer respectively to the class of all men and to the class of all natural numbers between 1 and 4294967297. In each case, the statements says that there is at least one member of the class having a certain property. It does not say how many such members are there. Nor does it say which ones they are. Thus the first statement tells us nothing by way of the name and the address of the person with that height, and the second one does not say what this divisor is. These statements are, therefor, not as strong as, respectively, the statements, say,

Mr. X in London is 5 feet 7 inches tall

or

641 divides 4294967297

which are very specific. A statement which merely asserts the existence of something without naming it or without giving any method for finding it is called an *existence statement*. In the bivalued logic setting of mathematics, existence statements are either true or false, even if they are not specific.

3.3. **Negation of a statement.** A *negation* of a statement is a statement which is true precisely when the original statement is false and vice-versa. The simplest way to negate a statement is to precede it with the phrase ‘It is not the case that ...’. Thus the negation of

Mr. X is rich

is

It is not the case that Mr. X is rich.

Note that the negation of

All men are mortal

is not

All men are immortal.

In view of the comments that we have made about the truth of statements about a class, it is false, even when it fails to hold just in one case, that is, when there is even one man who is not mortal. So the correct logical negation is

There exists a man who is immortal.

Not surprisingly, the negation of an existence statement is a statement asserting that every member of the class (to which the existence statement refers) fails to have the property asserted by the existence statement. Thus, the negation of

There exists a rich man

is

No man is rich

that is,

Every man is poor.

If we keep in mind these simple facts, we can almost mechanically write down the negation of any complicated statement. For example, if $a_1, a_2, a_3, \ldots$ is a sequence of real numbers, then the negation of

$$\forall \epsilon \in \mathbb{R} \text{ such that } \epsilon > 0,\ \exists N \in \mathbb{N} \text{ such that } \forall n \in \mathbb{N} \text{ satisfying } n > N,\ |a_n - L| < \epsilon$$

is

$$\exists \epsilon \in \mathbb{R} \text{ such that } \epsilon > 0 \text{ and } \forall N \in \mathbb{N},\ \exists n \in \mathbb{N} \text{ such that } n > N \text{ and } |a_n - L| \geq \epsilon.$$

If a statement is denoted by some symbol P, then the negation of P is denoted by $\neg P$.

3.4. **Vacuous truth.** An interesting point arises while dealing with statements about a class. A class which contains no elements at all is called a *vacuous* or *empty* class. For example, the class of all six-legged men is empty because there is no man who has six legs. But now consider the statement

(3) Every six-legged man is happy.

Is this statement true or false? We cannot call it meaningless. It has a definite meaning, just like the statement

Every rich man is happy.

We may call the statement (3) useless, but that does not debar it from being true or false. Which way is it then? Here the reasoning goes as follows. Because of bivalued logic, the statement (3) has to be either true or false, but not both. If it is false, then its negation is true. But the negation is the statement

There exists a six-legged man who is not happy.

But this statement can never be true because there exists no six-legged man whatsoever (and so the question of his being happy or unhappy does not arise at all). So the negation has to be false, and hence the original statement is true!

A layman may hesitate in accepting the above reasoning, and we give some recognition to his hesitation by calling such statements as being *vacuously true*, meaning that they are true simply because there is no example to render them false.

Note, by the way, that the statement

(4) Every six-legged man is unhappy

is also true (albeit vacuously). There is no contradiction here because the statements (3) and (4) are *not* negations of each other.

What is the use of vacuously true statements? Certainly, no mathematician goes on proving theorems which are known to be vacuously true. But such statements sometimes arise as special cases of a more general case.

3.5. **Logical precision in mathematics.** The importance of logic in mathematics cannot be over-emphasized. Logical reasoning being the soul of mathematics, even a single flaw of reasoning can thwart an entire piece of research work. We already pointed out that in mathematics every theorem has to be deduced from the axioms in a strictly deductive manner. Every step has to be justified, and this is the rule for all mathematics. But it deserves to be emphasized here, since in high-school mathematics, the concern was usually with numbers. Consequently the required justifications were based upon some very basic properties of numbers and their specific mention was rarely made. For example if we are to 'solve $(x+3)3(x-3)=30$', we mechanically solve it in the following steps:

$$\begin{aligned}&(x+3)(x-3)=10\\&x^2-9=10\\&x^2=19\\&x=\sqrt{19}\text{ or }-\sqrt{19}.\end{aligned}$$

Although no justification is given for these steps, they require various properties of real numbers such as associativity, commutativity and cancellation laws for multiplication and addition, distributivity of multiplication over addition, and finally, the existence of square roots of real numbers. So far in high-school, we ignored these. But in mathematics, one sometimes considers 'abstract' algebraic systems where some of these laws of associativity, distributivity and so on do not hold. Then the justification for each step will have to be given carefully, starting from the axioms. Hence a proof is needed even when the statement may seem 'obvious'.

4. Sets and functions

It is sometimes said that mathematics *is* the study of sets and functions. This is an oversimplification of matters, but it is true.

4.1. **Sets.** A *set* is a collection of objects considered together. For instance the set of all positive integers, or the set of all rational numbers and so on. The set comprising no element is called the *empty set*, and it is denoted by $\emptyset$. The objects belonging to the set are called its *elements*. For example, 2 is an element belonging to the set of positive integers.

There are two standard methods of specifying a particular set.

Method 1. Whenever it is feasible to do, we can list its elements between braces. Thus $\{1,2,3\}$ is the set comprising the first three positive integers.

This manner of specifying a set, by listing its elements, is unworkable in many circumstances. We then use the second method, which is to use a property that characterizes the elements of the set.

Method 2. If P denotes a certain property of elements, then $\{a \mid P\}$ stands for the set of all elements a for which the property P is true[1]. The set then contains all those elements (and

[1]The symbol '|' is read as 'such that'. Some authors use ':' instead of '|'.

no others) which possess the stated property. For example,

$$\{a \mid a \text{ is real and irrational}\}$$

is the set of all a such that a is real and irrational, that is those real numbers that cannot be written as a quotient of two integers. The set $\{1, 2, 3\}$ can also be described as

$$\{n \mid n \text{ is an integer and } 0 < n < 4\}.$$

We usually denote elements by small letters and sets by capital letters. If a is an element of the set A, then we abbreviate this by writing

$$a \in A.$$

Similarly, if a does not belong to the set A, then we write $a \notin A$.

We say that a set A is a *subset of* a set B if every element belonging to A also belongs to B. We then write

$$A \subset B.$$

(Some authors use the notation $A \subseteq B$ instead of $A \subset B$.) If $A \subset B$, we sometimes say that 'A is contained in B' or that 'B contains A'. For example the set of integers is contained in the set of rational numbers: $\mathbb{Z} \subset \mathbb{Q}$.

For any set A, $A \subset A$ and $\emptyset \subset A$. Two sets A and B are said to be *equal* if they consist of exactly the same elements, and we then write $A = B$. A is equal to B iff $A \subset B$ and $B \subset A$. If $A \subset B$ and $A \neq B$, then we say that 'A is strictly contained in B', or that 'B strictly contains A.'

The *intersection* of sets A and B, denoted by $A \cap B$, is the set of all elements that belong to A and to B:

$$A \cap B = \{a \mid a \in A \text{ and } a \in B\}.$$

If $A \cap B = \emptyset$, then the sets A and B are said to be *disjoint*. For example, the intersection of the set of all integers divisible by 2 and the set of all integers divisible by 3 is the set of all integers divisible by 6. More generally, if $A_1, \dots, A_n$ are sets, then their intersection is the set

$$\{a \mid \text{for all } i \in \{1, \dots n\},\ a \in A_i\}.$$

The intersection of the sets $A_1, \dots, A_n$ is denoted by $\bigcap_{i=1}^{n} A_i$. If we have an infinite family of sets, say, A_n, $n \in \mathbb{N}$, then their intersection is the set

$$\{a \mid \text{for all } i \in \mathbb{N},\ a \in A_i\},$$

which is denoted by $\bigcap_{i=1}^{\infty} A_i$.

The *union* of sets A and B denoted by $A \cup B$, is the set of all elements that belong to A or to B:

$$A \cup B = \{a \mid a \in A \text{ or } a \in B\}.$$

For example, the union of the set of even integers and the set of odd integers is the set of all integers. More generally, if $A_1, \dots, A_n$ are sets, then their union is the set

$$\{a \mid \text{there exists an } i \in \{1, \dots n\} \text{ such that } \ a \in A_i\}.$$

The union of the sets $A_1, \dots, A_n$ is denoted by $\bigcup_{i=1}^{n} A_i$. If we have an infinite family of sets, say, A_n, $n \in \mathbb{N}$, then their union is the set

$$\{a \mid \text{there exists an } i \in \mathbb{N} \text{ such that } a \in A_i\},$$

which is denoted by $\bigcup_{i=1}^{\infty} A_i$.

Given sets A and B, the *product* of A and B is defined as the set of all ordered pairs (a, b), such that a is from A and b is from B. The product of the sets A and B is denoted by $A \times B$. Thus,

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

We do not define an ordered pair, but remark that unless $a = b$, (a, b) is not the same as (b, a). The name 'product' is justified, since if A and B are finite, and have m and n elements, respectively, then the set $A \times B$ has mn elements. Note that as sets $A \times B$ and $B \times A$ are not equal, even though they have the same number of elements. Similarly given sets $A_1, \dots, A_n$, we define $A \times \cdots \times A_n$ by

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

If all the A_i's are equal o the set A, then we denote $A \times \cdots \times A$ by A^n.

4.2. **Functions or maps.** Let A and B be two nonempty sets. A *function* (or a map) is a rule which assigns to each element $a \in A$, an element of the set B.

The set A is called the *domain*, and B is called the *codomain* of the function f. We write

$$f : A \to B$$

where A is the domain and B is the codomain. If $a \in A$, then f takes a to an element in B, and this element from B is denoted by $f(a)$. The element $f(b)$ ($\in B$) is called the *image of a under* f. We sometimes also say that 'f maps a to $f(a)$'. The set of all images is called the *image of* f, and this set is denoted by $f(A)$:

$$f(A) = \{b \in B \mid \text{there exists an } a \in A \text{ such that } f(A) = b\}.$$

Clearly $f(A) \subset B$.

For example, if we take $A = B = \mathbb{Z}$ and consider the rule f that assigns to the integer n the integer n^2, then we obtain a function $f : \mathbb{Z} \to \mathbb{Z}$ given by

$$f(n) = n^2, \quad n \in \mathbb{Z}. \tag{5}$$

We observe that the image of f, is the set $f(\mathbb{Z}) = \{0, 1, 2, \dots\}$ comprising the nonnegative integers, which is strictly contained in the codomain $\mathbb{Z}$. Thus $f(\mathbb{Z}) \subset \mathbb{Q}$, but $f(\mathbb{Z}) \neq \mathbb{Z}$.

Note that while talking about a function, one has to keep in mind that a function really consists of three objects: its domain A, its codomain B and the rule f. Thus for example, if the function $g : \mathbb{Z} \to \mathbb{Q}$ is given by $g(n) = n^2$, $n \in \mathbb{Z}$, then g is a different function from the function $f : \mathbb{Z} \to \mathbb{Z}$ given by (5) above, since the codomain of f is $\mathbb{Z}$, while that of g is $\mathbb{Q}$.

Functions can be between far more general objects than sets comprising numbers. The important thing to remember is that the rule of assignment is such that for each element from the domain there is only one element assigned from the codomain. For example, if we take the set A to be the set of all human beings in the world, and B to be the set of all females on the planet, and $f : A \to B$ to be the function which associates to a person his/her mother. Then we see that f is a function. However, if g is rule which assigns to each person a sister he/she has, then clearly this is not a function, since there are be people with more than one sister (and also there are people who do not have any sister).

Properties of functions play an important role in mathematics, and we highlight two very important types of functions.

A function $f : A \to B$ is said to be *injective* (or *one-to-one*) if

$$\text{for all } a_1 \text{ and } a_2 \text{ in } A \text{ such that } f(a_1) = f(a_2), a_1 = a_2. \tag{6}$$

Equivalently, a function $f : A \to B$ is injective iff

$$\text{for all } a_1 \text{ and } a_2 \text{ in } A \text{ such that } a_1 \neq a_2, f(a_1) \neq f(a_2).$$

This means that for each point b in the image $f(A)$ of a function $f : A \to B$, there is a *unique* point a in the codomain A such that $f(a) = b$. The function $f : \mathbb{Z} \to \mathbb{Z}$ given by (5) is not injective, since for instance for the points $1, -1$ in the domain $\mathbb{Z}$, we see that $1 \neq -1$, but $f(1) = 1 = f(-1)$. However, the function $g : \mathbb{Z} \to \mathbb{Z}$ given by

$$g(n) = 2n, \quad n \in \mathbb{Z},$$

is injective.

A function $f : A \to B$ is said to be *surjective* or *onto* if

$$\text{for all } b \in B, \text{ there exists an } a \text{ in } A \text{ such that } f(a) = b. \tag{7}$$

Note that (7) is equivalent to $f(A) = B$. In other words, a function is surjective if every element from the codomain is the image of some element. The map $f : \mathbb{Z} \to \mathbb{Z}$ given by (5) is not surjective. Indeed, -1 is an element from the codomain, but there is no element n from the domain $\mathbb{Z}$ such that $(f(n) =)\ n^2 = -1$. Consider the map $g : \mathbb{Z} \to \{0, 1, 2, 3, \dots\}$ given by

$$g(n) = n^2, \quad n \in \mathbb{Z}.$$

Then clearly g is surjective.

A function $f : A \to B$ is said to be *bijective* if it is injective and surjective. Thus to check that a map is bijective we have to check two things, injectivity and surjectivity. Consider the map h from the set of all integers to the set of all even integers, given by

$$h(n) = 2n, \quad n \in \mathbb{Z}.$$

Then h is injective and surjective, and so it is bijective.

Finally we mention a situation that arises frequently in mathematics. Let $f : A \to B$ be a function, and suppose that S is a nonempty subset of A. Then one can construct a new function from f as follows. Consider the function $g : S \to B$, defined by

$$g(a) = f(a), \quad a \in S.$$

This function is called the *restriction of* f to S, and it is denoted by $f|_S$.

5. Some common mathematical notation

$\mathbb{N}$ the set of natural numbers $1, 2, 3, \dots$
$\mathbb{Z}$ is the set of integers $\dots, -3, -2, -1, 0, 1, 2, 3, \dots$
$\mathbb{Q}$ is the set of rational numbers
$\mathbb{R}$ is the set of real numbers
$\mathbb{C}$ is the set of complex numbers.

$\forall$ means ‘for all’ or ‘for every’
$\exists$’ means ‘there exists’
$:=$ means ‘is defined to be’ or ‘defined by’