UNIT 14 METHODS OF PROOF

Structure **Page No.**

14.1 INTRODUCTION

In the previous unit you studied about statements and their truth values. In this unit, we shall discuss ways in which statements can be linked to form a logically valid argument. Throughout your mathematical studies you would have come across the terms 'theorem' and 'proof'. In Sec. 14.2, we shall talk about what a theorem is and what constitutes a mathematically acceptable proof.

In Sec.14.3, we shall discuss some ideas formalised by the English mathematician Boole and the German logician Frege (1848 – 1925). These are the different methods used for proving or disproving a statement. As you go through the different types of **valid arguments**, please try and find connections with what we discussed in Block 1.

Fig.1: George Boole (1815 – 1864)

The principle of mathematical induction has a very special place in mathematics because of its simplicity and vast applicability. You will revisit this tool for proving statements in Sec.14.4.

Please go through this unit carefully. You need to be able to convince your learners that its contents are part of the foundation on which all mathematical knowledge is built.

Objectives

After reading this unit, you should be able to develop in your learners the ability to

- explain the terms 'theorem', 'proof' and 'disproof';
- describe the direct method and some indirect methods of proof;
- state and apply both forms of the principle of induction.

14.2 WHAT IS A PROOF?

Suppose I tell somebody, "I am stronger than you." The person is quite likely to turn around, look menacingly at me, and say, "Prove it!" What she or he really wants is to be convinced of my statement by some evidence. (In this case it would probably be a big physical push!)

Convincing evidence is also what the world asks for before accepting a scientist's predictions, or a historian's claims.

In the same way, if you want a mathematical statement to be accepted as true, you would need to provide **mathematically acceptable** evidence to support it. This means that you would need to show that the statement is **universally true**. And this would be done in the form of a logically valid argument.

Definition: An **argument** (in mathematics or logic) is a finite sequence of statements $p_1, \cdots, p_n$, p such that $(p_1 \wedge p_2 \wedge \cdots \wedge p_n) \rightarrow p$.
Each statement in the sequence, $p_1, p_2, \ldots, p_n$ is called a **premise** (or an **assumption**, or a **hypothesis**). The final statement p is called the **conclusion.**

Let's consider an example of an argument that shows that a given statement is true.

Example 1: Give an argument to show that the mathematical statement 'For any two sets A and B, $A \cap B \subseteq A$' is true.

Solution: One argument could be the following.

Let x be an arbitrary element of $A \cap B$.
Then $x \in A$ and $x \in B$, by definition of '$\cap$'.
Therefore, $x \in A$.
This is true for every x in $A \cap B$.
Therefore, $A \cap B \subseteq A$, by definition of '$\subseteq$'.

* * *

The argument in Example 1 has a peculiar nature. The truth of each of the 4 premises and of its conclusion follows from the truth of the earlier premises in it. Of course, we start by assuming that the first statement is true. Then, assuming the definition of 'intersection', the second statement is true. The third one is true, whenever the second one is true because of the properties of logical implication. The fourth statement is true whenever the first three are true, because of the definition and properties of the term 'for all'. And finally, the last statement is true whenever all the earlier ones are. In this way we have shown that the given statement is true. In other words, we have proved the given statement, as the following definitions show.

Definitions: We say that a proposition p **follows logically from** propositions $p_1, p_2, \cdots, p_n$ if p must be true whenever $p_1, p_2, \cdots, p_n$ are true, i.e., $(p_1 \wedge p_2 \wedge \cdots \wedge p_n) \Longrightarrow p$.
[Here, **note** the use of the implication arrow '$\Rightarrow$'. For any two propositions r and s, **'r $\Rightarrow$ s' denotes 's is true whenever** r **is true.'** Note that, using the contrapositive, this also denotes 'r is false whenever s is false'. Thus '$r \rightarrow s$' and '$r \Longrightarrow s$' are different except when both r and s are true or both are false.]

A **proof** of a proposition p is a mathematical argument consisting of a sequence of statements $p_1, p_2, \cdots, p_n$ from which p logically follows. So, p is the conclusion of this argument.

The statement that is proved to be true is called a **theorem**.

Sometimes, as you will see in Sec.14.3.3, instead of showing that a statement p is true, we try to prove that it is false, i.e., that $\sim p$ is true. Such a proof is called a **disproof** of p. In the next section you will read about some ways of disproving a statement.

Sometimes it happens that we feel a certain statement is true, but we don't succeed in proving it. It may also happen that we can't disprove it. Such statements are called **conjectures**. If and when a conjecture is proved, it would be called a theorem. If it is disproved, then its negative will be a theorem!

In this context, there's a very famous conjecture which was made by a mathematician **Goldbach** in 1742. He stated that :
For every $n \in \mathbf{N}$, if n is even and $n > 2$, then n is the sum of two primes.

To this day, no one has been able to prove it or disprove it. To disprove it several people have been hunting for an example for which the statement is not true, i.e., an even number $n > 2$ such that n cannot be written as the sum of two prime numbers.

Now, as you have seen, a mathematical proof of a statement consists of one or more premises. These premises could be of four types:

i) a proposition that has been proved earlier (e.g., to prove that the complex roots of a polynomial in $\mathbf{R}[x]$ occur in pairs, we use the division algorithm); or

ii) a proposition that follows logically from the earlier propositions given in the proof (as you have seen in Example 1); or

iii) a mathematical fact that has never been proved, but is universally accepted as true (e.g., two points determine a line). Such a fact is called an **axiom** (or **a postulate**);

iv) the definition of a mathematical term (e.g., assuming the definition of '$\subseteq$' in the proof of $A \cap B \subseteq A$).

You will come across more examples of each type while doing the following exercises, and while going through proofs in this course and other courses.

E1) Write down an example of a theorem, and its proof (of at least 4 steps), taken from school-level algebra. At each step, indicate which of the four types of premise it is.

E2) Is every statement a theorem? Why?

So far we have spoken about valid, or acceptable, arguments. Now let us see an example of a sequence of statements that will **not** form a valid argument. Consider the following sequence.

If Maya sees the movie, she won't finish her homework.
Maya won't finish her homework.
Therefore, Maya sees the movie.

Looking at the argument, can you say whether it is valid or not? Intuitively you may feel that the argument isn't valid. But, is there a formal logical tool that you can apply to check if your intuition is correct? What about truth tables? Let's see.

The given argument is of the form
$[(p \to q) \wedge q] \Rightarrow p$
where
p: Maya sees the movie, and
q: Maya won't finish her homework.

Let us look at the truth table related to this argument (see Table 1).

Table 1

p	q	$p \to q$	$(p \to q) \wedge q$
T	T	T	T
T	F	F	F
F	T	T	T
F	F	T	F

The last column gives the truth values of the premises. The first column gives the corresponding truth values of the conclusion. Now, **the argument will only be valid if whenever both the premises are true, the conclusion is true.** This happens in the first row, but **not** in the third row.
Therefore, the argument is **not valid**.

Why don't you check an argument for validity now?

E3) Check whether the following argument is valid.
$(p \to q \vee \sim r) \wedge (q \to p) \Rightarrow (p \to r)$

You have seen that a proof is a logical argument that verifies the truth of a theorem. There are several ways of proving a theorem, as you will see in the next section. All of them are based on one or more **rules of inference**, which are different forms of arguments. We shall now present four of the most commonly used rules.

'Modus ponens' is a Latin term which means 'method of affirmation'.

i) **Law of detachment** (or **modus ponens**)
Consider the following argument:

If Kali can draw, she will get a job.
Kali can draw.
Therefore, she will get a job.

To study the form of the argument, let us take p to be the proposition 'Kali can draw.' and q to be the proposition 'Kali will get a job.'. Then the premises are $(p \to q)$ and p. The conclusion is q.
So, the form of the argument is

$$\begin{array}{l} p \to q \\ \underline{p} \\ \therefore q \end{array} \quad , \quad \text{i.e., } [(p \to q) \wedge p] \Rightarrow q.$$

$\therefore$ denotes 'therefore'.

Is this argument valid? To find out, let's construct its truth table (see Table 2).

Table 2: Truth table for $[(p \to q) \wedge p] \Rightarrow q$

p	q	$p \to q$	$(p \to q) \wedge p$
T	T	T	T
T	F	F	F
F	T	T	F
F	F	T	F

In the table, look at the second column (the conclusion) and the fourth column (the premises). Whenever the premises are true, i.e., in Row 1, the conclusion is true. Therefore, the argument is valid.
This form of valid argument is called the law of detachment because the conclusion q is detached from a premise (namely, $p \to q$). It is also called the **law of direct inference**.

'Modus tollens' means 'method of denial'.

ii) **Law of contraposition** (or **modus tollens**)
To understand this law, consider the following argument:

If Kali can draw, then she will get a job.
Kali will not get a job.
Therefore, Kali can't draw.

Taking p and q as in (i) above, you can see that the premises are $p \to q$

and $\sim q$. The conclusion is $\sim p$.

So the argument is

$$\begin{array}{l} p \to q \\ \underline{\sim q} \\ \therefore \sim p \end{array}, \quad \text{i.e., } [(p \to q) \wedge \sim q] \Rightarrow \sim p.$$

If you check, you'll find that this is a valid form of argument.

There are two more rules of inference that most commonly form the basis of several proofs. The following exercise is about them.

E4) You will find three arguments below. Convert each of them into the language of symbols, and check if they are valid.

i) Either the eraser is white or oxygen is a metal.
The eraser is black.
Therefore, oxygen is a metal.

ii) If Madhu is a 'sarpanch', she will head the 'panchayat'.
If Madhu heads the 'panchayat', she will decide on property disputes.
Therefore, if Madhu is a 'sarpanch', she will decide on property disputes.

iii) Either Munna will cook or Munni will practise Karate.
If Munni practises Karate, then Munna studies.
Munna does not study.
Therefore, Munni will practise Karate.

E5) Write down one example each of modus ponens and modus tollens.

As you must have discovered, the arguments in E4(i) and (ii) are valid. The first one is an example of a **disjunctive syllogism**. The second one is an example of a **hypothetical syllogism**.

Thus, a disjunctive syllogism is of the form

$$\begin{array}{l} p \vee q \\ \underline{\sim p} \\ \therefore q \end{array}, \quad \text{i.e., } [(p \vee q) \wedge \sim p] \Rightarrow q.$$

And, a hypothetical syllogism is of the form

$$\begin{array}{l} p \to q \\ \underline{q \to r} \\ \therefore p \to r \end{array}, \quad \text{i.e., } [(p \to q) \wedge (q \to r)] \Rightarrow (p \to r).$$

Let us now see how different forms of arguments can be put together to prove or disprove a statement.

14.3 DIFFERENT METHODS OF PROOF

In this section we shall consider three different strategies for proving a statement. We will also discuss a method that is used only for disproving a statement.

Let us start with a proof strategy based on the first rule of inference that we discussed in the previous section.

14.3.1 Direct Proof

This form of proof is based entirely on modus ponens. Let us formally spell out the strategy.

Definition : A **direct proof** of $p \Rightarrow q$ is a logically valid argument that begins with the assumption that p is true and, in one or more applications of the law of detachment, concludes that q must be true.

So, to construct a direct proof of $p \Rightarrow q$, we start by assuming that p is true. Then, in one or more steps of the form $p \Rightarrow q_1, q_1 \Rightarrow q_2, \ldots, q_n \Rightarrow q$, we conclude that q is true. Consider the following examples.

Example 2: Give a direct proof of the statement 'The product of two odd integers is odd.'.

Solution: Let us clearly analyse what our hypotheses are, and what we have to prove.
We start by considering any two odd integers x and y. So our hypothesis is
p: x and y are odd.
The conclusion we want to reach is
q : xy is odd.
Let us first prove that $p \Rightarrow q$.
Since x is odd, $x = 2m + 1$ for some integer m.
Similarly, $y = 2n + 1$ for some integer n.
Then $xy = (2m + 1)(2n + 1) = 2(2mn + m + n) + 1$.
Therefore, xy is odd.
So we have shown that $p \Rightarrow q$.

Now we can apply modus ponens to $p \wedge (p \Rightarrow q)$ to get the required conclusion.

Note that the essence of this direct proof lies in showing $p \Rightarrow q$.

* * *

Example 3: Give a direct proof of the theorem 'The square of an even integer is an even integer.'.

Solution: First of all, let us write the given statement symbolically, as
$(\forall\, x \in \mathbf{Z})(p(x) \Rightarrow q(x))$
where p(x) : x is even, and
$q(x) : x^2$ is even, i.e., q(x) is the same as $p(x^2)$.

The direct proof, then goes as follows:

Let x be an even number (i.e., we assume p(x) is true).
Then $x = 2n$, for some integer n (we apply the definition of an even number).
Then $x^2 = (2n)^2 = 4n^2 = 2(2n^2)$.
$\therefore\ x^2$ is even (i.e., q(x) is true).

Note that we have proved the statement for every x since we have treated x as an arbitrary even number and not a particular value.

* * *

Why don't you try an exercise now?

E6) Give a direct proof of the statement 'If x is a real number such that $x^2 = 9$, then either $x = 3$ or $x = -3$.'.

Let us now consider another proof strategy.

14.3.2 Indirect Proofs

In this sub-section we shall consider two roundabout methods for proving $p \Rightarrow q$.

Proof by contrapositive: In the first method, we use the fact that the proposition $p \Rightarrow q$ is logically equivalent to its **contrapositive** $(\sim q \Rightarrow \sim p)$, i.e.,
$(p \Rightarrow q) \equiv (\sim q \Rightarrow \sim p)$.

For instance, 'If Ammu does not agree with communalists, then she is not orthodox.' is the same as 'If Ammu is orthodox, then she agrees with communalists.'.

Because of this equivalence, to prove $p \Rightarrow q$, we can, instead, prove $\sim q \Rightarrow \sim p$. This means that we can assume that $\sim q$ is true, and then try to prove that $\sim p$ is true. In other words, **what we do to prove $p \Rightarrow q$ in this method is to assume that q is false and then show that p is false.** Let us consider an example.

Example 4: Prove that 'If $x, y \in \mathbf{Z}$ such that xy is odd, then both x and y are odd.', by proving its contrapositive.

Solution: Let us name the statements involved as below.
p : xy is odd
q : both x and y are odd.

So,
$\sim p$: xy is even, and
$\sim q$: x is even or y is even, or both are even.

We want to prove $p \Rightarrow q$, by proving that $\sim q \Rightarrow \sim p$. So we start by assuming that $\sim q$ is true, i.e., we suppose that x is even.
Then $x = 2n$ for some $n \in \mathbf{N}$.
Therefore, $xy = 2ny$.
Therefore, xy is even, by definition.
That is, $\sim p$ is true.

So, we have shown that $\sim q \Rightarrow \sim p$. Therefore, $p \Rightarrow q$.

* * *

Why don't you ask your students to try some related exercises now?

E7) Write down the contrapositive of the statement 'If f is a 1-1 function from a finite set X into itself, then f must be surjective.'.

E8) Prove the statement 'If x is an integer and x^2 is even, then x is also even.' by proving its contrapositive.

And now let us consider the other way of proving a statement indirectly.

Proof by contradiction: In this method, to prove q is true, we start by assuming that q is false (i.e., $\sim q$ is true). Then, by a logical argument we arrive at a situation where a statement is true as well as false, i.e., we reach a contradiction $r \wedge \sim r$ for some statement r. This means that the truth of $\sim q$ implies a contradiction, a statement that is always false. This can only happen when $\sim q$ is false also. Therefore, q must be true.

This method is called **proof by contradiction**. It is also called **reductio ad absurdum** (a Latin phrase) because it relies on reducing a given assumption to an absurdity.

Let us consider an example of the use of this method.

Example 5: Show that $\sqrt{5}$ is irrational.

Solution: Let us try and prove the given statement by contradiction. For this, we begin by assuming that $\sqrt{5}$ is rational. This means that there exist positive integers a and b such that $\sqrt{5} = \frac{a}{b}$, where a and b have no common factors.

This implies $a = \sqrt{5}b \Rightarrow a^2 = 5b^2 \Rightarrow 5|a^2 \Rightarrow 5|a$.
Therefore, by definition, $a = 5c$ for some $c \in \mathbf{Z}$.
Therefore, $a^2 = 25c^2$.

But $a^2 = 5b^2$ also.
So $25c^2 = 5b^2 \Rightarrow 5c^2 = b^2 \Rightarrow 5|b^2 \Rightarrow 5|b$.

But now we find that 5 divides both a and b, which **contradicts** our earlier assumption that a and b have no common factor.
Therefore, we conclude that our assumption that $\sqrt{5}$ is rational is false, i.e., $\sqrt{5}$ is irrational.

* * *

We can also use the method of contradiction to prove an implication $r \Rightarrow s$. Here we can use the equivalence $\sim (r \rightarrow s) \equiv r \wedge \sim s$. So, to prove $r \Rightarrow s$, we can begin by assuming that $r \Rightarrow s$ is false, i.e., r is true and s is false. Then we can present a valid argument to arrive at a contradiction.

Consider the following example from plane geometry.

Example 6: Prove the following:
If two distinct lines L_1 and L_2 intersect, then their intersection consists of exactly one point.

Solution: To prove the given implication by contradiction, let us begin by assuming that the two distinct lines L_1 and L_2 intersect in more than one point. Let us call two of these distinct points A and B. Then, both L_1 and L_2 contain A and B. This contradicts the axiom from geometry that says 'Given two distinct points, there is exactly one line containing them.'.

Therefore, if L_1 and L_2 intersect, then they must intersect in only one point.

* * *

The contradiction rule is also used for solving many logical puzzles by discarding all solutions that reduce to contradictions. Consider the following example.

Example 7: There is a village that consists of two types of people — those who always tell the truth, and those who always lie. Suppose that you visit the village and two villagers A and B come up to you. Further, suppose
A says, "B always tells the truth," and
B says, "A and I are of opposite types."
What types are A and B?

Solution: Let us start by assuming A is a truth-teller.
$\therefore$ What A says is true.
$\therefore$ B is a truth-teller.
$\therefore$ What B says is true.
$\therefore$ A and B are of opposite types.
This is a contradiction, because our premises say that A and B are both truth-tellers.
$\therefore$ The assumption we started with is false.
$\therefore$ A always tells lies.
$\therefore$ What A has told you is a lie.
$\therefore$ B always tells lies.
$\therefore$ A and B are of the same type, i.e., both of them always lie.

* * *

Here are a few exercises for you now. While doing them you would realise that there are situations in which all the three methods of proof we have discussed so far can be used.

E9) Use the method of proof by contradiction to show that
i) $\sqrt{p}$ is irrational, for any prime p.
ii) For $x \in \mathbf{R}$, if $x^3 + 4x = 0$, then $x = 0$.

E10) Prove E 9(ii) directly as well as by the method of contrapositive.

There can be several ways of proving a statement.

E11) Suppose you are visiting the village described in Example 7 above. Another two villagers C and D approach you. C tells you, "Both of us always tell the truth," and D says, "C always lies." What types are C and D?

Let us now consider the problem of showing that a statement is false.

14.3.3 Counterexamples

Suppose I make the statement 'All human beings are 5 feet tall.'. You are quite likely to show me an example of a human being standing nearby for whom the statement is not true. And, as you know, the moment we have even one example for which the statement $(\forall x)p(x)$ is false [i.e.,$(\exists x)(\sim p(x))$ is true], then the statement is false.

An example that shows that a statement is false is a **counterexample** to such a statement. The name itself suggests that it is an example to counter a given statement.

A common situation in which we look for counterexamples is to disprove statements of the form $p \rightarrow q$. From Unit 13, you know that $\sim (p \rightarrow q) \equiv p \wedge \sim q$. Therefore, a counterexample to $p \rightarrow q$ needs to be an example where $p \wedge \sim q$ is true, i.e., p is true and $\sim q$ is true, i.e., the hypothesis p holds but the conclusion q does not hold.

For instance, to disprove the statement 'If n is an odd integer, then n is prime.', we need to look for an odd integer which is not a prime number. 15 is one such integer. So, $n = 15$ is a counterexample to the given statement.

Notice that **a counterexample to a statement p proves that** p is false, i.e., **$\sim$ p is true.**

Let us consider another example.

Example 8: Disprove the following statement:
$(\forall\, a \in \mathbf{R})(\forall\, b \in \mathbf{R})[(a^2 = b^2) \Rightarrow (a = b)]$.

Solution: A good way of disproving it is to look for a counterexample, that is, a pair of real numbers a and b for which $a^2 = b^2$ but $a \neq b$. Can you think of such a pair? What about $a = 1$ and $b = -1$? They serve the purpose.

In fact, there are infinitely many counterexamples. (Why?)

* * *

Now, an exercise!

E12) Disprove the following statements by providing a suitable counterexample.

i) $\forall\, x \in \mathbf{Z}, x \in \mathbf{Q} \setminus \mathbf{N}$.
ii) $(x+y)^n = x^n + y^n \;\forall\, n \in \mathbf{N}, x, y \in \mathbf{Z}$.
iii) $f : \mathbf{N} \rightarrow \mathbf{N}$ is 1-1 iff f is onto.
(**Hint:** To disprove $p \Leftrightarrow q$ it is enough to prove that $p \Rightarrow q$ is false or $q \Rightarrow p$ is false.)

There are some other strategies of proof, like a **constructive** proof, which you must have come across in other mathematics courses. We shall not discuss this method here.

Other proof-related adjectives that you will come across are **vacuous** and **trivial**.

A **vacuous proof** makes use of the fact that if p is false, then $p \to q$ is true, regardless of the truth value of q. So, to vacuously prove $p \to q$, all we need to do is to show that p is false. For instance, suppose we want to prove that 'If $n > n+1$ for $n \in \mathbf{Z}$, then $n^2 = 0$.'. Since '$n > n+1$' is false for every $n \in \mathbf{Z}$, the given statement is **vacuously true**, or **true by default**.

Similarly, a **trivial proof** of $p \to q$ is one based on the fact that if q is true, then $p \to q$ is true, regardless of the truth value of p. So, for example, 'If $n > n+1$ for $n \in \mathbf{Z}$, then $n+1 > n$' is trivially true since $n+1 > n \;\forall\; n \in \mathbf{Z}$. The truth value of the hypothesis (which is false in this example) does not come into the picture at all.

Here's a chance for you to think up such proofs now!

E13) Give one example each of a vacuous proof and a trivial proof.

And now let us study a very important technique of proof for statements that are of the form $p(n), n \in \mathbf{N}$.

14.4 PRINCIPLE OF INDUCTION

In a discussion with some students the other day, one of them told me very cynically that all Indian politicians are corrupt. I asked him how he had reached such a conclusion. As an argument he gave me instances of several politicians, all of whom were known to be corrupt. What he had done was to formulate his general opinion of politicians on the basis of several particular instances. This is an example of **inductive logic**, a process of reasoning by which general rules are discovered by the observation of several individual cases. Inductive reasoning is used in all the sciences, including mathematics. But in mathematics we use a more precise form.

Precision is required in mathematical induction because, as you know, a statement of the form $(\forall\; n \in \mathbf{N})p(n)$ is true only if it can be shown to be true for each n in **N**. (In the example above, even if the student is given an example of one clean politician, he is not likely to change his general opinion.)

How can we make sure that our statement p(n) is true for each n that we are interested in? To answer this, let us consider an example.

Suppose we want to prove that $1+2+3+\cdots+n = \frac{n(n+1)}{2}$ for each $n \in \mathbf{N}$. Let us call p(n) the predicate '$1+2+\cdots+n = \frac{n(n+1)}{2}$'. Now, we can verify that it is true for a few values, say, $n = 1, n = 5, n = 10, n = 100$, and so on. But we still can't be sure that it will be true for some value of n that we haven't tried.

But now, suppose we can show that if p(n) is true for some $n, n = k$ say, then it will be true for $n = k+1$. Then we are in a very good position because we already know that p(1) is true. And, since p(1) is true, so is $p(1+1)$, i.e., p(2), and so on. In this way we can show that p(n) is true for every $n \in \mathbf{N}$. So, our proof boils down to two steps, namely,

i) Checking that p(1) is true;
ii) Proving that whenever p(k) is true, then $p(k+1)$ is true, where $k \in \mathbf{N}$.

This is the principle that we will now state formally, in a more general form.

Principle of Mathematical Induction (PMI): Let p(n) be a predicate involving a natural number n. Suppose the following two conditions hold:

i) p(m) is true for some $m \in \mathbf{N}$;
ii) If p(k) is true, then $p(k+1)$ is true, where $k(\geq m)$ is any natural number.

Then p(n) is true for every $n \geq m$.

Looking at the two conditions in the principle, can you make out why it works? (As a hint, put $m = 1$ in our example above.)
Well, (i) tells us that p(m) is true. Then putting $k = m$ in (ii), we find that $p(m+1)$ is true. Again, since $p(m+1)$ is true, $p(m+2)$ is true, and so on.

Going back to the example above, let us complete the second step. We know that p(k) is true, i.e., $1 + 2 + \cdots + k = \frac{k(k+1)}{2}$. We want to check if $p(k+1)$ is true. So let us find

$$\begin{aligned} 1 + 2 + \cdots + (k+1) &= (1 + 2 + \cdots + k) + (k+1) \\ &= \frac{k(k+1)}{2} + (k+1), \text{ since } p(k) \text{ is true} \\ &= \frac{(k+1)(k+2)}{2} \end{aligned}$$

So, $p(k+1)$ is true.

And so, by the principle of mathematical induction, we know that p(n) is true for every $n \in \mathbf{N}$.

What does this principle really say? It says that if you can walk a few steps, say m steps, and if at each stage you can walk one more step, then you can walk any distance. It sounds very simple, but you may be surprised to know that the technique in this principle was first used by Europeans only as late as the 16th century by the Venetian F. Maurocylus (1494 – 1573). He used it to show that $1 + 3 + \cdots + (2n-1) = n^2$. Pierre de Fermat (1601 – 1665) improved on the technique and proved that this principle is equivalent to the following often-used principle of mathematics.

The Well-ordering Principle: Any non-empty subset of **N** contains a smallest element.

You may be able to see the relationship between the two principles if we reword the PMI in the following form.

Principle of Mathematical Induction (Equivalent form): Let $S \subseteq \mathbf{N}$ be such that

i) $m \in S$
ii) For each $k \in \mathbf{N}, k \geq m$, the following implication is true: $k \in S \Rightarrow k+1 \in S$.

Then $S = \{m, m+1, m+2, \cdots\}$.

The term 'mathematical induction' was first used by De Morgan.

Can you see the equivalence of the two forms of the PMI? If you take
$S = \{n \in \mathbf{N} \mid p(n) \text{ is true }\}$,
then you can see that the way we have written the principle above is a mere rewrite of the earlier form.

Now, let us consider an example of proof using PMI.

Example 9: Use mathematical induction to prove that
$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n}{6}(n+1)(2n+1) \ \forall \ n \in \mathbf{N}.$$

Solution: We call p(n) the predicate
$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n}{6}(n+1)(2n+1).$$
Since we want to prove it for every $n \in \mathbf{N}$, we take $m = 1$.

Note that p(n) is a predicate, not a statement, unless we know the value of n.

Step 1: p(1) is $1^2 = \frac{1}{6}(1+1)(2+1)$, which is true.

Step 2: Suppose, for an arbitrary $k \in \mathbf{N}, p(k)$ is true, i.e.,
$1^2 + 2^2 + \cdots + k^2 = \frac{k}{6}(k+1)(2k+1)$ is true.

Step 3: To check if the assumption in Step 2 implies that $p(k+1)$ is true. Let's see.

$$p(k+1) \text{ is } 1^2 + 2^2 + \cdots + k^2 + (k+1)^2 = \frac{k+1}{6}(k+2)(2k+3)$$

$$\Longleftrightarrow (1^2 + 2^2 + \cdots + k^2) + (k+1)^2 = \frac{k+1}{6}(k+2)(2k+3)$$

$$\Longleftrightarrow \frac{k}{6}(k+1)(2k+1) + (k+1)^2 = \frac{k+1}{6}(k+2)(2k+3),$$

since $p(k)$ is true.

$$\Longleftrightarrow \frac{k+1}{6}[k(2k+1) + 6(k+1)] = \frac{k+1}{6}(k+2)(2k+3)$$

$\Longleftrightarrow 2k^2 + 7k + 6 = (k+2)(2k+3)$, dividing throughout by $\frac{k+1}{6}$, which is true.

So, $p(k)$ is true implies that $p(k+1)$ is true.

So, both the conditions of the principle of mathematical induction hold. Therefore, its conclusion must hold, i.e., $p(n)$ is true for every $n \in \mathbf{N}$.

* * *

Have you gone through Example 9 carefully? If so, you would have noticed that the proof consists of three steps:

Step 1 (called the **basis of induction**): Checking if $p(m)$ is true for some $m \in \mathbf{N}$.

Step 2 (called the **induction hypothesis**): Assuming that $p(k)$ is true for an arbitrary $k \in \mathbf{N}, k \geq m$.

Step 3 (called the **induction step**): Showing that $p(k+1)$ is true, by a direct or an indirect proof.

Now let us consider an example in which $m \neq 1$.

Example 10: Show that $2^n > n^3$ for $n \geq 10$.

Solution: We write $p(n)$ for the predicate '$2^n > n^3$'.

Step 1: For $n = 10, 2^{10} = 1024$, which is greater that 10^3. Therefore, $p(10)$ is true.

Step 2: We assume that $p(k)$ is true for an arbitrary $k \geq 10$. Thus, $2^k > k^3$.

Step 3: Now, we want to prove that $2^{k+1} > (k+1)^3$. Note that

$$\begin{aligned} 2^{k+1} &= 2.2^k > 2.k^3, \text{by our assumption} \\ &> \left(1 + \frac{1}{10}\right)^3 .k^3, \text{ since } 2 > \left(1 + \frac{1}{10}\right)^3 \\ &\geq \left(1 + \frac{1}{k}\right)^3 .k^3, \text{ since } k \geq 10 \\ &= (k+1)^3. \end{aligned}$$

Thus, $p(k+1)$ is true if $p(k)$ is true for $k \geq 10$.

Therefore, by the principle of mathematical induction, $p(n)$ is true $\forall\, n \geq 10$.

* * *

Why don't you try to apply the principle now?

E14) Use mathematical induction to prove that

$$1 + \frac{1}{4} + \frac{1}{9} + \cdots + \frac{1}{n^2} \leq 2 - \frac{1}{n} \;\forall\, n \in \mathbf{N}.$$

E15) Show that for any integer $n > 1$, $\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \cdots + \frac{1}{\sqrt{n}} > \sqrt{n}$.
(**Hint** : The basis of induction is p(2).).

Before going further **a note of warning**! To prove that p(n) is true $\forall\, n \geq m$, **both** the basis of induction **as well as** the induction step must hold.
If even one of these conditions does not hold, we cannot arrive at the conclusion that p(n) is true $\forall\, n \geq m$.

For example, suppose p(n) is $(x + y)^n \leq x^n + y^n \;\forall\, x, y \in \mathbf{R}$. Then p(1) is true. But Steps 2 and 3 do not hold. Therefore, p(n) is not true for every $n \in \mathbf{N}$. (Can you find a value of n for which p(n) is false?)

As another example, take p(n) to be the statement '$1 + 2 + \cdots + n < n$'. Then, if p(k) is true, so is p(k + 1) (prove it!). But the basis step does not hold for any $m \in \mathbf{N}$. And, as you can see, p(n) is false.

Now let us look at a situation in which we may expect the principle of induction to work, but it doesn't. Consider the sequence of numbers $1, 1, 2, 3, 5, 8, \cdots$. These are the **Fibonacci numbers**, named after the Italian mathematician Fibonacci. Each term in the sequence, from the third term on, is obtained by adding the previous 2 terms. So, if a_n is the nth term, then $a_1 = 1, a_2 = 1$, and $a_n = a_{n-1} + a_{n-2} \;\forall\, n \geq 3$.
Suppose we want to show that $a_n < 2^n \;\forall\, n \in \mathbf{N}$ using the PMI. Then, if p(n) is the predicate $a_n < 2^n$, we know that p(1) is true.
Now suppose we know that p(k) is true for an arbitrary $k \in \mathbf{N}$, i.e., $a_k < 2^k$. We want to show that $a_{k+1} < 2^{k+1}$, i.e., $a_k + a_{k-1} < 2^{k+1}$. But we don't know anything about a_{k-1}. So, how can we apply the principle of induction in the form that we have stated it? In such a situation, a stronger, more powerful, version of the principle of induction comes in handy. Let's see what this is.

Principle of Strong Mathematical Induction: Let p(n) be a predicate that involves a natural number n. Suppose we can show that
i) p(m) is true for some $m \in \mathbf{N}$, and
ii) whenever $p(m), p(m + 1), \cdots, p(k)$ are true, then p(k + 1) is true, where $k \geq m$.
Then we can conclude that p(n) is true for all natural numbers $n \geq m$.

Why do we call this principle stronger than the earlier one? This is because, in the induction step we are making more assumptions, i.e., that **p(n) is true for every n lying between m and k**, not just that p(k) is true.

In using the strong form we often need to check Step 1 for more than one value of n.

Let us now go back to the Fibonacci sequence. To use the strong form of the PMI, we take m = 1. We have seen that p(1) is true. We also need to see if p(2) is true. This is because we have to use the relation $a_n = a_{n-1} + a_{n-2}$, which is valid for $n \geq 3$.

Now that we know that both p(1) and p(2) are true, let us go to the next step. In Step 2, for an arbitrary $k \geq 2$, we assume that p(n) is true for every n such that $1 \leq n \leq k$, i.e., $a_n < 2^n$ for $1 \leq n \leq k$.

Finally, in Step 3, we must show that p(k + 1) is true, i.e., $a_{k+1} < 2^{k+1}$. Now

$$\begin{aligned} a_{k+1} &= a_k + a_{k-1} \\ &< 2^k + 2^{k-1}, \text{ by our assumption in Step 2.} \\ &= 2^{k-1}(2 + 1) \\ &< 2^{k-1}.2^2 \\ &= 2^{k+1} \end{aligned}$$

$\therefore$ p(k + 1) is true.
$\therefore$ p(n) is true $\forall\, n \in \mathbf{N}$.

Though the "strong" form of the PMI appears to be different from the "weak" form, **the two are actually equivalent.** This is because each can be obtained from the other. So, we can use either form of mathematical induction. In a given problem we use the form that is more suitable. For instance, in the following example, as in the case of the one above, you would agree that it is better to use the strong form of the PMI.

Example 11: Use induction to prove that any integer $n \geq 2$ is either a prime or a product of primes.

Solution: Here $p(n)$ is the predicate 'n is a prime or n is a product of primes.'.

Step 1 **(basis of induction)** : Since 2 is a prime, $p(2)$ is true.

Step 2 **(induction hypothesis)**: Assume that $p(n)$ is true for any integer n such that $2 \leq n \leq k$, i.e., $p(3), p(4), \cdots, p(k)$ are true.

Step 3 **(induction step)**: Now consider $p(k+1)$. If $k+1$ is a prime, then $p(k+1)$ is true. If $k+1$ is not a prime, then $k+1 = rs$, where $2 \leq r \leq k$ and $2 \leq s \leq k$. But, by our induction hypothesis, $p(r)$ is true and $p(s)$ is true. Therefore, r and s are either primes or products of primes. And therefore, $k+1$ is a product of primes. So, $p(k+1)$ is true.

Therefore, $p(n)$ is true $\forall\, n \geq 2$.

* * *

Why don't you try some exercises now?

E16) If $a_1, a_2, \cdots$ are the terms in the Fibonacci sequence, use the weak as well as the strong forms of the principle of mathematical induction to show that $a_n > \frac{3}{2}\ \forall\, n \geq 3$. Which form did you find more convenient?

E17) Consider the following "proof" by induction of the statement 'Any n marbles are of the same size.', and say why it is wrong.

Basis of induction : For $n = 1$, the statement is clearly true.

Induction hypothesis : Assume that the statement is true for $n = k$.

Induction step : Now consider any $k+1$ marbles $1, 2, \cdots, k+1$. By the induction hypothesis the k marbles $2, 3, \cdots, k+1$ are of the same size. Therefore, all the $k+1$ marbles are of the same size.

Therefore, the given statement is true for every n.

E18) Prove that the following result is equivalent to the principle of mathematical induction (strong form):

Let $S \subseteq \mathbf{N}$ such that

i) $m \in S$

ii) If $m, m+1, m+2, \cdots, k$ are in S, then $k+1 \in S$.

Then $S = \{n \in \mathbf{N} | n \geq m\}$.

E19) To prove that $\sum_{i=1}^{n} \frac{1}{\sqrt{i}} \leq 2\sqrt{n} - 1\ \forall\, n \in \mathbf{N}$, which form of the principle of mathematical induction would you use, and why? Also, prove the inequality.

With this we come to the end of our discussion on various techniques of proving or disproving mathematical statements. Let us take a brief look at what you have read in this unit.

14.5 SUMMARY

In this unit you have studied the following points.

1. What constitutes a proof of a mathematical statement, including 4 commonly used rules of inference, namely,
   i) law of detachment (or modus ponens) : $[(p \to q) \wedge p] \Rightarrow q$
   ii) law of contraposition (or modus tollens) : $[(p \to q) \wedge \sim q] \Rightarrow \sim p$
   iii) disjunctive syllogism : $[(p \vee q) \wedge \sim p] \Rightarrow q$
   iv) hypothetical syllogism : $[(p \to q) \wedge (q \to r)] \Rightarrow (p \to r)$
2. The description and examples of a direct proof, which is based on modus ponens.
3. Two types of indirect proofs : proof by contrapositive and proof by contradiction.
4. The use of counterexamples for disproving a statement.
5. The "strong" and "weak" forms of the principle of mathematical induction, and their equivalence with the well-ordering principle.

14.6 COMMENTS ON EXERCISES

E1) For example,
Theorem: $(x+y)^2 = x^2 + 2xy + y^2$ for $x, y \in \mathbf{R}$.
Proof: For $x, y \in \mathbf{R}$, $(x+y)^2 = (x+y)(x+y)$ (by definition of 'square')
$(x+y)(x+y) = x(x+y) + y(x+y)$ (by distributivity, which has been proved earlier)
$x(x+y) + y(x+y) = x^2 + 2xy + y^2$ (again by distributivity, and by definition of addition and multiplication of algebraic terms).
Therefore, $(x+y)^2 = x^2 + 2xy + y^2$ (using an earlier proved statement that $a = b$ and $b = c$ implies that $a = c$).

E2) No, not unless it has been proved to be true.

E3) premises ↓ ↓ conclusion ↓

p	q	r	$\sim r$	$q \vee \sim r$	$p \to q \vee \sim r$	$q \to p$	$p \to r$
T	T	T	F	T	T	T	T
T	T	F	T	T	T	T	F
T	F	T	F	F	F	T	T
T	F	F	T	T	T	T	F
F	T	T	F	T	T	F	T
F	T	F	T	T	T	F	T
F	F	T	F	F	T	T	T
F	F	F	T	T	T	T	T

The premises are true in Rows 1, 2, 4, 7, 8. So, the argument will be valid if the conclusion is also true in these rows. But this does not happen in Row 2, for instance. Therefore, the argument is invalid.

E4) i) Let p : The eraser is white,
q : Oxygen is a metal.
Then the argument is

$$\frac{\begin{array}{l} p \vee q \\ \sim p \end{array}}{\therefore q}$$

Its truth table is given below.

	conclusion	premises	premises
p	q	$\sim$ p	p $\vee$ q
T	T	F	T
T	F	F	T
F	T	T	T
F	F	T	F

All the premises are true only in the third row. Since the conclusion in this row is also true, the argument is valid.

ii) The argument is $(p \to q) \wedge (q \to r) \Longrightarrow (p \to r)$
where p: Madhu is a 'sarpanch',
q : Madhu heads the 'Panchayat',
r : Madhu decides on property disputes.
This is valid because, whenever both the premises are true, so is the conclusion (see the following table.)

			premises	premises	conclusion
p	q	r	p $\to$ q	q $\to$ r	p $\to$ r
T	T	T	T	T	T
T	T	F	T	F	F
T	F	T	F	T	T
T	F	F	F	T	F
F	T	T	T	T	T
F	T	F	T	F	T
F	F	T	T	T	T
F	F	F	T	T	T

iii) The argument is
$[(p \vee q) \wedge (q \to r) \wedge \sim r] \Longrightarrow q$
where p: Munna will cook.
q: Munni will practise Karate.
r: Munna studies.
This is **not valid**, as you can see from Row 4 of the following truth table.

p	q	r	$\sim r$	$p \vee q$	$q \rightarrow r$
T	T	T	F	T	T
T	T	F	T	T	F
T	F	T	F	T	T
T	F	F	T	T	T
F	T	T	F	T	T
F	T	F	T	T	F
F	F	T	F	F	T
F	F	F	T	F	T

(conclusion: q; premises: $\sim r$, $p \vee q$, $q \rightarrow r$)

E6) We need to prove $p \Longrightarrow q$, where
p: $x \in \mathbf{R}$ such that $x^2 = 9$, and
q: $x = 3$ or $x = -3$.
Now, $x^2 = 9 \Longrightarrow \sqrt{x^2} = \pm\sqrt{9} \Longrightarrow x = \pm 3$.
Therefore, p is true and $(p \Longrightarrow q)$ is true, allows us to conclude that q is true.

E7) If f is not surjective, then f is not a 1-1 function from X into itself.

E8) We want to prove $\sim q \Longrightarrow \sim p$, where
p: $x \in \mathbf{Z}$ such that x^2 is even,
q: x is even.
Now, we start by assuming that q is false, i.e., x is odd.
Then $x = 2m + 1$ for some $m \in \mathbf{Z}$.
Therefore, $x^2 = 4m^2 + 4m + 1 = 2(2m^2 + 2m) + 1$
Therefore, x^2 is odd, i.e. p is false.
Thus, $\sim q \Longrightarrow \sim p$, and hence, $p \Longrightarrow q$.

E9) i) This is on the lines of Example 5.
ii) Let us assume that $x^3 + 4x = 0$ and $x \neq 0$. Then $x(x^2 + 4) = 0$ and $x \neq 0$. Therefore, $x^2 + 4 = 0$, i.e., $x^2 = -4$. But $x \in \mathbf{R}$ and $x^2 = -4$ is a contradiction. Therefore, our assumption is false. Therefore, the given statement is true.

E10) **Direct proof**: $x^3 + 4x = 0 \Longrightarrow x(x^2 + 4) = 0$
$\Longrightarrow x = 0$ or $x^2 + 4 = 0$
$\Longrightarrow x = 0$, since $x^2 \neq -4 \; \forall \, x \in \mathbf{R}$.
Proof by contrapositive: Suppose $x \neq 0$. Then $x(x^2 + 4) \neq 0$ for any $x \in \mathbf{R}$.
$\therefore x^3 + 4x \neq 0$ for every $x \in \mathbf{R}$.
So we have proved that 'For $x \in \mathbf{R}, x \neq 0 \Longrightarrow x^3 + 4x \neq 0$.'.
That is, 'For $x \in \mathbf{R}, x^3 + 4x = 0 \Longrightarrow x = 0$.'.

E11) Suppose C tells the truth. Therefore, D always tells the truth. Therefore, C always lies, which is a contradiction. Therefore, C can't be a truth-teller, i.e., C is a liar. Therefore, D is a truth-teller.

E12) i) What about $x = 1$?
ii) Take $n = 2, x = 1$ and $y = -1$, for instance.
iii) Here we can find an example f such that f is 1-1 but not onto, or such that f is onto but not 1-1.
Consider $f : \mathbf{N} \rightarrow \mathbf{N} : f(x) = x + 10$. Show that this is 1-1, but not surjective.

E13) i) **Theorem**: The area of every equilateral triangle of side a and perimeter 2a is divisible by 3.
Proof: Since there is no equilateral triangle that satisfies the hypothesis, the proposition is vacuously true.

ii) **Theorem**: If a natural number c is divisible by 5, then the perimeter of the equilateral triangle of side c is 3c.
Proof: Since the conclusion is always true, the proposition is trivially true.

E14) Let p(n) be the given predicate.

Step 1: $p(1): 1 \le 2-1$, which is true.

Step 2: Assume that p(k) is true for some $k \ge 1$, i.e., assume that $1+\frac{1}{4}+\cdots+\frac{1}{k^2} \le 2-\frac{1}{k}$.

Step 3: To show that $p(k+1)$ is true, consider

$$\begin{aligned} 1+\frac{1}{4}+\cdots+\frac{1}{k^2}+\frac{1}{(k+1)^2} &= \left(1+\frac{1}{4}+\cdots+\frac{1}{k^2}\right)+\frac{1}{(k+1)^2} \\ &\le \left(2-\frac{1}{k}\right)+\frac{1}{(k+1)^2}, \text{ by Step 2.} \end{aligned}$$

Now, $\left(2-\frac{1}{k}\right)+\frac{1}{(k+1)^2} \le 2-\frac{1}{(k+1)}$

iff $\frac{1}{(k+1)^2} \le \frac{1}{k}-\frac{1}{k+1}$

iff $k \le k+1$, which is true.

Therefore, $\left(2-\frac{1}{k}\right)+\frac{1}{(k+1)^2} \le 2-\frac{1}{(k+1)}$

Therefore, $p(k+1)$ is true.

Thus, by the PMI, p(n) is true $\forall\, n \in \mathbf{N}$.

E15) $p(2): \frac{1}{\sqrt{1}}+\frac{1}{\sqrt{2}} > \sqrt{2}$, which is true.

Now, assume that p(k) is true for some $k \ge 2$. Then

$$\begin{aligned} \frac{1}{\sqrt{1}}+\frac{1}{\sqrt{2}}+\cdots+\frac{1}{\sqrt{k}}+\frac{1}{\sqrt{k+1}} &> \sqrt{k}+\frac{1}{\sqrt{k+1}}, \text{ since p(k) is true.} \\ &= \frac{\sqrt{k(k+1)}+1}{\sqrt{k+1}} \\ &> \sqrt{k+1}, \text{ since } \sqrt{k+1} > \sqrt{k}. \end{aligned}$$

Hence $p(k+1)$ is true.

$\therefore$ p(n) is true $\forall\, n \ge 2$.

E16) We shall apply the strong form of the PMI here.

Let $p(n): a_n > \frac{3}{2}$.

Step 1: p(3) and p(4) are true.

Step 2: Assume now that for $k \in \mathbf{N}, k \ge 3, p(n)$ is true for every n such that $3 \le n \le k$.

Step 3: We want to show that $p(k+1)$ is true. Now

$$\begin{aligned} a_{k+1} = a_k + a_{k-1} &> \frac{3}{2}+\frac{3}{2}, \text{ by Step 2} \\ &> \frac{3}{2}. \end{aligned}$$

$\therefore p(k+1)$ is true.

Thus, p(n) is true $\forall\, n \ge 3$.

In this case, you will be able to use the weak form conveniently too since $a_k > \frac{3}{2}$ is enough for showing that $p(k+1)$ is true.

Thus, **in this case the weak form is more appropriate** since fewer assumptions give you the same result.

E17) The problem is at the induction step. The first marble may be a different size from the other k marbles. So, we have not shown that $p(k+1)$ is true whenever $p(k)$ is true.

E18) With reference to the statement of the strong form of the PMI, let $S = \{n \in \mathbf{N} | p(n) \text{ is true}\}$.
Then you can show how the form in this problem is the same as the statement of the strong form of the PMI.

E19) Let $p(n) : \sum_{i=1}^{n} \frac{1}{\sqrt{i}} \leq 2\sqrt{n} - 1$.
The weak form suffices here, since the assumption that $p(k)$ is true is enough to prove that $p(k+1)$ is true. We don't need to assume that $p(1), p(2), \cdots, p(k-1)$ are also true to show that $p(k+1)$ is true. Let's prove that $p(n)$ is true $\forall\, n \in \mathbf{N}$.

Now, $p(1) : 1 \leq 2 - 1$, which is true.
Next, assume that $p(k)$ is true for some $k \in \mathbf{N}$.
Then $\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \cdots + \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{k+1}} \leq (2\sqrt{k} - 1) + \frac{1}{\sqrt{k+1}}$, since $p(k)$ is true.
Now $2\sqrt{k} - 1 + \frac{1}{\sqrt{k+1}} \leq 2\sqrt{k+1} - 1$
$\Longleftrightarrow 2(\sqrt{k+1} - \sqrt{k}) \geq \frac{1}{\sqrt{k+1}}$
$\Longleftrightarrow 2(k+1 - \sqrt{k(k+1)} \geq 1$
$\Longleftrightarrow 1 \geq 0$, which is true.
$\therefore p(k+1)$ is true.
$\therefore p(n)$ is true $\forall\, n \in \mathbf{N}$.