Assistant Elementary Educational Officer

Mathematics

Unit-III

Calculus

PG-TRB Materials

Tamil/English/Maths/Commerce/Botany
Chemistry/Physics/History/Economics/Zoology

10% Discount for all PGTRB Materials

Contact

8072230063
UNIT-III-CALCULUS

The Mean Value Theorem

A secant line is a line drawn through two points on a curve. The Mean Value Theorem relates the slope of a secant line to the slope of a tangent line.

The Mean Value Theorem. If $f$ is continuous on $a \leq x \leq b$ and differentiable on $a < x < b$, there is a number $c$ in $a < x < b$ such that

$$
\frac{f(b) - f(a)}{b - a} = f'(c).
$$

The picture below shows why this makes sense. I’ve drawn a secant line through the points $(a, f(a))$ and $(b, f(b))$. The Mean Value Theorem says that somewhere in between $a$ and $b$, there is a point $c$ on the curve where the tangent line has the same slope as the secant line.

Lines with the same slope are parallel. To find a point where the tangent line is parallel to the secant line, take the secant line and “slide” it (without changing its slope) until it’s tangent to the curve.

If you experiment with some curves, you’ll find that it’s always possible to do this (provided that the curve is continuous and differentiable as stipulated in the theorem).

Example. Consider $f(x) = x^3 + 3x^2$ on the interval $-5 \leq x \leq 1$. Since $f$ is a polynomial, $f$ is continuous on $-5 \leq x \leq 1$ and differentiable on $-5 < x < 1$. Moreover,

$$
\frac{f(1) - f(-5)}{1 - (-5)} = \frac{4 - (-50)}{1 - (-5)} = 9.
$$

Hence, there is a number $c$ — maybe more than one — between $-5$ and $1$ such that $f'(c) = 9$. I’ll try to find one.

$$
f'(x) = 3x^2 + 6x, \quad f'(c) = 9c + 6, \quad \text{Set } f'(c) = 9 \quad \text{and solve for } c:
$$

$$
3c^2 + 6c = 9, \quad c^2 + 2c - 3 = 0, \quad (c + 3)(c - 1) = 0, \quad c = -3 \quad \text{or} \quad c = 1.
$$

$c = 1$ is not in the interval $-5 < x < 1$ — it’s an endpoint — but $c = -3$ is. $c = -3$ is a number satisfying the conclusion of the Mean Value Theorem.

Notes:

1. There may be more than one value for $c$ which works.
2. In general, finding a value of \( c \) that works may be difficult. But the theorem only guarantees that such a \( c \) exists, not that you'll be able to find it.

You have to ensure that the hypotheses of the theorem are satisfied before you apply it.

**Example.** Consider \( f(x) = \frac{1}{x^2} \) on the interval \(-1 \leq x \leq 1\). Then

\[
\frac{f(1) - f(-1)}{1 - (-1)} = 0.
\]

However, \( f'(x) = -\frac{2}{x^3} \), and \( f'(c) = -\frac{2}{c^3} = 0 \) has no solution.

This does not contradict the Mean Value Theorem, because \( f \) blows up at \( x = 0 \), which is in the middle of the interval \(-1 \leq x \leq 1\).

**Example.** Calvin Butterball runs a 100 yard dash in 20 seconds. Assume that the function \( s(t) \) which gives his position relative to the starting line is continuous and differentiable. Show that Calvin must have been running at 5 yards per second at some point during his run.

When \( t = 0 \), he’s at the starting line, so \( s = 0 \). When \( t = 20 \), he’s at the finish line, so \( s = 100 \). Applying the Mean Value Theorem to \( s \) for \( 0 \leq t \leq 20 \), I find that there is a point \( c \) between 0 and 20 such that

\[
s'(c) = \frac{100 - 0}{20 - 0} = 5.
\]

That is, Calvin’s velocity at \( t = c \) was 5 yards per second, which is what I wanted to show.

The Mean Value Theorem is often used to prove mathematical results. Here’s an example. You know that the derivative of a constant is zero. The converse is also true.

**Theorem.** If \( f \) is continuous on the closed interval \([a, b]\) and \( f'(x) = 0 \) for all \( x \) in the open interval \((a, b)\), then \( f \) is constant on the closed interval \([a, b]\).

**Proof.** To prove this, let \( d \) be any number such that \( a < d \leq b \). The Mean Value Theorem applies to \( f \) on the interval \([a, d]\), so there is a number \( c \) such that \( a < c < d \) and

\[
\frac{f(d) - f(a)}{d - a} = f'(c).
\]

By assumption, \( f'(c) = 0 \). Therefore,

\[
\frac{f(d) - f(a)}{d - a} = 0, \quad \text{so} \quad f(d) - f(a) = 0, \quad \text{and} \quad f(d) = f(a).
\]

Since \( d \) was an arbitrary number such that \( a < d \leq b \), it follows that \( f(a) = f(x) \) for all \( x \) in \([a, b]\). This means that \( f \) is constant on the interval.

**Example.** I know that \( \frac{d}{dx} x^3 = 3x^2 \). If \( f(x) \) is any other function such that \( \frac{d}{dx} f(x) = 3x^2 \), then

\[
\frac{d}{dx} (f(x) - x^3) = 3x^2 - 3x^2 = 0.
\]
By the theorem, \( f(x) - x^3 = c \), where \( c \) is a constant. Therefore, \( f(x) = x^3 + c \). In other words, the only functions whose derivatives are \( 3x^2 \) are functions like 
\[
x^3, \quad x^3 + 13, \quad x^3 - \sqrt{2}, \quad \text{and so on.}
\]
When I discuss antiderivatives later on, I’ll express this fact by writing 
\[
\int 3x^2 \, dx = x^3 + c.
\]

In the case where the Mean Value Theorem applies and \( f(a) = f(b) \), I get 
\[
\frac{f(b) - f(a)}{b-a} = 0.
\]
The MVT says there is a point \( c \) in \( a < x < b \) such that \( f'(c) = 0 \). This is called Rolle’s Theorem, and a special case may be stated more informally as follows:
- For a “nice” function, there is at least one horizontal tangent between every pair of roots.

In the picture above, there are three critical points between the roots at \( a \) and \( b \).

Example. By the Mean Value Theorem, the function \( f(x) = x(x - 20)(x - 200)(x - 2000) \) has critical points — places where \( f' = 0 \) — between 0 and 20, between 20 and 200, and between 200 and 2000.

Example. Prove that the function \( f(x) = x^5 + 7x^3 + 13x - 18 \) has exactly one root.

Note that a graph is not a proof!

**Step 1.** Since \( f(10) = 107112 \) and \( f(-10) = -107148 \), and since \( f \) is continuous, the Intermediate Value Theorem implies that there is a root between \(-10\) and \(10\). Thus, \( f \) has at least one root.

**Step 2.** Suppose that \( f \) has more than one root. Suppose, in particular, that \( a \) and \( b \) are distinct roots of \( f \).

By Rolle’s theorem, \( f \) must have a horizontal tangent between \( a \) and \( b \).

However, the derivative is \( f'(x) = 5x^4 + 21x^2 + 13 \). Since the powers of \( x \) are even, \( f'(x) > 0 \) for all \( x \): There are no horizontal tangents.

This contradiction shows that there can’t be more than one root.

Step 1 shows that there’s at least one root. Step 2 shows there can’t be more than one. Therefore, there must be exactly one root.
Example. Here is another mathematical result which follows from the Mean Value Theorem; it will be useful in graphing curves.

- If \( f \) is differentiable on \( a < x < b \) and \( f'(x) > 0 \) on \( a < x < b \), then \( f \) increases on \( a < x < b \).

  To say that \( f \) increases means that \( f \) goes up from left to right.

  To see this, take \( p \) and \( q \) between \( a \) and \( b \); say \( a < p < q < b \). I want to show \( f(p) < f(q) \). By the MVT,
  \[
  \frac{f(q) - f(p)}{q - p} = f'(c)
  \]
  for some \( c \) between \( p \) and \( q \).

  But \( f'(c) > 0 \), so
  \[
  \frac{f(q) - f(p)}{q - p} > 0, \quad f(q) - f(p) > 0, \quad f(q) > f(p).
  \]

  This proves that \( f \) increases on the interval.

Example. Here's another example of how the Mean Value Theorem can be used to prove a mathematical result — in this case, an inequality. Apply the MVT to \( f(x) = \tan x \) on the interval \( 0 \leq x \leq k \), where \( k < \frac{\pi}{2} \) to avoid running into the vertical asymptote. Then for some \( c \) between \( 0 \) and \( k \),
\[
\frac{\tan k - \tan 0}{k - 0} = (\sec c)^2.
\]

Now \((\sec c)^2 \geq 1\) and \(\tan 0 = 0\), so
\[
\tan k \geq k.
\]

A picture which illustrates this (not to scale) follows:

The curve is the graph of \( y = \tan x \) and the line is \( y = x \). You can see that the curve appears to lie above the line.

Example. (Using the Mean Value Theorem to estimate a function value) Suppose that \( f \) is a differentiable function,
\[
f(3) = 2, \quad 3 \leq f'(x) \leq 4\quad \text{for all } x.
\]

Prove that \( 8 \leq f(5) \leq 10 \).
Apply the Mean Value Theorem to $f$ on the interval $3 \leq x \leq 5$:

$$\frac{f(5) - f(3)}{5 - 3} = f'(c) \text{ where } 3 < c < 5.$$  

Then since $3 \leq f'(c) \leq 4$, I have

$$3 \leq \frac{f(5) - f(3)}{2} \leq 4$$

$$6 \leq f(5) - 3 \leq 8$$

$$8 \leq f(5) \leq 10$$

SRIMAAN : 8072230063
1. Introduction

There are numerous mathematical generalizations of the natural numbers. On the one hand, there are the integers, the rational numbers, the real numbers, and the complex numbers. On the other hand, there are the cardinal numbers and the ordinal numbers. In the present chapter, we concentrate on cardinal numbers, leaving the other kinds of number to later chapters.

Natural numbers have both a cardinal use and an ordinal use. The former pertains to the quantitative notion of “how many”. The latter pertains to the order of the numbers rather than their sizes. In ordinary language syntax, the cardinal numerals are ‘1’, ‘2’, ‘3’, etc., whereas the ordinal numerals are ‘1st’, ‘2nd’, ‘3rd’, etc.

The cardinal (i.e., chief) use of numbers pertains to measuring how big sets are. This yields the derivative concept of the cardinality of a set A, which is how big A is, or how many elements A has. This suggests introducing a function sign ‘#(_)', defined unofficially as follows.

\[(d) \quad #(A) = \text{the number of elements in } A\]

We need to provide a formal version of the definiens. Whatever we choose, the definition should have the following consequences. [Recall the numerical predicates from the previous chapter.]

\[(t0) \quad #(A) = 0 \iff 0[A] \]
\[(t1) \quad #(A) = 1 \iff 1[A] \]
\[(t2) \quad # (A) = 2 \iff 2[A] \]

etc.

\[
(t0) \quad \forall n \ ( #(A) = n \iff n[A] )
\]

But this is ungrammatical; in first order logic, the same variable cannot appear in both singular term position and predicate position, as ‘n’ does in (t0). We have used the numerals ambiguously, as quantifiers, as predicates, and as singular terms; only numerals-as-singular-terms can be quantificationally generalized in first order logic.

However, we can simulate (t0), by defining the expression ‘n[A]’, not to be the result of applying predicate variable ‘n’ to singular term ‘A’ (which cannot be accomplished in first order logic), but as applying a two-place predicate, written ‘[_][_]’, to singular terms ‘n’ and ‘A’.

In the next section, we discuss how one can define ‘[_][_]’ as a two-place predicate expression in set theory.
Equipollence

We know how to say that \(A\) has two, three, four, etc., elements. How do we say that \(A\) has \(n\) elements, or \(m+n\) elements?

First of all, we know (intuitively at least) that every natural number has the appropriate number of elements – 0 has 0 elements, 1 has 1 element, and in general, \(m\) has \(m\) elements. Accordingly, we know (intuitively) that a set has \(m\) elements iff it has the same size as the number \(m\).

How do we define sameness of size? Usually, to decide whether sets \(A\) and \(B\) have the same size, we count \(A\), and we count \(B\), and we compare the results. In particular, we deem \(A\) and \(B\) to be equal in size if they are both empty, or if they both have 1 element, or if they both have 2 elements, etc.

But this produces a circular definition! We define number of elements in terms of sameness of size, and we define sameness of size in terms of number of elements.

Fortunately, counting is not the only way of comparing the sizes of \(A\) and \(B\). We can also compare them directly. For example, I don’t have to count the fingers on my right hand, and on my left hand, to know that my two hands have equally many fingers. All I have to do is pair up the fingers of the two hands. More dramatically perhaps, I do not have to count the cars on the road, and the drivers on the road, to know there are equally many cars and drivers [let us presume there are no towed cars, teaching cars, fire trucks, etc.]

The reason is that, in each case, the objects in the two sets can be paired one-to-one.

This idea of pairing the elements of the two sets leads to the official definition for equality-of-size.

\[(D1) \quad A \text{ and } B \text{ are equal in size } \Rightarrow \text{ there is a bijection between } A \text{ and } B.\]

Recall that a bijection between \(A\) and \(B\) is a 1-1 function from \(A\) onto \(B\). We use a special symbol for this predicate, and we give the corresponding concept a special name – equipollence, also equipotence. To say that \(A\) and \(B\) are equipollent is to say they have the same size. The word derives from ‘pollence’ which means power. The power of a set is its size. For example, the power set \(\mathcal{P}(A)\) of a set \(A\) is always bigger than the set, as we later prove.

The formal definition goes as follows.

\[(D1^*) \quad A \approx B \equiv \exists f \{ f: A \approx B \}\]

First of all, let us observe that \(\approx\) is an equivalence relation, in the sense that the following are theorems.

\[(T1) \quad A \approx A\]
\[(T2) \quad A \approx B \implies B \approx A\]
\[(T3) \quad A \approx B \land B \approx C \implies A \approx C\]

The proofs are quite simple. \((T1)\) is shown by showing that the identity function on \(A\) is a bijection from \(A\) onto itself. \((T2)\) is shown by showing that if \(f\) maps \(A\) onto \(B\) 1-1, then the inverse function \(f^{-1}\) maps \(B\) onto \(A\) 1-1. \((T3)\) is shown by showing that if \(f\) maps \(A\) onto \(B\) 1-1, and \(g\) maps \(B\) onto \(C\) 1-1, then the composite function \(g \circ f\) maps \(A\) onto \(C\) 1-1.
We are now in a position, finally, to define ‘\(n[A]\)’, at least in the finite case.

\[(d)\quad n[A] =_{\equiv} n \in \omega \land n \approx A\]

In other words, \(A\) has \(n\) elements iff there is a bijection from the number \(n\) onto \(A\). Notice that, \(t_1[t_2]\) is well-formed for any singular terms \(t_1, t_2\), even if \(t_1\) does not refer to a natural number. However, given the definition, \(t_1[t_2]\) cannot be true unless \(t_1\) denotes a natural number.

Notice the following natural consequences of our definitions.

\[
\begin{align*}
(T4) & \quad \forall n (n \in \omega \rightarrow n(n)) \\
(T5) & \quad \forall n (n \in \omega \rightarrow \forall y z (n[y] \land n[z] \rightarrow y \equiv z))
\end{align*}
\]

In other words, every natural number \(n\) has \(n\) elements, and if \(A\) and \(B\) both have \(n\) elements (for any \(n\)), then \(A\) and \(B\) are equal in size.

We are now in a position to go back and define the function sign ‘\(#(\_\)\)’.

\[(d)\quad #(A) = n =_{\equiv} n[A]\]

This has the following consequence, in conjunction with the earlier theorem.

\[
\begin{align*}
(t) & \quad #(A) = #(B) \rightarrow A \approx B
\end{align*}
\]

This says that if the number of \(A\)-elements is the same as the number of \(B\)-elements, then \(A\) and \(B\) are equipollent.

What we would like is the corresponding biconditional.

\[
\begin{align*}
(??) & \quad #(A) = #(B) \iff A \approx B
\end{align*}
\]

But this is not true, at least the way we have defined ‘\(#(\_\)\)’ so far. The problem is that our definition of ‘\(#(\_\)\)’ applies only to finite sets.

Finitude and infinitude are the topics of the next section.

**Finite and Infinite Sets**

At this point, we are in a position, finally, to define the term ‘finite’ and the complementary term ‘infinite’.

\[
\begin{align*}
(D2) & \quad \text{fin}[A]) =_{\equiv} \exists n (n \in \omega \land n[A])
\end{align*}
\]

In other words, \(A\) is finite iff it has \(n\) elements for some natural number \(n\), which is to say that it is equipollent to some natural number.

To say that a set is infinite is simply to say that it is not finite:

\[
\begin{align*}
(D3) & \quad \text{inf}[A] \iff \sim \text{fin}[A]
\end{align*}
\]
Clearly, there are finite sets; for example, every natural number is finite, since every natural number is equipollent to a natural number – itself. Also, the singleton of any natural number is finite, having exactly one element.

The following are expected (if not easily proved) theorems about finite sets.

(T6) \( \forall x \forall y \left( \text{fin}[x] \land y \subseteq x \rightarrow \text{fin}[y] \right) \)

(Any subset of a finite set is finite.)

(T7) \( \forall x \left( \text{fin}[x] \rightarrow \text{fin}[\varnothing(x)] \right) \)

(The power set of any finite set is finite.)

(T8) \( \forall x \forall y \left( \text{fin}[x] \land \text{fin}[y] \rightarrow \text{fin}[x \cup y] \right) \)

(The union of two finite sets is finite.)

(T9) \( \forall x \forall y \left( \text{fin}[x] \land \text{fin}[y] \rightarrow \text{fin}[x \times y] \right) \)

(The Cartesian product of two finite sets is finite.)

(T10) \( \forall x \left( \text{fin}[x] \land \forall y (y \in x \rightarrow \text{fin}[y]) \rightarrow \text{fin}[\bigcup x] \right) \)

(The union of a finite collection of finite sets is finite.)

The existence of finite sets is obvious. The existence of infinite sets is less obvious. However, the Axiom of Infinity (appropriately so called) does yield the existence of at least one infinite set. In particular, we have the following theorem.

(T11) \( \omega \) is infinite

Proof: we proceed by induction, proving that no natural number is equipollent to \( \omega \). Base case: \( \omega \) is not equipollent to 0; this is because only 0 is equipollent to 0, and \( \omega \) is not 0. Inductive hypothesis: \( \omega \) is not equipollent to \( m \), to show: \( \omega \) is not equipollent to \( m+1 \). Suppose otherwise, to show a contradiction. Then there is a function that maps \( \omega \) 1-1 onto \( m+1 \), call it \( f \). There is a unique element \( x \) in \( \omega \) such that \( f(x) = m \), call it \( b \). Claim: \( f \) maps \( \omega \setminus \{b\} \) 1-1 onto \( \omega \setminus \{b\} \). Define function \( g \) from \( \omega \) to \( \omega \setminus \{b\} \) as follows. For all \( x \leq b \), \( g(x) = x \); for all \( x > b \), \( f(x) = x^* \). Claim: \( g \) maps \( \omega \) 1-1 onto \( \omega \setminus \{b\} \). Claim: \( f \circ g \) maps \( \omega \) 1-1 onto \( m \), which means that \( \omega \) is equipollent to \( m \), which contradicts the inductive hypothesis.

**Denumerable Sets**

A set is finite iff it is equipollent to some element of \( \omega \); otherwise, it is infinite. On the other hand, a set is said to be denumerable (or denumerably infinite) if it is equipollent to the set \( \omega \) of all natural numbers. Formally speaking,

(D4) \( \text{den}[A] \iff A \approx \omega \)

Notice the following immediate theorem.

(T12) \( \forall X \left( \text{den}[X] \rightarrow \text{infin}[X] \right) \)

I.e., every denumerable set is infinite. Reason: if \( X \) is denumerable, it is equipollent to \( \omega \); if \( X \) is finite, it is equipollent to some natural number. So if \( X \) is both denumerable and finite, then \( \omega \) is equipollent to some natural number, and hence is finite, which we have already proved is not true.
Examples of denumerable sets abound. For example, the set
\[ \{ x : x \in \omega & x > m \}, \]
where \( m \) is any natural number, is denumerable. To show a set is denumerable, it is sufficient to produce a bijection from \( \omega \) onto \( A \). In the above case, the function is defined so that \( f(x) = m + x \).

Another example, the set of multiples of \( m \),
\[ \{ x : \exists y(y \in \omega & x = my) \}, \]
where \( m \) is any natural number, is also denumerable. The bijection is defined so that \( f(x) = mx \).

Similarly, the set of powers of \( m \),
\[ \{ x : \exists y(y \in \omega & x = m^y) \}, \]
where \( m \) is any natural number, is also denumerable. The bijection is defined so that \( f(x) = m^x \).

The sets in question are all subsets of \( \omega \); indeed, they are all infinite subsets. The upshot of this is that, although it would seem that there are twice as many numbers as even numbers (multiples of 2), there are in fact equally many. The two sets can be paired up one-to-one.

SRIMAAN AEEO MATERIALS: MATHS/ ENGLISH AVAILABLE - CONTACT: 8072230063

The sets in question are all infinite subsets of \( \omega \). One can actually prove that every infinite subset of \( \omega \) is denumerable.

\[(T13) \forall X (X \subseteq \omega & \text{infin}[X] \rightarrow \text{den}[X])\]

Intuitive proof sketch: suppose \( A \subseteq \omega \), and \( \text{infin}[A] \). Then \( A \) is well-ordered by \( \leq \). That means every subset \( X \) of \( A \) has a first element, denoted \( \text{first}(X) \). Define \( f(x) \) as follows (using strong induction).

\[ f(n) = \text{first}(A \cup \{ \{m \} : n < \infty \} ) \]

Intuition: \( f(0) \) is first element of \( A \), \( f(1) \) is the next element of \( A \) after \( f(0) \), \( f(2) \) is the next element of \( A \) after \( f(1) \), etc.

Claim: \( f \) is a 1-1 function from \( \omega \) onto \( A \). Hence, \( A \) is denumerable.

So far, we have concentrated on subsets of \( \omega \). We can also look at selected supersets of \( \omega \). For example, \( \omega^+, \omega^{++}, \omega^{+++} \), etc. [which are the first few transfinite \( \text{ordinals} \)]. Recall the definition of successor: \( A^+ = A \cup \{A \} \). So \( \omega^+ = \omega \cup \{\omega \} \). These sets are also denumerable. For example, define \( f: \omega^+ \rightarrow \omega \) as follows: \( f(x) = 0 \), if \( x = \omega \), otherwise \( f(x) = x^+ \).

The Cartesian product \( \omega \times \omega \) of \( \omega \) with itself seems like a candidate for a set that is considerably bigger than \( \omega \), but it too is denumerable. The function (\( \omega \)-sequence) may be pictured as follows, using a technique that traces to Cantor.
An alternative picture might be visually helpful.

\[
\begin{array}{cccccc}
(0,0) & (0,1) & (0,2) & (0,3) & (0,4) & (0,5) \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
(1,0) & (1,1) & (1,2) & (1,3) & (1,4) & (1,5) \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
(2,0) & (2,1) & (2,2) & (2,3) & (2,4) & (2,5) \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
(3,0) & (3,1) & (3,2) & (3,3) & (3,4) & (3,5) \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
(4,0) & (4,1) & (4,2) & (4,3) & (4,4) & (4,5) \\
\end{array}
\]

In this diagram, we simply count the first row, the second row, etc. It is evident that every ordered pair is on some row, and each row can be counted.

The same technique can be used to prove that the rational numbers are denumerable, which means that there are just as many natural numbers as there are rational numbers!

The above technique can also be used to prove the following.

(T14) If \( C \) is denumerable, and every element of \( C \) is denumerable, then \( \bigcup C \) is denumerable. I.e., the union of any denumerable collection of denumerable sets is itself denumerable.

(T15) If \( C \) is a finite family of denumerable sets, then \( \prod C \) is denumerable. I.e., the Cartesian product of any finite family of denumerable sets is denumerable.

From what has transpired so far, one would naturally draw the conclusion that all infinite sets are denumerable, which implies in particular that all infinite sets have the same size. This is quite definitely not true, as we see in the next section.
Uncountable Sets

So far, we have discussed finite sets and denumerable sets, which are a species of infinite sets. It was suggested at the end of the last section that there are infinite sets that are not denumerable. Such sets are called uncountable (also, uncountably infinite, also nondenumerable). The traditional usage is actually somewhat confusing, so let us chart the relationships.

(D2) A set is finite iff it is equipollent to some natural number.
(D3) A set is infinite iff it is not finite.
(D4) A set is denumerable iff it is equipollent to ω.
(D5) A set is countable iff it is finite or denumerable.
(D6) A set is uncountable (non-denumerable) iff it is infinite but not denumerable.

Notice in particular that there is overlap of the concepts; in particular, Infinite + Countable = Denumerable. Also notice the mildly illogical fact that ‘nondenumerable’ does not mean ‘not denumerable’. A set that is not denumerable (versus nondenumerable) may be either uncountable or finite, whereas a nondenumerable set cannot be finite.

SRIMAAN AEEO MATERIALS:MATHS/ENGLISH AVAILABLE-CONTACT:8072230063
Are there any uncountable sets? The affirmative answer was first proved by Cantor, who proved two examples, namely

(c1) the set of real numbers (rationals ∪ irrationals) is uncountable.
(c2) the power set of ω is uncountable.

We have not discussed, nor even defined, the real numbers so far, so we must postpone the proof of (c1) until later. However, we already have the machinery to prove (c2).

(T16) Ψ(ω) is uncountable.

Proof: Let f be any function from ω into Ψ(ω). We wish to prove that f is not onto. For each number n, there is a unique subset A = A(n). Call a number n normal if it is not an element of f(n). Consider the subset of normal numbers, call it N = {n : n /∈ f(n)}. Claim: N /∈ ran(f), and hence f is not onto Ψ(ω). Proof: Suppose N /∈ ran(f). Then N = f(m), where m ∈ ω, i.e., f(m) = {n : n /∈ f(n)}. It follows that m ∈ f(m) ↔ m /∈ f(m), which is a contradiction. This proves that Ψ(ω) is not denumerable; to prove that Ψ(ω) is uncountable (nondenumerable; recall distinction from above), we must further prove that ω is infinite. This is left as an exercise.

Infrapollence

The theorem that Ψ(ω) is uncountable is actually a special case of a more general theorem, which says that no set is equal in size to its power set, that every set is smaller in size than its power set. In order to state this theorem, we need to develop some additional notation and terminology.

Just as the identity of two sets, A and B, can be decomposed into two inclusions, A≤B and B≤A, the equipollence (equality of size) of two sets can be similarly decomposed. The relevant notion is what I propose to call infrapollence, which means “of lower (or equal) power”. It is officially defined as follows.

(D7) A ≤ B =⇒ ∃f [ f : A → B(1-1) ]
The definiendum is read: \( A \) is smaller than or equal in size to \( B \). The definition says this means that there is a 1-1 function from \( A \) into \( B \). Recall that such a function is sometimes called an *injection*. Thus, \( A \) is smaller than or equal in size to \( B \) iff there is an injection from \( A \) into \( B \).

The definition seems intuitively correct. It basically says that for every \( A \)-element, \( a \), there is a unique \( B \)-element, \( f(a) \), and maybe there are some \( B \)-elements left over; in other words, there are enough \( B \)'s to "go around".

The following theorems are obvious consequences of the definition.

- (T17) \( A \leq A \)
- (T18) \( A \leq B \) \& \( B \leq C \) \( \rightarrow \) \( A \leq C \)
- (T19) \( A \subseteq B \rightarrow A \leq B \)
- (T20) \( A \approx B \rightarrow A \leq B \)

In (T17), the identity function injects \( A \) into itself. In (T18), if \( f \) injects \( A \) into \( B \), and \( g \) injects \( C \) into \( D \), then the composite function \( g \circ f \) injects \( A \) into \( C \). In (T19), the identity map on \( A \) injects \( A \) into any superset of \( A \). In (T20), if \( f \) maps \( A \) 1-1 onto \( B \), then \( f \) automatically maps \( A \) 1-1 onto \( B \).

(T17) and (T18) say that \( \leq \) is a quasi-order relation, so there is an associated strict order relation, called strict infrapollence, defined as follows.

- (D8) \( A < B \equiv A \leq B \) \& \( \sim[B \leq A] \)

As with every quasi-order relation, there is an affiliated equivalence relation, defined in the usual way.

- (T21) \( A < B \) \& \( B < C \) \( \rightarrow \) \( A < C \)
- (T22) \( A < B \rightarrow \sim[B < A] \)
- (T23) \( \sim[A < A] \)

In other words, strict infrapollence is transitive, asymmetric, and irreflexive.

### The Schroeder-Bernstein Theorem

As we saw in the previous section, infrapollence is a quasi-order relation, which is to say that the following obtains.

- (T17) \( A \leq A \)
- (T18) \( A \leq B \) \& \( B \leq C \) \( \rightarrow \) \( A \leq C \)

As with every quasi-order relation, there is an affiliated equivalence relation, defined in the usual way.

- (d) \( A \approx B \equiv A \leq B \) \& \( B \leq A \)

Recall that a quasi-order is a partial order iff the affiliated equivalence relation is the identity relation. Needless to say, \( \leq \) is not a partial order relation; the following is *not* true.

- (false) \( A \leq B \) \& \( B \leq A \) \( \rightarrow \) \( A = B \)
For example, consider \( \{ a \}, \{ b \} \), where \( a \neq b \). Then \( \{ a \} \not\subseteq \{ b \} \) and \( \{ b \} \not\subseteq \{ a \} \), but \( \{ a \} \neq \{ b \} \). On a more esoteric level, the set \( \omega^+ \) of positive natural numbers can be injected into \( \omega \), by the identity function, and \( \omega \) can be injected into \( \omega^+ \), by the successor function. But \( \omega \neq \omega^+ \).

But what is the equivalence relation \( \simeq \)? The answer is exactly what one might expect – equipollence. In other words, we have the following theorem, which is usually called the Schroeder-Bernstein Theorem.

\[
(T24) \quad A \leq B \& B \leq A \implies A \simeq B
\]

It is also called the Cantor-Schroeder-Bernstein Theorem; Cantor conjectured it, and Schroeder and Bernstein independently proved it in the 1890’s.

In order to prove The Schroeder-Bernstein Theorem, (T24), we first prove another important theorem, the Fixed Point Theorem (or rather, a special case).

**Fixed Point Theorem** (special case):

\[
(T25) \quad \text{Let } A \text{ be any set, and let } f \text{ be any function from } \mathcal{P}(A) \text{ into } \mathcal{P}(A) \text{ satisfying the following condition, called monotonicty.}
\]

\[
(m) \quad \forall X, Y [X \subseteq Y \implies f(X) \subseteq f(Y)]
\]

Then there is a "fixed point", i.e., a set \( Z \) such that \( f(Z) = Z \).

**Proof**: Suppose \( f \) is such a function on \( \mathcal{P}(A) \). Consider the collection \( \{ X : X \subseteq f(X) \} \), call it \( C \). Consider the union, \( \cup C \). We wish to show that \( \cup C = f(\cup C) \).

Claim: \( \forall X [X \in C \implies X \subseteq f(\cup C)] \). For suppose \( D \in C \); then \( D \subseteq f(D) \), also \( D \subseteq \cup C \), so by (m), \( f(D) \subseteq f(\cup C) \). So \( D \subseteq f(\cup C) \). So \( \cup C \subseteq f(\cup C) \). We are half done. Now \( \cup C \subseteq f(\cup C) \), by (m), \( f(\cup C) \subseteq f(\cup C) \). But the latter means that \( f(\cup C) \subseteq C \), from which we obtain the converse inclusion, \( f(\cup C) \subseteq \cup C \).

Thus, \( f(\cup C) = \cup C \).

Next, we state a few relevant lemmas, whose proofs are left as exercises.

**SRIMAAN AEEO MATERIALS: MATHS/ ENGLISH AVAILABLE-CONTACT:8072230063**

**Lemma 1:**

Let \( A \) and \( B \) be sets. Let \( A_1 \) and \( A_2 \) be disjoint subsets of \( A \) such that \( A_1 \cup A_2 = A \). Similarly, let \( B_1 \) and \( B_2 \) be disjoint subsets of \( B \) such that \( B_1 \cup B_2 = B \). Then:

\[
A_1 \simeq B_1 \& \ A_2 \simeq B_2 \implies A \simeq B
\]

**Lemma 2:**

Let \( A \) be any set, let \( X, Y \) be any subsets of \( A \), and let \( f \) be any function on \( A \). Then:

\[
(1) \quad X \subseteq Y \implies A-Y \subseteq A-X
\]

\[
(2) \quad X \subseteq Y \implies f^+(X) \subseteq f^+(Y)
\]

**Lemma 3** (a corollary to Lemma 2):

Let \( A, B \) be any sets, let \( f \) be any function from \( A \) to \( B \), let \( g \) be any function from \( B \) into \( A \), let \( X, Y \) be any subsets of \( A \). Then:

SRIMAAN AEEO MATERIALS: MATHS/ ENGLISH AVAILABLE-CONTACT:8072230063

\[ X \subseteq Y \rightarrow g^-(B-f^-(A-X)) \subseteq g^-(B-f^-(A-Y)) \]

**Lemma 4:**

if \( f \) is a 1-1 function, and \( A \subseteq \text{dom}(f) \), then \( A \approx f^+(A) \)

Having proven the fixed point theorem, and the relevant lemmas, we now prove

**The Schroeder-Bernstein Theorem.**

(T24) \( A \leq B \ & \ B \leq A \rightarrow A \approx B \)

**Proof:** Suppose \( A \leq B \ & \ B \leq A \), to show \( A \approx B \). Then there is an injection, \( f \), from \( A \) into \( B \). And there is an injection, \( g \), from \( B \) into \( A \). We use Lemma 1 to show \( A \approx B \); in particular, we divide \( A \) into disjoint parts \( A_1, A_2 \) and \( B \) into disjoint parts \( B_1, B_2 \), and we show the respective parts are equivalent. We divide \( A \) into parts by appealing to the Fixed Point Theorem. Specifically, let \( K \) be the inductively defined function that maps each subset \( X \) of \( A \) to the set \( g^{-1}(B-f^{-1}(A-X)) \) which satisfies the requirements of \((\text{FTP})\), so there is fixed point, namely a set, call it \( K \), such that \( g^+(B-f^+(A-K)) = K \). Since \( A-K \subseteq \text{dom}(f), A-K \approx f^+(A-K), \) by Lemma 4. Also, \( B-f^+(A-K) \subseteq \text{dom}(g), \) and \( K = g^+(B-f^+(A-K)) \), so \( K \approx B-f^+(A-K), \) by Lemma 4. Furthermore, the sets in question divide \( A \) and \( B \) respectively, so applying Lemma 1, we have \( A \approx B \).

**Hyper-Uncountable Sets; The Continuum Hypothesis**

We proved earlier that the power set of \( \omega \) is uncountable. In the present section, we use the same technique, in conjunction with the Schroeder-Bernstein theorem, to show that every set is strictly smaller than its power set. This is sometimes called Cantor’s Theorem.

(T25) \( A < \mathcal{P}(A) \)

**Proof:** Given the definition, we have to show \( A \leq \mathcal{P}(A) \), and we have to show \( \neg(\mathcal{P}(A) \leq A) \). The former is shown by noting that the function \( f \), defined so that \( f(x) = \{ x \} \) is an injection from \( A \) into \( \mathcal{P}(A) \). The latter is shown by showing \( \neg(\mathcal{P}(A) \approx A) \), and appealing to the Schroeder-Bernstein Theorem. So suppose that \( \mathcal{P}(A) \approx A \). Then there is a 1-1 function from \( A \) onto \( \mathcal{P}(A) \). We prove that, in fact, there is no function from \( A \) onto \( \mathcal{P}(A) \). Once again, we construct the set of normal elements of \( A \). \( N = \{ x: x \not\in f(x) \} \). Claim: \( N \notin \text{ran}(f) \). For suppose otherwise. Then \( N = f(a) \), where \( a \in A \). So \( f(a) = \{ x: x \not\in f(x) \} \), from which it follows that \( a \in f(a) \leftrightarrow a \notin f(a) \), a contradiction.

Since every set is smaller than its power set, we have the following infinite chain.

\[ A < \mathcal{P}(A) < \mathcal{P}(\mathcal{P}(A)) < \mathcal{P}(\mathcal{P}(\mathcal{P}(A))) < \text{etc.} \]

[What is worse, there is a set bigger than all of these – an infinitely-hyper-uncountable set! However, we do not prove that just yet.]
Before moving on, it is useful to consider the following questions.

(q1) \( \exists X [\omega < X \land X < \wp(\omega)] \) ??
(q2) \( \exists X [\wp(\omega) < X \land X < \wp(\omega)] \) ??
(q3) \( \exists X [\wp(\omega) < X \land X < \wp(\omega)] \) ??

etc.

The negative answer to (q1) is known as the **Continuum Hypothesis** (CH), and the negative answer to the entire series is known as the **Generalized Continuum Hypothesis** (GCH).

To prove the Continuum Hypothesis was the first problem posed by Hilbert in his famous list of outstanding problems in mathematics (1900). In the 30’s, Gödel proved CH could be consistently added to set theory. In the 60’s, Paul Cohen proved that CH cannot be derived from the usual axioms of set theory.

### The Cardinal Numbers

Finally, we return to our original problem – to identify (or postulate) the cardinal numbers in such a way that we can deduce the following theorem, which might be called the **Fundamental Principle of Cardinal Numbers**.

\[
(t) \quad \#(A) = \#(B) \iff A \approx B
\]

Perhaps it is useful to consider the solution proposed by Russell and Whitehead in *Principia Mathematica*, where \(\#(\_\_\_)\) is defined as follows.

\[
(RW1) \quad \#(A) = _\sigma \{ X : X \approx A \}
\]

\[
(RW2) \quad \text{card}[A] = _\sigma \exists X \ [ A = \#(X) ]
\]

In other words, the cardinality of a set \(A\) is simply the collection of all sets equipollent to \(A\), and cardinal numbers are equivalence classes, where the equivalence relation is equipollence. For example, the cardinality of the empty set \(\emptyset\), which is Russell’s definition of the cardinal number zero, and the cardinality of the singleton \(\{\emptyset\}\) is the set of all singletons, which is Russell’s definition of the cardinal number one. Infinite cardinal numbers are just equivalence classes whose members are infinitely large sets.

It is easy to see that, provided \(\#(A)\) is always a proper set (i.e., non empty), then \((t)\) is an immediate consequence of the Russell-Whitehead definition. It is just like factoring out an equivalence relation on a set. The equivalence classes are identical iff the objects are equivalent.

Unfortunately, as we have already seen, there is a serious difficulty with the Russell-Whitehead definition of the natural numbers, let alone their definition of general cardinal numbers. Within our framework, the equivalence classes in question do not exist (except the first one); they are too big.

As seen in an earlier chapter, in the case of the natural numbers (the finite cardinal numbers), we circumvent this problem by a special selection method. We select a representative singleton, and decree it to be the number 1, we select a representative doubleton, and decree it to be the number 2, etc.
It appears that we have to do the same thing for the transfinite cardinals. For each “cardinality class”, we have to select a representative set of that cardinality. This includes picking a representative of each of the following.

- an \(X\) such that \(X \approx \omega\)  
- an \(X\) such that \(X \approx \wp(\omega)\)  
- an \(X\) such that \(X \approx \wp(\wp(\omega))\)  
etc.

More generally, for each set \(A\), we must select a representative set equipollent to \(A\).

The official definition we adopt goes as follows.

\[(\text{D9}) \quad \#(A) =_\sigma \text{ the first ordinal number } \alpha \text{ such that } \alpha \approx A\]

[Ordinals are discussed in detail in the next chapter.] Observe that this definition does not work properly unless we assume the Axiom of Choice. The ordinals are well-ordered by \(\leq\), where \(a \leq b\) iff \(a \in b\) or \(a = b\). So if there is any ordinal equipollent to a set \(A\), there is a smallest such ordinal. The question, then, is whether every set equipollent to \(A\) is the same cardinal number. Here, we appeal to the Well-Ordering Theorem, which is a well-known equivalent of the Axiom of Choice. In particular, we well-order the set \(A\) (using AC), and then we appeal to a further theorem that every well-ordered set is isomorphic (and hence equipollent) to at least one ordinal.

For those who don’t accept the Axiom of Choice, other routes are available. One approach is to treat the expression ‘\(\#(A)\)’ as syncategorimatic, as a pseudo-term, by which I mean that it only appears in formulas of the form ‘\(\#(A) = \#(B)\)’, and these formulas are defined as follows.

\[\text{(d) } \quad \#(A) = \#(B) =_\sigma A \approx B\]

A more common alternative approach is to introduce ‘\(\#(\omega)\)’ as a further primitive item of set theory, and postulate the following further axiom.

\[\text{(a1) } \forall x \forall y [ \#(x) = \#(y) \iff x \approx y ]\]

This has the desirable consequence that the finite cardinals are coextensive with the natural numbers. It does not say what the infinite cardinal numbers are, however. For example, the infinite cardinal numbers may be finite sets, for all we know. This is ok, so long as they are not natural numbers. The following might be true (for all we know).

\[
\begin{align*}
\#(\omega) &= \{\{\emptyset\}\} \\
\#(\wp(\omega)) &= \{\{\{\emptyset\}\}\} \\
\#(\wp(\wp(\omega))) &= \{\{\{\{\emptyset\}\}\}\}
\end{align*}
\]

If we want the infinite cardinal numbers to be like the finite cardinal numbers, in the sense that they all have the “appropriate” size, then the following is a further plausible axiom.
This just says that the cardinal number of a set is the same size as the set. This rules out the above identifications.

Still another approach is to treat cardinal numbers as points (first elements), rather than as sets, in which case no cardinal number is identified with any set. This requires redoing the theory of natural numbers.

So long as we have the Fundamental Principle of Cardinal Numbers, either as a theorem deduced from the definition of cardinal number, or as an axiom within a theory that regards ‘#(_’) as a primitive expression, we can deduce many properties of cardinal numbers.

For example, we can define addition, multiplication, and less-than-or-equal, as follows.

\[ \forall X[#(X) \approx X] \]

For example, one can prove that addition and multiplication are both associative and commutative, that multiplication distributes over addition, that \( \leq \) is reflexive, transitive, and anti-symmetric. Curiously, one cannot prove that \( \leq \) is connected, that is, one cannot prove

\[ (C) \quad m \leq n \lor m \leq n, \]

without appealing to the Axiom of Choice; (C) is in fact equivalent to the Axiom of Choice. So, if one believes that infinite cardinal numbers are (or should be) arranged in a linear order just like the finite cardinal numbers, then one must further postulate (C), but this is tantamount to postulating the Axiom of Choice, in which case one might as well define the cardinal numbers to be first ordinal numbers, as originally suggested.

Mathematicians routinely employ the Axiom of Choice in their reasoning, because of its great power, although many use it apologetically. As shown by Cohen, AC is not a logical consequence of Zermelo-Fraenkel set theory, and hence is not an essential feature of the iterative conception of sets.

The latter is the idea that sets are generated from first elements by repeated (iterated) application of a few set-forming processes, including – pairs, unions, power sets, and replacement. These seem a necessary part of our conception of sets. By contrast, the epistemological status of the other axioms, including the Axiom of Choice, and the Continuum Hypothesis, remains unclear. This suggests the existence of alternative set theories on a par with alternative (non-Euclidean) geometries.
SRIMAAN COACHING CENTRE: PG-TRB-CHEMISTRY
STUDY MATERIALS AVAILABLE:

- PG TRB: TAMIL MATERIAL (QUESTION BANK)
- PG TRB: ENGLISH MATERIAL (QUESTION BANK)
- PG TRB: MATHEMATICS MATERIAL (QUESTION BANK) (E/M)
- PG TRB: PHYSICS MATERIAL (E/M)
- PG TRB: CHEMISTRY MATERIAL (QUESTION BANK) (E/M)
- PG TRB: COMMERCE (QUESTION BANK) (Tamil & English Medium)
- PG TRB: ECONOMICS (QUESTION BANK) (T/M)
- PG TRB: HISTORY (QUESTION BANK) (T/M)
- PG TRB: ZOOLOGY (QUESTION BANK) (E/M)
- PG TRB: BOTANY (QUESTION BANK) (T/M)

SRIMAAN COACHING CENTRE - PG- TRB MATERIALS: MATHS/ENGLISH/TAMIL/COMMERCE/CHEMISTRY/PHYSICS/BOTANY/ZOOLOGY/HISTORY/ECONOMICS
STUDY MATERIALS AVAILABLE - 8072230063
GOVT. POLYTECHNIC TRB MATERIALS:

> MATHEMATICS
> ENGLISH with Question Bank
> COMPUTER SCIENCE/IT with Question Bank
> ECE MATERIAL With Question Bank
> CHEMISTRY
> PHYSICS

GROUP 2A: GENERAL ENGLISH

AEEO EXAM: MATHEMATICS/ENGLISH

10% Discount for all materials. Materials are sending through COURIER

CONTACT: 80722 30063

THANK YOU