MATH 147
Personal Notes

Marcus Chan

Taught by Brian Forrest

UW Math '25

Chapter 1: A Short Introduction to Mathematical Logic and Proof

BASIC NOTIONS OF MATHEMATICAL LOGIC AND TRUTH TABLES

DEFINITIONS

1. A "statement" is a mathematical sentence that can be determined to be true or false.

2. A "conjecture" is a statement which is widely regarded to be true, but no proof for it exists yet. (there is evidence or strong speculation for its validity).

3. An "axiom" is a statement that is assumed to be true, with no prerequisite proof required.

STATEMENT "OPERATIONS"

NEGATION / NOT ($\neg$)

The negation of a statement p, or $\neg p$, is simply the opposite of p; i.e. NOT p.

p	$\neg p$
T	F
F	T

* the law of the excluded middle: p & $\neg p$ are never both T or F.

IMPLYING STATEMENTS ($\Rightarrow$)

If statement p implies statement q, or $p \Rightarrow q$, then the truth of p means q is also true.

"anticedent" "consequence"

p	q	$p \Rightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

$p \Rightarrow q$ can only be F if p is T but q is F!

* note: if we look at the truth table, we see that anything false can imply anything.

AND & OR ($\wedge$ & $\vee$)

The "AND" & "OR" operators are represented by the symbols "$\wedge$" & "$\vee$" respectively.

* "OR" is referred to as "inclusive OR".

EXCLUSIVE OR (XOR)

If statement $r = p$ XOR q, then r is true if p is true or q is true, but not both.

↳ p XOR q = $(p \vee q) \wedge \neg(p \wedge q)$.

p	q	(AND) $p \wedge q$	(OR) $p \vee q$	(NAND) $\neg(p \wedge q)$	(XOR) $(p \vee q) \wedge \neg(p \wedge q)$
T	T	T	T	F	F
T	F	F	T	T	T
F	T	F	T	T	T
F	F	F	F	T	F

EQUIVALENCE

Statements p and q are logically equivalent if the truth of one implies the truth of the other: ie. p holds if and only if q holds, or $(p \Rightarrow q) \wedge (q \Rightarrow p)$ is true. (ie $p \Leftrightarrow q$).

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

* this does NOT imply statement p is the same as statement q. Rather, it implies that p & q have the same "output" (T/F) under all inputs.

CONTRAPOSITIVE

The contrapositive of a statement $p \Rightarrow q$ is $\neg q \Rightarrow \neg p$; ie to get it, you switch the positions of p & q and negate both

p	q	(initial statement) $p \Rightarrow q$	$\neg q$	$\neg p$	(contrapositive) $\neg q \Rightarrow \neg p$
T	T	T	F	F	T
T	F	F	T	F	F
F	T	T	F	T	T
F	F	T	T	T	T

identical!

★ From the truth table, we can see that the contrapositive of $p \Rightarrow q$ is logically equivalent to it.

→ the above can also be proven by the fact that $(p \Rightarrow q) \Leftrightarrow (\neg q \Rightarrow \neg p)$ is a tautology: the statement is true regardless of the inputs (the truth values assigned to p & q).

VARIABLES & QUANTIFIERS

1. There are mathematical sentences whose truth values are dependent on additional parameters.

2. Take, for instance, the statement "$x > 0$".
   - The statement may be true or false depending on the value of x, as it can vary; hence we call x a "variable".
   - We can use the functional notation $p(x) : x > 0$ to represent this statement.
   - So $p(-4)$ is false, but $p(3)$ is true, for example.

QUANTIFIERS

FOR ALL ($\forall$)

1. The symbol "$\forall$" can be used to denote "for all".

* "for all" is known as the <u>universal quantifier</u>.

For example, to say "for all x, $p(x)$ is true", we would write

$\forall x : p(x)$.

* this is a statement that can be true or false.

2. If we want to prove the above statement to be <u>false</u>, we need only to show one counterexample.

eg the statement "for all natural numbers n, n factors as a product of primes" is false as it does not hold for $n=1$.

THERE EXISTS ($\exists$) & SUCH THAT ($\ni$)

1. If we say that "there exists an x such that $p(x)$ is true", it implies that we can find at least 1 value of x_0 such that $p(x_0)$ is true.

* "there exists" is known as the "existential quantifier".

↳ in mathematical notation, the above statement would be written as

$$\underbrace{\exists x}_{\text{there exists an } x} \ \underbrace{\ni}_{\text{such that}} \ \underbrace{p(x)}_{p(x) \text{ is true.}} .$$

2. To prove that "$\exists x \ni p(x)$" is <u>true</u>, we just need to find an example of a value for x that makes $p(x)$ true.

eg to prove "there exists an integer greater than 5" we just need to show that the integer 6 is greater than 5.

3. To prove that "$\exists x \ni p(x)$" is <u>false</u>, we need to show the statement is false for all possible values of x.

eg to disprove "there exists a natural number less than 1" we just show every natural number is greater than it.

COMPOUND SENTENCES

ORDER MATTERS

The order of quantifiers in a statement is very important; any slight alteration to it can change the statement's meaning drastically.

For instance, consider the statements

① $\forall x : \exists z \ni x \leq z$, and

for all x, there exists a z such that $x \leq z$.

→ This statement is <u>true</u>. why? → for any value of x, observe that the statement is true simply if $z = x$.

② $\exists x \ni \forall z : x \leq z$,

there exists a value for x such that $x \leq z$ for all values of z.

→ This statement is <u>false</u>. why? → it implies a least real number exists, which we know is not true.

where x, z are real numbers.

NEGATION

1. There are often many ways to write the negative of a statement with quantifiers.

2. Take, for instance, the statement "for every x, $p(x)$ is true".

↳ to prove it is false, we must find that $p(x)$ is false for a value of x: ie

$$\underbrace{\neg(\forall x : p(x))}_{\text{the negative of "for all } x,\ p(x) \text{ is true"}} \ \underbrace{\Leftrightarrow}_{\text{is logically equivalent to}} \ \underbrace{\exists x \ni \neg p(x)}_{\text{there exists an } x \text{ such that } p(x) \text{ is false.}} .$$

3. Another example: take the statement "there exists an x such that $p(x)$ is true".

↳ to prove it is false, we need to show $p(x)$ is false for all values of x; ie

$$\underbrace{\neg(\exists x \ni p(x))}_{\text{the negative of "there exists an } x \text{ such that } p(x) \text{ is true"}} \ \underbrace{\Leftrightarrow}_{\text{is logically equivalent to}} \ \underbrace{\forall x : \neg p(x)}_{\text{for all } x,\ p(x) \text{ is false.}}$$

RULES OF INFERENCE & THE FOUNDATIONS OF PROOF

DEDUCTIVE REASONING

Deductive reasoning is a strategy for proofing: we start with a hypothesis/something we know to be true, and apply "rules of inference" to reach our desired conclusion.

MODUS PONENS

Modus Ponens tells us that if p is true, and p implies q, then q must also be true; ie

$$p \wedge (p \Rightarrow q) \Rightarrow q$$

if p is true & p implies q it implies that q is true.

"if Fido is a dog, and dogs are animals, then Fido is an animal."

MODUS TOLLENS

Modus Tollens tells us that if we know that $p \Rightarrow q$ is true, and q is false, then p must also be false;

ie

$$(p \Rightarrow q) \wedge \neg q \Rightarrow \neg p$$

if p is true implies q is true & q is false it implies that p is false.

"if a spoon is a dog, it must be an animal. However a spoon is not an animal and so if must not be a dog."

HYPOTHETICAL SYLLOGISM

Hypothetical syllogism is basically an "associativity law" for implication: if $p \Rightarrow q$ is true, & $q \Rightarrow s$ is true, then $p \Rightarrow s$ must be true too;

ie

$$(p \Rightarrow q) \wedge (q \Rightarrow s) \Rightarrow (p \Rightarrow s)$$

if p implies q is true & if q implies s is true it implies if p is true, s is also true.

"if Fido is a dog, it must be an animal. If Fido is an animal, it must eat. Since Fido is a dog, it must eat."

DISJUNCTIVE SYLLOGISM

Disjunctive syllogism says that if we know either p or q is true, and we can show p is false, then q must be true;

ie

$$(p \vee q) \wedge \neg p \Rightarrow q$$

if p or q is true & p is false then it implies that q is true.

"Fido is either a cat or a dog. If Fido isn't a cat, it must be a dog".

CONSTRUCTIVE DILEMMA

Constructive dilemma states that if $p \Rightarrow q$ & $r \Rightarrow s$ is true, and we know either p or r is true, then either q or s must be true;

ie

$$[[(p \Rightarrow q) \wedge (r \Rightarrow s)] \wedge (p \vee r)] \Rightarrow (q \vee s).$$

if $p \Rightarrow q$ & $r \Rightarrow s$ are true & p or r is true it implies that q or s is true.

"if Fido is a dog, he is an animal. if Fido is a potato, he is a vegetable. hence, if we know Fido is a dog, he must be an animal."

DESTRUCTIVE DILEMMA

Destructive dilemma states that if we know $p \Rightarrow q$ & $r \Rightarrow s$ are true, and we know that either q or s is false, then we can conclude that either p or r must also be false;

ie

$$[[(p \Rightarrow q) \wedge (r \Rightarrow s)] \wedge [\neg q \wedge \neg s]] \Rightarrow \neg p \wedge \neg r.$$

if $p \Rightarrow q$ & $r \Rightarrow s$ are true & q or s is false it implies that either p or r is false.

"if Fido is a dog, it must be an animal. if Fido is a potato, it must be a vegetable. Since Fido is not an animal, he cannot be a dog."

SIMPLIFICATION, ADDITION AND CONJUNCTION

1. Simplification tells us that if we know p & q are both true, then p is true; ie

$$p \wedge q \Rightarrow p.$$

2. Addition tells us that if we know p is true, then either p or q is true for any q; ie

$$p \Rightarrow (p \vee q).$$

3. Conjunction tells us that if we know p & q are true separately. then p and q must be true together.

INDUCTIVE REASONING

In inductive reasoning, we instead begin with some specific observations and then try to draw out a more general conclusion.

There is another type of reasoning called *inductive reasoning*. In inductive reasoning, we begin with some specific observations and then try to draw a more general conclusion. For example, if we knew that the first few terms of an infinite sequence were $\{2, 4, 6, 8, 10, \ldots\}$, then we might guess from the pattern of these five terms that this was the sequence of all even natural numbers. From this we could speculate that the next term in the sequence would be 12. Unlike, most instances of deductive reasoning that we will see in this course, inductive reasoning most often does **not** result in a proof. Indeed, it is possible that if we were to be told a few more terms in the sequence above we might find that we have $\{2, 4, 6, 8, 10, 0, 2, 4, 6, 8, 10, \cdots\}$ where the general formula for the n-th term is $a_n = 2n \mod 12$. We see that our inductive conclusion was wrong. This does not make inductive reasoning useless. In fact, inductive reasoning is the foundation for much of science, particularly experimental science. Even in mathematics, inductive reasoning often leads us to an understanding of what is actually going on. It helps us to formulate conjectures, mathematical statements that we believe to be true, and for which we might later find proofs. It is also a key element in problem solving. Moreover, in the next chapter we will see how to employ an important formal technique of proofs that is based on inductive reasoning called *Proof by Induction*.

Chapter 2:
Sets, Relations and Functions

NOTATION

NUMBER SETS

1. $\mathbb{N}$ denotes the set of natural numbers: $\{1, 2, 3 \dots\}$
2. $\mathbb{Z}$ denotes the set of integers: $\{\dots -2, -1, 0, 1, 2 \dots\}$
3. $\mathbb{Q}$ denotes the set of rational numbers: $\{\frac{a}{b} : a \in \mathbb{Z},\ b \in \mathbb{N}\}$
4. $\mathbb{R}$ denotes the set of real numbers.

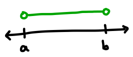
INTERVALS

An interval is any set S which obeys the following property: if $\forall x, y \in S$, and $x \leq z \leq y$, then $z \in S$.

OPEN, HALF-OPEN & CLOSED INTERVALS

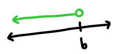
1. An open interval is an interval where both endpoints are excluded; ie in the form

(a, b) , or $\{x : a < x < b\}$.



Open intervals can also be unbounded – meaning on one or both sides – they can also take the form

$(-\infty, b)$ or $\{x : x < b\}$;

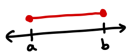


(a, ∞) or $\{x : x > a\}$;

& $(-\infty, \infty)$ or $\{x : x \in \mathbb{R}\}$.

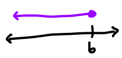
2. A closed interval is an interval where both endpoints are included, ie in the form

$[a, b]$ or $\{x : a \leq x \leq b\}$.



Similarly, closed intervals can also be unbounded on one side, which take the form

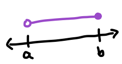
$(-\infty, b]$ or $\{x : x < b\}$



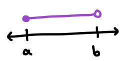
& $[a, \infty)$ or $\{x : x > a\}$.

3. Half-open intervals are intervals where one endpoint is included and one is excluded; ie in the form

$(a, b]$ or $\{x : a < x \leq b\}$



& $[a, b)$ or $\{x : a \leq x < b\}$



DEGENERATE INTERVALS

An interval is degenerate if it either:
1. is the empty set; or
2. consists of only one element (a singleton set), ie $\{a\}$ where $a \in \mathbb{R}$.

The singleton set

☆ Observe $\{a\}$ is an interval.

Why? → $\{a\} \Leftrightarrow [a, a]$.

Going back to our definition of an interval, and letting $x = a$, $y = a$; we see the only possible value z can take is a, which satisfies the condition.

The empty set

☆ Observe the empty set, $\emptyset$ or $\{\}$, is also an interval.

Why? → proof by contradiction.

Suppose it isn't. That implies that $\exists x, y, z \in \emptyset \ni [\neg(x \leq z \leq y)]$.

This is impossible because no such x, y exist in $\emptyset$! Contradiction : therefore $\emptyset$ must be an interval.

SETS & THEIR PRODUCTS

SUBSETS ($\subset$ / $\subseteq$)

1. A set A is a subset of another set B if every element of A is also contained in B.

eg if A is $\{1,2,3\}$, & B is $\{1,2,3,4,5\}$, A is a subset of B.

→ To specify that A is a subset of B, we can use the notation

$A \subset B$ or $A \subseteq B$.

2. The above notation indicates A might equal B.

→ If we want to write A as a proper subset of B, then we can:

① either further specify $A \neq B$; or

② use the notation $A \subsetneq B$.

POWER SET ($\mathbb{P}$)

1. Given a set X, we define the power set of X, or $\mathbb{P}(X)$, to be the set

$$\mathbb{P}(X) = \{ A \mid A \subseteq X \} ;$$

ie $\mathbb{P}(X)$ consists of all subsets of X, including $\emptyset$ & X itself.

eg if $X = \{1,2\}$,
$\mathbb{P}(X) = \{ \emptyset, \{1\}, \{2\}, \{1,2\} \}$.

CARDINALITY

1. The cardinality of a set X is the number of elements in X, and denoted by $|X|$.

eg if $X = \{1,2,3\}$, $|X| = 3$.

PRODUCT

1. The product of 2 sets X & Y, denoted as $X \times Y$, is defined to be

$$X \times Y = \{ (x,y) \mid x \in X, y \in Y \}.$$

*(x,y) is known as the "ordered pair" of x & y.

2. For n sets $X_1, X_2 .. X_n$, we define the product to be

$$X_1 X_2 X_3 .. X_n = \prod_{i=1}^{n} X_i = \{ (x_1, x_2 .. x_n) \mid x_i \in X_i \}.$$

*$(x_1, x_2 .. x_n)$ is known as an n-tuple; x_i is known as the ith coordinate.

3. The cardinality of the product of n sets $X_1, X_2 .. X_n$ is the product of the cardinalities of the products of the n sets; ie

$$|X_1 X_2 .. X_n| = |X_1||X_2| .. |X_n| .$$

SET DIFFERENCE & COMPLEMENT ($\setminus$, ^{c})

1. The set difference of a set B minus another set A, written as "$B \setminus A$", is defined to be

$$B \setminus A = \{ x \in B \mid x \notin A \} ;$$

ie $B \setminus A$ consists of all the elements in B that are not in A.

2. Let $\mathcal{E}$ denote the universal set. Then, $U \setminus A$ is known as the complement of A and is written as

A^c or A'.

UNION ($\cup$)

1. The union of two sets A & B ($A \cup B$) where $A, B \subseteq \mathcal{E}$, is defined to be

$$A \cup B = \{ x \mid x \in A \vee x \in B \}.$$

2. If we want to represent the union of multiple sets, we can use the notation

$$\bigcup_{1 \leq \alpha \leq i} A_\alpha = A_1 \cup A_2 \cup A_3 \ldots \cup A_i ,$$

where $A_1, A_2 .. A_i \subseteq \mathcal{E}$.

INTERSECTION ($\cap$)

1. The intersection of two sets A & B ($A \cap B$), where $A, B \subseteq \mathcal{E}$, is defined to be

$$A \cap B = \{ x \mid x \in A \wedge x \in B \}.$$

2. Similarly, when talking about the intersection of multiple sets, we can use the notation

$$\bigcap_{1 \leq \alpha \leq i} A_\alpha = A_1 \cap A_2 \cap .. \cap A_i ,$$

where $A_1, A_2 .. A_i \subseteq \mathcal{E}$.

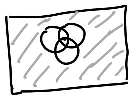
DE MORGAN'S LAWS

① $$\left[\bigcup_{1 \leq \alpha \leq i} A_\alpha \right]^c = \bigcap_{1 \leq \alpha \leq i} A_\alpha^c .$$

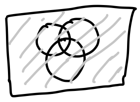
Proof.
Observe LHS is the set of all the elements in $\mathcal{E}$, but not in A_α (for $1 \leq \alpha \leq i$).
For the RHS, it will also be the same, since no element in any A_α can be in the RHS.
Hence LHS = RHS.

② $$\left[\bigcap_{1 \leq \alpha \leq i} A_\alpha \right]^c = \bigcup_{1 \leq \alpha \leq i} A_\alpha^c .$$

Proof.
Observe LHS is the set of everything except the common intersection of A_α, where $1 \leq \alpha \leq i$.
Then, the RHS will consist of everything too, bar the common intersection since no A_α^c can consist of it.
Hence LHS = RHS.



The grey area represents the set represented by dM's laws in a simplified case (3 sets only).



CHOICE FUNCTIONS

1. We can use choice functions to define a "product" for an infinite collection of sets, which we can denote as "$\{X_\alpha\}_{\alpha \in I}$", where I is some infinite set.

2. We begin by noting that each n-tuple, $(x_1, x_2 .. x_n) \in \prod_{i=1}^{n} X_i$, directly corresponds to some function $f_{(x_1, x_2 .. x_n)}$, where

$$f_{(x_1, x_2, .. x_n)} : \{1,2,3 .. n\} \rightarrow \bigcup_{i=1}^{n} X_i .$$

this function directly corresponds w/ $(x_1, x_2 .. x_n)$

the set of all inputs are the natural numbers from 1 to n.

the set of all outputs is the union of all the "X_i" sets.

and

$$f_{(x_1, x_2, x_3 .. x_n)}(i) = x_i .$$

if the input of the function is i, the output of the function is x_i.

3. Hence, since there exists an 1-1 correspondence between the elements of

$\prod_{i=1}^{n} X_i$ & $\{ f : \{1,2 .. n\} \rightarrow \bigcup_{i=1}^{n} X_i \mid f(i) \in X_i \}$

(the product of all the sets X_i) (the set of all possible choice functions)

we can write

$$\prod_{i=1}^{n} X_i \cong \{ f : \{1,2, .. n\} \rightarrow \bigcup_{i=1}^{n} X_i \mid f(i) \in X_i \}.$$

4. Since f allows us to choose 1 element from each of our sets, we call f a choice function.

5. We can then use choice functions to denote a definition for infinite sets:

Given a collection $\{X_\alpha\}_{\alpha \in I}$ of sets, define

$$\prod_{\alpha \in I} X_\alpha = \{ f : I \rightarrow \bigcup_{\alpha \in I} X_\alpha \mid f(\alpha) \in X_\alpha \}.$$

*1 If $X_\alpha = X \ \forall \alpha \in I$, X_α is written as X^I.

*2 A function $f \in \prod_{\alpha \in I} X_\alpha$ is called a choice function on $\{X_\alpha\}_{\alpha \in I}$.

RELATIONS & FUNCTIONS

RELATIONS

1. A relation R on 2 sets X & Y is any set $R \subseteq X \cdot Y$.

2. If $x \in X$ & $y \in Y$ satisfy $(x,y) \in R$, we say x is R-related to y, described using the notation "xRy".

3. Subsequently, the domain and range of R, denoted by $dom(R)$ & $ran(R)$ respectively, are defined by

$$dom(R) = \{x \in X \mid \exists y \in Y \ni (x,y) \in R\}$$

& $$ran(R) = \{y \in Y \mid \exists x \in X \ni (x,y) \in R\}.$$

4. Lastly, the set Y is known as the codomain of R, or $codom(R)$.

5. Note: if $X = Y = \mathbb{R}$, we say R is a relation on $\mathbb{R}$.

FUNCTIONS

1. A function f on X with values in Y, denoted by $f: X \to Y$, is a relation $f \subseteq X \cdot Y$ such that $\forall x \in X$, there exists exactly one $y \in Y$ such that $(x,y) \in f$.

2. In this case, we denote the value of y by $f(x)$ and write $y = f(x)$.

3. By using our earlier definitions for relations and applying them to f, observe that
   i) $dom(f) = X$;
   ii) $codom(f) = Y$; &
   iii) $range(f) = \{y = f(x) \mid x \in X\}$.

Additionally, define the graph of f by
   iv) $graph(f) = \{(x,y) = (x, f(x)) \mid x \in X\}$.

*notice $graph(f) \subseteq X \cdot Y$.

4. Alternative definition of a function: If $R \subseteq X \cdot Y$ is a relation, then R determines a function f if and only if $\exists[(x,y_1) \in R \wedge (x,y_2) \in R]$, it implies $y_1 = y_2$.

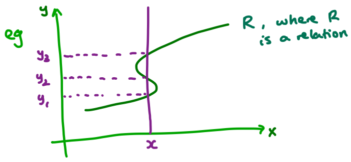
5. In this case, we let $dom(f) = \{x \in X \mid (x,y) \in R$ for some $y \in Y\}$, and write $y = f(x)$ if $x \in dom(f)$ & $(x,y) \in R$.

THE VERTICAL LINE TEST

1. The premise of the vertical line test is as follows: if, for any graph in the Cartesian plane, any vertical line can be drawn that intersects the graph more than once, the graph cannot be a function.

2. Why? Because this implies two or more y values are mapped to the same value of $x \in dom(R)$.

eg



The test implies $\{(x,y_1), (x,y_2), (x,y_3)\} \subseteq graph(R)$; hence R cannot be a function.

COMPOSITION OF FUNCTIONS

1. Let X, Y and Z be non-empty sets. Then, let the functions f & g be such that $f: X \to Y$ & $g: Y \to Z$. The composition of g by f is the function

$$h(x) = g \circ f(x) = g(f(x)).$$

2. However, it is imperative that $ran(f) \subseteq dom(g)$; otherwise, $g(f(x))$ cannot exist.

TRANSFORMATION OF FUNCTIONS

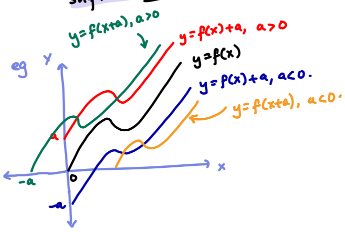
TRANSLATIONS

Let $f: \mathbb{R} \to \mathbb{R}$. Then:

① $f(x) + a$ corresponds to the graph of f shifted upwards by a units.

② $f(x+a)$ corresponds to the graph of f shifted to the left by a units.

eg



*notice how if $a<0$, it corresponds to a shift of the graph of f downwards (for ①) and to the right (for ②) by $|a|$ units.

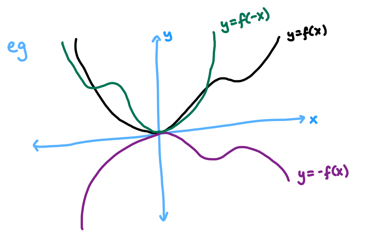
REFLECTIONS

Let $f: \mathbb{R} \to \mathbb{R}$. Then:

① $-f(x)$ corresponds to the graph of f being flipped through the x-axis.

② $f(-x)$ corresponds to the graph of f being flipped through the y-axis.

eg

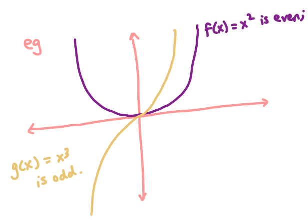


EVEN & ODD FUNCTIONS

1. A function f is even if $\forall x: f(x) = f(-x)$; ie its graph is symmetric across the y-axis.

2. A function g is odd if $\forall x: g(x) = -g(-x)$; ie its graph is symmetric about the origin.

eg



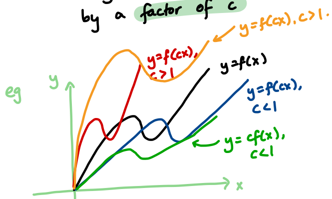
SCALING

Let $f: \mathbb{R} \to \mathbb{R}$. Then:

① $f(cx)$ corresponds to the graph of f being "compressed" in the x direction by a factor of c.

② $cf(x)$ corresponds to the graph of f being "stretched" in the y direction by a factor of c.

eg



*similarly, notice if $a<0$, ① & ② correspond to a stretch in the x direction and a compression in the y direction by a factor of $|a|$ respectively.

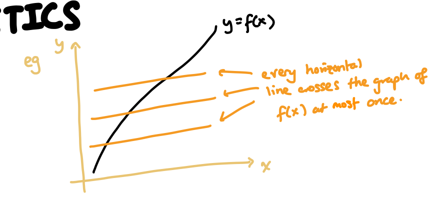
FUNCTION CHARACTERISTICS

1-1 FUNCTIONS

1. A function $f: X \to Y$ is 1-1 (one-to-one) if f assigns different x's to different y's; ie if $\forall x_1, x_2 \in X$ & $x_1 \neq x_2$, then $f(x_1) \neq f(x_2)$.

2. Visually, a function f is 1-1 if every horizontal line crosses the graph of f at most once.

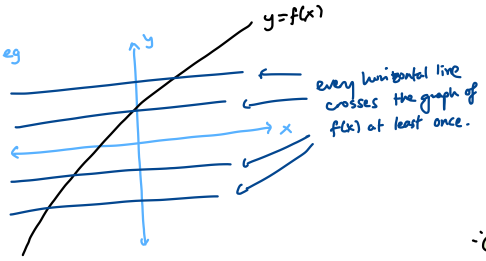
3. We say f is "1-1 on an interval I" if $\forall x_1, x_2 \in I$ & $x_1 \neq x_2$, then $f(x_1) \neq f(x_2)$.



ONTO FUNCTIONS

1. A function $f: X \to Y$ is onto if $\operatorname{ran}(f) = Y$; ie $\forall y_0 \in Y$, $\exists x_0 \in X$ such that $f(x_0) = y_0$.

2. Visually, a function f is onto if every horizontal line crosses the graph of f at least once.



INVERSE FUNCTIONS

1. If a 1-1 & onto function $f: X \to Y$ exists, the inverse function of f, $g: Y \to X$, can be defined by

$g(y) = x$ if and only if $f(x) = y$.

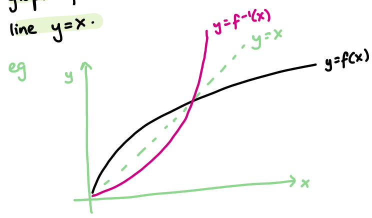
*note: $g(y)$ is often denoted by $f^{-1}(y)$.

2. If $f^{-1}(y)$ exists, we say f is invertible on X.

3. f is invertible on some interval I if there exists a function $g(y): f(I) \to I$ by $x = g(y)$ if and only if $x \in I$ & $y = g(x)$.

GRAPHING INVERSE FUNCTIONS

If f is invertible, where $f: X \to Y$, then the graph of f^{-1} is the reflection of $f(x)$ through the line $y = x$.



INCREASING & DECREASING FUNCTIONS

1. Let f be a function defined on some interval. Also, let $x_1, x_2 \in I$ be such that $x_1 < x_2$.

Then:

① f is increasing if $\forall x_1, x_2 \in I: f(x_1) < f(x_2)$;

② f is non-decreasing if $\forall x_1, x_2 \in I: f(x_1) \leq f(x_2)$;

③ f is decreasing if $\forall x_1, x_2 \in I: f(x_1) > f(x_2)$;

④ f is non-increasing if $\forall x_1, x_2 \in I: f(x_1) \geq f(x_2)$;

⑤ f is monotonic on I if either ①, ②, ③ or ④ is true;

⑥ f is strictly monotonic on I if either ① or ③ is true.

2. If f is strictly monotonic, then it is 1-1 on I.

PULLBACK

1. Let $f: X \to Y$. Then the pullback of f is defined to be the function $f^{-1}: \mathcal{P}(Y) \to \mathcal{P}(X)$ by

$f^{-1}(B) = \{x \in X \mid f(x) \in B\}$ for each $B \in \mathcal{P}(Y)$.

2. The pullback of a subset B tells us all the elements in X mapped into B by f.

MATHEMATICAL INDUCTION

PRINCIPLE OF MATHEMATICAL INDUCTION

The Principle of Mathematical Induction states that if, for any set S,

1) $1 \in S$, &
2) For each $k \in \mathbb{N}$, if $k \in S$, then $k+1 \in S$,

then $S = \mathbb{N}$.

PROOF BY MATHEMATICAL INDUCTION

The goal of proving by mathematical induction is to prove a statement $p(n)$ is true $\forall n \in \mathbb{N}$, or similar.

There are 3 general steps to the proof:

① Identify $p(n)$;

② Show $p(1)$ is true;

③ Show that if $p(k)$ is true, $p(k+1)$ also is.

PRINCIPLE OF STRONG INDUCTION

The Principle of Strong Induction states that, for any set $S \subseteq \mathbb{N}$, if

① $1 \in S$, &

② For each $k \in \mathbb{N}$, if $1, 2, \ldots, k \in S$, then $k+1 \in S$,

then $S = \mathbb{N}$.

WELL-ORDERING PRINCIPLE

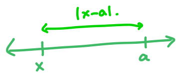
The Well-Ordering Principle states that every set $S \subseteq \mathbb{N}$, where $S \neq \emptyset$, has a least element.

*note: these 3 principles are equivalent — they imply each other.

Chapter 3: Real Numbers

ABSOLUTE VALUES

1. Given any $x, a \in \mathbb{R}$, the absolute value of $x-a$, denoted as $|x-a|$, equals the distance between them on the number line.

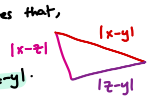


2. Formally, $|x| = \begin{cases} x, & x \geq 0 \\ -x, & x < 0 \end{cases}$.

INEQUALITIES INVOLVING ABSOLUTE VALUES

TRIANGLE INEQUALITY

1. The triangle inequality states that, given any $x, y, z \in \mathbb{R}$,
$|x-y| \leq |x-z| + |z-y|$.



Proof.
We assume $x \leq y$ without any loss in generality.

case 1 : $z < x$. $\Rightarrow |z-y| > |x-y|$.
$\therefore |z-y| + |x-z| > |x-y|$ as $|x-z| > 0$. QED

case 2 : $x \leq z \leq y$. $\Rightarrow |x-y| = |x-z| + |z-y|$. QED.

case 3 : $z > y$. $\Rightarrow |x-y| < |x-z|$.
$\therefore |x-y| < |x-z| + |z-y|$ as $|z-y| > 0$. QED.

2. Another variant of the Triangle Inequality is that, if $x, y \in \mathbb{R}$,
$|x+y| \leq |x| + |y|$.

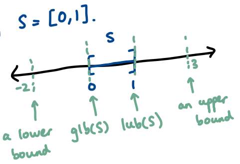
BOUNDS

1. For any set S :

① α is an upper bound of S,
if $x \leq \alpha \;\; \forall x \in S$;
→ if α is the smallest such upper bound
(that is, if $x \leq \gamma \;\; \forall x \in S$, then $\alpha \leq \gamma$)
it is called the least upper bound of S
and is denoted as $\mathrm{lub}(S)$ or $\sup(S)$.
→ if S contains $\mathrm{lub}(S)$, α is also called the maximum of S and is denoted by $\max(S)$.

② β is a lower bound of S,
if $x \geq \beta \;\; \forall x \in S$.
→ if β is the greatest such lower bound
(that is, if $x \geq \gamma \;\; \forall x \in S$, then $\beta \leq \gamma$)
it is called the greatest upper bound of S
and is denoted as $\mathrm{glb}(S)$ or $\inf(S)$.
→ if S contains $\mathrm{glb}(S)$, β is also called the minimum of S and is denoted by $\min(S)$.

eg $S = [0,1]$.



LEAST UPPER BOUND PROPERTY

LUBP states that every set $S \subset \mathbb{R}$, which is bounded above, has a least upper bound.

GREATEST LOWER BOUND PROPERTY

GLBP states that every set $S \subset \mathbb{R}$, which is bounded below, has a greatest lower bound.

IS ∅ BOUNDED?

We can show $\emptyset$ is bounded above or below.

→ Observe that any $\alpha \in \mathbb{R}$ can be an upper bound & lower bound for $\emptyset$.

Why? Proof by contradiction.
Suppose α is not an upper bound for $\emptyset$.
Then it implies that
$\exists x \in \emptyset : x > \alpha$.
But $\emptyset$ is empty and has no elements;
$\therefore$ a contradiction.

1ST ARCHIMEDEAN PROPERTY

1. The 1AP states that $\mathbb{N}$ has no upper bound in $\mathbb{R}$; ie it is not bounded above.

Proof by contradiction.
- Suppose $\exists M \in \mathbb{R} \ni x < M \;\; \forall x \in \mathbb{N}$.

→ By LUBP, $\mathbb{N}$ must have some least upper bound α.
Since $\alpha = \mathrm{lub}(\mathbb{N})$, $\alpha - 1 \neq \mathrm{lub}(\mathbb{N})$.
$\Rightarrow \exists n \in \mathbb{N} \ni (\alpha - 1 < n < \alpha.)$

→ By the principle of mathematical induction,
$n \in \mathbb{N} \Rightarrow n+1 \in \mathbb{N}$.
However this implies that $\alpha < n+1$
and so an element of $\mathbb{N}$ is greater than α;
a blatant contradiction.

2ND ARCHIMEDEAN PROPERTY

2. The 2AP states that
$\forall \varepsilon > 0 : \exists n \in \mathbb{N} \ni 0 < \frac{1}{n} < \varepsilon$.

Proof. $\forall \varepsilon : \exists \varepsilon \ni 0 < \varepsilon < n, \; n \in \mathbb{N}$.
since a violation of this implies
$\varepsilon = \mathrm{lub}(\mathbb{N})$, which cannot be true by 1AP.

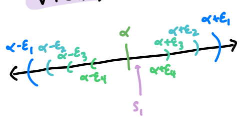
If $0 < \varepsilon$, then $0 < \frac{1}{\varepsilon}$.

If $\frac{1}{n} < \varepsilon$, then $n > \frac{1}{\varepsilon}$.

$0 < \frac{1}{\varepsilon} \wedge n > \frac{1}{\varepsilon} \Rightarrow 0 < \frac{1}{\varepsilon} < n$. ∎

DENSITY IN ℝ

A set $S \subseteq \mathbb{R}$ is dense in $\mathbb{R}$ if
$\forall \alpha \in \mathbb{R}, \varepsilon > 0 : \exists s \in S \ni (\alpha - \varepsilon < s < \alpha + \varepsilon)$



ie for any $\alpha \in \mathbb{R}$, no matter how small ε is, an element of S will always be in the interval.

DENSITY OF ℚ & ℝ\ℚ in ℝ

$\forall a, b \in \mathbb{R} \ni (a < b) : \exists (r \in \mathbb{Q}, s \notin \mathbb{Q}) \ni r, s \in (a,b)$;
ie the rationals and irrationals are dense in $\mathbb{R}$.

Chapter 4: Sequences and Convergence

SEQUENCES

1. A sequence is simply an ordered list.
   eg 1, 2, 4, 7 .. or 3, 5, 13, 18 .. etc.

2. To denote an (infinite) sequence, we use the notation
   $\{a_1, a_2, \ldots, a_n, \ldots\}$ or $\{a_n\}_{n=1}^{\infty}$ or $\{a_n\}$.
   * n = the "index" of a_n.
   * some authors use round instead of curly brackets.

SUBSEQUENCES & TAILS

1. Let $\{a_n\}$ be a sequence, and $\{n_1, n_2 \ldots n_k, \ldots\}$ be a sequence where $n_i \in \mathbb{N}$ & $n_i < n_{i+1}$ $\forall\, i \in \mathbb{N}$.
   A subsequence of $\{a_n\}$ is a sequence of the form
   $\{a_{(n_k)}\} = \{a_{n_1}, a_{n_2}, a_{n_3} \ldots\}$

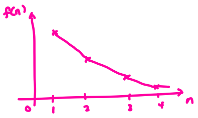
2. Given a sequence $\{a_n\}$, $k \in \mathbb{N}$, the subsequence
   $\{a_n\}_{n=k}^{\infty} = \{a_k, a_{k+1} \ldots\}$
   is called the tail of $\{a_n\}$ with cutoff k.

METHODS OF DEFINING SEQUENCES

LISTING

1. We can use a list to specify a sequence.
   eg if $a_n = \frac{1}{n}$,
   the sequence is $\{1, \frac{1}{2}, \frac{1}{3}, \ldots\}$

2. We can also use a function to identify a sequence, which also allows us to plot its graph.
   eg $f(n) = \frac{1}{n}$:



VIA RECURSION

We can use recursion to define a sequence — ie, the formula for a_n utilises previous terms (a_{n-1}, a_{n-2} etc).

eg¹ $a_{n+1} = \frac{1}{a_n + 1}$, $a_1 = 1$.

eg² $a_1 = a_2 = 1$, $a_n = a_{n-1} + a_{n-2}$
(Fibonacci)

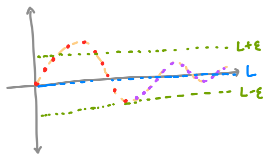
eg³ $a_{n+1} = \frac{1}{2}\left(a_n + \frac{\alpha}{a_n}\right)$
(Heron's formula for finding $\sqrt{\alpha}$)

LIMITS OF SEQUENCES

1. Formally, L is the limit of $\{a_n\}$ as $n \to \infty$, if $\forall \varepsilon > 0 : \exists N \in \mathbb{N}$ $\ni$ if $n \geq N$, $|a_n - L| < \varepsilon$.

2. $\exists L$, we say the sequence is convergent, and write
$$\lim_{n\to\infty} a_n = L.$$

3. An alternative $\text{def}^{\underline{n}}$ for L: L is a limit of $\{a_n\}$ if $\forall \varepsilon > 0$, the interval $(L-\varepsilon, L+\varepsilon)$ contains a tail of $\{a_n\}$.



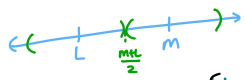
* notice that all the purple points are within the interval.

UNIQUENESS OF LIMITS

If $\{a_n\}$ is a convergent sequence, then it has one and only limit L.

Proof by contradiction.

Suppose $\{a_n\}$ has 2 limits L & M, w/ $L \neq M$.



Then, both intervals must contain a tail of the sequence.

$\Rightarrow$ at least one element of $\{a_n\}$ must be in both intervals simultaneously;

however since they are disjoint, this is impossible. ☒

DIVERGENCE TO $\pm\infty$

1. We say a sequence $\{a_n\}$ diverges to $+\infty$ if $\forall M > 0$, $\exists N \ni$ if $n \geq N$, then $a_n > M$, and write
$$\lim_{n\to\infty} a_n = +\infty.$$

2. Similarly, we say a sequence $\{a_n\}$ diverges to $-\infty$ if $\forall M < 0$, $\exists N \ni$ if $n \geq N$, then $a_n < M$, and write
$$\lim_{n\to\infty} a_n = -\infty.$$

LIMIT ARITHMETIC

Let $\{a_n\}$ & $\{b_n\}$ be sequences, and $\lim_{n\to\infty} a_n = L$ & $\lim_{n\to\infty} b_n = M$. Then:

1. $\forall c \in \mathbb{R} : (a_n = c \ \forall n) \Rightarrow c = L.$
2. $\forall c \in \mathbb{R} : \lim_{n\to\infty} c a_n = cL.$
3. $\lim_{n\to\infty} (a_n + b_n) = L + M.$
4. $\lim_{n\to\infty} (a_n b_n) = LM.$
5. $\lim_{n\to\infty} \left(\frac{a_n}{b_n}\right) = \frac{L}{M}$ if $M \neq 0$.

   * if $M=0$, the limit may exist, or it might not.
6. $[(\forall n : a \geq 0) \wedge (\alpha > 0)] \Rightarrow \lim_{n\to\infty} a_n^{\alpha} = L^{\alpha}.$
7. $\lim_{n\to\infty} a_{n+k} = L \quad \forall k \in \mathbb{N}.$

THE LIMIT OF $\frac{p(n)}{q(n)}$, WHERE $p(n), q(n)$ ARE POLYNOMIALS

Let $a_n = \dfrac{b_0 + b_1 n + b_2 n^2 \cdots + b_j n^j}{c_0 + c_1 n \cdots + c_k n^k}$.

Then
$$\lim_{n\to\infty} a_n = \begin{cases} \frac{b_j}{c_k}, & j = k \\ 0, & j < k \\ +\infty, & j > k \wedge \frac{b_j}{c_k} > 0 \\ -\infty, & j > k \wedge \frac{b_j}{c_k} < 0. \end{cases}$$

LIMIT POINTS

1. An $\alpha \in \mathbb{R}$ is called a limit point of a sequence $\{a_n\}$ if there is a subsequence $\{a_{n_k}\}$ of $\{a_n\}$ such that
$$\lim_{k\to\infty} a_{n_k} = \alpha$$

2. We denote the set of limit points of $\{a_n\}$ by $\text{LIM}(\{a_n\})$.

eg^1 $\text{LIM}(\{-1, 1, -1 \cdots\}) = \{-1, 1\}$

eg^2 $\text{LIM}(\{1, \frac{1}{2}, 2, \frac{1}{4}, 3, \frac{1}{8} \cdots\}) = \{0\}.$

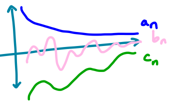
*note: in eg^2, a limit point exists BUT the series is divergent!

$\Rightarrow$ having an unique LP is not enough to show the sequence converges.

LIMIT THEOREMS

SQUEEZE THEOREM

The ST states that if $a_n \leq b_n \leq c_n$, & $\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = L$, then $\{b_n\}$ converges and $\lim_{n\to\infty} b_n = L$.



Proof.

$(\lim_{n\to\infty} a_n = L = \lim_{n\to\infty} c_n)$

$\Rightarrow \exists N \in \mathbb{N} \ni [(L-\varepsilon < a_n < L+\varepsilon) \wedge (L-\varepsilon < c_n < L+\varepsilon)] \ \forall n \geq N$,

$\Rightarrow L-\varepsilon < a_n < b_n < c_n < L+\varepsilon$,

$\Rightarrow L-\varepsilon < b_n < L+\varepsilon$, and we are done.

MONOTONE CONVERGENCE THEOREM

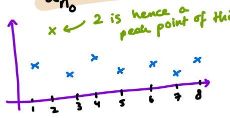
1. A sequence $\{a_n\}$ is monotonic if it is either non-decreasing or non-increasing,
ie $(\forall n \in \mathbb{N} : a_n \leq a_{n+1}) \vee (\forall n \in \mathbb{N} : a_n \geq a_{n+1})$.

2. The MCT states that, for a non-decreasing sequence $\{a_n\}$,
① If $\{a_n\}$ is bounded above, then $\lim_{n\to\infty} a_n = \text{lub}(\{a_n\})$;
② otherwise, $\lim_{n\to\infty} a_n = +\infty$;

& For a non-increasing sequence $\{b_n\}$,
① If $\{b_n\}$ is bounded below, then $\lim_{n\to\infty} b_n = \text{glb}(\{b_n\})$;
② otherwise, $\lim_{n\to\infty} b_n = -\infty$.

THE PEAK POINT LEMMA

1. A peak point for a sequence $\{a_n\}$ is an index $n_0 \in \mathbb{N}$, such that
$a_{n_0} > a_n \quad \forall n > n_0$.



2. Then, the PPL states that every sequence $\{a_n\}$ has a monotonic subsequence $\{a_{n_k}\}$, where n_k is a peak point.

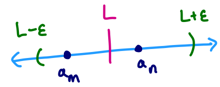
BOLZANO-WEIERSTRASS THEOREM

The BWT states that every bounded sequence has a convergent subsequence.

Proof. Assume $\{a_n\}$ is bounded.
By PPL, $\{a_{n_k}\}$ is monotonic,
where n_k is a peak point.
But $\{a_{n_k}\}$ is also bounded.
$\therefore$ By MCT, $\{a_{n_k}\}$ is convergent.

CAUCHY SEQUENCES

1. A sequence $\{a_n\}$ is Cauchy if
$\forall \varepsilon > 0 : \exists N \in \mathbb{N} \ni (|a_n - a_m| < \varepsilon \ \forall m,n \geq N)$.



* note: just because a sequence satisfies $\lim_{n\to\infty}(a_{n+1} - a_n) = 0$, it does not guarantee it is Cauchy.

CE : Let $a_n = 1 + \frac{1}{2} + \frac{1}{3} \cdots + \frac{1}{n}$.
$\Rightarrow \lim_{n\to\infty}(a_{n+1} - a_n) = \lim_{n\to\infty}\left(\frac{1}{n+1}\right) = 0$,
but $\{a_n\}$ is divergent and not Cauchy.

2. All convergent sequences are Cauchy.

Proof. Assume $\{a_n\}$ converges with limit L.
Then for all $\varepsilon > 0$,
$\exists N \ni |a_k - L| < \frac{\varepsilon}{2} \quad \forall k \geq N$.

Now let $m, n \geq N$. Then
$|a_m - L|$, $|a_n - L| < \frac{\varepsilon}{2}$ also.

By the Δ inequality
$|a_n - a_m| \leq |a_n - L| + |L + a_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$,
ie $|a_n - a_m| < \varepsilon$, and we are done.

COMPLETENESS THEOREM FOR $\mathbb{R}$

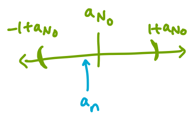
1. The CTFR states that every Cauchy sequence is convergent.

Proof.

- Lemma ①. Every Cauchy sequence is bounded.

Proof. Choose a $N_0 \ni [(n, m \geq N_0) \Rightarrow |a_n - a_m| < 1]$.

Let $m = N_0$. This tells us that
$|a_n - a_{N_0}| < 1$



$\Rightarrow |a_n| < |a_{N_0}| + 1$.

So, if $M = \max\{|a_1|+1, |a_2|+1, \ldots |a_{N_0}|+1\}$,
$|a_n| \leq M$, and we are done.

- Lemma ②. If a Cauchy sequence has a convergent subsequence $\{a_{n_k}\}$ that converges to L, $\{a_n\}$ also converges to L.

Proof. Choose $N_0 \ni [(m, n \geq N_0) \Rightarrow |a_n - a_m| < \frac{\varepsilon}{2}]$.

Now, since $\{a_{n_k}\}$ converges to L,
$\exists k_0 \ni n_{k_0} > N_0$.

Using the fact that $|a_{n_{k_0}} - L| < \frac{\varepsilon}{2}$,
we now choose any $n \geq N_0$.

Then $|a_n - L| < |a_n - a_{n_{k_0}}| + |a_{n_{k_0}} - L| = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$,
and we are done.

- Putting it all together:
→ If a sequence is Cauchy, by ① it is bounded.
→ By the BWT, $\{a_n\}$ contains a subsequence $\{a_{n_k}\}$ that converges.
→ Then by ②, $\{a_n\}$ must also converge, and we are done.

SERIES

1. Given a sequence $\{a_n\}$, the formal sum

$$a_1 + a_2 \cdots + a_n + \cdots$$

is called a series.

("formal" ⇒ there is no numerical meaning to the series yet).

2. The k^{th} partial sum S_k, where $k \in \mathbb{N}$, is defined by

$$S_k = \sum_{n=1}^{k} a_n.$$

CONVERGENCE OF A SERIES

1. A series converges if $\{S_k\}$ converges, where S_k is the kth partial sum.

2. If it does, and $\lim_{k\to\infty} S_k = L$, then we write

$$\sum_{n=0}^{\infty} a_n = L.$$

Otherwise, we say it diverges.

DIVERGENCE TEST

If a series $\sum_{n=0}^{\infty} a_n$ is such that

$$\lim_{n\to\infty} a_n \neq 0,$$

the series diverges.

* Note: the test cannot show a series converges! (eg $a_n = \frac{1}{n}$).

THE GEOMETRIC SERIES

1. A geometric series is a series of the form

$$\sum_{n=0}^{\infty} r^n = 1 + r + r^2 \cdots,$$

where r is denoted as the ratio of the series.

2. The kth partial sum of a as is given by

$$S_k = \frac{1-r^{k+1}}{1-r}.$$

3. $\lim_{k\to\infty} S_k$ exists (and hence the as converges) if $|r|<1$.

(otherwise $|r^{k+1}| \to \infty$ as $k \to \infty$)

4. If $|r|<1$, then

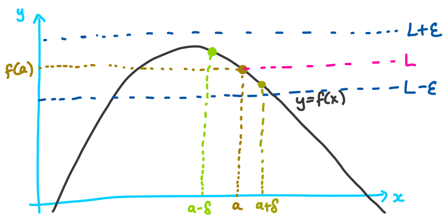
$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}.$$

Chapter 5: Limits & Continuity

LIMIT OF A FUNCTION

1. For any given function f, and $a \in \mathbb{R}$, we say f has a limit L as x approaches a, or that L is the limit of $f(x)$ at $x=a$, if for any tolerance $\varepsilon > 0$, we can find a cutoff distance $\delta > 0$ such that if $0 < |x-a| < \delta$, then $|f(x) - L| < \varepsilon$.

2. We write $\lim_{x\to a} f(x) = L$ to describe the above.



Notice how, if $x \in (a-\delta, a+\delta)$, $f(x) \in (L-\varepsilon, L+\varepsilon)$;
ie if $0 < |x-a| < \delta$, $|f(x) - L| < \varepsilon$.

* Also, note for any ε, we do not have to find the largest δ; we just need to find one value which works.

3. Also note:
   ① For $\lim_{x\to a} f(x)$ to exist, f must be defined on an open interval (α, β) containing $x=a$, except possibly at $x=a$.
   ② The value of $f(a)$, even if defined, does not affect the existence of a limit or its value.
   ③ If two functions are equal, except possibly at $x=a$, their limiting behaviour at a is identical.

SEQUENTIAL CHARACTERISATION OF LIMITS

Let f be defined on an open interval containing $x=a$, except possibly at $x=a$. Then,

$L = \lim_{x\to a} f(x)$ if and only if $\lim_{n\to\infty} f(x_n) = L$,

where $\{x_n\}$ is a sequence with $x_n \neq a$ & $x_n \to a$.

UNIQUENESS OF LIMITS FOR FUNCTIONS

If $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} f(x) = M$, $L = M$;
ie, the limit of a function is unique.

STRATEGIES TO SHOW A LIMIT DOES NOT EXIST

To show $\lim_{x\to a} f(x)$ does not exist, we can do either of the following:

① Find a sequence $\{x_n\}$ with $x_n \to a$, $x_n \neq a$, but $\lim_{n\to\infty} f(x_n)$ does not exist; or

② Find two sequences $\{x_n\}$ & $\{y_n\}$ with $x_n \to a$, $x_n \neq a$ & $y_n \to a$, $y_n \neq a$, and $\lim_{n\to\infty} x_n = L$ & $\lim_{n\to\infty} y_n = M$, but $L \neq M$.

ARITHMETIC RULES FOR FUNCTIONS

Let f & g be functions, and $a \in \mathbb{R}$.
Assume $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$.
Then:

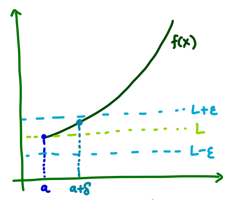
① If $f(x) = c$ $\forall x \in \mathbb{R}$, $\lim_{x\to a} f(x) = c$.
② $\forall c \in \mathbb{R}$: $\lim_{x\to a} cf(x) = cL$.
③ $\lim_{x\to a} [f(x) + g(x)] = L + M$.
④ $\lim_{x\to a} [f(x)\,g(x)] = LM$.
⑤ $\lim_{x\to a} \left[\frac{f(x)}{g(x)}\right] = \frac{L}{M}$ if $M \neq 0$.
*note: if $M = 0$ and $\lim_{x\to a} \left[\frac{f(x)}{g(x)}\right]$ exists, then $L = 0$.
⑥ $\lim_{x\to a} [(f(x))^\alpha] = L^\alpha$ $\forall \alpha > 0$, $L > 0$.

LIMITS OF POLYNOMIALS

If $p(x) = \alpha_0 + \alpha_1 x^1 \ldots + \alpha_n x^n$, then
$\lim_{x\to a} p(x) = p(a)$.

ONE-SIDED LIMITS

1. Let f be a function & $a \in \mathbb{R}$. We say f has a limit L as x approaches a from the right if for any $\varepsilon > 0$, $0 < |x-a| < \delta$, we can find a $\delta > 0$ such that if and $x > a$, then $|f(x) - L| < \varepsilon$.
   → we use the notation $\lim_{x\to a^+} f(x) = L$ to describe the above.



Notice $\forall x \in (a, a+\delta)$, $f(x) \in (L-\varepsilon, L+\varepsilon)$.
So $\lim_{x\to a^+} f(x) = a$.
(A similar example exists for if x approaches a from the left)

2. Similarly, f has a limit L as x approaches a from the left if
   $\forall \varepsilon > 0$: $\exists \delta > 0$ such that if $0 < |x-a| < \delta$, and $x < a$, then $|f(x) - L| < \varepsilon$.
   → similarly, we use the notation $\lim_{x\to a^-} f(x) = L$ to describe the above.

3. Lastly, note that
   $\lim_{x\to a} f(x) = L \iff \lim_{x\to a^-} f(x) = L = \lim_{x\to a^+} f(x)$.

*note: this is valid for $L = \pm\infty$

SQUEEZE THEOREM FOR FUNCTIONS

Assume functions f, g & h are defined on an open interval I containing $x=a$, except possibly at $x=a$.
Suppose, then, that $\forall x \in I$, $g(x) \leq f(x) \leq h(x)$, and $\lim_{x\to a} g(x) = L = \lim_{x\to a} h(x)$.
We can then imply $\lim_{x\to a} f(x)$ exists, and $\lim_{x\to a} f(x) = L$.

*note: it is possible for $L = \pm\infty$.

FUNDAMENTAL LOG LIMIT

$\lim_{x\to\infty} \frac{\ln(x)}{x} = 0.$

Proof. Observe $0 \leq \frac{\ln(x)}{x} = \frac{2\ln(\sqrt{x})}{x} = \frac{2}{x^{1/2}}\left[\frac{\ln(x^{1/2})}{x^{1/2}}\right] \leq \frac{2}{x^{1/2}}$ (since $\frac{\ln(x)}{x} < 1.$)

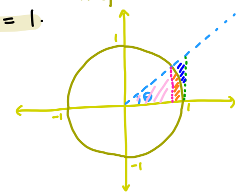
But $\lim_{x\to\infty} \frac{2}{x^{1/2}} = 0.$

Hence, by the Squeeze Theorem, $\lim_{x\to\infty} \frac{\ln(x)}{x} = 0.$ ▨

THE FUNDAMENTAL TRIGONOMETRIC LIMIT

$\lim_{\theta\to 0} \frac{\sin\theta}{\theta} = 1.$

Proof. (For the + direction).



Clearly, $\square < \square + \square < \square + \square + \square.$

$\therefore \frac{\sin(\theta)\cos(\theta)}{2} < \frac{\theta}{2} < \frac{\tan(\theta)}{2}.$

$(\cdot \frac{2}{\sin\theta})$ $\therefore \cos\theta < \frac{\theta}{\sin\theta} < \frac{1}{\cos\theta}.$

$(\frac{1}{\cdot})$ $\therefore \frac{1}{\cos\theta} > \frac{\sin\theta}{\theta} > \cos\theta \quad \forall \theta \in \mathbb{R}.$

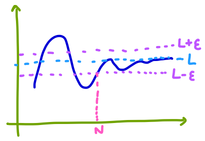
But $\lim_{\theta\to 0} \frac{1}{\cos\theta} = \lim_{\theta\to 0} \cos\theta = 1.$ Hence $\lim_{\theta\to 0} \frac{\sin\theta}{\theta} = 1.$ ※

LIMITS AT INFINITY & ASYMPTOTES

LIMIT AT INFINITY

1. We say a function f has a limit L as $x\to\infty$ if $\forall \varepsilon > 0$: $\exists N > 0$ such that if $x > N$, then $|f(x) - L| < \varepsilon$.

*note: we can define limits at $-\infty$ in a similar manner.

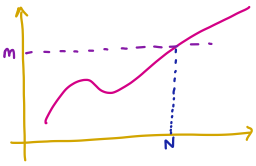


Notice $\forall x > N$, $f(x) \in (L-\varepsilon, L+\varepsilon)$.

2. We write $L = \lim_{x\to\infty} f(x)$ to denote the above.

3. Also, if $L = \lim_{x\to\infty} f(x)$ or $L = \lim_{x\to-\infty} f(x)$, L is called a horizontal asymptote of $f(x)$.

4. Moreover, if $\forall M > 0$, $\exists N > 0$ such that $x > N \Rightarrow f(x) > M$, then $f(x)$ approaches ∞ as x tends to ∞, and write $\lim_{x\to\infty} f(x) = \infty$.

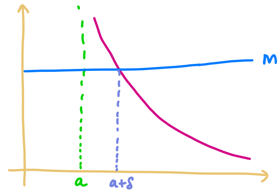


Notice $\forall x > N$, $f(x) > M$.

INFINITE LIMITS

1. We say f has a limit of ∞ as x approaches a from above if $\forall M > 0$, $\exists \delta > 0$ such that if $x > a$ and $0 < |x - a| < \delta$, then $f(x) > M$.

*note: a similar def^n exists for if f has a limit of $-\infty$.

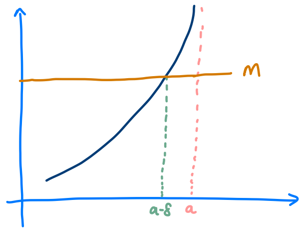


Notice if $x \in (a, a+\delta)$, then $f(x) > M$.

2. We write $\lim_{x\to a^+} f(x) = \infty$ to denote the above.

3. Similarly, f has a limit of ∞ as x approaches a from below if $\forall M > 0$, $\exists \delta > 0$ such that if $x < a$ & $0 < |x - a| < \delta$, then $f(x) > M$.

*again, a similar def^n exists for $-\infty$.



Notice if $x \in (a-\delta, a)$, then $f(x) > M$.

4. We write $\lim_{x\to a^-} f(x) = \infty$ to denote the above.

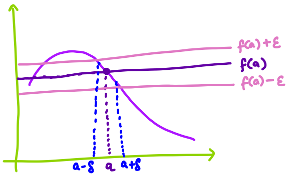
5. If any of $\lim_{x\to a^\pm} f(x) = \pm\infty$ occur, we say the line $x = a$ is a vertical asymptote for the function f.

CONTINUITY

1. A function f is continuous at $x=a$ if $\lim_{x\to a} f(x)$ exists, and $\lim_{x\to a} f(x) = f(a)$.

(Otherwise, we say it is discontinuous at $x=a$, or that $x=a$ is a point of discontinuity for f.)

2. Alternatively, f is continuous at $x=a$ if $\forall \varepsilon > 0$: $\exists \delta > 0$ such that if $|x-a|<\delta$, then $|f(x) - f(a)| < \varepsilon$.



Notice if $x \in (a-\delta, a+\delta)$, $f(x) \in (f(a)-\varepsilon, f(a)+\varepsilon)$.

3. Lastly, f is continuous at $x=a$ if and only if $\lim_{h\to 0} f(a+h) = f(a)$.

SEQUENTIAL CHARACTERISATION OF CONTINUITY

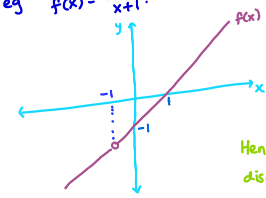
1. A function f is continuous at $x=a$ if and only if all sequences $\{x_n\}$ that converge to a satisfy $\lim_{n\to\infty} f(x_n) = f(a)$.

TYPES OF DISCONTINUITIES

REMOVABLE DISCONTINUITY

A removable discontinuity is a point $x=a$ on f such that if we let $f(a) = \lim_{x\to a} f(x)$, the function f becomes continuous.

eg $f(x) = \frac{x^2-1}{x+1}$. ($f(x)$ is discontinuous at $x=-1$).



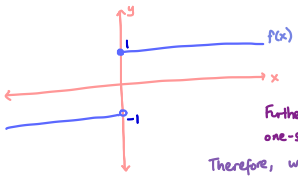
Notice that if we let $f(-1) = \lim_{x\to -1} f(x) = -2$, $f(x)$ simply becomes $f(x) = x-1$, which is continuous $\forall x \in \mathbb{R}$.

Hence $x=-1$ is a removable discontinuity.

FINITE JUMP DISCONTINUITY

A finite jump discontinuity at a point $x=a$ on f occurs when $\lim_{x\to a^+} f(x)$ and $\lim_{x\to a^-} f(x)$ both exist, but $\lim_{x\to a^+} f(x) \neq \lim_{x\to a^-} f(x)$, & $\left|\lim_{x\to a^+} f(x) - \lim_{x\to a^-} f(x)\right|$ is finite.

eg $f(x) = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases}$ ($f(x)$ is discontinuous at $x=0$.)



Notice that we cannot simply "replace" the discontinuity, as $\lim_{x\to a} f(x)$ does not exist!

Furthermore, the "jump" between the one-sided limits is finite.

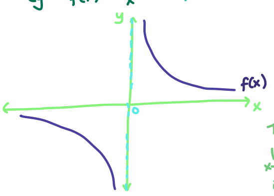
Therefore, we call the point at $x=0$ a finite jump discontinuity.

☆ note: both jump discontinuities and the oscillatory discontinuity are classified as essential discontinuities, as there is no way to "replace" the discontinuity to make $f(x)$ continuous!

INFINITE JUMP DISCONTINUITY

An infinite jump discontinuity at a point $x=a$ on f occurs when $\lim_{x\to a^+} f(x)$ & $\lim_{x\to a^-} f(x)$ both exist, but $\lim_{x\to a^+} f(x) \neq \lim_{x\to a^-} f(x)$, and $\left|\lim_{x\to a^+} f(x) - \lim_{x\to a^-} f(x)\right|$ is infinite.

eg $f(x) = \frac{1}{x}$. ($f(x)$ is discontinuous at $x=0$.)



Notice, again, we cannot simply "replace" the discontinuity, as $\lim_{x\to a} f(x)$ does not exist!

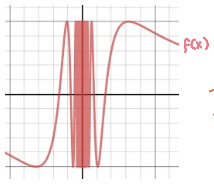
Then, $\lim_{x\to a^+} f(x) = +\infty$, and $\lim_{x\to a^-} f(x) = -\infty$; so, the "jump" is infinite.

Therefore, the point at $x=0$ is an infinite jump discontinuity.

OSCILLATORY DISCONTINUITY

An oscillatory discontinuity is a point $x=a$ on f where although f is bounded near $x=a$, it does not have a limit because of infinitely many oscillations near a.

eg $f(x) = \sin(\frac{1}{x})$. ($f(x)$ is discontinuous at $x=0$.)



Notice that, near $x=0$, $f(x)$ oscillates infinitely many times!

Therefore, the point at $x=0$ is called an oscillatory discontinuity.

CONTINUITY OF STANDARD FUNCTIONS

POLYNOMIALS

Earlier, we showed that for any polynomial $p(x)$, $\lim_{x\to a} p(x) = p(a)$ $\forall a \in \mathbb{R}$.

Hence, $p(x)$ is continuous for all $x \in \mathbb{R}$.

SIN(X) AND COS(X)

1. Note $\lim_{x\to 0} \sin(x) = 0$ & $\lim_{x\to 0} \cos(x) = 1$.

2. Then,
$$\lim_{x\to a} \sin(x) = \lim_{h\to 0} (\sin(a+h))$$
$$= \lim_{h\to 0} [\sin(a)\cos(h) + \cos(a)\sin(h)]$$
$$= \sin(a) + 0 \cdot \cos(a)$$
$\therefore \lim_{x\to a} \sin(x) = \sin(a)$,

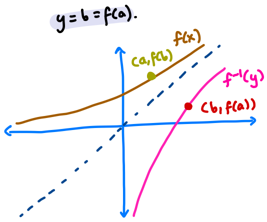
and hence $\forall x \in \mathbb{R}$ $f(x) = \sin(x)$ is continuous.

3. Similarly,
$$\lim_{x\to a} \cos(x) = \lim_{h\to 0} (\cos(a+h))$$
$$= \lim_{h\to 0} [\cos(a)\cos(h) - \sin(a)\sin(h)]$$
$$= \cos(a) - 0 \cdot \sin(a)$$
$\therefore \lim_{x\to a} \cos(x) = \cos(a)$,

and so $\forall x \in \mathbb{R}$ $f(x) = \cos(x)$ is continuous.

INVERSE FUNCTIONS

If $y=f(x)$ is invertible, with inverse $x=f^{-1}(y)$, then if $f(a)=b$ and f is continuous on an open interval containing $x=a$, then f^{-1} is continuous at $y=b=f(a)$.



Recall $f^{-1}(y)$ is the reflection of $f(x)$ across $y=x$.

Hence, if there were no breaks in $f(x)$, there would be no breaks in $f^{-1}(y)$, since reflection does not produce breaks.

e^x & $\ln(x)$

1. Observe that
$$\lim_{x\to a} e^x = \lim_{h\to 0} e^{a+h}$$
$$= \lim_{h\to 0} (e^h \cdot e^a)$$
$$= 1 \cdot e^a$$
$$= e^a,$$
and hence $\forall x \in \mathbb{R}$ $f(x)$ is continuous.

2. Since $\ln(x)$ is the inverse function of e^x, $\ln(x)$ must also be continuous $\forall x \in \mathbb{R}$.

ARITHMETIC RULES FOR CONTINUOUS FUNCTIONS

1. If functions f & g are continuous at $x=a$, then:
   ① $f+g$ is continuous at $x=a$;
   ② fg is continuous at $x=a$; and
   ③ $\frac{f}{g}$ is continuous at $x=a$, provided $g(a) \neq 0$.

2. If f is continuous at $x=a$, and g is continuous at $x=f(a)$, then $h=gf(x)$ is continuous at $x=a$

Proof. Let f be continuous at $x=a$, and g be cont. at $x=f(a)$.
Let $h=gf(x)$.
Then $\exists \{x_n\}$, where $x_n \to a$, that satisfies $\lim_{n\to\infty} f(x_n)=f(a)$.
However, since $f(x_n) \to f(a)$, we can now imply
$\lim_{n\to\infty} g[f(x_n)] = g[f(a)]$, by utilising the Sequential Characterisation of Limits again.
Hence $\lim_{n\to\infty} h(x_n) = h(a)$. ※

CONTINUITY ON AN INTERVAL

1. We say f is continuous on (a,b) if $\forall x \in (a,b)$, f is continuous.
   (if $(a,b)=\mathbb{R}$, then we say f is continuous on $\mathbb{R}$.)

2. We say f is continuous on $[a,b]$ if:
   ① f is continuous $\forall x \in (a,b)$;
   ② $\lim_{x\to a^+} f(x) = f(a)$; and
   ③ $\lim_{x\to b^-} f(x) = f(b)$.

INTERMEDIATE VALUE THEOREM

Assume f is continuous on $[a,b]$, and either $f(a) < \alpha < f(b)$ or $f(a) > \alpha > f(b)$.
Then $\exists c \in (a,b)$ such that $f(c) = \alpha$.

Proof. We first assume $\alpha = 0$; so, $f(a) < 0 < f(b)$.
Let $S = \{x \in [a,b] \mid f(x) \leq 0\}$.
Since $S \neq \emptyset$ as $a \in S$ & S is bounded above by b,
$\exists c = \text{lub}(S)$.
Then, $\exists \{x_n\} \subseteq S$ with $x_n \to c$
$[a,b]$, and $f(x_n) \leq 0$ $\forall n \in \mathbb{N}$,
Since f is continuous on $[a,b]$, and $f(x_n) \leq 0$ $\forall n \in \mathbb{N}$,
the Sequential Characterisation of Continuity shows that
$f(c) = \lim_{n\to\infty} f(x_n) \leq 0$. (So $f(c) \leq 0$.)

Next, let $y_n = c + \frac{b-c}{n}$.
Clearly, $c < y_n \leq b$, so $f(y_n) > 0$.
Since $y_n \to c$, $\therefore f(c) = \lim_{n\to\infty} f(y_n) \geq 0$. (So $f(c)=0$).

To obtain the more general result, we consider the functions $g(x) = f(x) - \alpha$ and $h(x) = \alpha - f(x)$ in the cases $f(a) < \alpha < f(b)$ & $f(a) > \alpha > f(b)$ respectively. ※

*note: this proof relies on the fact that $\mathbb{R}$ is continuous.

APPROXIMATE SOLUTIONS OF EQUATIONS

ROOTS OF A POLYNOMIAL

1. We can use the IVT to show whether a polynomial has any real roots.

eg $p(x) = x^5 + x - 1$.
We first note that $p'(x) = 5x^4 + 1 > 0$ $\forall x \in \mathbb{R}$; and so $p(x)$ is increasing $\forall x \in \mathbb{R}$.
Then, $p(0) = -1$ (<0) & $p(1) = 1$ (>0).
Since f is continuous over $[0,1]$, the IVT implies $\exists c \ni p(c) = 0$.
Lastly, since $p(x)$ is increasing $\forall x \in \mathbb{R}$, c is the only real root of $p(x)$.

2. We can also use the IVT to approximate the value of a polynomial's real roots.

eg $p(x) = x^5 + x - 1$.
From above, we know the root $c \in [0,1]$.
Then, we can test different values in this interval;
eg $p(\frac{1}{2}) = -\frac{15}{32} < 0$. (So $c \in (\frac{1}{2}, 1)$. since $p(\frac{1}{2})<0$ but $p(1)>0$.)
We can then proceed to test the midpoints of the successive intervals, and repeat this process to find smaller intervals in which c resides.
eg $p(.75) = -0.0126.. < 0$. ($\therefore c \in (0.75, 1)$).
$p(.875) = 0.3879.. > 0$ ($\therefore c \in (0.75, 0.875)$)
etc.

THE BISECTION METHOD

In fact, the method described to the left is called the Bisection Method, which can help us find an approximate solution to $f(x) - g(x) = 0$.

Steps:
① Let $F(x) = f(x) - g(x)$.
Find a_0, b_0 such that $F(a_0) < 0$ & $F(b_0) > 0$.
② By IVT, $a_0 < c < b_0$, where $F(c) = 0$.
③ Then, evaluate $F(\frac{a_0 + b_0}{2})$. $(= F(d))$.
If $F(d)$ & $F(a_0)$ have the same sign, let $a_1 = a_0$ & $b_1 = d$ to obtain a new interval $[a_1, b_1]$, which contains a solution to the eq^{n}.
Otherwise, let $a_1 = d$ & $b_1 = b_0$.
④ We can repeat step ③ to obtain smaller intervals in which c is contained.

THE EXTREME VALUE THEOREM

GLOBAL MAXIMA & MINIMA

Suppose $f: I \to \mathbb{R}$, where I is an interval.

Then,

① c is a global maximum for f on I if $c \in I$, and $f(x) \leq f(c)$ $\forall x \in I$.

② c is a global minimum for f on I if $c \in I$, and $f(x) \geq f(c)$ $\forall x \in I$.

THE EVT

Suppose f is <u>continuous</u> on $[a,b]$.
Then $\exists c, d \in [a,b]$ such that
$f(c) \leq f(x) \leq f(d)$ $\forall x \in [a,b]$.

Proof.

Step 1: We show $f([a,b]) = \{f(x) \mid x \in [a,b]\}$ is bounded.

How? Assume this is not the case.
Then $\forall n \in \mathbb{N}$ $\exists x_n \in [a,b]$ such that $|f(x_n)| > n$.
Since $\{x_n\} \subset [a,b]$, BWT tells us $\exists \{x_{n_k}\}$ which converges to some point $t \in [a,b]$.
Then, the SCC tells us that $f(x_{n_k}) \to f(t)$.
However, this is impossible since $|f(x_{n_k})| > n_k$,
so $\{f(x_{n_k})\}$ is not bounded. Hence f is bounded on $[a,b]$. #

Step 2. We show $\exists d \in [a,b] \ni f(x) \leq f(d)$ $\forall x \in [a,b]$.

How? First, let $M = \text{lub}(\{f(x) \mid x \in [a,b]\})$.
Then, $\forall n \in \mathbb{N}$: $(M - \frac{1}{n} < M) \Rightarrow \exists x_n \in [a,b] \ni \left[f(x_n) \in (M - \frac{1}{n}, M]\right]$.
(Since f is continuous.)
By BWT, $\exists \{x_{n_k}\}$, with $x_{n_k} \to d \in [a,b]$.
Then, by SCC & the Squeeze Theorem,
$f(d) = \lim_{k \to \infty} f(x_{n_k}) = M$.
$\therefore f(x) \leq f(d)$ $\forall x \in [a,b]$, and we are done. ✲

Step 3. We show $\exists c \in [a,b] \ni f(c) \leq f(d)$ $\forall x \in [a,b]$.

How? Let $L = \text{glb}(\{f(x) \mid x \in [a,b]\})$. We can show $\exists c \in [a,b]$ such that $f(c) = L$, using a similar argument as stage 2

$\therefore f(c) \leq f(x) \leq f(d)$. $\forall x \in [a,b]$. ⊠

CURVE SKETCHING (PART 1)

We can use a strategy to sketch basic curves:

① Determine the domain of f.
② Determine whether f has any symmetries (ie if f is even or odd.)
③ Determine where f changes sign, and plot these points.
④ Find any discontinuity points for f.
⑤ Identify the nature of these points, and evaluate the relevant one/two-sided limits at these points.
⑥ Draw any vertical asymptotes.
⑦ Identify whether any horizontal asymptotes exist, and draw them.
⑧ Sketch the graph. If needed, plot some sample points.

UNIFORM CONTINUITY

We say f is uniformly continuous on $S \subseteq \mathbb{R}$ if $\forall \varepsilon > 0$. $\exists \delta > 0$ such that if $x, y \in S$ and $|x - y| < \delta$, then $|f(x) - f(y)| < \varepsilon$.

SEQUENTIAL CHARACTERISATION OF UNIFORM CONTINUITY

Assume $f(x)$ is defined on $S \subseteq \mathbb{R}$. Then, the following 2 statements are <u>equivalent</u>:

i) $f(x)$ is uniformly continuous on S; and
ii) if $\{x_n\}, \{y_n\} \subseteq S$ with $\lim_{n \to \infty} |x_n - y_n| = 0$, then $\lim_{n \to \infty} |f(x_n) - f(y_n)| = 0$.

UNIFORM CONTINUITY ON [a,b]

If f is continuous on $[a,b]$, then f is also uniformly continuous on $[a,b]$.

Chapter 6: Derivatives

We say the function f is differentiable at $t=a$ if $\lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$ exists, and denote the above limit as $f'(a)$.

THE TANGENT LINE

1. Assume f is differentiable at $x=a$. Then, the tangent line to the graph of f at $x=a$ is the line passing through $(a, f(a))$ with slope $f'(a)$.

2. It follows that the equation of the tangent line is

$$y = f(a) + f'(a)[x-a].$$

*note: we cannot define the derivative as "the slope of the tangent line!"

DIFFERENTIABILITY VS CONTINUITY

If f is differentiable at $t=a$, f is continuous at $t=a$.

Proof. Since f is differentiable at $t=a$,

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} \text{ exists.}$$

Let $h=t-a$. Then we also know

$$\lim_{t\to a} \frac{f(t)-f(a)}{t-a} \text{ exists.}$$

However, since the denominator approaches zero, the numerator must also approach zero, ie

$$\lim_{t\to a} (f(t)-f(a)) = 0 \quad \text{or} \quad \lim_{t\to a} f(t) = f(a).$$

This, in turn, implies $f(x)$ is continuous at a. #

*note that continuity does not imply differentiability!
(eg $f(x) = |x|$ at $x=0$)

* note that "sharp points" are not differentiable.

THE DERIVATIVE FUNCTION

1. We say f is differentiable on an interval I if $f'(a)$ exists $\forall a \in I$.

2. Then, we define the derivative function f' to be

$$f'(t) = \lim_{h\to 0} \frac{f(t+h)-f(t)}{h}.$$

($f'(t)$ is simply the derivative of f at t for each $t \in I$.)

3. Leibniz Notation: note that

$$\frac{df}{dt} = \frac{d}{dt}(f) = f',$$

and

$$\left.\frac{df}{dt}\right|_{x=a} = f'(a).$$

HIGHER DERIVATIVES

1. The second derivative of f, denoted as f'', is defined to be

$$f'' = \frac{d}{dx}(f').$$

Similarly, the 3rd derivative of f, f''', is defined as

$$f''' = \frac{d}{dx}(f'').$$

2. In general, $\forall n \geq 1$,

$$f^{(n+1)} = \frac{d}{dx}(f^{(n)}),$$

and $f^{(n)}$ is called the n^{th} derivative of f.

DERIVATIVES OF ELEMENTARY FUNCTIONS

CONSTANT FUNCTION

We can show if $f(x)=c$ for a constant $c\in\mathbb{R}$, then $f'(x)=0 \quad \forall x\in\mathbb{R}$.

Proof. $f'(x) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$

$= \lim_{h\to 0}\frac{c-c}{h}$

$= 0.$ ∎

LINEAR FUNCTION

Let $f(x)=mx+b$. Then $f'(x)=m \quad \forall x\in\mathbb{R}$.

Proof. $f'(a) = \lim_{h\to 0}\frac{f(a+h)-f(a)}{h}$

$= \lim_{h\to 0}\frac{(ma+mh+c)-(ma+c)}{h}$

$= \lim_{h\to 0}\frac{mh}{h}$

$= m.$ ∎

SIN(X) & COS(X)

1. First, we can derive that $\lim_{h\to 0}\frac{\cos(h)-1}{h} = 0.$

Proof. $\lim_{h\to 0}\frac{\cos(h)-1}{h} = \lim_{h\to 0}\left(\frac{\cos(h)-1}{h}\right)\left(\frac{\cos(h)+1}{\cos(h)+1}\right)$

$= \lim_{h\to 0}\left(\frac{\cos^2(h)-1}{h(\cos(h)+1)}\right)$

$= \lim_{h\to 0}\left(\frac{-\sin^2(h)}{h(\cos(h)+1)}\right)$

$= \lim_{h\to 0}\left(\frac{\sin(h)}{h}\right)\cdot\lim_{h\to 0}\left(\frac{-\sin(h)}{\cos(h)+1}\right)$

$= 1\cdot 0 \quad \because$ as $h\to 0$, $\sin(h)\to 0$ & $\cos(h)+1\to 2$

$\therefore \lim_{h\to 0}\frac{\cos(h)-1}{h} = 0.$ ∎

2. Now, we can show if $f(x)=\sin(x)$, then $f'(x)=\cos(x)$.

Proof. $f'(x) = \lim_{h\to 0}\frac{\sin(x+h)-\sin(x)}{h}$

$= \lim_{h\to 0}\frac{(\sin(x)\cos(h)+\cos(x)\sin(h))-\sin(x)}{h}$

$= \lim_{h\to 0}(\sin(x))\left(\frac{\cos(h)-1}{h}\right) + \lim_{h\to 0}(\cos(x))\left(\frac{\sin(h)}{h}\right)$

$= \sin(x)\cdot 0 + \cos(x)\cdot 1$

$\therefore \frac{d}{dx}(\sin(x)) = \cos(x).$ ∎

3. Similarly, if $f(x)=\cos(x)$, then $f'(x)=-\sin(x)$.

Proof. $f'(x) = \lim_{h\to 0}\frac{\cos(x+h)-\cos(x)}{h}$

$= \lim_{h\to 0}\frac{(\cos(x)\cos(h)-\sin(x)\sin(h))-\cos(x)}{h}$

$= \lim_{h\to 0}(\cos(x))\left(\frac{\cos(h)-1}{h}\right) - \lim_{h\to 0}(\sin(x))\left(\frac{\sin(h)}{h}\right)$

$= 0\cdot\cos(x) - 1\cdot\sin(x)$

$\therefore \frac{d}{dx}(\cos(x)) = -\sin(x).$ ∎

a^x & e^x

1. We can show if $f(x)=a^x$, then $f'(x) = C_a(a^x)$, where $C_a = f'(0)$.

Proof. $f'(x) = \lim_{h\to 0}\frac{a^{x+h}-a^x}{h}$

$= \lim_{h\to 0}\frac{a^x(a^h-1)}{h}$

$= a^x\cdot\lim_{h\to 0}\frac{a^h-1}{h}$

$= a^x\cdot f'(0)$

$\therefore f'(x) = C_a\cdot(a^x).$ ∎

2. Then, "e" ($=2.718\ldots$) is the unique value such that if $f(x)=e^x$, then $f'(0)=1$.

3. We can also prove $\lim_{h\to 0}\frac{e^h-1}{h} = 1.$

Proof. If $f(x)=e^x$, then $f'(0)=1$.

$\Rightarrow 1 = \lim_{h\to 0}\frac{e^{0+h}-e^0}{h}$

$\therefore 1 = \lim_{h\to 0}\frac{e^h-1}{h}.$ ∎

4. Finally, we can prove if $f(x)=e^x$, then $f'(x)=e^x$.

Proof. $f'(x) = \lim_{h\to 0}\frac{e^{x+h}-e^x}{h}$

$= \lim_{h\to 0}(e^x)\left(\frac{e^h-1}{h}\right)$

$= 1\cdot e^x$

$\therefore \frac{d}{dx}(e^x) = e^x.$ ∎

5. Additionally, we can show if $f(x)=a^x$, then $f'(x) = a^x\ln(a)$.

Proof. $f(x) = a^x = (e^{\ln(a)})^x = e^{x\ln(a)}$.

$\therefore f'(a) = \lim_{h\to 0}\frac{a^{x+h}-a^x}{h}$

$= \lim_{h\to 0}(a^x)\left[\frac{a^h-1}{h}\right]$

$= \lim_{h\to 0}(a^x)\left[\frac{e^{h\ln(a)}-1}{h}\right]$

$= \lim_{h\to 0}(a^x)(\ln(a))\left[\frac{e^{h\ln(a)}-1}{h\ln(a)}\right]$

$= 1\cdot a^x\ln(a)$

$\therefore f'(x) = a^x\ln(a).$ ∎

LINEAR APPLICATION

1. Let $y=f(x)$ be differentiable at $x=a$.
Then, the "linear approximation" to f at $x=a$ is the function

$$L_a^f(x) = f(a) + f'(a)(x-a).$$

* $L_a^f(x)$ is also called the "linearisation" or the "tangent line approximation" to f at $x=a$.

* Why is it a linear "approximation"?
→ for values of x close to a, we have that

$$f'(a) \approx \frac{f(x)-f(a)}{x-a}.$$

So $(x-a)f'(a) \approx f(x) - f(a)$

or $f(x) \approx f(a) + (x-a)f'(a)$.

2. There are 3 main properties of L_a:

① $L_a(a) = f(a)$.

② L_a is differentiable and $L_a'(a) = f'(a)$.

③ L_a is the only first degree polynomial with properties ① & ②.

ERROR TO LINEAR APPROXIMATION

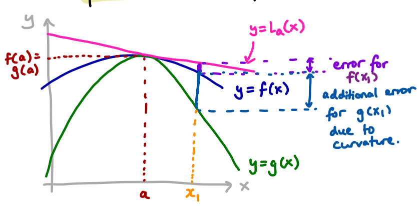
1. Let $y=f(x)$ be differentiable at $x=a$.
Then, the error in using linear approximation to estimate $f(x)$ is given by

$$\text{error} = |f(x) - L_a(x)|.$$

2. There are multiple factors that affect the error in linear approximation:

① Generally, as $|x-a|$ increases, the error in L_a increases.

② Generally, the more "curved" the graph is at $x=a$, the greater the potential error of L_a.



→ alternatively, we can use this more precise definition:

if $|f''(x)| < M \;\; \forall x \in I$, where I is an interval that contains a point a,
then $$|f(x) - L_a(x)| \leq \frac{M}{2}(x-a)^2.$$

*since $f''(x)$ is a measure of the "curvature" of the graph.

APPLICATIONS OF LINEAR APPROXIMATION

ESTIMATING CHANGE

1. Assume we know $f(a)$ at some point a.
We can use L_a to figure out what change in the value of $f(x)$ we can expect if we move to a point x_1 near a.

2. In other words, we want to know what $\Delta y = f(x_1) - f(a)$ will be if we change the variable by $\Delta x = x_1 - a$ units.

3. Subsequently, after using L_a, we find that

$$\begin{aligned} \Delta y &= f(x_1) - f(a) \\ &\cong L_a(x_1) - f(a) \\ &= (f(a) + f'(a)(x_1-a)) - f(a) \\ &= f'(a)(x_1-a) \end{aligned}$$

$\therefore \Delta y \cong (f'(a))\Delta x.$

QUALITATIVE BEHAVIOUR OF FUNCTIONS

1. We can also use L_a to study the "qualitative" behaviour of functions;
e.g. $y = e^{-x^2}$.

2. First, we begin with a simpler function;
ie $h(u) = e^u$.
Then by definition $h(0) = h'(0) = 1$.

3. So $L_0^h(u) = 1+u$ is the tangent line to $h(u)$ at $(0,1)$.
Hence $e^u \approx 1+u$ if u is near 0.

4. However, if x is close to 0, then $-x^2$ is very close to 0.
So, if we let $u = -x^2$, we get that

$$y = e^{-x^2} = h(-x^2) \approx 1 - x^2$$

* so $y = 1-x^2$ is a very good approximation to $y = e^{-x^2}$ at values close to 0.

*note: $1-x^2$ is NOT the linear approximation for $y = e^{-x^2}$

NEWTON'S METHOD

Newton's Method uses the linear approximation to a differentiable function to approximate the solution to an equation of the type $f(x)=0$.

METHOD

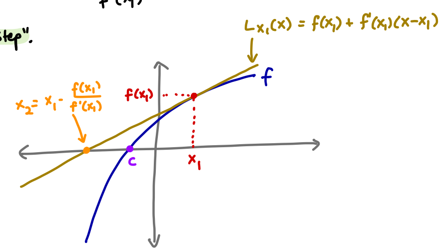
1. First, pick a point x_1 that is reasonably close to a point c with $f(c)=0$.

   * we can use the IVT to help us find such an x_1.

2. If f is differentiable at $x=x_1$, then we can approximate f near x_1 by using
$$f(x) \approx L_{x_1}(x) = f(x_1) + f'(x_1)(x-x_1).$$

3. Since $f(x) \approx L_{x_1}(x)$, we can infer $f(0) \approx L_{x_1}(0)$. So if $f'(x_1) \neq 0$, we can approximate c by x_2, where $L_{x_1}(x_2) = 0$.

4. After expanding and simplifying, we will get that $x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$, and so this is our next "step".



5. We continue this process indefinitely, which results in a sequence $\{x_n\}$ such that
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad \forall n \in \mathbb{N},$$
where x_{n+1} is the point at which the tangent line to f through $(x_n, f(x_n))$ crosses the x-axis.

6. Ultimately, we will observe that (generally) $\{x_n\}$ converges to a c with $f(c)=0$.

ARITHMETIC RULES OF DIFFERENTIATION

CONSTANT MULTIPLE RULE

Let f be differentiable at $x=a$. Then if $h(x) = cf(x)$, then $h'(a) = cf'(a)$.

Proof.
$$(cf)'(a) = \lim_{h\to 0} \frac{(cf)(a+h) - (cf)(a)}{h} = c \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$
$$\therefore (cf)'(a) = cf'(a). \quad \blacksquare$$

SUM RULE

Let f & g be differentiable at $x=a$. Then $h(x) = f(x) + g(x)$ is differentiable at $x=a$ and $h'(a) = f'(a) + g'(a)$.

Proof.
$$(f+g)'(a) = \lim_{h \to 0} \frac{(f+g)(a+h) - (f+g)(a)}{h}$$
$$= \lim_{h\to 0} \frac{f(a+h) + g(a+h) - f(a) - g(a)}{h}$$
$$= \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} + \lim_{h\to 0} \frac{g(a+h) - g(a)}{h}$$
$$\therefore (f+g)'(a) = f'(a) + g'(a). \quad \blacksquare$$

PRODUCT RULE

Let f & g be differentiable at $x=a$. Then $h(x) = f(x)g(x)$ is also differentiable at $x=a$, and
$$h'(a) = f'(a)g(a) + f(a)g'(a).$$

Proof.
$$(fg)'(a) = \lim_{h\to 0} \frac{(fg)(a+h) - (fg)(a)}{h}$$
$$= \lim_{h \to 0} \frac{f(a+h)g(a+h) - f(a+h)g(a) + f(a+h)g(a) - f(a)g(a)}{h}$$
$$= \lim_{h\to 0} f(a+h)\left[\frac{g(a+h) - g(a)}{h}\right] + \lim_{h \to 0} g(a)\left[\frac{f(a+h) - f(a)}{h}\right]$$
$$\therefore (fg)'(a) = f(a)\cdot g'(a) + f'(a)\cdot g(a). \quad \blacksquare$$

RECIPROCAL RULE

Let g be differentiable at $x=a$. If $g(a) \neq 0$, then $h(x) = \frac{1}{g(x)}$ is differentiable at $x=a$ also and
$$h'(a) = \frac{-g'(a)}{[g(a)]^2}.$$

Proof.
$$\left(\tfrac{1}{f}\right)'(a) = \lim_{h\to 0} \frac{\frac{1}{f(a+h)} - \frac{1}{f(a)}}{h}$$
$$= \lim_{h \to 0} \frac{f(a) - f(a+h)}{hf(a)f(a+h)}$$
$$= -\lim_{h\to 0} \frac{f(a+h) - f(a)}{h} \cdot \lim_{h \to 0} \frac{1}{f(a+h)f(a)}$$
$$= -f'(a) \cdot \frac{1}{[f(a)]^2} \quad \because f \text{ is continuous at } x=a$$
$$\therefore \left(\tfrac{1}{f}\right)'(a) = \frac{-f'(a)}{[f(a)]^2}. \quad \blacksquare$$

QUOTIENT RULE

Let f & g be differentiable at $x=a$. Then if $g(a) \neq 0$, $h(x) = \frac{f(x)}{g(x)}$ is also differentiable at a and
$$h'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{[g(a)]^2}.$$

Proof.
$$\left(\tfrac{f}{g}\right)'(a) = \left(f \cdot \tfrac{1}{g}\right)'(a)$$
$$= f(a)\cdot\left(\tfrac{1}{g}\right)'(a) + f'(a)\cdot\left(\tfrac{1}{g}\right)(a) \quad \text{(by the Product Rule)}$$
$$= f(a) \cdot \frac{-g'(a)}{[g(a)]^2} + f'(a)\cdot\frac{1}{g(a)} \quad \text{(by the Reciprocal Rule)}$$
$$\therefore \left(\tfrac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{[g(a)]^2}. \quad \blacksquare$$

DERIVED RESULTS

POWER RULE

Assume that $\alpha \in \mathbb{R} \setminus \{0\}$ and $f(x) = x^{\alpha}$.
Then f is differentiable and

$$f'(x) = \alpha x^{\alpha - 1}$$

wherever $x^{\alpha-1}$ is defined.

POLYNOMIALS

Let $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$.
Then p is always differentiable,
and $p'(x) = a_1 + 2a_2 x + 3a_3 x^2 + \cdots + n a_n x^{n-1}$.

CHAIN RULE

Assume $y = f(x)$ is differentiable at $x = a$
and $z = g(y)$ is differentiable at $y = f(a)$.
Then $h(x) = (g \circ f)(x)$ is also differentiable at $x = a$ and

$$h'(a) = g'(f(a)) \cdot f'(a)$$

* this can also be written as

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}.$$

We can also show

$$L_a^h(x) = L_{f(a)}^g \circ L_a^f(x).$$

"UPGRADED" VERSION OF THE CHAIN RULE

Assume $f: I \to \mathbb{R}$, where $I \subseteq \mathbb{R}$, and $g: J \to \mathbb{R}$, where $f(I) \subseteq J$, and that I & J are open intervals such that I contains some $x = a$ and J contains $f(a)$.
Then, if $f(x)$ is differentiable at $x = a$ and $g(y)$ is differentiable at $y = f(a)$, necessarily $h(x) = (g \circ f)(x)$ is differentiable at $x = a$ with $h'(a) = g'(f(a))(f'(a))$.

Proof. Let $\phi: J \to \mathbb{R}$ be defined by

$$\phi(y) = \begin{cases} \dfrac{g(y) - g(f(a))}{y - f(a)}, & y \neq f(a) \\ g'(f(a)), & y = f(a). \end{cases}$$

Then, since $f(a) \in J$, so

$$\lim_{y \to f(a)} \phi(y) = \lim_{y \to f(a)} \frac{g(y) - g(f(a))}{y - f(a)} = g'(f(a)),$$

and so $\phi(y)$ is continuous at $y = f(a)$.

Next, $\forall y \in J$:

$$g(y) - g(f(a)) = \phi(y)[y - f(a)],$$

even when $y = f(a)$.
Hence $g(f(x)) - g(f(a)) = \phi(f(x))[f(x) - f(a)]$ $\forall x \in I$, since $f(I) \subset J$.

Therefore

$$\begin{aligned} \lim_{x \to a} \frac{g(f(x)) - g(f(a))}{x - a} &= \lim_{x \to a} \frac{\phi(f(x))[f(x) - f(a)]}{x - a} \\ &= \lim_{x \to a} \phi(f(x)) \left[\frac{f(x) - f(a)}{x - a} \right] \\ &= \left(\lim_{x \to a} \phi(f(x)) \right) \cdot \left(\lim_{x \to a} \frac{f(x) - f(a)}{x - a} \right) \\ &= \phi(f(a)) \cdot f'(a) \end{aligned}$$

$$\therefore (g \circ f)'(a) = g'(f(a)) \cdot f'(a). \qquad \blacksquare$$

DERIVATIVES OF INVERSE FUNCTIONS

ALL CONTINUOUS FUNCTIONS THAT ARE 1-1 ARE EITHER INCREASING OR DECREASING

We claim that if f is continuous & 1-1 on $[a,b]$, then f is either increasing or decreasing on $[a,b]$.

Proof. Suppose this was not true.
Then $\exists c,d,e \in [a,b]$ with $c<d<e$ such that either
① $f(c) < f(d)$ & $f(d) > f(e)$; or
② $f(c) > f(d)$ & $f(d) < f(e)$.
w/o loss in generality, assume Case 1 is true.
Then by the IVT, $\exists \gamma \in \mathbb{R}$ such that
$f(c) < \gamma < f(d)$ & $f(e) < \gamma < f(d)$.
Hence $\exists s \in (c,d)$ & $t \in (d,e)$ such that
$f(s) = \gamma = f(t)$ since f is continuous,
but this is impossible if f is 1-1. ※

MONOTONE CONVERGENCE THEOREM FOR FUNCTIONS

Suppose f is increasing on $[a,b]$. Then:

① $\lim_{x\to c^+} f(x)$ exists $\forall c \in [a,b)$, and $\lim_{x\to c^+} = \text{glb}(\{f(x) \mid x \in (c,b]\})$.

② $\lim_{x\to c^-} f(x)$ exists $\forall c \in (a,b]$, and $\lim_{x\to c^-} = \text{lub}(\{f(x) \mid x \in [a,c)\})$.

Proof. We prove ① as the proof for ② is similar.
Let $S = \{f(x) \mid x \in (c,b]\}$.
Then S is bounded below by $f(c)$.
Let $L = \text{glb}(S)$, and let $\varepsilon > 0$ be arbitrary.
Then since $(L+\varepsilon)$ is not a lower bound for S, hence there exists a $d \in (c,b]$ such that
$L \leq f(d) < L+\varepsilon$.
So if $x \in (c,d)$, then
$L \leq f(x) < f(d) < L+\varepsilon$,
implying $\lim_{x\to c^+} f(x) = L$. ∎

CONTINUITY OF MONOTONIC FUNCTIONS

Suppose f is increasing on $[a,b]$. Then f is continuous on $[a,b]$ if and only if

$$f([a,b]) = \{f(x) \mid x \in [a,b]\} = [f(a), f(b)].$$

Proof. Since f is increasing, each $x \in [a,b]$ satisfies
$f(a) \leq f(x) \leq f(b)$.
and so $f(a) < \gamma < f(b)$.
We first prove the forward argument.
Assume f is continuous, and that $f(a) < \gamma < f(b)$.
By the IVT, $\exists c \in (a,b)$ such that $f(c) = \gamma$.
Thus $f([a,b]) = [f(a), f(b)]$, as required. ※
Then, we prove the backward argument.
Assume f is discontinuous at some point $c \in (a,b)$.
Then $\lim_{x\to c^-} f(x) = L < M = \lim_{x\to c^+} f(x)$.
However this implies $[L,M] \cap f([a,b]) = \{f(c)\}$, and so $f([a,b]) \neq [f(a), f(b)]$.
Moreover, if f is discontinuous at $x=a$, then $f(a) < M = \lim_{x\to a^+} f(x)$.
So $(f(a), M) \cap f([a,b]) = \emptyset$.
Similarly, if f is discontinuous at $x=b$, then $\lim_{x\to b^-} f(x) = L < f(b)$.
So $(L, f(b)) \cap f([a,b]) = \emptyset$.
Hence if f is not continuous, then $f([a,b]) \neq [f(a), f(b)]$. ∎

CONTINUITY FOR INVERSE FUNCTIONS

Suppose $f: [a,b] \to \mathbb{R}$ is continuous and 1-1, with $f([a,b]) = [c,d]$.
Let $g: [c,d] \to [a,b]$ be the inverse function of f on $[c,d]$.
Then g is continuous on $[c,d]$.

Proof. Since f is either increasing or decreasing on $[a,b]$, hence g is also either increasing or decreasing.
So, as $g([c,d]) = [a,b]$, thus g is continuous on $[c,d]$.

THE INVERSE FUNCTION THEOREM

Assume $y=f(x)$ is continuous and invertible on $[c,d]$ with inverse $x=g(y)$, and f is differentiable at $a \in (c,d)$.
If $f'(a) \neq 0$, then g is differentiable at $b=f(a)$, and
$$g'(b) = \frac{1}{f'(a)} = \frac{1}{f'(g(b))}.$$
Moreover, L_a^f is also invertible and
$$(L_a^f)^{-1}(x) = L_b^g(x) = L_{f(a)}^g(x).$$

Proof. Let $g(y)$ be the inverse function of $f(x)$.
$$\text{Then } g'(b) = \lim_{y\to b} \frac{g(y)-g(b)}{y-b}$$
$$= \lim_{y\to b} \frac{g(y)-g(b)}{y-f(a)}$$
$$= \lim_{x\to a} \frac{x-a}{f(x)-f(a)} \quad \because g \text{ is continuous at } y=b$$
$$= \lim_{x\to a} \frac{1}{\left(\frac{f(x)-f(a)}{x-a}\right)}$$
$$= \frac{1}{f'(a)}.$$

DERIVATIVES OF INVERSE TRIGONOMETRIC FUNCTIONS

We can use the Inverse Function Theorem to calculate the derivatives of inverse trigonometric functions.

For instance, we can calculate that $\frac{d}{dx}(\arcsin(x)) = \frac{1}{\sqrt{1-x^2}}$.

Proof. $\forall x \in [-1,1]$, if $y = f(x) = \arcsin(x)$ & $x = g(y) = \sin(y)$ with $y \in [-\frac{\pi}{2}, \frac{\pi}{2}]$,
then $g(f(x)) = \sin(\arcsin(x)) = x$.
By the chain rule,
$$g'(f(x))\, f'(x) = 1,$$
and so
$$f'(x) = \frac{1}{g'(f(x))} = \frac{1}{\cos(f(x))} = \frac{1}{\sqrt{1-\sin^2(f(x))}}$$
$$\therefore f'(x) = \frac{1}{\sqrt{1-x^2}}.$$ ∎

IMPLICIT DIFFERENTIATION

We can use implicit differentiation to find derivatives of "relations" that are not written in the form $y = f(x)$.

eg $x^2 + y^2 = 1$.

$\Downarrow$

$\left(\cdot \frac{d}{dx}\right)$ $2x + 2y\frac{dy}{dx} = 0$.

$\therefore \frac{dy}{dx} = \frac{-x}{y}$.

WHEN NOT TO USE IMPLICIT DIFFERENTIATION

We must always ensure the implicitly defined function is defined before we implicitly differentiate.

eg $x^4 + y^4 = -1 - x^2y^2$.

$\Rightarrow \frac{dy}{dx} = \frac{-2xy^2 - 4x^3}{4y^3 + 2x^2y}$.

But LHS ≥ 0 and RHS ≤ -1, so the inequality is never satisfied!

LOCAL EXTREMA

1. A point c is a "local maximum" for a function f if there exists an open interval (a,b) containing c such that $f(x) \leq f(c)$ $\forall x \in (a,b)$.

2. Similarly, a point c is called a "local minimum" for f if there exists an open interval (a,b) containing c such that $f(c) \leq f(x)$ $\forall x \in (a,b)$.

THE LOCAL EXTREMA THEOREM

The Local Extrema Theorem states that for f if c is a local minimum or maximum and $f'(c)$ exists, then $f'(c) = 0$.

CRITICAL POINT

1. A point c in the domain of a function f is called a critical point for f if either
   i) $f'(c) = 0$; or
   ii) $f'(c)$ does not exist.

2. Critical points are generally local extrema.

exception: eg $f(x) = x^3$

$f'(0) = 3(0)^2 = 0$, but $x = 0$ is not a local extrema for f.

RELATED RATES

We can use derivatives to solve "related rate" problems; ie problems involving a mathematical relationship between the respective rates of change between various quantities.

eg It is given that $pV = kT$,
where p = pressure, V = volume, k = constant, T = temp. of the gas.

> Assume that the gas is heated so that the temperature is increasing. Suppose also that the gas is allowed to expand so that pressure remains constant. If at a particular moment the temperature is 348 Kelvin, but is increasing at a rate of 2 Kelvin per second while the volume is increasing at a rate of 0.001 cubic meters per second, what is the volume of the gas?

we know $pV = kT$.
Differentiating both sides w.r.t t, we get that

$$p\frac{dV}{dt} + \frac{dp}{dt}V = k\frac{dT}{dt}.$$

But since p is constant, hence $\frac{dp}{dt} = 0$.

Hence $p\frac{dV}{dt} = k\frac{dT}{dt}$.

Then, as $\frac{dT}{dt} = +2K°$ and $\frac{dV}{dt} = +0.001\,m^3s^{-1}$, we get that $p = 2000k$.

We can substitute this back into the original formula to get that $V \approx 0.174\,m^3$. (using the fact that $T = 348K°$.)

Chapter 7:

The Mean Value Theorem

Assume f is continuous on $[a,b]$ and f is differentiable on (a,b).
Then there always exists a $c \in (a,b)$ such that

$$\underbrace{f'(c)}_{\substack{\text{instantaneous rate of}\\ \text{change at } x=c.}} = \underbrace{\frac{f(b)-f(a)}{b-a}}_{\substack{\text{"average" rate of change}\\ \text{over } (a,b).}}.$$

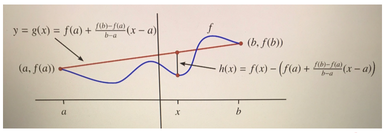
ROLLE'S THEOREM

Assume f is continuous on $[a,b]$, f is differentiable on (a,b), and $f(a) = 0 = f(b)$.
Then there exists a $c \in (a,b)$ such that
$f'(c) = \frac{f(b)-f(a)}{b-a} = 0$.

Proof.
Case ①: $f(x) = 0 \quad \forall x \in [a,b]$.
The claim follows trivially.
Case ②: $\exists x_0 \in [a,b]$ such that $f(x_0) > 0$.
Then necessarily the global maximum occurs at some point $x = c \in (a,b)$ by EVT.
But since c is also a local maximum, it follows that $f'(c) = 0$, and we are done.
Case ③: $\exists x_0 \in [a,b]$ such that $f(x_0) < 0$.
Then necessarily the global minimum occurs at some point $x = c \in (a,b)$ by EVT.
But since c is also a local minimum, hence $f'(c) = 0$, so we are done. ▨

We can use Rolle's theorem to prove the Mean Value Theorem.

Proof. Let $h(x) = f(x) - \left(f(a) + \frac{f(b)-f(a)}{b-a}(x-a)\right)$.
Note that $g(x) = f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$ is linear, and $g(a) = f(a)$ & $g(b) = f(b)$.
So g passes through $(a, f(a))$ and $(b, f(b))$.
But since $h(x) = f(x) - g(x)$, it implies $h(x)$ is the vertical distance between f and the secant line.



Subsequently, since f is continuous on $[a,b]$, differentiable on (a,b) and $h(a) = 0 = h(b)$, by Rolle's theorem there must exist a point c with $c \in (a,b)$ such that $h'(c) = 0$.

But $h'(c) = f'(c) - g'(c)$
$\Rightarrow \quad 0 = f'(c) - \frac{f(b)-f(a)}{b-a}$ — since g is a line with slope $\frac{f(b)-f(a)}{b-a}$

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

APPLICATIONS OF MVT

ANTIDERIVATIVES

Given a function f, an antiderivative is a function F such that $F'(x) = f(x)$.

* if $F'(x) = f(x) \quad \forall x \in I$, we say F is an antiderivative for f on I.

THE CONSTANT FUNCTION THEOREM

Assume $f'(x) = 0 \quad \forall x \in I$.
Then there exists an α such that $f(x) = \alpha \quad \forall x \in I$.

Proof. Let $x_1 \in I$ be arbitrary. Let $\alpha = f(x_1)$.
Then, for any other $x_2 \in I$, we know that (by MVT) there exists a $c \in (a,b)$ such that
$$f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$
But since $f'(c) = 0$, we get that $f(x_1) = f(x_2) = \alpha$.
Thus $f(x) = \alpha \quad \forall x \in I$. ▨

THE ANTIDERIVATIVE THEOREM

Assume that $f'(x) = g'(x) \quad \forall x \in I$.
Then $\exists \alpha \in \mathbb{R}$ such that
$f(x) = g(x) + \alpha \quad \forall x \in I$.

Proof. Let $h(x) = f(x) - g(x)$.
Then $h'(x) = f'(x) - g'(x) = 0 \quad \forall x \in I$,
and so the Constant Function Theorem tells us that $\exists \alpha$ such that $h(x) = \alpha$.
Hence $f(x) - g(x) = \alpha$,
or $f(x) = g(x) + \alpha \quad \forall x \in I$. ▨

LEIBNIZ NOTATION

We denote the "family of antiderivatives" of f by $\underbrace{\int}_{\substack{\text{indefinite}\\ \text{integral}\\ (\text{of } f)}} \underbrace{f(x)}_{\text{the "integrand".}}\,dx$.

COMMON ANTIDERIVATIVES

Here are the antiderivatives of several basic functions:
(we can use differentiation to verify)

① $\int x^\alpha \,dx = \frac{x^{\alpha+1}}{\alpha+1} + C, \quad \alpha \neq -1$

② $\int \frac{1}{x} \,dx = \ln(|x|) + C$

③ $\int a^x \,dx = \frac{a^x}{\ln(a)} + C$

④ $\int \sin(x) \,dx = -\cos(x) + C$

⑤ $\int \cos(x) \,dx = \sin(x) + C$

⑥ $\int \sec^2(x) \,dx = \tan(x) + C$

⑦ $\int \frac{1}{1+x^2} \,dx = \arctan(x) + C$

⑧ $\int \frac{1}{\sqrt{1-x^2}} \,dx = \arcsin(x) + C$

⑨ $\int \frac{-1}{\sqrt{1-x^2}} \,dx = \arccos(x) + C$.

INCREASING/DECREASING FUNCTION THEOREM

Let I be an interval, and $x_1, x_2 \in I$ be arbitrary.
Assume $x_1 < x_2$. Then:

① if $f'(x) > 0 \quad \forall x \in I$, then $f(x_1) < f(x_2)$.
· ie f is increasing on I

② if $f'(x) \geq 0 \quad \forall x \in I$, then $f(x_1) \leq f(x_2)$.
· ie f is non-decreasing on I.

③ if $f'(x) < 0 \quad \forall x \in I$, then $f(x_1) > f(x_2)$.
· ie f is decreasing on I.

④ if $f'(x) \leq 0 \quad \forall x \in I$, then $f(x_1) \geq f(x_2)$.
· ie f is non-increasing on I.

Proof. We prove ①, since the proofs for the others are similar.
Let $x_1, x_2 \in I$ be such that $x_1 < x_2$.
Then, if f is differentiable on I, MVT holds for $[x_1, x_2]$, and so there exists a $c \in (x_1, x_2)$ such that

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(c) > 0.$$

Consequently (since $x_2 - x_1 > 0$) we have that $f(x_2) > f(x_1)$, which we wanted to show. ▣

FUNCTIONS WITH BOUNDED DERIVATIVES

Assume f is continuous on $[a,b]$ and differentiable on (a,b) with

$$m \leq f'(x) \leq M \quad \forall x \in (a,b).$$

Then

$$f(a) + m(x-a) \leq f(x) \leq f(a) + M(x-a) \quad \forall x \in [a,b].$$

Proof. We know $m \leq f'(x) \leq M$.
Let $x \in [a,b]$ be arbitrary.
Then since MVT is true on $[a,x]$, it implies there exists a $c \in [a,x]$ such that

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

Hence $$m \leq \frac{f(x) - f(a)}{x-a} \leq M$$

or $$f(a) + m(x-a) \leq f(x) \leq f(a) + M(x-a).$$ ▣

COMPARING FUNCTIONS USING THEIR DERIVATIVES

Assume f and g are continuous at $x=a$ with $f(a) = g(a)$.
Then:

① if both f and g are differentiable for $x > a$, and $f'(x) \leq g'(x) \quad \forall x > a$, then
$f(x) \leq g(x) \quad \forall x > a$.

② if both f and g are differentiable for $x < a$, and $f'(x) \leq g'(x) \quad \forall x < a$, then
$f(x) \geq g(x) \quad \forall x < a$.

Proof. We prove ①, as the proof for ② is similar.
Let $h(x) = g(x) - f(x)$.
Then h is continuous at $x=a$ and differentiable for $x>a$ with $h'(x) = g'(x) - f'(x) \geq 0 \quad \forall x > a$.
So, by MVT, if $x > a$, it follows that $\exists c \in (a,x)$ such that

$$0 \leq h'(c) = \frac{h(x) - h(a)}{x-a}.$$

But since $h(a) = 0$ and $x - a > 0$, hence $h(x) \geq 0$, implying that $g(x) \geq f(x) \quad \forall x > a$. ▣

THE FUNDAMENTAL EXPONENTIAL LIMIT

1. We can show that $\lim_{n\to\infty} \left(1+\frac{1}{n}\right)^n = e$.

Proof. First, let $f(x) = x - \frac{1}{2}x^2$;
$g(x) = \ln(1+x)$; and
$h(x) = x$.

Then $0 = f(0) = g(0) = h(0)$,
and $f'(x) = 1 - x$,
$g'(x) = \frac{1}{1+x}$ and
$h'(x) = 1$.

So, if $x > 0$, then
$g'(x) < \frac{1}{1+0} = 1 = h'(x)$,
and since $(1-x)(1+x) = 1 - x^2 < 1$,
so $f'(x) = 1 - x < \frac{1}{1+x} = g'(x)$.

Therefore $f'(x) < g'(x) < h'(x) \quad \forall x > 0$, and so by the Increasing Function Theorem

$$f(x) < g(x) < h(x) \quad \forall x > 0.$$

ie $x - \frac{1}{2}x^2 < \ln(1+x) < x \quad \forall x > 0$.

Then, if $x > 0$, we can divide all the terms by x to get that

$$1 - \frac{1}{2}x < \frac{\ln(1+x)}{x} < 1.$$

In particular, if $x = \frac{1}{n}$, then

$$1 - \frac{1}{2n} < n \ln\left(1 + \frac{1}{n}\right) < 1.$$

$$\Rightarrow 1 - \frac{1}{2n} < \ln\left(\left(1+\frac{1}{n}\right)^n\right) < 1.$$

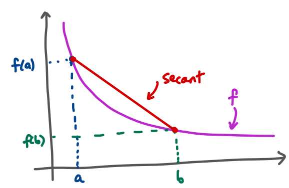
$$\Rightarrow e^{1-\frac{1}{2n}} < \left(1+\frac{1}{n}\right)^n < e.$$

But as $n \to \infty$, $1 - \frac{1}{2n} \to 0$, so $e^{1-\frac{1}{2n}} \to e$,
hence by the Squeeze Theorem,
$\lim_{n\to\infty} \left(1+\frac{1}{n}\right)^n = e$. ▣

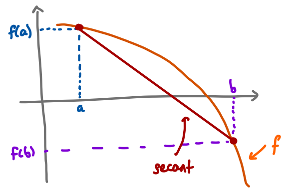
2. We can use a similar proof to show that for any $\alpha \in \mathbb{R}$, $e^\alpha = \lim_{n\to\infty} \left(1 + \frac{\alpha}{n}\right)^n$.

CONCATIVITY

1. We say the graph of f is "concave upwards" on I if $\forall a, b \in I$, the secant line (ie the line segment joining $(a, f(a))$ and $(b, f(b))$) lies above the graph of f.



2. Similarly, we say that the graph of f is "concave downwards" on I if $\forall a, b \in I$, the secant line lies below the graph of f.



SECOND DERIVATIVE TEST FOR CONCAVITY

Recall $f''(x)$ is the second derivative of f.
① If $f''(x) > 0 \quad \forall x \in I$, then f is concave upwards on I.
② If $f''(x) < 0 \quad \forall x \in I$, then f is concave downwards on I.

* f'' is a measure of how "quickly" the slopes are changing.

INFLECTION POINTS

For any function f, a point $(c, f(c))$ is an "inflection point" for f if
① f is continuous at $x=c$; and
② the concavity of f changes at $x=c$.

*note: this usually occurs when $f''(x) = 0$.
But $f''(c) = 0$ does not guarantee that c is an inflection point!
(eg $f = x^4$)

CLASSIFYING CRITICAL POINTS

THE FIRST DERIVATIVE TEST

Assume c is a critical point of f, and f is continuous at c.

① If there exists an interval (a,b) containing c such that
i) $f'(x) < 0 \quad \forall x \in (a,c)$; and
ii) $f'(x) > 0 \quad \forall x \in (c,b)$;
then f has a local minimum at c.

② Similarly, if there exists an interval (a,b) containing c such that
i) $f'(x) > 0 \quad \forall x \in (a,c)$; and
ii) $f'(x) < 0 \quad \forall x \in (c,b)$;
then f has a local maximum at c.

THE SECOND DERIVATIVE TEST

Assume $f'(c) = 0$ and f'' is continuous at $x=c$. Then:
① If $f''(c) < 0$, then f has a local maximum at c; and
② if $f''(c) > 0$, then f has a local minimum at c.

L'HÔPITAL'S RULE

Assume that $f'(x)$ and $g'(x)$ exist near $x=a$, $g'(x) \neq 0$ except possibly at $x=a$, and that $\lim_{x\to a} \frac{f(x)}{g(x)}$ is an indeterminate form of type $\frac{0}{0}$ or $\frac{\infty}{\infty}$.

Then,
$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)},$$
provided the latter limit exists (or is ∞ or $-\infty$).

* this rule also works for one-sided limits and for limits at $\pm\infty$.

* note: LH only works if the form of the limit is $\frac{0}{0}$ or $\frac{\infty}{\infty}$!!

USING LHR TO FIND LIMITS OF OTHER INDETERMINATE FORMS

$0 \cdot \infty$

The indeterminate form "$0\cdot\infty$" usually arises from $h(x) = f(x)\,g(x)$, where $\lim_{x\to a} f(x) = 0$ and $\lim_{x\to a} g(x) = \infty$.

Example: $\lim_{x\to 0^+} x\ln(x)$

Since $\lim_{x\to 0^+} x = 0$ and $\lim_{x\to 0^+} \ln(x) = -\infty$, this is of the form $0\cdot\infty$.

But, observe that $x\ln(x) = \frac{\ln(x)}{\left(\frac{1}{x}\right)}$,

and this would be of the form $\frac{\infty}{\infty}$.

Hence, we can apply L'Hôpital's Rule:
$$\lim_{x\to 0^+} \frac{\ln(x)}{\left(\frac{1}{x}\right)} = \lim_{x\to 0^+} \frac{\frac{d}{dx}(\ln(x))}{\frac{d}{dx}(x^{-1})}$$
$$= \lim_{x\to 0^+} \frac{\left(\frac{1}{x}\right)}{\left(-\frac{1}{x^2}\right)}$$
$$= \lim_{x\to 0^+} \frac{\left(\frac{1}{x}\right)}{\left(\frac{1}{x^2}\right)} \cdot \frac{x}{-1}$$
$$= \lim_{x\to 0^+} (-x) = 0,$$
and so $\lim_{x\to 0^+} x\ln(x) = 0$. ∎

1^∞

Example: $\lim_{x\to\infty} \left(1+\frac{1}{x}\right)^x$.

Note that since $\lim_{x\to\infty}\left(1+\frac{1}{x}\right) = 1$ and $\lim_{x\to\infty} x = \infty$, this limit is of the indeterminate form 1^∞.

We first write the function as follows:
$$\left(1+\frac{1}{x}\right)^x = e^{x\ln\left(1+\frac{1}{x}\right)}.$$

Then, $\lim_{x\to\infty} x\ln\left(1+\frac{1}{x}\right)$ is of the form $0\cdot\infty$, so we can use the same trick as above to convert this into a $\frac{\infty}{\infty}$ form, and so solve it using LHR:
$$\lim_{x\to\infty} \frac{\ln\left(1+\frac{1}{x}\right)}{\left(\frac{1}{x}\right)} = \lim_{x\to\infty} \frac{\frac{d}{dx}\left(\ln\left(1+\frac{1}{x}\right)\right)}{\frac{d}{dx}\left(\frac{1}{x}\right)}$$
$$= \lim_{x\to\infty} \frac{\frac{1}{1+\frac{1}{x}}\left(-\frac{1}{x^2}\right)}{\left(-\frac{1}{x^2}\right)}$$
$$= \lim_{x\to\infty} \frac{1}{1+\frac{1}{x}}$$
$$= 1.$$

Finally, since e^x is continuous, we get that
$$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = \lim_{x\to\infty} e^{x\ln\left(1+\frac{1}{x}\right)} = e^1 = e. \ ∎$$

CAUCHY'S MEAN VALUE THEOREM (CMVT)

Suppose f and g are continuous on $[a,b]$ and differentiable on (a,b) with $g'(x)\neq 0$ $\forall x\in(a,b)$. Then $g(a)\neq g(b)$ and there exists a $c\in(a,b)$ such that
$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}.$$

* if $g(x) = x$, this is the MVT.

Proof. Since $g'(x)\neq 0$, so $g(a)\neq g(b)$.

Let
$$H(x) = \frac{f(b)-f(a)}{g(b)-g(a)}[g(x)-g(a)] - (f(x)-f(a)).$$

Then $H(x)$ is continuous on $[a,b]$ and differentiable on (a,b) with
$$H(a) = \frac{f(b)-f(a)}{g(b)-g(a)}[g(a)-g(a)] - (f(a)-f(a)) = 0$$
and
$$H(b) = \frac{f(b)-f(a)}{g(b)-g(a)}[g(b)-g(a)] - (f(b)-f(a)) = 0.$$

So, by Rolle's Theorem, there exists a $c\in(a,b)$ such that
$$0 = H'(c) = \frac{f(b)-f(a)}{g(b)-g(a)}\,g'(c) - f'(c).$$

It follows that
$$\frac{f'(c)}{g'(c)} = \frac{f(b)-f(a)}{g(b)-g(a)}.$$

CURVE IN $\mathbb{R}^2$

A curve in $\mathbb{R}^2$ is a function $\vec{F}:[a,b]\to\mathbb{R}^2$ given by
$$\vec{F}(t) = (g(t), f(t)),$$
where $g(t)$ and $f(t)$ are called the coordinate functions of $\vec{F}$.

GEOMETRIC INTERPRETATION OF CMVT

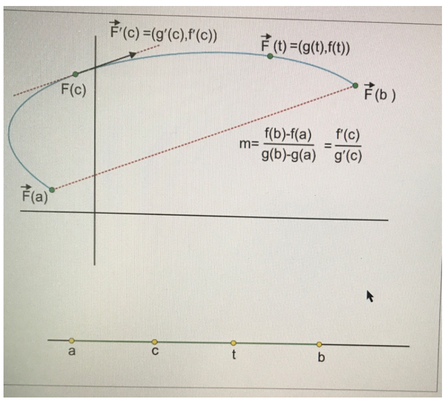
Note that for any curve in $\mathbb{R}^2$ $\vec{F}$,
$$\vec{F}'(t) = (g'(t), f'(t)),$$
and the line in the direction of $\vec{F}'(t)$ through the point $\vec{F}(t)$ is the tangent line to the curve at $\vec{F}(t)$.

Then, the slope of this line is $m = \frac{f'(t)}{g'(t)}$.

Similarly, the slope of the secant line through $\vec{F}(a)$ and $\vec{F}(b)$ is $\frac{f(b)-f(a)}{g(b)-g(a)}$.

So, using CMVT, we can deduce that $\exists c\in(a,b)$ such that $\frac{f'(c)}{g'(c)} = \frac{f(b)-f(a)}{g(b)-g(a)}$;

ie the secant line through $\vec{F}(a)$ and $\vec{F}(b)$ is parallel to the tangent line to the curve through $\vec{F}(c)$.



PROOF OF L'HÔPITAL'S RULE

INDETERMINATE FORMS

1. We call $\mathbb{R}^* = \mathbb{R} \cup \{-\infty, \infty\}$ the set of extended real numbers.

2. Next, suppose $f, g : I \to \mathbb{R}$, where I is an open interval containing some $a \in \mathbb{R}^*$ as an endpoint. Also assume $g(x) \neq 0 \ \forall x \in I$.

Then:

① $\lim_{x\to a^\pm} \frac{f(x)}{g(x)}$ is called an "indeterminate form of type $\frac{0}{0}$" if

$$\lim_{x\to a^\pm} f(x) = 0 = \lim_{x\to a^\pm} g(x); \quad \text{and}$$

② $\lim_{x\to a^\pm} \frac{f(x)}{g(x)}$ is called an "indeterminate form of type $\frac{\infty}{\infty}$" if

$$\lim_{x\to a^\pm} f(x) = \pm\infty \quad \text{and} \quad \lim_{x\to a^\pm} g(x) = \pm\infty.$$

L'HÔPITAL'S RULE FOR $\frac{0}{0}$

Assume $f, g : (a, b) \to \mathbb{R}$, where $a, b \in \mathbb{R}^*$ with $a < b$. Also assume f and g are differentiable on (a, b) and that both $g(x) \neq 0$ & $g'(x) \neq 0 \ \forall x \in (a, b)$.

① Assume $\lim_{x\to a^+} f(x) = 0 = \lim_{x\to a^+} g(x)$. Then,

i) if $\lim_{x\to a^+} \frac{f'(x)}{g'(x)} = L \in \mathbb{R}$, then $\lim_{x\to a^+} \frac{f(x)}{g(x)} = L$.

ii) if $\lim_{x\to a^+} \frac{f'(x)}{g'(x)} = \pm\infty$, then $\lim_{x\to a^+} \frac{f'(x)}{g'(x)} = \pm\infty$.

② Similarly, assume $\lim_{x\to b^-} f(x) = 0 = \lim_{x\to b^-} g(x)$. Then,

i) if $\lim_{x\to b^-} \frac{f'(x)}{g'(x)} = L \in \mathbb{R}$, then $\lim_{x\to b^-} \frac{f(x)}{g(x)} = L$.

ii) if $\lim_{x\to b^-} \frac{f'(x)}{g'(x)} = \pm\infty$, then $\lim_{x\to b^-} \frac{f'(x)}{g'(x)} = \pm\infty$.

Proof. We prove ① i, as the other proofs are similar.

Let $\varepsilon > 0$ Choose $a < \beta$ in I such that if $a < \xi < \beta$, then $\left| \frac{f'(\xi)}{g'(\xi)} - L \right| < \varepsilon$

(We can do this because $\lim_{x\to a^+} \frac{f'(x)}{g'(x)} = L$.)

Next, let $a < x < \beta$ be arbitrary. Let the sequence $\{y_n\}$ be such that $a < y_n < x$ and $y_n \to a$.

Then, by the CMVT, for each $n \in \mathbb{N}$ there exists a point $\xi_n \in (x, y_n)$ such that

$$\left| \frac{f(x) - f(y_n)}{g(x) - g(y_n)} - L \right| = \left| \frac{f'(\xi_n)}{g'(\xi_n)} - L \right|.$$

Since $a < \xi_n < \beta$, it follows that

$$\left| \frac{f(x) - f(y_n)}{g(x) - g(y_n)} - L \right| < \varepsilon.$$

This is true $\forall n \in \mathbb{N}$; but since $\lim_{n\to\infty} f(y_n) = 0 = \lim_{n\to\infty} g(y_n)$, it follows that

$$\lim_{n\to\infty} \left| \frac{f(x) - f(y_n)}{g(x) - g(y_n)} - L \right| = \left| \frac{f(x)}{g(x)} - L \right| \leq \varepsilon$$

and so $\lim_{x\to a^+} \frac{f(x)}{g(x)} = L$. ∎

*There are similar proofs for the cases where the form of the function is $\frac{\infty}{\infty}$.

BASIC CURVE SKETCHING: PART 2

1. We can use derivatives to derive certain characteristics of graphs of functions.

2. Steps:

① Complete steps in Part 1.

② Calculate $f'(x)$.

③ Identify any critical points; ie where $f'(x) = 0$ or f' does not exist.

④ Determine whether f is increasing or decreasing by analysing the sign of $f'(x)$ between critical points.

⑤ Test the critical points to determine if they are local maxima, minima or neither.

⑥ Find $f''(x)$.

⑦ Locate where $f''(x) = 0$ or where $f''(x)$ does not exist. Use these points to divide $\mathbb{R}$ into intervals, and determine the concavity of f by analysing the sign of $f''(x)$ inside these intervals (if possible).

⑧ Find any points of inflection.

⑨ Incorporate the info into the graph.

Chapter 8: Taylor Polynomials and Big-O Notation

TAYLOR POLYNOMIALS

1. Assume that f is n-times differentiable at $x=a$. Then, the "n-th degree Taylor polynomial for f centered at $x=a$" is the polynomial

$$T_{n,a}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

$$= f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

* note: $f^{(k)}(a)$ represents the k^{th} derivative of a. (if $k=0$, $f^{(k)}(a) = f(a)$).

2. Note as $n \to \infty$, $T_{n,a}(x)$ becomes a better approximation to the curve near $x=a$.

3. Also, $\forall k \in \{0, 1, \dots, n\}$: $T_{n,a}^{(k)}(a) = f^{(k)}(a)$, and is the only polynomial in which this is true.

TAYLOR POLYNOMIALS OF COMMON FUNCTIONS

Here are the Taylor polynomials of common functions:

① $f(x) = e^x$: $T_{n,0}(x) = \sum_{k=0}^{n} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!}$

② $f(x) = \sin(x)$: $T_{n,0}(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k \left(\frac{x^{2k+1}}{(2k+1)!}\right) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \dots + (-1)^{\lfloor \frac{n}{2} \rfloor}\left(\frac{x^{2\lfloor \frac{n}{2} \rfloor + 1}}{(2\lfloor \frac{n}{2} \rfloor + 1)!}\right)$

③ $f(x) = \cos(x)$: $T_{n,0}(x) = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \left(\frac{x^{2k}}{(2k)!}\right) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \dots + (-1)^{\lfloor \frac{n}{2} \rfloor}\left(\frac{x^{2\lfloor \frac{n}{2} \rfloor}}{(2\lfloor \frac{n}{2} \rfloor)!}\right)$

TAYLOR'S THEOREM

TAYLOR REMAINDER

1. Assume $f^{(n)}(a)$ exists. Then the "n^{th} degree Taylor remainder function centered at $x=a$" is the function

$$R_{n,a}(x) = f(x) - T_{n,a}(x).$$

2. Then, the error in using the Taylor polynomial to approximate f is given by

$$\text{error} = |R_{n,a}(x)|.$$

TAYLOR'S THEOREM

Assume $f^{(n+1)}(x)$ exists $\forall x \in I$, where I is an interval containing $x=a$.

Let $x \in I$ be arbitrary. Then there exists a $c \in (x, a)$ such that

$$R_{n,a}(x) = f(x) - T_{n,a}(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$$

Proof. Let $x \in I$ such that $x \neq a$.

Then there exists a M such that

$$R_{n,a}(x) = f(x) - T_{n,a}(x) = M(x-a)^{n+1}.$$

Next, let

$$F(t) = f(t) + f'(t)(x-t) + \frac{f''(t)}{2!}(x-t)^2 + \dots + \frac{f^{(n)}(t)}{n!}(x-t)^n + M(x-t)^{n+1}.$$

Notice $F(x) = f(x) = F(a)$. So, by MVT, $\exists c \in (x, a)$ such that $F'(c) = 0$.

Since
$$\frac{d}{dt}\left(\frac{f^{(k)}(t)}{k!}(x-t)^k\right) = \frac{d}{dt}\left[\frac{f^{(k)}(t)}{k!}\right](x-t)^k + \frac{f^k(t)}{k!} \cdot \frac{d}{dt}\left[(x-t)^k\right]$$
$$= \frac{f^{(k+1)}(t)}{k!}(x-t)^k + \frac{f^k(t)}{k!} \cdot k(x-t)^{k-1}(-1)$$
$$= \frac{f^{(k+1)}(t)}{k!}(x-t)^k - \frac{f^k(t)}{(k-1)!}(x-t)^{k-1},$$

it follows that
$$F'(t) = \frac{f^{(n+1)}(t)}{n!}(x-t)^n - M(n+1)(x-t)^n.$$

So
$$0 = F'(c) = \frac{f^{(n+1)}(c)}{n!}(x-c)^n - M(n+1)(x-c)^n$$

and hence $M = \frac{f^{(n+1)}(c)}{(n+1)!}$, as required. ∎

ERROR IN LINEAR APPROXIMATION

1. Note that $T_{1,a}(x) = L_a(x)$, and so $|R_{1,a}(x)|$ shows the error in using the linear approximation.

2. Then, by Taylor's Theorem, it follows that there exists a c such that

$$|R_{1,a}(x)| = \left|\frac{f''(c)}{2}(x-a)^2\right|.$$

* This shows the error in linear approximation depends on the (potential) size of $f''(x)$ and on $|x-a|$.

TAYLOR'S THEOREM IMPLIES MVT

1. Observe that when $n=0$, Taylor's theorem requires f be differentiable on I and its conclusion states $\forall x \in I$, $\exists c \in (x, a)$ such that

$$f(x) - T_{0,a}(x) = f'(c)(x-a).$$

2. But $T_{0,a}(x) = f(a)$, so the above expression simplifies to

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

and so Taylor's Theorem is the MVT if $n=0$.

TAYLOR'S APPROXIMATION THEOREM I

Assume $f^{(k+1)}$ is continuous on $[-d,d]$ for $d>0$.
Then there exists a constant $M>0$ such that
$$|f(x) - T_{k,0}(x)| \leq M|x|^{k+1} \quad \forall x\in[-d,d].$$

Proof. Let $g(x) = \left|\frac{f^{(k+1)}(x)}{(k+1)!}\right|$.

Note that since $f^{(k+1)}$ is continuous, g is also continuous.

Then, by the EVT, g has a maximum on $[-1,1]$.

Thus, there exists a M such that
$$\left|\frac{f^{(k+1)}(x)}{(k+1)!}\right| \leq M \quad \forall x\in[-1,1].$$

Let $x\in[-1,1]$ be arbitrary. Then, by Taylor's Theorem, we know that there exists a c between 0 and c such that
$$|R_{k,0}(x)| = \left|\frac{f^{(k+1)}(c)}{(k+1)!}x^{k+1}\right|.$$

Therefore
$$\begin{aligned}|f(x) - T_{k,0}(x)| &= |R_{k,0}(x)| \\ &= \left|\frac{f^{(k+1)}(c)}{(k+1)!}x^{k+1}\right| \\ &\leq M|x^{k+1}|,\end{aligned}$$
and we are done. ◻

BIG-O

We say f is "Big-O" of g as $x\to a$ if there exists $\varepsilon>0$ and $M>0$ such that $|f(x)| \leq M|g(x)|$ $\forall x\in(a-\varepsilon, a+\varepsilon)$ (except possibly at $x=a$).

In this case, we write
$f(x) = O(g(x))$ (as $x\to a$). optional.

*we assume $0<\varepsilon\leq 1$.

*we say $f(x)$ has "order of magnitude that is less than or equal to that of $g(x)$" near $x\to a$.

$f(x)=O(x^n) \Rightarrow \lim\limits_{x\to 0} f(x) = 0$

Suppose $f(x)=O(x^n)$ for some $n\in\mathbb{N}$.
Then $\lim\limits_{x\to 0} f(x) = 0$.

Proof. By def^n, $-M|x^n| \leq f(x) \leq M|x^n|$ on $(-\varepsilon,\varepsilon)$ except possibly at $x=0$.
Then since $\lim\limits_{x\to 0} -M|x^n| = 0 = \lim\limits_{x\to 0} M|x^n|$ by the Squeeze Theorem we get that $\lim\limits_{x\to 0} f(x) = 0$. ◻

EXTENDED BIG-O NOTATION

Suppose f,g,h are defined on an open interval containing $x=a$, except possibly at $x=a$.
Then, we write
$$f(x) = g(x) + O(h(x)) \quad \text{as } x\to a$$
if
$$f(x) - g(x) = O(h(x)) \quad \text{as } x\to a.$$

*this tells us at values near $x=a$, $f(x)\approx g(x)$ with an error that is an order of magnitude at most that of $h(x)$.

TAYLOR'S APPROXIMATION THEOREM II

Let $r>0$. Assume $f^{(n+1)}(x)$ exists $\forall x\in[-r,r]$ and $f^{(n+1)}$ is continuous on $[-r,r]$.
Then $f(x) = T_{n,0}(x) + O(x^{n+1})$ as $x\to 0$.

Proof. By the EVT, $f^{(n+1)}$ is bounded on $[-r,r]$.
Let M be such that $|f^{(n+1)}(x)| \leq M$ $\forall x\in[-r,r]$.
Then, by Taylor's Theorem, there exists a c between x and 0 such that
$$|f(x) - T_{n,0}(x)| = \left|\frac{f^{(n+1)}(c)}{(n+1)!}x^{n+1}\right| \leq \left|\frac{M}{(n+1)!}x^{n+1}\right| = \frac{M}{(n+1)!}|x^{n+1}|.$$
This shows $f(x) - T_{n,0}(x) = O(x^{n+1})$ as $x\to 0$, and the result of the theorem follows. ◻

ARITHMETIC OF BIG-O

Assume $f(x)=O(x^n)$ and $g(x)=O(x^m)$ as $x\to 0$, for some $m,n\in\mathbb{N}$.
Let $k\in\mathbb{N}$. Then:

① $c(O(x^n)) = O(x^n)$; ie $c\cdot f(x) = O(x^n)$.
② $O(x^n)+O(x^m) = O(x^k)$, where $k=\min\{n,m\}$; ie $f(x)\pm g(x) = O(x^k)$.
③ $O(x^n)O(x^m) = O(x^{n+m})$; ie $f(x)g(x) = O(x^{n+m})$.
④ If $k\leq n$, then $f(x) = O(x^k)$.
⑤ If $k\leq n$, then $\frac{1}{x^k}O(x^n) = O(x^{n-k})$; ie $\frac{1}{x^k}f(x) = O(x^{n-k})$.
⑥ $f(u^k) = O(u^{kn})$; ie we can simply substitute $x=u^k$.

CALCULATING TAYLOR POLYNOMIALS

ORDER OF A POLYNOMIAL IS AT MOST ITS DEGREE

Let p be a polynomial with degree n or less.
Suppose $p(x) = O(x^{n+1})$. Then $p(x)=0$ $\forall x\in\mathbb{R}$.

Proof. Let $Q(n)$ be the statement
"if $p(x)$ is a polynomial with degree n or less, and $p(x) = O(x^{n+1})$, then $p(x)=0$ identically."

First, assume $n=0$.
Then $p(x) = c_0 = O(x)$ for some $c_0\in\mathbb{R}$.
Since $p(x)=O(x)$ and $p(x)$ is continuous, it follows that
$c_0 = \lim\limits_{x\to 0} f(x) = 0$, and so $Q(0)$ holds,
proving that $p(x)=0$.

Next, suppose $Q(k)$ is true for some $k\geq 1$.
Let $p(x) = c_0+c_1x+c_2x^2+\cdots+c_kx^k+c_{k+1}x^{k+1} = O(x^{k+2})$.
Then, once again, $c_0 = \lim\limits_{x\to 0} p(x) = 0$.
So $q(x) = \frac{p(x)}{x} = c_1+c_2x+c_3x^2+\cdots+c_kx^{k-1}+c_{k+1}x^k = O(x^{k+1})$.
It follows from the inductive hypothesis that $q(x)=0$, and so $p(x) = xq(x) = 0$ also, proving the claim for $k+1$.
Hence, by induction, the claim is true $\forall n\in\mathbb{N}\cup\{0\}$. ◻

CHARACTERISATION OF TAYLOR POLYNOMIALS

Let $r>0$ be arbitrary. Assume $f^{(n+1)}(x)$ exists $\forall x\in[-r,r]$ and $f^{(n+1)}$ is continuous on $[-r,r]$.
Then if p is a polynomial of degree n or less with
$$f(x) = p(x) + O(x^{n+1}),$$
then $p(x) = T_{n,0}(x)$.

Proof. By assumption, $f(x)-p(x) = O(x^{n+1})$.
Then, by the Taylor Approximation Theorem II, necessarily $f(x) - T_{n,0}(x) = O(x^{n+1})$.

Hence
$$\begin{aligned}h(x) &= p(x) - T_{n,0}(x) \\ &= [f(x) - T_{n,0}(x)] - [f(x) - p(x)] \\ &= O(x^{n+1}) + O(x^{n+1}) \\ &= O(x^{n+1}).\end{aligned}$$

But since h is a polynomial with degree n or less, it follows that $0 = h(x) = p(x) - T_{n,0}(x)$, and therefore $p(x) = T_{n,0}(x)$. ◻