Lecture 1:

We start with some trivial definitions.

A mathematical statement is a sentence like 1+1=2 that can be true or false. For example 2+2=5 is a statement, just not a true statement.

A proof is a sequence of true statements establishing a conclusion. You've all seen them before. (Well, I guess except for proofs by contradiction, those are a sequence of false statements establishing a conclusion)

We prove things to know they are true and why. When we prove things, we need to be careful to prove the right things. We should not prove "if B then A" instead of "if A then B".

A lemma is a statement we prove first then use for a larger proof.

Notation:

The set of natural numbers {1, 2, 3, ...} is denoted by $\mathbb{N}$

The set of integers {..., -3, -2, -1, 0, 1, 2, 3, ...} is denoted by $\mathbb{Z}$

The set of rational numbers (Numbers that can be written as an element of $\mathbb{N}$ divided by an element of $\mathbb{Z}$) is denoted by $\mathbb{Q}$. However, the ancient greeks realized that not all numbers were of this form. The length of the diagonal of a unit square is $\sqrt{2}$, which is a real number that when you square it you get 2, and this provably cannot be written as $\frac{a}{b}$ where a and b are integers. Apparently, and this may be a legend, someone in the ancient greek times was drowned as a punishment for saying that irrational numbers existed, because all numbers were thought to be rational.

The set of real numbers is denoted by $\mathbb{R}$.

The set of complex numbers is denoted by $\mathbb{C}$.

The square root of 2 is a solution to the equation $x^2 - 2 = 0$. A natural question is whether every real number is the solution to a polynomial with integer coefficients. It turns out that this is not the case – There exist transcendental numbers, which are defined as numbers where this does not hold. We will prove this later.

One principle with proofs is that if we can find a counterexample to a statement, we immediately know that it is false, for example:

CLAIM: If 9 divides n^2 then 9 divides n.

COUNTEREXAMPLE: 9 divides 3^2 but 9 does not divide 3, so the claim is false.

More notation:

If A is a mathematical statement, then we write ¬A to mean "Not A".

If A and B are mathematical statements, we write A∨B for "A or B" and A∧B for "A and B". We write A⇒B for "A implies B", A⇐B for "B implies A", and A⇔B for when A and B imply eachother, which means A and B are equivalent. This is often written as A iff B or A if and only if B.

If A⇒B, then a counterexample is an example of A∧¬B. In plain english, this is saying that a counterexample to A implies B is a situation where A is true and B is not true.

Also, if we want to prove A⇒B we can often prove instead that ¬B⇒¬A. This is because if A implies B, then whenever B is false, it must be that A is also false, since A implies B. Similarly, ¬B⇒¬A implies A⇒B so the 2 notions are indeed equivalent. This is called the contrapositive

Example:

Claim: If 2 divides n^2 then n is even.

Proof: We will prove the contrapositive, ie if n is odd then n^2 is odd.

Since n is odd, there is an integer k such that n=2k+1. Then $n^2 = (2\mathrm{k}+1)^2 = 2(2\mathrm{k}^2+2)+1$ by simple algebra, which is also odd. So done.

More notation:

∀ means “for all”

∃ means “there exists”

s.t. means “such that”.

Recall that for sets, A∪B is the union, ie the set of everything in A or B, A∩B is the intersection, ie all elements in A and B, and a∈A means a is an element of the set A, and A⊆B means A is a subset of B, and A⊂B means A is a proper subset of B (ie A is a subset of B and A does not equal B).

If a<b and a and b are real numbers, The interval (a,b) is all numbers between a and b not including a and b. This is called an open interval. [a,b) is the same interval with a included. (a,b] is the same interval with b included. [a,b] is the same interval with both a and b included, ie a closed interval.

Example:

∀x∈ℝ(∃y∈ℝ s.t. y<x) means that for all real numbers x there exists a real number y such that y is less than x.

You have to be careful about the order of these quantifiers. The statement

∃y∈ℝ(∀x ∈ℝ s.t. y<x) is saying that there exists a real number y that is less than all real numbers x, which is obviously false (if there were such a y, the case when for example x=y-1 causes a contradiction), while the above statement is obviously true (as you can pick for example y=x-1).

Also sets are collections of mathematical objects that are equal if they have the same elements. This is obvious

Lecture 2:

When I say a statement, I need to unpack the hidden meaning. For example, if I say that $x^2 - 5x + 6 = 0$ has solutions x=2 and x=3, I am implicitly saying

1. Those are solutions
2. Those are the only solutions

Also, I will now give a false proof that will give some important lessons.

Claim (which is obviously nonsense): Every positive real number is greater than or equal to 1.

"Proof": Let r be the least positive real. Then either r<1, r>1 or r=1. If r<1, then $0 < r^2 < r$, contradicting the assumption that r is the smallest positive real. Since this is a contradiction, we cannot have r<1.

The reason this is wrong is because:

1. What I have really proven is that the smallest real number does not exist, so we cannot say "Let r be the least positive real".
2. For this reason, it is important that we carefully justify everything.

Now, let's make a truth table for statements A and B:

A	B	A∨B	A∧B	¬A	A⇒B	B⇒A	A⇔B
False	False	False	False	True	True	True	True
False	True	True	False	True	True	False	False
True	False	True	True	False	False	True	False
True	True	True	True	False	True	True	True

Now hopefully columns 3-5 are obvious. In column 6, we have that if A is false, then A implies B is true because A implies B is saying that B is true if A is true but that if A is false it does not matter so the statement is vacuously true. Similar for column 7. Then the last column is saying that A implies B and B implies A.

In fact, now one can check using a truth table like this that indeed "A implies B" is the same as "Not B implies Not A".

Definition: Difference

If A and B are sets, A \ B means the set of all elements of A not contained in B.

Proposition about sets :

1. Union is associative, ie A∪(B∪C)=(A∪B)∪C
2. Intersection is associative, ie A∩ (B∩C)=(A∩B) ∩C
3. Union and intersection are distributive, ie A∪(B∩C)=(A∪B) ∩ (A∪C)

Proof: You can easily verify these by shading in the appropriate areas of a venn diagram, or by checking using a truth table. I'll do #3 as an example.

X in A	X in B	X in C	X in B∩C	X in A∪(B∩C)	X in A∪B	X in A∪C	X in (A∪B)∩(A∪C)
No	No	No	No	No	No	No	No
No	No	Yes	No	No	No	Yes	No
No	Yes	No	No	No	Yes	No	No
No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yes	No	No	No	Yes	Yes	Yes	Yes
Yes	No	Yes	No	Yes	Yes	Yes	Yes
Yes	Yes	No	No	Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

As you can see, columns 5 and 8 match, as required.

Lecture 3:

Two more identities about sets that can be verified from the table

A \ (B ∪ C) = (A \ B) ∩ (A \ C)

$A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$

What this means that when we're in A, (Not in B or C) is the same as (Not in B and Not in C), and that (Not in B and C) is the same as (Either not in B, or not in C) which makes sense.

Also, for statements A and B,

$A \wedge B = \neg(\neg A \vee \neg B)$, and

$A \vee B = \neg(\neg A \wedge \neg B)$

The reason why is the same idea as the two set identites above. These are variations of De Morgan's laws.

Also, for set identites like $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ we can prove it by showing that $A \cap (B \cup C)$ is a subset of $(A \cap B) \cup (A \cap C)$ and that $(A \cap B) \cup (A \cap C)$ is a subset of $A \cap (B \cup C)$.

Example:

If $x \in A \cap (B \cup C)$ then $x \in A$ and $x \in B \cup C$ so either x is in A and B or x is in A and C, which means $x \in (A \cap B) \cup (A \cap C)$.

Conversely, if $x \in (A \cap B) \cup (A \cap C)$ then x is either in A and B or x is in A and C. Either way, X is in both A, and either B or C, so $x \in A \cap (B \cup C)$, so done.

Notation (I know this is boring we'll eventually stop notation and start doing interesting stuff):

If $A_1, A_2, \dots, A_n$ are sets, with n possibly infinite, then we can write

$\bigcap_{r=1}^{n} A_r$ for the intersection of the sets, ie the set of elements in all the A's.

We can also write $\bigcup_{r=1}^{n} A_r$ for the union of the sets, ie the set of elements in any of the A's.

Given an index set I and a collection of sets denoted by A_i we can write $\bigcup_{i \in I} A_i$, and same for intersections, kind of like we do sometimes for sums.

Given sets A and B we can form their **cartesian product** AxB, which is the set of all ordered pairs (a, b) with a in A and b in B.

We could also have sets of ordered triples like (a, b, c). In fact, the set of ordered pairs (x, y) with x and y real numbers corresponds exactly to the standard two dimensional plane.

Definition: The **power set** of a set A is the set of all subsets of A. For example, the power set of the set {1,2} is the set {∅, {1}, {2}, {1,2}}.

Note: We can define subsets of other sets by saying X = {x ∈ A such that something}, but we cannot define X = {x such that something}. There is no universal set of any kind, and we should only construct sets from finite sets or known sets. The reason is to avoid stuff like defining Y = {x st x is a set and $x \notin x$} as this is known as russell's paradox. If $Y \in Y$ then by definition $Y \notin Y$ but if $Y \notin Y$ then by definition $Y \in Y$.

Also a finite set has size n if it has n elements. A set is defined to be finite if its size is a natural number.

A function from a set A to a set B is a rule that assigns each element of A to a single element of B, and yeah we've seen before what a function is. We can also think of a function as a subset of AxB consisting of all the ordered pairs of the form {a, f(a)} for a in A.

We can write a function like $f: \mathbb{R} \to \mathbb{R}: x \mapsto x^2$. This is a function. However, $f: \mathbb{R} \to \mathbb{R}: x \mapsto x^{-1}$ is not a function since 0 is not mapped to anything, and : $\mathbb{R} \to \mathbb{R}: x \mapsto \pm\sqrt{|x|}$ is not a function because each input except for 0 maps to 2 outputs.

Definition: If we have a function that maps elements a in a set A to elements b in a set B, then the image of an element a is f(a). The pre-image of an element b in B is the set of elements x in A such that f(x)=B.

Example: If $f(x) = x^2$ then the image of 2 is 4, but the pre-image of 4 is 2 and -2.

Also, we can define images and pre-images of sets in the way you would expect. For example, if $f(x) = x^2$ then the image of (2,3) is (4,9) and the pre-image of (4,9) is (-3,-2)∪(2, 3). Notice that the pre-image of an image is not the original thing.

Lecture 4:

More examples of images:

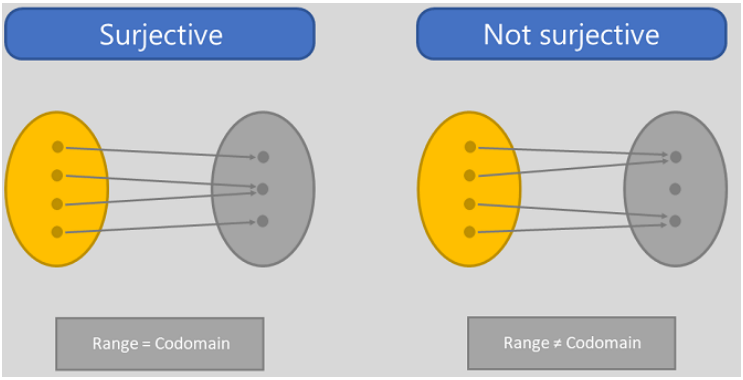
The image of (-1, 4) under the function x^2 is [0, 16) and the pre-image of (-1, 4) under that same function is (-2, 2).

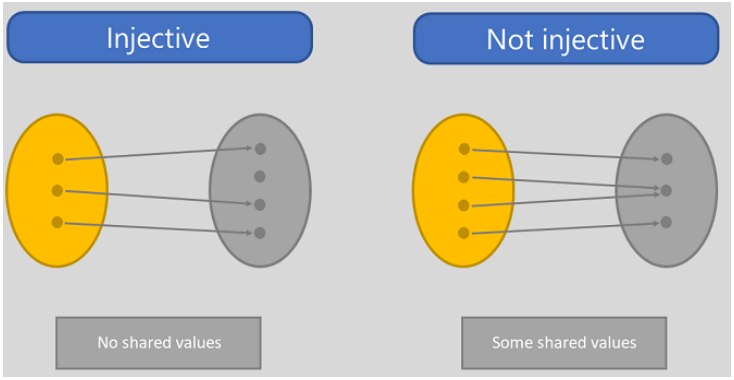
Definition: An **ordered pair** is a set (a, b), or {(a, b), a} with the second element to specify which comes first. We can also have ordered triples, like (a, b, c), or more.

Definition: A function is **injective** if every output that is mapped to is only mapped to by a single input. Equivalently, f is injective if f(b)=f(a) implies b=a. These functions are nice because we can cancel them while doing algebra without running into issues.

Definition: A function is **surjective** if every output in the codomain gets mapped to at least once.

Here are images from the internet to illustrate this:





Definition: A **bijection** is a function which is both injective and surjective. This is a one-to-one mapping where every output is mapped to by exactly one input.

Definition: A **permutation** is a bijection from a set X to itself. This can be thought of as re-ordering the elements of X.

Note that it is important to specify the domain and range of a function. This is often implied, but a question like "is x^2 injective" is meaningless, since the answer is yes in the natural numbers and no in the real numbers.

We denote the size of a set A using |A| or #A. We observe some facts which are obvious for finite sets:

- |A|<|B| if there is no surjection from A to B
- |A|>|B| if there is no injection from A to B
- By definition, |A|=|B| if and only if there is a bijection from A to B, and if the sets are finite then any function from A to B is injective if and only if it is surjective.

Important note: A being a subset of B that is not B itself only implies A is smaller than B if A and B are finite. For example: if A is the even integers and B is the integers, then A is a subset of B, but they are the same size as simply mapping each element of A to the element in B equal to half of that element gives a bijection.

In level 6, in order to show something to do with intervals in probability, we proved that $|\mathbb{R}| > |\mathbb{N}|$ using Cantor's diagonal argument.

A function X to X where every element is mapped to itself is called the identity function.

A sequence can be interpreted as a function $a_n: \mathbb{N} \to \mathbb{R}$.

Addition (+) can be thought of as a function from $\mathbb{N} * \mathbb{N} \to \mathbb{N}$ since it takes in two natural numbers and outputs a third. In general, a binary operation is a function like this on a set X from $X * X \to X$.

Definition: For a set X and a subset Y of X, the **indicator function**, denoted i_A is defined to be 1 if A is in Y and 0 otherwise.

Here are some facts:

i) If the functions i_A and i_B are the same then A and B are equal, and the converse of this is also true.
ii) $i_{A\cap B} = i_A i_B$. This can be verified by going through the different possibilities.
iii) $i_{A\cup B} = i_A + i_B - i_A i_B$. Again one can easily verify this.
iv) $i_{X\setminus A} = 1 - i_A$

Another way to see (iii) is the following:

$$i_{A\cup B} = i_{X\setminus((X\setminus A)\cap(X\setminus B))} = 1 - i_{(X\setminus A)\cap(X\setminus B)} = 1 - i_{X\setminus A} i_{x\setminus B} = 1 - (1 - i_A)(1 - i_B) = i_A + i_B - i_A i_B$$

Lecture 5:

We can have composite functions. This is denoted by $f \circ g$ to mean $f(g(input))$. This is clearly associative: $(f \circ g) \circ h = f\left(g\left(h(input)\right)\right) = f \circ (g \circ h)$

Definition: A function f is **invertible** if there exists a function g such that $f \circ g = g \circ f = identity$, and it is the case that if f goes from A to B then g goes from B to A.

Example: the function x+1 on the natural numbers is not invertible because if we try to make x-1 the inverse, g(1) would have to be 0, as otherwise you couldn't add 1 and get 1, so $f \circ g$ could not be the identity.

If $g \circ f = identity$ then g is called a **left inverse** of f. If this is the case, f must be injective, since if $f(a) = f(b)$ then $g\big(f(a)\big) = g\big(f(b)\big)$ so a=b since $g \circ f = identity$, so we satisfy the injectivity criterion of being able to cancel the f.

Conversely, if f is injective then there is a left inverse. We construct this as follows: For each element in the range of f, send it back. For elements not in the range, pick any element to send it to. One can easily check that this works: $g\big(f(a)\big)$ has a sent to something in f's range then sent back.

If we have a right inverse, ie $f \circ g = identity$ then we can conclude that f must be surjective. If B is the domain of g and this is equal to the image of $f \circ g$, then since $f \circ g = identity$ this implies everything in the domain of B gets hit, and thus everything in the image of $f \circ g$, and thus f gets hit, so f is surjective.

Conversely, if f is surjective, define g to take every element in the range of f back to any element (which we pick) that maps to that element in the range. Doing so will give a right inverse.

Therefore, a function is invertible exactly when it is bijective.

$f^{-1}(A)$ where A is a set is often used as notation for the pre-image of A.

Definition: A **relation** on a set X can be thought of as a subset R of X*X. We write aRb if (a,b) is in R.

Examples of relations on the natural numbers:

1. aRb if a and b have the same last digit
2. aRb if a<b
3. aRb if a does not equal b
4. aRb if a=b=1
5. aRb if |a-b|<3

Definition: A relation is **reflexive** if every element is related to itself (ie aRa for all a).

Definition: A relation is **symmetric** if for every a and b in our set aRb implies bRa.

Definition: A relation is **transitive** if for every a, b, c in our set aRb and bRc implies aRc.

Definition: A relation is an **equivalence relation** if it satisfies all three of the properties above.

Here is a table based on the 5 examples above which you should stare at until you can convince yourself that it is correct:

	Example 1	Example 2	Example 3	Example 4	Example 5
Symmetric	Yes	No	No	No	Yes
Reflexive	Yes	No	Yes	Yes	Yes
Transitive	Yes	Yes	No	Yes	No

Lecture 6:

Theorem: If ~ is an equivelence relation on a set X we have a bunch of pairwise disjoint subsets of X whose union is all of X called equivelence classes such that two elements are related if and only if they are in the same class.

Proof:

Suppose we have such sets. One easily checks that a relation defined as being satisfied if two elements are in the same class is an equivalence relation. Now suppose we have an equivelence relation. Suppose [x] is the set of elements y that are related to x, ie y~x or x~y by the symmetry property. And x~x by reflexivity, so every element is in an equivelence class, and we just have to prove that they are pairwise disjoint. Now suppose there is an element z, then we just have to prove that either [x]=[z] or [x]∩[z] is the empty set. Suppose there is an element y in both [x] and [z], then y is related to both x and z, which by transitivity must be related to eachother, meaning x must be related to z. If an element w is in [z] then w is related to z and thus related to [x] so it is also in [x]. Therefore, [z] is contained in [x] and by the same logic [x] is contained in [z], so if they have non-empty intersection they are the same.

We define the quotient of X by an equivalence relation R as X/R=The set of equivelence classes.

In example 1 from lecture 5, the equivelence classes were the set of natural numbers with a particular last digit.

The map (which just means function) $X \to X/R$ is called the projection map.

Example: On the set $\mathbb{Z} * \mathbb{N}$ we can define an equivelence relation by (a,b)R(c,d) if ad=bc, or equivelently a/b=c/d, which is an equivelence relation since it can be easily seen that the equivelence classes correspond exactly to the rational numbers.

More on binary operations:

A binary operation . is commutative if it is always true that A.B=B.A

It is associative if it is always true that A.(B.C)=(A.B).C

A binary operation . is distributive over a binary operation ~ if it is always true that A.(B~C)=(A.B)~(A.C), ie they interact the same way multiplication interacts with addition.

Now we will show the peano axioms to define the natural numbers:

The set of natural numbers $\mathbb{N}$ has an element which we call 1. It has a function called S that we think of as adding 1 such that S is injective (So we cannot loop back around to a previous element by repeatedly applying S) and every element that is not 1 is S(something in $\mathbb{N}$). Also, so that we ensure we only have the natural numbers (as something like the set $\mathbb{N} \cup (\mathbb{Z} + 0.5)$ satisfies these axioms), we have induction as an axiom: If A is a set with 1 ∈ A and the property that if a ∈ A then s(a) ∈ A then A=$\mathbb{N}$.

We define addition as n+1=S(n) [A1] and n+S(m)=S(n+m) [A2] and multiplication defined as n*1=n and n*S(m)=n*m+n. We define 0 such that a+0=a always. In theory, we can prove all the obvious properties of these operations. Although we do not need to do that in this course as it is ok to do math considering associativity of addition and multiplication as well as the distributive property to all be given, I will prove that (a+b)+c=a+(b+c) always from the Peano axioms as an example, by showing a proof from wikipedia. The proof is in the image below:

For the base case $c = 0$,

$$(a + b) + 0 = a + b = a + (b + 0)$$

Each equation follows by definition [A1]; the first with $a + b$, the second with b.

Now, for the induction. We assume the induction hypothesis, namely we assume that for some natural number c,

$$(a + b) + c = a + (b + c)$$

Then it follows,

$$\begin{aligned}
&(a + b) + S(c) \\
&= S((a + b) + c) \text{ [by A2]} \\
&= S(a + (b + c)) \text{ [by the induction hypothesis]} \\
&= a + S(b + c) \quad \text{[by A2]} \\
&= a + (b + S(c)) \text{ [by A2]}
\end{aligned}$$

In other words, the induction hypothesis holds for $S(c)$. Therefore, the induction on c is complete.

Lecture 7:

Definition: An **Ordering** on a set is a way to compare elements, which satisfies that exactly one of a=b, a<b and a>b holds, and just to make sure it is actuallty an ordering in the way we expect, we have that ordering is transitive (a<b and b<c implies a<c, and similarly for >). An ordering is a **Total Ordering** if any two elements can be compared. An example of a partial ordering is an ordering on sets such that A<B if A is a subset of B and not equal to B, which satisfies the rules above but {1} and {2} cannot be compared.

The axiom of induction given above is called the weak principle of induction (WPI). The strong principle of induction (SPI), which we saw in the level 4 proof for uniqueness of prime factorization, states that if it is the case that {1, 2, 3, … ,n} being a subset of a set A implies that n+1 is in the set A, then all natural numbers are in the set A. We will prove some things that seem useless but are useful because they can apply to other sets with other axioms defined.

This is confusing, as we have a (Statement 1 implies Statement 2) implies (Statement 3) situation.

Theorem: WPI implies SPI

Proof: Apply WPI to the set of numbers n such that {1, 2, 3, … ,n} is a subset of A.

Theorem: SPI implies WPI

Proof: If n is in A implies n+1 is in A then clearly {1, 2, 3, … ,n} being a subset of A also implies n+1 is in A so by strong induction A has all the natural numbers.

Definition: A total ordering on a set is called **well-ordered** if it has the property that any subset has a least element with respect to this ordering.

The well-ordering principle (WOP) states that the usual ordering on $\mathbb{N}$ is a well-ordered. This is obvious, and we will prove it shortly. An example of a total ordering that is not a well ordering is $\mathbb{Z}$ with the usual ordering: The set of, say, even integers does not have a least element: they keep going down forever. However, we can define an ordering on $\mathbb{Z}$ that is a well ordering, where we say the order of the elements is (0, -1, 1, -2, 2, …)

Theorem: SPI implies WOP

Proof: Suppose there is no least element of a subset P of a set where SPI holds. Then consider the set Q of elements that are not in the set: 1 is in Q otherwise 1 would be the minimal element of P. 2 is also in Q or else 2 would be a minimal element of P, and so on. {1, 2, 3, … ,n} being in Q implies n+1 is in Q otherwise that would be the least element of P. So by strong induction, we're done.

WOP does not imply SPI because many "proofs" of this state the principle of strong induction as "n being in a set A if everything <n is in a set A imples everything is in the set A" when the proper statement is that "{1, 2, 3, … ,n} being in a set A impying S(n) is in the set A implies everything in the set A", wherethe problem is 1 may not always be the unique number that is not a successor, this is just the case for the natural numbers. Other "proofs" also assume 1 is the unique number that is not a successor. This statement is not always true when we get onto ordinals – Ordinals are beautiful and one of my favourite things in maths but unfortunately they do not show up until later on in the course so I strongly encourage you to look into them yourself.

Example: We can prove using the well ordering principle that every number can be factored into primes (although we will not prove uniqueness). Let S be the set of counterexamples, then S has a least element m. If m is prime, this is a contradiction, and if not then by definition we can write m as ab, where a and b are less than m and thus can be factored into primes, so we also have a contradiction.

Theorem: A power set of a set with size n has 2^n elements.

Proof: For each of the n elements, we can either put it or not put it in a subset which there are 2 ways to do. Choosing between 2 things to do n times gives a total of 2^n choices.

Corollary: $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n$ because we can interpret LHS as the number of subsets with 0 elements + the number with 1 elements + the number with 2 and so on.

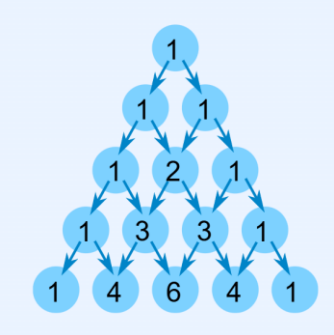
Lecture 8:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

Since we can either choose or not choose the last element.

I think I put this in a previous level but I'm putting it here just in case.

We can obtain Pascal's Triangle by having each element be the sum of the two above it.



Because of this, the number in each circle counts the number of ways to get there, since we can get there from either of the circles above it.

In fact, the k'th number in the n'th row is $\binom{n}{k}$. This is because we have to take k right arrows in order to get there and there are n rows that can happen in and we need to choose k of those rows to take our right turn, if that makes sense.

And the formula $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ as well as the binomial theorem is known from A level. When n is large, this is well approximated by $\frac{n^k}{k!}$, since it can be written as $\frac{n(n-1)(n-2)\ldots(n-k+1)}{k!}$.

Observe that |A∪B|=|A|+|B|-|A∩B|

Lets try to count |A∪B∪C|. If we write this as |A|+|B|+|C| we double count the stuff in the intersections.

If we then write |A∪B∪C|=|A|+|B|+|C|-|A∩B|-|B∩C|-|C∩A|

But now we would not be counting anything in A∩B∩C. The correct expression is

|A|+|B|+|C|-|A∩B|-|B∩C|-|C∩A|+|A∩B∩C|

Theorem (Inclusion-Exclusion principle):

Let $S_1, S_2, S_3, \ldots, S_n$ be finite sets, then

$$|S_1 \cup S_2 \cup S_3 \cup \ldots \cup S_n| = \sum|S_A| - \sum|S_A \cap S_B| + \sum|S_A \cap S_B \cap S_C| - \sum|S_A \cap S_B \cap S_C \cap S_D| + \cdots$$

where the sums range over the unordered tuples (A, B, C, ...)

Equivelently, $\sum i_{S_1 \cup S_2 \cup S_3 \cup \ldots \cup S_n} = \sum i_{S_A} - \sum i_{S_A} i_{S_B} + \sum i_{S_A} i_{S_B} i_{S_C} - \cdots$

Proof:

Let $x \in (S_1 \cup S_2 \cup S_3 \cup \ldots \cup S_n)$. Suppose x lives in exactly k of the sets.

Then the number of sets that x lives in is k (obviously). The number of distinct pairs of sets that x lives in is $\binom{k}{2}$. The number of distinct triples of sets that x lives in is $\binom{k}{3}$, and so on.

So the value of $\sum i_{S_A} - \sum i_{S_A} i_{S_B} + \sum i_{S_A} i_{S_B} i_{S_C} - \cdots$ is $k - \binom{k}{2} + \binom{k}{3} - \binom{k}{4} + \cdots - (-1)^k \binom{k}{k}$

$= 1 - \left(1 - k + \binom{k}{2} - \binom{k}{3} + \binom{k}{4} - \cdots + (-1)^k \binom{k}{k}\right) = 1 - \left(1 + (-1)\right)^k$ by the binomial theorem, which therefore equals 1 as required. The LHS of the indicator function equation is 1 trivially since x is in the union. If x is not in the union both sides are trivially 0.

Notation: We say a|b to mean a divides b, meaning there is an integer c such that ac=b. We can also say that a is a divisor or factor of b, or b is a multiple of a.

Definition: A composite number is a number that is not prime.

Lecture 9:

We saw a few examples of strong induction at the beginning of Level 4, including the statement that every number is the product of primes in a unique way up to ordering. We have seen a proof by contradiction example in A level that there are infinitely many primes.

Proposition: Let n and k be natural numbers, then we can divide n by k and get a remainder less than k, meaning we can write n=qk+r where r is non-negative less than k. This is unique as if we change q r is no longer in that range.

Hopefully you can see that this is true, if not just divide both sides by k, and then use the fact that any real number has an integer part and a fractional part less than 1.

Consider two numbers a and b, then they have a highest common factor (hcf), also called a greatest common divisor (gcd).

Now we will demonstrate Euclid's algorithm by example to show how to find the highest common factor of two numbers.

Example: Suppose we want to find hcf(420, 792), which we will call c.

This procedure will resemble what we did for polynomials in Level 4 to prove partial fraction decomposition.

Clearly, c divides 792 and 420, so it also divides their difference which is 372. Our new number (372) is constructed as the remainder of the bigger number (a or 792) when divided by the smaller number (b or 420). This is equal to r=a-qb by rearranging the equation above, therefore any common factor of a and b also divides r and the converse is true as well. Specifically, c is hcf(420, 372) because if there was a larger common factor of 420 and 372 it would be a common factor of 420 and 792 as well which contradicts c being the highest. At each step when we repeat this, the remainder will be smaller than our smaller number so we will eventually reach 0 in finitely many steps since everything is an integer.

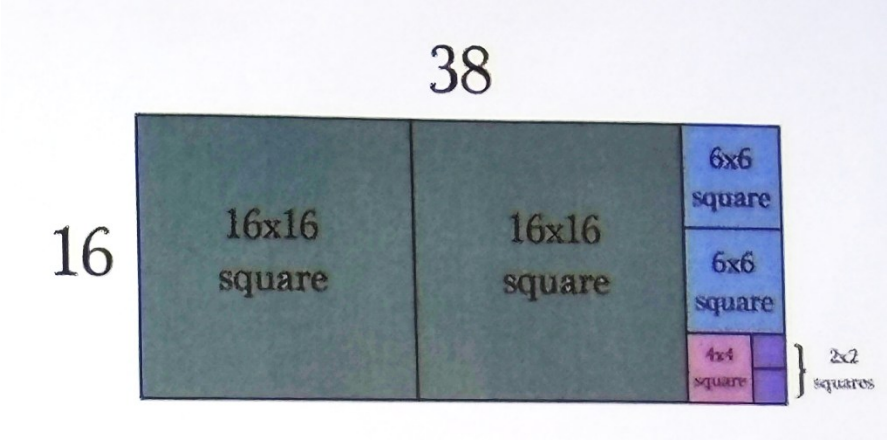
Now we do the same for 372 and 420: We find the difference which is 48, then hcf(372,48)=c. Now we keep subtracting 48 from 372 until we can't anymore, then the final remainder is 36, so we get hcf(36, 48)=c. Now continuing, we get hcf(36, 12)=c then hcf(0, 12)=c, so c=12.

Corollary: Any common factor of a and b divides c:=gcd(a,b)

Proof: By doing Euclid's algorithm, we can see that all common factors of a and b are common factors of 0 and c since we explained above that the common factors stay the same at each step, thus any common factor of a and b is a common factor of 0 and c, and therefore divides c.

Therefore, we can define gcd(a,b) as the unique number dividing a and b such that any common factor of a and b divides gcd(a,b), as it is now clear that this exists and is unique.

Also, here is a visual interpretation of the Euclid algorithm with 38 and 16 as an example.



Lecture 10:

Definition: We say numbers are **coprime** or **relatively prime** if their highest common factor is 1.

Geometric interpretation: if we have an A*B grid rectangle with A and B coprime its diagonal will not intersect any grid corners.

Now we will do something similar to what we did for polynomials in Level 4:

An example of the euclid algorithm on coprime numbers (52 and 87) is as follows:

87 = 1*52 + 35

52 = 1*35 + 17

35 = 2*17 + 1

Now we get to 1. Let's work backwards from 1:

1 = 35 – 2 * 17

1 = (52 – 17) – 2 * (52 – 35)

1 = (52 – (52 – 35) – 2 * (52 – 35)

1 = 52 – (52 – (87 – 52) – 2 * (52 – (87 – 52))

Collecting like terms will give

1 = 3 * 87 – 5 * 52

In general, we can always do a procedure like this: If A and B are coprime we can run Euclid's algorithm and work backwards from 1 to get integers x and y with Ax+By=1. More generally, we can find integers x and y with Ax+By=hcf(A,B). The fact that we can do this is called Bezout's identity. This is exactly what we did for polynomials in Level 4 when we showed that a linear combination of two polynomials can equal their highest common factor. This proves existence of such x and y and also gives a nice way to find them.

Alternative proof of Bezout's identity:

Let h be the least positive integer that can be written as Ax+By for integers x and y. h exists by the well ordering principle. Clearly, all common factors of A and B are a factor of Ax+By and thus h. Now suppose that h does not divide A: Then we can write a=qh+r with 0<r<h and q an integer. So r=A-qh=A-(Ax+By), contradicting the definition of h since h was the smallest positive integer that can be written as Ax+By for integers x and y but r is smaller. Therefore h divides A, and by similar logic h also divides B. Therefore, h is a common factor of A and B and all common factors of A and B divide h, so h is the highest common factor of A and B, so done. This proof tells us that the integers x and y exist but does not give us a way to find them.

Now we can answer questions like: Does there exist x and y with 160x+72y=33. I mean, this one is obviously no because the left hand side and the right hand side is odd, but the point is such integers exist exactly when the number on the right is a multiple of the highest common factor of 160 and 72, or in general the numbers A and B we start with. The reason is if the right is divisible by h=hcf(A,B) we can write it as fh so we can find a solution for when it equals h and multiply it by f. Conversely, if h

does not divide the right hand side, then we can write it as fh+r with r less than h, but then fh+r and fh has a solution, so the difference gives r as a solution which is a contradiction.

Proposition: If p is a prime and p divides ab then p divides a or p divides b.

Proof: Suppose p does not divide a. Then hcf(p,a) is 1 since the only factors of p are 1 and p and p does not divide a. Then there exists integers x and y with px+ay=1, so pbx+aby=b. But p divides ab, so therefore p divides the left hand side, and therefore p divides b.

Corrolary: if $p|a_1a_2a_3 \dots a_n$ and p is prime then by a simple induction, $p|a_i$ for some i from 1 to n.

Corollary: Now we can prove uniqueness of prime factorization in an alternative way from the level 4 way. Suppose $n = p_1p_2 \dots p_k = q_1q_2 \dots q_l$, for primes p and q, and suppose these primes are in order from least to graetest. Then p_1 divides $q_1q_2 \dots q_l$ and thus divides one of the q's by the previous corollary. Thus, one of the q's equals p_1 since p_1 divides it and they are prime. So we can cancel that q and p_1 from both sides, and we can keep doing this and eventually we will be left with 1=1, and so it will become clear that the p's and q's must have been the same at the begininng. If we do this by strong induction, then by cancelling one of the p's we end up using the strong induction hypothesis to deduce uniqueness of the rest.

Lecture 11:

Note: There exist number systems in which factorization is not unique. An example is the set of numbers $x + i\sqrt{3}y$ with x and y integers. In this system $4 = 2 * 2 = (1 + \sqrt{-3})(1 - \sqrt{-3})$. Therefore the theorem above should not be considered "obvious", since 2 and $1 \pm \sqrt{-3}$ are "prime" in this world since they cannot be further factored into integers that are not just 1 or -1.

Example: We can easily list all the factors of $2^3 3^7 11$: They are just all the numbers $2^{0\leq x\leq 3}3^{0\leq y\leq 7}11^{0\leq z\leq 1}$. If there were other factors, this would imply a factorisation of $2^3 3^7 11$ that is different, contradicting uniqueness. In fact, the number of factors of $2^3 3^7 11$ must therefore be (3+1)(7+1)(1+1)=64, since we multiply together the number of choices for each of the powers. This is how we can find the number of factors of any number.

Example: Recall from math GCSE that the common factors and common multiples of numbers a and b can be found by taking the prime factorisation of a and b and taking the minimum and maximum of the powers respectively.

Example: Lcm(a,b)Hcf(a,b)=ab because the power of some p in Lcm(a,b)Hcf(a,b) is the larger of the power of p in a and b plus the smaller of the power of p in a and b, which is exactly ther power of p in ab. Note that every other common multiple of a and b must me a multiple of Lcm(a,b), since for each p, the exponent of p in the prime factorisation must be at least that of the power of p in Lcm(a,b). Alternatively, say for example 10 was the lcm of 2 numbers and so was 45 which is not a multiple of 10, then so would 45-4*10 contradicting the assumption that 10 is the lcm, essentially the remainder would be a common multiple.

Example: (Another proof of infinite primes by paul erdos the GOAT) Suppose there are k primes $p_1, p_2, \dots, p_k$. Now consider $N \coloneqq p_1^{j_1}p_2^{j_2}p_3^{j_3} \dots p_k^{j_k} = m^2p_1^{i_1}p_2^{i_2}p_3^{i_3} \dots p_k^{i_k}$ where each of the i's are 0 or 1. Now we know that $m \leq \sqrt{N}$, so there are at most $\sqrt{N}$ possibilities for m, so there are at most $\sqrt{N}2^k$ numbers of the form $m^2p_1^{i_1}p_2^{i_2}p_3^{i_3} \dots p_k^{i_k}$ that are less than N, so for $N > 4^k$ we contradict the fact that

every number less than N can be written in this way, so pick a number that cannot, then this number must have a prime factor not amongst our finite set of primes.

This is stronger than Euclid's proof as it tells us that the k'th prime must come before 4^k.

Definition: We define $a = bmodn$ if a and b differ by a multiple of n. Alternatively, b is the remainder when a is divided by n.

Examples: $16 = 2mod7, 83497 = 7mod10$

Proposition: a+b mod n = (a mod n) + (b mod n)

Proof Any number which is the sum of an a mod n and b mod n number is (a+xn)+(b+yn)=(a+b)+(x+y)n and therefore is an a+b mod n number.

Proposition: a*b mod n = (a mod n) * (b mod n)

Proof: Similar to above, (a+xn)(b+yn)=ab+(bx+ay+xyn)n.

Example: $2a^2 + 3b^3 = 1$ does not have a solution with a, bin $\mathbb{Z}$. The reason is as follows:

Consider $2a^2 + 3b^3 mod 3$. Then $3b^3 = 0mod3$ so we just have to consider $2a^2 mod3$. If a is a multiple of 3, this is 0 and not 1 which is not possible. If a is 1 mod 3, we have $2(3k + 1)^2 mod3 =$ $18k^2 + 12k + 2mod3 = 3(6k^2 + 4k) + 2\ mod\ 3 = 2\ mod\ 3$ which is again not 1 mod 3. If a is 2 mod 3, a similar rgument shows that $2a^2 + 3b^3 mod3 = 2mod3$ and therefore $2a^2 + 3b^3$ can never be 1mod3 which is a contradiction so $2a^2 + 3b^3 = 1$ must have no integer solutions. Note that if two numbers are the same mod something, we say they are congruent mod that something.

Lecture 12:

Example: Lets solve $7x = 2mod10$. We can multiply both sides by 3 to get $21x = 6mod10$. But then 20x is 0 mod 10 so we can subtract it to get $x = 6mod10$.

Given integers a and b, we say that b is an inverse of a mod n if ab = 1 mod n. We say a is invertible mod n, or a is a unit mod n, if it has an inverse. From the previous example, 3 and 7 are inverses of eachother mod 10. These either come in pairs or a number is its own inverse.

Example: 4 does not have an inverse mod 10, such an inverse would have to be a number that can be multiplied by an even number to get an odd number.

Remark: If a is a unit mod n then its inverse is unique mod n.

Proof: Suppose there exists 2 inverses b and b' such that ab=ab'=1 mod n. We know that b=bab=bab'=b' mod n, since ba=1 mod n.

Corollary: If a is a unit mod n and ab=ac mod n then b=c mod n (multiply by the inverse of a). ie we can cancel units.

This is not true for non-units: 4*3 and 4*8 are the same mod 10 but we cannot divide out the 4 because 3 is not congruent to 8 mod 10.

Proposition: Let p be prime. Then every non-zero residue mod p is a unit mod p.

Proof: If a is coprime to p, then Bezout's identity asserts the existence of integers x and y with ax+yp=1, so x is an inverse to a mod p.

In fact, by the same logic, we deduce that a is a unit mod n if and only if a is coprime to n.

Corollary: If a is coprime to n, then the congruence ax=b mod n has a unique solution x mod n, where x=b*(inverse of a mod n).

Example: Because 365 and 7 are coprime, it means that Christmas can fall on any day of the week. This is because the equation 365x=kmod7 has asolution for any k.

If a and n are not coprime, then if ax=b mod n then this has a solution if and only if b is a multiple of hcf(a,n), since hcf(a,n) must divide ax-b as n divides ax-b. But hcf(a,n) divides a and ax-b and thus must divide b. Conversely, for the same reason as above Bezout's identity asserts the existence of a solution in the case that b is a multiple of hcf(a,n). If hcf(a,n)=d, then the solution to ax=b mod n is exactly the solution to $\frac{a}{d}x = \frac{b}{d} mod \frac{n}{d}$. This is because we can divide everything by d.

Example: Lets solve 7x=4mod30. Hcf(7,30)=1 so 7 has an inverse. It turns out that this inverse is 13, so we get that x=4*13mod30=22mod30.

This is unique: Suppose x' is also a solution with x' not equal to 22 mod 30, then we have that 7x'=4mod30. But then 7x and 7x' are the same mod 30, so multiplying by 13 on both sides we have that x and x' are equal mod 30.

Example: Lets solve 10x=12mod34, then hcf(10,34)=2 so the solutions are the same as the solutions to 5x=6mod17, multiplying by 7 on both sides gives x=8mod17, so this is our solution.

Example: Lets solve the simaltaneous equations x=1mod4 and x=2mod3.

We know that x=1, 5 or 9 mod 12 and that x=2, 5, 8 or 11 mod 12 from each of the above equations. Therefore x=5mod12 is the only possibility.

However, the simaltaneous congruences x=1mod4 and x=2mod6 do not have a solution because x has to be both odd and even.

Theorem (Chinese Remainder Theorem): Let m,n be coprime and a,b be integers. Then the simaltaneous congruences x=a mod m and x=b mod n has a unique solution mod mn.

Proof: We will do this next lecture.

Lecture 13:

Proof of Chinese Remainder Theorem: By Bezout's identity, there exists integers s and t satisfying sm+tn=1. Therefore sm=1 mod n, and tn=1 mod m. Now consider x=a(tn)+b(sm). Then this is congruent to a mod m and b mod n. To prove uniqueness mod mn, suppose y is also a solution to the simaltaneous congruences. Then y-x must be 0 mod m and 0 mod n, and since m and n are coprime, y-x is congruent to 0 mod mn, and therefore x=y mod mn.

This can be extended by induction: If $m_1, m_2, \dots, m_k$ are pairwise coprime, then the simaltaneous congruences $x = a_1 mod m_1 = a_2 mod m_2 = \cdots = a_k mod m_k$ has a unique solution mod $m_1 m_2 \dots m_k$

Definition: $\phi(n)$ is the number of integers from 1 to n-1 coprime to n. Equivelently, this function is counting the number of units mod n. We define $\phi(1) = 1$. Eg, $\phi(9) = 6$ since the units are 1,2,4,5,7,8. When p is a prime, $\phi(p) = p - 1$ since everything from 1 to p-1 is a unit and $\phi(p^2) = p^2 - p$, since there are p non-units.

Powers mod n: $2^1 mod 7 = 2, 2^2 mod 7 = 4, 2^3 mod 7 = 1, 2^4 mod 7 = 2$ and we're back where we started. In general, you can see that if we take powers mod a number, we must eventually get back to somewhere we have been to before, so it will be eventually periodic.

Theorem (Fermat's Little Theorem): Let p be a prime and a not divisible by p. Then $a^{p-1} = 1\ mod\ p$.

Proof: Since a is a unit mod p, $ax = ay\ mod\ p$ if and only if $x = y\ mod\ p$.

Now consider the numbers $a, 2a, 3a, \dots, (p-1)a$. These must be pairwise incongruent mod p: if $ia = ja\ mod\ p$, then (i-j)a is divisible by p, but that is not possible – i-j<p and a is not divisible by p. Therefore, $a, 2a, 3a, \dots, (p-1)a$ mod p must be 1, 2, 3, … ,p-1 in some order.

Therefore, $a * 2a * 3a * \dots * (p-1)a = 1 * 2 * 3 * \dots * p-1\ mod\ p$. Therefore, $a^{p-1}(p-1)! = (p-1)!\ mod\ p$. Since (p-1)! Is not divisible by p, it is a unit and we can multiply through by the inverse of (p-1)! mod p to get the desired result.

Proposition: If p and q are primes, $\phi(pq) = (p-1)(q-1)$.

Proof: mod pq, the ones that are not coprime are the p multiples of q and the q multiples of p, but we have double counted the multiple of pq so we need to add that back. Then we get pq-p-q+1 which is indeed (p-1)(q-1).

Theorem (Fermat-Euler theorem): $a^{\phi(m)} = 1\ mod\ m$ for a coprime to m.

Proof: The same way as in the previous theorem: Let $n_1, n_2, \dots, n_{\phi(m)}$ be the units mod m. Then since a is coprime to m, if we multiply everything by a we still have units mod m. We in fact have distinct units since if $an_j - an_i = 0\ mod\ m$ then $n_j = n_i$ mod m so j must equal i.

$a^{\phi(m)}\left(n_1 n_2 n_3 \dots n_{\phi(m)}\right) = n_1 n_2 n_3 \dots n_{\phi(m)} mod\ m$. Since $n_1 n_2 n_3 \dots n_{\phi(m)}$ is a product of units, we can multiply out by the inverse to get $a^{\phi(m)} = 1\ mod\ m$ as required.

Consider $(p-1)!\ mod\ p$.

Examples:

If p = 3, we want 2 mod 3 which is 2

If p = 5, we want 24 mod 5 which is 4

If p = 7, we want 720 mod 7 which is 6

If p = 11, we want 3628800 mod 11 which is 10

Interestingly, it seems to always be -1 mod p. We will now prove that this is the case:

Proof: Suppose p is an odd prime. 1*2*3*…*(p-1) is a product of units, and the terms come in unit-inverse pairs unless they are self-inverse. We want to determine which of these are self-inverse.

Lemma: If p is prime (This is not necessarily true otherwise, eg for n=8 this is not true), then $x^2 = 1\ mod\ p$ implies x = -1 or 1 mod p. ie, only 1 and p-1 are self inverse units mod p.

Proof of lemma: $x^2 - 1 = 0\ mod\ p$ implies $(x+1)(x-1) = 0\ mod\ p$. This means that in the product (x+1)(x-1), p must appear in the prime factorization of either x+1 or x-1, since if p divides a product ab then p divides a or p divides b. Therefore we get the desired result.

So, in 1*2*3*...*(p-1), all the terms 2, 3, 4, ... , p-2 can be split into unit-inverse pairs which have a product of 1 mod p and therefore may cancel, and so we end up getting just p-1 mod p. This gives the desired result. Example: For p=11, the pairs are 3 and 4, 2 and 6, 7 and 8, 5 and 9.

Lecture 14:

We will now investigate when a number squares to give -1 mod p.

Example: When p=5, 2 squares to 4 which is -1 mod 5. When p=7, you cannot square a number and get 6 mod 7 – You can see this by trying all the possibilities. When p=13, 5 squares to 25 which is -1 mod 13. No luck when p=19.

Proposition: If p is an odd prime, then -1 is a square mod p if and only if p is congruent to 1 mod 4.

Proof: Suppose p = 1 mod 4. Then by Wilson's theorem, I know that -1 is congruent to (p-1)! mod p. We can write (p-1)! as $(1)(2)(3)\dots\left(\frac{p-1}{2}\right)\left(\frac{p+1}{2}\right)\dots(p-2)(p-1)$. Lets subtract p from all terms in the second half of the expression. You can see that this will not change the value mod p. $(1)(2)(3)\dots\left(\frac{p-1}{2}\right)\left(-\left(\frac{p-1}{2}\right)\right)\dots(-2)(-1)$. But since p = 1 mod 4, there are an even number of these minus signs so we can cancel them. Therefore we have $\left(\left(\frac{p-1}{2}\right)!\right)^2$ mod p is -1. So we have constructed such a number.

Example: For p=29, $x = 14!$ Is a number that squares to -1 mod 29.

Conversely, suppose p=3 mod 4, or equivelently that p=(-1 mod 4). Suppose there exists z such that $z^2 = -1\ mod\ 4$. Now lets add p-3, which is divisible by 4, to the exponent. Since $z^{4k} = ((z^2)^2)^k = (-1)^{2k} = 1\ mod\ p$, we know that $z^{p-1} = z^{4\left(\frac{p-3}{4}\right)+2} = -1\ mod\ p$, contradicting Fermat's little theorem.

Note: Fermat's little theorem says that $2^p = 2\ mod\ p$. Conversely, if we find that $2^n mod\ n = 2$, then there is a good chance that n is prime. However, it is not always the case. However, under 1000, the only composite numbers which pass the test are 341, 561 and 645.

We used 2 as the base, however we could use other bases like 3 or 5 to try to weed out more candidates. But it turns out that some numbers like 561 and 1729 are always a false positive, and these are called Carmichael numbers. We won't prove this, it's just a cool thing.

Let us agree to take a message and convert it to a number somehow. Perhaps by using letters base 26. Then we will use the RSA scheme to encrypt the message in such a way that only someone who has a certain key to decrypt it. Here is how to do that:

First, we think of two very large primes. There are ways to test for large primes, but that is beyond this course. Fun fact: A few years ago I found $260370566^{65536} + 1$ is prime using a computer search. As of this lecture, it is the 9358th largest known prime.

Then let n be the product of these primes so n=pq. We will pick an exponent e which should be coprime to $\phi(n)$. We will publish the pair (n, e), but it is very hard to factor n if you don't know the factors. We can split our message into pieces M which are less than n in such a way that each M is coprime to n. We can send the message $M^e mod\ n$, which we can compute quickly by repeated squaring mod n. Since we can reduce mod n, the numbers don't get too big.

To decrypt this, we need to work out the decoding exponent d. By fermat euler, we want that $ed = 1\ mod\ \phi(n)$. Since e was coprime to $\phi(n)$ we can find this using Euclid's algorithm. Then we can use this number d we can find $M^{e*d} mod\ n = M^{1+k\phi(n)} mod\ n$, but by the Fermat-Euler theorem this is just M mod n. Note that in order to decrypt in this way, we need to know $\phi(n)$ which involves knowing how to factor n. But when n is large it can take too long for a computer to factor in reasonable time.

Recall how we constructed the natural numbers using the Peano axioms. We can obtain the integers from the natural numbers by allowing for subtraction. Formally, we could view $\mathbb{Z}$ as the set of equivelence classes of $\mathbb{N}x\mathbb{N}$ where (a,b)R(c,d) if a+d=b+c. We can think of (a,b) as a-b. We can construct 0 for (1,1) and -a for (1, 1+a). We can define the rules multiplication and addition by (a,b)+(c,d)=(a+c,b+d) and (a,b)*(c,d)=(ac+bd,bc+ad). We will not check that this works, you can do that if you want but we all know how arithmetic works in the integers. We can view $\mathbb{Q}$ as the set of equivelence relations in $\mathbb{Z}$x$\mathbb{N}$ (a,b)R(c,d) if and only if ad=bc, and we think of (a,b) as (a/b). We can define multiplication, addition, subtraction and division on this using formulas as above.

We can also define an ordering on the rationals in the way we expect. (ie, depending on whether ad<bc). Next lecture we will construct the reals – This construction and its properties is fundamental for all of real analysis, which we have done in previous levels and will do next level.

Lecture 15:

Note that for any 2 rational numbers, we can find one between them. We can do this by taking their average, ie $\frac{p+q}{2}$. The term for this is that the rational numbers are densely ordered. However, this is not enough due to the existence of irrational numbers, for example we know from A level that $\sqrt{2}$ cannot be a rational number.

Note: if $\sqrt{x}$ is rational and x is an integer then it must be a square number because then we get $x = \frac{a^2}{b^2}$ so $a^2 = xb^2$ so x must be square by uniqueness of prime factorization. Similarly, if $\sqrt[n]{x}$ is rational and x is an integer, x must be a perfect power of n, and in fact $\sqrt[n]{x}$ must also be an integer.

The idea is that $\mathbb{Q}$ has gaps which we need to fill in. Let A be the set of positive rationals p such that $p^2 < 2$. Then we see that A contains no largest element and no least upper bound/supremum. The reason is that for any rational number, if we suppose it is the least upper bound, we can find one closer to $\sqrt{2}$, which there are many ways to do, giving us a contradiction. Previously, we have taken the fact that every bounded set has a least upper bound as obvious, but the idea is that we can construct the reals by saying "I have decided that I want the least upper bound of the set A to exist.", and then the least upper bound to any bounded set of the reals will exist by construction. This is why I have said it is an axiom.

So here is how we construct the reals: We start with 0 and 1 and have an operation + and * and an ordering < satisfying the following conditions:

1. + is commutative and associative and adding 0 does nothing
2. * is commutative and associative and multiplying by 1 does nothing
3. It is always true that a*(b+c)=a*b+a*c
4. For every real number, we have -x and if x is not 0, its reciprocal is real
5. Adding or multiplying 2 real numbers gives another real number

6. For any a and b, we have exactly one of a=b, a<b or a>b, and this ordering is transitive and antisymmetric.
7. For any c, a<b implies a+c<b+c and, if c>0, a<b implies ac<bc

Now this gives us all the rational numbers – We can make integers by adding and make rationals by taking reciprocals and adding them together. In fact, the ordering above is unique. This is not obvious, but lets get an idea for why.

Suppose $1 < 0$, then $0 < -1$ since we can add to both sides, which means that -1 is positive, so we can multiply both sides of $1 < 0$ by -1 to get $-1 < 0$ which is a contradiction. So we know that $1 > 0$.

Now suppose $\frac{1}{3} < 0$. Then $\frac{2}{3} < \frac{1}{3}$ and thus $1 < \frac{2}{3}$ by additivity. This contradicts that $1 > 0$. Also, again by additivity, we know that $0 < 1 < 2 < 3 < \cdots$. The point is for any rational numbers we can argue from these axioms that the ordering is as we expect.

Now we get to the most important property:

8. For any non-empty subset of the reals that is boudned above (ie, there exists x greater than or equal to everything in the set), the least upper bound is in the set (ie, any upper bound is greater than or equal to x). This is important because we can define, say, the set of rationals less than pi, or the square root of 2, or whatever, then just take the least upper bound.

Note that non-empty is important or the least upper bound would be negative infinity. And bounded above is important since otherwise the least upper bound would be positive infinity.

Now given the real numbers and the ordering, we can add the square root of -1 and its multiples and its sum with the real numbers to get the complex numbers. A decimal expansion is, for example, we can define 3.14159265... as the least upper bound of {3, 3.1, 3.14, 3.141, 3.1415, ...}.

We can see that the integers (as defined by the peano axioms) and rational numbers are in the real numbers. For anything, we can add 1 to get its successor, and everything is the successor of some element as we can subtract 1. By repeatedly adding 1 we do not just get back to where we started because otherwise we would not be able to have the unique ordering as defined above. This is the essence of why we indeed have the integers and the rational numbers contained in the reals.

Lecture 16:

Example: The least upper bound of both (0, 1) and [0, 1] is 1. In both cases, 1 is an upper bound (everything in the sets is less than or equal to 1) and any smaller number is not an upper bound (since there is a number above it in the set – eg 0.99999 is not an upper bound because of 0.999995). Note that the least upper bound is sometimes an element of the set itself and sometimes not, and sets like (0, 1) does not even have a largest element, perhaps contrary to intuition. However, for non empty bounded sets the supremum always exists: recall that for a set S we can write sup(S) to mean the least upper bound of S, since we had to go all the way through all this and beyond just to prove that A level stats works. Recall? (Have I said this yet?) that (a, b) is called the open interval a to b and [a, b] is called the closed interval a to b. In future analysis courses we will see that open and closed intervals have surprisingly different properties – we have seen this a bit already with uniform continuity.

The set $\{0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \dots\}$ also has a least upper bound of 1: Numbers of the form $1 - \frac{1}{n}$ get arbitrarily close to 1 but never reach it. We will do some things and then come back to this example.

Proposition: The natural numbers are not bounded above. This is obvious, but we will prove it from the definitions above.

Proof: Suppose on the contrary that the natural numbers are bounded above. Let c=sup($\mathbb{N}$). By definition, c is a least upper bound, so c-1 is not an upper bound for $\mathbb{N}$. This means there is a natural number greater than c-1. But then the successor of that number is greater than c, which is a contradiction. So done. We see that the construction of the real numbers is kind of genius in the sense that we need all of these obvious things to actually hold.

Corollary: For any real number t>0, there exists a natural number n such that 1/n<t. This and the proposition above is known as the archimedean property.

Proof: $\frac{1}{t}$ is real since we have defined it to be the case that every non-zero number has a real reciprocal. This is not an upper bound for the natural numbers, so pick a natural number n greater than it. Then consider $\frac{1}{n}$ which must be less than t. If we're being pedantic, $\frac{1}{n} < t$ since $\frac{nt^{-1}}{n} < \frac{nt^{-1}}{t}$ since $t^{-1} < n.\, nt^{-1}$ is positive because we can multiply it by t which is positive and then get a positive number n.

We have of course seen what it means for a sequence to converge, since we have had to do this to justify some A level maths things, but we will come back to this.

Remark: A set S is bounded below if -S is bounded above. This is obvious. And if this happens S must have a highest lower bound (minus the least upper bound of -S). We can denote a highest lower bound of S by inf(S) (I might have shown this before I'm not sure).

Corollary: inf(1/n)=0 because by the corollary above because 0 is a lower bound and if t>0 is a lower bound this is not possible since we can find something smaller than it. Therefore the set above which we said we would come back to indeed has 1 as its supremum.

Note: We think the real numbers go on forever, but in fact there are no infinitely large or infinitely small numbers that are real. This is also counterintuitive but it is very important to get used to these counterintuitive ideas.

Note: $\sqrt{2}$ actually exists. The set of rational numbers which squares to less than 2 has a least upper bound. What is this least upper bound squared? We need to prove that it is actually 2. Of course, this number is c exists and is between 1 and 2.

Suppose $c^2 < 2$. For $0 < t < 1$, consider $(c+t)^2 = c^2 + 2ct + t^2 < c^2 + 4t + t$ $(since\ t < 1, c < 2)$

$= c^2 + 5t$. In fact, if t is sufficiently small (eg, $t = \frac{2-c^2}{10}$), then c is no longer an upper bound for our set.

Similarly if $c^2 > 2$, then consider $(c-t)^2 = c^2 - 2ct + t^2 < c^2 - 4t + t^2 < c^2 - 3t$, so if we pick $t = \frac{c^2-2}{6}$, then we have an upper bound $c - t$ smaller than c, which is again a contradiction.

Note: A similar proof shows that many other gaps in the real numbers are filled.

Note that one can easily see that a number is rational if and only if adding it or multiplying it by any rational number preserves the fact that it is rational. We can see this since adding or multiplying integer fractions gives another integer fraction.

Lecture 17:

Proposition: The rational numbers are dense in the reals. This means they are everywhere, ie between any 2 numbers we can find a rational number. Again, this should be obvious to you, but we will prove it.

Proof: Let a to b be an interval which we can shift by an integer until a and b are positive. Then by the archimedean property there exists an n with $\frac{1}{n} < b - a$. Now let T be the set of natural numbers k such that $\frac{k}{n} \geq b$. This is non-empty because there exists a k with $k \geq bn$. T is a non-empty set of natural numbers so by the well ordering principle it has a minimal element m. Now consider $\frac{m-1}{n}$. Then we know that $\frac{m-1}{n} < b$ since m-1 is less than the least k such that $\frac{k}{n} \geq b$ by construction. So we just need to show that $\frac{m-1}{n} > a$. Suppose that $\frac{m-1}{n} \leq a$, then $\frac{m}{n} = \frac{m-1}{n} + \frac{1}{n} \leq a + \frac{1}{n} < b$ since we picked n such that $\frac{1}{n} < b - a$, but m is the least element such that $\frac{m}{n} \geq b$, so this is a contradiction.

The irrationals are also dense in this sense. Proof: Between any two numbers we can find a rational number, and we can find a second rational number between that one and one of our endpoints. Let these rational numbers be p and q. Then $p + \frac{1}{\sqrt{2}}(q - p)$ is between p and q so certainly between our original two numbers, but it is irrational.

Now we will talk about sequences, which are ennumerated collections of real numbers, or formally we can consider them to be functions from the natural numbers to the real numbers. Sequences are often written as $a_1, a_2, a_3, \ldots$, or we can write the whole sequence as (a_n)

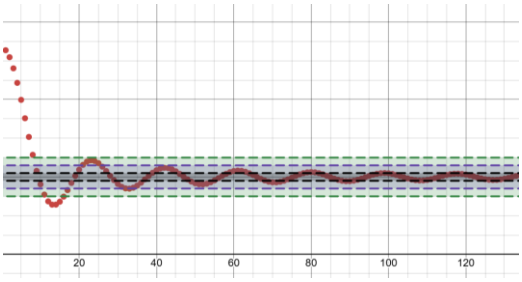
Recall that a sequence tends to a limit if it eventually gets arbitrarily close to some value and stays that close. The definition is elegant but it took mathematicians a long time to come up with a definition that works.

For example (as we saw in previous levels), the sequence 0, 0.9, 0.99, 0.999, … does indeed get closer to 35, but we don't want to say 35 is the limit because it does not get arbitrarily close to 35, we clearly want the limit to be 1.

And we do not want things to just get arbitrarily close, we want them to stay that close. If we have a sequence like 1, 2, 1, 1, 2, 1, 1, 1, 2, 1, 1, 1, 1, 2, … does not stay arbitrarily close to 1. So we want to make it precise that the sequence gets to and stays arbitrarily close to a limit.

A sequence converges if for any tolerance ε>0, no matter how small, we have that there is an N such that whenever n>N, there exists a limit L such that $L - \varepsilon < a_n < L + \varepsilon$. Then we say that the sequence tends to the limit L. Or we can write that $|a_n - L| < \varepsilon$.

We want to say that if we put a small buffer zone around the limit, the terms in the sequence will eventually go to that buffer zone. As an example, here is what that looks like, we eventually stay in any of the tolerance bands:



And recall that we want to think of $|a-b|$ as the distance between a and b on our number line or on the complex plane. And recall the triangle inequality: $|a-c| \leq |a-b| + |b-c|$ which says that the distance from A to C is no more than the distance from A to B plus the distance from B to C. So an important trick we will use a lot is $|a-c| = |a-b+b-c| \leq |a-b| + |b-c|$.

A sequence that does not converge diverges. A sequence goes to infinity if $\frac{1}{a_n} \to 0$. A sequence is bounded if there exists an M>0 such that $|a_n| \leq M$ always, but bounded does not imply convergent. However convergent does imply bounded: If we fix any ε then after a_k we are at most $L+\varepsilon$ then the sequence is bounded after a_k so the whole thing is bounded by $\max(a_1, a_2, \dots, a_k)$.

Examples:

$$0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots \to 1$$

Proof: The difference between a term and 1 is $\frac{1}{n}$, but by the archimedean property for any ε we can pick an n such that $\frac{1}{n} < \varepsilon$. So done.

Example:

$$0, \frac{1}{2}, 0, \frac{1}{4}, 0, \frac{1}{6}, 0, \dots \to 0$$

Proof: Given $\varepsilon > 0$, pick $n > \frac{1}{\varepsilon}$ by the archimedean property, then for any k>n, $|a_k - 0|$ is either 0 or $\frac{1}{k}$, so less than ε, so done.

Example:

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots \to 0$$

Proof: Given $\varepsilon > 0$, pick $n > \frac{1}{\varepsilon}$ by the archimedean property. $|a_n| = \frac{1}{2^n} < \frac{1}{n} < \varepsilon$ again by the archimedian property.

Proof of the inequality $2^n > n$ used above: For natural numbers 2^n is the size of the power set of {1, 2,…,n} which contains more than n elements since it contains {1}, {2}, … ,{n} plus some more.

Example: $(-1)^n = -1, 1, -1, 1, \dots$ does not converge to any limit. This is because for a band of width less than -2, we cannot possibly make the sequence stay there. For all $\varepsilon < 1$, this is never going to work: We can never have it be such that both -1 and 1 are both within 0.5 of something.

We can write the definition of convergence as $\forall \varepsilon > 0\ \exists N \in \mathbb{N}, L \in \mathbb{R} : k > n \Rightarrow |a_k - L| < \varepsilon$. Its negation says there exists an ε such that this is not true, therefore we can write $\exists \varepsilon > 0\ \forall N \in \mathbb{N}, L \in \mathbb{R} : k > N \not\Rightarrow |a_k - L| < \varepsilon$ to mean a sequence does not converge. For $k > N \not\Rightarrow |a_k - L| < \varepsilon$ we can write that there is always a term where $|a_{>k} - L| > \varepsilon$ but I'm too lazy to also do that

Lecture 18:

Notice how this definition of a limit is more rigorous and precise than the hand wavy "infinitesimal" that past mathematicians like Newton used to use.

Theorem: Limits are unique. This is fairly straight forward and intuitive enough that in the past we have used it. But we will prove this from the definition. Let it be such that a and b are both limits of a sequence, then for large enough n, $|a_n - a| < \varepsilon, |a_n - b| < \varepsilon$. Idea: We cannot stay simaltaneously within two disjoint bands. We want to set $\varepsilon < \frac{|b-a|}{2}$, then we have $|a_n - a| + |a_n - b| \geq |b - a| > 2\varepsilon$ by the triangle inequality so we cannot have both $|a_n - a| < \varepsilon, |a_n - b| < \varepsilon$. So done.

Definition: A sequence is eventually bounded if there exists an M>0 such that for all n>N, $|a_n| < M$.

Proposition: Every eventually bounded sequence is bounded.

Proof: It is bounded by $Max(|a_1|, |a_2|, |a_3|, \dots, |a_N|, M)$.

Every convergent sequence is bounded as we proven in my notes from last lecture.

Any unbounded sequence does not necessarily tend to infinity in the limit: A counterexample is a sequence like 0, 1, 0, 2, 0, 3, 0, 4, 0, 5, …

Definition: A sequence is monotonic if it is not strictly decreasing or not strictly increasing.

Theorem (monotone convergence theorem **for sequences** – the monotone convergence theorem for integrals is different): Every bounded monotonic sequence converges. See level 4 for the proof (It is in the existence and uniqueness of e proof: we can add the detail to show it from the definition that if l is the least upper bound of a non-decreasing sequence then fix ε, for any $l - \varepsilon$ we have a greater term since it is the least upper bound and all our terms are less than $l + \varepsilon$ since that's also an upper bound, and similarly for non-increasing).

Remark: Boundedness is necessary: 1, 2, 3, 4, … does not converge. We need a bound so the sequence has something to converge to. We even know that convergence implies boundedness in general.

We know now that the least upper bound property implies the monotone convergence theorem. Theorem: The MCT implies the least upper bound property so we can use it as a definition in the reals since they are equivalent, and we will prove this and come back to this in the analysis course.

Here is another "obvious" theorem: If $a_n \leq d$ for all d and $a_n \to c$ then $c \leq d$. (It is not necessarily the case that if $a_n < d$ and $a_n \to c$ then $c < d$: consider -1/n with d=0)

Idea: If c>d we will have to get into a band above d which is not possible.

Proof: Let $\varepsilon < |d - c|$. Then $|a_n - c| \leq |d - c|$ implies $a_n > d$ so we have a problem.

Proposition: The limit of the sum of convergent sequences is the sum of the limit

Proof: If $a_n \to a, b_n \to b$ then eventually $|a_n - a|, |b_n - b| < \frac{\varepsilon}{2}$, so by the triangle inequality $|a_n + b_n - a - b| < \varepsilon$.

We could have put ε instead of $\frac{\varepsilon}{2}$ first and got 2ε at the end, but that is ok because it does not change the spirit of the convergence: it preserves the "arbitrarily small" idea.

Lecture 19:

As mentioned, an infinite sum is defined as the limit of the partial sums. Ie,

$\sum_{n=1}^{\infty} f(n) = \lim_{r\to\infty} \sum_{n=1}^{\infty} f(r)$ when it exists.

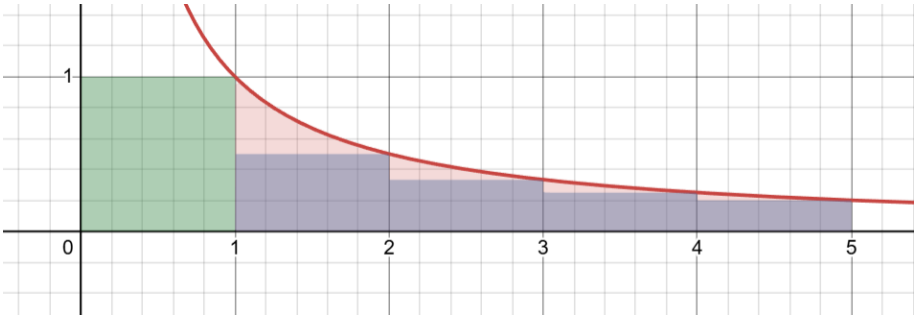
We review the geometric series as an example and reproduce the proof that it converges if and only if the common ratio is less than 1 in absolute value (cf level 4).

Proposition: $\sum_{n=1}^{\infty} \frac{1}{n}$ is not finite. This may seem surprising – The terms go to 0, and even if you add 1 million terms you only get about 14 as the sum, so it is reasonable to think that this converges. However, we will see an argument for why it does not converge, and why it diverges so slowly.

Proof: $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \cdots > 1 + \frac{1}{2} + 2\left(\frac{1}{4}\right) + 4\left(\frac{1}{8}\right) + 8\left(\frac{1}{16}\right) + \cdots = 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \cdots$ which clearly diverges off to infinity. So done.

Proposition: $\ln(x) < \sum_{n=1}^{x} \frac{1}{n} < 1 + \ln(x)$

Proof: If you're being pedantic you might say we need to define the logarithm. But don't worry: From all the previous levels we know well what the logarithm is. The right hand side of the inequality was done in the Level 6 pure mathematics document during a radius of convergence argument for proving series solutions to differential equations work. The same argument can be adapted to get the left hand side of the inequality. We will again use x=5 as an example, recalling that $\ln(x) = \int_1^x \frac{1}{t} dt$.



Red < Green: Stack them on top of eachother. Therefore Red + Purple < Green + Purple. But then this inequality is exactly saying $\ln(x) < \sum_{n=1}^{x} \frac{1}{n}$. So done.

So, the series diverges – and diverges slowly – because it is approximately the logarithm. In fact, the limit as x goes to infinity of $\sum_{n=1}^{x} \frac{1}{n} - \ln(x)$ exists and is equal to one minus the area of what the infinite red regions in the diagram above would be if we kept extending it: This limit is often called Y and is approximately equal to 0.5772.

Personal story: When I was probably around 5 or 6 years old and playing around with my calculator and tried the $\sum_{n=1}^{x} \frac{1}{n}$ series, which is known as the harmonic series, I noticed that the sums like $\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}$ or $\frac{1}{9} + \frac{1}{10} + \cdots + \frac{1}{16}$ used in the standard powers of 2 proof that were known to be at least a half converged to a value around 0.693. It turns out from the logarithmic behavior from the proof above that this value is exactly ln(2).

Example: Lets consider $\sum_{n=1}^{\infty} \frac{1}{n^2}$.

We can bound this as follows:

$$1 + \left(\frac{1}{2^2} + \frac{1}{3^2}\right) + \left(\frac{1}{4^2} + \frac{1}{5^2} + \frac{1}{6^2} + \frac{1}{7^2}\right) + \cdots < 1 + \frac{2}{2^2} + \frac{4}{4^2} + \frac{8}{8^2} + \cdots = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 2$$

Where we now group from 2^k to $2^{k+1}-1$. So we bound the series above by 2 and thus it converges (bounded increasing sequences converge). But what does it converge to? This is a surprisingly deep question and the answer turns out to be $\frac{\pi^2}{6}$. We do not need to prove this since it's just a cool fact and not something we will build work on – the point of the example was to prove convergence. A proof can be found in the misc results section of the website (this infinite sum is called the basel problem).

To do proper rigorous proofs, we should really get bounds on the partial sums and not manipulate infinite sums.

Now here is an important example: Let (d_n) be a sequence with $d_n \in \{0,1,2,3,4,5,6,7,8,9\}$. Then look at the series $\sum_{n=1}^{\infty} \frac{d_n}{10^n}$. This is the decimal expansion of the number we can write out as $0.d_1d_2d_3$ … Now note that the series is bounded below by 0 and bounded above by 1 since the partial sums are increasing and each bounded above by 1. Also the series is bounded by its largest possible value which is the series $\sum_{n=1}^{\infty} \frac{9}{10^n}$, which by a geometric series we can show that this is 1. We sketched out a proof that this is 1 all the way back in Level 1.

We will now prove the obvious statement that every real number x between 0 and 1 has a decimal expansion. As an example, we will try to reach $x = \frac{1}{3}$. What happens is it is 0.3(stuff) because it is between $\frac{3}{10}$ and $\frac{4}{10}$. We can carry on by bounding it between consecutive integers divided by 100, then 1000, etc then we get that it is 0.33333……. What happens is the difference between our x and the fraction such as $\frac{3}{10}, \frac{33}{100}, \frac{333}{1000}$, … gets arbitrarily small: it is less than $\frac{1}{10^n}$, which can be made smaller than any ε>0.

Now suppose $0.a_1a_2a_3 \ldots = 0.b_1b_2b_3 \ldots$ and we want to know when this implies $a_i = b_i$. Lets suppose that k is the first place where a and b differ. Lets say $a_k < b_k$ since it could be the other way and the same idea holds. We know $\sum_{j=k+1}^{\infty} \frac{a_j}{10^j} \leq \sum_{j=k+1}^{\infty} \frac{9}{10^j} = \frac{1}{10^k}$. Therefore the only possibility is $b_k = a_k + 1$, since if they differed by more then this tail sum would not be able differ enough to compensate. In fact, if $b_k = a_k + 1$ then the tails have to differ by $\frac{1}{10^k}$ and thus must be 00000… and 99999… Therefore decimal expansions are unique unless they end with trailing 0's or trailing 9's or equivalently if they can be written as $\frac{integer}{10^{integer}}$ that is the only way their decimal expansion is not unique.

Lecture 20:

Theorem: A decimal expansion is periodic (eventually repeats) if and only if the number is rational

Proof: Lets see one direction with an example, by doing something similar to what we would do in GCSE maths:

$$x = 0.7832147147147147 \ldots = \frac{7832 + 0.147147147147 \ldots}{10000} = \frac{7832 + \frac{147}{999}0.999999999}{10000}$$

$$= \frac{7832 + \frac{147}{999}}{10000}$$

Which we can see must be rational.

Conversely, we want to show that any rational number has an eventually periodic decimal expansion. If $x = \frac{a}{b}$. We will rewrite this as $x = \frac{p}{2^a 5^b q}$ with a,b,p at least 0 and q a positive integer. And we want q to be coprime to 10, ie we take out all factors of 2 and 5. We will now write:

$10^{a+b}x = \frac{p2^b 5^a}{q}$ for non-negative integers c and d. If we show that this has a periodic decimal expansion we are done since we just shifted by a+b units. We will write this as $\frac{nq+c}{q} = n + \frac{c}{q}$. But now q and 10 are coprime which means that $10^{\phi(q)} = 1\ mod\ q$ by the fermat-euler theorem. Therefore we know that $10^{\phi(q)} - 1 = kq$. We want to show that $\frac{c}{q}$ has a periodic decimal expansion, but this is the same as $\frac{kc}{kq} = \frac{something}{999999.....9999999}$ which is periodic, for example $\frac{124683}{999999} = 0.124683124683124683124683$ ... so done. And we are ok since the numerator is not greater than the denominator so we have enough space.

Proposition: e is irrational

Proof: We know from the taylor series of e^x evaluated at x=1 that $e = 1 + 1 + \frac{1}{2} + \frac{1}{3!} + \frac{1}{4!} + \cdots$

Now suppose $e = \frac{a}{b}$. Then we know that $1 + 1 + \frac{1}{2} + \frac{1}{3!} + \frac{1}{4!} + \cdots + \frac{1}{b!} = \frac{b! + b! + \frac{b!}{2} + \frac{b!}{3!} + \cdots 1}{b!}$. It is easy to see that the numberator is an integer: $\frac{b!}{x!} = (x+1)(x+2)\ldots(b-1)b$ if x is an integer less than b. But then since $e = \frac{a}{b}$, we know that $e = \frac{a(b-1)!}{b!}$, so e is an integer divided by b!. Therefore,

$e - \left(1 + 1 + \frac{1}{2} + \frac{1}{3!} + \frac{1}{4!} + \cdots + \frac{1}{b!}\right)$ must also be an integer divided by b!. But then $\frac{1}{(b+1)!} + \frac{1}{(b+2)!} + \frac{1}{(b+3)!} + \cdots$ is what that is equal to, but that is between 0 and $\frac{1}{b!}$ which we will formally show shortly, so we will have our contradiction.

Basically, b is at least 1. Therefore $\frac{1}{(b+1)!} + \frac{1}{(b+2)!} + \frac{1}{(b+3)!} + \cdots$

$= \frac{1}{b!}\left(\frac{1}{b+1} + \frac{1}{(b+1)(b+2)} + \frac{1}{(b+1)(b+2)(b+3)} + \cdots\right)$. But the stuff in the brackets is between 0 and 1, since it is strictly less than $\frac{1}{b+1} + \frac{1}{(b+1)^2} + \frac{1}{(b+1)^3} + \cdots$, but b is at least 1 so this is at most $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots$ which is 1. Therefore we have our contradiction. So done.

Proposition: The number $L = \sum_{n=1}^{\infty} \frac{1}{10^{n!}} = 0.110001000000000000000001000\ldots$ is transcendental

Proof: We will use the fact that for any polynomial p there exists a constant k such that $|p(x) - p(y)| \leq k(x-y)$ for x and y between 0 and 1, and this is because p has a maximum derivative so k is at most that (by the mean value theorem, see differential equations lecture 2). We will use also the fact that a polynomial of degree d that is non-zero has at most d real roots (it can't have more than d factors by the factor theorem). We will finish the proof next lecture.

Lecture 21:

Suppose L is the root of a polynomial p with integer coefficients and suppose this has finite degree d. We see that 0<L<1. We know that there exists a k such that $|p(x) - p(y)| \leq k(x-y)$ for x and y between 0 and 1. Set $L_n := \sum_{k=1}^{n} \frac{1}{10^{k!}}$ so that $L_n \to L$. Note that $|L - L_n| = \sum_{k=n+1}^{\infty} \frac{1}{10^{k!}} \leq \frac{2}{10^{(n+1)!}}$. Now write d with integer coefficients as $a_d x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0$ with $a_d \neq 0$. Notice also that

$L_n = \frac{s}{n!}$ For some natural number s. But then $p(L_n) = \frac{t}{10^{d*n!}}$ for some integer t. This will be non-zero for sufficiently large values of n since otherwise there would be more than d roots if t is 0 infinitely many times. Since p(L)=0 by assumption, we know that $|p(L_n) - p(L)| = |p(L_n)| \geq \frac{1}{10^{d*n!}}$. But then we have for sufficiently large n that that $\frac{1}{10^{d*n!}} \leq |p(L_n) - p(L)| \leq k|L_n - L| \leq \frac{2k}{10^{(n+1)!}}$. But then when n is big enough, what will happen is that $\frac{2k}{10^{(n+1)!}}$ is much smaller because the denominator is much larger, so we have a contradiction, so we are done.

This does not prove that $e = \sum \frac{1}{n!}$ is transcendental but it turns out that it is (although this result is beyond the course).

A similar proof actually shows that if x is too good of a rational approximation, ie for all n there exists a rational number such that $\frac{p}{q}$ is within $\frac{1}{q^n}$ of x, then x is transcendental.

We can define the complex numbers as real-imaginary part pairs satisfying:

- (a,b)+(c,d)=(a+c,b+d)
- (a,b)*(c,d)=(ac-bd,ad+bc)

Which satisfies all the usual properties.

Definition: A set is **countably infinite** if it is infinite but in bijection with $\mathbb{N}$. It is **countable** if it is countably infinite or finite. Countable basically means we can list or ennumerate the elements of the set.

Proposition: $\mathbb{Z}$ is countable.
Proof: 0,-1,1,-2,2,-3,3,...

Lemma: Any subset of $\mathbb{N}$ is countable.

Proof: Let S be a non empty subset of $\mathbb{N}$. By the well ordering principle there is a least element s_1. There is also (if there are more elements) a second least element s_2, then s_3, etc. If at some point this process ends we have a finite set which is countable. If not, the map $n \to s_n$ is the bijection we need. This is injective because these minimal elements are all different, and this is surjective because if k is in S then k is a natural number and there are less than k elements in S less than k so we must hit it at some point.

Lecture 22:

Threorem:

i) X is countable if and only if there exists an injection from X->$\mathbb{N}$
ii) If X is non-empty, then X is countable If and only if there is a surjection from $\mathbb{N}$->X

Proof:

i) If X is finite it obviously injects into $\mathbb{N}$. If X is countable it bijects to $\mathbb{N}$ so it injects into $\mathbb{N}$. Conversely, suppose X injects into $\mathbb{N}$, then trivially X has a bijection to its image s where s is a subset of $\mathbb{N}$. If s is finite then so is X so done. If s is infinite then surely s is not uncountable as this would make no sense, but the proof is that we can use our lemma from last lecture that subsets of $\mathbb{N}$ are countable.

iii) If X is countable then either it is finite in which there is a trivial surjection $\mathbb{N}$->X or it is countably infinite so it is a bijection. So suppose f is a surjection $\mathbb{N}$->X. So for each a in X pick the least element in $\mathbb{N}$ that maps to a which exists as it is a surjection then send each element in X to that one, then we have an injection as we cannot have two things in $\mathbb{N}$ send to the same thing in X under the surjection and that would be the only way we have a many-to-one under the map we just constructed. But then we have an injection X->$\mathbb{N}$ so X is countable by (i).

Corollary: Any subset of a countable set is countable

Proof: Pick a bijection from our countable set to $\mathbb{N}$. Then the image in $\mathbb{N}$ of our subset is countable by the fact that any subset of $\mathbb{N}$ is countable and in bijection with our subset, so done.

We can view countability as saying a set is not fundamentally bigger than $\mathbb{N}$.

Theorem: $\mathbb{N}\times\mathbb{N}$ is countable.

Proof:

(Image of a proof. Note that for every element we can pick a finite number x and say this is the x'th element, which is why we can ennume)

Alternative proof:

Lets use the previous theorem and construct an injection from $\mathbb{N}\times\mathbb{N}$ to $\mathbb{N}$. We can do this by sending each pair (a,b) to $2^a 3^b$ which is an injection by uniqueness of prime factorization.

Proposition: $\mathbb{Z}\times\mathbb{Z}$ is countable.

Proof:

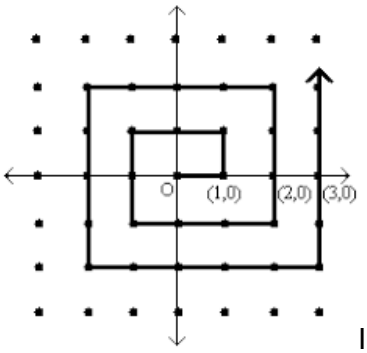


Image of a proof

Theorem: $\mathbb{Q}$ is countable

Proof: $\mathbb{Q}$ is in bijection with the subset of $\mathbb{Z}\times\mathbb{Z}$ where (a,b) written as $\frac{a}{b}$ is in lowest terms and b is not 0. And subsets of countable sets are countable.

Theorem: A countable union of countable sets $A_{1,2,\ldots}$ is countable. Ie, we've been given a list of countably many countable sets and we wish to show the union is countable.

Idea: It is clearly a subset of $\mathbb{N}\times\mathbb{N}$

Proof: For each A_i, list the elements like $A_{i1}, A_{i2}, A_{i3}, \ldots$*

*(For vectors and matrices students, no we are not doing the summation convention)

For each $x = A_{ij}$ in the union, pick an instance of x as your representative, eg by doing the least i such that it is in the i'th set (since only one instance of the same element is allowed in the union) and send it to $2^i 3^j$, so we have an injection to $\mathbb{N}$ so done.

Theorem: $\mathbb{A}$ (the set of algebraic numbers) is countable.

Proof: We just need to show that the set of all polynomials with integer coefficients is countable because each one just has a finite number of roots because then $\mathbb{A}$ would be a countable union of finite sets. We just need to show that the set of polynomials with integer coefficients of a fixed degree is countable because then we just have to take the countable union of all those. But then the set of degree k polynomials is exactly the same as $\mathbb{Z}^{k+1}$, but this is countable because by induction, $\mathbb{N}$ is countable so if $\mathbb{Z}^k$ is countable then $|\mathbb{Z}^{k+1}| = |\mathbb{Z}^k \times \mathbb{Z}| = |\mathbb{Z} \times \mathbb{Z}| = |\mathbb{N}|$ where when I say these are equal I mean there is a bijection between them, by the induction hypothesis and a previous theorem on $\mathbb{Z} \times \mathbb{Z}$. So done.

Theorem: $\mathbb{R}$ is uncountable

Proof: See level 6 technical results

Corrolary: Transcendental numbers exist, in fact there are uncountably many of them

Proof: If all real numbers were algebraic they would be countable so this is a contradiction. If there were countably many then the reals would be a union of two countable sets which is also a contradiction.

Lecture 23:

Recall that we can use a diagonal argument to show that $\mathbb{R}$ (and in fact any interval) is uncountable. We can use the same argument to show that a set does not have a bijection to its power set. We will do this proof now.

So suppose, for example, that $\mathbb{N}$ has a bijection to its power set. Then in each subset in our list, each natural number will either be in (Y for yes) or not (N for no) in the list. So as an example we can write it in a table like this:

Y	N	Y	…
N	N	Y	…
N	Y	N	…
…	…	…	…

Now highlight the diagonal elements. Then flip them, so in this example we get NYY… . Then this is not in our list. So contradiction, so done.

We can write this formally as $S = \{n \in \mathbb{N} : n \notin S_n\}$ and we can apply this to any set. This gives a pretty elegant way to see the inequality $2^n > n$.

It turns out that $|P(\mathbb{N})| = |\mathbb{R}|$. This is an interesting fact and we will prove this next lecture.

Note that $|\mathbb{R}| = |(0,1)|$ because we can do $\tan\left(\pi\left(x+\frac{1}{2}\right)\right)$ as a bijection.

Alternative proof that $P(\mathbb{N})$ is uncountable: We want to find an injection from (0,1) to $P(\mathbb{N})$. To do this we pick something between 0 and 1 and write x in binary or base 2 (as in the place values differ by a factor of two instead of ten and the digits are 0 or 1). Then consider the terminating binary expansion (ie, if there is a not unique one, pick the one without an infinite string of 1's). Then if at the k'th digit there is a 1, make it so k is in the subset of $\mathbb{N}$ that that maps to. We can think of this as turning 0's to N's and 1's to Y's. As an example, 0.11101 in binary will map to the subset {1, 2, 3, 5}. Therefore if there was an injection from $\mathbb{N} \to P(\mathbb{N})$ we could pick the elements in $P(\mathbb{N})$ that get mapped to from our injection $(0,1) \to P(\mathbb{N})$ and pick the elements in $\mathbb{N}$ that map to those to get a surjection from a subset of $\mathbb{N}$ to $(0,1)$, which we can turn into a surjection $\mathbb{N} \to (0,1)$, which we know is impossible. So done.

Note that there is no set of all sets or else it would contain its power set as that is a set.

Example: Consider a family of open pairwise disjoint intervals A_i of the reals. Then this is countable because we cannot have an uncountable set of non-zero things add up to a finite number (Level 6 technical results), which would have to happen because we could arctan everything in the reals including the open intervals and then get such a sum that contradicts this. An alternative proof is that each interval contains a rational number but then if the set was uncountable we would have uncountably many rational numbers. Another alternative proof is a variation of the level 6 technical results proof of the related fact that we cannot add uncountably many numbers and get a finite number: The set of sets with a length at least 1 is countable, so is the set of intervals with length $\geq \frac{1}{2}$, $\geq \frac{1}{3}$, and so on, then we have a countable union of countable sets (every interval is in one by the archimedian property), so the set of all the intervals is countable.

Lecture 24:

Summary of what we've done so far:

X is countable if

i) There is an injection X to $\mathbb{N}$
ii) You can ennumerate it
iii) It is a countable union of countable sets

X is uncountable if

i) We can run a diagonal argument to show that there is no surjection $\mathbb{N}$ to X
ii) X has an injection to another uncountable set

Intuitively we think of existence of a bijection to mean sets have the same size. It's like we have an equivalence relation of sizes by existence of a bijection.

Intuitively we think of "A injects to B" to mean "A is at most as big as B" and "A surjects to B" to mean "A is at least as big as B". For this to make sense, we need to formally prove that if A has both an injection and a surjection to B then there is a bijection A to B, and also that that "A injects to B" is equivalent to "B surjects to A".

Lemma: Given non-empty sets A and B, the existence of an injection f from A to B is equivalent to the existence of a surjection g from B to A.

Proof: Define $g: B \to A$ by sending $f(a_0) \to a_0$ for each a_0 in A and for an element not in the image of f sending it to anything. This gives a surjection g from B to A. Conversely, if g is a surjection, then partition B into the pre-images of each element in A. Then pick each a in A to get set to something in the corresponding pre-image to get an injection.

Lemma: If there is both an injection (f) and a surjection (t) from $A \to B$ then there is also a bijection (h). We can interpret this to mean that $|A| \leq |B| \leq |A| \implies |A| = |B|$

Proof: Note that by the previous lemma there is an injection g from $B \to A$. For each a in A, pick the element $g^{-1}(a)$ in B in the pre-image of A if it exists, then pick the pre-image of that in A if it exists, ie $f^{-1}\big(g^{-1}(a)\big)$, then consider the sequence of these pre-images: ... $g^{-1}(f^{-1}(g^{-1}($... and we will call this the ancestor sequence of a, as it is talking about where a came from. These sequences will either continue infinitely, or have an even number of steps, or have an odd number of steps. Let A_0 be the set of a in A whose ancestor sequences terminate after an even number of steps, ie the last point is in our set A. Now let A_1 be the set of a in A that end in an odd number of steps, and A_∞ be those such that the ancestor sequence does not terminate. Do the same for B. Now note that we can construct bijections between the following partitions: $A_0 \to B_1, A_1 \to B_0, A_\infty \to B_\infty$, then we can construct a bijection out of those three. f must biject $A_0 \to B_1$: it is an injection because f is an injection, and it is a surjection because everything in B_1 has an odd number of ancestors so for each b in B_1, $f^{-1}(b)$ exists and is in A_0, so we have a surjection as b was arbitrary. Now note that for every a in A_1, we must have that $g^{-1}(a)$ exists and is in A_0, and we have a surjection because each b comes from g(b) and an injection because $g^{-1}(a) = g^{-1}(b)$ implies $a = b$ because if $a \neq b$ then since g is a function we must have that $g^{-1}(a) \neq g^{-1}(b)$. Now we have a bijection $A_\infty \to B_\infty$ by f, which is clearly an injection, and it is a surjection because for each b in B_∞ by definition there is an $f^{-1}(b)$ in A_∞. So we can write our bijection as follows: $h(x) = \begin{cases} f(x): x \in A_0 \\ g^{-1}(x): x \in A_1 \\ f(x): x \in A_\infty \end{cases}$

Example: We have a bijection $[0,1] \to [0,1] \cup [2,3]$ because we have a surjection by (3x or something else if this is in (1/3, 2/3)) and an injection by the identity map.

And that is the end of the course.