Mathematics

For Data Science I

WEEK 1 - WEEK 4

Prof. Neelesh Upadhye
Department of Mathematics
Indian Institute of Technology Madras

Prof. Madhavan Mukund
Department of Computer Science,
Chennai Mathematical Institute

IIT Madras
ONLINE DEGREE

INDEX

S. No	Topic	Page No.
	Week 0	
Lec 0	Introduction	01
	Week 1	
Lecture 1	Natural Numbers and their operations	04
Lecture 2	Rational Numbers	21
Lecture 3	Real and Complex Numbers	32
Lecture 4	Set Theory	40
Lecture 5	Construction of Subsets and set operations	57
Lecture 6	Sets: Examples	70
Lecture 7	Examples of Set Operations and Counting Problems	83
Lecture 8	Relations	92
Lecture 9	Functions	110
Lecture 10	Relations: examples	127
Lecture 11	Function: examples	144
Lecture 12	Prime Numbers	151
Lecture 13	Why is a number irrational?	158
Lecture 14	Set versus Collections	164
Lecture 15	Degrees of infinity	171
	Week 2	
Lecture 16	Rectangular Coordinate system	188
Lecture 17	Distance formula	194
Lecture 18	Section formula	199
Lecture 19	Area of triangle	204
Lecture 20	Slope of a line	208
Lecture 21	Parallel and perpendicular lines	213
Lecture 22	Representation of a Line-1	220
Lecture 23	Representation of a Line-2	232
Tutorial 2.1	Tutorial 2.1	238
Tutorial 2.2	Tutorial 2.2	245
Tutorial 2.3	Tutorial 2.3	250
Tutorial 2.4	Tutorial 2.4	253
Tutorial 2.5	Tutorial 2.5	258
Tutorial 2.6	Tutorial 2.6	260

Tutorial 2.7	Tutorial 2.7	264
Tutorial 2.8	Tutorial 2.8	270
Tutorial 2.9	Tutorial 2.9	275

Week 3

Lecture 24	General equation of line	279
Lecture 25	Eq of parallel and perpedicular lines in general form	287
Lecture 26	Equation of a perpendicular line passing through a point	290
Lecture 27	Distance of a line from a given point	292
Lecture 28	Straight line fit	300
Tutorial 3.0	Tutorial 3.0	308
Tutorial 3.1	Tutorial 3.1	315
Tutorial 3.2	Tutorial 3.2	320
Tutorial 3.3	Tutorial 3.3	328
Tutorial 3.4	Tutorial 3.4	330
Tutorial 3.5	Tutorial 3.5	333
Tutorial 3.6	Tutorial 3.6	335
Tutorial 3.7	Tutorial 3.7	337
Tutorial 3.8	Tutorial 3.8	340

Week 4

Lecture 29	Quadratic functions	342
Lecture 30	Examples of Quadratic functions	352
Lecture 31	Slope of quadratic function	357
Lecture 32	Solution of quadratic equation using graph	361
Lecture 33	Slope: Line & Parabola	368
Tutorial 4.1	Tutorial 4.1	380
Tutorial 4.2	Tutorial 4.2	383
Tutorial 4.3	Tutorial 4.3	384
Tutorial 4.4	Tutorial 4.4	386
Tutorial 4.5	Tutorial 4.5	388
Tutorial 4.6	Tutorial 4.6	390
Tutorial 4.7	Tutorial 4.7	394
Tutorial 4.8	Tutorial 4.8	396

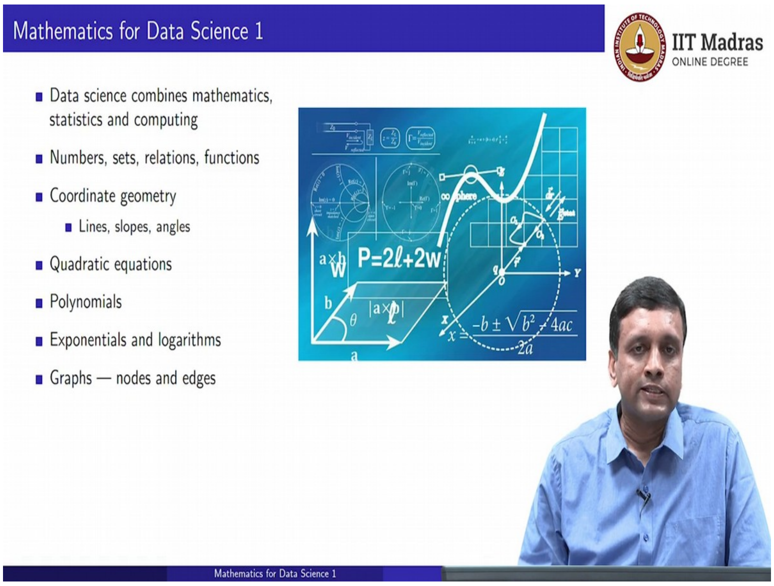
Mathematics for Data Science 1

Prof. Madhavan Mukund
Prof. Meghana Nasre
Prof. Neelesh S Upadhye
Department of Computer Science
Department of Computer Science & Engineering
Department of Mathematics
Chennai Mathematical Institute
Indian Institute of Technology, Madras

Lecture - 00
Introduction

(Refer Slide Time: 00:05)

(Refer Slide Time: 00:14)



So, welcome to the course on Mathematics for Data Science. This is the 1st course of two courses which are there in the foundational setting. So, why are we studying mathematics in this programming and data science course is because data science actually combines mathematics, statistics and computing. So, without a good background in mathematics, it is not possible to really appreciate many of the ideas that go into data science.

So, in this 1st course in mathematics for data science, we will basically be covering material which may be familiar to many of you. We will start with fairly basic things about numbers, sets, relations and functions. This is just to bring everybody onto the same page in terms of terminology and notation. Many of these concepts as we said you would already know or even if you have not seen it for some time, this refresher should tell you what you need to know.

Having got these basics under our belt, we will do some coordinate geometry. So, we will look at how to draw lines and how to get the slope of a line, how to calculate the angles between two lines and so on. So, these are again things which you might have studied in school and you may have forgotten. So, it is good to brush up and remind ourselves of how these things work.

We will move on from lines to quadratic equations. So, if you remember lines represent linear equations, quadratic equations have a square term if you draw them, they look like parabolas. So, we will look at quadratic equations and then, we will generalize to higher power so, these are what are called polynomials. So, these are all functions which we can draw as graphs in the sense of coordinate geometry, but we can also analyze them in many different ways and functions will be quite essential in our study of data science. So, moving on from polynomials, we have functions which are not polynomials; those that grow very fast, these are exponentials and those that grow very slowly, these are logarithms.

So, to summarize we will be looking at large variety of functions starting from lines and going through polynomials to exponentials and logarithms. And finally, we will move to something which perhaps you have not seen in school which is a different form of graph. So, this is not the kind of graph where you have an x axis and a y axis and you draw a curve, explaining the relationship between x and y rather this is a graph of the kind you see when you look at for example, a map of an airline timetable. So, in this graph, we have nodes representing points of interest and edges representing connections.

So, one example is a road network or an airline network, but these edges can also represent other relations. For example, we can think of an organization and we can think of employees and they are connected to the manager that they report to. So, we will look at graphs, how to represent data as graphs and some simple manipulation on graphs algorithmically.

So, I hope you will enjoy this course. I am sure that a lot of it will be familiar to you, but I hope that you will also find something new and a new perspective on things that you already know and with this, you should have a good foundation for all the courses that come up ahead.

Thank you.

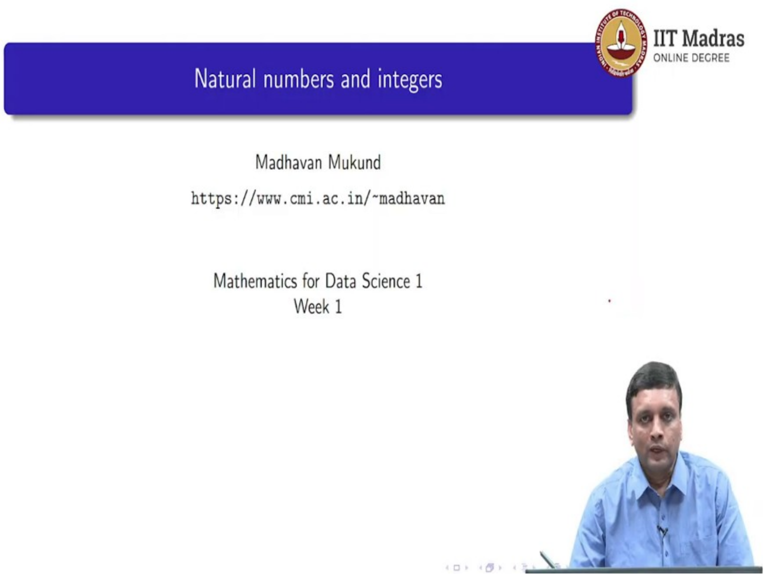
Mathematics for Data Science 1

Prof. Madhavan Mukund
Department of Computer Science
Chennai Mathematical Institute

Week - 01
Lecture – 01
Natural Numbers and their operations

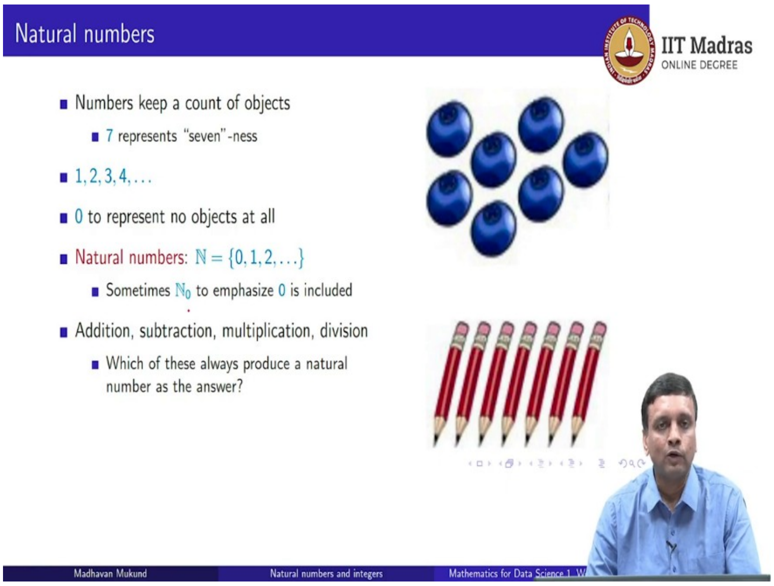
(Refer Slide Time: 00:06)

(Refer Slide Time: 00:14)



So, welcome to the 1st week of Mathematics 1 for Data Science. So, we are going to start with some very basic things which you probably know; right from the beginning we are going to start talking about numbers. So, in this 1st module what we are going to talk about is natural numbers and integers.

(Refer Slide Time: 00:30)



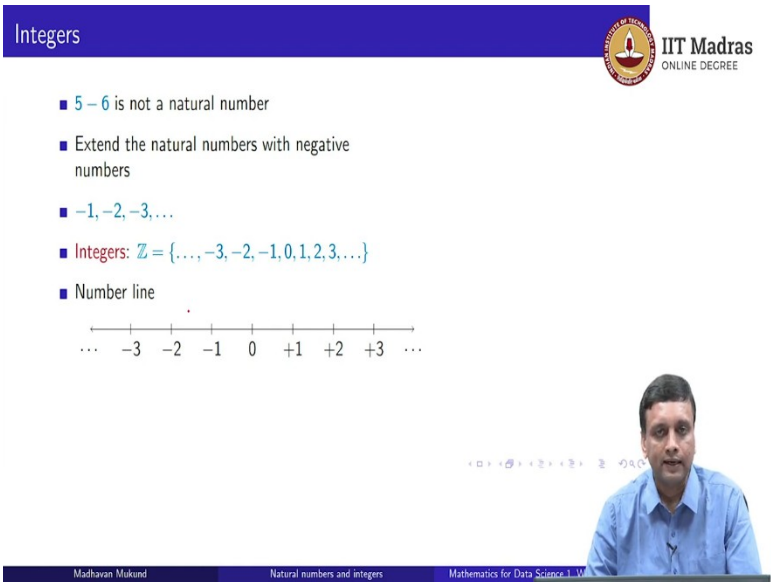
So, as you probably remember from as young as you were in school when you first came across numbers, we use numbers mainly for counting. So, for instance if we see 7 balls like this and then we see 7 pencils like this, then we need to know that these are the same number of things and for this we use this number 7. So, 7 represents what is common to these two objects that there are 7 balls and 7 pencils. So, 7 is an abstract concept in that sense and it refers to a quantity.

So, we all of course, know the numbers 1, 2, 3, 4 and all that. So, when we see a number of things, we can count them. But perhaps the most important number of all which is of Indian origin is 0. So, it is quite important to have a way to represent something when there is nothing to count because without a 0, we cannot use our place numbering system that we use to manipulate numbers.

So, these numbers starting with 0 are what are often called the natural numbers. Now there is some confusion in some books and many books will actually use only 1, 2, 3, 4 to represent the natural numbers. So, we use N to represent the set of natural numbers and in case there is any confusion whether 0 is included in this set or not, now sometimes people will not include 0 in the set of natural numbers.

So, sometimes to emphasize that we are using 0, we will actually put the subscript 0 below the N right. So, we will write either N or N_0, but whenever we are talking about natural numbers, it always includes a 0. Now what can we do with natural numbers? Well we can add them, we can subtract them, we can multiply them, we can divide them. So, these are the normal arithmetic operations which you have studied in school.

(Refer Slide Time: 02:43)



But what is really interesting from a mathematics perspective is, when we take natural numbers and we perform an operation on them, do we always get a natural number? So, if we add two natural numbers, do we get a natural number? If we subtract a number from another, we get a natural number? If we multiply them, do we get a natural number? If we divide one by another, do we get a natural number?

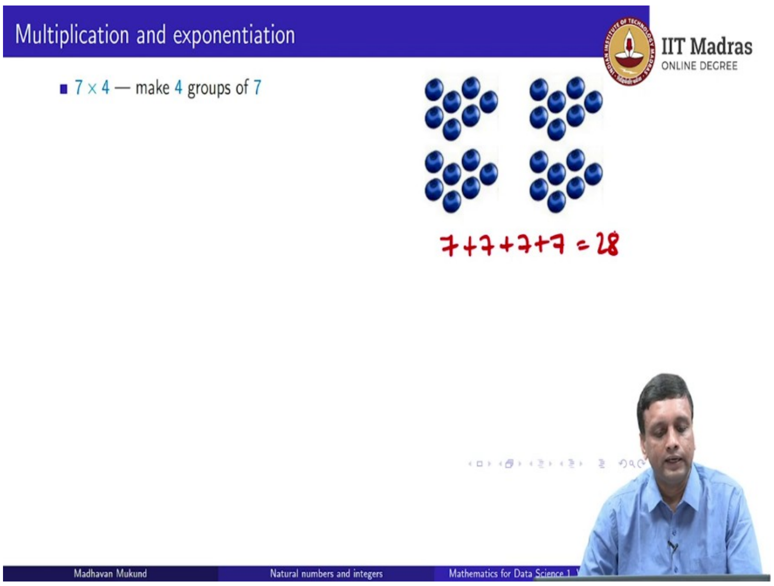
So, the first operation which fails this test is subtraction because if we subtract a larger number from a smaller number, so supposing we take 6 and subtract it from 5; then we go below 0 right. If you have 5 things and we take away 6 things, we will be cannot take away 6 things that is what subtraction means. So, we need to expand the scope of our numbers to allow these operations to work sensibly and this is how we get the negative numbers.

So, we had the positive numbers 0, 1, 2 the non-negative technically because 0 is neither positive nor negative. So, we had the positive numbers 1, 2, 3, 4. We added a 0 to account for the fact that we are counting nothing and now we add symmetrically on the other side negative numbers -1, -2, -3. So, this is just to illustrate why we get them of course, this is something that you should know from school.

So, this set which is the natural numbers extended with a negative numbers is what we call the integers and we use Z to indicate the set of integers. So, we have N the set of natural numbers which starts at 0 and goes forward 0, 1, 2, 3, 4 and we have the integers which start at no at minus infinity and go to plus infinity. So, these are both infinite sets, but the natural numbers have a starting point 0 and the integers extend to infinity in both directions.

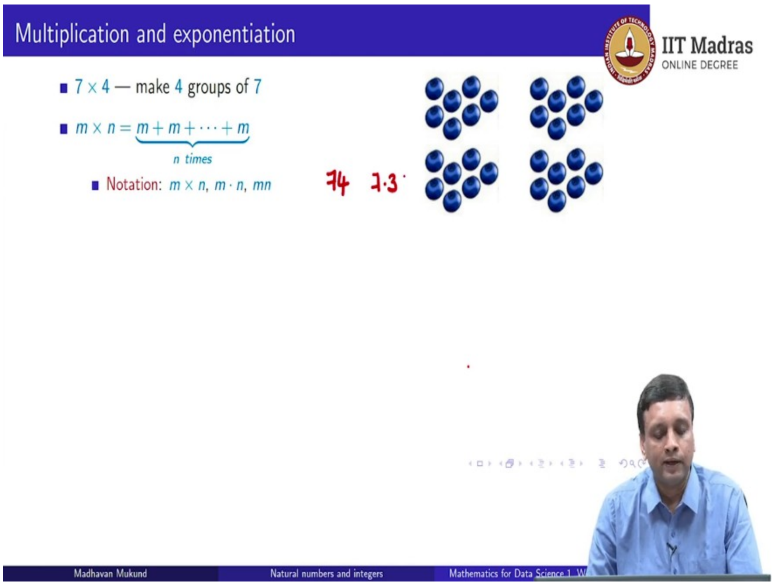
So, it is very convenient mentally to think of the integers as forming this kind of a sequence where on the left you have the very small ones and on the right you have the very long ones and this is normally called the number line. So, as you go from left to right, the numbers are increasing and this is how the integers are arranged.

(Refer Slide Time: 04:27)



So, we said that subtraction takes us away from natural numbers and we brought the integers. So, now, let us look at the other two operations that we talked about multiplication and division. So, let us start with multiplication. So, when we say 7 × 4 what we are really saying is take 7 objects and make 4 copies of them. So, for instance on the right, we have those 7 balls that we started with and then we have made 4 copies of them. So, if we want to know how many balls are here, then we have 7 from the first group, 7 from the second group and so on. So, we have 4 groups of 7 and this is if we add it up going to be 28.

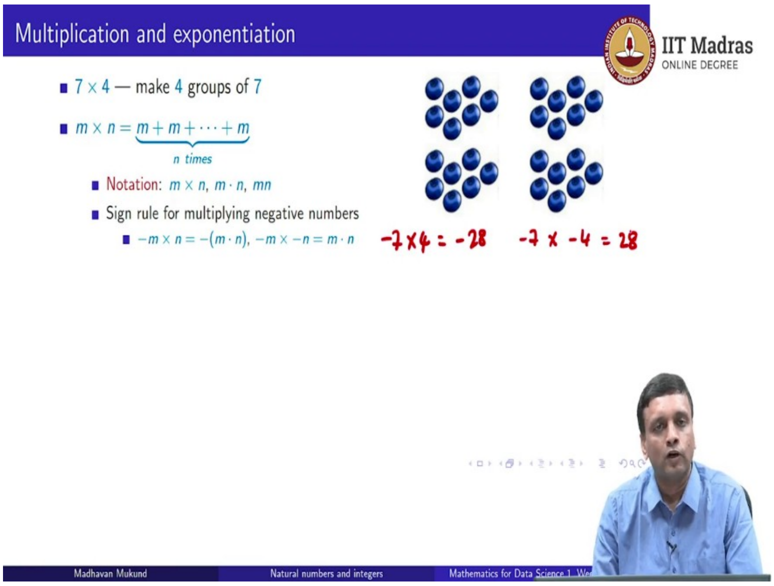
(Refer Slide Time: 05:02)



So, in general this is how we multiply when we take a number m and multiply it by n, what we are doing is we are making n copies of m. So, we are taking m + m + m… n times. So, in this sense multiplication is repeated addition.

So, we often use this time sign the × sign for multiplication, but this is often cumbersome when we write out equations. So, sometimes we replace this time sign by a . and sometimes we write nothing at all. So, if we just write two symbols together, we do not write this normally for numbers because imagine that if I write 7 4 like this, then you do not know whether it is a number 74 or its 7 × 4. So, if we have numbers, we will normally write a dot explicitly between them like 7 × 3. But when we have a names like m or n standing for numbers, then if we write mn; we assume that it is one number m multiplied by another number n.

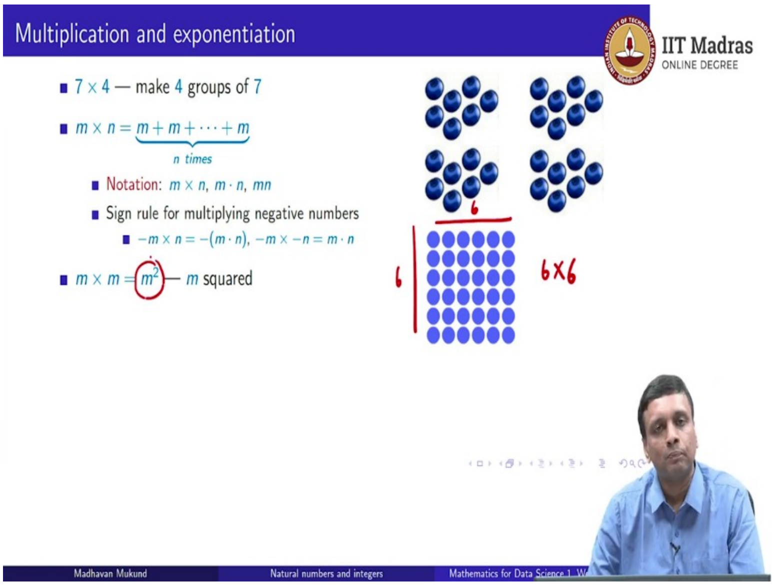
(Refer Slide Time: 05:56)



Now, we have integers, an integers have signs they are positive and negative numbers. So, we have to remember that when we multiply numbers with signs, the resulting number also has a sign and there is a sign rule which basically says that if we have one negative number multiplied by one positive number, then the result is a negative number. So, let us assume that m is a positive number so, - m is a negative number. So, say - 7 × 4 would be - 28. On the other hand if I take (- 7) × (- 4), then the two negations will cancel, and I will get 28.

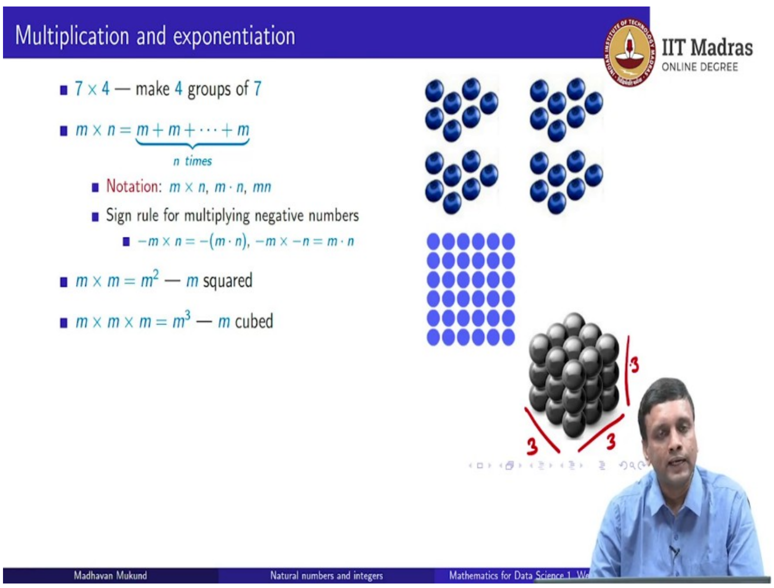
So, if you have an even number of minus signs, you get a positive number; if you have an odd number of minus signs, you get a negative number.

(Refer Slide Time: 06:40)



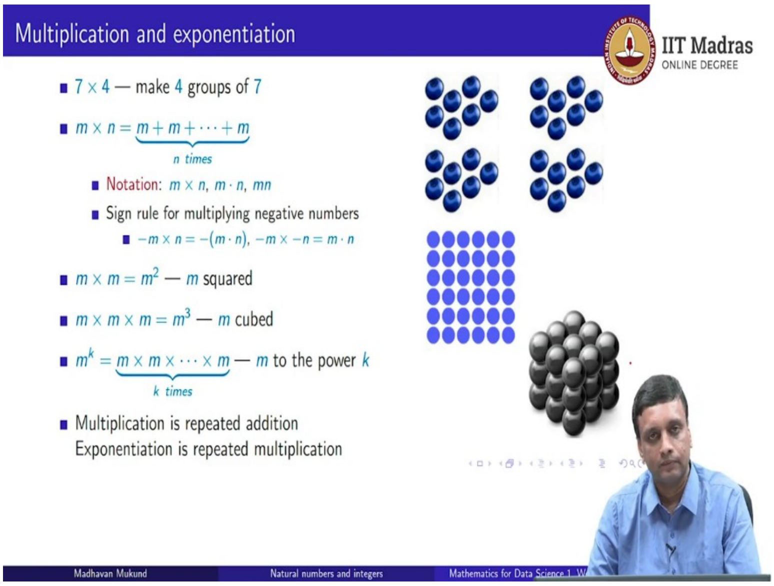
Now just like we have repeated addition, we can also do repeated multiplication. So, instead of doing m plus m, we can takes m times m and this is called m squared and the reason that it is called m squared is visible in the picture here. So, we have now here 6 balls and 6 balls. So, we have 6 × 6 right. So, this means that we can arrange these 6 times 6 balls in a square and this is why we call this square.

(Refer Slide Time: 07:15)



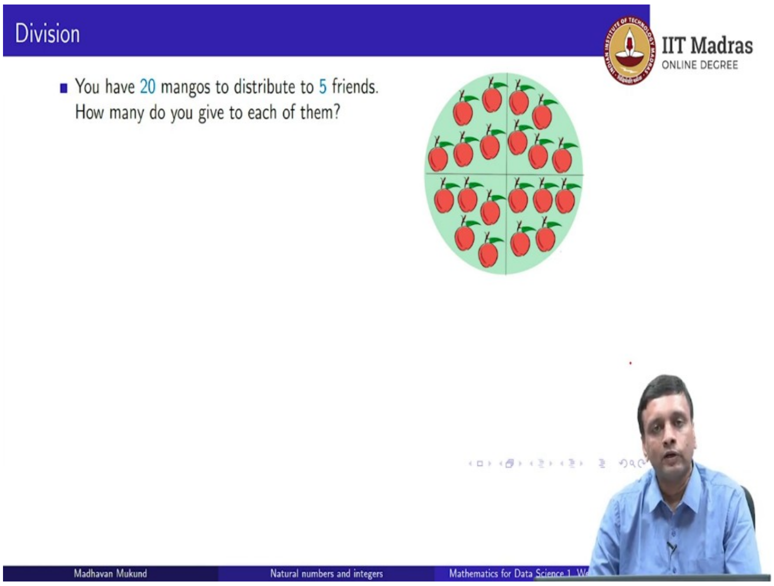
So, this notation m^2 stands to the fact that, m is multiplied by itself twice. Now if you multiply it by self 3 times, then we get a cube. So, here for instance we have 3 balls by 3 balls and then we have a height a stack of 3 such balls. So, we have a square of 3 by 3, 9 balls and we have 3 stacks of these one on top of the other. So, this naturally forms a cube so, m × m × m is written m^3.

(Refer Slide Time: 07:37)



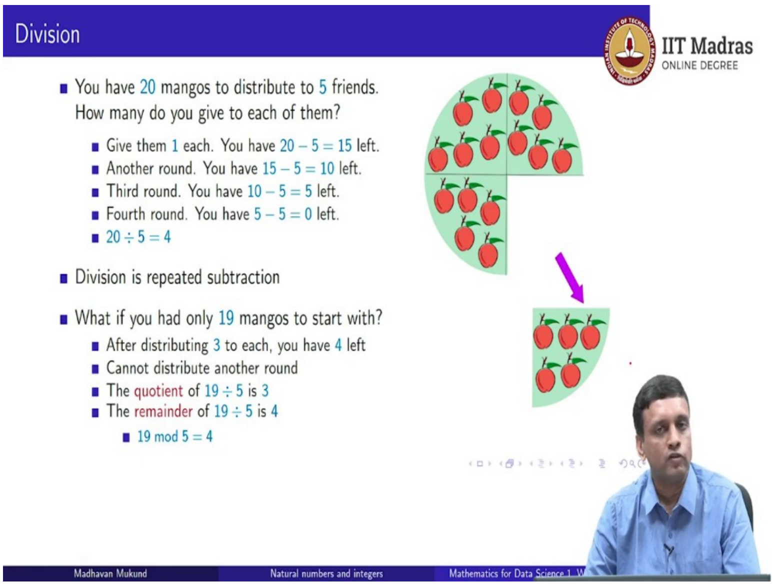
Now, unfortunately we live in a 3-dimensional world and we cannot imagine objects which have more than 3 dimensions. So, our vocabulary stops with cube. So, in general if we have m^k, then we write m × m × m…, k times and we just say it is m^k, we do not have a fancy name for it. We just say it is the kth power of m ok. So, to emphasize multiplication is repeated addition and exponentiation as we have seen is repeated multiplication.

(Refer Slide Time: 08:07)



So, now let us come to division. So, you would have seen this familiar problem in school. You have a certain number of objects and you want to divide them among certain number of people. So, for example, supposing you have 20 mangoes and you want to give them to 5 friends. So, how many mangoes does each friend get?

(Refer Slide Time: 08:31)



So, here on the right we have this picture and then, what you do is well you start by distributing one mango to each friend right. So, you take out 5 mangoes and you give them to each of your friends. So, now, you have given away 5 mangoes and you have only 15 mangoes left so, you repeat the process. Among the 15 mangoes, you give away 5 to your friends one each and now your 15 mangoes have become 10 and do it one more time and your 10 mangoes have become 5, do it a third time or fourth time rather and the 5 mangoes are now gone.

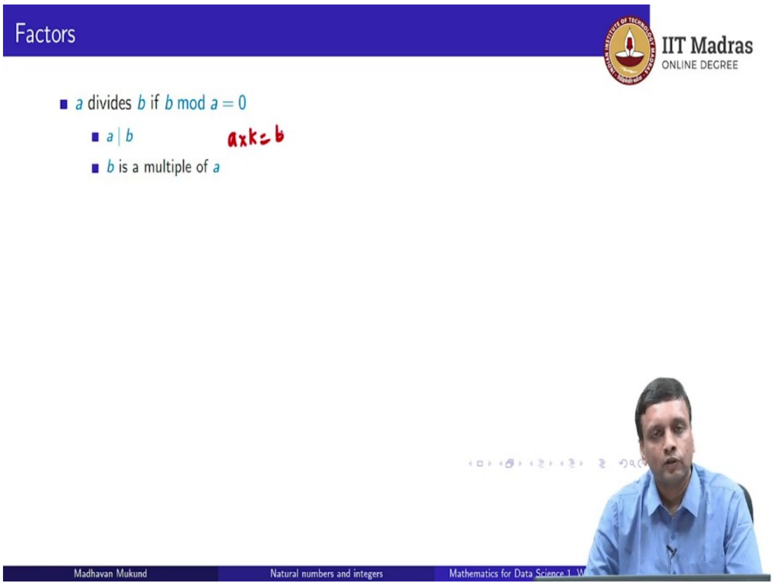
So, after 4 rounds of distributing mangoes, each time giving one mango each so, 5 mangoes per round, you have got rid of your 20 mangoes so, 20 ÷ 5 is 4. So, here as we have illustrated, division is actually repeated subtraction. You keep subtracting by the number you are trying to divide and finally, if you hit 0, then you have divided it exactly.

Well, what if you had only 19 mangoes? Now you know very well that 19 mangoes cannot be evenly divided into 4 into 5 groups. So, if you would start distributing like we had above the first three rounds would go fine; you would come from 19 to 14 from 14 to 9 and then you will come from 9 to 4 and now you have only 4 mangoes left and you have 5 friends so, you cannot give 1 each.

So, we have managed to distribute 3 times and we have 4 left over. So, formally this is written as saying that the quotient the number of times you can actually divide without getting into a fractional part is 3 and the remainder that is after you have a little bit left over which you cannot subtract one more time is the remainder is 4. So, for 19 ÷ 5, the quotient is 3 and the remainder is 4.

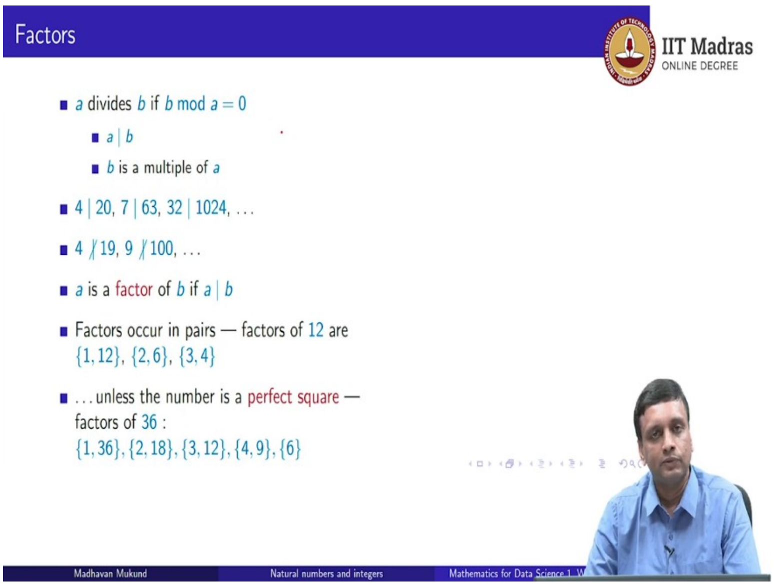
Now, very often we will need to use this remainder and there is a notation for remainder. So, this is this notation called modulus. So, modulus is another word for remainder and it is written as mod. So, 19 mod 5 is the same as the remainder when 19 is divided by 5. So, instead of saying the remainder of 19 divided by 5 is 4, we will often say 19 mod 5 is 4.

(Refer Slide Time: 10:38)



So, with this notation, we can now define what is a factor. So, a factor is a number which divides a bigger number evenly without any remainder. So, a | b, if b mod a is 0. Remember what this mean is means is that if b is divided by a, there is no remainder and we write this with this vertical bar |. So, on the left is the smaller number, on the right is the bigger number. So, a divides b this is what this is supposed to say and the other way of thinking about it is that b is some multiple of a. So, b if a | b then a.k=b ok. So, we have some multiple the some number of times that a goes into b. So, therefore, b is a multiple of a.

(Refer Slide Time: 11:27)



So, here are some examples we have already seen that $4 \mid 20$ because 4×5 is 20, $7 \mid 63$ because 7×9 is 63, $32 \mid 1024$ because 32×32 is 1024 and so on.

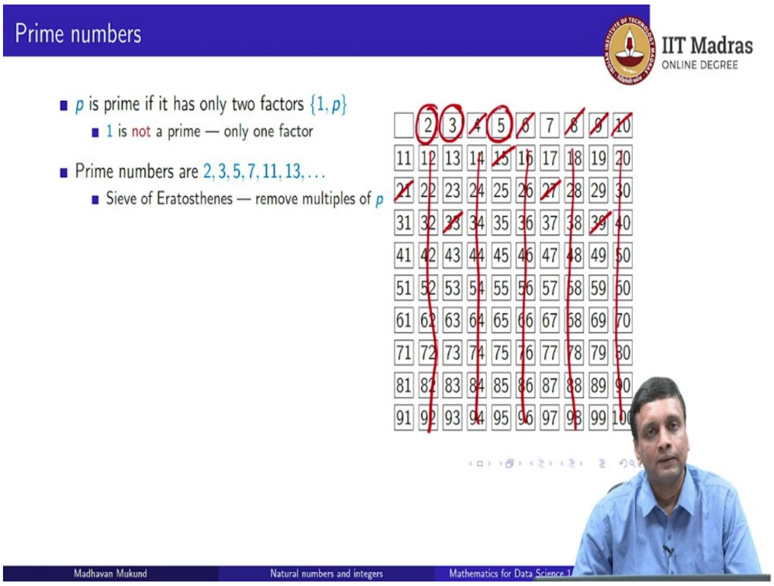
Now, the symbol that we use for not being a divisor is just to put a stroke across that vertical line. So, 4 does not divide 19 because there is no way to multiply anything by 4 and get 19. Similarly, 9 does not divide 100 evenly because we get $9 \times 11 = 99$ and then we go 108.

So, we say formally that a is a factor of b if $a \mid b$ right. So, $a \mid b$ is the same as saying that a is a factor of b and it is easy to see that factors must come in pairs because if $a \mid b$ then, a goes into b some k times. So, $k \mid b$ right so, $k \times a = b$ so, both k is a factor and a is a factor. So, for instance, if you take a number 12 then 1 is a factor because 1 divides everything and in fact, for every number n, $1 \times n$ is n so, the pair for 1 is always the number itself.

Now in this case, 12 is divisible by 2 and 2 goes in 6 times. So, the pair 2, 6 form 2 factors 6 times 2 is 12, 2×6 is 12 and similarly 3×4. Now, of course, there is an important side condition which is that sometimes the pair is the same as the number itself and this happens when the number actually happens to be a perfect square that is, it is some number multiplied by itself. So, for instance consider 36 so, 36 is 6×6. So, if you look at the factors of 36 and group them in pairs, then we have 1 and 36, we have 2 and 18, we have 3 and 12, we have 4 and 9 and finally, we have the factor 6, but 6 is multiplied by 6. So, 6 does not produce a new factor as its pair, it is just itself.

So, another way of thinking about it is that, if you have something which is not a square you will have an even number of factors, you will have 2 + 2 + 2 +2. If something is a square, you will have an odd number of factors, you will have 2 + 2 +2 and finally, when you come to the number of which it is a square that number will come only once in the list of factors.

(Refer Slide Time: 13:45)



So, once we talk about factors, we come to a very interesting class of numbers which are the prime numbers. So, a prime number is one which has no factors other than 1 and itself. So, 1 is a factor always and 1 × n is n. So, for we try to usually write p for a prime number. So, a prime number has only two factors 1 and p.

Now, it is important that it must have two factors, two separate factors. So, one technically is not a prime because it has only one factor one itself because 1×1 is 1 and so, the only factor that 1 has is 1. So, the smallest prime actually is 2 because it has two factors 1 and itself 2 and no other factors. 3 is also a prime because it has only 2 factors 1 and 3, 2 does not go into 3 and so on.

So, we are all familiar with the smaller prime numbers. So, 2 is the first prime number, 3 is the next prime number, then 5, then 7. Notice that, after 2 no even numbers can be primes because they are all multiples of 2 and so, 2 divides them. Now we come to 9 and 9 is not a prime number because it is a multiple of 3, but 11 is a prime number and so on.

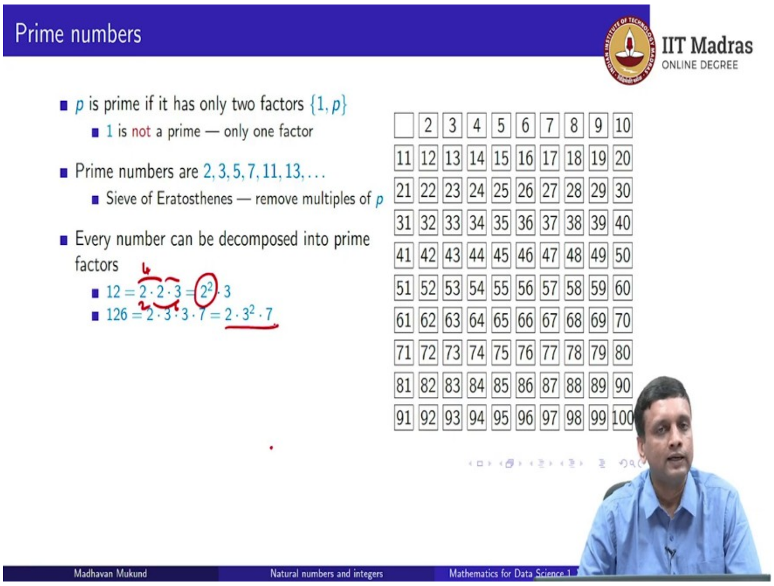
So, there is actually one clever way which is call the sieve of Eratosthenes to generate prime numbers which is whenever you discover a prime, you knock off all the numbers which are multiples of it. So, we can do this for instance to get all the prime numbers from 1 to 100. So, what we do is we first lay out a grid like this right, we know that 1 is not a prime so, the first prime that we have as a candidate is 2 right. So, this is how the sieve of Eratosthenes works, you lay out the numbers in a grid and now we can try and mark off all the prime numbers which are up to 100.

So, we know that 1 is not a prime so, we leave 1 off the grid and we start with 2. So, 2 is our first prime number and what the sieve of Eratosthenes says is you knock off all multiples of 2. So, you knock off all the even numbers and of course, now you can do it in one shot so, you can knock off this whole column, this whole column so, all these numbers are not prime ok.

So, now once you have you have a target so, we are looking only up to 100. So, up to 100 we have knocked off all the powers of 2 or all the multiples of 2. So, now, we look at the first number which is not been marked off and we notice that 3 is a prime because 3 is not yet marked off. So, now, we start mark off multiples of 3, some of them are already marked off because they are multiples of 2. So, 6 is already gone, but 9 is also gone, 12 is already gone, but 15 is also gone and so on.

So, we can mark off all the other multiples of 2 which are not multiples of 3 and so on right. So, we get this kind of a picture and now having done this assuming we have done it all the way, then we will come and find that 5 is a prime right. So, this is the process by which if you want to know count all the primes up to a certain number n, you can write out all the numbers up to n and starting at the left you can take the first unmarked number, call it a prime and mark all its multiples to the right as non primes and the next unmarked number will be the next prime and so on.

(Refer Slide Time: 16:56)

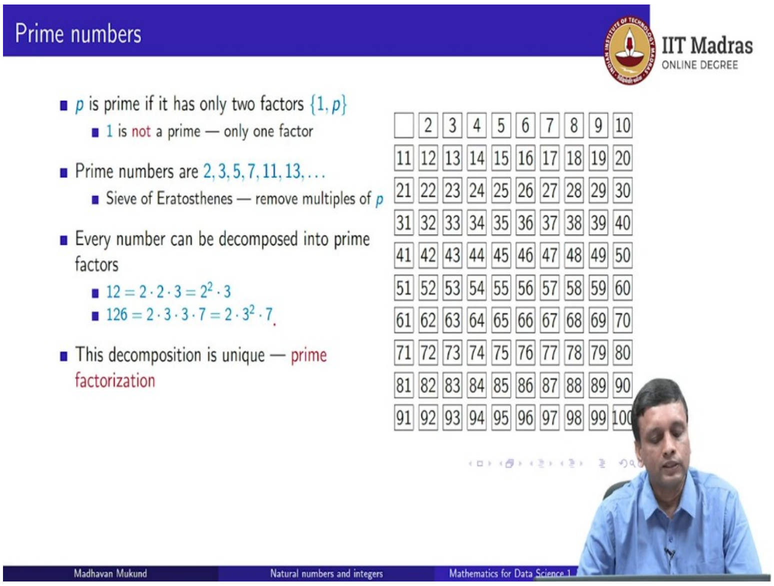


Now, this is not necessarily an efficient way to do the prime numbers, but this is a good way to generate them without missing out any. One of the important facts that we use all the time is that every number can not only be factorized as we have seen into a number of different pairs of factors it can actually factorize uniquely into the prime numbers that form it.

So, for instance if we look at 12, we said that 12 was 2 times 6, it was also 4 times 3, it was 1 times 12 and so on, but fundamentally it has 3 prime factors 2 2 again and 3. So, depending on how we combine them for instance, we can get 4 × 3 or we can get 2 × 6 and so on, but 2 × 2 × 3 is the absolute unique way of writing 12 as a product of prime numbers and using our exponentiation notation, we can condense this and put the 2 2's together and say it is $2^2 \times 3$.

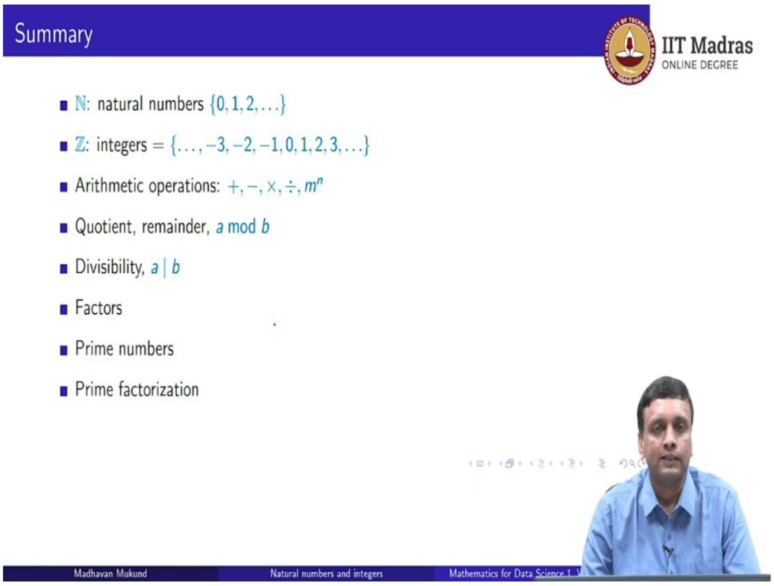
Similarly, if we take a number like 126, then it is 2 × 3, 6 × 3, 18 ×7 ok. So, the prime factors are precisely 2 3 twice and 7 and we can write this as $2 \times 3^2 \times 7$. So, this is very important because we use it implicitly along a lot and we will see later how we use this.

(Refer Slide Time: 18:10)



So, this is called the prime factorization right. So, every integer can be decomposed into a product of primes in a unique way.

(Refer Slide Time: 18:20)



So, to summarize we started with a natural numbers which we use for counting which are the numbers 0, 1, 2, 3, 4 and so on. Then, we extended these numbers with a negative numbers and this gave us the set of integers. So, the integers include all the natural numbers as well as the negative numbers 0, 1, 2, 3 and so on -1, -2, -3 and so on. We saw some basic arithmetic operations on these the usual addition, subtraction, multiplication, division and exponentiation.

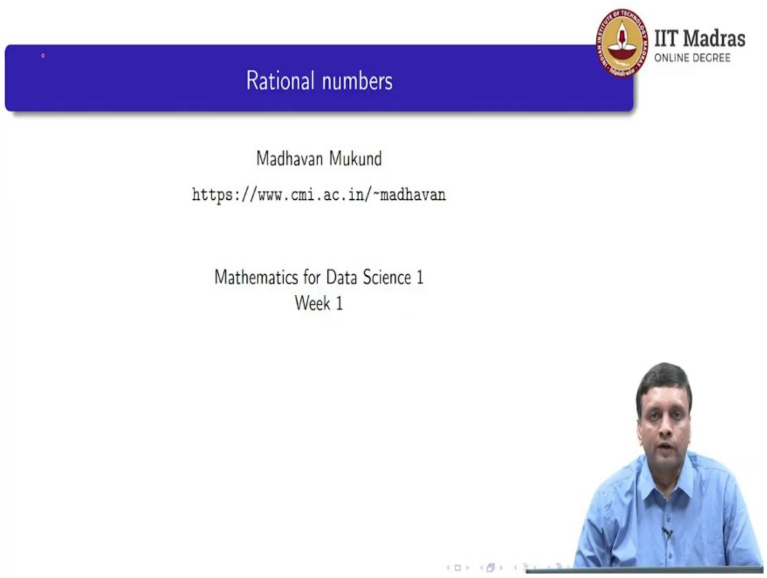
We also looked at what happens when we divide integers and we do not want to look at fractions, then we talk about the quotient which is the integer number of times that the dividend goes into the number and the remainder is also written as a mod b. So, using this notation of a mod b, we can talk about divisibility which we write with a vertical bar. So, a | b if a mod b is 0. So, the factors of a number are those numbers which divide it and a prime number has exactly two factors 1 and itself and we can always decompose any integer uniquely into the list of factors, prime factors that multiply out to form that number.

Mathematics for Data Science 1
Prof. Madhavan Mukund
Department of Computer Science
Chennai Mathematical Institute

Week - 01
Lecture - 02
Rational Number

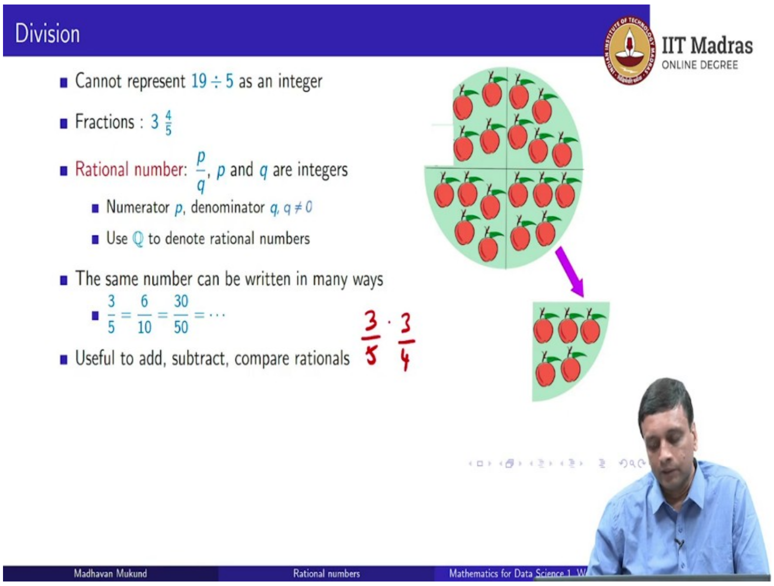
(Refer Slide Time: 00:06)

(Refer Slide Time: 00:14)



So, now first lecture on Numbers; we looked at natural numbers and integers. So, now, let see what happens when we try to divide. So, let us look at the rational numbers.

(Refer Slide Time: 00:23)



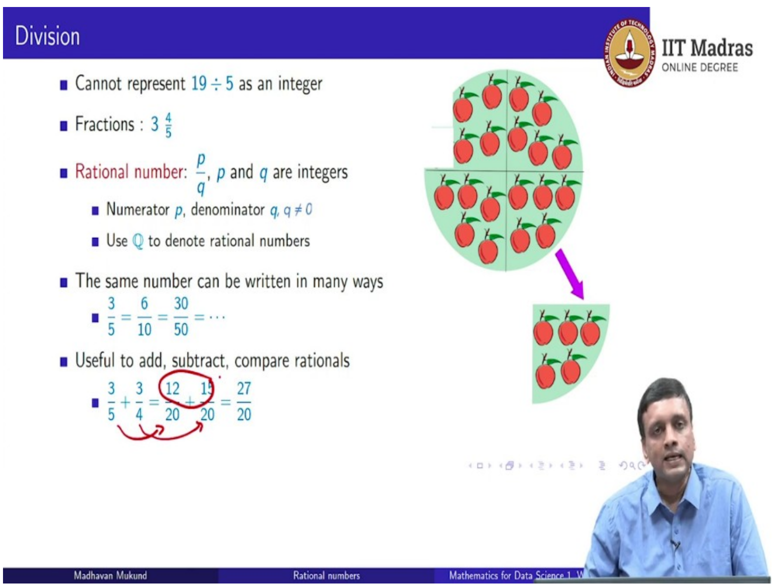
So, we said that we cannot represent 19 / 5 as an integer because we cannot find a number k such that 5 × k is 19. So, as we know the way we deal with this is to represent this quantity as a fraction. So, we say that 19 /5 is $3\frac{4}{5}$. So, this number is an example of a rational number. So, rational number what we usually called fractions in school, a rational number is something that can be written as ,$\frac{p}{q}$; where, p and q are both integers. So, as you probably remember from school, the number on the top is called the numerator. So for,$\frac{p}{q}$; p is called the numerator and q is called the denominator.

So, just like we had the symbols N and Z to represent the natural numbers and the integers, we have a special symbol which is somewhat unusual which is Q. So, Q stands for the rational numbers and again, to just say it is a special Q, we write these double lines on sides. So, this Q with these fat boundaries denotes the rational numbers. So, one thing about the rational numbers is that the same number can be written in many different ways. Now, this is not true of integers. Of course, we are not talking about changing base from binary to decimal or something.

But if you write a 7, there is only one way to write 7 fix, if you are fix the notation that you are using for writing numbers. With rational numbers, this is not true because there are many ways of writing $\frac{p}{q}$such that $\frac{p}{q}$is actually a same number. So, for instance if we take the number$\frac{3}{5}$, then we all know that $\frac{3}{5}$is the same as $\frac{6}{10}$and this is the same as $\frac{30}{50}$. So, when we take a rational number and multiply it by something the same quantity on the top and the bottom so, $\frac{3}{5}$, 3×2 and 5×2, we get the same number; $\frac{6}{10}$ or 3×10 and 5×10, we get the same number $\frac{30}{50}$. So, this is sometimes a nuisance, but it is also sometimes useful.

Now, there is no reasonable way to compare two numbers like say $\frac{3}{5}$and $\frac{3}{4}$ or $\frac{2}{5}$ and $\frac{3}{4}$. If we have two fractions which have different denominators, there is no way to directly compare them. So, the only way to compare them is to somehow convert them into equivalent fractions such that they have the same denominator. So, the usual way is just to find a number such that both the denominators multiply into that number rather factors of that number. Now, you can find the smallest such number which is called the least common multiple; but you can find any number of this form.

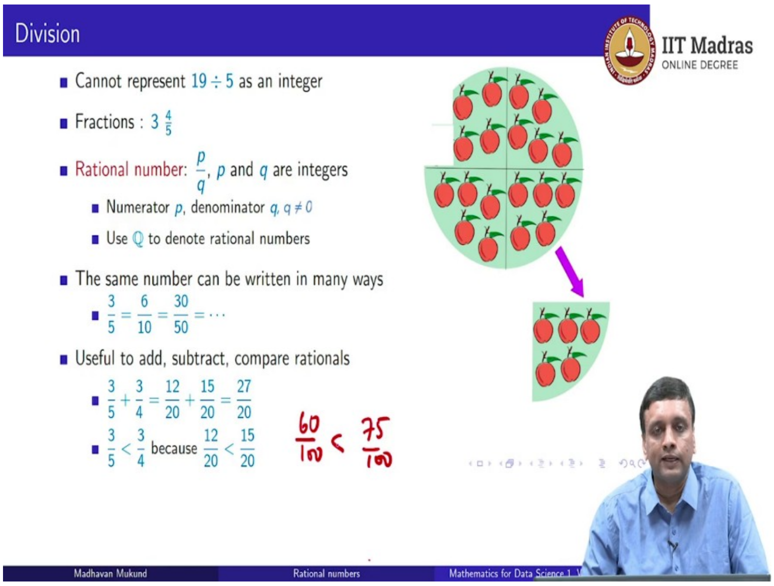
(Refer Slide Time: 03:05)



So, for instance, if you want to add $\frac{3}{5}$ and $\frac{3}{4}$, now you cannot do that directly; but you know that 20 is a number which divides which is a multiple of both 5 and 4. So, you can represent $\frac{3}{5}$ as equivalently as $\frac{12}{20}$; you can represent $\frac{3}{4}$ equivalent. So, this is equivalent and this is equivalent. So, you have converted these numbers into a different fraction of the same number; but this new representation has the same denominator.

And now once, the two denominator that the same, you can add the numerators and you can get (12 + 15)/20 is $\frac{27}{20}$. So, this kind of manipulation requires the denominators to be the same and therefore, it is actually extremely useful that we can write the same rational number in many different ways. The same is to we want to compare two numbers.

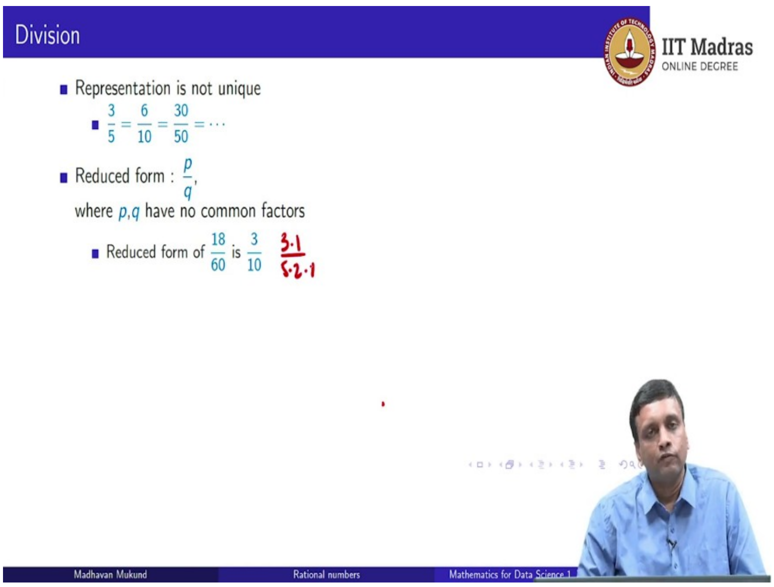
(Refer Slide Time: 03:54)



If we want to check whether $\frac{3}{5}$is bigger or smaller than $\frac{3}{4}$, there is no way to do it directly. What we have to do is again take the denominators and make them the same and then, say that $\frac{12}{20}$ is less than $\frac{15}{20}$ because you are dividing something 20 parts and you are taking 12 of them that is less than taking 15. Now, as I said there is no reason why this must be the smallest one. So, for instance you could take a bigger number like 100, right. So, 5 goes into 100 and 5 goes 4 also goes into 100.

So, we could also say that $\frac{3}{5}$ is the same as $\frac{60}{100}$ and, $\frac{3}{4}$ is the same as $\frac{75}{100}$ and therefore, since 60 is less than 75; $\frac{60}{100}$ is less than $\frac{75}{100}$ and therefore, $\frac{3}{5}$ is less than $\frac{3}{4}$. So, it is not really important that the denominator is the smallest common multiple of the two denominators; but it must be some common multiple so that you can bring it all to a common number that you can then compare.

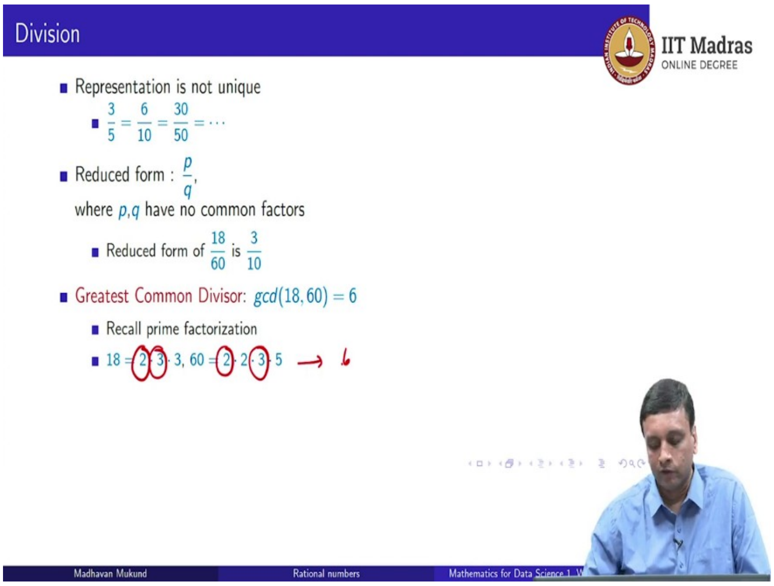
(Refer Slide Time: 04:49)



So, we saw that representation is not unique for rational numbers. So, how do we find actually the best way to represent a rational number? So, normally if you are not using it for some arithmetic operation of some comparison, we would prefer to have it in a reduced form. So, the reduced form of a rational number is one, where there are no common factors between the top and the bottom. So, $\frac{p}{q}$ is of the form, where we cannot find any factor f such that f | p and f | q.

So, for instance, if we take $\frac{18}{60}$, then its reduced form will be $\frac{3}{10}$. Notice that 3 is of the form 3 × 1 and 10 is of the form 5 × 2 × 1. So, therefore, there is no common factor between the top and the bottom and therefore, this is in reduced form.

(Refer Slide Time: 05:42)

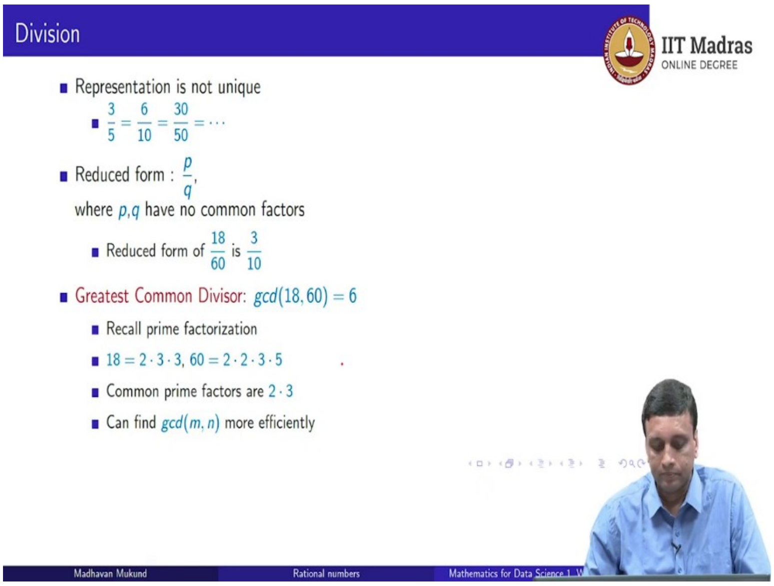


So, this is called the greatest common divisor problem. So, we want to find the largest number which divides both the top and the bottom; both the numerator and the denominator; divide them both by this and then come to something in the reduced form. So, in this case, what we are saying is that the gcd of 18 and 60 is actually 6 and we can do this using our prime factorization that we talked about before.

So, if we look at prime factorization for 18, then 18 is 2 × 3 × 3 right; its 2 × 3 is 6 and 6 × 3 is 18 and the prime factorization of 60 is 2 × 2 ×3 × 5; its 4 × 3, 12 and 12 × 5. So, now, you can look at what are common. So, we have one 2 here and one 2 here. So, we can say that this is part of the same factor, we have one 3 here and another 3 there. The second 2 is not present in the first term.

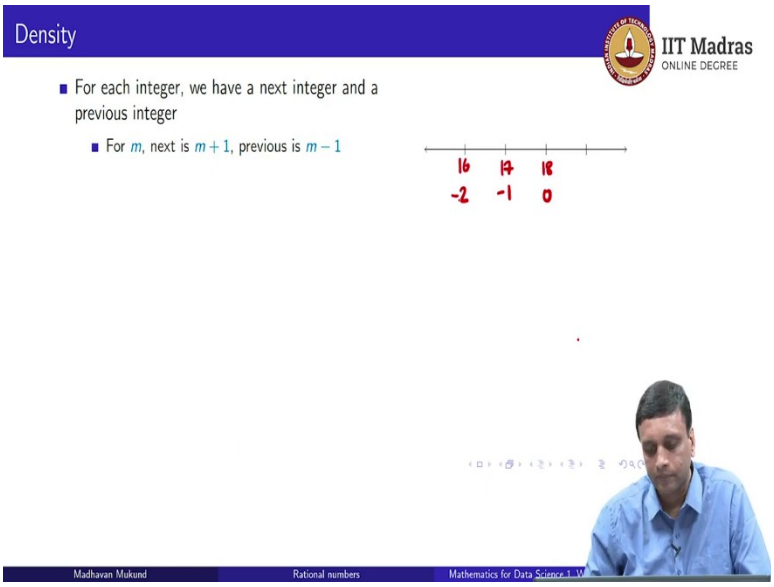
So, we have a 2 and 3 and 18 which are factors. We have a 2 and 3 in 60 which are factors and this gives us the fact that 6 is a common factor. There is no bigger common factor because we want to assemble a bigger common factor, we have to pull out one more prime from each side; but there is no prime left which is present on both sides. 3 is there in 18; 2 and 5 are there on 60, but we do not have a matching one of the other side right.

(Refer Slide Time: 06:59)



So, this way, the common prime factors are one 2 and one 3 and so, 2 × 3 equal to 6 is the gcd. Now, this is not the best way to find the gcd, there are more efficient ways to find the gcd. But this intuitively tells us what the gcd is. You take the prime factorization of both the numbers and you collect together all the primes that occur in both the numbers, the same number of times.

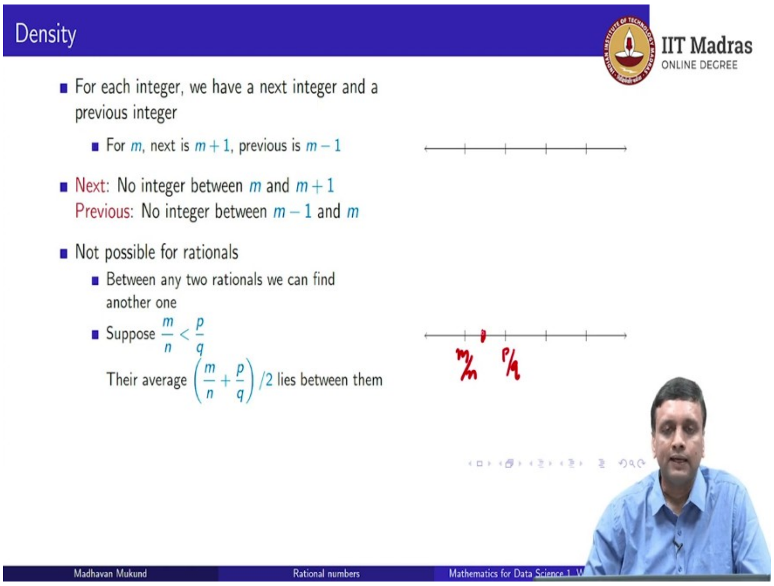
(Refer Slide Time: 07:23)



So, here is another interesting property about rational numbers. Now, for each integer, we know intuitively that there is something which is the next integer and the previous integer. If

I tell you 22 and ask you what is the next integer? Then, you will know it is 23. What is the previous one? It will be 21. So, for every integer m, the next one is m + 1 and the previous one is m - 1 and it does not matter, if this is positive or negative. So, for instance if I am at 17, then the next integer is 18, the previous one is 16; right. If I am at -1, then the next integer is 0 and the previous integer is -2. So, I can always take the integer that I am at, add 1 and get the next integer, subtract 1 you will get the previous integer.

(Refer Slide Time: 08:10)



So, the property of this next and previous is that there is nothing in between right. So, there is no integer between m and m +1, there is no integer between m and m -1. So, that is what next means, it is not some bigger integer or some smaller integer. It is the immediate neighbor in the integer of the in this number line. Now, what about rationals? Is it possible to talk about the next and the previous rational number? Now, it turns out that this is not possible for a very simple reason.

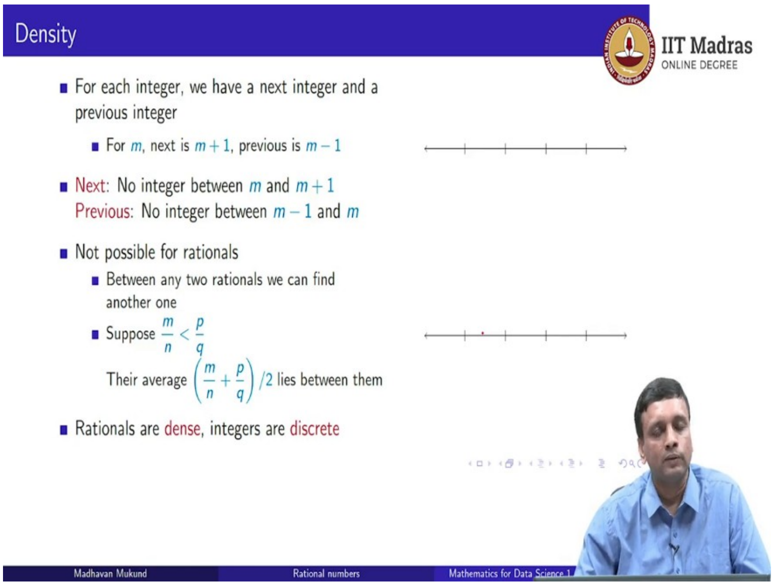
So, between any two rationals, we can always find another one because we can always take the average of 2 numbers. So, remember that if you take the average of any 2 numbers, then it must be between those 2 numbers right because it is the sum of the numbers divided by 2. So, the average cannot be smaller than both or cannot bigger than both. So, if the 2 numbers are not the same, then it must lie strictly between them. If the numbers are the same, then the average is the same.

So, if somebody has 37 marks and 37 marks, then their average marks is 37. But if they have 37 marks and 52 marks, even without calculating the average, you know that their average is bigger than 37, but smaller than 52; right. So, in the same way, if I give you 2 fractions $\frac{m}{n}$ and $\frac{p}{q}$ and I tell you that $\frac{m}{n}$is smaller than $\frac{p}{q}$. Remember that in order to do this, we would have to normally get the denominators to be the same and so on.

But supposing I know that $\frac{m}{n}$ is smaller than $\frac{p}{q}$. So, I know that say $\frac{m}{n}$is here and I know that say $\frac{p}{q}$is here and supposing you claim that $\frac{m}{n}$ and $\frac{p}{q}$ are adjacent, that is $\frac{p}{q}$is the next rational after $\frac{m}{n}$. Well, I will say no; let me take these 2 numbers and find its average right. So, this average now is also a rational number because you can also represent it as $\frac{a}{b}$ right. If you just workout this $\frac{m}{n}$ plus $\frac{p}{q}$ divided by 2, you can simplify this whole expression and you will get a new number which is also of the form $\frac{a}{b}$.

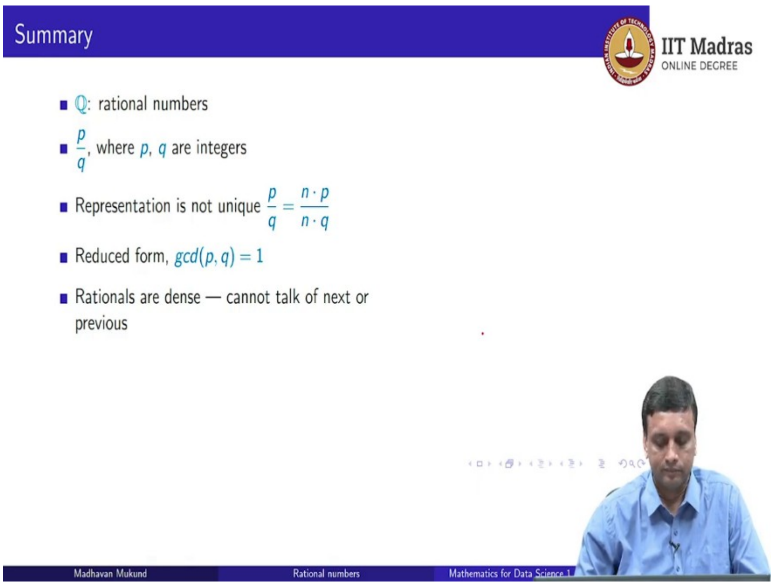
So, this is also a rational number and this rational number as we argued must be between the 2 numbers and therefore, between any 2 rational numbers by just taking the average of the mean of the 2 numbers, I can find another one.

(Refer Slide Time: 10:17)



So, in other words, the rational numbers are dense right. So, dense in the usual sense, so dense just means that they are closely packed together. So, basically you cannot find any gaps in the rational numbers because any between any 2 rational numbers, you will find another rational number and this is not true of the integers because we saw that in the number line, there is a gap between m and m + 1, there is no integer there right. So, we say that the rational numbers are dense and conversely, we say that the integers and the natural numbers are discrete. So, a discrete set has this kind of next property and a dense set has no next property between any 2 numbers, will find another number right.

(Refer Slide Time: 10:56)



To summarize, we use this funny symbol Q to denote the rational numbers and a rational number is just the ratio. So, that is where it comes from actually; so, ratio. So, rational number comes from the word ratio and so, it is a ratio of 2 integers p divided by q. Now, there is no unique representation of a rational number because we can multiply both the numerator and the denominator by the same quantity and get a new rational number which is exactly the same in terms of the quantity that it represents.

And we use this fact for things like arithmetic and comparisons, but if we really want to talk about rational numbers in a canonical way, in a unique way; then, we get this reduced form, where we cancel out the common factors using prime factorization. So, that we get a number whose gcd of the numerator and the denominator is 1.

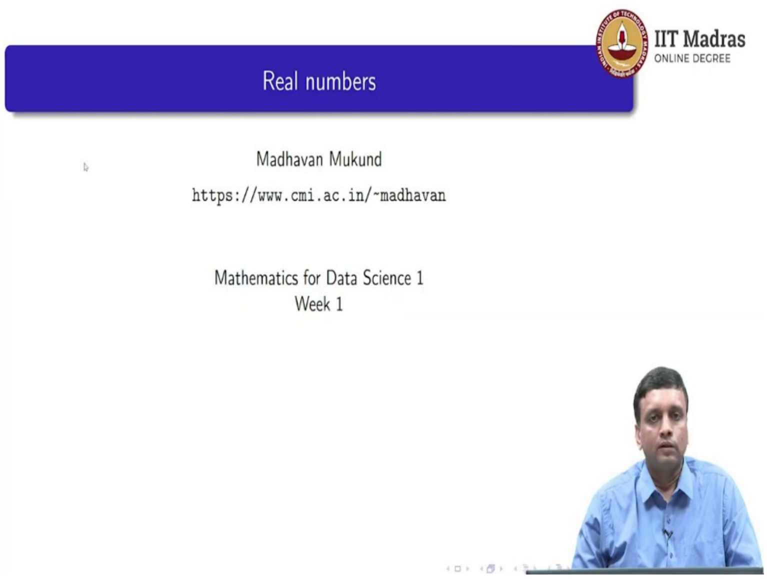
And finally, we saw that we cannot talk about the next or the previous rational number because between any 2 rational numbers, there is another rational number. In particular, if you take the average of the 2 numbers, you will find a number that is in between. So, unlike the integers and the natural numbers which are discrete for which next and previous makes sense; for the rational numbers, there is no such quantity.

Mathematics for Data Science 1
Prof. Madhavan Mukund
Department of Computer Science
Chennai Mathematical Institute

Week - 01
Lecture - 03
Real and Complex Number

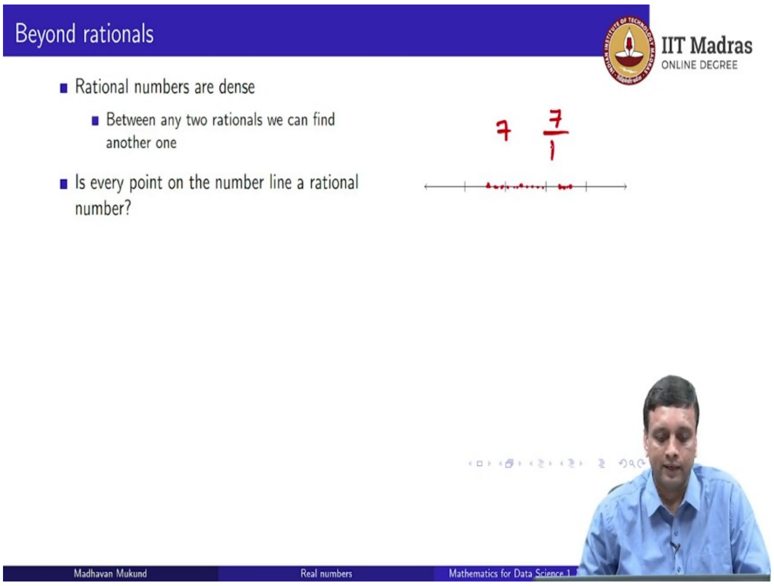
(Refer Slide Time: 00:06)

(Refer Slide Time: 00:14)



So, we started with the natural numbers and the integers and then, we moved on to the rational numbers which are defined as $\frac{p}{q}$; where p and q are both integers.

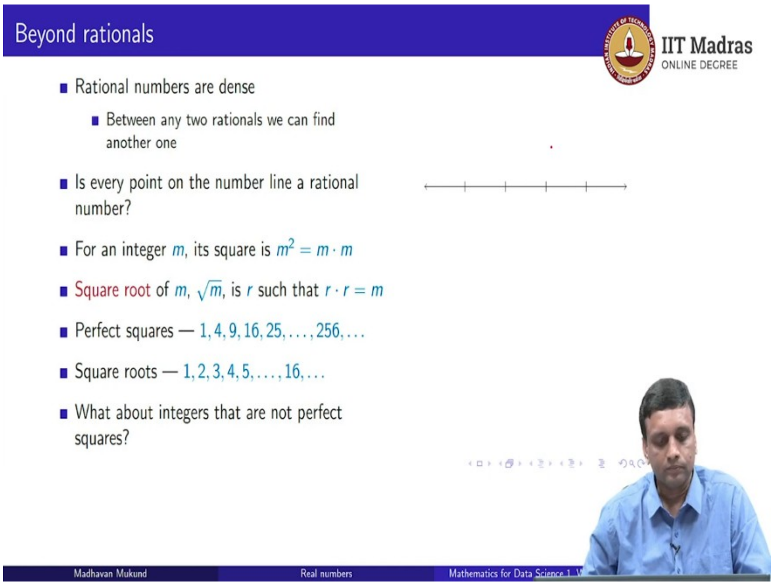
(Refer Slide Time: 00:23)



So, we decided that the rational numbers are dense right and that means that on this number line between any two rationals, you can find a rational. So, if I want to now talk about this number line, then I know that if I take any two positions, then I will find a rational between them and I will find a rational between them and so on. So, it makes sense to ask this question which is that if I take any two points and the rational between them any two points, then is this entire number line composed only of rational numbers. Of course, some of those rational numbers are integers.

So, an integer is a rational number because I can write 7; for instance, as $\frac{7}{1}$right. So, this is of the form $\frac{p}{q}$. So, any rational number which in reduced form as denominator 1 is an integer; so, an integer is a special case of a rational number. So, do all the rational numbers fill up this number line? That is the question.

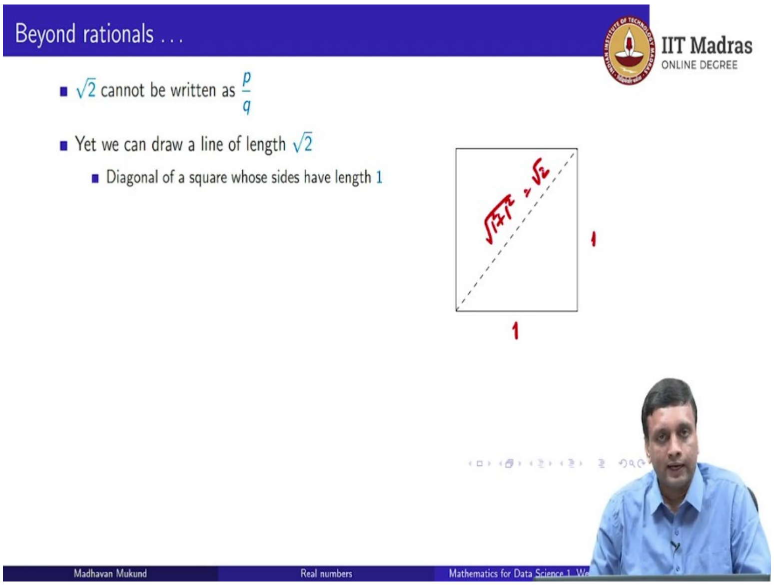
(Refer Slide Time: 01:12)



So, it turns out this is not the case. So, remember that a square of a number is the number multiplied by itself. So, if I take a number m and multiply it by itself, I get m^2 which is m × m and if I take this operation and turn it around, then the square root of a number is that number r such that r × r is equal to m right. So, I want to find out which number, I have to square in order to get m and that is called the square root.

So, if we take the so called perfect squares, like 1, 4, 9, 16, 25 and so on their square roots are integers. So, 1^2 is 1. So, the √1 is 1; $2^2$is 4, so the √4 is 2; $5^2$is 25, so √25 is 5; $16^2$is 256, so √256 is 16 and so on. So, some integers are clearly squares of other integers and so, you can get the square root and find an integer. Now, what happens if something is not a square right? So, supposing I take a number which is not a square like 10 and I take its square root, I know that the square root is not an integer, its somewhere between 3 and 4 because $3^2$is 9 and $4^2$is 16. Question is, is it a rational number or not?

(Refer Slide Time: 02:36)

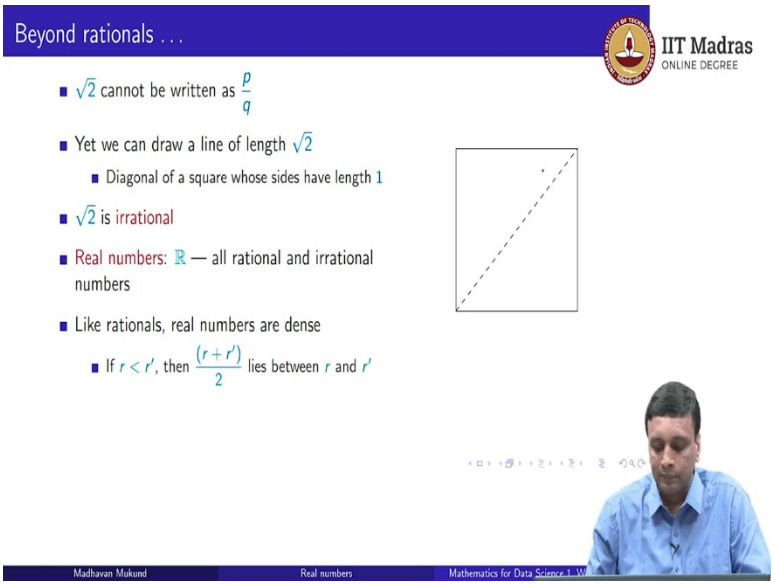


So, what happens to the square roots of integers that are not perfect squares? So, the smallest such number which is not a perfect square because 1 remember is a perfect square, 1 × 1 is 1. The smallest such number that is not a perfect square is actually 2 and it is one of the very old results that the $\sqrt{2}$ cannot be written as $\frac{p}{q}$. This was certainly known to the ancient Greeks, in fact, to Pythagoras and one way to do this is to see that you can actually draw a line of; so, this is not an unreal number in that sense right.

So, you can actually draw a line of this length because if you take a square, whose sides are 1 right. So, this is 1, then if you remember your Pythagoras theorem; then, the hypotenuse of this triangle is going to be $\sqrt{1^2+1^2}$,technically which is $\sqrt{2}$. So, I can actually physically draw a line whose length is $\sqrt{2}$. So, this is a very real quantity.

On the other hand, for reasons that we will not described here, but there will be a separate lecture explaining this for if you are interested. $\sqrt{2}$ cannot be written as a rational number $\frac{p}{q}$. So, here is a number which is a very measurable quantity, I can actually draw this quantity as a length. At the same time, it does not fit into this number line of rational numbers which seems to cover all the rational numbers, all the numbers because they are dense.

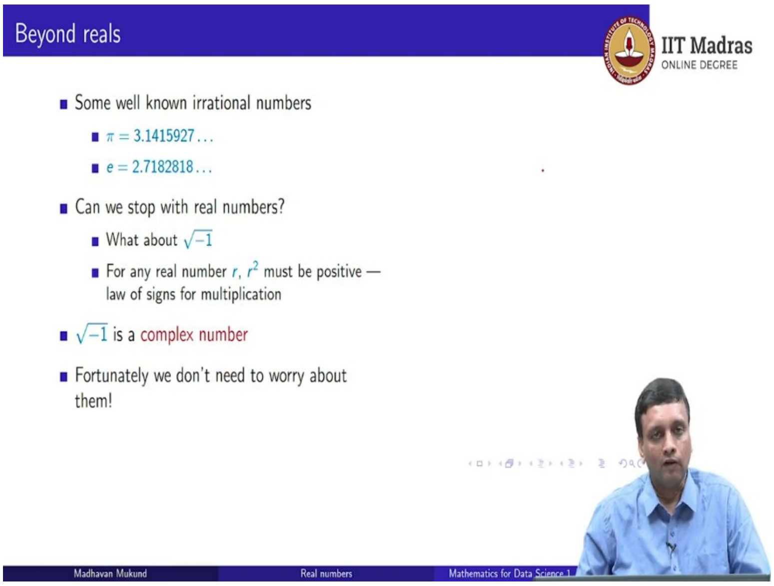
(Refer Slide Time: 03:56)



So, √2, since it is not a rational number right and yet it exists is called an irrational number and these numbers which constitute all the rational numbers and the real irrational numbers together are called the real numbers. So, the real numbers are denoted by this double line R. So, we had N for the natural numbers, Z for the integers, Q for the rational numbers and now, we have the real numbers R.

So, the real numbers extend the rational numbers by these so called irrational numbers which are very much on the number line, but which cannot be written on the form $\frac{p}{q}$. Now, it is not difficult to argue that like the rationals, the real numbers are dense for the very same reason. Because if you have two real numbers r and r' such that r is smaller than r prime, then you can just take their average r + r' divided by 2. This must be a number πwhich is bigger than r and it is smaller than r' and therefore, it must lie between them. So, between any 2 real numbers, you will find another real number. So, the real numbers are also dense.

(Refer Slide Time: 05:00)



So, there are some irrational numbers which we use a lot in mathematics and which you have probably come across; one of them is this famous number π which comes when we are talking about circles. Because it is the ratio of the circumference to the diameter and this is an invariant. π is always; the circumference divided by diameter for any circle is π ok.

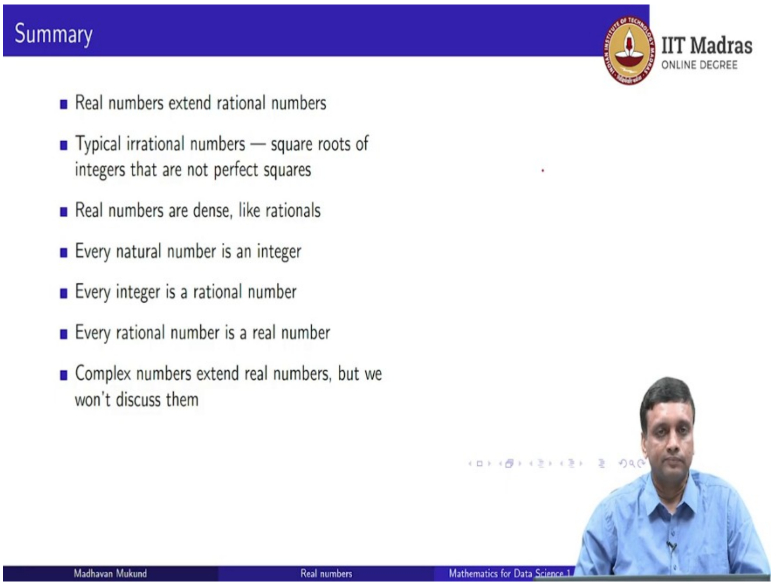
So, π is an irrational number. We cannot write it in the form $\frac{p}{q}$ and it has this. If you write it in this decimal form, it has this infinite decimal expansion. Another number which is very popular as an irrational number is this number e which is used for natural logarithms. So, it is 2.7182818 and so on right. So, there are a lot of irrational numbers. So, $\sqrt{2}$ as we have seen as an irrational number. It will turn out that square root of anything, $\sqrt{3}$ is also an irrational number, $\sqrt{6}$ is also an irrational number.

Anything which is not a perfect square, its square root is actually an irrational number. But many of these numbers are not very useful to us, but π and e are certainly very useful irrational numbers. So, now, we have seen that we can find more numbers on the line than just the rationals and these are the real numbers. So, do we stop here? Well, let us look at the square root operation which we use in order to claim that there are irrational numbers. So, what happens if we now take the square root of a negative number like -1?

So, remember that we had a sign rule for multiplication. The sign rule for multiplication said that if I multiply any two numbers, then if the two signs are the same that is their two negative signs or two positive signs, I will get a positive sign in the answer. Only if the two signs are different, if I have one minus sign and one plus sign, will I get a negative answer. So, if I want to multiply two numbers and get a -1, one of them must be negative and one must not be negative. But by definition, a square root is a number which is multiplied by itself, the same number has to be multiplied by itself. So, it will have the same sign.

So, any square root which multiplies by itself must give me a positive number. So, if I take a negative number, there is no way to find a square root for it. So, if we want to find square roots for negative numbers, we have to create yet another class of numbers called complex numbers. So, complex numbers extend the real numbers, just like real numbers extend the rational numbers and rational numbers extend the integers and so on. But the good news for you is that we do not have to look at complex numbers for this course.

(Refer Slide Time: 07:35)



So, to summarize, a real numbers extend the rational numbers by adding the so called irrational numbers which cannot be represented of the form $\frac{p}{q}$ and a typical example of an irrational number is the square root of an integer that is not a perfect square. So, $\sqrt{2}$ for example is not a rational number and this is also of the case was $\sqrt{3}$, $\sqrt{5}$, $\sqrt{6}$ and so on. So, except for the perfect squares, none of the square roots are actually rational numbers. Now, just like we said that the rational numbers are dense because the average of any two rational numbers is a rational number. Similarly, the real numbers are dense because the average of any two real numbers is a real number.

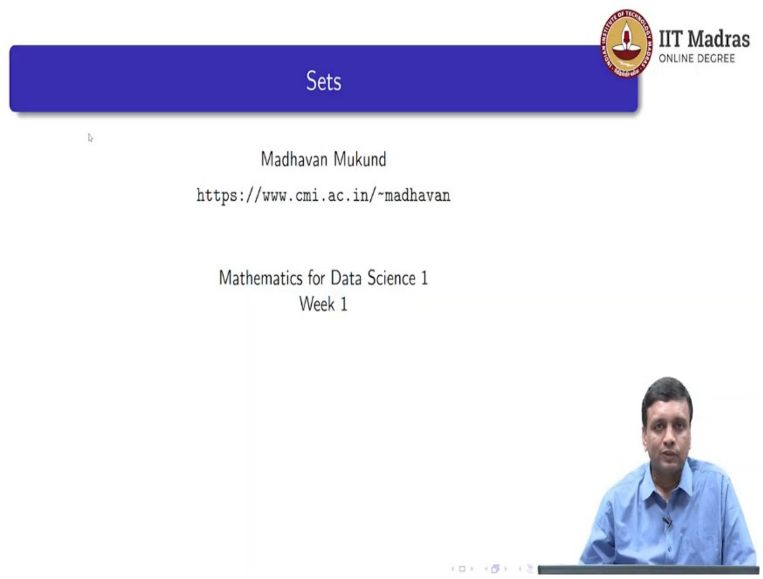
So, we have a progression in terms of numbers. So, every natural number that we started with is also an integer because the integers extend the natural numbers with negative quantities. Now, every integer is also a rational number because we can think of every integer as a ratio $\frac{p}{q}$; where, the denominator is 1. And finally, every rational number is a real number because we said that the real numbers include all the rational numbers plus all the irrational numbers. And finally, we said that there are even things beyond rational numbers like complex numbers, but we will not discuss them.

Mathematics for Data Science 1
Prof. Madhavan Mukund
Department of Computer Science
Chennai Mathematical Institute

Week - 01
Lecture - 04
Set Theory

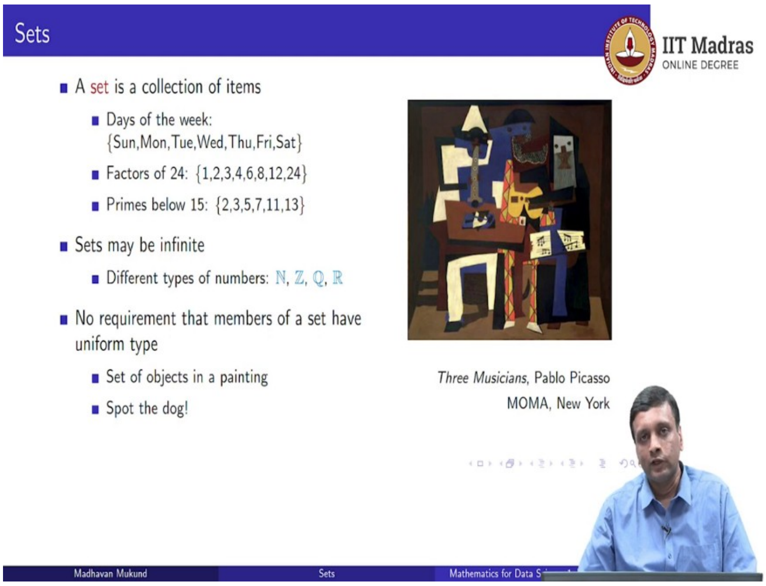
(Refer Slide Time: 00:06)

(Refer Slide Time: 00:14)



So, we have seen numbers; we have seen natural numbers, we have seen integers, rationals, reals and we have loosely talked of them as sets of numbers. So, let us try to understand little more clearly what we mean by a set.

(Refer Slide Time: 00:26)



So, at its basic level a set is a collection of items. So, for instance, we could have a set called the days of the week which has 7 members; Sunday, Monday, Tuesday, Wednesday, Thursday, Friday and Saturday or we could take a number like 24 and list out the factors of 24 and call this a set. So, we have 1, 2, 3, 4, 6, 8, 12 and 24.

So, if you count, there are 8 factors that 24 has or we could take all the prime numbers up to a certain limit. Supposing, we want to know the prime numbers below 15, then we know that we do not have 1; but 2, 3, 5, 7 are the single digit prime numbers and then, 7, 11 and 13.

So, 2, 3, 5, 7, 11, 13 are all the primes below 15. So, this is how we talk about sets informally as just collections of items. Of course, as we have seen sets can be infinite and in particular, the infinite sets that we deal with very commonly are those which consists of the different types of numbers.

Remember that N this funny N stands for the natural numbers that is 0, 1, 2, 3, 4. Z stands for the integers. So, that is the natural numbers along with the negative integers like -1, -2, -3 and so on. Q is a peculiar symbol for the rational numbers, these are the fractions those numbers which we can write as $\frac{p}{q}$; where, p and q are both integers.

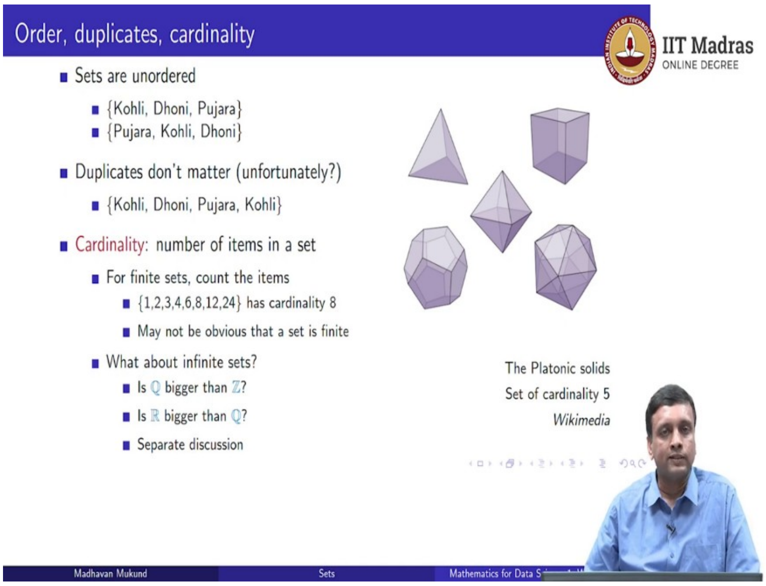
And finally, R is a set of real numbers. So, the real numbers includes all the rationals all the fractions, but also numbers that cannot be represented as fractions, such as the square root of 2 and other irrational numbers like π and e.

So, in all these things that we have seen above, it looks like there is some kind of condition which requires a set to have some uniformity; either a set consists of numbers or a set consists of days of the week or something like that. But actually mathematically there is no constraint on this. A set can have any kind of members, even a mixed membership; there is no uniformity of type.

So, for instance, we could enumerate the set of objects that appear in a painting. Now, here is a particularly famous painting, where it is not so easy to enumerate the objects because its drawn in a very abstract way. This is a painting called Three Musicians by Pablo Picasso. But we could see roughly that there are three people and that there are some musical instruments and so on and if you look very carefully, you will even find a dog.

So, notice that there is no commonality. There are people, there are musical instruments, there are chairs, there are tables, there are animals and so on. So, a set in particular can have any kind of members, it does not matter if they are mixed in type.

(Refer Slide Time: 03:00)



So, one of the important differences between say set and a sequence or a list is that the order in which we identify a set does not matter. So, normally when we talk of numbers, we tend to list them in a particular way; but as the set it does not matter. So, for instance, if you take the set of cricketers; Kohli, Dhoni and Pujara. If you reorder this set as Pujara, Kohli and Dhoni, it is the same set right.

So, the sequence in which you list the members of a set does not matter and for that matter, if you happened to accidentally write the same member twice, it does not change the set. So, in this particular set if we add Kohli a second time, as the set it does not matter. Though of course, if you are a cricket fan maybe you would like Kohli to bat twice for you.

So, when we look at a set, we might ask a basic question as to how many members it has. So, the cardinality of a set is the number of items in the set and if it is a finite set, we can just count the items. So, for instance, if you look at the factors that we listed of 24, then we can count them and say that this has cardinality 8.

Sometimes, it may not be obvious that a set is finite. You might remember from geometry that a regular polygon is one, where the all the sides are equal and all the angles are equal. So, the smallest regular polygon is an equilateral triangle in which we have 3 sides all equal and 3 internal angles of 60 degrees each. Then, we move to four sides we get a square, then we get regular pentagons, hexagons, heptagons, octagons and so on.

So, for any number of sides, you can draw a regular polygon with that many sides with equal angles on the inside. So, there is no limit. The set of regular polygons is infinite. But if we move to three dimensions, the corresponding notion to a regular polygon is what is called a platonic solid. In a platonic solid, first of all you have surfaces or sides each side is a regular polygon and all these regular polygons meet at the same angle in three dimensions.

Now, it turns out that though you might imagine that there are infinitely many regular polygons in two dimensions, there are only 5 platonic solids in three dimensions. So, this is an example of a set which turns out to be finite, even though there is no reason for it to be finite. So, these 5 platonic solids are the tetrahedron which has triangles.

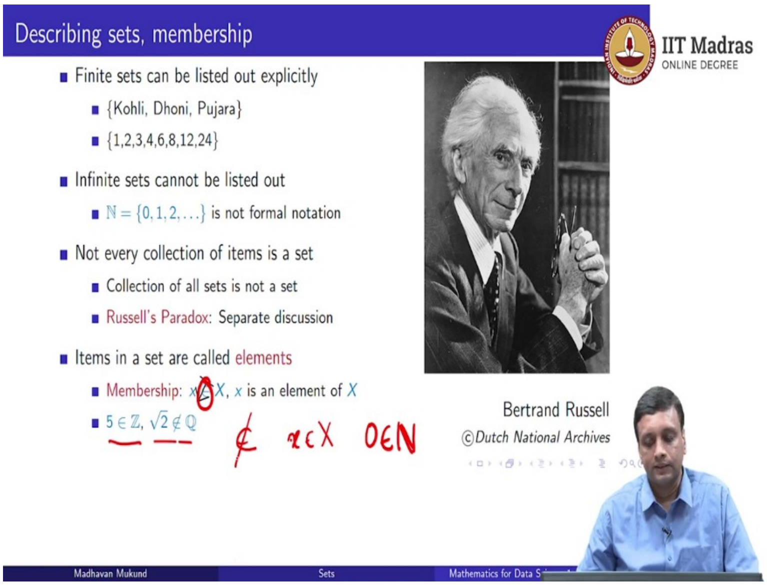
The cube which we have which has squares and then, we have an octahedron which has 8 sides which are triangles. Then, we have a dodecahedron with 12 sides and an icosahedron with 20 sides and there are no other regular solids, surprisingly it turns out.

Now, cardinality is quite easy to determine for a finite set, but what about for an infinite set? Remember that, we said that we wanted to go from integers to rational numbers because we want to talk about what happens when we divide 2 integers and the answer is not an integer and it is clear to us from our discussion that integers were discrete, we can talk about a next number and a previous number. So, there are gaps in the integers and rational were dense; between any 2 rational numbers, there is another rational number.

So, intuitively, it seems like we are adding things to the integers to get rational numbers. But can we make it formal in terms of cardinality? Are there more rational numbers than there are integers? And what happens, when we go from Q to R when we go from rational numbers to real numbers? So, remember that the real numbers, we had introduced because they were numbers such as the $\sqrt{2}$ which could not be represented as a fraction.

So, clearly there are some rational numbers which are real and some real numbers which are not rational and therefore, we have a bigger set; but again, R is really bigger than Q . So, this is a separate discussion, there will be a small separate lecture about this. But there is a way to measure cardinality of infinite sets, but it is not as straight forward as it is for finite set as you would imagine.

(Refer Slide Time: 06:58)



So, how do we describe a set? Well, we have already seen that for a finite set, we can just list out the members of the set explicitly. So, we can write out 3 numbers; Kohli, Dhoni, Pujara or 8 members the factors of 24. So, the normal notation for a list of items which form a set is to use these curly braces and to separate the items by commas.

Now, in many books and even in our lectures we will see notation like 0, 1, 2 …indicating that there is an infinite set of elements to be added which follows some kind of a pattern. So, this looks a way of listing out an infinite set, but you must understand that this is only an informal notation, this is not a formal notion.

So, you cannot write … and claim that you are listing out a infinite set. So, in fact, you need some other way of doing it and we will come to that as we go along in this lecture.

Now, it said seems reasonable that if a set is a collection of items, then we can collect anything and make it a set. It turns out that this is not quite true and this is particularly, a problem when we move to infinite sets. So, we have seen some infinite sets of numbers like naturals and reals and so on; but in general, if you take an infinite collection of objects, it may or may not form a set. In particular, Bertrand Russell showed that there is a problem, if we collect all the sets together and call it a set.

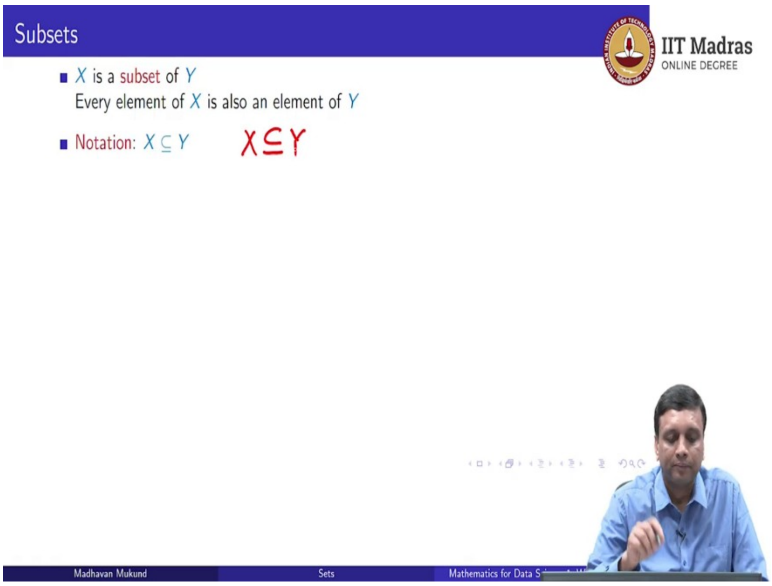
So, if we have a set of all sets, then we have a problem and this is something which is called Russell’s Paradox which we will discuss in the separate lecture, but you must be careful to note that though the notion of a set is intuitive and it seems natural that any collection of objects is a set, we have to actually be a little careful in mathematics, if we are using sets in order to define what is a set and what is not a set.

But given that whatever we will see in our course, we will be fairly straight forward. So, whenever we see a collection of numbers or a collection of objects of mathematical description, we can safely assume that they are sets.

So, again some terminology. So, we have talked of different things items in a set, members of a set and so on. So, the most formal notation for the members of a set is an element. So, a set consists of elements and we write this membership of an element in a set using this $\in$ notion. So, we have this $\in$ notation which stands for element of. So, when we write $x \in X$, we mean that small x is a member of the set capital X.

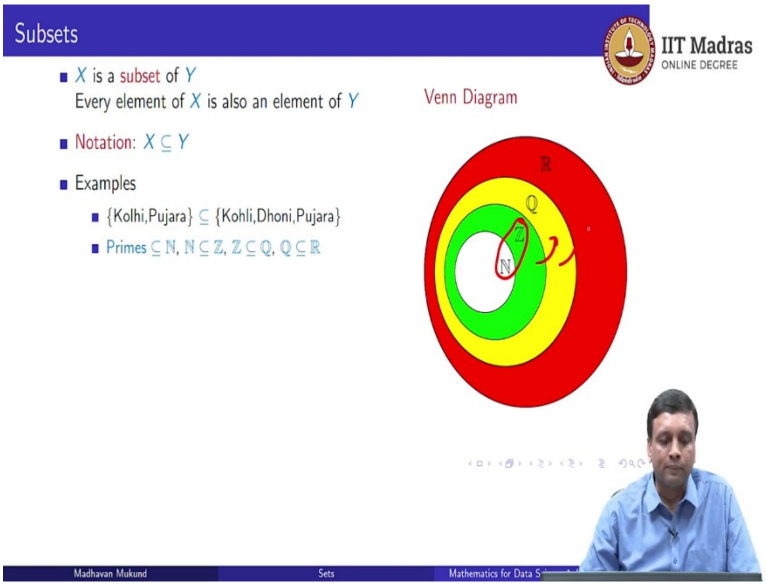
So, example 0 is a member of the natural numbers right. So, $0 \in N$is what we use. So, we can see for instance that 5 is an integer, but $\sqrt{2}$ as we claimed is not a rational number. So, an element of symbol with the line across it, means not an element of. So, 5 is an element of integer set and $\sqrt{2}$ is not a member of the set of rationals.

(Refer Slide Time: 10:02)



So, moving on from elements, we can compare sets by asking whether one set is included in another set and this is called a subset. So, $X \subseteq Y$, if every element of X is also an element of Y and this is written using this subset notation $\subseteq$. So, you have this familiar notation $X \subseteq Y$.

(Refer Slide Time: 10:27)



So, for example, if we take just 2 out of the 3 players were listed before saying Kohli and Pujara; then, this set forms the subset of our original set Kohli, Dhoni and Pujara. Similarly, if we take all the natural numbers and collect only the prime numbers. So, remember that the prime number is a number whose only factors are 1 and the number itself. So, it has exactly 2 factors; 1 and p and then, p is a prime number.

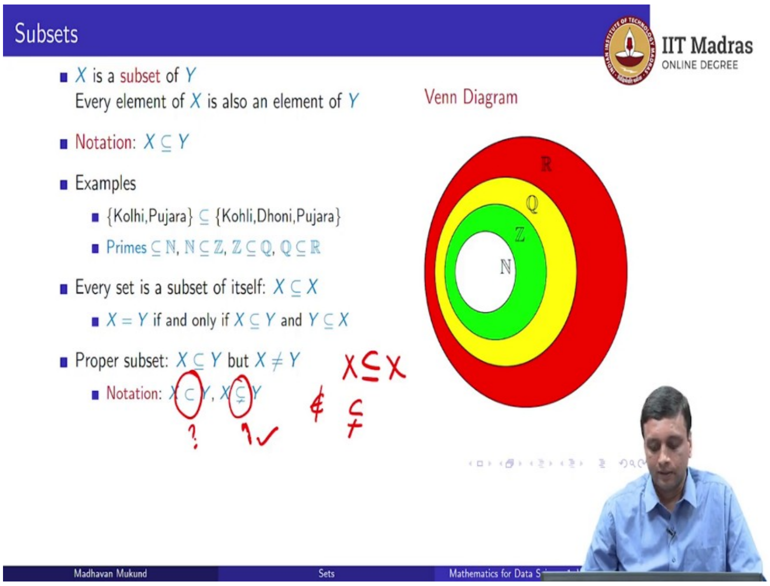
So, since some many numbers are not prime, primes is a subset of natural numbers. Since, the integers extend the natural numbers with the negative numbers, we can say that the natural numbers are included in the integers. So, $N \subseteq Z$. Similarly, we extended Z to Q. So, the set of integers is a subset of the rationals and the set of rationals is a subset of reals.

So, if you wanted to draw it, we could draw it in this particular way. So, we can raw a large circle representing the reals, a small circle inside right in the center representing the natural numbers and if one circle is included in another circle, it means that this circle is a subset of the circle outside it. So, here you can see that the natural numbers are a subset of the integers and then, from the integers, we can say that there are subset to the rationals and the rationals are a subset of the real numbers.

So, this kind of a diagram, where we represent a set by a boundary. So, this is a very abstract diagram. We are not in this case for example, listing out the elements of the set we are just indicating the extent of the set saying that the set are extends beyond Q and everything that is in Q is sitting inside R.

So, these are what are called Venn diagrams. So, a Venn diagram is a very useful way to picturize a set and relationships between sets; is one set a subset of another, is one set not a subset of another and so on.

(Refer Slide Time: 12:20)



So, we often use Venn diagrams pictorially in order to represent sets. So, notice that every set is a subset of itself because remember the definition of a subset set that X ⊆Y, if every member of X is also a member of Y. So, since every element of X is also an element of X, trivially as a extreme case of this definition, every set is a subset of itself.

So, this in fact, gives us an important notion which looks obvious; but it is not so obvious, when are two sets equal. So, two sets are equal if and only if, they are actually the same set of elements. So, one way to check that two sets are equal is to check that everything in the first set belongs to the second set. So, X ⊆Y and everything in the second set belongs to the first set. So, Y ⊆X.

So, often this happens when we have two different ways of looking at the same set of objects. We have two different descriptions of the same set of objects and we want to check whether they are equal or not. Then, using the first description, we argue that everything which satisfies the first description also satisfies the second description and vice versa.

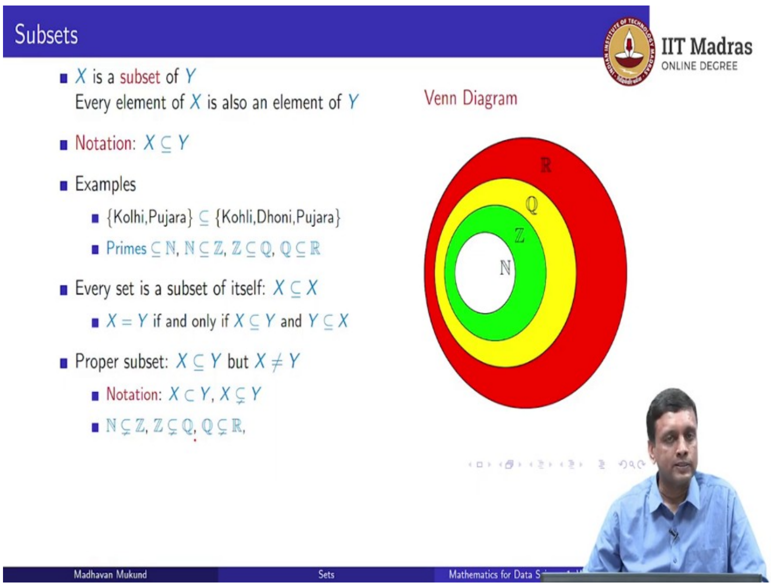
So, though this looks fairly obvious for finite sets, when it comes to infinite sets we have sometimes have to argue in an indirect way. So, this although it is an obvious statement is very important that X = Y provided X ⊆Y and Y ⊆X. So, sometimes we want to distinguish between the case, when X is really a proper subset of Y; that means, it does not include all of Y and that it is possibly equal to Y.

So, the subset equal to notation that we have right allows both. When we write X ⊆X, what we are saying is that it is a subset, but it is actually equal. So, we are allowing both cases. So, if you want to talk about proper subsets, sometimes we use a different notation.

So, we might either drop the equal to sign , just write the subset sign ⊂or we might explicitly like we said not element of right. So, we are saying that this is not equal to. So, we are dropping the equal to from below the subset.

Now, this is a bit dangerous. Second symbol this not equal to this is always correct. This is sometimes used both ways. So, you have to be bit careful when we look at books when you see the single subset without the equal to whether they mean subset and equal to or proper subset.

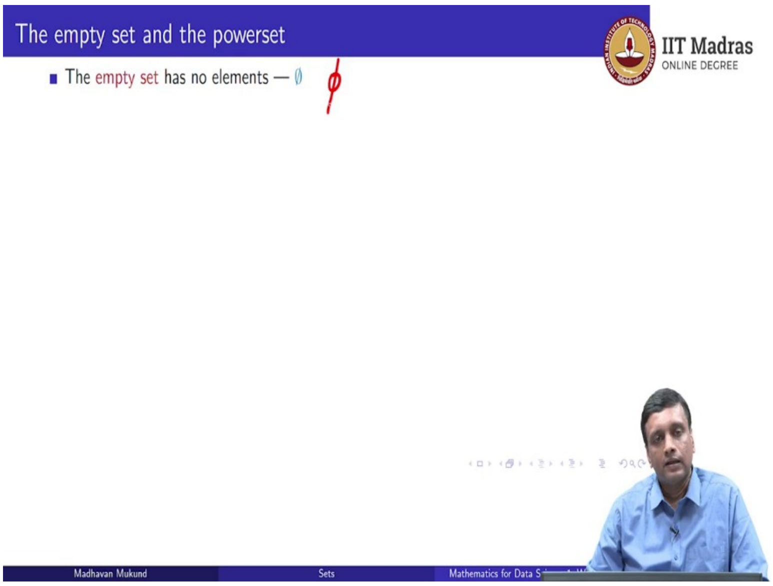
(Refer Slide Time: 14:45)



So, we know for instance that the natural numbers is a proper subset of the integers because the negative numbers are not there. Similarly, the integers are clearly a proper subset of the rationals and because the irrational numbers are not rational, the rational are a proper subset of the real numbers.

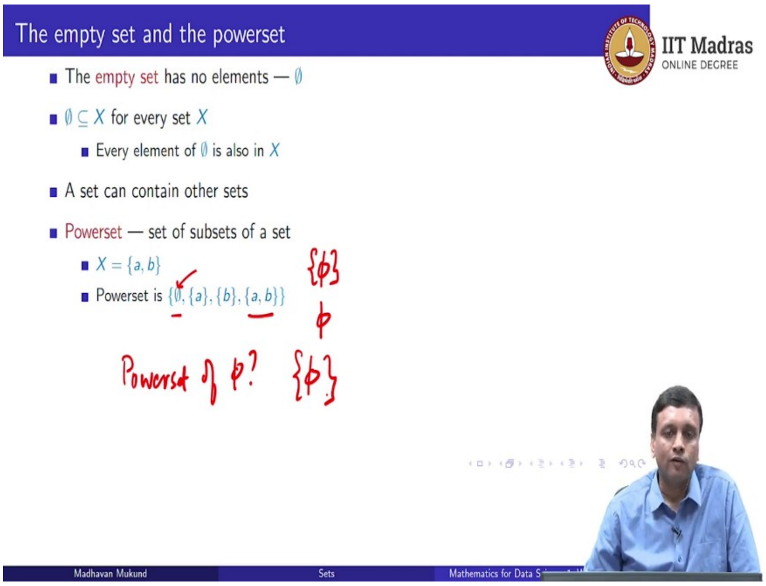
So, in most interesting cases, we will be looking at proper subsets. Sometimes, we will emphasize it by adding this cross against the equal to and sometimes, we will not and very often from context we will know whether we are talking about proper subsets or we are talking about subset which allow the full set.

(Refer Slide Time: 15:16)



Now, there is a very important set just like the 0 is very important in numbers, there is a very important set which is important set theory. It is the equivalent of 0. It is the set which has no elements. So, the set which has no elements is called the empty set and is written ∅. It is basically you can think of it as a 0 with a line across it. So, this Greek letter phi, symbolizes the empty set; so, it has no elements.

(Refer Slide Time: 15:44)



Now, what may not be very obvious is that this empty set is actually a subset of any set. Remember that we said that X ⊆ Y, if every element of X is also every is also an element of Y. Now, you might argue that an empty set has no elements. So, why is this true? Well, when we say for every and there is nothing in the set, then for every something is true right.

So, if I say that all birds with 3 legs have pink beaks, then this is actually true because we can imagine that there are no birds with three legs and therefore, every bird which actually has 3 legs will have a pink beak. But since, there are no birds with 3 legs this is actually true.

So, these kinds of vacuous statements as they are called will hold for sentences which use the word all where the set is empty. So, in particular, every element of the empty set because there are none. So, every element that could be in the empty set is also an any set X that you build. So, this empty set is a subset of every possible set. Now, though we have talked about elements and sets.

So, they are two different categories of objects. So, we have numbers and the numbers belong to a set of the type N or Q or R or Z; a set can clearly contain other sets. So, there is no restriction saying that the members of a set or the elements of a set must be some kind of discrete and indivisible objects.

So, one of the important sets of sets that we would like to look at is what is called the Powerset. So, we talked a subset. So, supposing we want to enumerate all the subsets. So, here is a two element set a comma b. So, what are all the subsets? Well, we already saw that the empty set is always a subset. So, that is one subset.

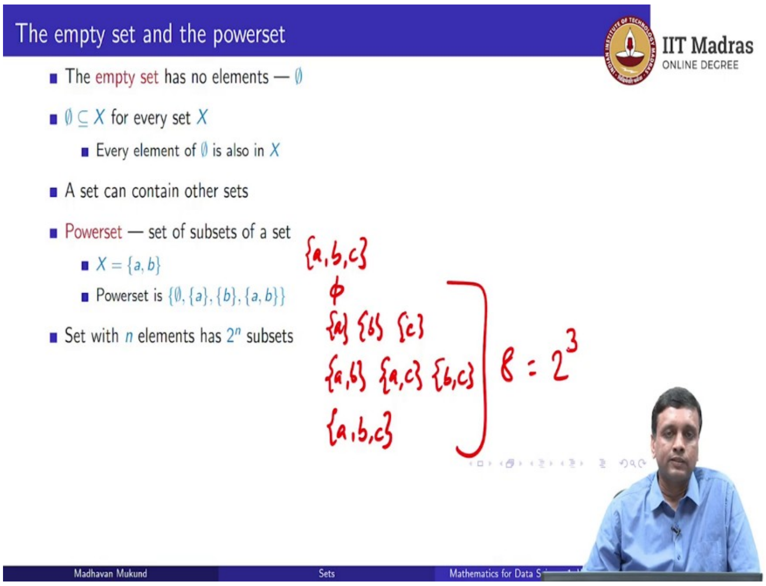
The set itself for any X, $X \subseteq X$. So, X equal to $\{a, b\}$. So, we have these two subsets which come just from the fact that empty is the subset of every set and the set itself is a subset. And then, we have two proper subsets either we can include the a and exclude the b or include the b and exclude the a. So, there are four subsets of X and if we group together these four subsets into a larger set, then we get the Powerset.

Now, notice that this itself is the set right. So, we do not write. So, this is different from this. The first is a set consisting of one element, namely the set consisting of the empty set. The lower thing is the empty set alone which is the set with no elements. So, if we put a brace around the empty set symbol, then we create a set with one element.

So, for instance, if you ask what is the power set of the empty set right. So, we know that the empty set has a power set which contains the empty set. So, we have at least one empty set as one element of the power set and there is nothing else right.

So, the full set is also the empty set, but if you duplicate an element, it is a same thing. So, in fact, the power set of the empty set is a set consisting of just one element, namely the empty set itself. So, just remember this, that the empty set on its own denotes a set with no elements, but an empty set with the brace around it is not the same thing. It is a set consisting of one element, namely the empty set.

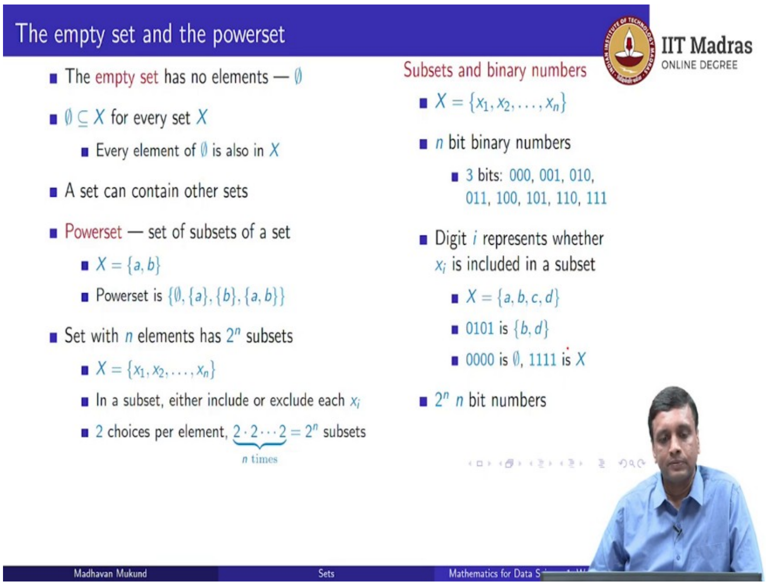
(Refer Slide Time: 19:03)



So, we saw above that if we have two elements, then the power set had four elements. So, in fact, one can generalize this and say that if we have n elements, then we would have 2^n subsets. So, for instance, if we had a, b, c right, then we would have 1 subset which is empty. We would have 3 subsets which are one element each and then, we would have 3 more subsets which are 2 elements each a, b a, c and b, c and finally, we would have the set itself right.

So, these are the only subsets. If you add these up, this is 8 which is 2^3. You can check that if you do it for a, b, c, d; then, you would have 2^4 , 16. So, why is it that a set with n element should have 2^nsubsets, no more no less?

(Refer Slide Time: 19:55)



So, here is one argument. Supposing we have n elements in the set. So, let us just call these without describing what they are specifically as x_1, $x_2$up tox_n. So, we have n distinct elements $x_1$to x_n. Remember these must be different because you cannot duplicate elements in the set. So, now, we want to construct a subset.

So, how do you construct a subset? Well for each elementx_i, we have to either include the set include $x_i$in the subset or exclude $x_i$subset. So, we have to make a choice for each $x_i$right.

So, overall, we have to make n choices right. For each $x_i$we have to decide whether to include it or exclude it from the subset. So, we have two different choices for each elements. So, we have two ways to decide whether to do something withx_1, keep it or leave it; $x_2$keep it or leave it. So, then we have two times two choices for $x_1$and $x_2$together; two times two choices forx_1, x_2, $x_3$together.

So, in general, if we have n such choices where each choice involves two options, then we have 2 into 2 into 2, n times 2^nchoices. So, each of these choices gives us different subset. So, whenever we make a different choice, we will either leave out i from the set or put anx_i. So, it will differ from the choice, where we do the other thing. So, each choice generates a separate subset. So, there are exactly 2^n subsets.

Here is another way of looking at subsets and getting to the same result. So, we can actually think of subsets in terms of binary numbers. So, let us again think of our n element set $x_1$to x_n right. So, now, supposing we look at n digit binary number. So, digit actually is comes from decimal. So, we say bit for binary digit. So, n bit binary number. So, remember in a binary number system, we have 0's and 1's and the place values represent powers of two.

So, we have the unit digits is units as usual. The next digit 2 to the power 0 is a is number of twos, number of fours, number of eights. So, it is like the decimal thing is in base 10. This is in base 2. So, now, if we look at n bit binary numbers, then for instance, if we look at 3 bit binary numbers, then we have 8 of them.

We can start with 0 0 0, then 0 0 1, 0 1 0 and so on up to 1 1 1 and again, the reason that there are 2 to the n, n bit numbers is because for each bit we can choose to put 0 or 1. So, we have two choices for the first bit, two choice for second bit and so on.

So, it is not surprising that an n bit binary number can represent 2 to the n different values from 0 to 2 to the n minus 1, if we think of them as numbers. Now, we are interested in n bit binary numbers as representing subsets. So, what we will look at is the ith bit and say that the ith bit represents the choice that we made.

If we chose to keep $x_i$in our subset, we will call it 0. If we chose to we will call it 1 for example. And if we choose to omit $x_i$from our set, we will call it 0. So, 0 represents the choice, where we leave outx_i; 1 represent the choice, where we keep x_i

So, supposing we have this four elements set a, b, c, d; then, if we look at the binary sequence or the bit sequence 0 1 0 1, the first 0 corresponds to a, so it says leave out a. The second 0 corresponds to c, so it says leave out c and for b and d we have put a 1. So, it says keep b and keep b. So, it says leave out a, keep b, leave out c, keep d. So, this 0 1 0 1 as a binary sequence corresponds to the set b comma d.

What does 0 0 0, the all 0 sequence say? The all 0 sequence says every $x_i$in the set is omitted from the subset. So, this is precisely the subset which is the empty set because it has no elements and what about the all 1 sequence? Well, the all 1 sequence says every x_i that we have is included in the final subset. So, this is the set itself. So, remember that these are the two extreme subsets; the empty set and the set itself and all the other ones come in between.

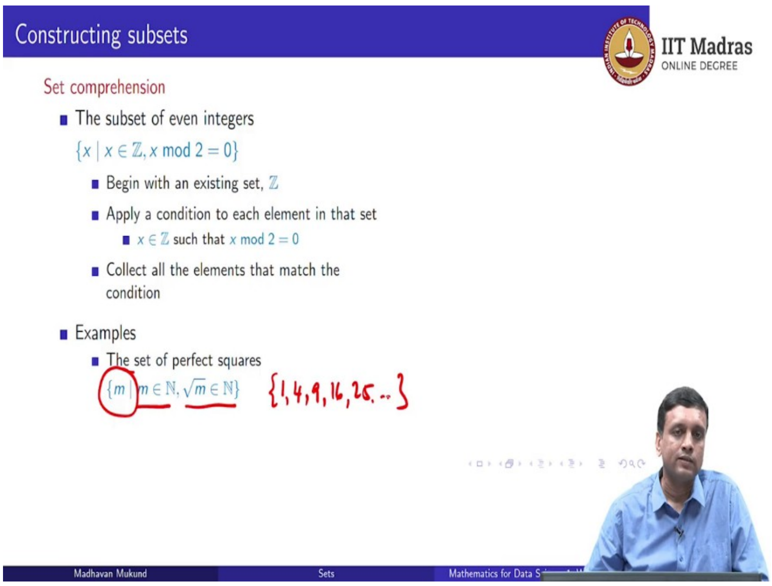
So, from this, we can see that every n bit number represents one sequence of choices. So, this gives us 2^nchoices because there are precisely2^n, n bit numbers. So, hopefully with this, you are now clear about the fact that any finite set with n elements has exactly 2^n subsets.

Mathematics for Data Science 1
Prof. Madhavan Mukund
Department of Computer Science
Chennai Mathematical Institute

Lecture - 05
Construction of Subsets and set operations

(Refer Slide Time: 00:06)

(Refer Slide Time: 00:14)



Now, let us talk about subsets in the infinite context. So, how do we talk about subsets of the numbers in a precise way? So, this is something called set comprehension. So, this is just some jargon. So, a set comprehension is just a term used for this which we have sometimes seen and which we will now review. So, if we want to talk about the set of even integers, the set of even integers are those integers which when divided by 2 have a remainder 0. So, remember that the remainder is called mod. So, x mod 2 is the remainder when divided by 2.

So, if x mod 2 is 0 it means that when we divide x by 2 there is no remainder. So, any such x is an even number. So, this notation that we have written is actually the set comprehension notation. So, let us try and separate out the different parts and understand what is going on.

So, when we use set comprehension first of all we can only do set comprehension when we have a starting set. So, we have to begin with a set and construct a subset of that set. So, the first thing says that we want to take all x in Z. So, this here says that we are looking at elements from an existing set in this case this set is a set of integers. Then it says I want to take all elements and apply some condition to decide whether to keep that number or not. So, that is the second part of the right hand side.

So, we have the first part which tells us which set we are looking at the second part which tells us what condition we want. So, we are really saying x in Z such that x mod 2 is 0 and finally, with this bar and this left hand side we are saying collect together all the x which satisfy this. So, this overall this notation says collect all the x for which x is in Z such that x mod 2 is 0 or in other words x is even. So, this is set comprehension notation and this is formally how you define a subset of an infinite set. Remember that we cannot list out the elements in an infinite set.

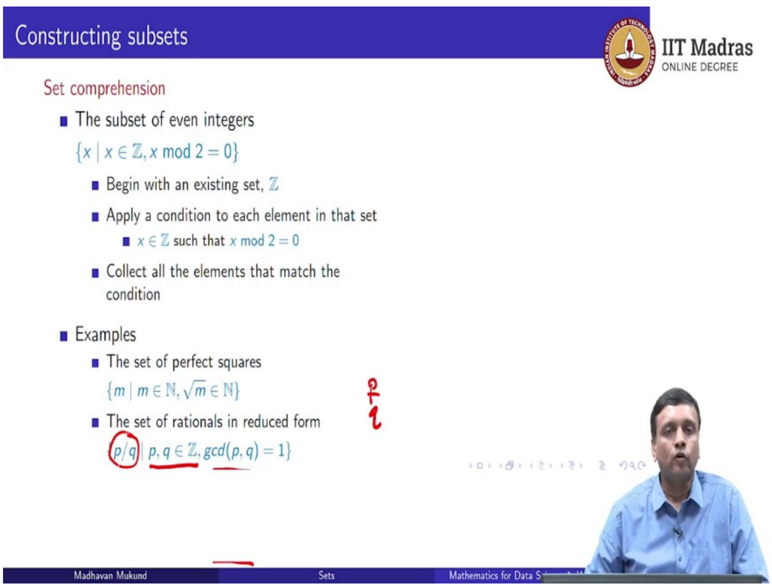
Now we assume that we already have a set like Z or N or Q or R for which we know what elements are. So, we do not have to describe how to pick out element we know what those elements are. What we are now giving is a description of how to choose elements which satisfy a given property. So, let us look at some more examples. So, for instance let us look at perfect squares.

So, remember that we said an integer is a perfect square if its square root is also an integer. So, for instance 25 is a perfect square because the square root is 5, but 26 is not a perfect square because there is no integer which multiplied by itself is 26. So, here is a set comprehension notation of the perfect square.

So, first of all remember square number has to be positive. We already discussed that negative numbers cannot be squares because when we multiply 2 numbers by them to the same numbered by itself the, 2 numbers will have the same sign. So, either it will be minus into minus is plus or it will be plus into plus is plus because the multiplication rule says that if the 2 numbers you are multiplying have the same sign the outcome is always positive. So, first of all we can only have positive numbers. So, instead of looking at integers, it suffices to look at the natural numbers.

So, we say for all m which are natural numbers such that the square root of m is also a natural number. So, this is that the square root of m also belongs to the set N collect all such m right. So, we are collecting all the m. So, this will give us if we write it out explicitly 1 will fall into this set, the next number that will fall into the set is 4, then 9 and then 16 and then 25 and so on right. So, the notation in blue is a succinct way of writing this informal infinite list which starts with 1 and goes on. So, we are pulling out the numbers from N one by one; checking if they are perfect squares and if so we are enumerating them.

(Refer Slide Time: 04:07)

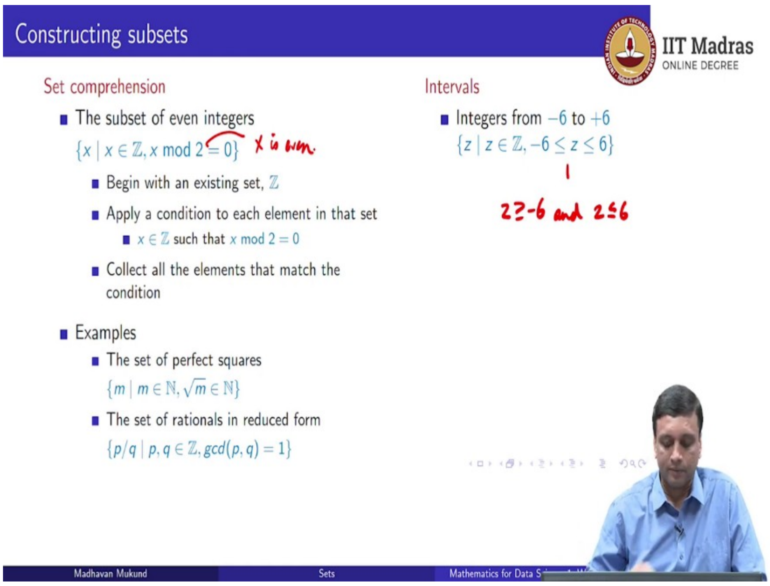


We also talked about rationals in reduced form. We said that there are many different ways of writing the same rational number because if we multiply the numerator and the denominator by the same quantity, the number we are representing does not change. And we use this fact in order to make denominators same when we did comparisons or arithmetic like addition and subtraction. So, what are the actual rationals in reduced form. So, this is a subset of the rationals. For example, $\frac{3}{5}$ is in reduced form, $\frac{6}{10}$ is not in reduced form because I can; cancel the 2 and get $\frac{3}{5}$.

So, if we want numbers and rationals in reduced form first of all we pick up any 2 numbers which are integers. Remember that a rational is actually a pair a numerator and a denominator which are integers. So, every rational looks like this $\frac{p}{q}$ right, but we do not want any such $\frac{p}{q}$. We want $\frac{p}{q}$ such that they do not have any common divisors other than 1. So, recall the gcd is the greatest common divisor; it is the largest number that divides both p and q and what we want is that p and q have no numbers which can be divided into them other than 1. And if the gcd of p and q is 1 then $\frac{p}{q}$ is a rational and it is in reduced form because the gcd is 1 right.

So, this is another example of set comprehension in order to define an interesting subset of the rationals.

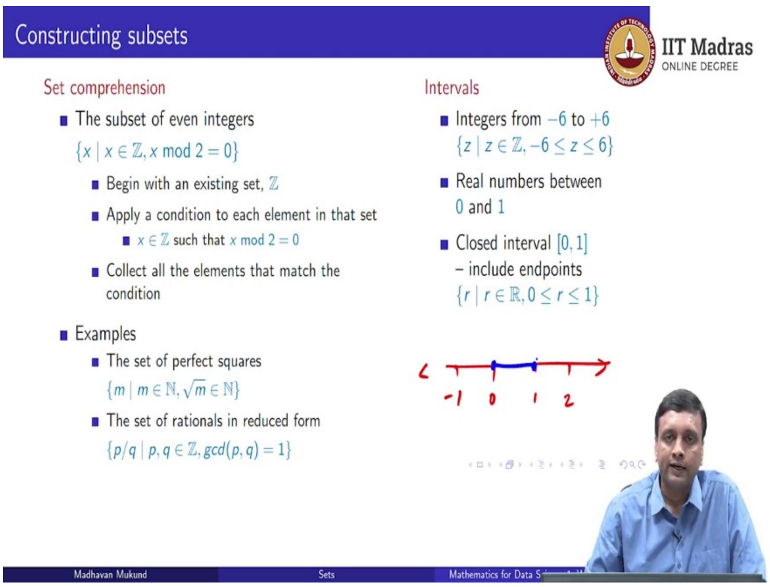
(Refer Slide Time: 05:34)



One of the things that we will often use with respect to numbers is to define intervals of numbers between something and something else. So, for instance if you are looking at the integers; we might want the integers from some lower limit to some upper limit. This for example, is an expression which describes the integers between -6 and +6 right. So, it says I want all z which belong to the set of integers such that z is above -6 greater than equal to -6 and less than or equal to 6. Now, we could split this for instance into two conditions. We could also say z is bigger than -6 and z is smaller than 6 and so on

So, the way in which we write this condition which applies to the thing may vary and all of them could be equivalent to each other. So, we will not be very pedantic about what syntax we used to write there. So, for instance in the previous case here, we could have just read written x is even instead of x mod 2 is 0 ok. So, we will not worry too much, but it is just that we have this format where we take the underlying set, we pick out all elements, make it satisfy condition. If it satisfy the condition, it belongs to the subset.

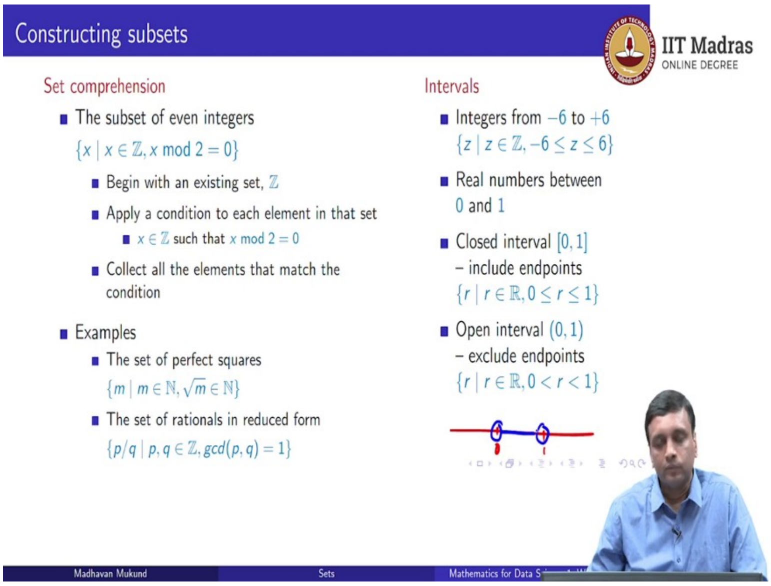
(Refer Slide Time: 06:44)



So, intervals are more interesting when we talk about real numbers and one of the intervals that we really often want to talk about is the interval between 0 and 1. So, 0 to 1 is quite interesting because we will often talk about probabilities for instance and probabilities range between 0 and 1. So, what can we do between 0 and 1? Well first of all we can take all the real numbers between 0 and 1 including both 0 and 1 and this is called the closed interval.

Closed interval means in this case, it includes the endpoints. So, if I draw this as a number line for instance. So, normally I have 0 1, 2 ,-1 and so on. So, this is my number line. So, then this closed interval says I want all the numbers from 0 to 1 including 0 and 1. So, this is my closed interval right. So, what we write is take all r in the set of reals such that $0 \leq r \leq 1$.

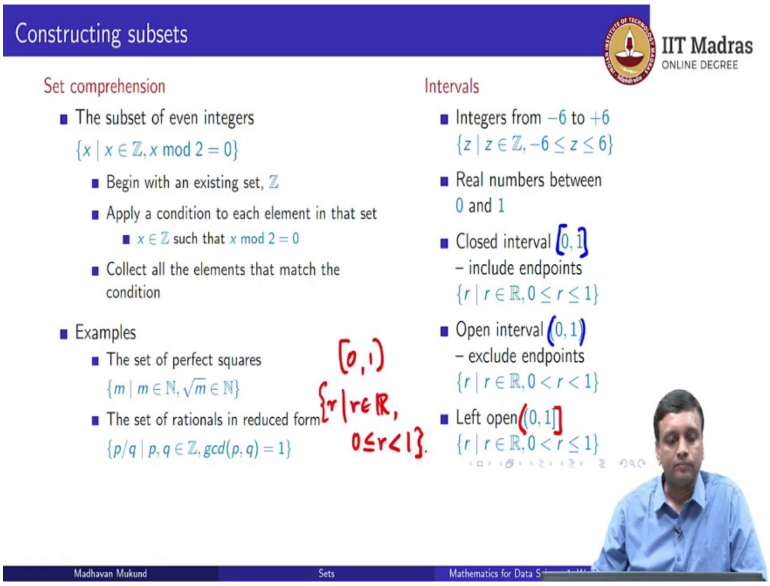
(Refer Slide Time: 07:46)



So, r must be between 0 and 1 it could be 0 and it could be 1. If we want to exclude the endpoints, then we get what is called an open interval and the way we draw an open interval; if we want to draw it in a pictorial way is to emphasize that the endpoints are missing by drawing a circle there.

So, we draw a circle to indicate that those are not included. So, if we so I have to fill in the circle corresponding to the endpoints that endpoint is included in our interval. If we do not fill it in it is not included, but formally it is just a set defined using set comprehension and whether it is open or closed depends on whether the inequality has an equal to or not whether it is strictly less than or it is less than equal to whether it is strictly greater than or greater than equal to.

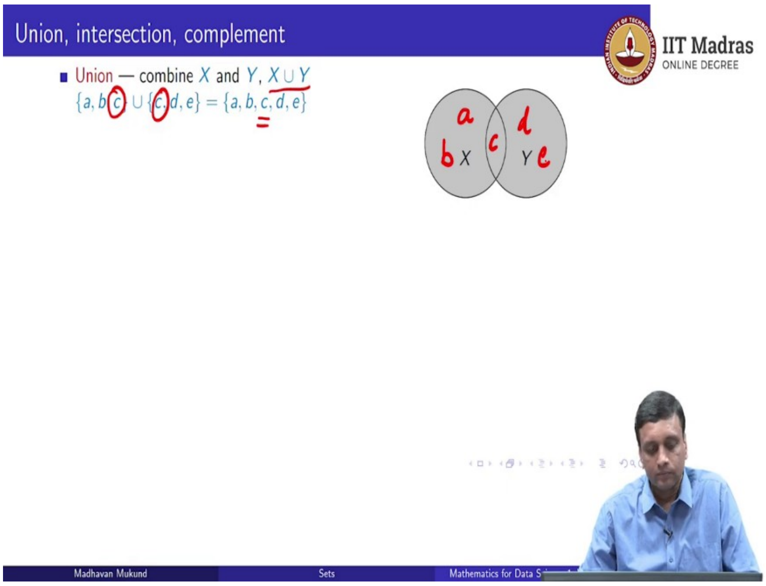
(Refer Slide Time: 08:29)



Now, there is nothing to stop us from including one endpoint and not including the other. So, we had an closed interval which had both endpoints, we had an open interval which had both endpoints missing. And we could say for instance that an interval is left open. So, it is all numbers between 0 and 1; it does not allow us to use 0, but 1 is included. So, in notation we will use this. So, the notice that we use this round bracket for open and we use the square bracket for closed. So, here obviously we will use a round bracket for the open end and a square bracket for the closed end. So, the left is open. So, we call this a left open interval.

So, left open interval has all numbers which are strictly bigger than 0, but less than equal to 1. So, correspondingly you could have a right open interval. And what would this be? This would be all the r such that r belongs to a set of reals. Now, $0 \le r$ we are allowed to include 0, but we should not include 1 right. So, this is the right open interval. So, this will be an important part of many discussions. So, you should be aware of these intervals as representing sets of points in particular a subset of the reals which can be defined using set comprehension.

(Refer Slide Time: 09:41)



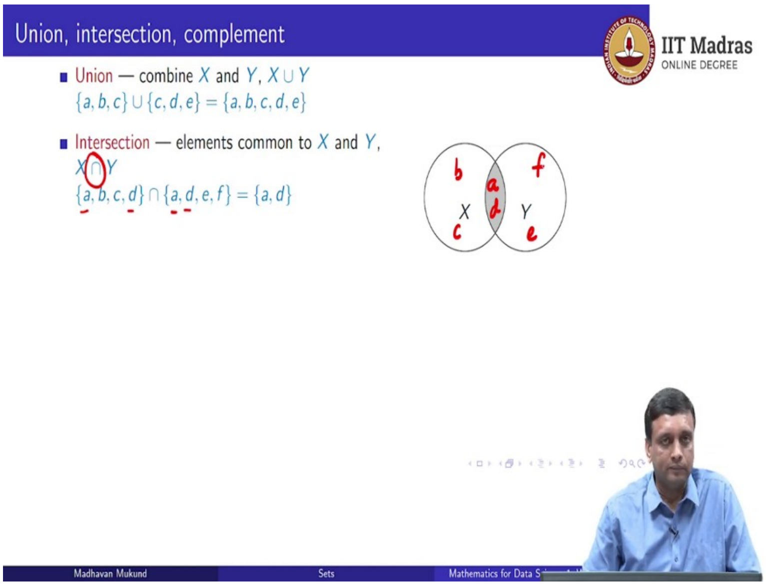
So, finally, let us look at some simple operations on sets which we are all familiar with. So, the first one is union. So, the union of two sets just combines them into a single set. So, suppose we have a, b, c as one set and we combine it with c, d, e then we get a single set. And notice that we have some elements which may appear in both sets and they appear only once in the final set because remember that a set has no duplicates right. So, in the union if we take sets which have some common elements across the two sets, they get represented exactly once in the final set.

So, therefore, the cardinality of the union will in general be less than the cardinality of the two sets put together. So, here we have two-three element sets, we take the union we get a five element set not a six element set because there are some elements which are common and the symbol for union is this ∪ right. So, $X \cup Y$ and if we go back to our Venn diagram; so, remember that we used when diagrams in order to informally look at sets and we talked about subsets. So, here we have a Venn diagram which represents the left hand side set is X, the right hand set is Y and the picture suggests that X is not a subset of Y and Y is not a subset of X, but there may be some overlap. So, this is the general case right.

Generally speaking if I give you two sets, there will be some elements which belong only to X some elements should belong only to Y and some which belong to both. So, this kind of a picture with two overlapping circles or ellipses is a particularly general picture of two sets represented as Venn diagrams. Even though we are not specifying what the elements are this is a picture. So, here for instance if we wanted to write out these elements in this particular set if you wanted to write we have a here, b here, c here, d here and e here.

So, what this means is that if we look at the circles a, b, c belongs to the left circle c, d, e belongs to the right circle, but we put c in the portion which is covered by both circles to indicate that it is in the common portion. So, this grey shaded area in this particular case represents the union of two sets.

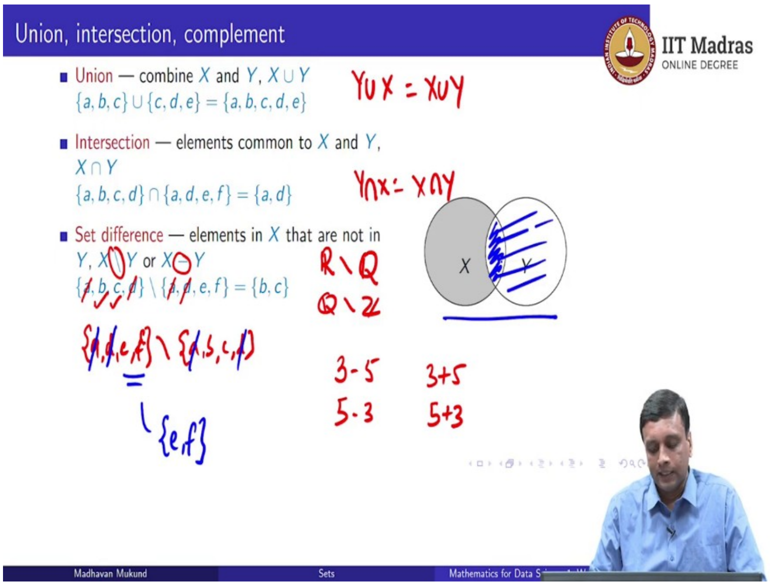
(Refer Slide Time: 11:43)



So, the corresponding thing which takes up only the elements which occur in both sets as you know is called intersection. So, intersection is written with the upside down version of the union sign right. So, X intersection Y is written like $X \cap Y$. So, here for instance we look at elements which are on both sides. So, we have a, b, c, d intersection a, d, e, f. So, a is common to both, b is not there on the right hand side, c is not there on the right hand side, d is common to both and if you go to the right hand side e is not there on the left hand side f is not there. So, only a and d are surviving intersection.

So, again if we draw this out as a Venn diagram on the right, the shaded portion which is the area which is overlapped by both the circles is the intersection. So, in this particular case we would write a here because it is in both b here. Notice the order is not important and in an Venn diagram if we actually put the elements the position is not important. So, I can put them anywhere and then I put e here and f there for instance. So, this is a pictorial representation of the two sets on the left. The shaded area corresponds to the intersection and the non-shaded portions are those which are in one set, but not in the other.

(Refer Slide Time: 12:53)



Another operation on sets is called set difference. So, in set difference we take two sets and we want to know what is there in the first set that is not there in the second set. So, for instance we want to know which are the real numbers which are not rational right. So, then we would write in this notation which are the real numbers which are not rational right or which are the rational numbers which are not integers. So, this is a common thing that we might want to do.

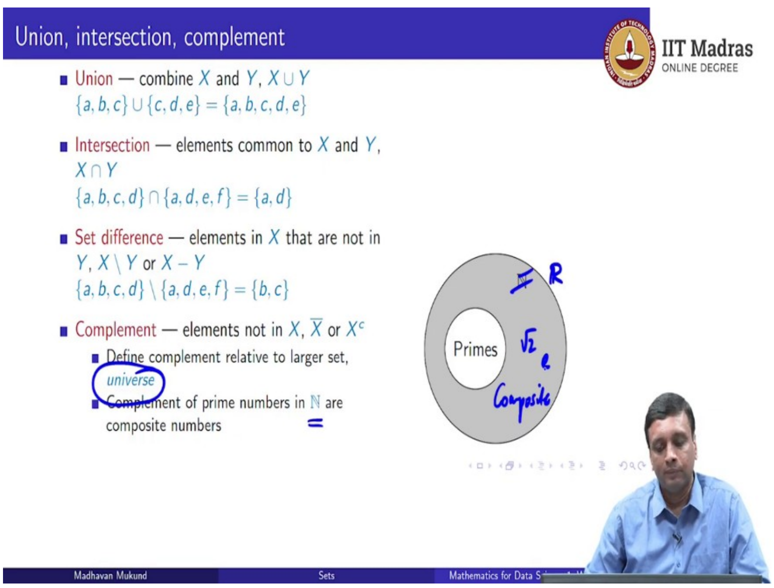
So, we write either this direct subtraction which is the normal minus sign or we write this back slash kind of notation \ to indicate the set difference. So, it is all elements in the first set which are not in the second set. So, here for instance if you look at the first set a is there, but a is also there in the second set. So, a is not counted, b is there, but b is not there in the second set. So, b is in the set difference c is there c is not there in the second set. So, c is in the set difference, but d for instance appears here. So, d is not counted.

So, here we have that the first set minus the second set has b and c because those are the 2 elements in the first set which are not in the second set. Now, this is like subtraction not symmetric in the sense that you know that 3 - 5 is not the same as 5 - 3 unlike 3 + 5 right. So, 3 + 5 is the same as 5 + 3, but 3 - 5 is not the same as 5 - 3. So, if I take union for instance, then Y ∪X = X ∪Y right and Y ∩X = X ∩Y because this it does not matter which side you take from.

Because finally, you are going to look at all elements which I has a common to both side or included in both sides. Now here if I take the reverse if I take a, d, e, f right and I subtract out the elements from a, b, c, d; then I would see that again a would disappear. So, the same elements disappeared because the common part is the same. So, a would disappear and d would disappear because these are the parts which are on both sides, but what survives now is e, f right.

So, when I do it in the other way around, I get the elements on the right hand side which are not on the left hand side. So, in the set difference the order of the sets in the expression matters. X \ Y is not the same as Y \ X just like in subtraction and here we have a picture right. So, this shows us this picture. It says that you take everything in X and you remove everything that all includes. So, in particular you remove all these elements which are in the intersection and that gives us X \ Y.

(Refer Slide Time: 15:19)



And finally, we often talk about the complement. We say those numbers that are not prime. So, those numbers that are not prime in particular are called composite numbers. So, composite number is defined to be a number which has factors other than 1 and same. So, any number which is not prime has more than 2 factors. So, such a number is called a composite number. So, clearly a number is either a prime or it is not a prime.

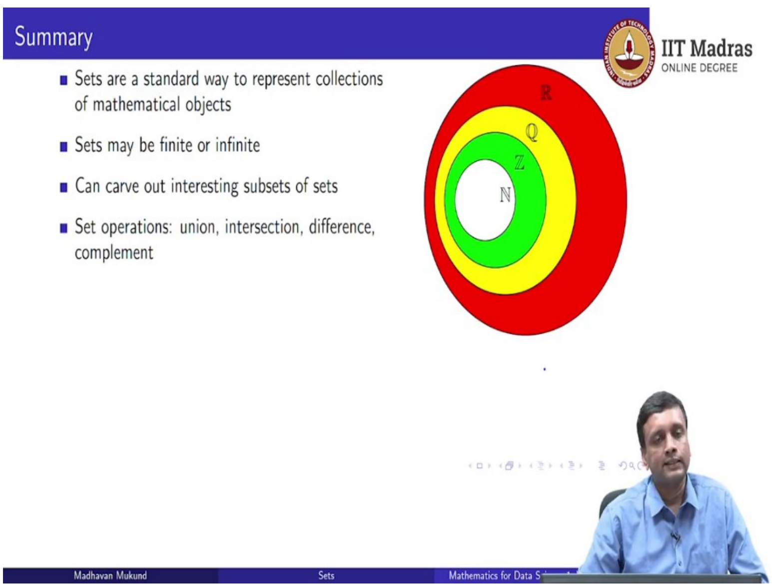
So, either it is prime or it is a composite. So, the composite numbers are disjoint from the primes and they are all the numbers that are not prime. So, this is what we mean by complement. Complement means the opposite side it means everything else, but complement is not very straightforward in set theory because complement with respect to what.

So, if I say numbers that are not prime, but I do not tell you in what set I am talking about this thing. If I look at complement in for example, in the reals; it will include all numbers like π and e and $\sqrt{2}$and so on and that is not what you mean right. When I say the complement of the primes; you are not thinking of rational numbers, irrational numbers and so on. You are thinking of integers or in particular you are talking about natural numbers which are not primes right. So, we would always want to define what is called a universe ok.

So, we need a universe with respect to which we are going to complement. So, if we say that the complement of prime numbers in the universe of natural numbers, then we get the composite numbers. So, when we say primes for instance we see this Venn diagram on the right, we see primes as a subset of the natural numbers. So, then the grey shaded area is all the composite numbers right. But if this was not this, but R then we would have various thing we would have $\sqrt{2}$, e and so on sitting here which is not what we intend.

So, whenever you use the word complement, you must make sure that you have specified complement with respect to what. What is the overall set with respect to which you are negating the set that you have and that is very important.

(Refer Slide Time: 17:18)



So, let us wrap up this lecture. So, we are all familiar with sets as an informal term which we have come across from school level and a set is a standard way to represent a collection of mathematical objects. So, it is very important to be familiar with the terminology of sets element of subset of and so on and also the notation the curly brace listing out the elements set comprehension and so on. So, sets may be finite or infinite. An infinite sets are actually quite tricky and interesting and most of the interesting sets that we are going to look at will be infinite because very often we will be thinking of sets in terms of numbers, but we will also be thinking in terms of finite things.

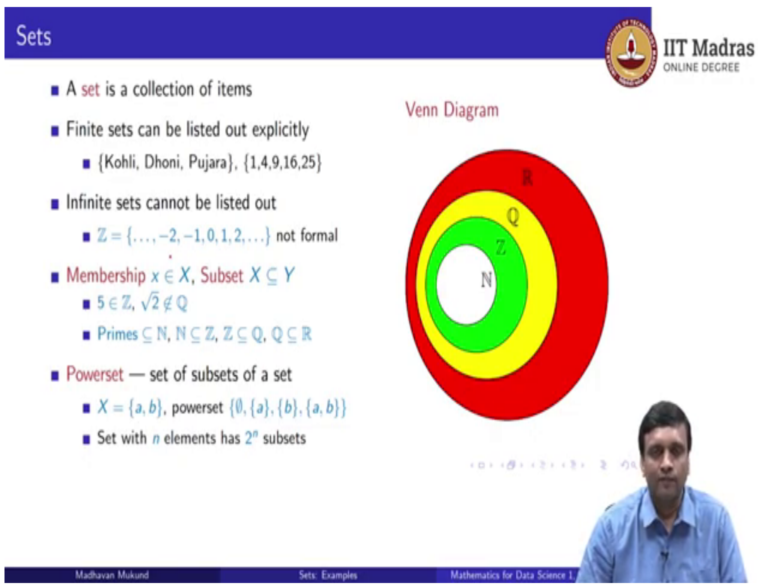
For instance we talked about we could talk about for instance a time table then we might want to know the set of stations at which the train stops or we might want to look at a shopping list and we might want to look at the set of items that the store has in its inventory. So, sets are a very useful way to talk about collections of objects infinite collections are important because numbers are infinite, but other finite collections are also important from a computational and data science point of view.

So, we saw that we have some useful notation like set comprehension which allows us to define subsets of infinite sets and we have these standard operations on sets like union, intersection, set difference and complement which allow us to take sets and combine them in many different ways. So, it is important that you get used to all these notions as I said because these notions are used implicitly throughout mathematics and these are not difficult notions is just a question of understanding the notation and understanding exactly what happens when you apply each of these operations.

Mathematics for Data Science 1
Professor. Madhavan Mukund
Department of Computer Science
Chennai Mathematical Institute
Lecture- 5A
Sets: Examples

So, we have seen some definitions of Sets and some operations on them. So, let us look at more examples to get familiar with the notation and the terminology of sets.

(Refer Slide Time: 00:25)



So, remember that a set is a collection of items and when we write out a set, if it is a finite set, then we can just enumerate the items in the set by writing them within curly braces. On the other hand, if we have an infinite set, we really can not write out all the elements even though informally, we put dot, dot, dot to indicate a sequence, if that sequence is not very regular. For example, supposing it is a set of prime numbers, which does not have a clear pattern, then it is not very easy to represent it explicitly like this.

So, we saw that there will be another notation called set comprehension that we will come to. But, before that let us talk about the two basic relationships between sets and membership of a set. So, membership is denoted by this element of relation. So, small x typically denotes a member or an element of a set, and capital X usually denotes a set itself. So, when we write x belongs to X like this, what we mean is the element x belongs to X.

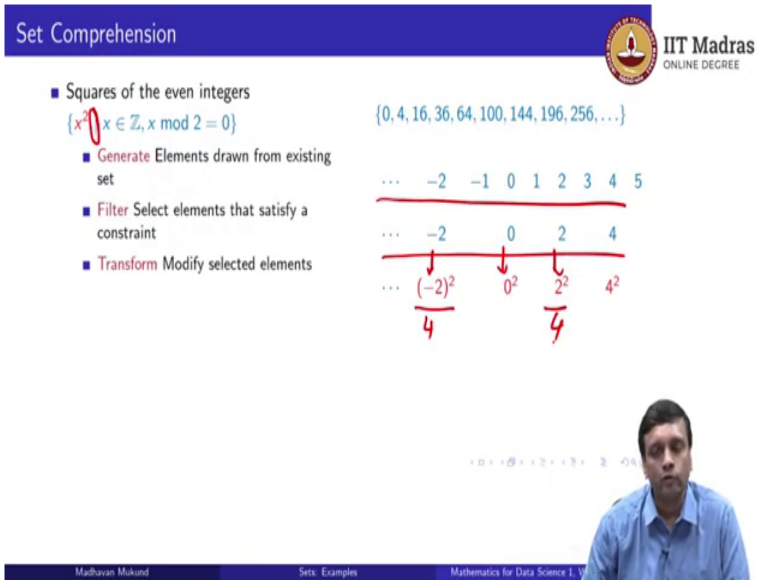
So, for example, the number 5 belongs to set of integers, and $\sqrt{2}$ does not belong to the set of rationals for instance. Subset on the other hand, says that one set is included in another set, so everything that belongs to X belongs to Y. So, for instance, all the prime numbers are natural numbers, so the primes are a subset of the naturals. Every natural number is an integer, so the natural numbers are a subset of the integers. Similarly, the integers are a subset of the rationals and the rationals are subset of the reals.

And we draw this using these Venn diagrams where we draw these ovals or circles or boxes representing the extent of a set, it is a picture of a set. And then depending on whether a box intersects another box or it sits inside a box, it indicates whether the first set is a subset of the other one or they overlap and so on. So, in this particular diagram which also has colors, we have indicated the subset relationship between the different types of numbers that we have studied, the naturals, the integers, the rationals, and the reals.

And finally, one very useful thing to know about sets is the power set. So, when we take a set, we can enumerate all its subsets. So, remember that we have just defined a subset. And in particular, we have this special subset called the empty set, which is a subset of every set. The empty set has no elements in it, but we needed it for technical reasons, and it is a subset of every set. And in addition, if you have 2 elements set $\{a,b\}$, then the subsets could be the individual elements, the set containing a and the set containing b or the entire set itself.

So once again, just like the empty set is a subset of everything the set itself is also a subset of itself. And we argued that for a finite set with n elements, we will always have 2^n subsets. So here for instance we have 2 elements, so we have $2^2 = 4$ subsets. So, this is just a review of what we have already seen.

(Refer Slide Time: 03:06)



Now, let us look at this set comprehension notation, which is what we said we would use when we have to describe infinite sets which cannot be written down explicitly. So, this was a typical example. So, supposing we want to write down the set of all the squares of the even integers. So, the even integers are -2, +2, 0 as an even, -4, +4, and so on. But if we square them, then we know that $(-2)^2$ is the same as 2^2 is 4.

So, the set on the right which is written in this informal dot, dot, dot notation has 0^2, 2^2, 4^2, 6^2 and so on. So, how would we write this out? Well, this is that notation on the left, which says that we take every x which belongs to the integers, check whether it is even, whether (x mod 2 = 0), and then square it. So, let us just break this up into parts so that we remember exactly what is happening.

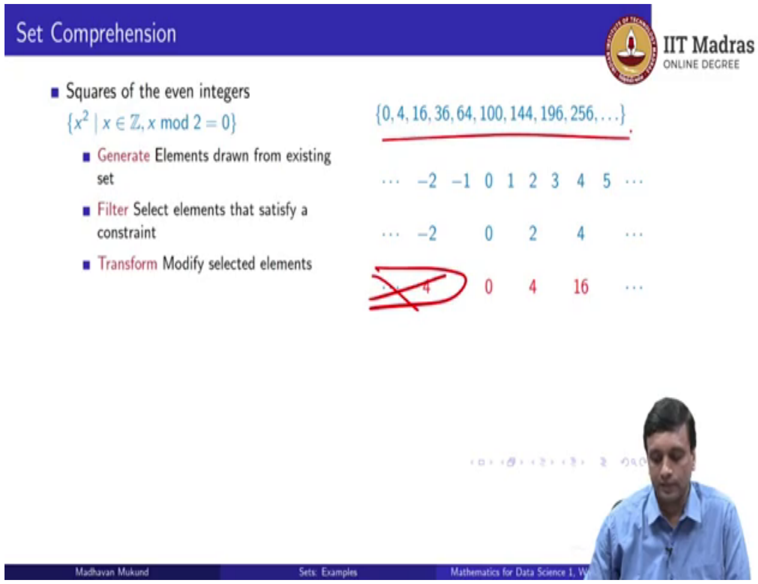
So first, in the set comprehension notation, we have a generator. A generator says that we are taking elements from an existing set, so we can only build new sets from old sets. So, we already have a set of integers, and we are going to try out every integer in the set, so, that is what x element of Z says, is try every x ∈ Z, so Z generates this set. Now, all the x's that come out are not interesting to us. So, we want to filter out those that are useful, that satisfy a given property.

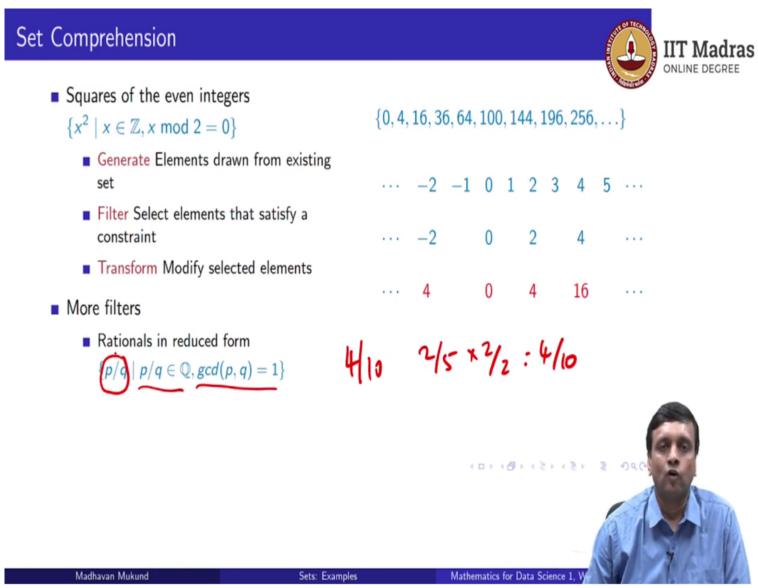
In this case, the property that we are looking for is that the number is even. So, we want those x which come out of Z through the generator, such that they satisfy the property that x when divided by 2 has remainder 0, which is the property that x is even. And finally, with these x, we do not want to keep them as they are, we want to transform them. So, on the left-hand side of this vertical bar, this is the left-hand side are the actual elements of the set.

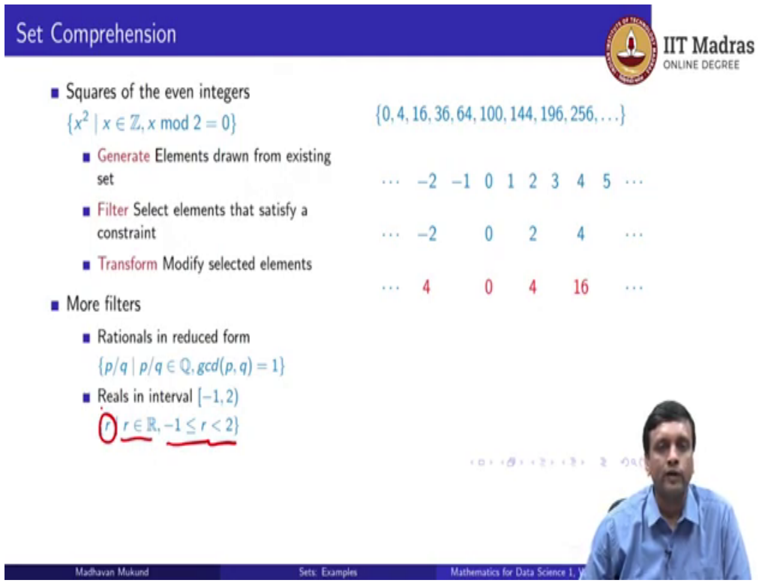
The elements of the set are generated right, then filtered through some conditions, which rule out the ones we do not want, and when the ones we keep, we can transform them. In this case, we want the squares, we do not want the even numbers, we want their squares. So, if you look on the right, this is what happened.

So, when we started the generating process, we had all the integers, then we filtered out, and we got only the even ones, and now we transform them. So, for each even number, we produced its square. And now in this process, you will notice that $(-2)^2 = 4$ and $2^2 = 4$ also. So, some elements will disappear because we do not keep duplicates.

(Refer Slide Time: 05:35)







So finally, when we go through this, we end up with this sequence 4, 0, 4. And then in this, we will throw away all the elements on the left, and we get the number sequence on the top. So, this is how set comprehension works.

So, we can write filters in many different ways as long as it is unambiguous, we will not be very particular about the language we use so long as there is no question about what we mean. So, for instance, we looked at this example, we have rational numbers, but some rational numbers are not in reduced form. For instance, if I write $\frac{4}{10}$, then I should actually think of this as $\frac{2}{5}$, because it is $\frac{2}{5} \times \frac{2}{2} = \frac{4}{10}$.

So, I have actually multiplied both the numerator and the denominator by 2, to go from $\frac{2}{5}$ to $\frac{4}{10}$, but it is the same rational number. So, we want the numerator and the denominator to not have any common divisors, which is the same as saying that their greatest common divisor is 1, that is nothing other than 1 divides both the top and the bottom of the fraction. So, if we take all the rational numbers, so we generate all the possible rational numbers $\frac{p}{q}$, which belong to the set of rationals.

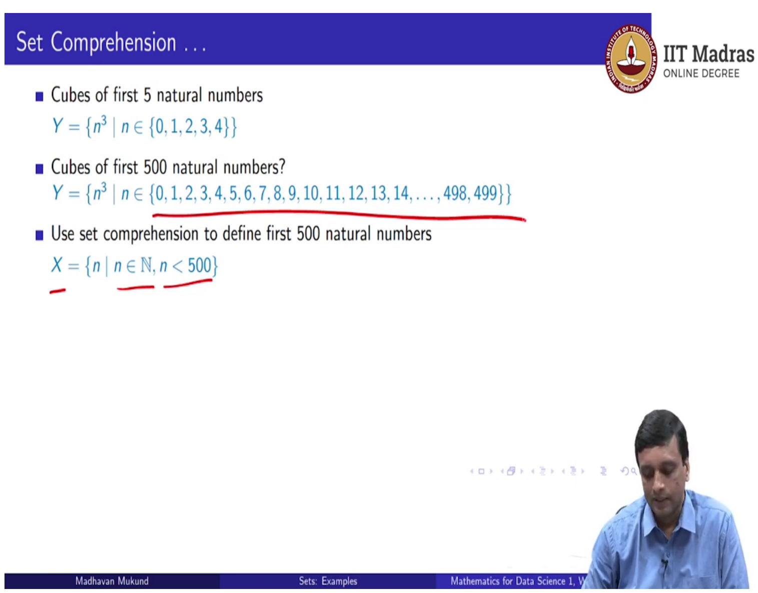
Then, we filter out those which have no common divisor between the numerator and the denominator and we keep only those, we do not transform it in any way, we just keep it here. So, here the transformation is just to keep it as it is, this is sometimes called the identity transformation. The identity just takes an input and produces the output the same as the input.

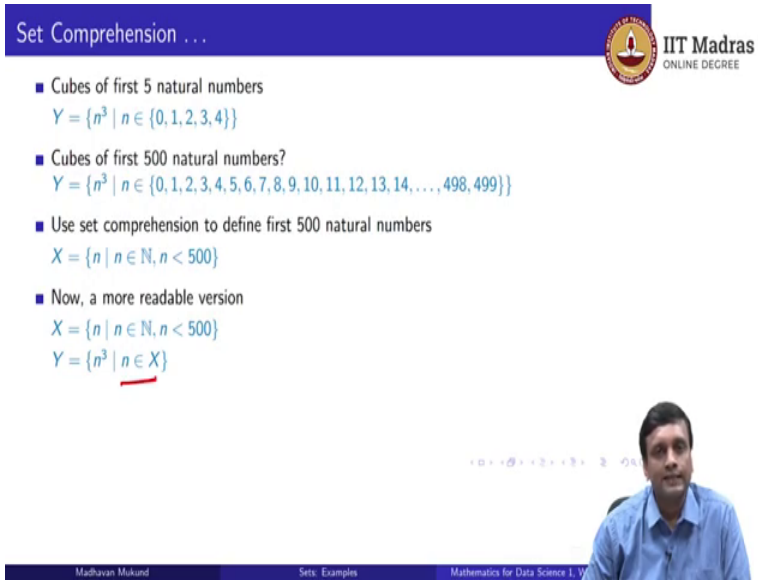
So, this gives the set of rationals in reduced form. So, here we have used a function, GCD. Even though we have not formally defined it here, we assume that people understand what GCD means. So, this is what we mean by saying that we can write the filter in any reasonable way, as long as people understand what it means.

Another example, we looked at are intervals. So, here we want the real numbers, which start from -1 including -1, and go up to but not including 2. So, in this case, we will use less than and less than equal to, so we will take all the reals. So, we take every possible real number, but we are not interested in all the reals, so we check whether it is greater than or equal to -1, so it includes -1 and everything above it. So, it cuts off everything which is strictly smaller than -1.

But, we also do not want it to cross 2, so we stop below 2, so it should be greater than equal to -1 or and less than 2 and if so, again we keep it without any transformation. And this notation on the top, the square bracket, and round bracket are indications of whether the endpoint is included or not. So, -1 endpoint is included, +2 endpoint is not included.

(Refer Slide Time: 08:14)





So, let us see why we would actually want set comprehension notation. So, let us extend our first example of squares of the even numbers to cubes. So, cube is just a number multiplied by itself 3 times. So, square is x×x, a cube is x×x×x, 3 times. So, if we want the cubes of the first 5 natural numbers, we can write it out explicitly like this, we can take this generator and generate the first 5 natural numbers as 0, 1, 2, 3, 4. Remember that, in our terminology natural numbers start with 0, even though in some books, you will find that natural numbers start with 1, we always assume natural numbers start with 0.

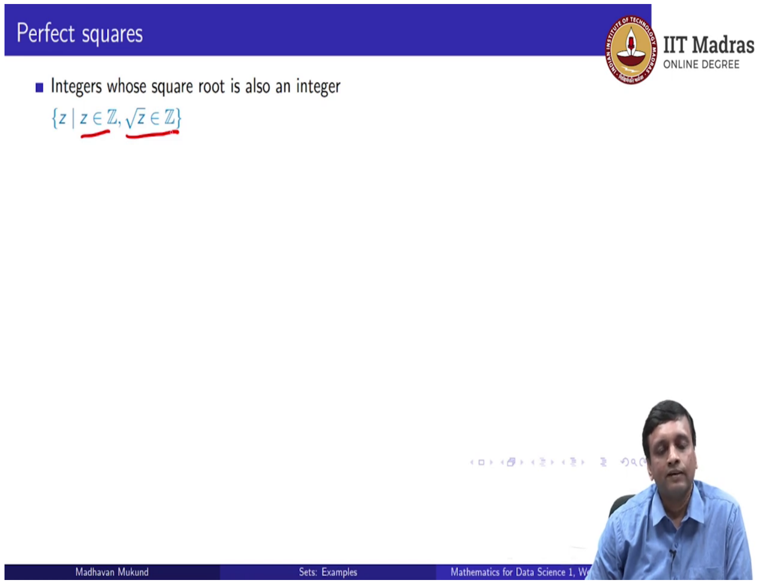
So, the first 5 natural numbers are 0, 1, 2, 3, 4. So, this is our generator, take every n in this and transform it to n^3 without doing any further filtering. We are not asking for the first 5 odd numbers or the first 5 numbers which have some other property, we just take, taking the first 5 numbers. Now, imagine that we change this question to the first 500 natural numbers, then though we can write it out explicitly, it is rather tedious.

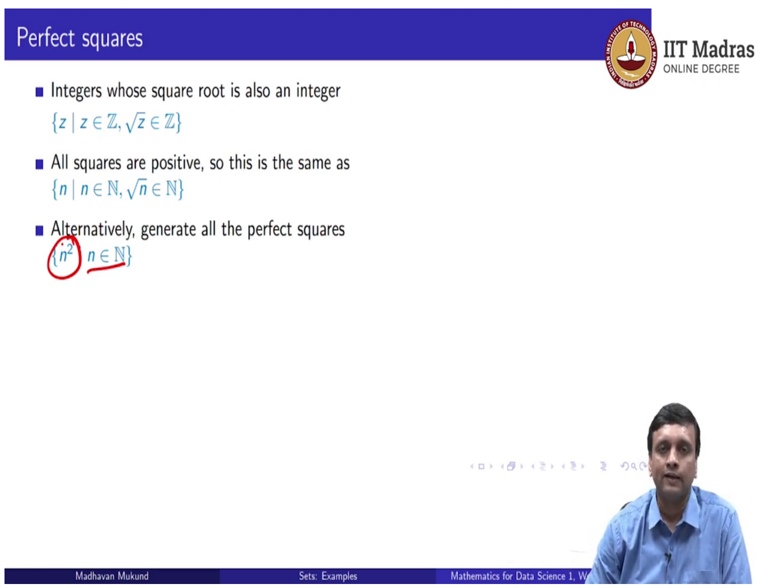
So, we have to replace the small list of 5 numbers by a long list of 500 numbers. And remember, we are not really allowed to write dot, dot, dot if we are being mathematically precise. So, we actually have to physically write out these 500 numbers. Now, this is not terribly convenient. On the other hand, we can define the first 500 numbers quite easily using set comprehension.

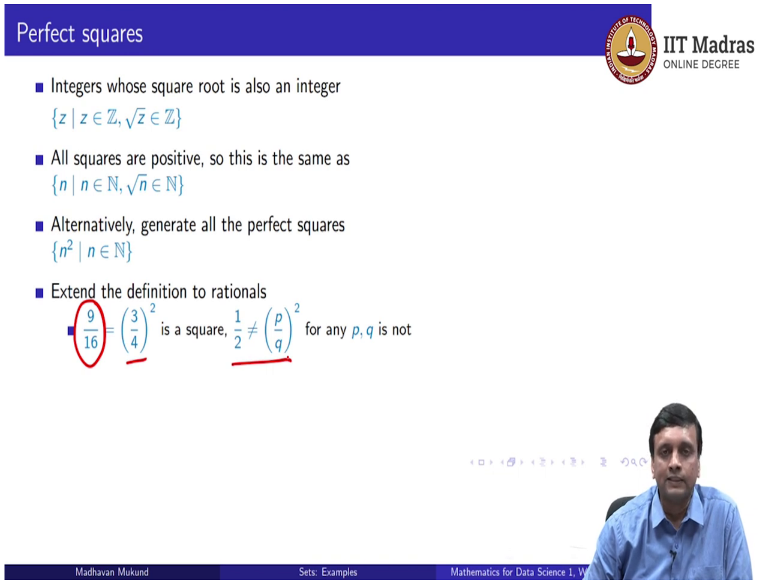
So, we can say, give me all the natural numbers, that is the generator, but restrict the natural number to be less than 500. So, remember that the first 500 natural numbers are going to be 0 up to 499. So now, this says that this set X is actually this long set here which we have written explicitly. So, we have replaced that very long and tedious expression by a much more compact expression, which captures exactly the same set. So now, we can have a much more readable version of these cubes of the first 500 natural numbers.

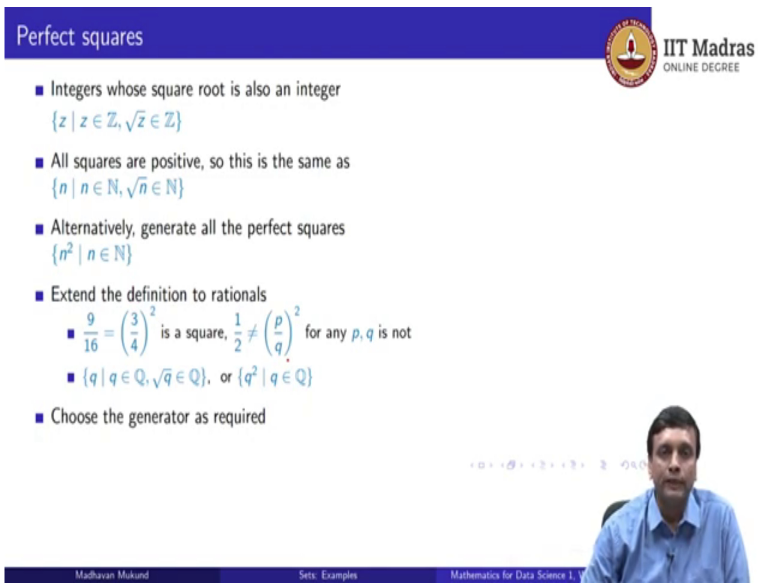
As an intermediate set, we generate the set X, set $X = \{n \mid n \in N, n < 500\}$. And then we take this as the generator and we say, okay, take every n which belongs to this X. So now, we know that x is restricted to 0 to 499. And then, take the cubes of these numbers, so we get n cubed in this range. So, this is one other use of set comprehension, which is to make our definitions more readable and understandable and less tedious to write.

(Refer Slide Time: 10:37)









So, let us look at one more round of examples. So, we saw this before, we talked about perfect squares. So, we said that some integers are squares of other integers and some integers are not squares. In particular, those which are not squares, their square roots are actually irrational. We proved for instance, in our supplementary lecture, that the $\sqrt{2}$ is irrational. So, perfect square is an integer such that its square root is also an integer. So, this is what this says, give me all the integers, which satisfy the condition that their square root is also an integer.

So, the square root of small z also belongs to a set of integers, give me all set Z and call it a perfect square. Now, notice that the square must be positive, we have already discussed this because you multiply 2 negative numbers, you get a positive number, you multiply 2 positive numbers, you again get a positive number. So, in fact, a perfect square must always be non-negative, it could be 0.

So, we could as well assume that the target set is generated by the set of natural numbers. And that, we are only interested in the positive square root, so remember that, 4 has 2 square roots, the number 4 is either $(-2)\times(-2)$, or 2×2, but it is sufficient to know that one of its square roots is an integer because the other one will just be the same with a minus sign.

So, we can as well define the same set of perfect squares in terms of the natural numbers, we generate all the natural numbers whose square roots are also a natural numbers. Now, we can turn this around and replace the filter by a condition. So, we know that every natural number when it is squared will give us a natural number. So, all the perfect squares will be generated in that form, take a natural number, square it. So, instead of looking for those numbers whose square root is a natural number, we can just take every natural number and square it.

So, we just generate all the natural numbers and without filtering them, we just take the output square. So, this also gives us 0, 1^2, 2^2, 3^2 and so on. So, these are all different ways of writing the same thing. In one case, we replace the generating set from integers to natural numbers because of the property of perfect squares. In another case, we transformed the filter into a transformation. So, instead of putting a condition on the numbers that we are generating, we took all the numbers and then squared them to get the actual perfect squares.

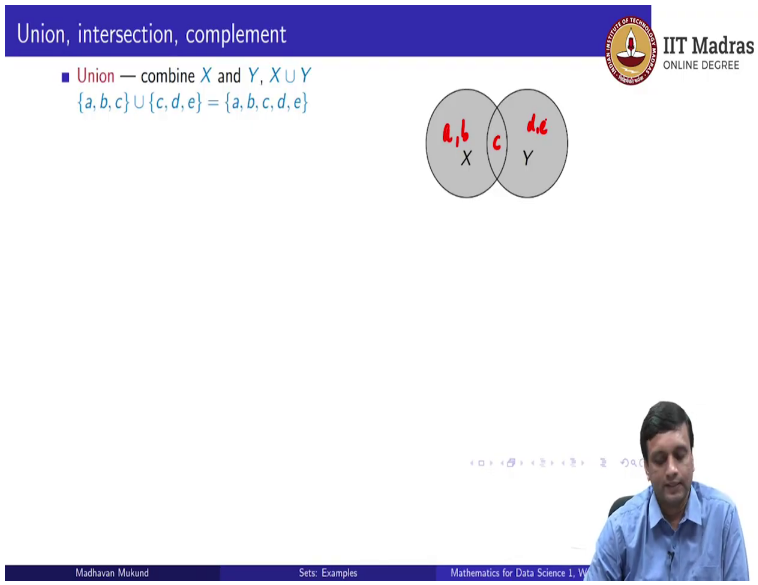
Now we could extend the notion of perfect squares to other sets of numbers. For instance, rationals can also admit a definition of a perfect square, so a rational will be a perfect square if it is a square of another rational. In particular, a rational could be an integer, but we will now integers can also be above and below the line, so we could have $\frac{9}{16}$, for instance as a rational number, which is $\frac{3^2}{4^2}$, so, $\frac{3}{4} \times \frac{3}{4} = \frac{9}{16}$.

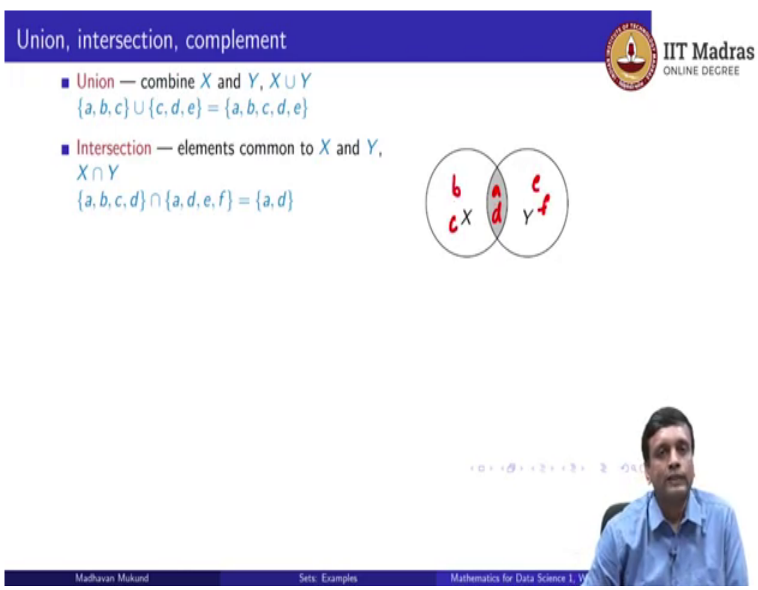
So, we might want to say that this is a perfect square in the world of rationals. And not everything is a perfect square because since $\sqrt{2}$ cannot be represented a rational, it is easy to check that half cannot be represented with a form $(\frac{p}{q})^2$. So, not every rational in this sense is a perfect square, some are, some are not. So, we can again change the definition above and replace Z and N by Q and get a reasonable definition of perfect squares in a different domain of numbers.

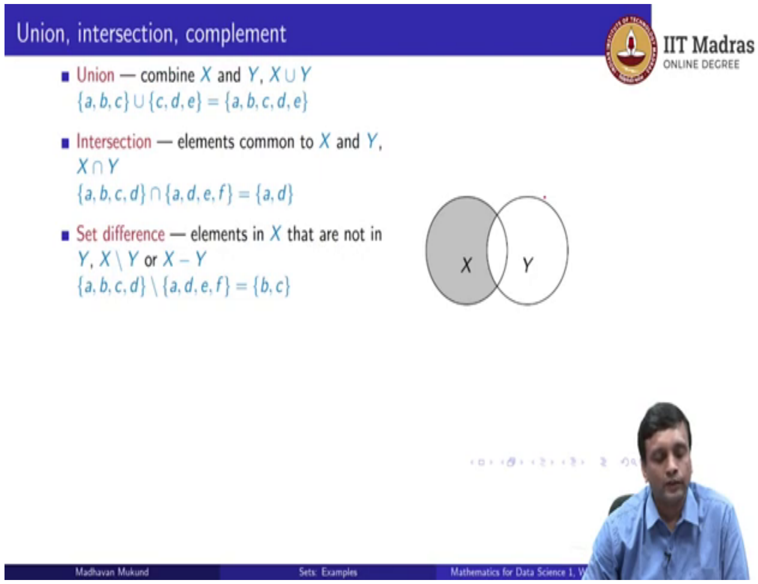
So, we can say give me all the rationals q such that $\sqrt{q}$ is also a rational. Or using the second form, we can say take all the rationals and square them. So, take every q, which is a rational q and give me q^2. So, this says that depending on how you choose the generator, you might generate the same set, or you might generate a different set. So, it is important to specify all the parts of a set comprehension correctly, so that there is no ambiguity and so that you get the set that you mean to get.

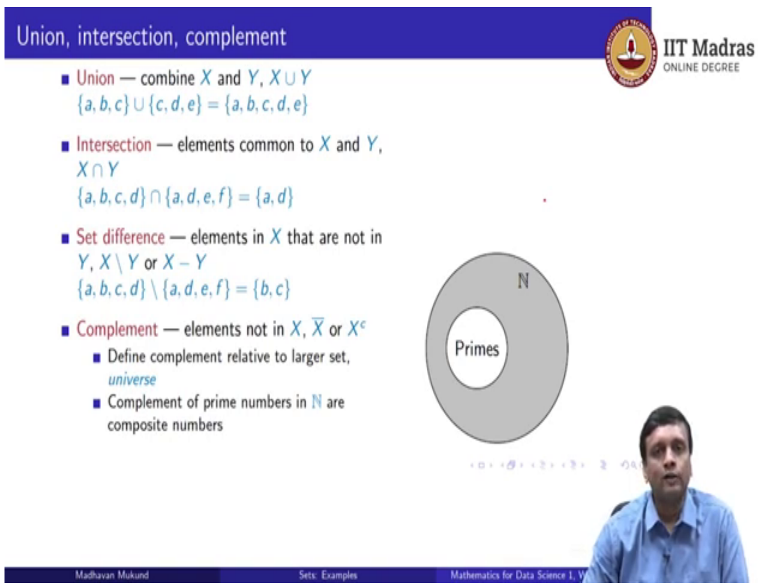
Mathematics for Data Science 1
Professor. Madhavan Mukund
Department of Computer Science
Chennai Mathematical Institute
Lecture- 5B
Examples of Set Operations and Counting Problems

(Refer Slide Time: 00:15)









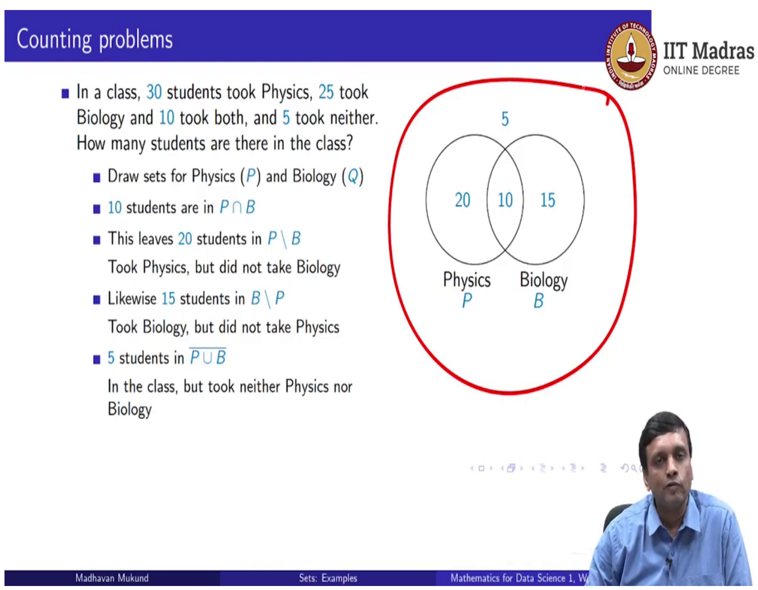
So, the other operations that we saw on sets are union, intersection, and complement, which we represented using Venn diagrams as shown here. So, the union takes two sets and combines them and removes the duplicates. So, the overlapping part between the two diagrams represents the common element. So, in this case, we would have this common element c over here, and then we have had a and b over here, and we would have d and e over here because d and e belongs only to Y, a, b belongs only to X.

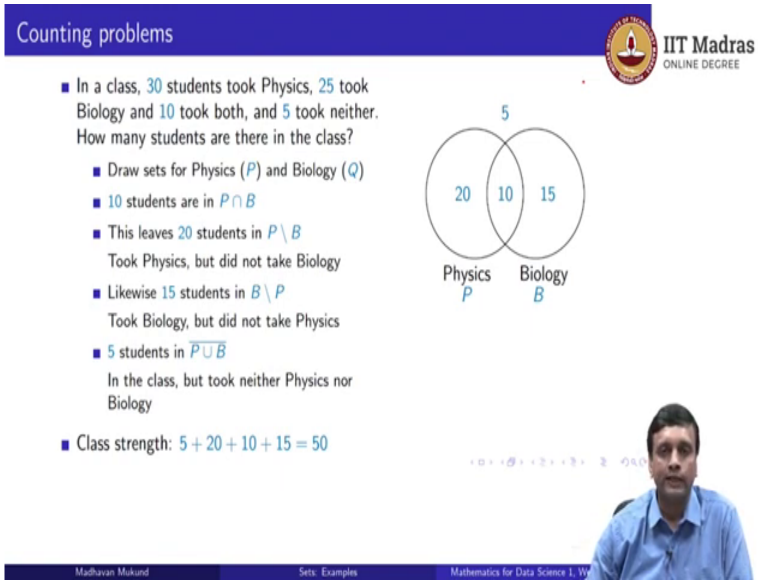
Conversely, we can take only those things which are common to the two and in this case, we have a and d over here, and then we know that b and c are only on the left and e and f are only on the right. So, the intersection tells us the elements which are common to the two sets. Set difference tells us what is on the left but not on the right.

And finally, the complement can be taken if we have an overall universe that is a full set to talk about. And with respect to that set, we can ask which elements are not in the set that we are looking at. So for instance, if we are looking at the natural numbers as a whole, the primes are a subset of the natural numbers, the complement of the primes are all those natural numbers that are not primes.

Now, remember that the complement matters, because if we take the complement of the primes, for example, with respect to the real numbers, we will get all sorts of other numbers which are not even integers. So, whenever we define the complement, we need to define the universe that we are talking about.

(Refer Slide Time: 01:37)





So, this leads us to a class of problems that you might come across, which can be solved nicely using these Venn diagrams. So, these Venn diagrams are not just pretty pictures, they are actually useful ways to reason about these problems. So, here is a typical problem that you could come across. So, you have a class in which 30 students have taken physics, and 25 students have taken biology, but 10 have actually taken both physics and biology, but there are also 5 who have taken neither of these two subjects.

So, these are the facts that are given to you. There are 30 students taking physics, 25 taken biology, 10 take both, 5 take neither, the question is how many students are there in the class. So, using Venn diagram notation, you can represent the fact that there are two sets of students, those who take physics and those who take biology by representing them by two sets, say P and Q. And we know that some take both, so, there is an intersection so these two sets overlap.

Now, from the data that we are given, we know that the overlap has 10 students, so we can write a number 10 in the intersection to indicate that there are 10 students who take physics and take biology. Now, we know that 30 students took physics overall and we have already accounted for 10 of them because they have all taken both physics and biology. So, there are 20 students who have taken physics, but have not taken biology.

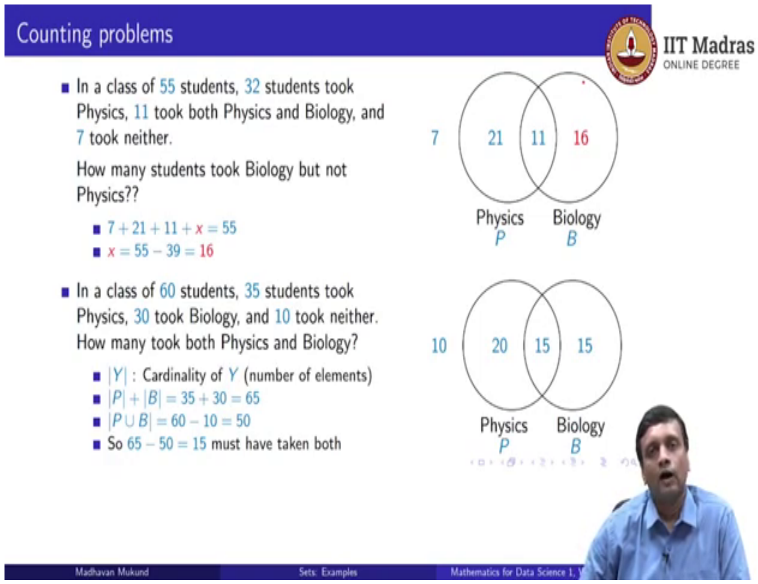
So, this in our set notation is the set difference, it is the difference between P and B, how many elements are in P which are not in B, how many students have taken physics who have not taken biology. And we have a symmetric thing on the right hand side. So, we know that there are 10 students who have taken both but 25 students take biology. So, there must be 15 students who are in $B \setminus P$, these are students who took biology and did not take physics.

So, in this way, we can populate the three regions of the Venn diagram with numbers indicating how many students are in each of these regions at 10 in the intersection, 20 on the left hand side, 15 on the right hand side. But, this is not the entire class because with respect to the entire class we have to take the number who are in the complement, those who have taken neither physics nor biology, and these are 5 students who are outside $P \cup B$.

Now, technically one should draw outside this the complement to indicate the entire class but just for convenience, I have not done that, but this entire complement outside this contains 5 elements. So, totally from this, we can see that there are 4 regions of interest. We have the $P \setminus B$ region physics but not biology, we have the $B \setminus P$ region, biology but not physics, we have the $P \cap B$ region taking both, and we have the complement, taking neither, and these are all disjoint from each other.

So, now if we add up the students across these, we get the exact number of students. And in this case it is $5 + 20 + 10 + 15 = 50$. So, there are actually 55 students taking physics and biology together, but the total class strength is only 50. And actually only 45 students are taking these subjects because 5 are not taken either of them.

(Refer Slide Time: 04:43)



So, here is a variation where the data for the problem is given in a different way. So now, you are told the class strength 55, you are told that 32 students took physics and of them 11 took physics and biology and you are also told that 7 took neither. So, the question is how many took biology but not physics. So again, we draw a Venn diagram and from the previous question, we know that we can put 11 in the intersection, because that is the number who took both.

And since there are 32 who took physics, we can subtract out these 11 and say that P\B is 21 and in the complement, we have 7. So, the question now is how many are in B\P, which I have marked by x, but now we know the total. So, we know that the four numbers together, add up to the total which is 55. So, 7 + 21 + 11+ x = 55. So, if we solve for x, we get that x = 16. So, we can deduce that 16 students have taken biology but not physics in this situation.

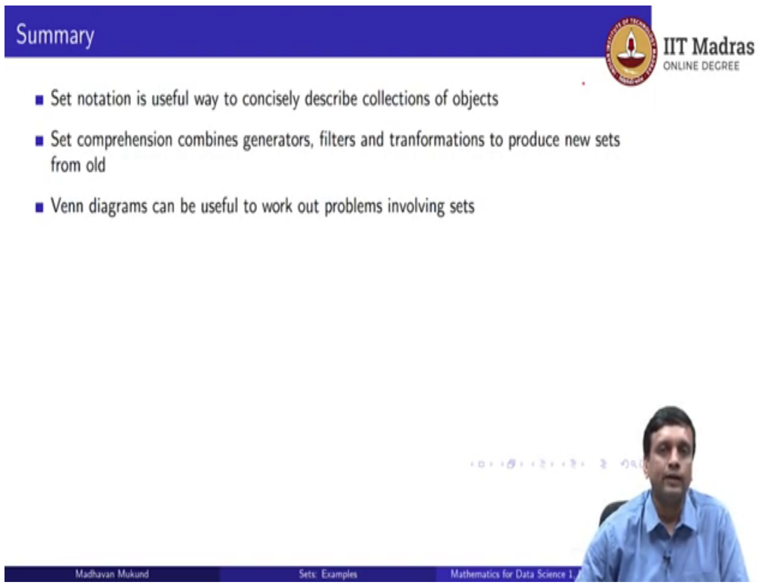
So here is yet another version of this. So, we have 60 students in the class. So again, we know the total number of students in the class, we are told that 35 students took biology, 35 students took physics, and 30 took biology, and 10 took neither. So now, we are trying to calculate the intersection, how many people took both subjects. So again, let us use this notation which we introduced when we first introduced sets.

So, this perpendicular bar on the side of a set indicates the size of the set. So, this is the cardinality of a set, cardinality is the number of elements, so the cardinality of Y is denoted by putting Y inside these bars. So, what we are told is that the set P has cardinality 35. That is a set of students who have taken physics overall, including those who have taken both, set B has 30 and 35 plus 30, there are 65 students who have taken in the union, of these I mean, have taken these together.

But we also know that there are 60 students in the class of whom 10 have taken neither. So, the actual union has only 50 elements. So, there are totally 65 people who are taking either physics or biology or both, but this total number actually spans only 50 students, so some of them must be taking both and are being counted twice. So, this must be the difference of the two.

So, 15 of these people must be counted twice, otherwise we would not have this mismatch. So, if we draw the diagram for this, this is how it comes out. We have 15, that we calculated for the intersection by taking the total number, realizing the 10 have taken neither, and then computing the difference between the number who should have taken both the subjects from those who are actually registered for either one or both of the subjects. So, these are three different examples using Venn diagrams to indicate how you can solve these kinds of counting problems.

(Refer Slide Time: 07:37)



So, to summarize, we use set notation because it is a very useful and precise way to talk about collections of objects. And if we use it nicely, it is also a concise way sometimes instead of writing out a long sequence of values, we can actually describe it using a condition. So, this is typically where we use set comprehension.

So, remember that set comprehension has three parts, some of which may not be used. So, you always have a generator, a basic set from which you are creating new sets, you may have a filter which takes out some elements from the generated set and throws them away and keeps only those that satisfy the condition.

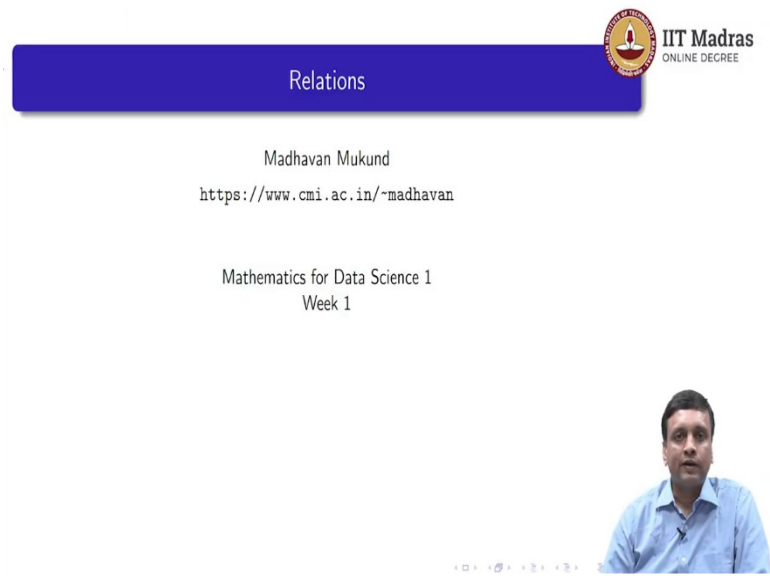
And finally, you may have a transformation which takes these filtered elements and does something to make them into the elements that you want, for example, the squares of the even numbers. And then we also saw that Venn diagrams are not just simple doodles that you draw to indicate sets, Venn diagrams can actually be very useful for calculating properties about sets, especially numerical problems about sets. So, it is important to be able to draw the proper Venn diagram to indicate which groups of sets overlap, how they overlap, and which parts are empty, and so on.

Mathematics for Data Science 1
Prof. Madhavan Mukund
Department of Computer Science
Chennai Mathematical Institute

Week - 01
Lecture – 06
Relations

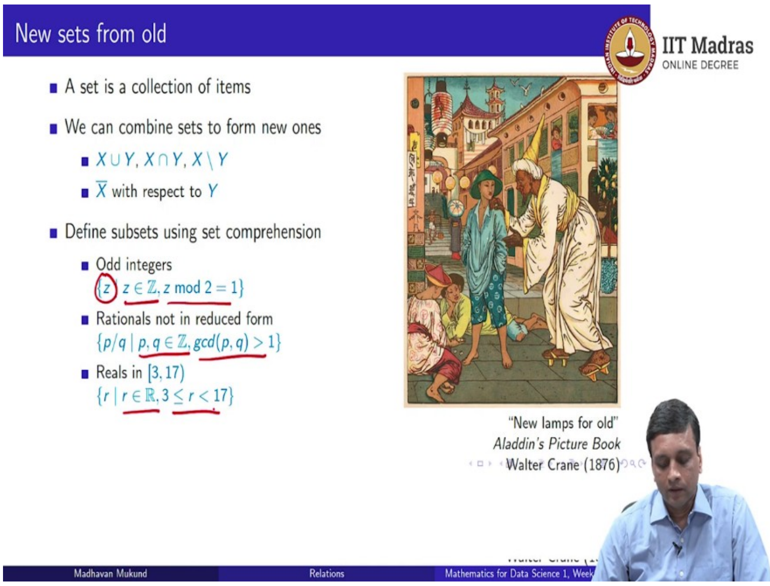
(Refer Slide Time: 00:05)

(Refer Slide Time: 00:14)



So, we have seen Sets, now let us move on to Relations.

(Refer Slide Time: 00:17)

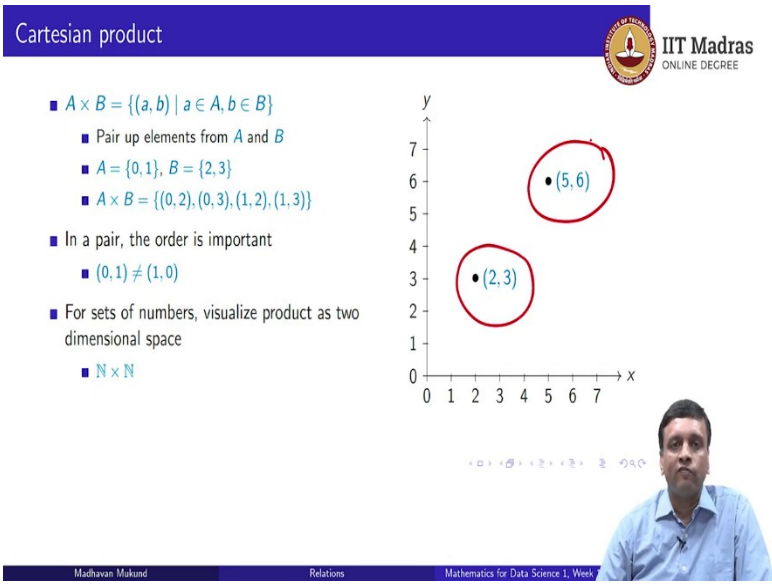


As we saw a set is a collection of items and we can construct new sets from old sets. So, we can take unions combine two sets into one. We can take intersections, take the common elements. We can take the difference that is take the elements of X which are not in Y and if we define the universe with respect to which we are working, we can define the complement those elements that are not in X.

Now, in general we are interested in carving out subsets of a set and so, we use the set comprehension notation. So, what this does is it takes a base set and takes elements of that set, then it applies some condition those elements we are interested in and then it collects them all together. So, we can take all the integers which are divisible by 2 or not divisible by 2 in this case, so we get the odd one; so, those where the remainder is 1.

Or we can take all fractions in which the numerator and the denominator have no common divisor or we can take for instance the real numbers which lie in an interval with [3,17).

(Refer Slide Time: 01:17)



So, now, we will see a new way to combine sets to form new sets and this is called the Cartesian product. And, in the Cartesian product basically what we do is we take two sets and we take one element from each and form a pair. So, A × B as it is called is the set of all pairs which we write with this normal bracket notation (a,b) such that the first element a comes from the big the set A and the second element comes from the set B.

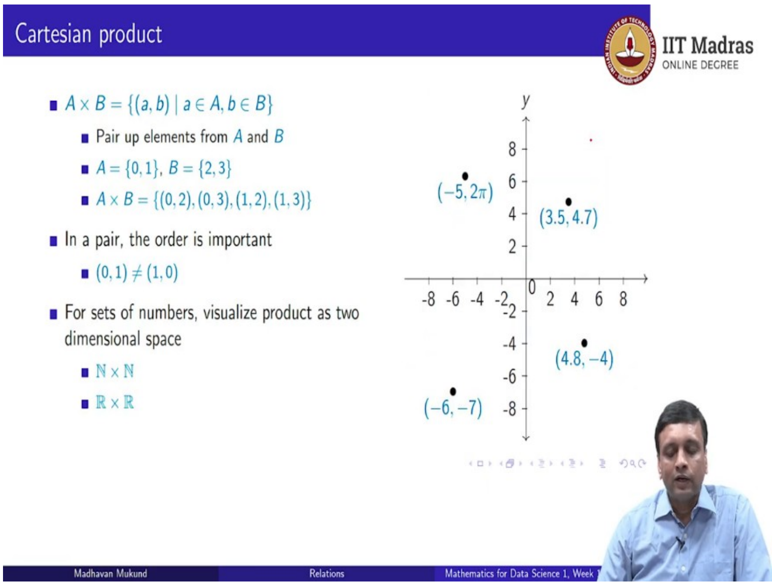
So, for instance, if A is the set {0, 1} and B is a set {2, 3} then all possible pairs we can form in the Cartesian product a 0 combined with 2. So, (0, 2), (0 ,3) and then 1 combined with 2, (1 , 2) and (1,3). So, we have four possible pairs.

Now, in sets we said that the order of the element is not important, but of course, when we are doing this kind of a pairing, then we know that the left set comes from the left part of the product and the right element comes from the right part of the product. So, for example, (0 ,1) is not equal to (1 , 0). So, here we have to respect the order when we talk about a pair.

Now, if we have sets of numbers right, then we normally visualize the product as a space which we draw familiarly as a graph. So, for instance if we take $N\times N$ then we draw $N \times N$ as this grid, where on the x-axis you have one copy of N, on the y-axis you have another copy of N. And, for example, if you want to look at the pair (2 , 3), then such that the x-coordinate is 2 and the y-coordinate is 3 and you get this point and similarly, if you look at the point (5 ,6); you get this point right.

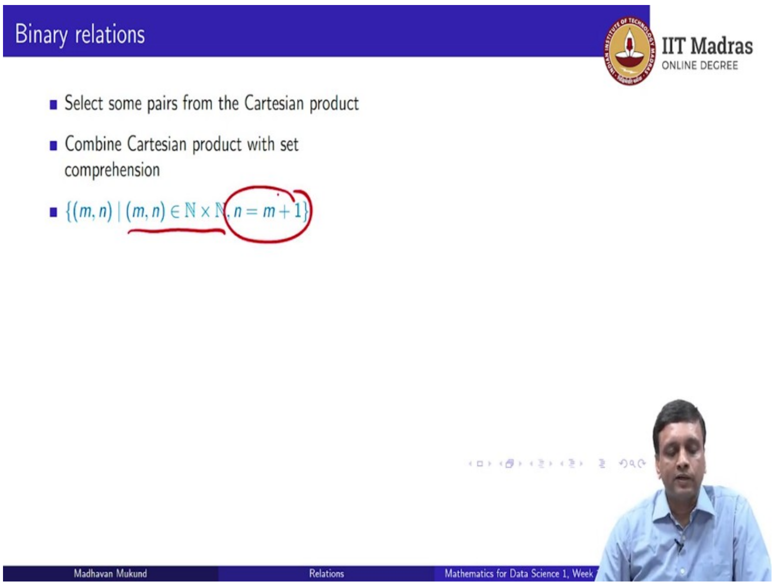
So, you can take the first coordinate plot it on the x-axis; take the second coordinate plot it on the y-axis and where those two points meet in the grid is the point that we are interested in. So, this is one way of visualizing a binary relation on numbers.

(Refer Slide Time: 03:00)



And, we can do the same thing if you are using say the reals, in which case the grid points that we are going to plot will have real coordinates and not just natural number coordinates.

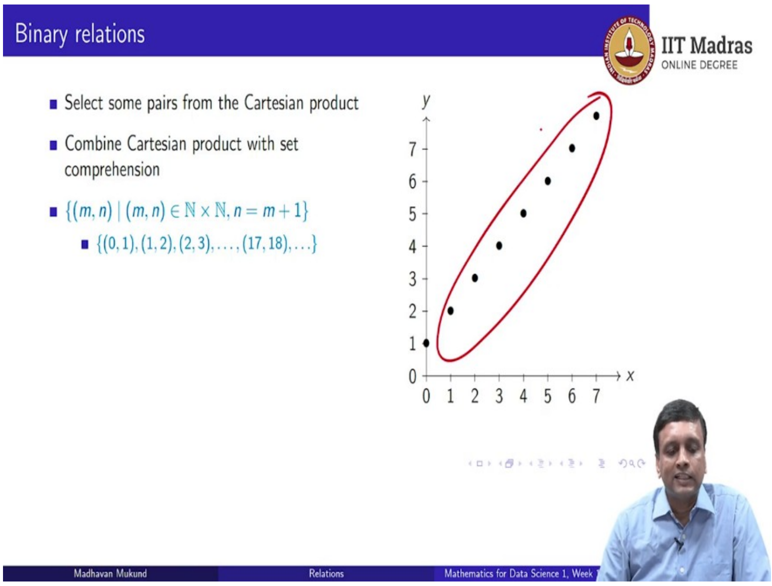
(Refer Slide Time: 03:10)



So, now we have this Cartesian product which consists of all possible pairs of the two sets and as we did with set comprehension we might want to pick out some of these sets some of these pairs and this is what we call a relation.

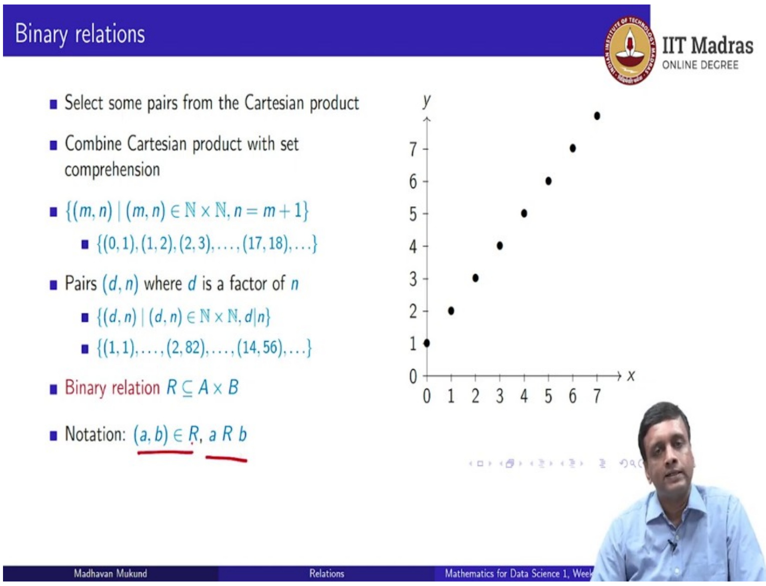
So, we combine this Cartesian product operation with set comprehension. So, for instance, we can take all pairs of numbers which are natural numbers (m ,n), but we want to insist that the second number is 1 plus the first number.

(Refer Slide Time: 03:40)



So, we get for instance (0, 1) because the second number one is 0 + 1; (2, 3) because 3 is 2 + 1, (17, 18) and so on. And, if we plot these points alone on the right then we get these so, we get a subset of the overall points and this these points satisfied this set comprehension condition.

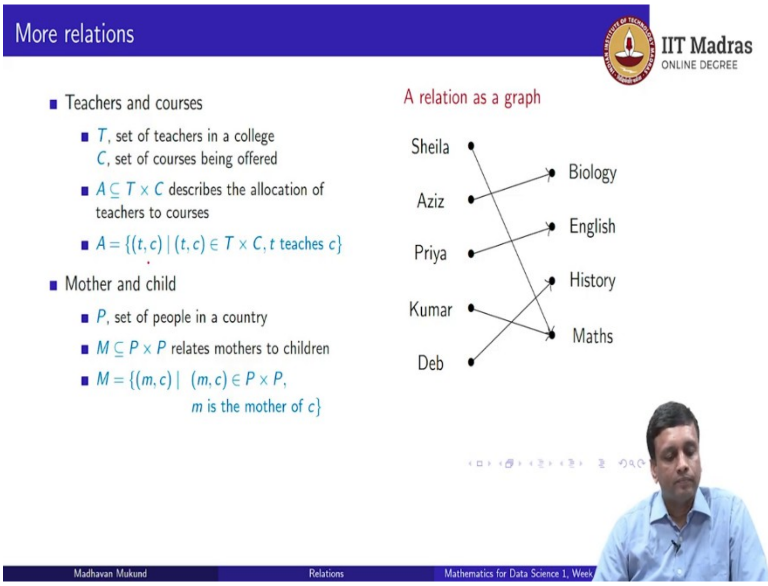
(Refer Slide Time: 03:59)



Another example would be pairs again of natural numbers (d,n), where d is a factor of n. Remember, d is a factor of n means that if I divide n by d, I get remainder 0. So, for instance 2 is a factor of 82, 14 is a factor of 56. So, these will be points in our relation. So, this is what is called a binary relation. So, formally it is a subset of the product. So, we take the Cartesian product all possible pairs and then we apply some kind of a condition which filters out the pairs of interest to us and it gives us therefore, a subset of pairs and this is what we call a relation.

Now, to denote the pairs that belonged to the relation either we can give the name of the relation as a set and say that (a,b) ∈R or sometimes to say that a is related to b, we use R as a kind of operator. We say a is related by R to b and so, we write a R b. So, these are two notations which you might see in different books and they mean exactly the same thing.

(Refer Slide Time: 04:57)



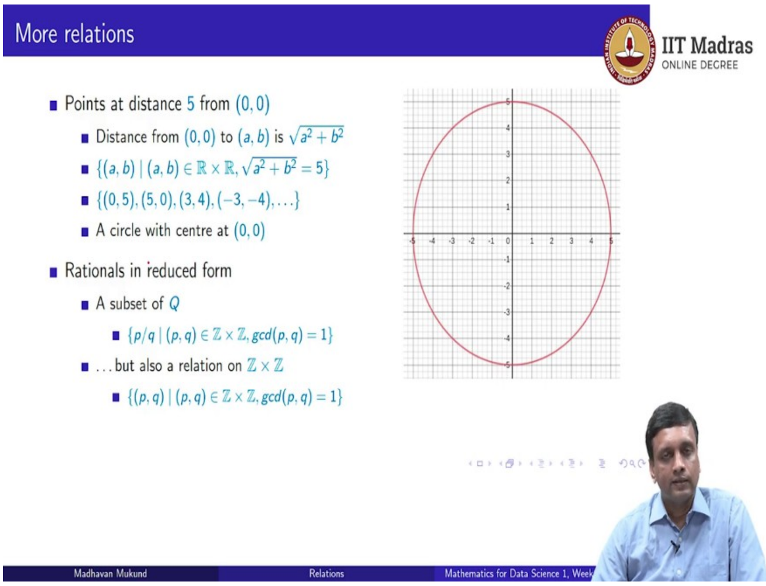
So, let us let us look at some other examples of relations outside the numbers. So, supposing you have a school in which there are some teachers and some courses to be taught. So, T is the set of teachers; C is the set of courses that are being offered in this term, then you need to describe which teachers are teaching which courses. So, we would have an allocation relation A which is a subset of all possible pairs T × C.

So, every teacher and principle could be teaching every course, but of course, this is not normally the case. We do not have all teachers teaching all courses, we have some teachers teaching some courses. So, we would specifically say take every pair of possible teacher course pairs, then we take out those were precisely the teacher T is actually teaching the course C and we collect those together to form this allocation relation.

So, here is a different graphical way of describing a relation not in terms of the grid and the graph that we have learned when we do graphs in school. So, this is also called a graph, but this is a graph in which we have some nodes representing the elements on the set. So, on the left hand side we have five teachers, on the right hand side we have four courses and the arrows from the left hand side to the right hand side connect the pairs which are in the relation. So, we see that Kumar teaches maths; Deb teaches history and so on.

So, this is a useful way of visualizing relations on finite sets and we will see this often as we go along. Another example of a similar type of a relation is that between a parent and a child specifically let us look at mothers and children. So, if we have a set of people in a country, then we can take the set of all pairs of people and then isolate from that pairs in which the first element of the pair is the mother of the second element. So, we want (m,c) which belongs to P × P such that m is the mother of c.

(Refer Slide Time: 06:49)



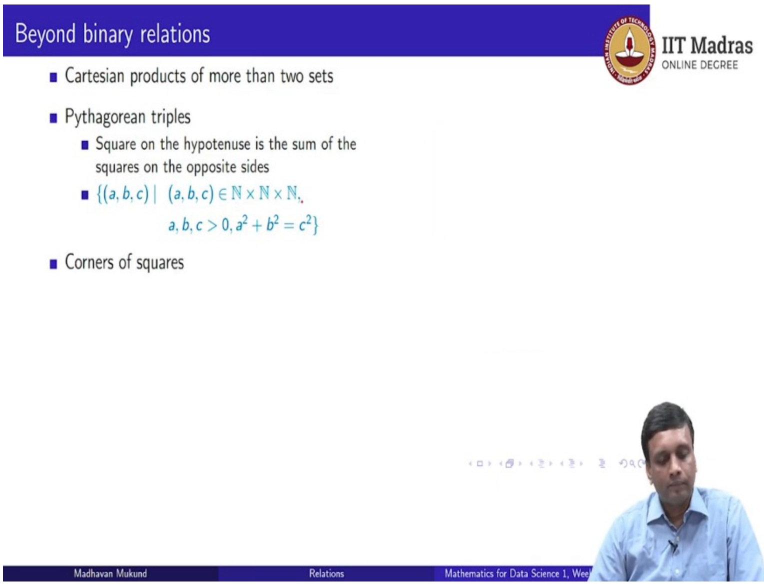
So, let us go back to numbers. So, supposing we want to plot all points which are in R × R which are at a distance 5 from (0, 0) which is normally called the origin. So, one thing you need to know for this we probably you should have learned this at some point is that if I take a point (a ,b) and calculate its difference from (0 , 0). So, this is calculated using the Pythagoras theorem and it comes out to be $\sqrt{a^2+b^2}$.

So, in other words, the relation we are looking for in this case is all (a,b) whose distance from (0 , 0) is 5. So, all (a, b) and R cross R such that $\sqrt{a^2+b^2}$is equal to 5. So, here are some of the points (0,5) for instance you can see (0, 5) is there because the sum is 0 plus 25 and the square root of that is 5. (3 ,4) is there because 3 squared is 9, 4 squared is 16, 9 plus 16 is 25, square root to 25 is again 5.

So, interestingly these points if we plot every such point in R × R which satisfies this actually defines a circle of radius 5 with center at (0 , 0). So, relations can define interesting geometric shapes and very often we do deal with geometric shapes in this relational form because it is easier to manipulate than looking at pictures. Now, depending on how we are going to view a relation, we can look at it in different ways.

So, remember that we looked at rationals in reduced form. So, we said that a rational in reduced form has p / q such that p and q are integers and the gcd is 1 right; that means, that they do not have a greatest common divisor other than 1. But, we can also think of this as a relation on integers itself. We want all pairs of integers. So, every rational is really a pair of integers, the numerator and the denominator and we want every pair of integers where the gcd is 1, that is, there is no common divisor.

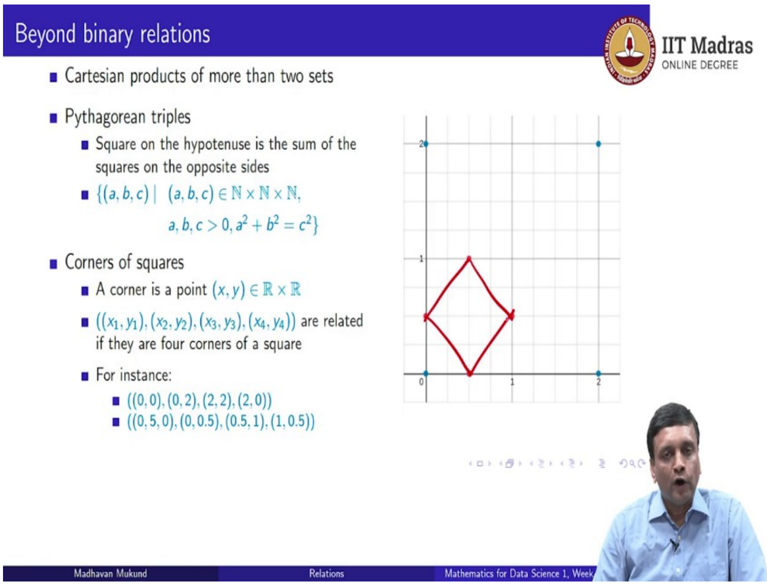
(Refer Slide Time: 08:51)



So, we do not have to restrict our self to binary relations. The Cartesian product notation extends to multiple sets. Let us look at three sets for instance. Remember, Pythagoras theorem which says that the square on the hypotenuse is the sum of the squares on the opposite sides. So, what values of a, b and c could be the sides of a right triangle are determined by Pythagoras's theorem.

So, we would say that a, b and c is a valid triple in the Pythagoras sense if (a, b, c) belongs to $N \times N \times N$. So, here we now have three copies of N and a, b, c must all be nonzero. They must all be positive length we do not want to have triangles in which one line one side is collapsed to a point and we want the constraint that $a^2+b^2=c^2$.

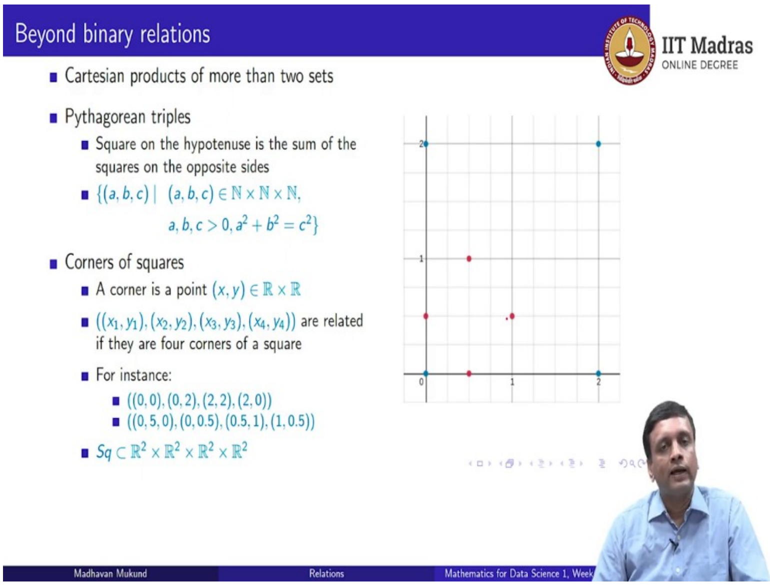
(Refer Slide Time: 09:41)



Here is another example. Suppose, we look at squares on the plane squares with real corners right. So, a corner is a point (x , y) which is in $R \times R$. So, we define the x coordinate and the y coordinate that defines the corner of a square and we want four corners which together form a square if we connect them by lines. So, for instance, if you look on the right the four blue dots correspond to a square which is cornered at (0, 0); (0, 2); (2, 0) and (2, 2).

The red square is also red points also define a square because this is a rotated square, but then if you rotate it vertically; you will turn out that this diamond is actually a square. So, there are many such four sets of points which form the corners of squares and we might be interested in all such four sets of points. So, now, we have a relation which involves four sets of points, but each point itself is a pair of real numbers; it is an x and a y.

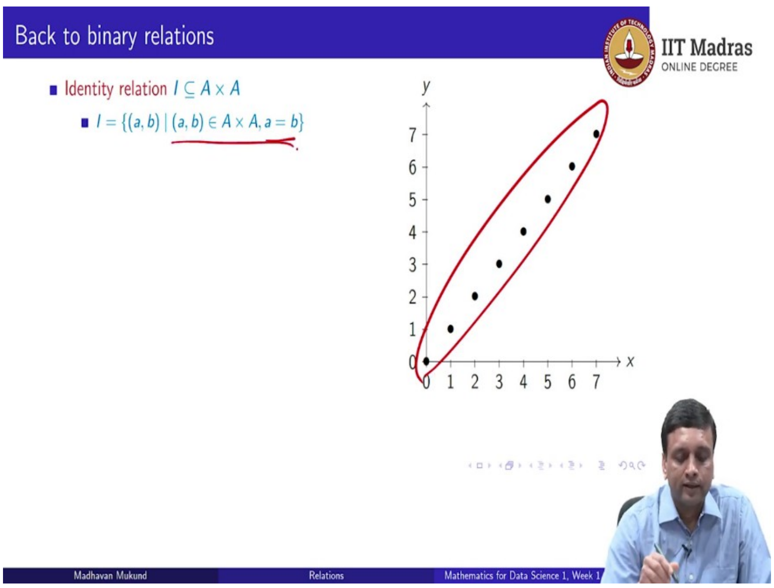
(Refer Slide Time: 10:35)



So, square if we think of it as a relation is actually a relation on R^2 that is the first corner times R^2 the second corner times R^2 the third corner and the fourth corner R^2 again. So, this is actually either a relation on eight copies of R or if you want to group it four copies of pairs of R.

So, this just says that we can take relations on arbitrary an arbitrary number of copies of a set and we get larger and larger from pairs, we move to triples we move to quadruples and in general if we have n copies we call this an n tuples.

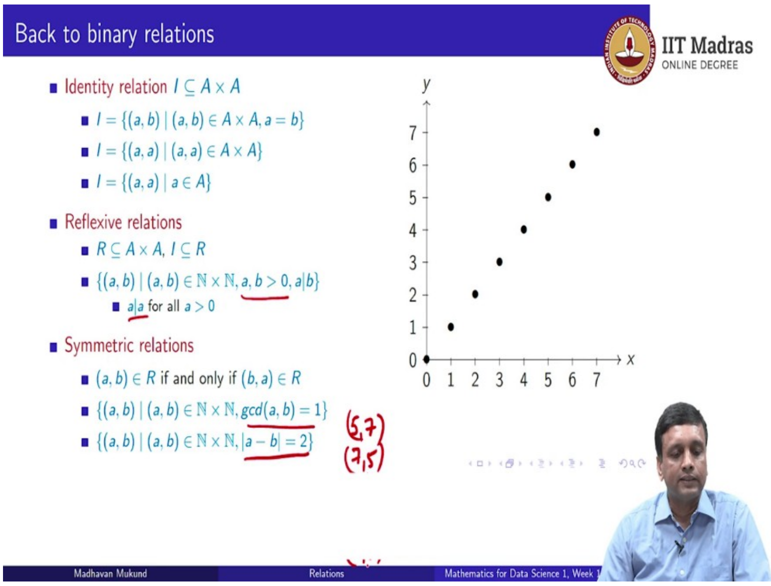
(Refer Slide Time: 11:11)



So, there are some special binary relations which pop up all over the place. So, it is useful to know their names. The first one is called the identity relation and as you would expect, the identity relation maps every element to itself. So, if I take A × A, so, first of all the identity relation is defined on two copies of the same set because identity means equality. So, I take A × A. So, this has all kinds of pairs (a , b), where both a and b belong to A and now, I want the condition that a = b.

So, in other words, I want things of the form (a , a). So, if I plot this for instance on the natural numbers and $N \times N$, then I get (0, 0); (1, 1) and so on, and these are the points which are drawn on the right in this grid.

(Refer Slide Time: 12:05)



Now, point of notation we sometimes it is tedious to write this notation as it is says us (a,b) time in n. So, we do not want to you know have to write out this long thing. So, sometimes we simplify this by saying I want all pairs such that a comma a belongs to A × A. So, what we are really saying is that the second day and the first day must be the same. So, we are collapsing the equality and this. Now, this is not technically correct, but this is often used in order to simplify the notation.

And, sometimes we might drop the product altogether. We might just say we want all pairs (a ,a)where a comes from the set A. So, in other words we are pulling out one copy of the element from the set and then we are constructing a pair by taking two copies of it. So, all of these are equivalent ways of writing this although only the first one technically follows the notation that we are using to introduce relations.

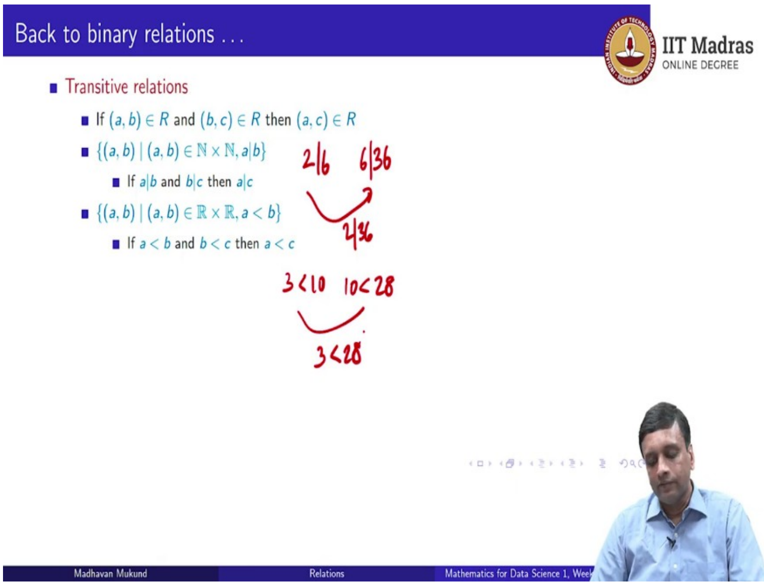
Now, there are some properties that relations may have. The first one is called reflexivity. So, reflexivity refers to the fact that an element is related to itself. So, a reflexive relation is one in which for every element a; (a,a) belongs to R. So, in other words based on what we just wrote above, it means that the identity relation is included in R. So, it does not mean that is the only thing. The identity relation has only the reflexive elements. A relation that is reflexive will have the identity pairs and it will have other pairs, but it must have all the identity pairs to be called reflexive.

A symmetric relation for instance is one where if (a, b) is there, then (b, a) must be there. So, for instance looking at reflexive relations, one example is the division relation. So, if we provided we make sure that the numbers are not 0, then we know it is reflexive because every number divides itself. So, if we take the reflect division relation as the relation that we introduced in the first part of this lecture that would be reflexive because a divides a for every a which is not 0.

Similarly, symmetric relations if we look at pairs where the greatest common divisor is 1, in other words they have no common divisors. This is what happens for example, in reduced fractions, then it does not matter whether we write it as (a , b) or (b, a). So, if (a, b) has greatest common divisor 1, so does (b , a). So, (a, b) and (b, a) must both either be there in the relation or neither will be there.

Similarly, if we look at this which is asking about the absolute value so, it is saying give me all numbers a and b such that a - b is either 2 or -2. So, the absolute value takes the difference and removes the negative sign. Now, we see that for instance if (5, 7) is there, then (7, 5) must be there because they both have the same difference depending on how we write it. Normally, in subtraction we have a sign difference, but because we are taking the absolute value there is no difference actually between these two. So, this absolute value relation also if we fix is a symmetric relation.

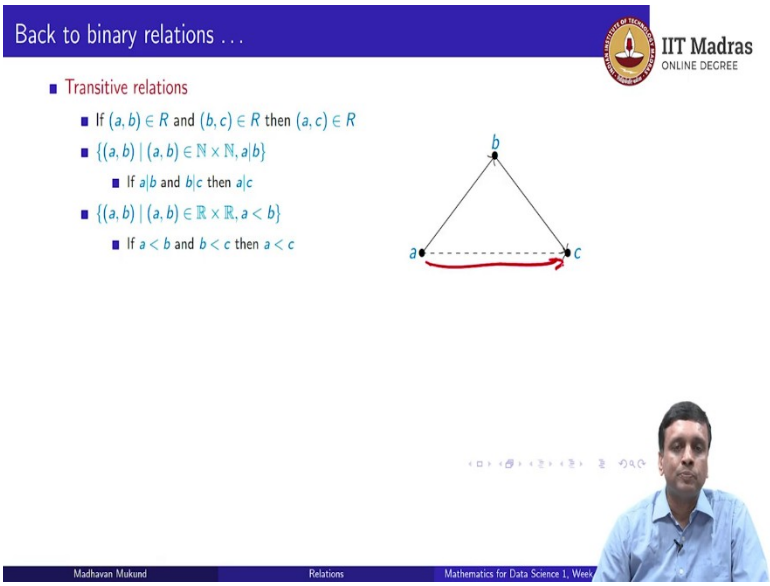
(Refer Slide Time: 14:51)



A third property that relations may have and which are useful is called transitivity. So, transitivity says that if we have two pairs which are related such that they share an elements. So, a is related to b and b is related to c, then a must be related to c. So, again our divisibility is a relation. So, supposing we say that 2 | 6 and we say that 6 | 36, then from this we can conclude that 2 | 36 as well, right.

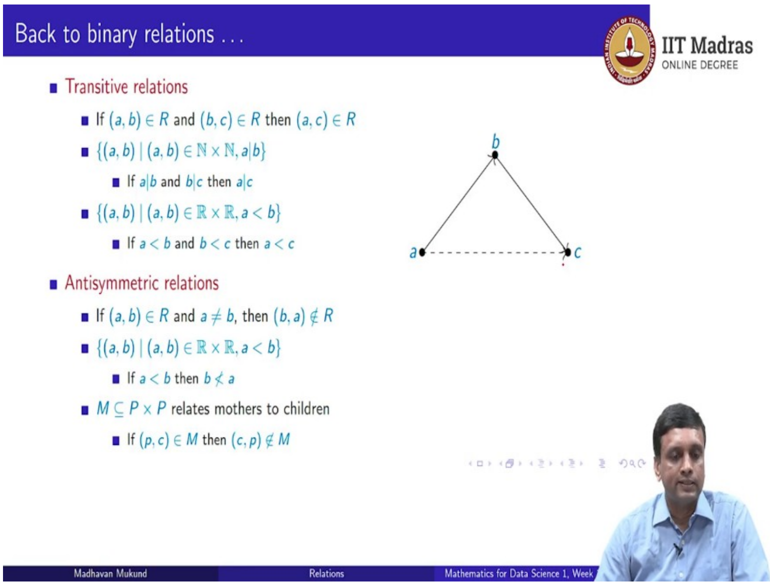
Similarly, if we take less than if we say that 3 < 10 and 10 < 28, then we know from this that 3 < 28. So, this is transitivity.

(Refer Slide Time: 15:34)



So, if we want to draw it pictorially if we have three elements a, b and c and this arrow remember we had this graph notation which says a is related to b and b is related to a, then this dashed line represents the requirement for transitivity a must be related to c.

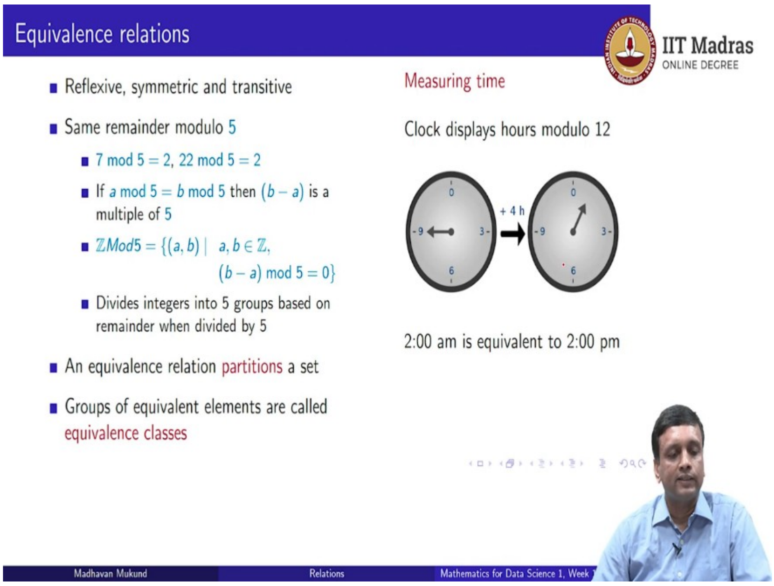
(Refer Slide Time: 15:49)



Now, we saw symmetry. So, symmetry says that if (a, b) is in R, then (b, a) must also be in R. Anti-symmetry says something different it says if (a, b) is in R, then (b, a) should not be in R. So, less than for example, which was transitive above is also anti-symmetric. If you take strictly less than, if a is strictly less than b; then it cannot be that be strictly less than a. So, this is an anti symmetric relation, but anti symmetry does not require that one of the two must be there. It only says that if one pair is there the opposite pair should not be there ok.

Similarly, if we look at our mother and children example; obviously, if p is the mother of c then c cannot be the mother of p ok. Now, there may be p and c such that neither p is the mother of c nor is c is the mother of p. So, that is allowed. We do not insist that every pair (p, c) must be related one way or another, but if it is related one way it should not be related the other way is what anti-symmetry says.

(Refer Slide Time: 16:46)



So, if we combine some of these conditions, we get an interesting class relations called equivalence relations. So, equivalence relation is something that is reflexive, symmetric and transitive. So, as an example supposing we connect together all numbers which have the same remainder modulo 5. So, for instance 7 has a remainder 2 with respect to 5 and so does 22. So, 7 and 5 would be related in this way if we define the relationship as having the same remainder modulo 5.

Now, notice that if two numbers have the same remainder modulo 5; that means, that going from one number to the other you are going in multiples of 5. So, for instance 22 -7 is 15 right. So, this is this modulo arithmetic. So, if you add the number that you are dividing by, then you get the same remainder and so, in set notation we can say that the integers modulo 5 are all pairs a, b such that b - a mod 5 is 0. In other words, we are not asking what is the actual remainder of b and a, we are just saying that b and a are separated by a multiple of 5 therefore, they must have the same remainder modulo 5.

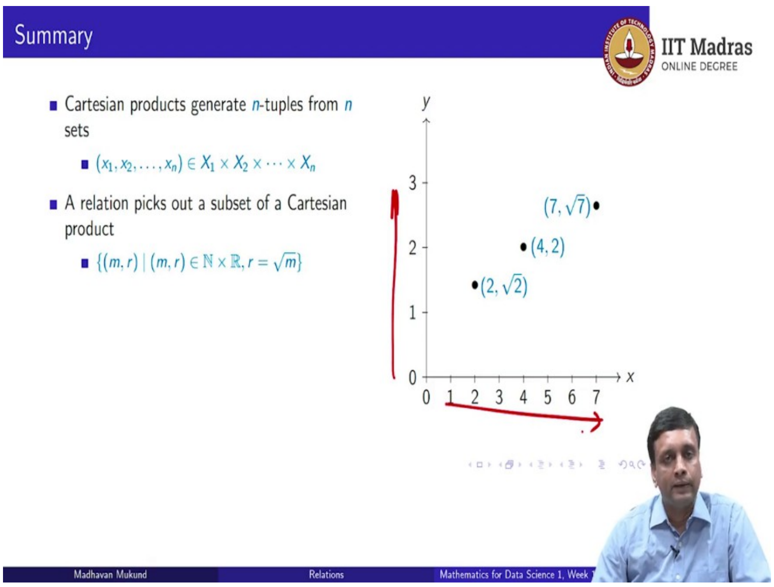
Now, this divides the integers into five groups if I based on the remainder. So, there are the group of numbers which are divisible by 5, they have remainder 0. Those like 6, 11 and all which have remainder 1; 7, 12 and all which one remainder 2 and so on. So, we have five possible remainders 0, 1, 2, 3, 4 and therefore, this divides the set of integers into five disjoint classes.

As an example of modulo arithmetic that we are all familiar with, consider what happens when we look at a normal clock. Now, a normal clock measures time from 0 to 12 and then cycles around again. So, though there are 24 hours in a day, the clock is actually partitioning these 24 into two sets where we have 0 and 12 as same, 1 and 13 as same and so on right. So, 2 am and 2 pm, there is no distinction on the clock.

So, the clock is actually showing us this equivalence class of hours regarding am and pm as being equal and we have to know from context whether the clock is showing am or pm. So, the main thing to note about an equivalence relation is that it partitions a set. It partitions a set into disjoint groups, all of the elements within a group are equivalent and all of the elements outside across groups are not equivalent to each other.

So, the groups of equivalent elements that we formed through an equivalence relation are called equivalence classes. So, this might look a little abstract now, but equivalence classes really represent a kind of equality and sometimes we are happy to work with this equality in terms of equivalence relations rather than actual equality and it has very much the same properties as equality does.

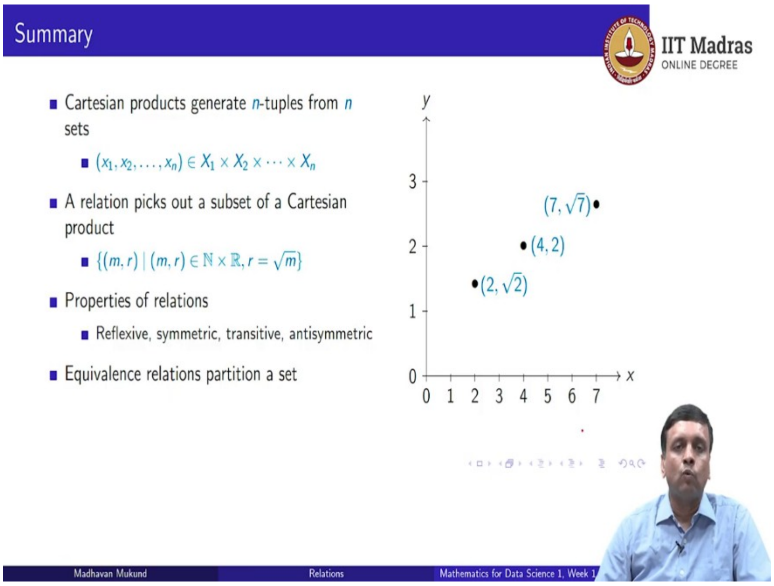
(Refer Slide Time: 19:28)



So, to summarize as we have seen a Cartesian product can generate n-tuples of elements from n sets. So, if we have X_1, X_2 , up to X_n, n sets these can be different or the same, then we can take one element from each set and form an n-tuple x_1, x_2, up to x_n. And, when we now pick out some particular subset of these n-tuples, we get a relation. So, for instance, if we take pairs from $N \times R$ and we want the second element of the pair the real number to be the square root of the first element, then we get $N \times R$ such that $r = \sqrt{m}$.

So, here on the right we have seen we show one picture of this. So, there are some elements like $(2, \sqrt{2})$, $(4, 2)$, $(7, \sqrt{7})$ and so on. Now, just notice that in this picture the y-axis is elongated compared to the x-axis. So, this is not in some sense to scale in both dimensions because the square root function behaves like this.

(Refer Slide Time: 20:27)



So, we have seen that there are some properties that we would like to record of binary relations – reflexivity, symmetry, transitivity and sometimes anti-symmetry. And, using reflexivity, symmetry and transitivity together we get what is called an equivalence relation, an equivalence relations partition sets into equivalence classes which behave like equality.

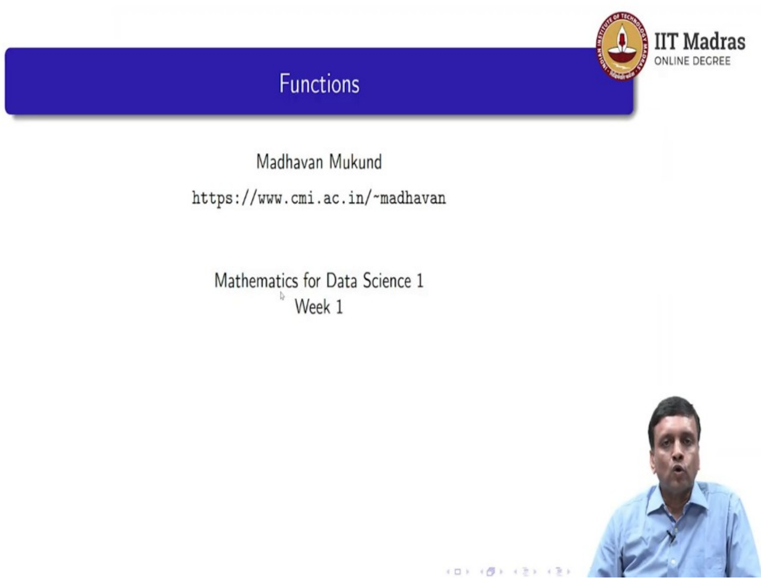
Thank you.

Mathematics for Data Science 1
Prof. Madhavan Mukund
Department of Computer Science
Chennai Mathematical Institute

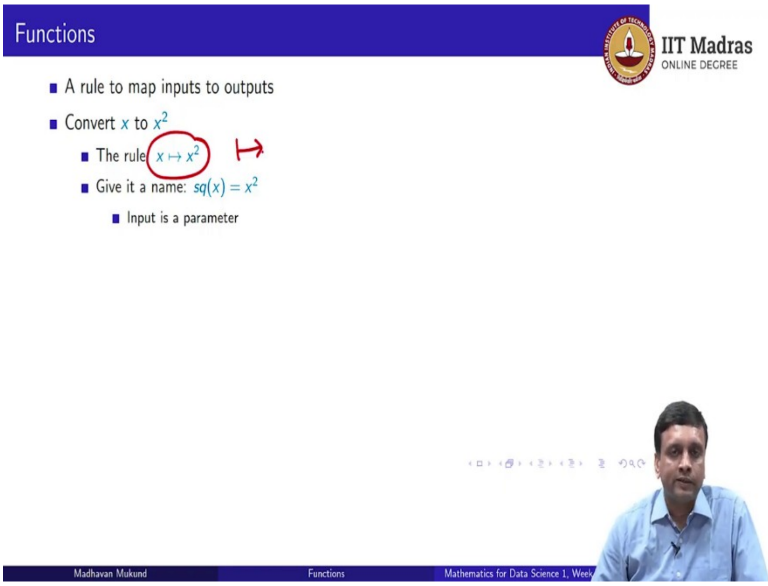
Week - 01
Lecture – 07
Functions

(Refer Slide Time: 00:05)

(Refer Slide Time: 00:14)



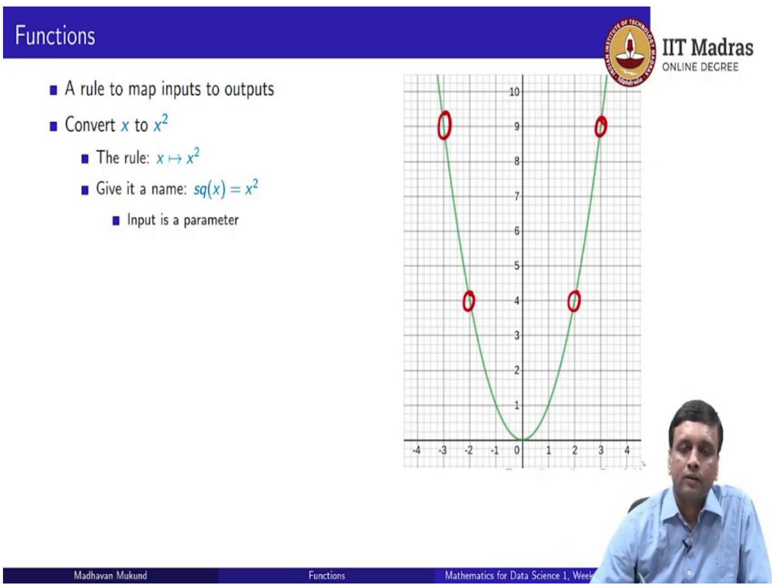
(Refer Slide Time: 00:17)



So, closely related to relations are functions. So, what is the function? A function is a rule that tells us how to convert an input into an output. So, for instance suppose we want a function that given an x returns as x^2, then this is one way to write the rule. We write this symbol which says x maps to x^2; given an x it is transformed to x^2, but more conventionally we also give a name to the function. So, in this case we can call it square(x).

So, square(x) takes a parameter x as input and it produces as output; some value which transforms this parameter, in this case x^2.

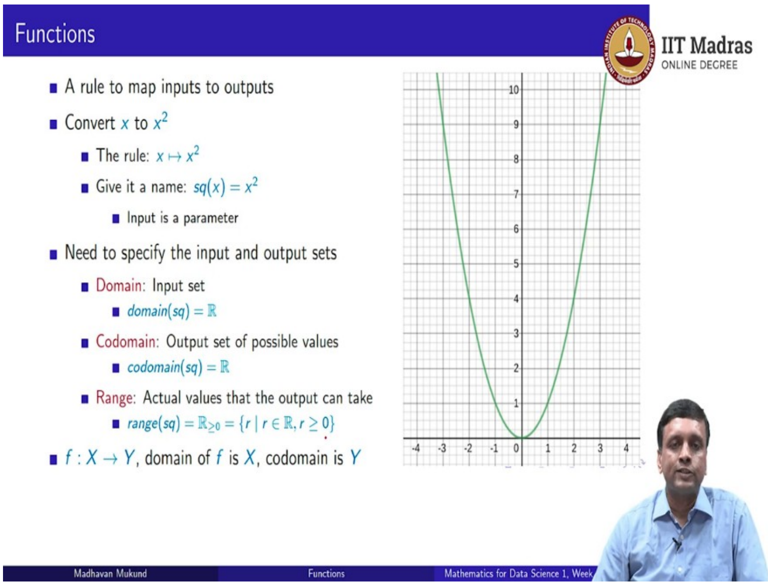
(Refer Slide Time: 00:56)



So, we can plot x versus x^2 by putting all the points where the second coordinate is the function value of the first coordinate. So, if we look at x^2 for instance, it forms this up you know inverted parabola shape which you should be familiar with. And notice that because for instance 2^2 is the same as $(-2)^2$, there is a symmetry about the y axis.

So, for instance 2^2 is the same as $(-2)^2$, and 3^2 would be the same as $(-3)^2$ and so on.

(Refer Slide Time: 01:31)



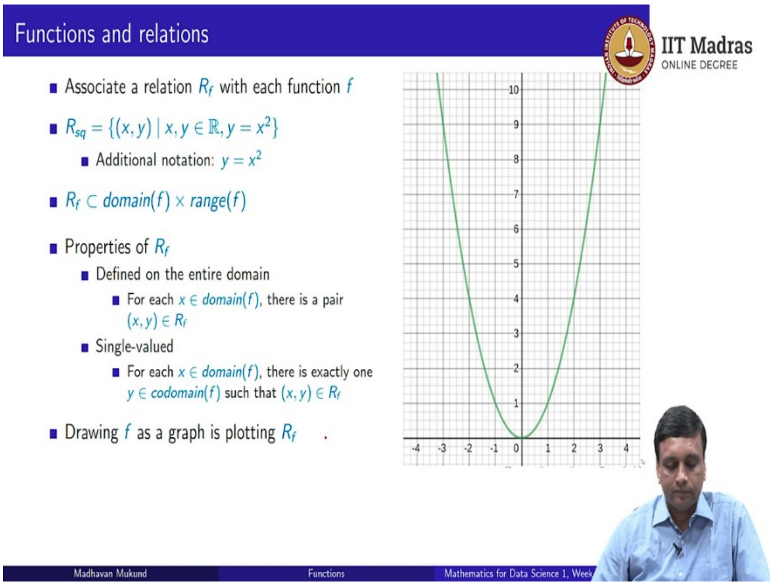
So, when we define a function, we have to be careful about specifying what set we take the input from and what sets the output produces. So, the input set is called the domain. So, for instance the domain of square as we have defined it above is a set of reals, so we can take the square of any real number.

Now the output when we apply square, we know that it is going to be a real number; so the codomain as it is called is the output set of possible values is called the codomain, in this case is the reals. But of course, we know that when we square a number; even if the input is negative, the output is going to be positive. So, even though the codomain is a set of all reals, we cannot get all reals as output of the square function. So, there is a separate name for that called the range.

So, the range of a function is a subset of the codomain; the range tells us what values the function can actually take. So, in this case the range of the square function is the non-negative reals. So, this is all real numbers greater than equal to 0 which is sometimes written like this and if you want to explicitly write it out; it is the set of all r in the set of reals such that r >= 0.

So, in order to specify a function abstractly and describe its domain and codomain, we usually write that f which is the name that we give to an arbitrary function is a function from X the domain to Y the codomain. So, this notation f : X to Y tells us without telling us what the function is actually doing; it tells us on what sets it operates, what is the input set and what is the output set.

(Refer Slide Time: 03:08)



So, the close connection between functions and relations is that we can associate with every function f a relation R_f; and R_f is merely all the pairs of inputs and outputs that the function allows. So, for example, with our square functions sq we have R_{sq} as all pairs (x,y), such that y is equal to x^2.

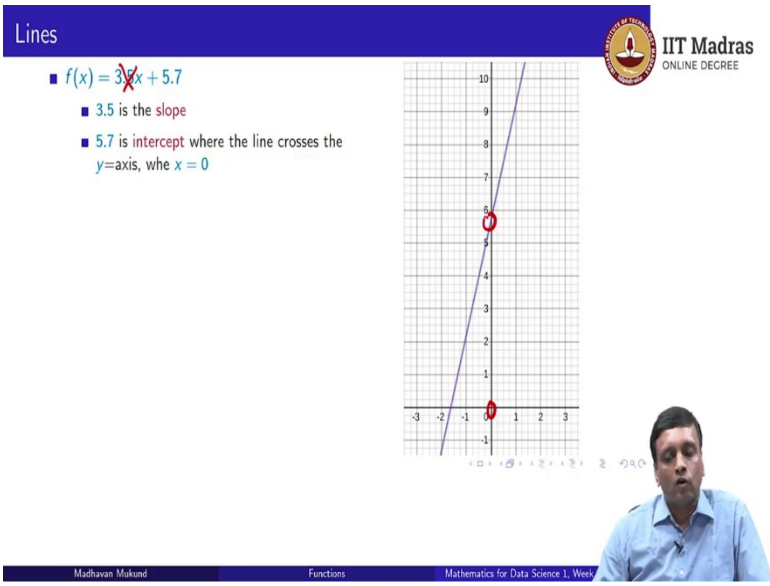
So, this is actually sometimes simplified by saying y is equal to x^2. So, we do not write out f(x) and then say f(x) is y; we just directly say y is equal to x^2 to denote that the output is the square of the input. So, this is an implicit notation, where we are implicitly naming the output for each x as y. So, notice that if we talk about it as a relation; remember that a relation is a subset of the Cartesian product of two sets. So, in this case, the Cartesian product is formed by the domain of the function and the range of the function, and then the relation is a subset of the domain X the range.

So, what are some properties of this relation? Well, first of all when we define the domain of a function, we really mean that the function is defined at every possible value in that domain. So, for every x and domain of the function f, there must be a valid value f(x); so there must be a y such that (x, y) belongs to the relation R_f. The other property is that this is a rule for producing an output from an input; so there can be no confusion about what the output is.

So, for each x that we feed in as a domain value to the function, there must be as exactly one output value f(x) that we get out. So, there is only one y in the codomain, such that (x,y) belongs to R_f. And in fact, we saw in the lecture on relations that, we would draw relations by plotting the points which form part of the relation. So, technically when we are drawing a graph of a function as we have done here for this parabola, we are actually drawing all the points which satisfy the relation R_f.

So, plotting a graph is the same for functions and relations; because implicitly we are plotting the relation that corresponds to a given function.

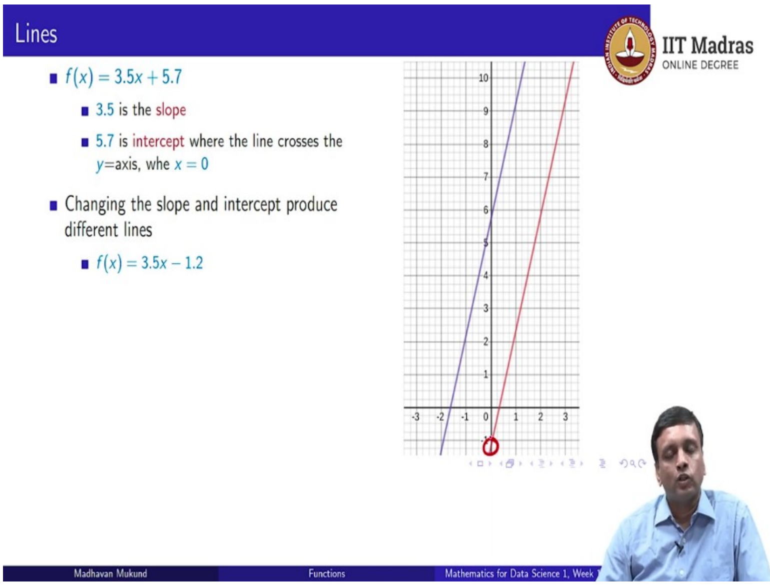
(Refer Slide Time: 05:23)



So, let us look at some other functions that we will encounter as we go along. So, if we have a function of the form something x + something. So, mx + c, then this defines a line. So, then the like we see a line 3.5x + 5.7. And what we will see as we go along in this course is that, the quantity which multiplies x is called the slope and it determines the angle at which the line goes; and the other quantity which is without x determines the intercept.

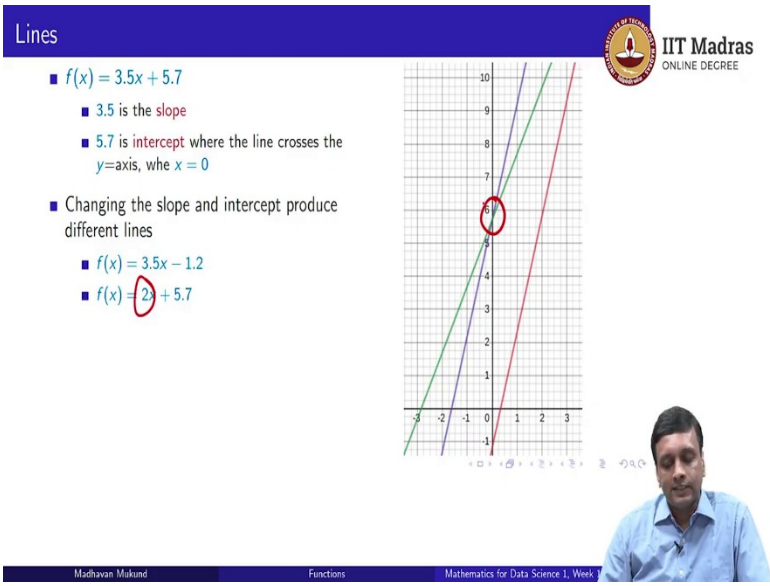
So, notice that if you set x = 0, then the first term goes to 0; this gets cancelled out, if x is 0. So, the answer will be 5.7. So, when x is 0, you get 5.7. So, what the second term tells us is where this line crosses the y axis.

(Refer Slide Time: 06:09)



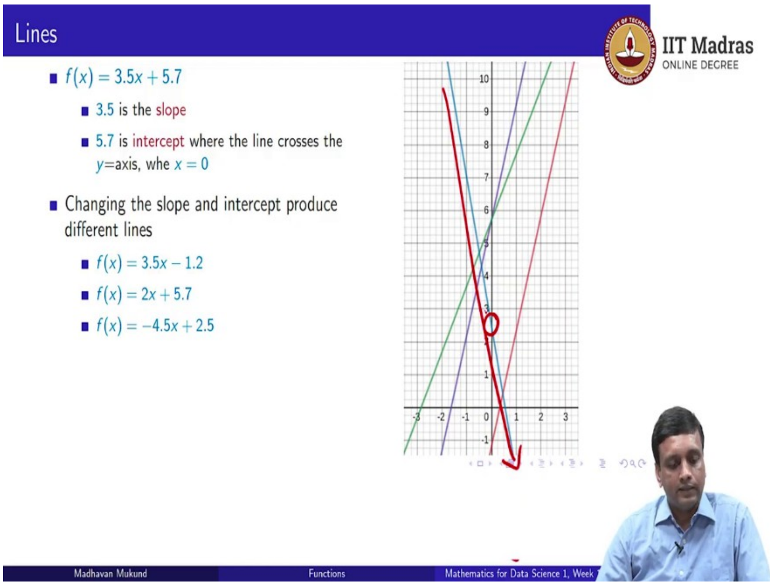
So, if we change these two values, we get different lines. So, for instance if we change the intercept and keep the slope the same; then we get a line which has the same slope it is parallel, it is at the same angle. But now the intercept is -1.2; so it crosses the y axis lower, so the whole line is shifted to the right.

(Refer Slide Time: 06:28)



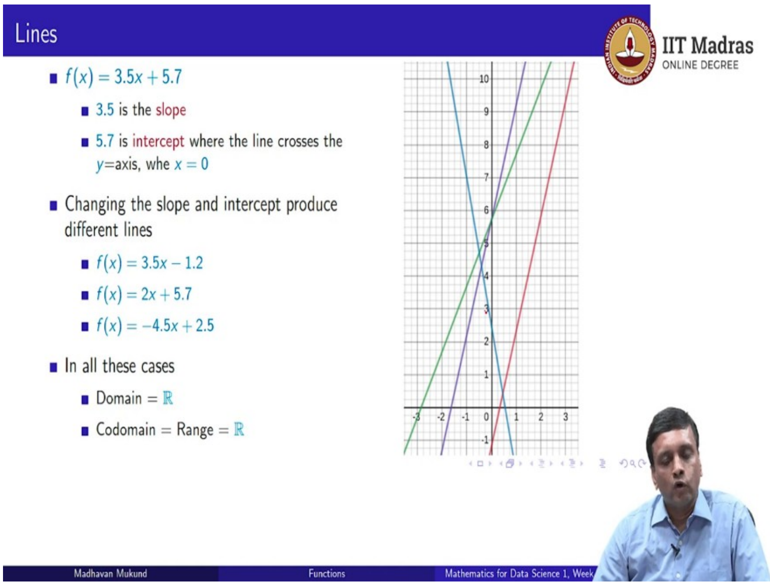
On the other hand if we keep the intercept the same; but we change the slope, we get a different slanted line. So, here we have reduced the slope from 3.5 to 2; so it is a shallower line and the green line passes through exactly the same point 5.7 as the previous one, but it has a shallower slope.

(Refer Slide Time: 06:49)



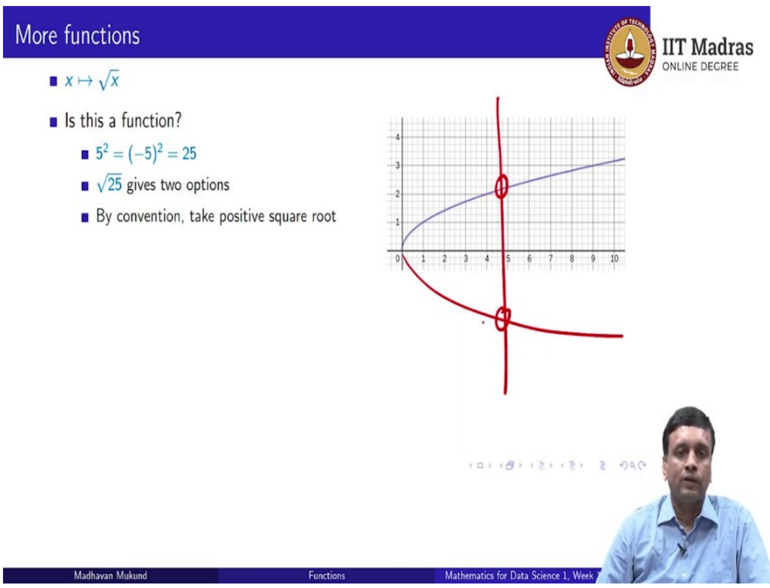
And we can change both and in fact, we can put a negative slope; so if you have a negative slope, it comes down rather than going up, so we have this line coming here. And notice that it crosses at 2.5, so that is the intercept. So, by changing the values of the slope and the intercept, we get many different lines and many different functions.

(Refer Slide Time: 07:07)



And for all of these functions that we have defined the domain is the set of reals, the codomain is the set of reals; but also because we can intuitively see that the line goes from way down -∞ to way up +∞ whether it is going up or down, it can take all values in the real. So, not only is the codomain equal to R, it is also the range.

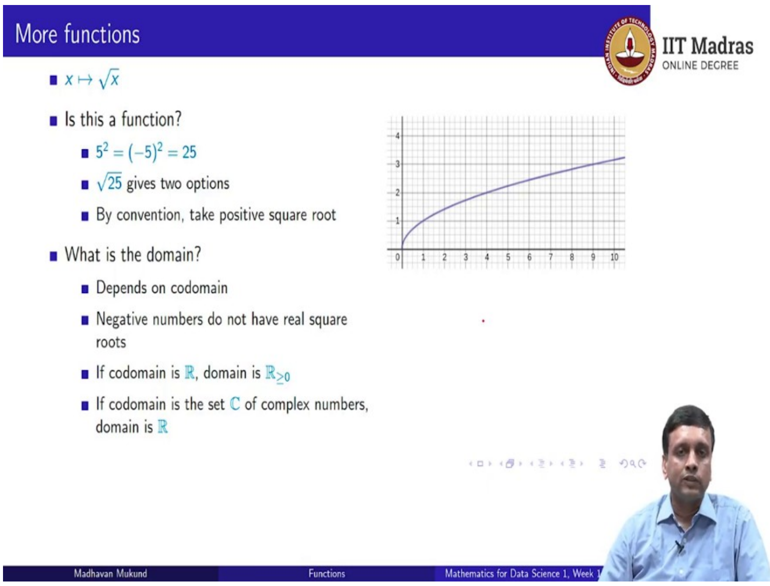
(Refer Slide Time: 07:28)



So, here is another function x maps to √x. The first question is, is this a function? So, remember that for a function, we need it to be defined on every input value and we also needed to have a unique output. So, remember that when we square a negative number, we get the same as when we square the positive version; so 5^2 and $(-5)^2$ are both 25.

So, technically if we take $\sqrt{25}$, we cannot determine whether we are talking about +5 or -5. So, when we write $\sqrt{x}$ as a function, our convention is that we are taking the positive square root. So, the function on the right plots the positive square root; if we were to take the negative square root, then it would be a symmetric curve going below. And now if we take both these together, then this is not a function; because if we take any x value, we have two possible outputs for this which is not allowed. So, we are taking by convention the positive square root.

(Refer Slide Time: 08:23)

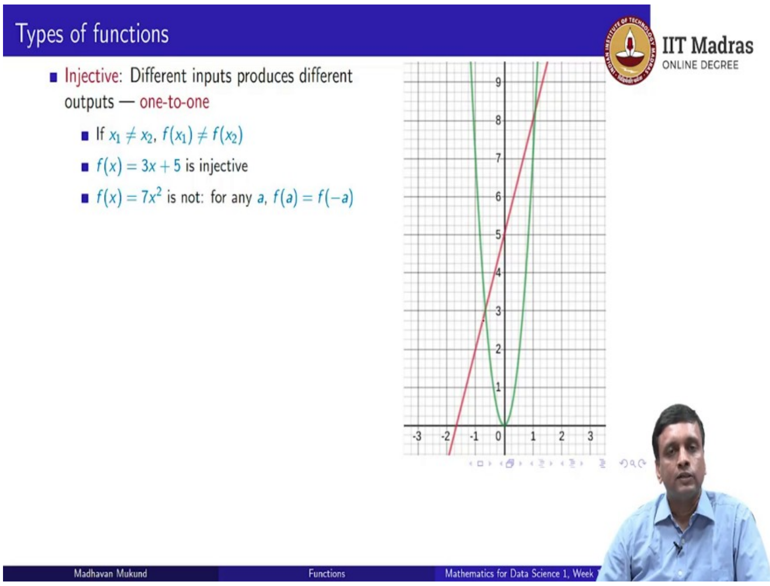


Now what is the domain of this function? Well it depends on what we allow the codomain to be. We have seen that negative numbers cannot have real square roots; no real number can multiply itself to produce a negative number, because of the law of signs for multiplication. So, if we insist that the output should be a real number, then the domain of this function, the function can only be defined when the input is not negative. So, we have this set which we defined before; the set of reals bigger than or equal to 0.

On the other hand, if we move to the set of complex numbers which we said we are not going to describe in detail; the set of complex numbers includes $\sqrt{-1}$ and implicitly through that the square root of all negative numbers. So, once we allow complex numbers as the output of our function, then we can define square root on all the real numbers.

So, the notion of domain and range is kind of flexible depending on how we are going to use the function. So, we have to be very clear when we are using a function what context we are using it in.

(Refer Slide Time: 09:24)

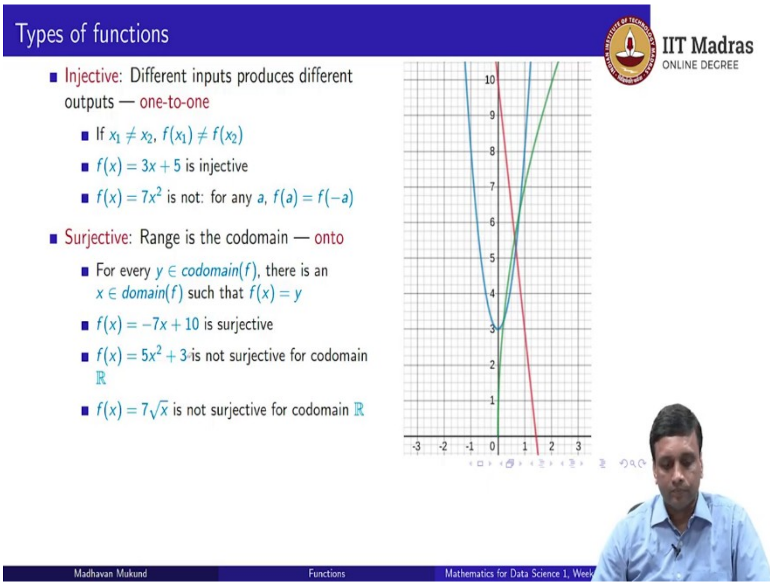


Now we saw when we looked at relations that there are some properties of relations which are interesting like reflexivity, symmetry and so on. Similarly there are properties of functions which are interesting; the first interesting property of function is whether it is one to one, whether it is injective.

What this means is; if I give you different inputs, does the function always produce different outputs? If $x_1 \neq x_2$, is it guaranteed that $f(x_1) \neq f(x_2)$? So, if we look at the linear function that we saw before the line, then we can see that it is injective; because if we change x, we move along the line to a new point. So, no two x points, point to the same y point; so therefore, this is an injective function.

If on the other hand, we take a parabola as function which of the other form something squared, so $7x^2$ for instance. Then we already saw that f(a) is the same as f (-a), so there will be two points; the plus version and the minus version, both of which has the same output. So, it is not the case that distinct outputs produce distinct, inputs produce distinct outputs; so the square function is not injective.

(Refer Slide Time: 10:32)



On the other side we talked about the distinction between the codomain and the range; we said that the codomain is the set of values into which the function produces answers, but the range is the actual set of values of the functions can take.

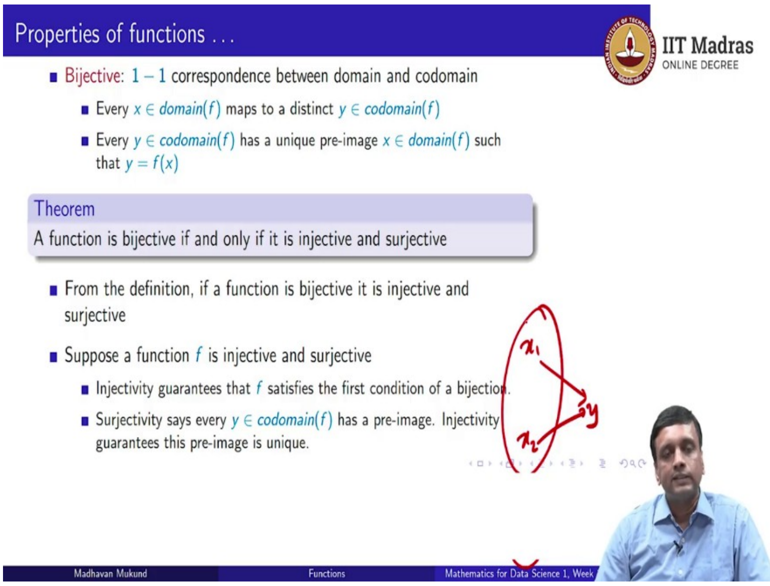
So, the question is, whether or not all values in the codomain are actually touched by the function and this is called surjectivity or onto. So, the range of a surjective function is in fact equal to the codomain, which says that for every y which is in the possible codomain of f; there is actually an x in the domain of f, such that f (x) = y.

Now, once again if we take a line, then this is surjective; because if I pick any point y, I can find a point x, I can solve for x for example, which gives me that y. On the other hand if I take a parabola, in this case we have shifted the parabola up, so it is $5x^2+3$. Then we can see that, first of all a parabola with no shifted, if I did not have this +3 term; then we know that it can only take positive values, because x^2will always be a non-negative number.

Now if I further add +3, it can only take values 3 and above; so this definitely is not surjective, the domain codomain is a set of all reals, but the actual range is only if the reals which are bigger than or equal to 3. Similarly if I take this $7\sqrt{\square}$function, then we know that even if we take the codomain to be R; so we only take square roots of positive numbers. We know that we will never get a negative answer, because by convention we have taken positive square roots.

So, this is again not a surjective function. So, these are two important properties of functions, are they injective is it one is to one; if I give you different inputs, do I get different outputs and is it surjective, is it onto, does every possible output have a corresponding input that maps to it.

(Refer Slide Time: 12:25)



So, if you combine these two, you get something called a bijective function. So, a bijective function is something with where there is a one to one correspondence between the domain and the codomain.

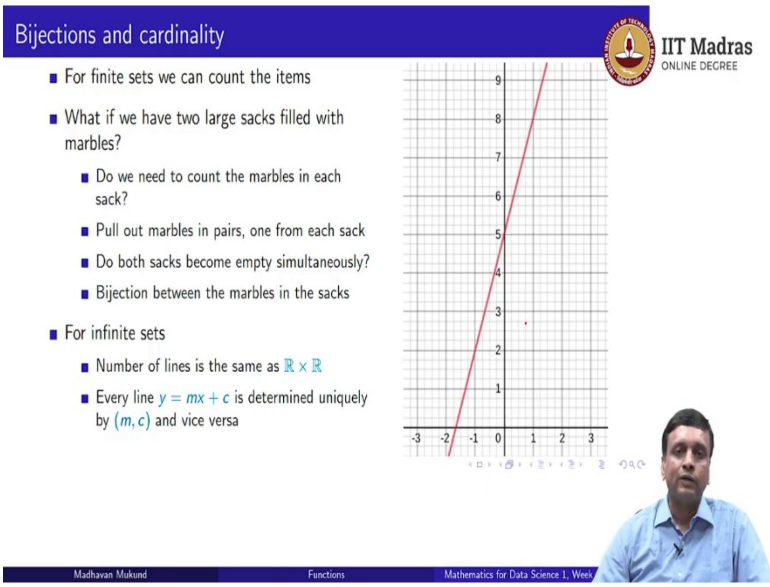
So, every x in the domain maps to a distinct y in the codomain and every y in the codomain has a unique x that maps to it. So, from the statement it looks clear that this corresponds to injectivity and surjectivity. So, actually this is the theorem that a function is bijectuve if and only if it is both injective and surjective.

Now this may look obvious, but actually only one direction is obvious, from the definition, we can see that if a function is bijective; it must be injective, because it says every x maps to a distinct y, so no two x will map to the same y.

It also says it is surjective, because it says every y in the codomain has a unique pre image. So, the fact that a bijection implies injectivity and surjectivity is part of the definition; the other way requires a small argument. So, supposing a function is injective and surjective, we have to show that it is bijective. So, for this, we have to guarantee first that every x maps to unique y; but this is guaranteed because the function is injective, injectivity says if I have two inputs x_1 and x_2 which are not the same, $f(x_1) \neq f(x_2)$. So, this is fine.

What about surjectivity? So, surjectivity says that everything in the output comes from some input not necessarily unique; but if two things map to the same output right, if two things map to the same output, if I have a y such that I have x_1 and x_2 mapping to the same y. So, if it has even, if a surjective function if the output has two pre images; then these two pre images do not satisfy injectivity. So, if I combine surjectivity in the presence of injectivity, I know that the pre image is unique; and therefore these two conditions guarantee that I have a bijection.

(Refer Slide Time: 14:15)



So, an important use of bijection is to count the items in a set. So, remember we said that the cardinality of a set is the number of items and if you have a finite set, we can count them. Now supposing somebody gives you two large sacks filled with marbles or balls and ask you to check whether the two sacks have the same number of balls each. So, think of these sacks as sets and these balls are a large number of elements.

Now, you could of course, count the marbles in each sack, but this is a bit tedious; because we know that as we are keeping track of these small objects, we often lose count or miss count or add one or plus one. So, at the end, we have to be doubly sure that we have counted correctly, so we will count it a number of times. So, counting the marbles in each sack and then checking if the two counts are equal is a tedious process and it is error prone, if we do it manually.

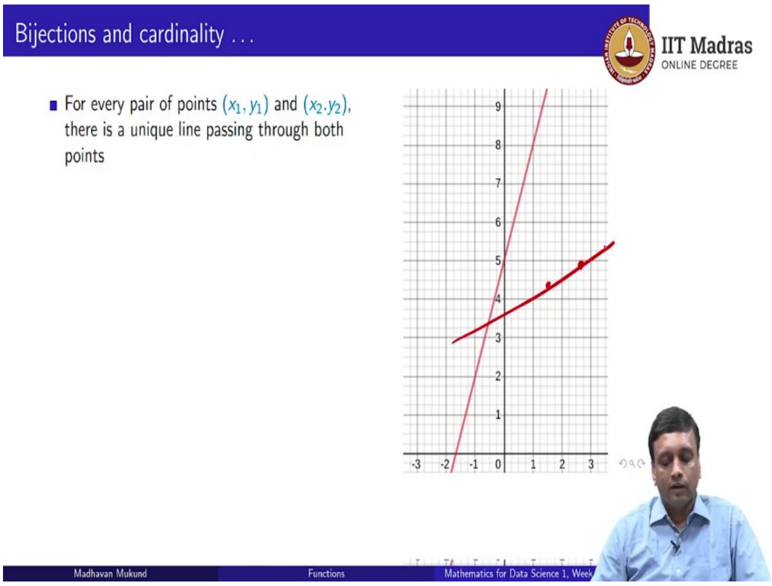
Now, here is a manual process which is less error prone. Supposing we put our hand into each sack and pull out a marble from each sack and put it away somewhere; then we put our hands again in and take out one marble each again and put it away somewhere. So, with each move, we are taking out one marble from each sack. So, what can we say; well if the two marbles sacks get empty together, then we pulled out one from each. So, we have actually established that there is a one to one correspondence between the marbles in the first sack and the marble in the second sack.

If on the other hand when we find one sack is empty and the other sack is not empty; this means that up to this point, we pulled out an equal number of marbles from both sacks and now one sack has extra marble, so they were not equal. So, in this way establishing a bijection is equivalent to saying that two sets have the same cardinality. So, for finite sets this is a convenience; but for infinite sets this is the only way in order to establish that the cardinality is the same.

So, for instance supposing we want to know whether the number of lines that we can draw is the same as the number of points on this plane R X R . So, R X R is a set of all points that you can draw on this plane and the number of lines we can draw is a number of such straight lines that we can draw; are these the same? Now it may not seem obvious how to argue this one way or another; but remember that we said that every line can be represented by a function of the form mx + c. And we also said that if you change m, you get a new line and if you change c, you get a new line. So, m and c together uniquely define a line.

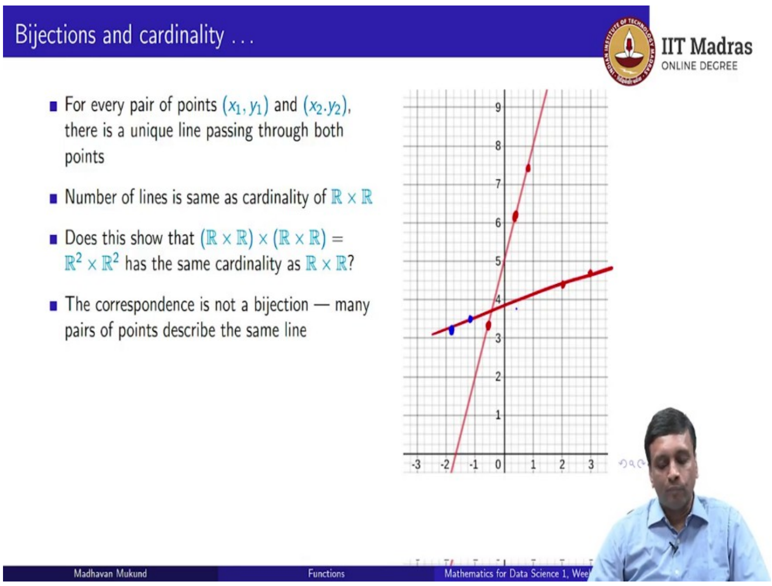
So, since m and c together uniquely define line; every pair (m, c) defines a line and every line defines a pair (m, c), so there is a one to one bijection between the lines and the pairs of points on this plane. So, actually the number of lines is the same as R X R. So, think about it, because this may not be obvious at first sight; but by establishing a bijection in this way, we can say that the number of lines that we can draw on a plane are equal to the number of points on a plane.

(Refer Slide Time: 17:09)



Now, suppose we extend this argument; if we take any two points right, if we take two points say $x_1$andx_2, we can draw a unique line passing through these points. So, this is a well known fact from geometry.

(Refer Slide Time: 17:23)

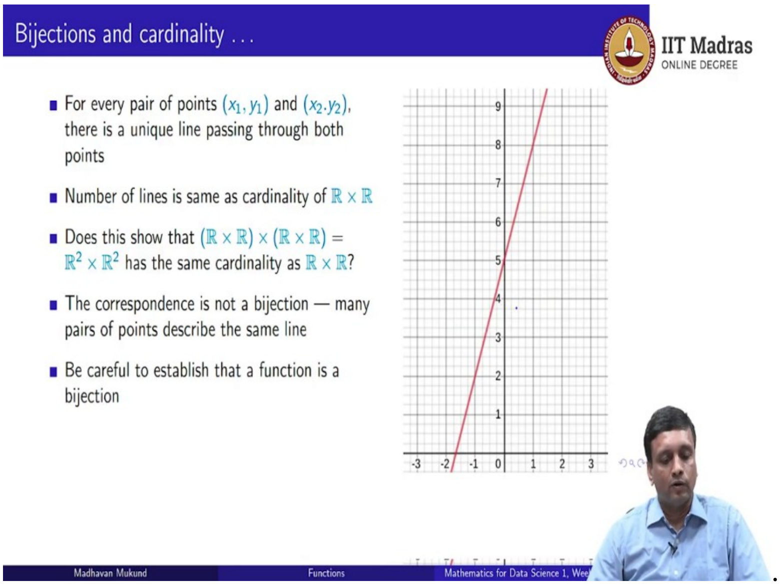


So, we know that the number of lines has the same cardinality as R X R that is what we claimed in the previous argument. Now we say that every pair of points defines a line. So, can we say that every pair of points therefore, has the same cardinality? So, remember this is a pair of points.

So, we have one point here and one point here. So, do we say that every pair of points has the same cardinality as the set of all points? So, it is R^2X R^2 the same as R XR, is this an argument for that? So, important thing is to ensure that we have a bijection; the problem is that this is not a bijection, because along any line we have many points, right.

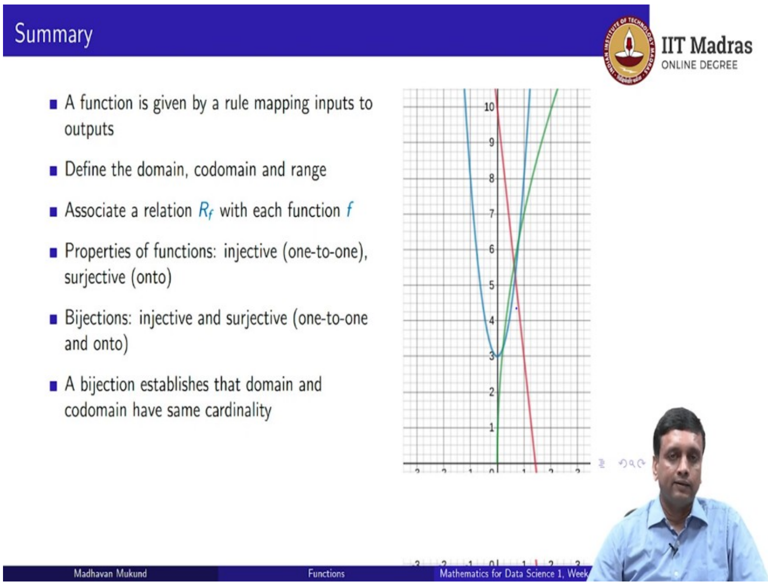
So, if I take these two points, indeed it forms a unique line; but I get the same line if I take these two points for instance. So, it is not the case that every pair of points that I pick generates a different line. So, unless I can show you that pairs of points, different pairs of points generate different lines; I do not get a one is to one correspondence between pairs of points and lines, and therefore this bijection breaks down.

(Refer Slide Time: 18:24)



So, whenever we are trying to use a bijection to describe some kind of a correspondence and count points especially in an infinite set, count elements of an infinite set, compare infinite sets against each other; you must make sure that the function you are defining is really a bijection.

(Refer Slide Time: 18:39)



So, to summarize a function gives us a rule to map inputs to outputs. And with each function we have to specify three sets; we have to specify the domain, so the function must be defined on every set in the element of the domain set, the codomain what are the output elements supposed to look like and the range which was actually the output assumed by the function once we applied.

So, not all elements in a codomain may actually be attainable by the function; the range is those elements which you can reach through the function. With each function we can associate a binary relation consisting of all pairs (x , y), such that y = f (x). Then we saw some interesting properties that we would like to prove for functions in order to make use of them; one is injectivity that is every pair of distinct inputs produces distinct outputs, so this is one is to one. And surjectivity which says actually that the codomain and the range match; everything that I could possibly generate, can in fact be generated by applying the function.

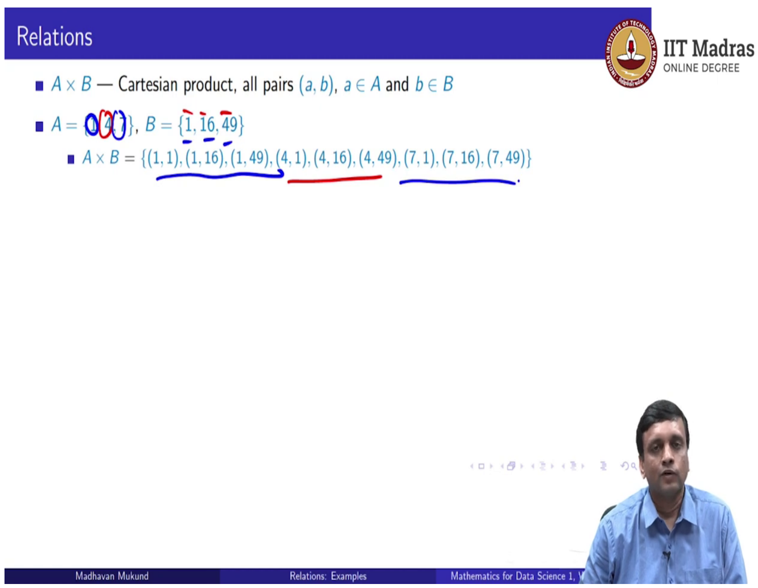
Then we saw that a bijection combines these two. So, a bijection gives us something which is an injection and a surjection; something that is one to one and onto. And once we have a bijection between two sets, we can actually argue that the two sets have the same cardinality and this is often the only way to prove that two infinite sets have the same cardinality.

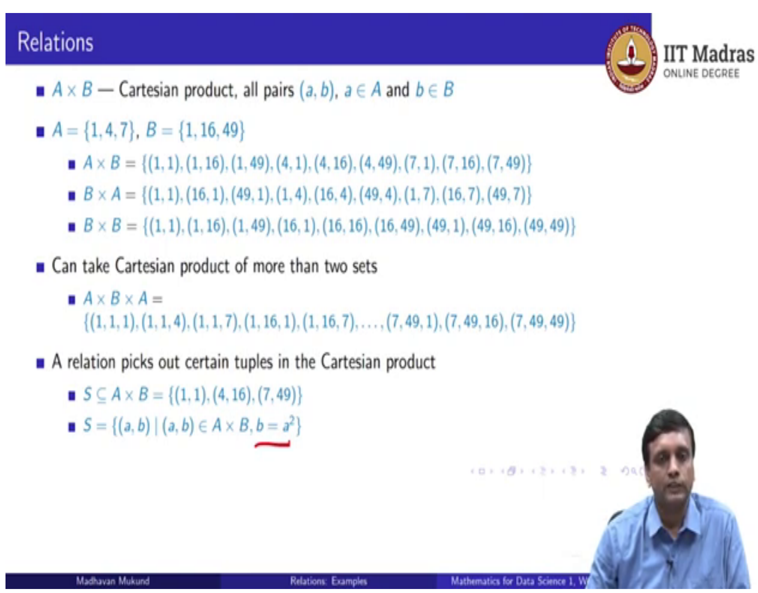
Thank you.

Mathematics for Data Science 1.
Professor Madhavan Mukund.
Department of Computer Science
Mathematical Institute, Chennai.
Lecture-7A.
Relations: Examples.

So, earlier we defined relations as subsets of elements of a Cartesian product which have special properties. So, let us take a look at relations again and understand why we are so interested in relations.

(Refer Slide Time: 00:25)





So, remember that a Cartesian product takes all pairs of elements from a collection of sets. In particular, if you say A cross B, you are taking 2 sets A and B, and you are taking every pair of elements of the form small a small b such that the first small a comes from capital A and small b comes from capital B. The order is important, the first element in the pair comes from the first set, the second comes from the second set. So concretely, let us look at these 2 sets.

So, suppose $A = \{1, 4, 7\}$, so it has 3 elements, and $B = \{1, 16, 49\}$. So, if you now look at $A \times B$, it looks at every pair. So, if you can take this one, and combine it with 1, 16 and 49 to get this. Then you can take this 4 and combine it again with 1, 16 and 49 to get these for 3 pairs, and finally you take 7 and then you combine it again with 1, 16 and 49 to get these pairs.

So, it is easy to see that if you have m elements in the first and n elements in the second, every one of those m elements is paired with every one of the n elements, so you get $m \times n$ pairs. Now, the first thing to remember is that the Cartesian product is ordered. So, there is a first and there is a second. So, if you reverse this and say $B \times A$, you do not get the same set of pairs, every pair is reverse. So, (16,1) replaces (1,16), (49,1) replaces (1,49). So, this is the first thing to remember about Cartesian products.

The other thing to remember is that there is no relation, there is no constraint on what you can take the Cartesian product of. You can very easily take the Cartesian product of a set with itself. So, the set to itself is not just pairs of identical elements, but also pairs of non identical elements. So, if you take $B \times B$, you get Of course, (1,1), (16,16), (49,49). But you also get the dissimilar pairs like (1,16), (16,49), (49,16), and so on.

So, this is an example with 2 sets, but there is nothing to restrict us to 2 sets. So, in general, a Cartesian product can take a large number of sets and gives us tuples. So, for instance, if we take 3 sets, we get these triples, each element has 3, each element in the Cartesian product has 3 elements in order.

So, here for instance, if I do $A \times B \times A$, I take every element in A, combine it with every element in B and then with A again. So, I have 1 from A, 1 from B and 1 from A. Then I have 1 from A, 1 from B and 4 from A, the second copy of A and the first copy of A are different.

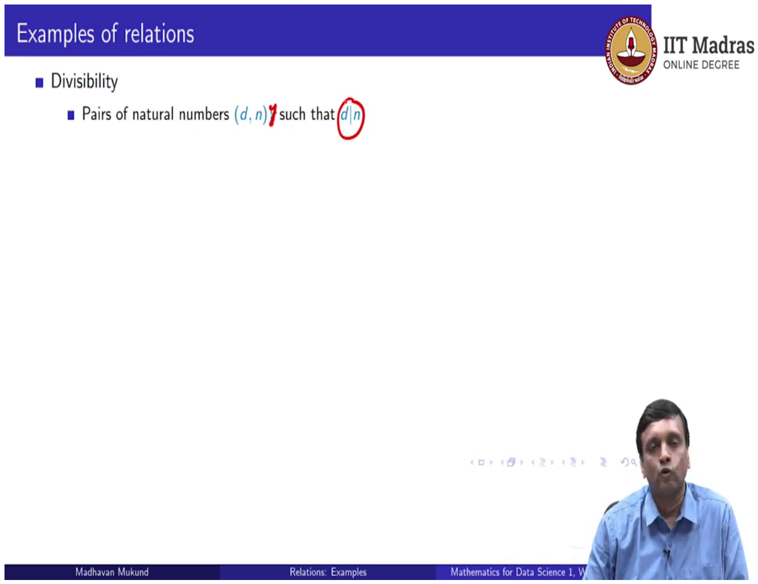
So, I have (1,1,1), (1,1,4), (1,1,7), then I move to the second element of B, I have (1,16,1), (1,16,7). Now, ultimately the Cartesian product is a set, so it does not matter in what order I write these triples. But to order to write them down systematically, it is convenient to write them down in this particular way, where we go through each set one by one, otherwise, we may miss out on something. So, the reason we need Cartesian products is because they are the building blocks of relations.

So finally, what we want is not all these pairs or triples, but some of them which are of interest to us. So, for example, from the first Cartesian product $A \times B$, we may be interested in the pairs where each element from A is paired with a corresponding position B. So, the first element in A is paired only with the first position in B, second with the second and so on. So, we might want to say that we want S, a set which is a subset of $A \times B$, which from those 9 different pairs picks out only 3 of them of interest, (1,1), (4,16) and (7,49).

Now, if as in this case, there is some way of describing this, which is more abstract, you can also use a set comprehension. So, we can talk in terms of positions or observe that in this particular case, the second element is always a square of the first element. So, we could also write this as the set of pairs (a,b), where (a,b) comes from $A \times B$, so we are generating every possible pair in the Cartesian product.

But then we are filtering, remember that we had these filters, so we are filtering it so that we only retain those pairs for which the second component B is the square of the first component. So, this is how relations are defined. They are typically defined as subsets of the Cartesian product. And we can either write out the subset explicitly or try to express it implicitly using the set comprehension notation.

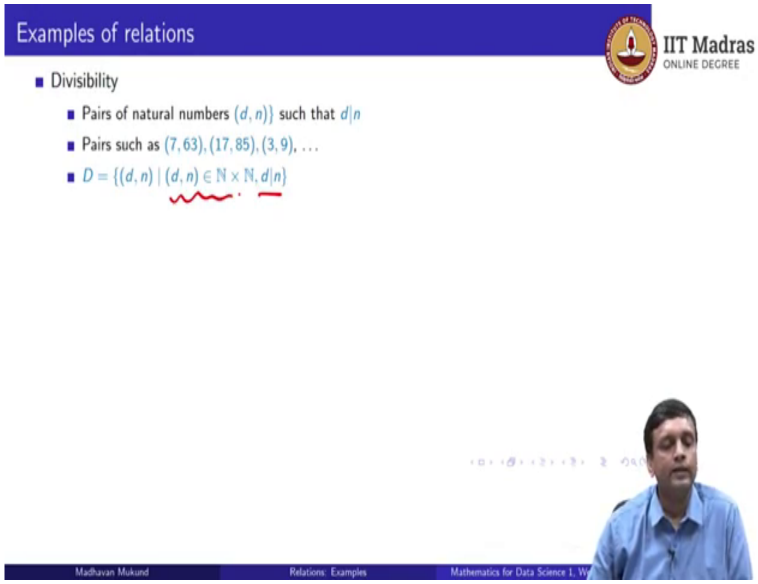
(Refer Slide Time: 4:57)



So, we saw some examples. So, let us look at these examples again more carefully, some examples from numbers. So, divisibility is an important relation when we are talking about natural numbers or integers. So, divisibility talks about pairs of natural numbers, such that the first one divides the second one. So, we want (d,n) such that d divides n, remember this notation, this perpendicular bar for numbers denotes, this is not the same as the one that we use in set comprehension.

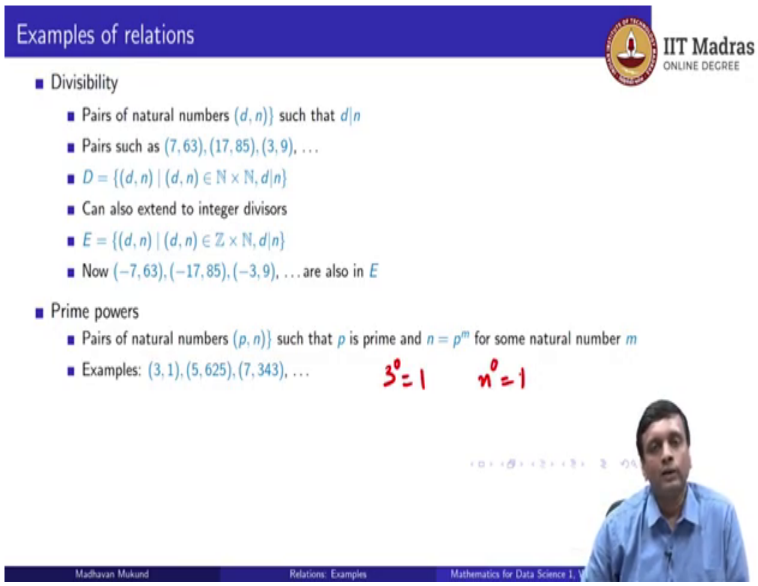
So, here it is an operation, arithmetic operation which says d divides n, so if I divide n by d, there is no remainder, it is a 0, d perfectly divides n. So, this would have this divisibility relation would have pairs like (7,63) because $7 \times 9 = 63$, or (17,85), because $17 \times 5 = 85$, and so on. So, we have a large number of pairs of divisors and numbers which the divisors divide equally, evenly. So, this we can write in our set comprehension notation because this is an infinite set, so we have no other way of listing everything.

(Refer Slide Time: 06:09)



So, we take all pairs N × N, (d,n), such that d | n. So, this is our filter. So, we want to generate everything of this form, but filter out under the condition that d must be a divisor of n and keep all such pairs. And this we can call d, the divisibility relation.

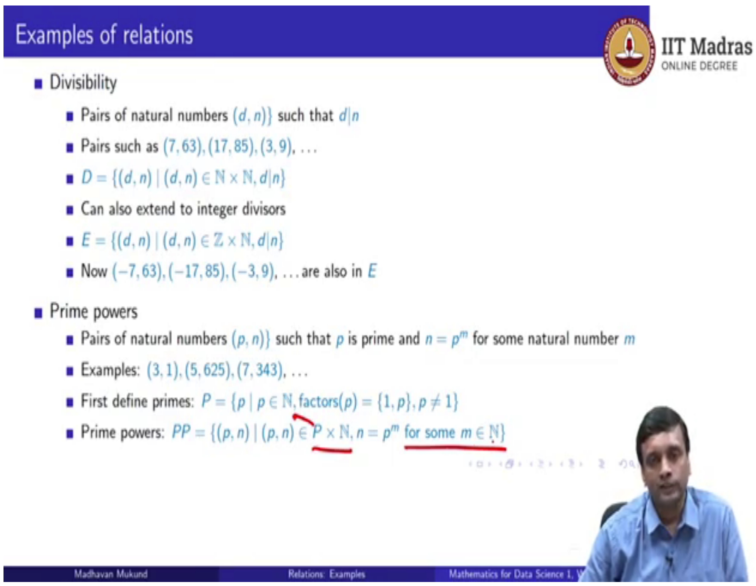
(Refer Slide Time: 6:2 1)



Now, this is the relation on pairs of natural numbers, so we only get positive divisors. If we extend it to integers, then we will get even negative divisors. We know that (-7) × (-9) is also 63, because the 2 negative signs will cancel out. So, if you extend the generating set from N to Z, from the natural numbers to the integers, then we get a larger set of divisor pairs. So, we get minus and plus elements for the same pairs that we had in the original relation.

Here is another example. Let us look at what we call prime powers. So, a prime power is something that is a prime multiplied by itself for a certain number of times. So, for instance, we can say that $5^5 = 3125$. So, $5^2 = 25$, 5^4 rather, 5^2 , $5^3 = 125$, and $5^4 = 625$. So, 625 is a prime power, similarly $343 = 7^4$, so it is a prime power and so on. Why is (3,1) in this relation because anything to the power 0 is 1 by definition. So, $3^0 = 1$, in fact, anything to the power, so any number to the power 0 is 1. This is by definition. So, for every number comma 1 will be a prime power.

(Refer Slide Time: 7:41)

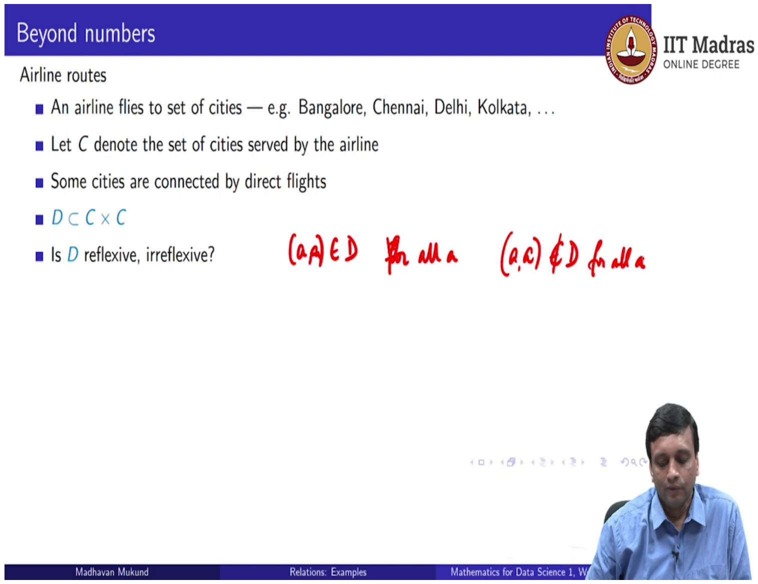


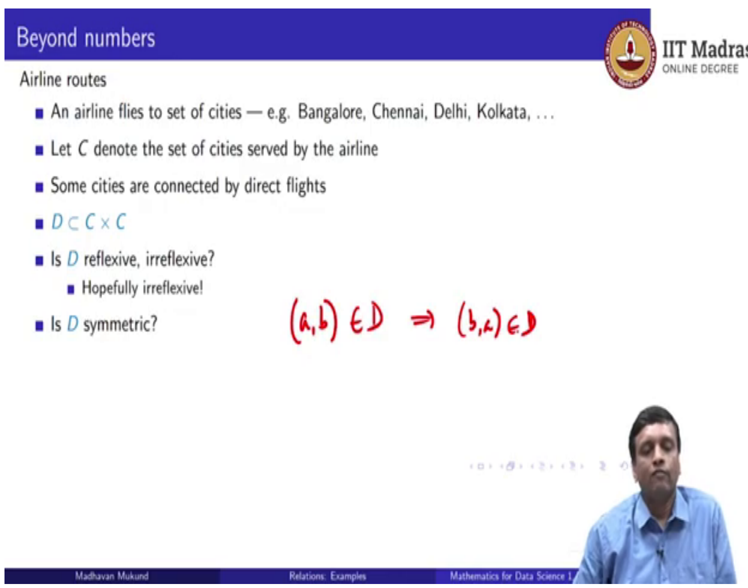
So, if you want to define prime powers, it is useful to first define primes. So, one way we can define primes is to say, give me a natural number, such as the factors of the natural number consists of exactly 2 elements, 1 and the number itself. And because in sets, we do not distinguish duplicates, in this definition, if I just say factors(p) = {1, p}, it includes a case where p is 1, because factors(1) = {1,1}, which is just 1. But I do not want to count 1 as a prime number. So, we also specify that P is not 1. So, this is the set of primes.

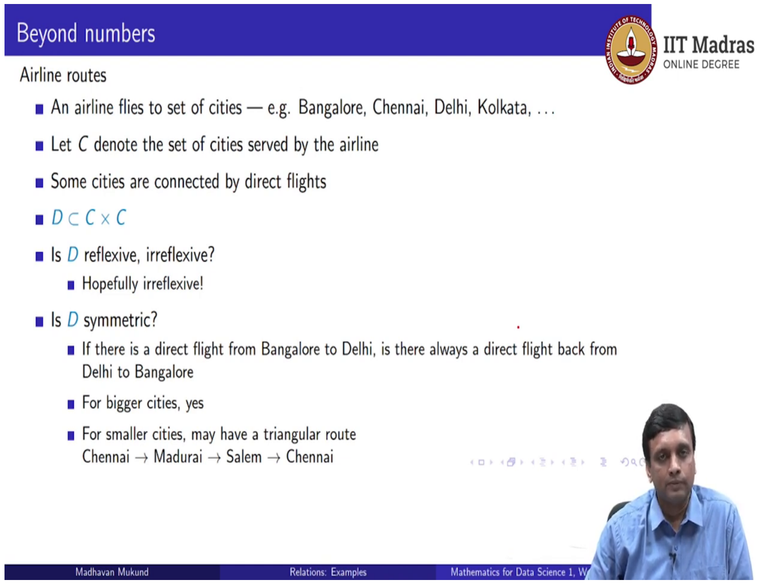
And now, we can say the set of prime powers is the set of all pairs in P × N, where P is defined above, P × N, such that n is the power of p. So, n = p^m for some m, which is a natural number, which could be 0. That is why we get (3,1). So, this is an example that we also talked about. It is saying that when you are writing the set comprehension, you can write these kinds of statements.

So, you do not have to be very precise about what you are writing mathematically in terms of notation, as long as the understanding is clear, there is no ambiguity about what you mean. So, you can write words like for some, you can also write it in a mathematical notation using symbols for there exists and for all and so on, but it is not necessary. As long as you are precise, you can use set comprehension notation in a flexible way.

(Refer Slide Time: 9:00)







So, these are relations in a formal sense. But why are we so interested in relations especially in the context of computing and data. So, let us look at relations which go beyond numbers. So, here is an example. Supposing we are talking about an airline, which serves a set of cities and we are interested in the routes that this airline serves. So, let us C be the set of cities where the airline operates. So clearly, the airline operates between some pairs of cities, but not all of them.

So, some of these cities are connected by direct flights and for other situations, you have to take a hopping flight which goes from city A to city B and then from city B to C. So, let us look at that subset D of direct flights between cities in C. So, this is an example of a relation. Not every pair of cities is connected by a direct flight. So, if you take all possible pairs of cities, some of them are connected by direct flights, and some are not. So, this way, information about an airline's route is really a relation in the sense that we mean.

Now, we have defined certain properties of relations, we said that the relation is reflexive. Now, this is useful to ask this question because we are talking about a relation between a set and itself. So, we can ask whether every element in the set is related to itself or is not related to itself. So, reflexive means that always we have (a,a) in D, for all, for every a. And irreflexive means, exactly the opposite of this is never in D and for all A.

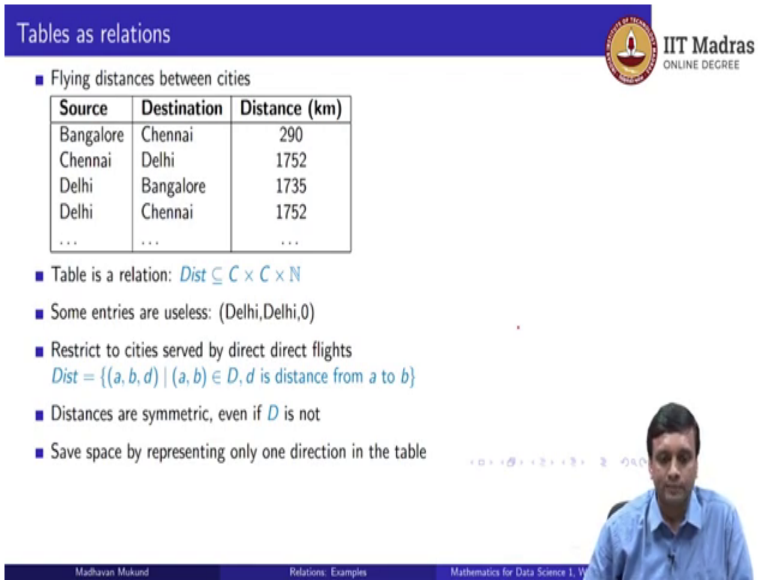
So, the question is, in terms of direct flights, is this going to be a reflexive relation and irreflexive relation or neither. Well, it is easy to see that this should not be reflexive. Because we do not expect an airline to actually operate a flight which takes off from an airport and then lands immediately in the airport. And in fact, we would precisely like it to be irreflexive, that is, this should never happens.

So, this should not be reflexive because we do not want every airport to serve itself and we want it to be irreflexive because we want no airport to serve itself. So, this is an example of an irreflexive relation. Now, is it a symmetric relation? So, symmetric relation says that whenever I have a pair of cities in the relation, then I will also have the reverse pair in the relation. So, if I can fly from one city to another directly, then I can also fly back.

So, concretely for instance, if I take any 2 cities a direct supposing there is a direct route from Bangalore to Delhi, then is there always a direct flight back from Delhi to Bangalore. Now, if you think about airlines, this is usually the case. But actually, if you look at domestic flights in particular, this is typically true only for the bigger cities, it will certainly be true for all the metro cities and the largest state capitals and so on. But if you look at smaller cities, this is not necessarily in the case.

For instance, it is quite common for airlines to serve 3 cities in a triangular route. So, you might have a flight that takes you from Chennai to Madurai, but if you want to come back from Madurai to Chennai, you cannot fly back directly, but you may have to fly to Salem and then come. So, between these 3 cities you can get from one to another, either directly or indirectly depending in which direction you are going. So, this relation is going to be irreflexive but not necessarily symmetric, it depends on the context.

(Refer Slide Time: 12:29)



Now, one thing you can do is to extend this to a table. So, here is a useful table that we might want to keep, which might be used to derive other things such as how long it takes to fly or how expensive a ticket is like to be. So, here we are just recording a fact which is what is the flying distance between a pair of cities. So, this table says that if the source is Bangalore and the destination is Chennai, it is 290 kilometers, whereas if the source is Chennai and the destination is Delhi, it is 1752 kilometers.

So, for every direct flight which our airline operates, you can record this distance and put it in a table. So, what is important to recognize and this is why relations are so useful in computing and data is a table is just a relation. So, every column represents a potential set of values. Here, the first column represents a possible city, so it is taken from the set C, the second column is also taken from the set C, the third column is a natural number.

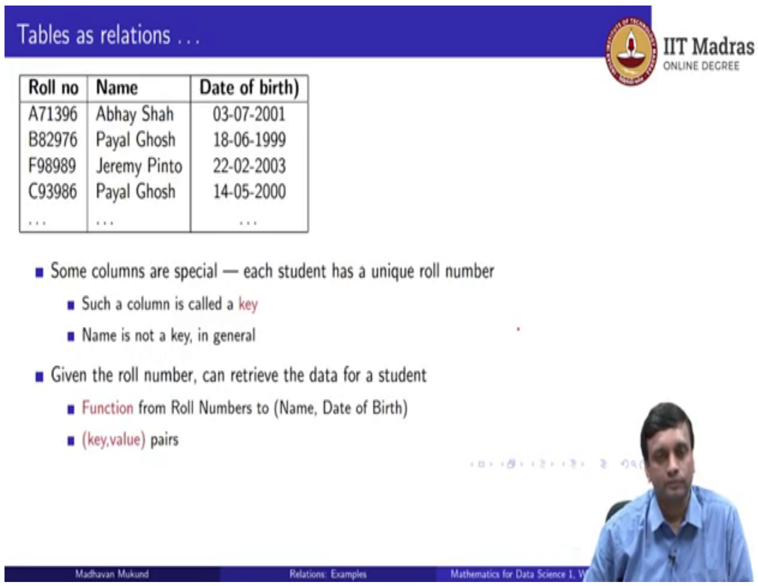
If you take pairs of cities which are the same, you could put 0, so it could be from Delhi to Delhi it is 0. So, in general, you have all possible pairs of cities and all possible numbers, but only some of them are interesting. Namely, when I have 2 cities which are actually connected by a flight and the distance the number is actually the real distance. So, it is a relation on $C \times C \times N$.

As we said, some relations are useless so we would not record them even though we know them. We know that for every city, the flying distance from the city to itself is 0, so there is no reason to record it in the table. The other thing is that unlike our direct flight's relation, this is actually a symmetric relation. So, first of all, we will only keep direct flights because we do not want indirect flights. But distances are definitely symmetric.

So, it doesn't really matter whether there is a direct flight from Chennai to Delhi and back or whether there is a direct flight from Chennai to Madurai and not back. It is enough to record the distance from Chennai to Delhi and Chennai to Madurai once each. I do not have to keep the distance from Delhi to Chennai separately as you can see above, in this example, Chennai to Delhi and Delhi to Chennai are both exactly the same distance 1752 because that is how distances work, distances are symmetric.

So, if we have symmetric entries, in a practical sense, when we represent a relation as a table, we can save on space by not recording the symmetric entries and making a note separately that this relation is symmetric. So, that is why it is important to know the property of the relation. It is not just an abstract question, is this reflexive, is this irreflexive, it is actually a practical consideration, a symmetric relation can be represented by only half the entries in the relation, the other half followed by symmetry.

(Refer Slide Time: 15:12)



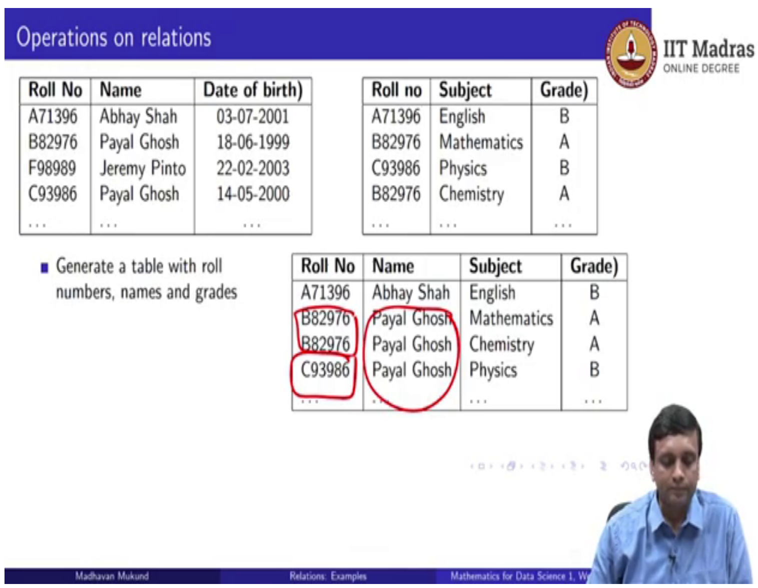
So, let us go further with this. So, another place where we often encounter tables are, for instance, when looking at data about people. Let us look at students. So, typically a college would record or a school would record information about students in this form. So, they would assign a roll number, then they would record maybe the name, the date of birth, and there would typically be other personal information like maybe their home address, phone number, and so on.

So here, what is important is that some columns are not natural in the sense. So, we know that everybody has a name and they are born on a particular date, but this roll number is actually assigned to them by the school or college. And this is something which is designed to be unique, so no 2 students get the same roll number. So, this kind of column is called a key. And this is because we want to identify, define each student directly and individually without getting confused about which student we are talking about.

And unfortunately, the other columns are not keys, 2 students could have the same name. And it is even possible for 2 students to have the same name and the same date of birth. So, we cannot rely on the fact that the other columns will uniquely distinguish. So now, if we have a unique roll number for every student, then each row is identified by the roll number. So, we can actually think about the row as being something where if I give you the roll number, you can tell me which row it is and give me the other values in that thing.

So, this is more like a function. A function says given an input give me a unique output. So, given a roll number, tell me all the values associated with the roll number, the name, the date of birth, and so on. So, this kind of a stored table is also called sometimes a set of key value pairs, given the key there is a unique value. I can change the value for a given key by updating it. But if I add a new entry, I have to add a new key so there is no confusion.

(Refer Slide Time: 17:03)

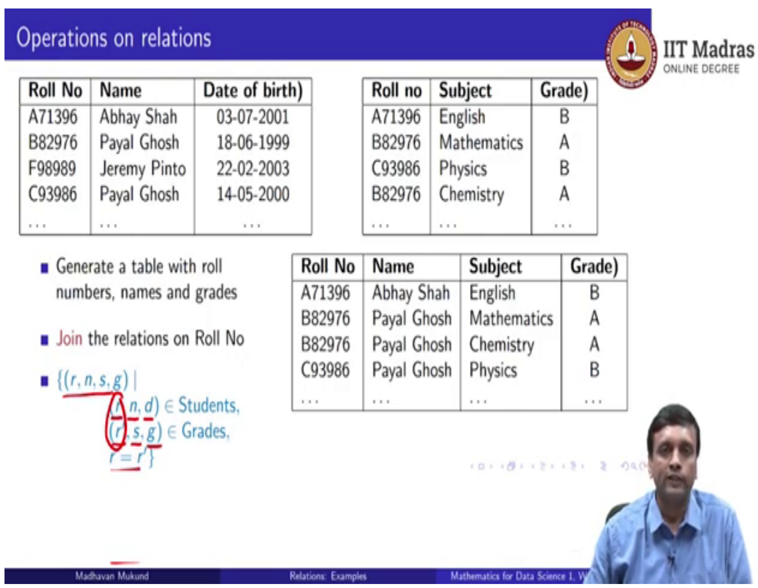


So, usually a school or college will maintain more than one table of this kind. For instance, there might be a separate table, where we maintain the marks of the student or the grades of a student in the courses that they do. And here for conciseness, we might keep only the roll numbers and the subject names and not the names of the students. So, for instance, in the second table, we have the roll number, subject and the grade. Here is a typical requirement when we have to generate a report card.

The grade card has, the grade table has the roll number and the subject and the grade but it does not tell us who the student is. And that is, for example, it may be difficult for an outsider who except for the student themselves to know whose roll number belongs to whom, because nobody would recognize these strange character sequences. So, we want a table that looks like this which has the roll number and extra column with the name which is not there in the grade table which is taken from the first table and then we want the subject and the grade.

And here, we see why it is important to have keys because we have this name Payal Ghosh, which is ambiguous, there are 2 Payal Ghosh'es. And in fact, they have 2 different entries in this table because they have 2 different roll numbers. So, the Payal Ghosh who got an A in mathematics is not the same as the Payal Ghosh who got a B in physics. So, this is an operation which combines these 2 tables. And remember that a table is a relation.

(Refer Slide Time: 18:23)

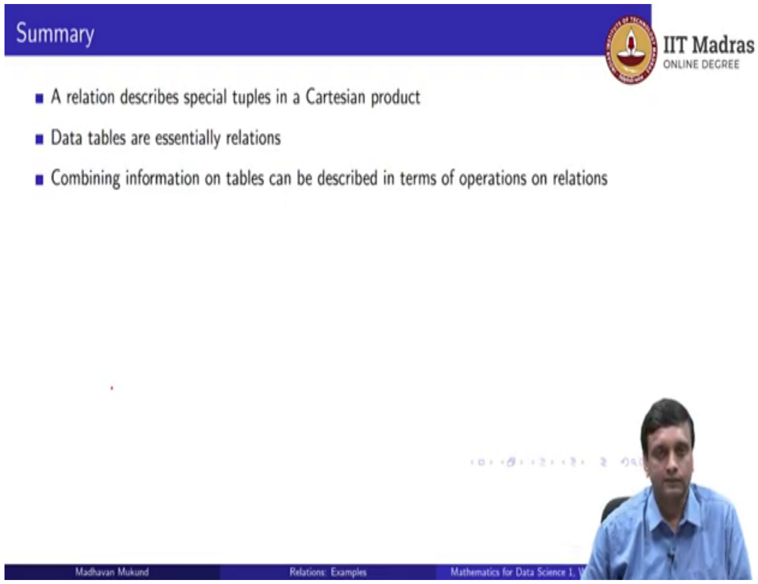


So, this operation, which combines 2 tables is also an operation which combines 2 relations, and it is an important operation in computing and in data science called a Join. So formally, a Join takes tuples from 2 relations and combines them on common values. So here, for instance, you take any arbitrary roll number, name and date of birth from students, you take any arbitrary roll numbers subject and grade from grades, but you want that the roll number in the roll number of the 2 sides belongs the same.

So, the r comes from students and the r′ comes from grades and you want r = r′. And if this is the case, then you put out a new tuple, which combines the n from the left hand side throws away the date of birth, we are not interested in preserving the date of birth, keeps the n and keeps the subject and the grade s and g and of course keeps the roll number which is the same on both sides.

So, this will ensure that we do not get rows merged, where they correspond to 2 different students. So, the marks for Abhay, or the grade for Abhay will not be merged with the name and date of birth for Jeremy Pinto, because they have 2 different roll numbers. So, this is called the Join and this is a very important operation on relations, and therefore on tables. And this is something that we use implicitly all the time.

(Refer Slide Time: 19:39)



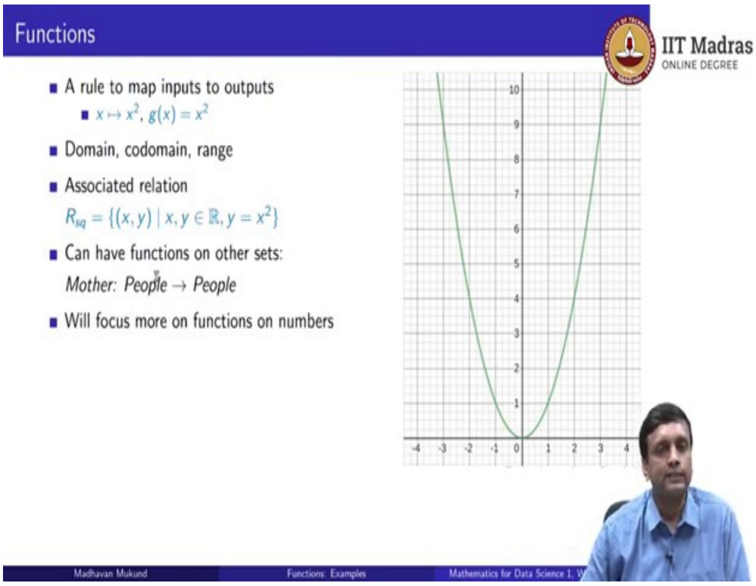
So, to summarize, a relation describes special tuples in a Cartesian product. And what is really important for us from a computing and data science point of view is that we work with tables all the time and tables are really relations. So, that is why relations play such a central role in many of the things that we are going to look at. So, it is important to get the terminology of relations right.

And when we combine information on tables, these are actually operations on relations such as the Join operation that we described, this is only one kind of Join we may have different types of operations, which we will see in other courses later on. But please, keep in mind that tables are relations. Thank you.

Mathematics for Data Science 1
Professor Madhavan Mukund
Department of Computer Science
Mathematical Institute, Chennai
Lecture-1.7B
Function: Examples

So, let us take a closer look at functions now.

(Refer Slide Time: 0:19)



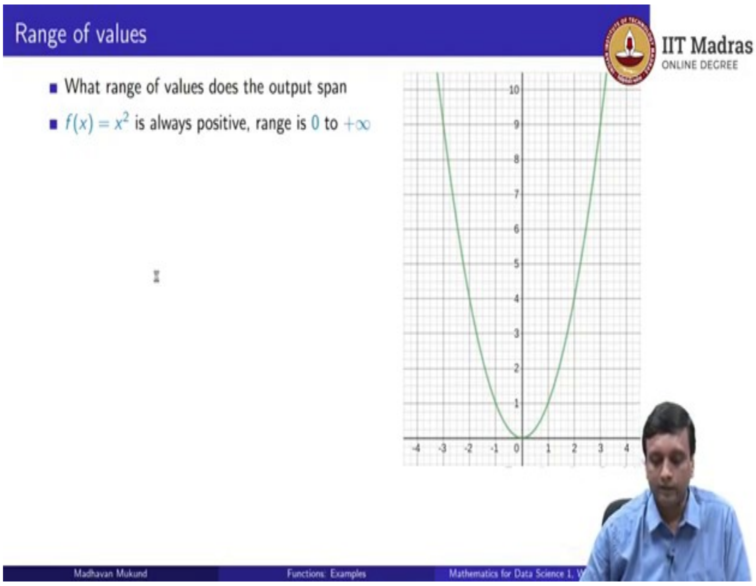
So, remember that a function is a rule that map's inputs to outputs. So for instance, if we are looking at numbers, a function could take an input x and map it to x^2, which can also write given a name saying $g(x)$ is equal to x^2, which says g is the name of a function, which when it takes an input of the form x produces an output of the form x^2.

And with such functions, we have a notion of a domain that is what are the inputs that are allowed, the set from which we take inputs. Codomain, what is the set to which the outputs belong and range which is the actual outputs that this input set generates for this given rule. So, for instance, we have for this function this relation associated with it, all pairs x comma y such that x and y are reals. So, the domain and the codomain are both reals, the rule is y equals x^2, so that is the filter that we put, we only want such pairs.

And if we plot all the points which belong to the relation, we get this graph on the right. And this actually tells us that the range of the function even though the codomain is all reals, the range of the function actually keeps this function above 0, so we only get non negative reals as outputs. Now, we are not restricted to looking at functions on numbers, we can also look at functions on other sets.

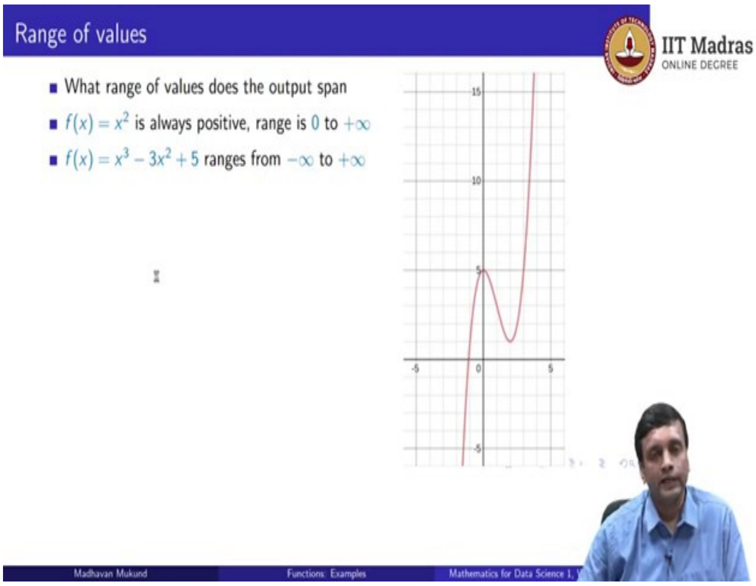
So, for instance, if we look at the set of all people in the universe, in the world, in the country, in any range of geographical regions, we can look for the function mother which says, given a person, this will map the person uniquely to the mother of that person. So, this is a function because every person has 1 mother. So, in this lecture, and in general, when we are talking about functions in this course, we will look more at functions on numbers. So, let us look at these a little more closely. What are the questions that we really want to ask about functions on numbers?

(Refer Slide Time: 2:09)



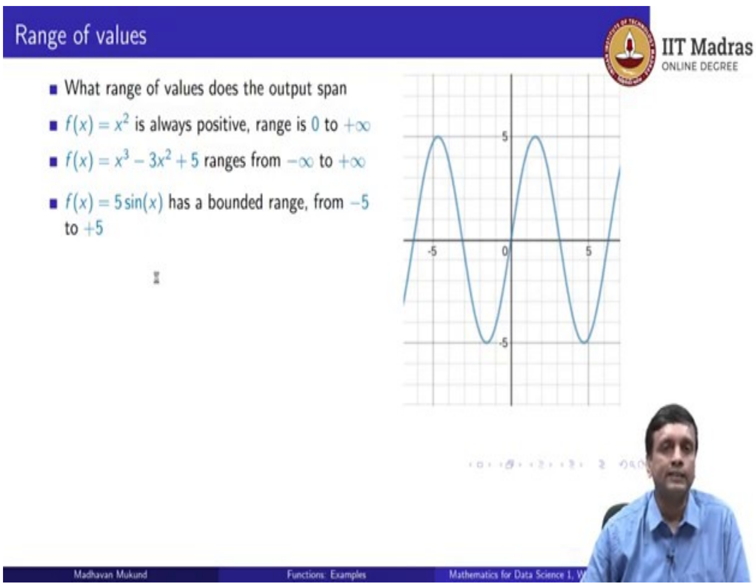
So, one of the basic questions is, what are the ranges of the values that we can get. So, in other words, we have a core domain. But what is the range of values that we can actually achieve through the function. So as we saw, this square function, $f(x)=x^2$ is always positive, so we always get something between 0 and $-\infty$, there is no upper bound, but we never get something which is negative.

(Refer Slide Time: 2:35)



On the other hand, if we take a cubic function of this form $f(x)=x^3-3x^2+5$, then when x becomes very small, the x^3 becomes very small because the cube of a negative number is a negative number. So, cube have a large negative number, I mean magnitude, the $(-1000)*(-1000)*(-1000)=-10^{-9}$. So, as we go into negative, large negative values, we can at large negative outputs, same for large positive values. So, this has a range from minus $-\infty$ to $+\infty$.

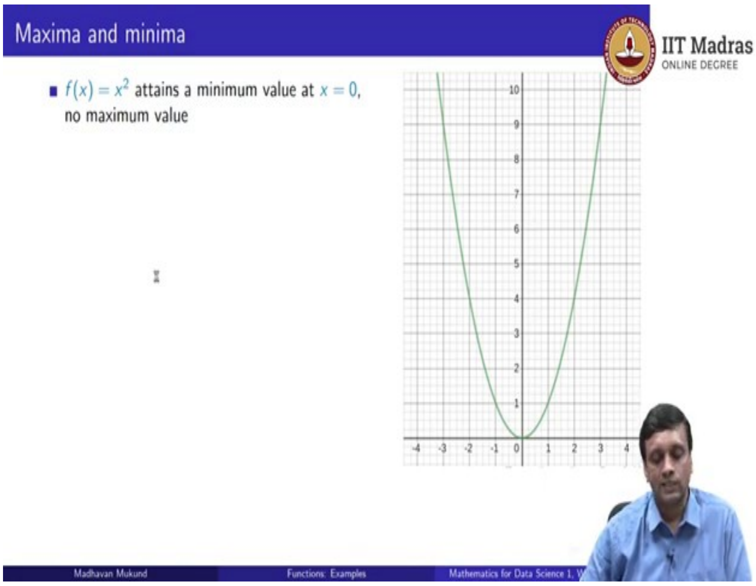
(Refer Slide Time: 3:05)



And then there are some functions like the trigonometric function $\sin x$, which oscillate between an upper bound and lower bound. So, if you take $\sin x$, usually it is between +1 and -

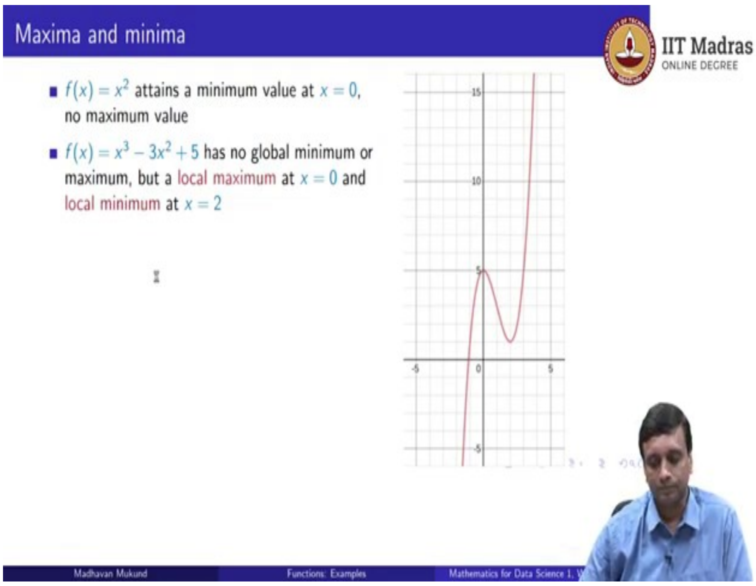
1. If we take $5\sin x$, then it will be between -5 and $+5$. So, this has a bounded range. Even though we consider all possible inputs, we never go outside this range from -5 and $+5$.

(Refer Slide Time: 3:28)



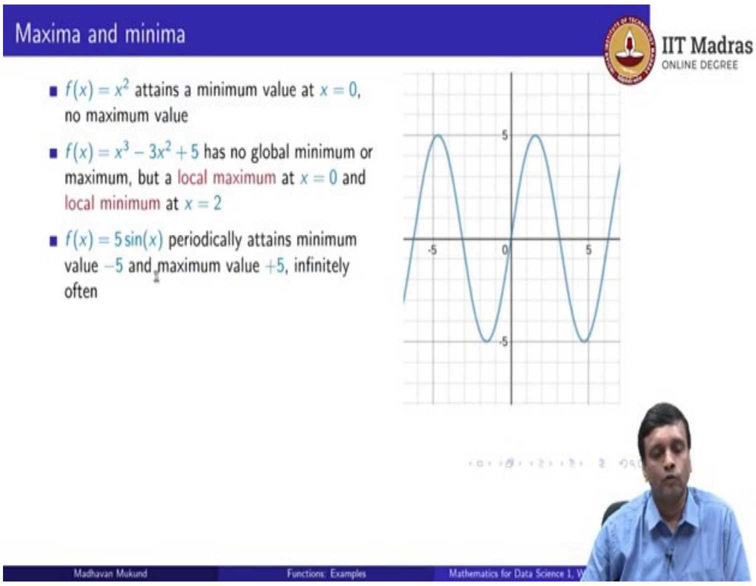
Now, within the range of values that it can take, we are often interested in specific points, in particular, where the value are a minimum and where they are maximum. So, for instance, this function that we have seen before $f(x)=x^2$, it is clear from the graph on the right that at 0 the output is 0 and at all other points is bigger than 0, so it attains its minimum value at 0. And because it keeps growing indefinitely in both sides, there is no maximum value.

(Refer Slide Time: 3:57)



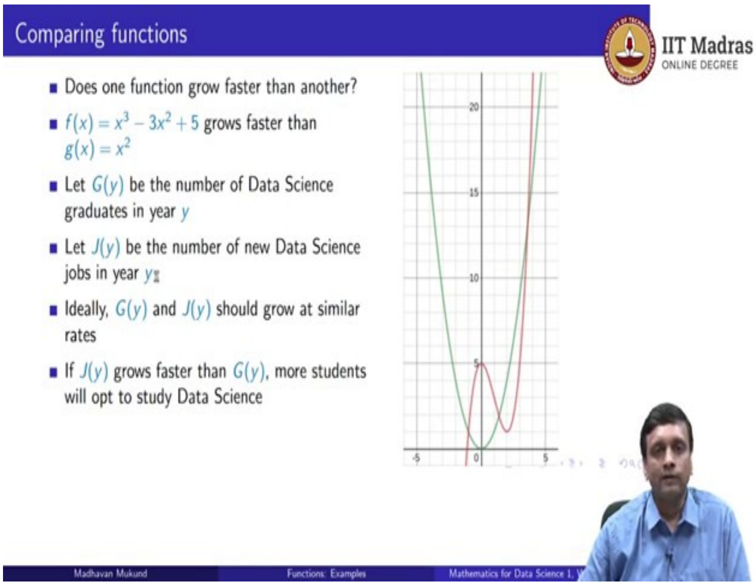
Now, the cubic function we said grows arbitrarily small as we go to the negative inputs and arbitrary large. So, there is actually no maximum and minimum, but it has an interesting behavior in between because it zigzags it goes up, comes down and goes up again. So, there is something called a local maximum and a local minimum. So, at $x=0$, it turns around, so it achieves a maximum value and starts falling briefly and then at $x=2$ it turns around again. So, it achieves a local minimum and goes up again. So, we are interested in finding out where these local maxima and minima are for various reasons.

(Refer Slide Time: 4:32)



And similarly, if we look at something like $\sin x$, then it has, of course, local minima and maxima, −¿5 is a local minimum and +¿5 is a local maximum, it is also a global minimum and maximum because these are the maximum and minimum values that the function can ever attain. And now, these values are actually attained infinitely often periodically as we go from left to right.

(Refer Slide Time: 4:53)



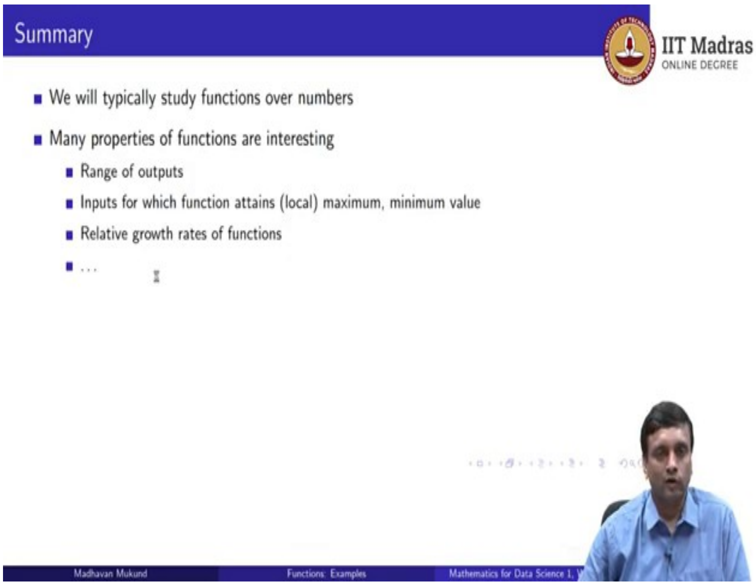
Another thing which we are interested in about functions is how fast they grow. Thus one function grow faster than another. So, if you look at our 2 functions, $f(x) = x^2$, and $f(x) = x^3 - 3x^2 + 5$, and we look at their 2 graphs, then it is very clear that the red line, although initially on the right, it is below the green line, it overtakes it, and after that, it is never going to be below the green line. So, in this way, the cubic function grows faster than the square function.

Now, why is this interesting? Well, we often see this informally stated in various contexts. So, let us look at a context which is relevant for you. So, let $G(y)$ be the number of data science students graduating in a year y. So, as the year increases, so we go from 2020 to 2021, and so on, the value G takes a certain number and hopefully because courses are growing, this number is increasing.

At the same time, there are jobs being created in data science. So, let $J(y)$ be the number of new data science jobs in a year . Now, ideally, you would like that these 2 are comparable, that the jobs are growing because the number graduates is growing and vice versa. If the number of jobs increases more than the number of graduates then there is a demand for graduates and of course, more graduates will opt to study data science. So, you would expect a demand for this kind of course.

Of course, the unfortunate case might happen the other way around, if suddenly there is a slump in demand, then people who graduate with a degree in data science will not be employable and then there will be a reverse trend. So, these are some of the reasons why when we look at data, we are interested in comparing the growth rate of functions and we will look at this in the context of the functions that we study mathematically.

(Refer Slide Time: 6:33)



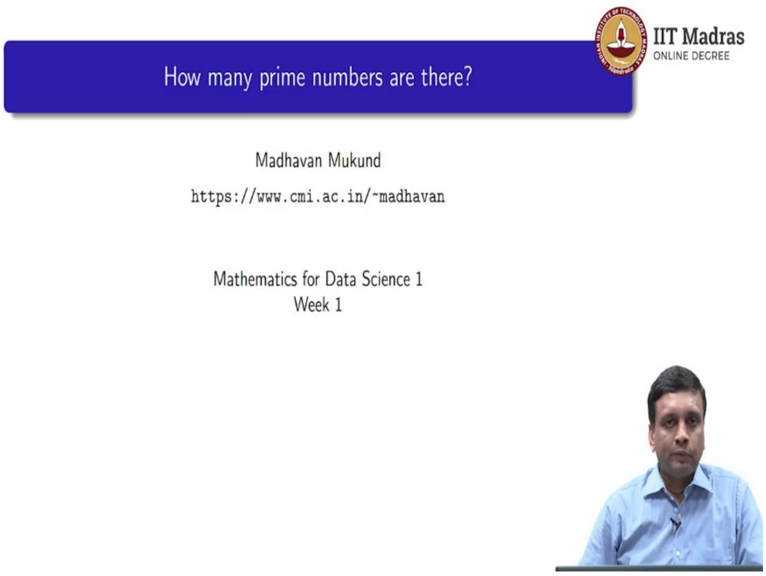
So, to summarize, we will typically study functions over numbers. And we are looking at many properties of these functions which are interesting to us, for instance, the range of outputs, where these functions attain local minima and local maxima and what are their relative growth rates and many other things which we will come across as we go along. Thank you.

Mathematics for Data Science 1
Prof. Madhavan Mukund
Department of Computer Science
Chennai Mathematical Institute

Week - 01
Lecture – 08
Prime Numbers

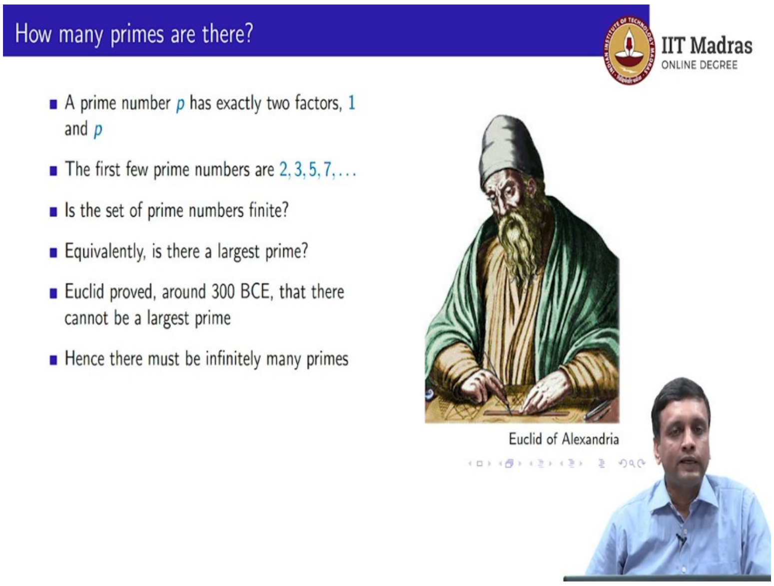
(Refer Slide Time: 00:06)

(Refer Slide Time: 00:14)



So, when we looked at the natural numbers, we talked about divisibility and we talked about the prime numbers. So, we know that the prime numbers start with 2, 3, 5, 7 and so on. So, how many prime numbers are there?

(Refer Slide Time: 00:25)

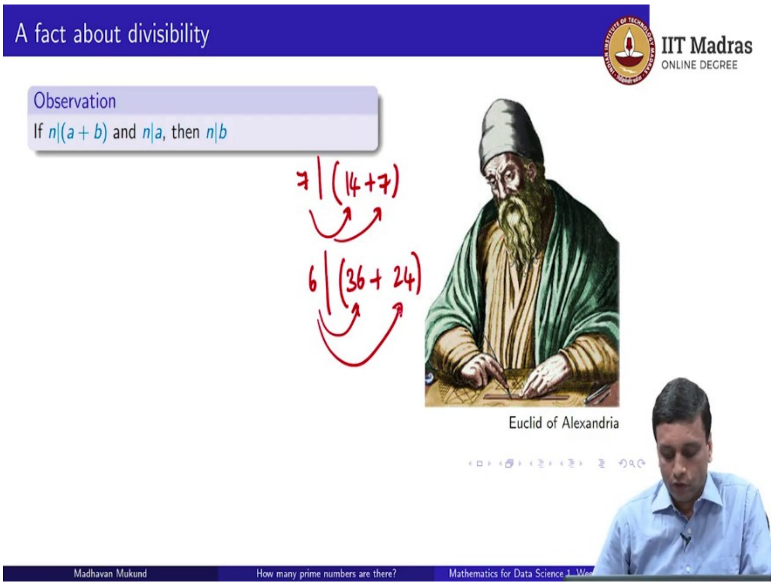


So, remember that a prime number is something that has only two factors 1 and itself. Now, it must have exactly two factors. So, 1 is not a prime. So, the first few prime numbers are 2, 3, then 4 is not a prime – because 4 is divisible by 2, then 5, again 6 is not a prime and so on. So, the question is, is this set of numbers these prime numbers is it a finite set or are there infinitely many prime numbers?

Now, if there is a finite set of prime numbers, there will be a largest prime number. So, the same question can be asked by asking is there a largest prime? So, if it is a finite set, in that finite set, there will be a largest one. And if there is a largest one, then below that largest one there are only finitely many numbers, so there can only be finitely many primes. So, asking whether the set of primes is finite is the same as asking whether there is a largest prime.

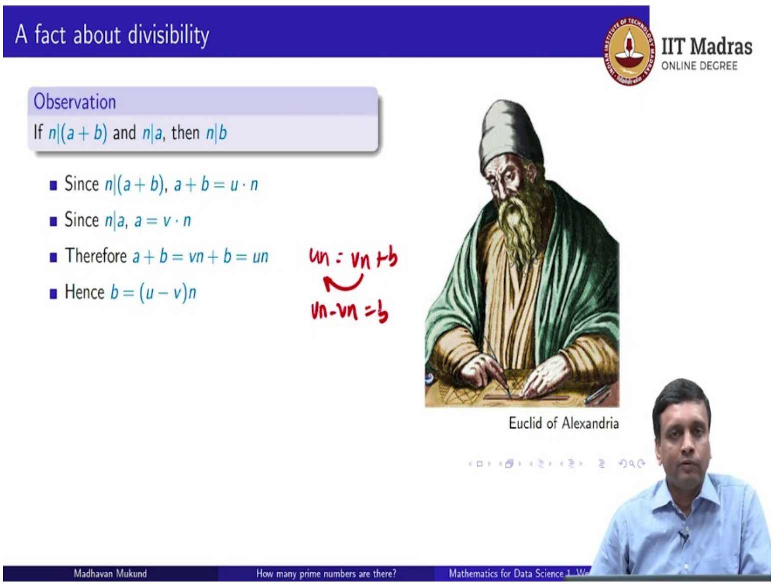
So, what we are going to see is a version of a proof that goes back to Euclid from about 300 BCE, which shows that there cannot be a largest prime. And as we argued if there is no largest prime, then it must be that the set of primes is actually an infinite set.

(Refer Slide Time: 01:35)



So, to go ahead with this we need a basic fact about divisibility. So, this says that if a number divides a+b and it also divides a, then it must divide b. So, let us look at an example. So, supposing you say that 7 divides 21, and I write 21 was 14+7 then 7 also divides 14, and therefore, it also divides 7. Similarly, if I say 6 divides 36+24 which is 60; then since 6 divides 36, it must also divide 24 right. So, this is not very difficult to prove. So, let us prove it just to get a feel of how such proofs go.

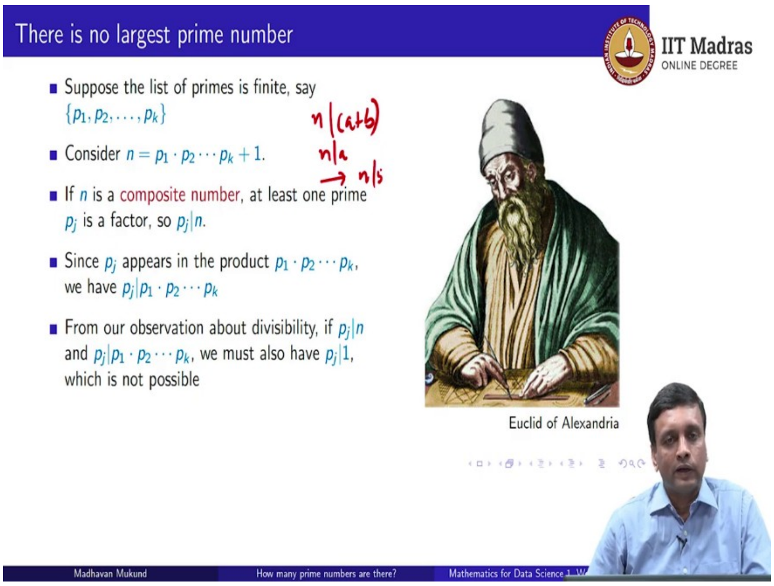
(Refer Slide Time: 02:17)



So, since n divides the sum a+b, a+b can be written as a multiple of n. So, let us call it u times n. Similarly, since we have assumed that n divides a, a can also be written as a multiple of n; let us call it v X n. So, what we are told is that any a + b is u X n for some u, a itself is v X n for some v. And the question is b also some multiple of n does n divide b?

Well, because of what we have just discussed a + b can be written as v n + b, because a is v n, and the sum v n + b which is the same as a + b is in fact u n. So, now, we can do some simple rearrangement. So, we can take u n = v n + b, and just take the v n to the other side and we get u n - v n = b and so b is (u – v) times n. So, this simply proves to us that if a number divides a sum and it divides one part of that sum, it also must divide the other part of the sum. And we will use this in order to show Euclid's result.

(Refer Slide Time: 03:21)



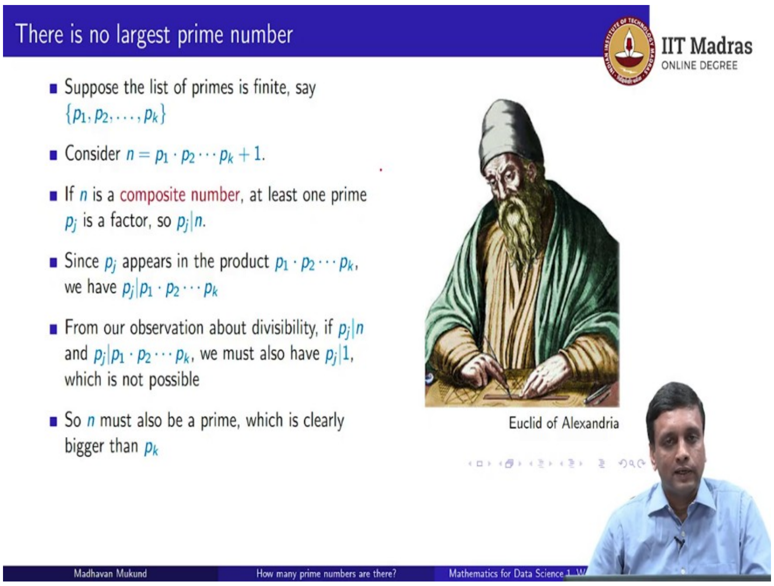
So, what Euclid said is that suppose the list of primes is finite. So, if it is finite, then we can list them out and it is a finite set, so it is some $p_1$to p_k. We do not have to be in any particular order. we can assume that p_1 is the smallest one; it is 2, $p_2$is 3 and so on. But it does not really matter as long as this exhaustively completes all the primes.

Now, we construct a new number which is the product of all these primes, you multiply all these primes by themselves to each other and then we add 1 right. So, n is p_1X p_2X ...X p_k+1. So, now, the question is what is the status of n? So, since we have assumed that the list of primes is finite, n must be a composite number, because this is not one of the primes that we had before right, it is bigger than all of them because it is the product of all of them plus 1.

Now, since it is a composite number it must have a factor other than 1 and itself. And because we have listed out all the primes one of the primes among them must be a factor. So, let us assume that $p_j$is a factor. So, $p_j$divides n right. So, there is one in this $p_1$to p_k, there is a p_j which divides n. But on the other hand, let us look at this part right. The first part the first part is the product of all the prime So, $p_j$appears in that product.

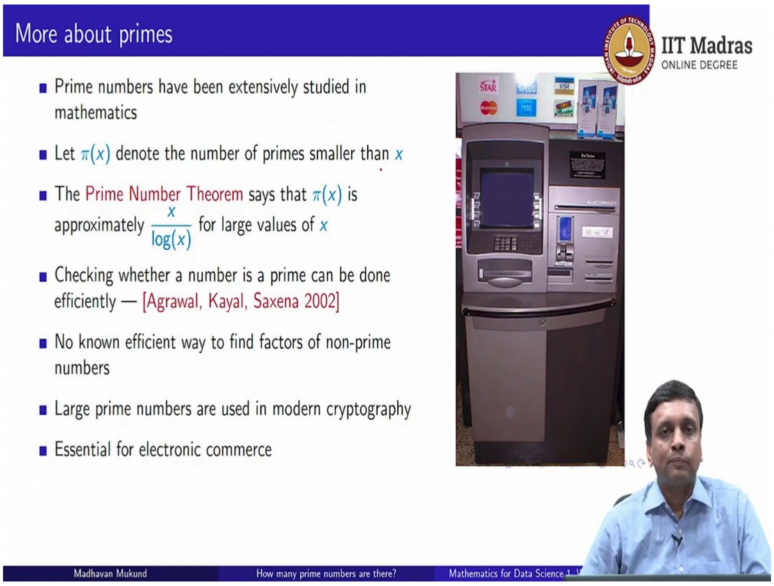
So, if it is one of the factors of the product, it must divide the product right. So, $p_j$divides n, and $p_j$also divides one part of the sum. So, remember what we said that if some number n divides a + b and if some number n divides a also, then n must divide b. So, in this case a+b is the product of the primes plus 1, and a itself is a product of the primes and we have argued that there is one prime p_j which divides both of these. So, therefore, by that divisibility result that we showed in the previous slide $p_j$must divide 1. But of course, we know that $p_j$is a number bigger than 1, it cannot divide 1. And so we have a contradiction right.

(Refer Slide Time: 05:21)



So, what is the contradiction? Well we assume that n was a new number was a composite number because we have exhausted all the primes, but in fact, it cannot be composite because then we cannot find a proper divisor for it among the primes. Therefore, n itself must be a prime. And notice by construction n is actually bigger than all these. So, it also shows that there is no largest prime, because for any set of primes we can always construct a larger prime. So, this is essentially what Euclid did.

(Refer Slide Time: 05:47)



So, we know more about prime numbers. So, prime numbers are very mysterious because their distribution is kind of unclear, but they also have important properties as we will see. So, prime numbers have been extensively studied in mathematics in an area called number theory. So, one of the things that is studied about prime numbers is how they are distributed. So, as we go a larger and larger in the set of natural numbers, how frequently do we find primes?

So, $\pi(x)$ is supposed to denote the number of primes that is smaller than any given number x. So, for instance, π (4) would be 2, because 2 and 3 are the only 2 primes below 4; π (10) would include 2, 3, 5 and 7. So, π (10) would be 4 and so on.

Now, as you go larger and larger, the gaps between the primes become larger. And in fact, you can prove amazing things like the prime number theorem which says that $\pi(x)$ is approximately x / log x for large values of x. Now, it does not matter if you do not understand what this means, but it is important to understand that this is a very significant type of argument that you can give about the distribution of a set of numbers which is quite in a way randomly distributed.

Now, in terms of modern applications of primes, it might seem that primes are very strange things, and we would only need to study them a number theory. In fact, the famous mathematician G. H. Hardy once said that he was very proud of the fact that he did number theory and nothing that he studied had any application. Well, it is not quite true because primes as we will see are actually quite useful.

So, one of the questions that you might want to ask is given a number check whether it is a prime. Now, of course, there is a brute force way of doing it which is to try and enumerate all the factors by looking at all the numbers below n and dividing n by them, but that is not considered to be an efficient way to do it. And in fact, this was proved by three Indian computer scientists from IIT Kanpur, Manindra Agrawal, Neeraj Kayal, and Nitin Saxena in 2002, and it is one of the breakthrough results in theoretical computer science in the history of the subject.

So, checking whether a number is prime can be done efficiently. But what about the other question, if I know a number is not a prime, can I factorize it? So, we know number is not a prime, but how do I find two non-prime, two non-trivial factors that is not 1 or itself. Now, it turns out that there is no efficient way to do this. So, this is quite paradoxical. We can check whether a number is prime or not, but if it is not a prime we can factorize it fast. And this in fact is the reason why we are so concerned about prime numbers, because we would like to find numbers which are not prime, but which are actually products of large primes. So, their factors are only large prime numbers, and this is a very important in cryptography.

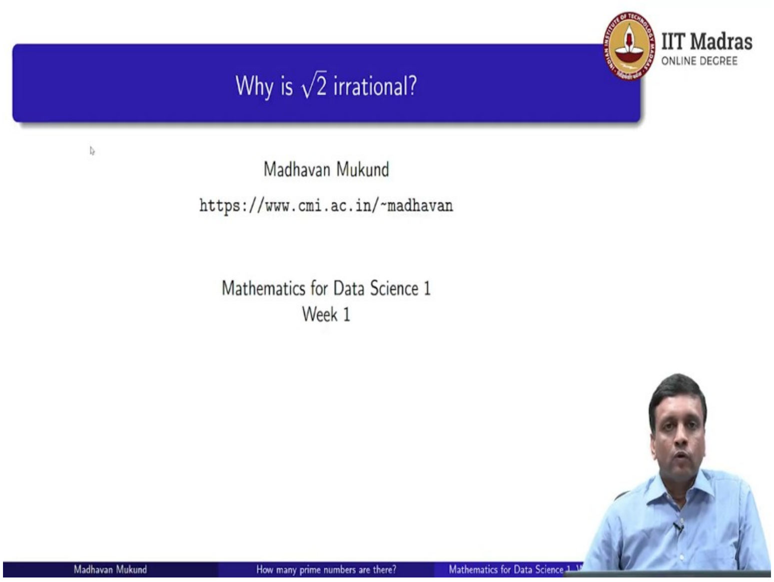
And cryptography in this sense is something which affects not just you know military secrets, but it affects us in day-to-day life because whenever we do electronic commerce our transactions are protected by cryptography to prevent unauthorized transactions from being executed on our behalf or to prevent them from being tampered with they are all encrypted. And a lot of this encryption is based on the existence of large prime numbers, and the fact that factorizing the product of two large primes is difficult. So, prime numbers though they are very exotic in number theory are actually a very, very important part of our day-to-day life.

Mathematics for Data Science 1
Prof. Madhavan Mukund
Department of Computer Science
Chennai Mathematical Institute

Week - 01
Lecture – 09
Why is a number irrational?

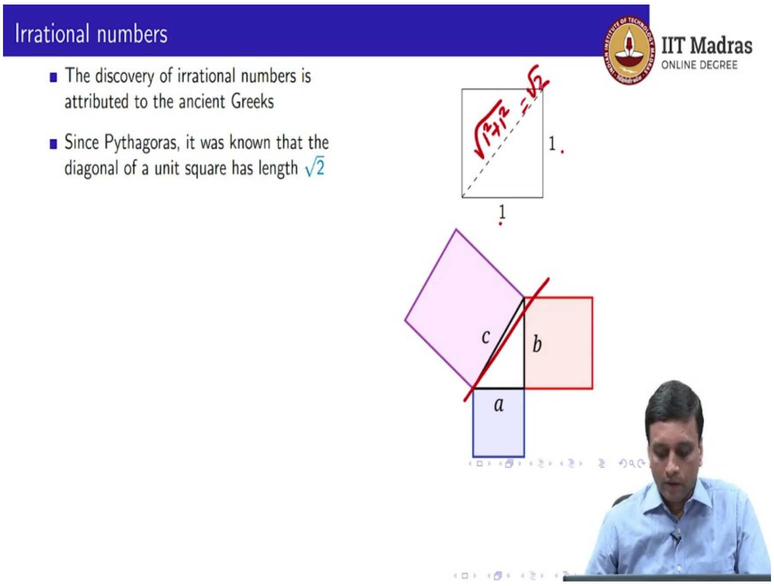
(Refer Slide Time: 00:06)

(Refer Slide Time: 00:14)



When we looked at the different types of numbers, we started with the natural numbers, move to the integers, then to the rationals which are expressed as $\frac{p}{q}$. And then we argued that the rationals do not exhaust all the numbers that we need; and in particular, we claim that the √❑ cannot be expressed as a rational numbers, so it is what is called an irrational number. So, let us try and ask why √❑is an irrational number.

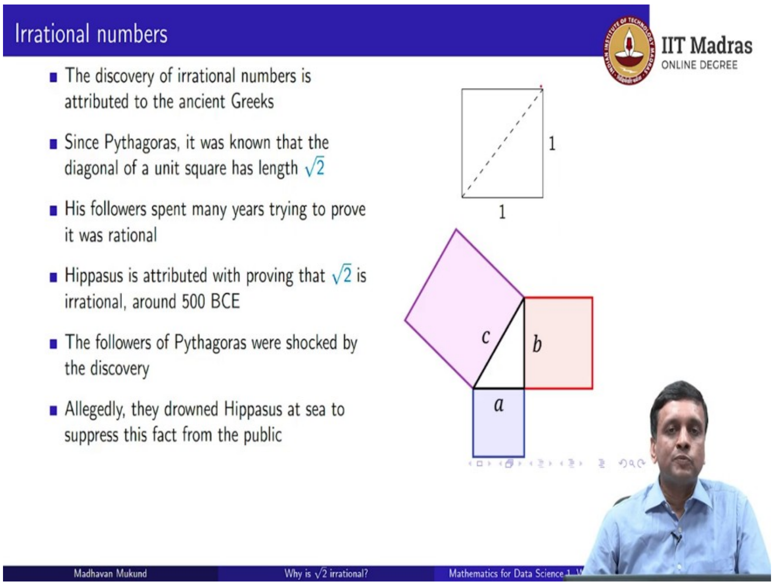
(Refer Slide Time: 00:37)



So, the discovery of irrational numbers actually is attributed to the ancient Greeks; and in particular, it comes from Pythagoras. So, remember that in Pythagoras's theorem which you must have studied in school. If you have a right angled triangle, then the square on the hypotenuse that is the square on the long diagonal side – this one, has an area which is the sum of the squares on the other side. So, in other words, if you have a right angled triangle and you measure the three sides, you get $a^2 + b^2 = c^2$. So, from this, knowing a and b, you can compute c.

So, in particular, if you draw a square which has one and one as its two sides, then this must be the √❑which is the √❑. So, you can actually physically draw if you assume that you can measure out a unit length using some kind of a measure, then by drawing a square, you can actually construct a length √❑.

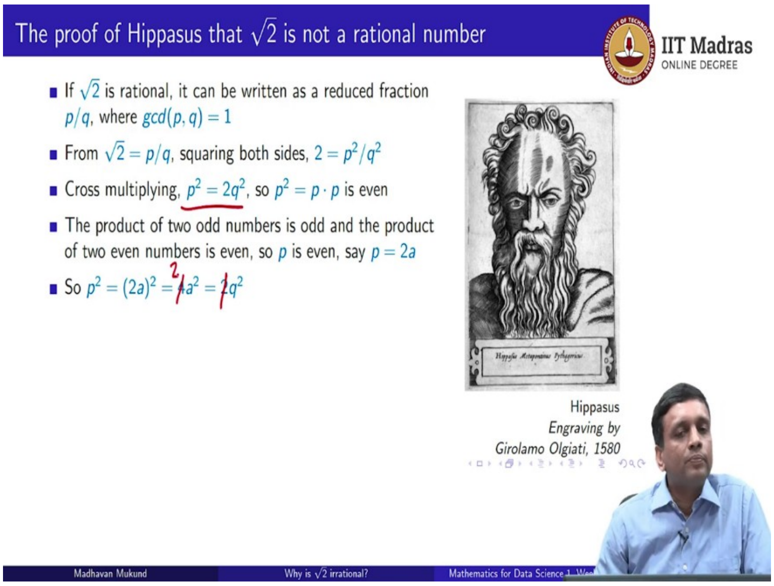
(Refer Slide Time: 01:35)



So, for Pythagoras it was very important to understand how to describe the √❑as a rational number, and he and his followers and many times many years trying to prove that in fact it could be expressed as a rational number. Much after Pythagoras, about 50-60 years after Pythagoras, one of his followers Hippasus is claimed to have proved that √❑is irrational this was around 500 BCE.

Now, the followers of Pythagoras had a very mystical idea about numbers, and they felt that numbers could solve everything. And in particular they were very keen that rational rational numbers should form the basis of all of what we could call it modern day time science and philosophy. So, the followers of Pythagoras were really shocked by this discovery of Hippasus, they found it to be a, I mean they could not argue with it; at the same time they felt that this discovery could not be revealed to the public because they felt it was very dangerous. So, in fact, it is said that they allegedly drowned him in the sea to prevent this from being made public. So, the √❑being irrational has a rather colorful history. And let us see now how Hippasus proved that this was actually the case.

(Refer Slide Time: 02:41)



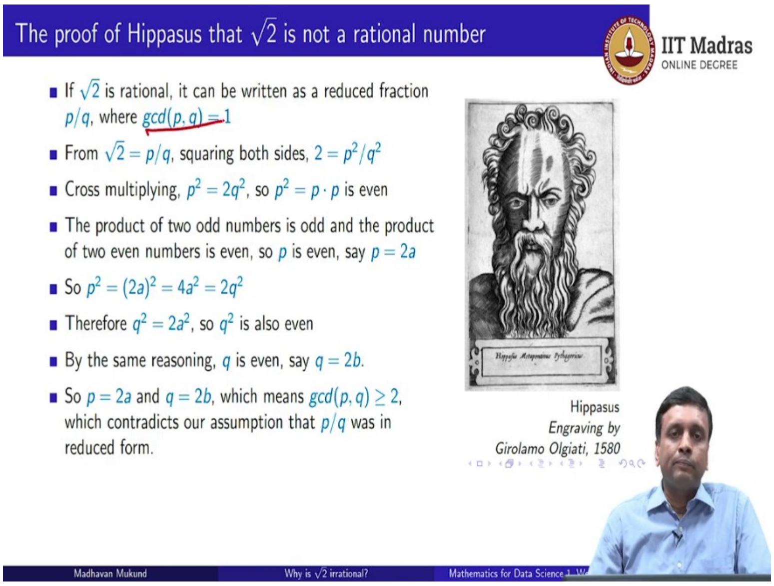
So, let us assume as in many of our arguments. Let us assume that √❑was rational. So, if it is rational, then we know that it can be written as a ratio or fraction of two integers p and q; and in particular we can assume that it is in reduced form. So, p and q have no common divisor, their gcd is 1. So, if we take √❑is equal to $\frac{p}{q}$, and we square both sides, then √❑times √❑is 2 on the left hand side, and $\frac{p}{q}$times $\frac{p}{q}$is $\frac{p^2}{q^2}$. So, we get 2 is equal to $\frac{p^2}{q^2}$. So, we can cross multiply as usual, take the $q^2$from the denominator on the right hand side to the left hand side numerator, and we get $2q^2$ is equal to p^2.

So, what is p^2? p^2 is p X p. And if it is of the form 2 times something, then it is an even number, because an even number is something which has 2 as a factor. So, p^2has 2 as a factor. So, $p^2$is an even number. Now, it is a basic fact about natural numbers that if you multiply two odd numbers, you get an odd number; and if you multiply two even numbers, you get an even number. So, if $p^2$is even, and p^2 is p X p, then both p and p – the two copies must both be even; so p must be an even number in other words.

So, if p is an even number, then we can write p as 2 times something because p is even p must be of the form two times something say 2a right. So, from this initial assumption, we have concluded that the numerator of this fraction which represents √❑is actually an even number of the form 2a.

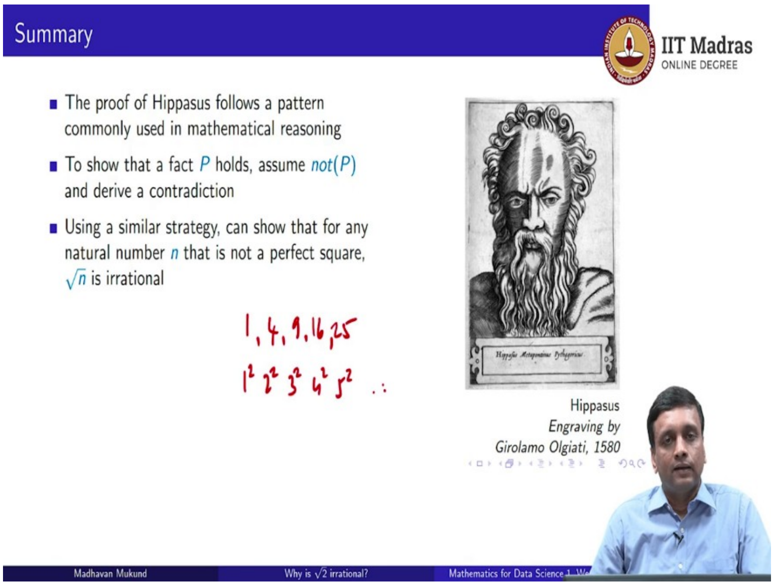
So, now, let us substitute in this equation for p^2right. So, $p^2$is $(2a)^2$ is 4 times a^2. So, now $4a^2$ is equal to $2q^2$. So, now, we can cancel right. So, we can take this 2, and this 2, and cancel it.

(Refer Slide Time: 04:39)



So, we have in other words that $q^2$is $2a^2$. And if $q^2$is $2a^2$, then by the same argument as before $q^2$is also even, and so q must be even. And therefore, q can be written as the form of 2 times some other number b. So, we have that p is of the form 2 times a and q was of the form 2 times b. But what this means is that the gcd of p and q must be at least 2, because both of them are even numbers. So, they are both multiples of 2. So, we claimed initially that the gcd of p and q is 1. We said that they were actually both in reduced form. So, there was no common factor other than 1. And now we have shown that if we assume that we in fact generate 2 as a common factor. So, this cannot be the case. So, the only contradiction that we can resolve with this is by assuming that $\frac{p}{q}$could not have been there. So, therefore, $\sqrt{\square}$ cannot be represented by any reduced fraction $\frac{p}{q}$.

(Refer Slide Time: 05:32)



So, this argument of Hippasus is a common way of arguing things in mathematics right. To show that some fact capital P holds you first assume that not P holds, it is negation holds. So, we wanted to show that there is no way that √❑cannot be expressed as rational. So, we said let us assume the negation. Let us assume that √❑can in fact be express as a rational, and then you take that assumption and derive a contradiction. And since you cannot accept a contradiction, your assumption must be wrong and therefore, what you tried to prove originally was correct.

So, in fact, it is not just √❑that is irrational, √❑is also irrational. Now, 4 is a perfect square. So, we know that √❑ is 2. What about√❑; that is also irrational. So, among the integers among the natural numbers we have the perfect squares 1, 4, 9, 16, 25 and so on which consists of 1^2, 2^2, 3^2, 4^2, $5^2$and so on. So, a perfect square is one whose square root is also a natural number.

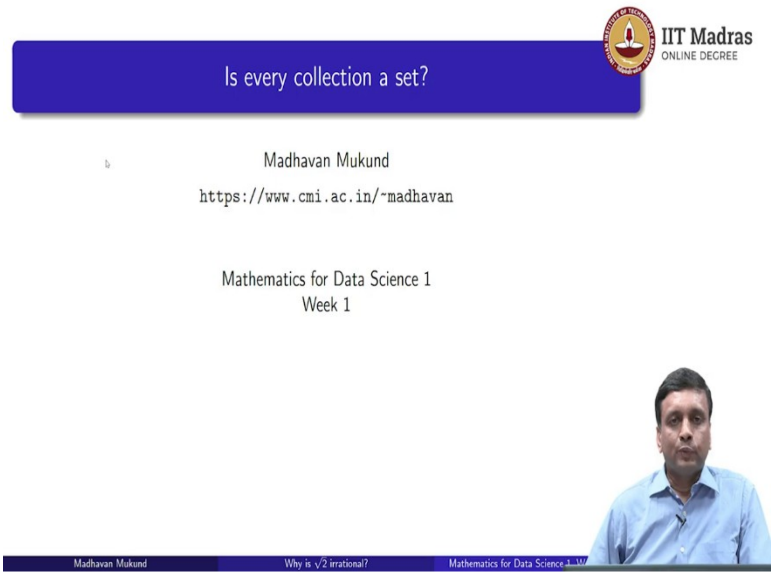
Now, it turns out that anything which is not a perfect square has an irrational square root, and the proof is not exactly the same because we have used a property of 2, and evenness in this proof, but with a very similar argument you can show this is the case. So, therefore, there are a lot of irrational numbers that you can generate just by taking square roots of non-perfect squares.

Mathematics for Data Science 1
Prof. Madhavan Mukund
Department of Computer Science
Chennai Mathematical Institute

Week - 01
Lecture – 10
Set versus Collections

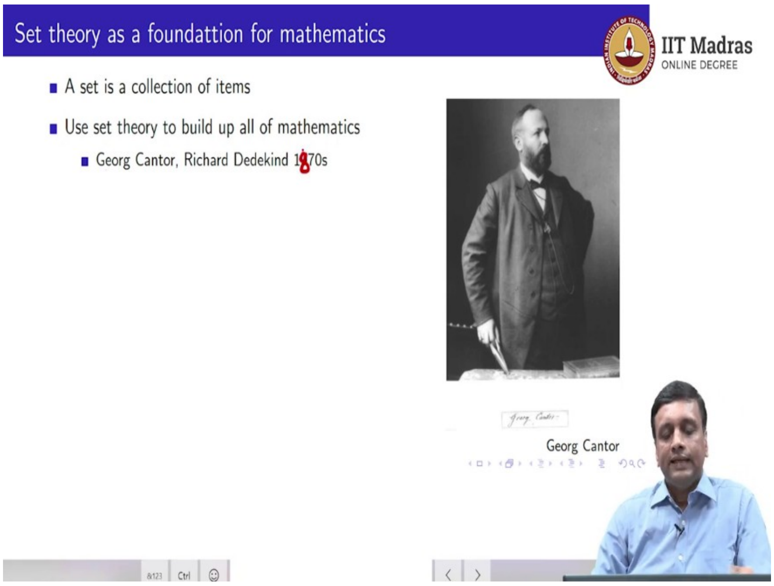
(Refer Slide Time: 00:06)

(Refer Slide Time: 00:14)



So, we have looked at sets, and we said that a set loosely speaking is a collection of items. And then we made some remarks in that lecture that not everything can be thought of with a set. So, let us ask whether every collection is in fact a set, and if not, why not?

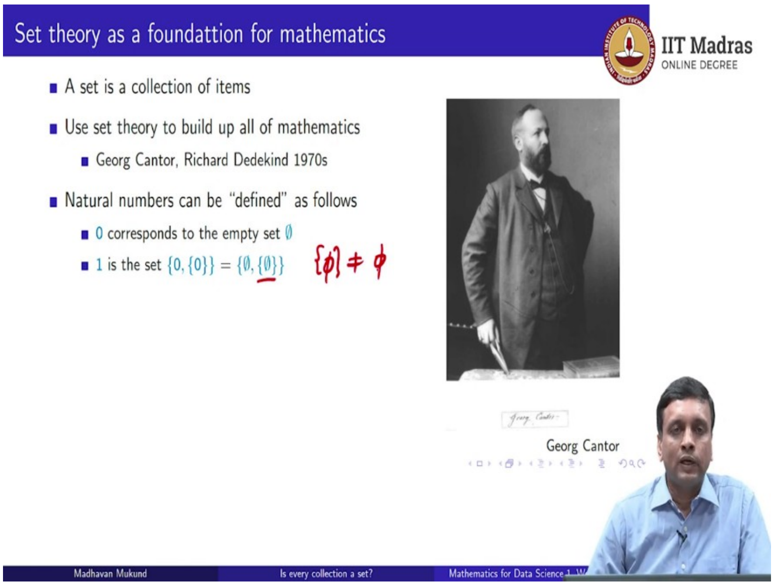
(Refer Slide Time: 00:29)



So, as we said a set is a collection of items. And when set theory was investigated formally starting from the late 1800s, the idea was to make set theory a foundation of mathematics. So, let us try to briefly understand what that means. So, we wanted to the mathematicians of the time wanted to start off with very basic things and build up all of mathematics from that, and they felt that set theory was a good place to start.

So, some of the mathematicians who are involved in this was Georg Cantor and Richard Dedekind from the 1870. So, this is a mistake, this is not the 1970s of course, but the 1870s.

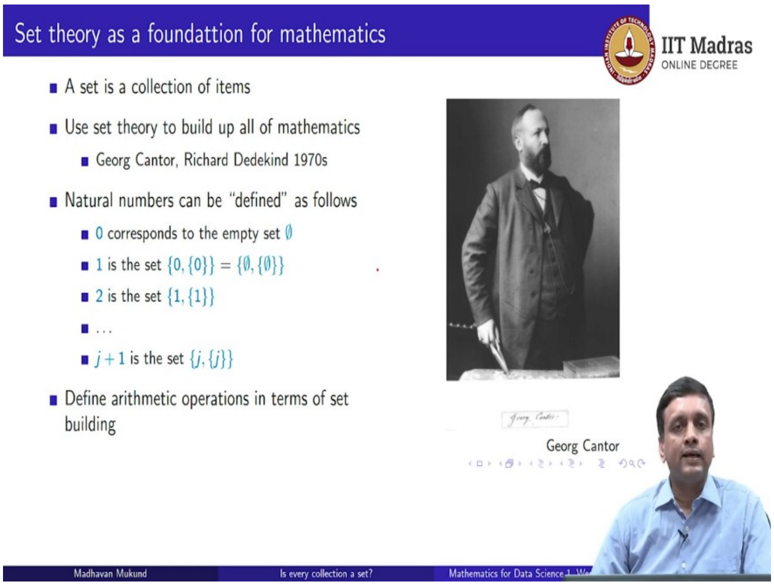
(Refer Slide Time: 01:05)



So, one is aspect of this foundational nature of set theory is for insists how do you generate numbers if you have only sets. So, one of the things that you need if you start with set theory is the empty set. So, you have it for free. So, what they said is that 0 can be thought of as the empty set.

So, we are going to use sets to represent numbers, and we are going to use the empty set to stand for 0. So, what is 1? Well, 1 is a set that consists of 0 and the set containing 0; in other words it is a set containing the empty set, and the set containing empty set. So, remember that the set containing empty set this is not the same as this right. The empty set has no elements; the set containing the empty set has one element.

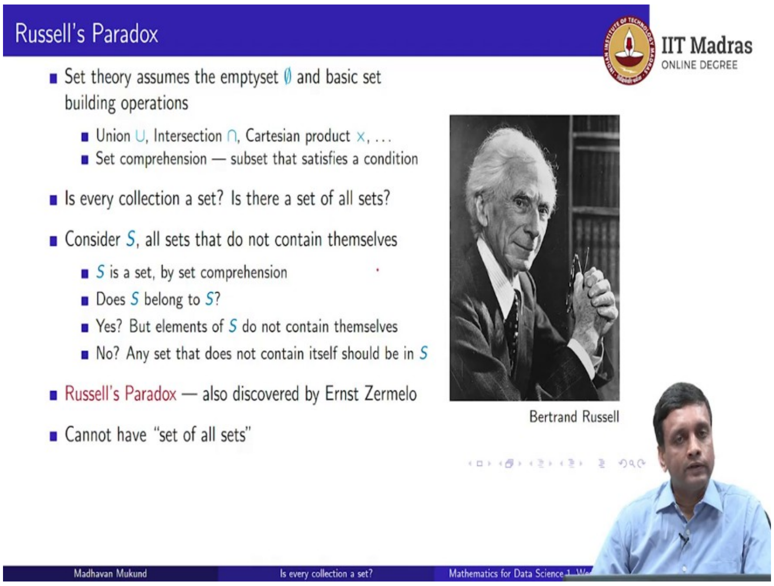
(Refer Slide Time: 01:52)



Similarly, 2 would be the set which contains 1 in the representation above, and the set containing 1. So, it is a bit tedious to write out. So, I have not expanded it. But you just take the expression for one in terms of the empty set replace it twice, and you get the number 2. And in this way for any number j plus 1, you can get it from the number j by taking the representation of j adding the set containing the representative j putting it into a new set.

So, these are the natural numbers as expressed using sets starting from the empty set. And then you can actually define set theoretic ways of combining these two define, the addition of two numbers and this format to get a new number which is the sum and the product and so on. So, this is what it means to use something like set theory as a foundation of mathematics.

(Refer Slide Time: 02:37)



So, basically set theory assumes that you have the empty set, and then you have basic set building operations. For instance, you can take the union of sets, you can take the intersection of sets, you can take the Cartesian product which we saw when we were looking at relations. And you can of course to set comprehension which is that you can take some elements from a set which satisfy a condition and build a subset.

So, now into this picture came Bertrand Russell and he asked whether this would make sense or not. So, here we come back to our fundamental question is every collection a set? In particular he asked can there be a set of all sets? So, remember that sets are objects just like anything else. So, we can collect them together. So, is this collection of all sets in fact a set?

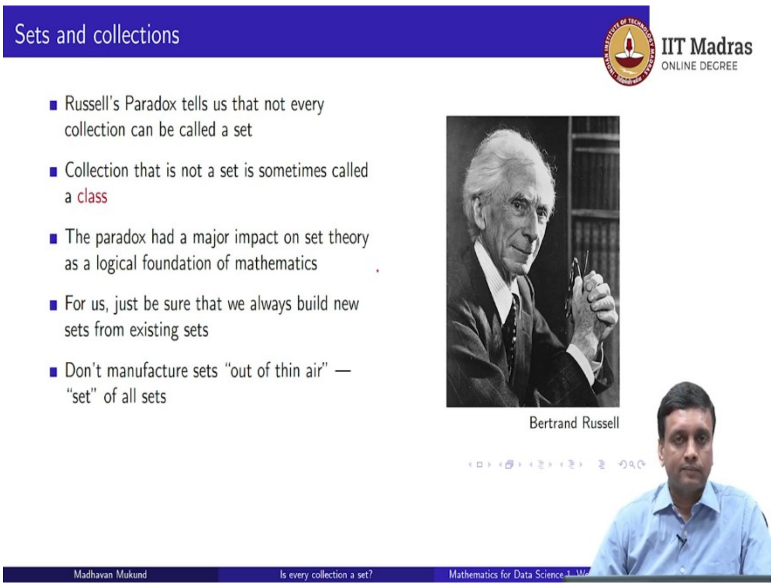
Well, supposing it is a set, then we can do the following. We can apply set comprehension right, and we can pick out some sets from this collection of all sets. So, we will call capital S, the subset of all sets that do not contain themselves. So, this is a subset of this hypothetical set of all sets. So, this capital S is a set because we have applied set comprehension to the set of all sets. So, we have the set of all sets. And among all sets we have pulled out those sets which do not contain themselves. So, this is the condition we have applied, and this is allowed by set comprehension.

Now, the question is does the set that we have constructed belong to itself, does S belong to S? Well, if it does belong to itself, then it does not satisfy its own definition because elements of S should not contain themselves. So, S cannot belong to itself, because if it did it would contradict to way we have pulled out S from the set of all sets. But if it does not belong to itself, then that is also a contradiction, because then S does not belong to S and by the condition that we have applied to pull out sets S must be included in that condition.

So, either way we have a paradox; we have a contradiction. So, S can neither belong to itself nor can it not belong to itself. And this is called Russell's Paradox. He was the first person who published this and made it publicly known, but this was also independently discovered by another well known set theorist of the time called Ernst Zermelo. So, what this really tells us? If you remember our argument is that we made some assumption, and then from that assumption we realized that we have a contradiction or an observed situation.

So, something must be wrong in one of our assumptions. And here it turns out that the assumption that goes wrong in all these is the assumption that there is a set of all sets. If we did not have a set of all sets, we could not have done the set comprehension, and therefore, we would not have reached this observed conclusion.

(Refer Slide Time: 05:23)



So, what Russell's Paradox really tells us is that, not every collection can be called as set in particular the set of all sets does not exist. So, he went through an exercise of trying to formulate a different version of set theory which he called type theory and so on, but in modern mathematics typically if you are not sure that what you are dealing with is a set then it is safer to just called such a collection a class. So, a class is just a collection of objects which does not have any of the implied properties that you expect from the sets.

So, this paradox as we said came in the context of set theory being used as a foundation of mathematics. And, this seem to casts doubts on whether it could be used at all. So, it had a major impact on this whole mathematical exercise of deriving mathematics from logical foundations which went on into the 20th century which we will not be able to discuss here unfortunately, but it is a fascinating subject in its own right.

For us what we have to be clear about is that whenever we use sets we must make sure that we always start with sets that we have and build new sets from existing sets. So, we can assume that the numbers are sets. So, we have the set of natural numbers, the set of integers, the set of rationals, the set of reals and so on. And, whenever we construct a new set we just have to verify that the set that we started with to construct the new set was already a set.

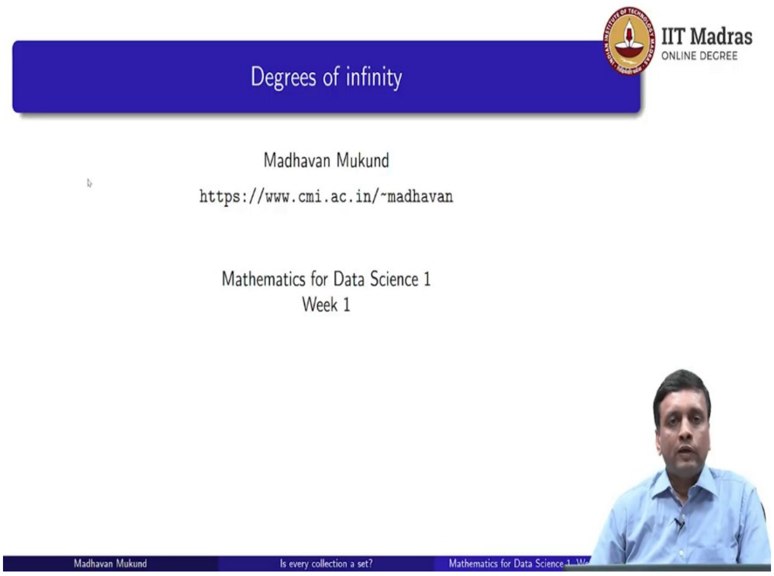
So, we take a Cartesian product or a union or set comprehension, we always start with old sets and make new sets. So, those old sets must be well-defined. So, in other words, we should not manufacture sets out of thin air such as the set of all sets.

Mathematics for Data Science 1
Prof. Madhavan Mukund
Department of computer science
Chennai Mathematical Institute

Week - 01
Lecture – 11
Degrees of infinity

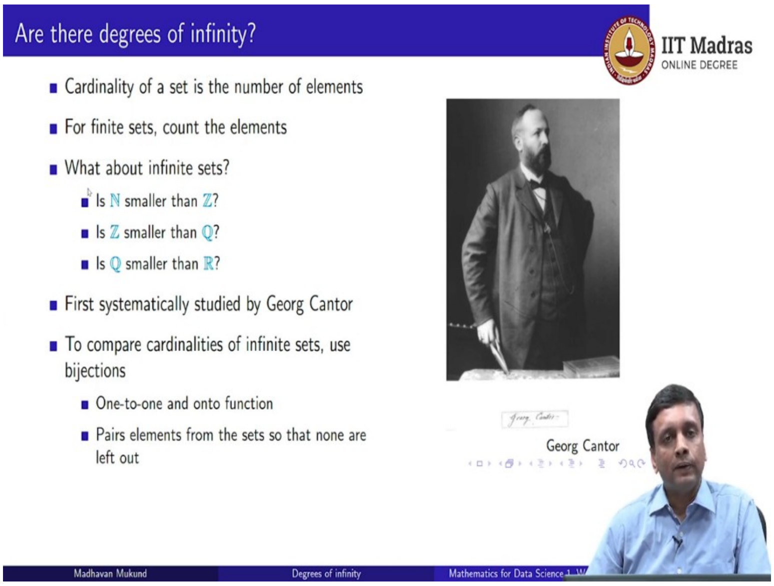
(Refer Slide Time: 00:06)

(Refer Slide Time: 00:14)



So, when we looked at the sets of numbers, we said that we have various kinds of infinite sets – the natural numbers, integers, reals, the rationals, some of them are discrete, some of them are dense. And the question that we asked was whether they all have the same size, or there are more of one than another?

(Refer Slide Time: 00:29)



So, the question that we want to ask is, are there degrees of infinity? So, we know that for a set the cardinality denotes a number of elements, and if it is a finite set we just have to count these elements. So, for a finite set, there is no problem about cardinality which is the count the number of elements and we are done.

We get a natural number which is the cardinality of the set. Now, the question is what do we do for infinite sets. So, let us look at the natural numbers for instance. So, in which we move from the natural number to the integers, we added negative numbers. So, clearly we have added an infinite set of numbers we roughly doubled the set. So, is the set of natural numbers are same as the integer number in size or not?

Similarly, when we move from the integers to the rationals, we move from a discrete set where we had a next and previous element to a dense set where between any two element there is an another element. So, this suggests that there should be more rational than reals rationals than integers, but is that true or not?

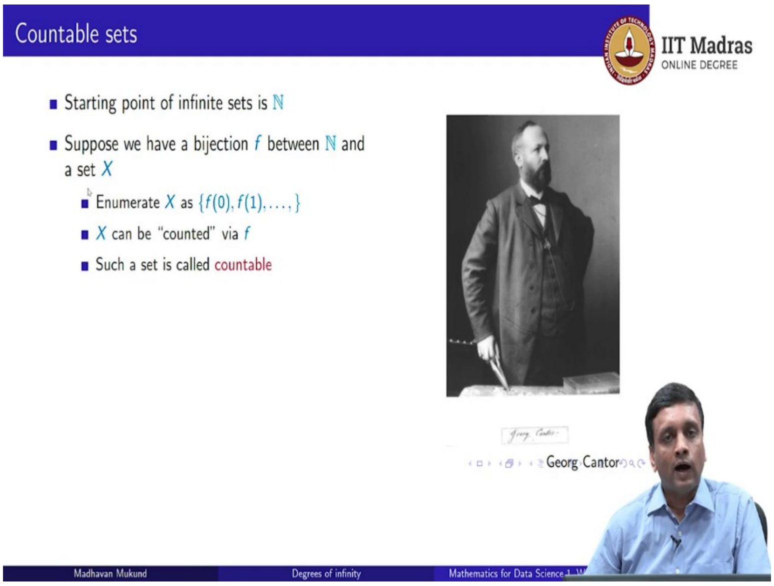
And finally, when we move from rationals to real numbers we added a whole bunch of irrational numbers which cannot be expressed in the form $\frac{p}{q}$. So, clearly the real numbers have a large number of new things which are not in the rationals. So, again is the set of reals larger than the set of rationals or not? So, this study of the cardinality of infinite sets was actually undertaken by Georg Cantor in the 1870s. And as we have seen when we studied functions the correct way to compare the cardinality of infinite sets is to use a bijection.

So, what is the bijection? The bijection is one-to-one and an onto function. In other words, it allows us to map one set to another set in such a way that two elements are always map to two different elements and everything on the other side is map 2 from something here that is the onto part. So, it is one-to-one no onto elements map to the same one, and it is onto no element on the right hand side is missed out.

So, intuitively what this allows us do through this function this bijection is to pair up the elements from the one side with the elements from the another side. So, I take an element on the left hand side, through the bijection I pair up it with an element of right hand side. And because it is one-to-one and onto, this pair this paring actually exhaustibly covers all the elements in both sides or nothing is left out.

So, we have paired up everything and therefore, the two sides have the same cardinality. So, this is the technique that we will investigate in order to resolve these questions about the cardinalities of the infinite sets of numbers that we have discussed above.

(Refer Slide Time: 02:55)

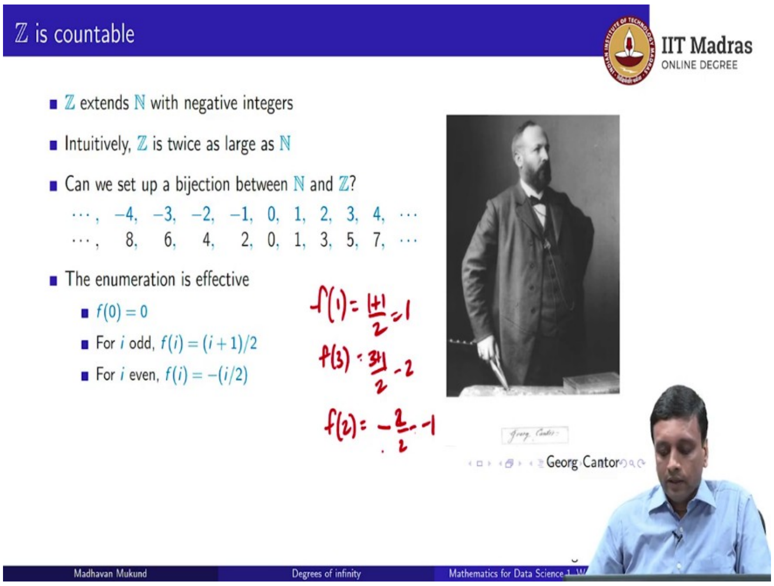


So, our starting point is the set of natural number, because this is the first infinite set that we have to begin with. When we start counting we realized that there is no largest number because we can always add 1. And so if we take all the finite numbers that we can used to count, we get an infinite set called the natural numbers.

Now, supposing we find a bijection between the set of natural numbers and some other set X, does not matter what this set is, but supposing there is a bijection. We can pair of the natural numbers with the elements of X. This means that we can actually effectively enumerate the elements of X, we can take the number paired with 0, f (0) and call that the beginning of X, then f (1) is an X element, f (2) and so on.

And because we are doing this kind of enumerating X, we can count X in a way via f and so we call any such set countable. So, countable set is one which can be bijectively paired up with the set of natural numbers. So, when we are looking at other sets, we will first check whether they are countable or if not we have to argue that they cannot be counted.

(Refer Slide Time: 04:02)



So, let us begin with set of integers and show that it is countable. So, why should be, why should it not be countable, or why should it be a surprise if it is countable? Well, because Z extends N with negative integers right. So, for every, if you do not count 0 in the calculation, for every positive natural number there is a corresponding negative integer in Z.

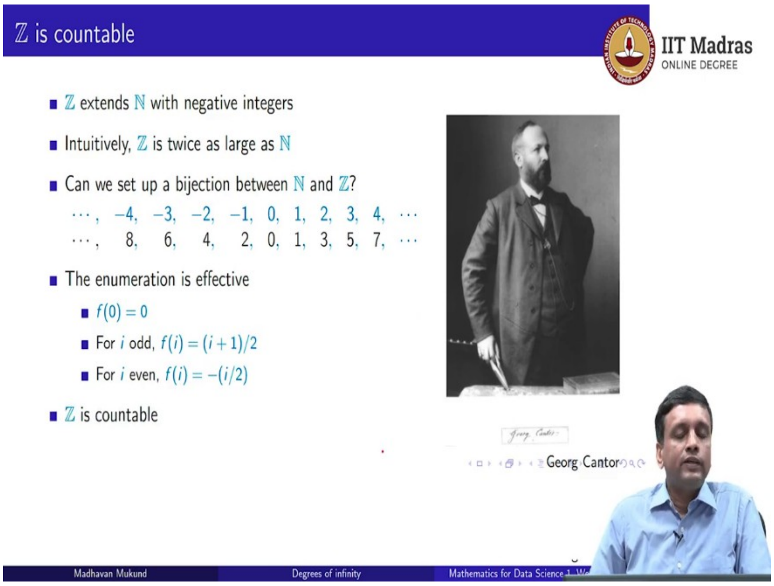
So, Z is referring twice as big as N; for +1 you have -1; +2 you have -2 and so on right. So, it seems contradictory that you can double the set, and still have the set of the same size that you started with. So, the question now is for Z to be countable, can we set up a bijection between the natural numbers and Z?

So, let us look at Z as we do on the number line. So, it starts from some -∞ and then it comes to -4, -3, -2, -1, 0, 1, 2, 3, 4 and continues. So, we start our enumeration at 0. So, we enumerate 0, the 0 of Z as the 0th element, then we map 1 to +1, map 2 to -1. What do we do next? Well, we map 3 to +2, and 4 to -2.

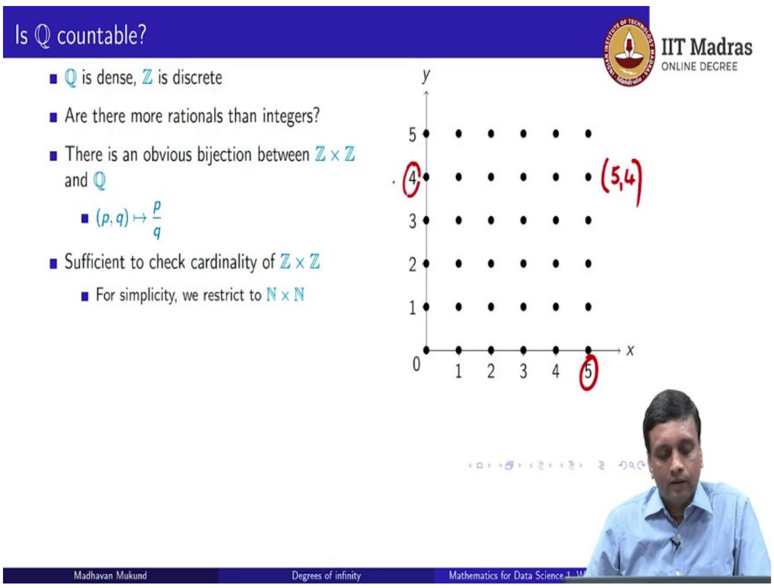
So, we keep zigzagging to the right hand to the left, we count Z by starting with the center moving right one, moving left one, moving right one, moving left one. So, in this way we could now enumerate the number +3 as 5, -3 as 6, +4 as 7, -4 as 8. So, in this way we can actually enumerate Z effectively. So, f (0) is 0 as we saw. If i is odd for example, 1 then f (i) is $\frac{i+1}{2}$.

So, f(1) for instance is (1 + 1)/ 2 = 1; f(3) = (3 + 1) / 2 = 2 and so on. So, if f is odd, I have $\frac{i+1}{2}$. And if it is even like 2, then I take - $\frac{i}{2}$. So, I take -2/2 which is -1. If it is 4, I get -4/2 which is -2 right. So, we have actually given an effective way of assigning a position in some sense or count to every number in Z, and this shows that the set of integers is actually countable.

(Refer Slide Time: 06:17)



(Refer Slide Time: 06:20)



Now, what about the rationals? One reason why we might suspect that the rationals are not countable is because the rationals we saw a dense between any two rational numbers there is an another rational number.

Whether the integers and the rational numbers are discrete, you can always find a next number; and in the case of integers you can always find a previous number. For natural numbers 0 has no previous number, every other number has a previous and a next. So, given that rationals are dense and the integers are discrete, the question is are there more rationals than there are integers?

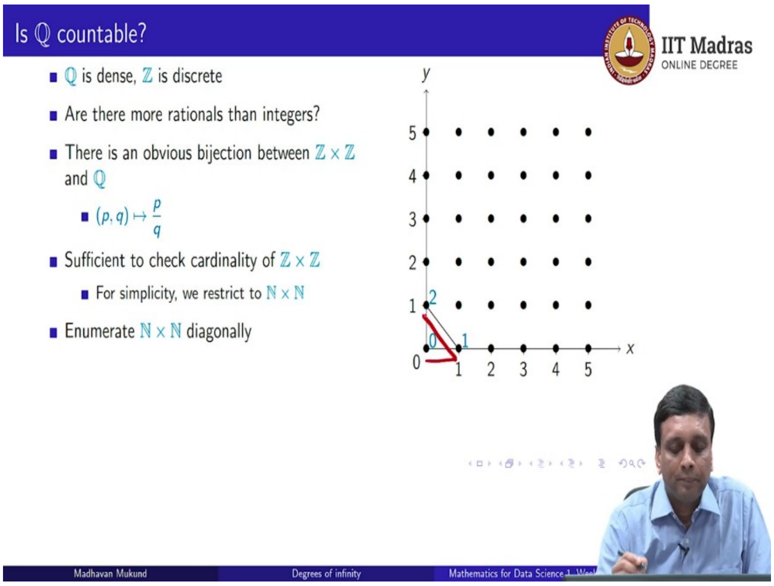
Now, there is an obvious bijection between pairs of integers and rationals because that is what a rational is, rational is a pair of integers p upon q. So, I can take a pair (p , q) and Z cross Z and directly connect it in an bijective way to the fraction $\frac{p}{q}$. So, every pair gives a unique rational number, every rational number gives me a unique pair.

There is no surprise here, there are no we are not talking about reduce forms of for example, we have different numbers like $\frac{1}{10}$, then we have $\frac{2}{20}$, and $\frac{3}{30}$, these are all different rational numbers they may represent the same value, but they represent different pairs. So, this is a clear bijection between Z cross Z and Q. So, Z cross Z has the same size of Q.

So, if we are looking at the cardinality of Q, we can also look at the cardinality of Z cross Z. Because if we can measure the cardinality the size of Z cross Z, then through this bijection, Q must have the same size, there is no need to separately measure the size of Q.

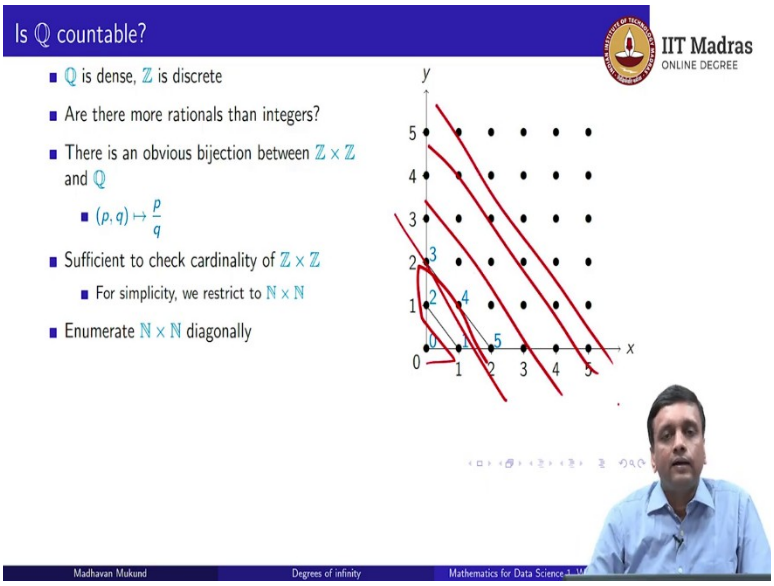
So, instead of Z cross Z just to make the picture easier to see, we will actually do N cross N, and then I will show you how to extend it to Z cross Z. So, here is a picture of N cross N. So, remember that we think of N cross N in a two-dimensional grid and at each point (i, j) I have a dot representing the pair (i, j). So, for instance this pair, this pair is (5 , 4), because it comes from the 5 and the 4 over here right. So, every dot in this pair in this grid is a pair in N cross N.

(Refer Slide Time: 08:22)



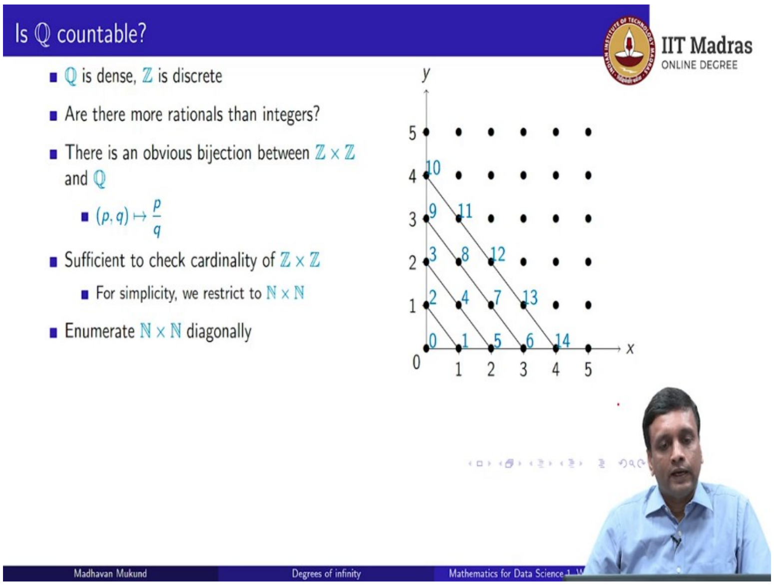
Now, I am going to enumerate this in a particular way. So, here is a one enumeration. So, you start with the 0th element as the element at the bottom left corner what is normally called the origin. Then you enumerate the first diagonal right, so you go from here and then you go right and then you go up. So, you enumerate in this way then you continue.

(Refer Slide Time: 08:47)



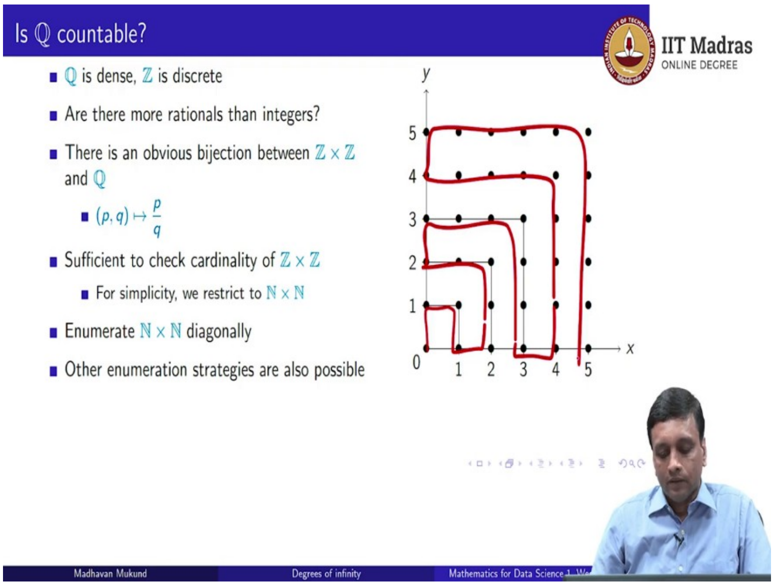
So, you started from here went up, then you up there, and come back down again right. So, you can slice this thing like this right. So, you can slice this gird like this, and enumerate it diagonal by diagonal.

(Refer Slide Time: 08:58)



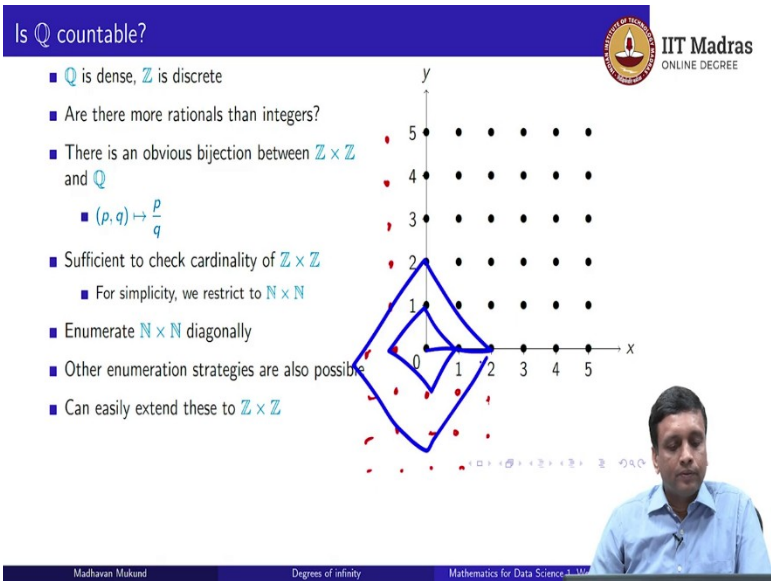
So, this gives us an effective enumeration of N cross N.

(Refer Slide Time: 09:04)



But we can also enumerate in different ways. For instance, we can enumerate in these larger and larger squares. So, we can start here, then finish this, then do this, then do this, then do this and so on right. So, long if we do not miss out any point in the grid we are done.

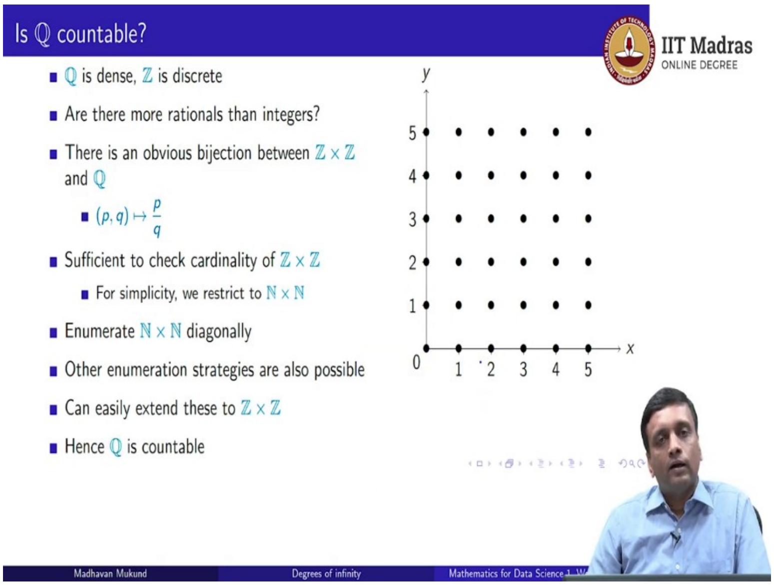
(Refer Slide Time: 09:21)



So, this shows us that N cross N is something that we can enumerate. Now, how would we do it for Z cross Z? Well, it is very simple. If I had Z cross Z, I would also have points on this side, and I would also have points below right. So, I would have points to the left and below 0 because I would had negative numbers.

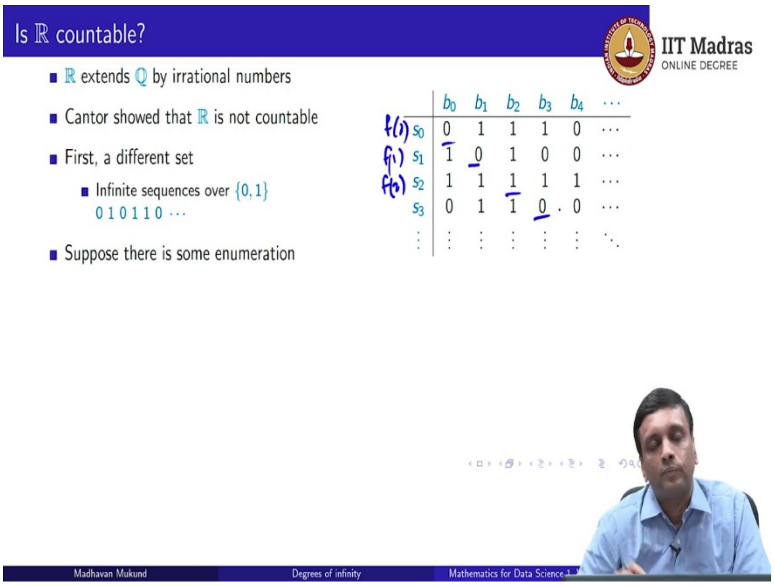
So, now, if I wanted to enumerate Z X Z, I would start here, then I would do this, and I would complete this diamond, then I would go here, and then go here, and then complete this diamond and so on right. So, instead of doing just the diagonal, I would extend the diagonal around to form a diamond, and in this way I would start from the center and spiral out so that I enumerate all the numbers in Z X Z. So, N X N can be enumerated as we saw, and this can be easily extended to Z X Z.

(Refer Slide Time: 10:10)



So, therefore, the set of rational numbers though it is dense and then it looks superficially to be much larger than the set of integers, actually both the integers and the rational numbers have the same number of elements which is quite surprising, but it is true.

(Refer Slide Time: 10:26)



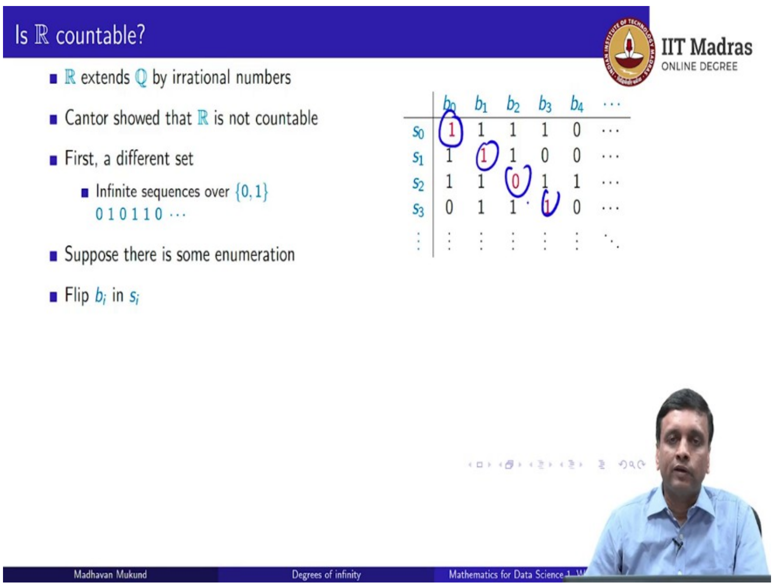
So, for all the infinite sets we have seen are countable right. Of course, the natural numbers are countable by definition, and then we saw integer are also countable, and the rational are also countable. So, what about the real numbers? So, how did we get to the real numbers? We took the rationals and then we added all these irrational numbers like √2, π, e and so on. So, Cantor showed that R actually is not countable. So, let us see how this proof works.

So, actually he did not, he did have a separate proof that R is not countable, but later on he made another proof which is easier to present which starts with the different set. So, instead of looking at R, we will look at something which looks quite different. We will look at infinite sequences over 0, 1. So, an infinite sequence of a 0, 1 is just something like you just keep writing down 0 or 1 infinitely many times without stopping right.

So, what Cantor argued is that this set is not something that you can count. So, supposing you can enumerate the infinite sequences over 0, 1, then on the right to see some enumeration; we are not looking at a particular enumeration in some particular order. We are just saying is there any enumeration at all, so that I can write down the 0-th sequence. So, this is the 0th sequences, this is f (0) in some sense, this is f (1), this f (2) and so on.

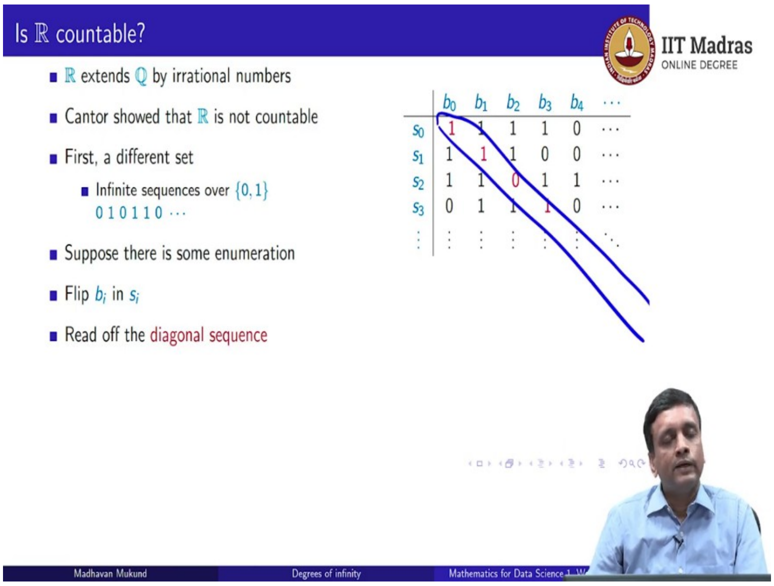
So, I have just written f (0) as s_0, and f (1) as s_1, and so on. And each sequence has positions which I have written b for bits because these are binary digits 0 or 1. So, each sequence has an infinite sequence of bits which characterize what it is, and no 2 rows are the same they are all different infinite sequences of 0s and 1s right. So, hypothetically this table is an enumeration of such sequences. So, if this is a enumeration of all such sequences, can be derive a contradiction? So, this is how Cantor derived a contradiction.

(Refer Slide Time: 12:26)



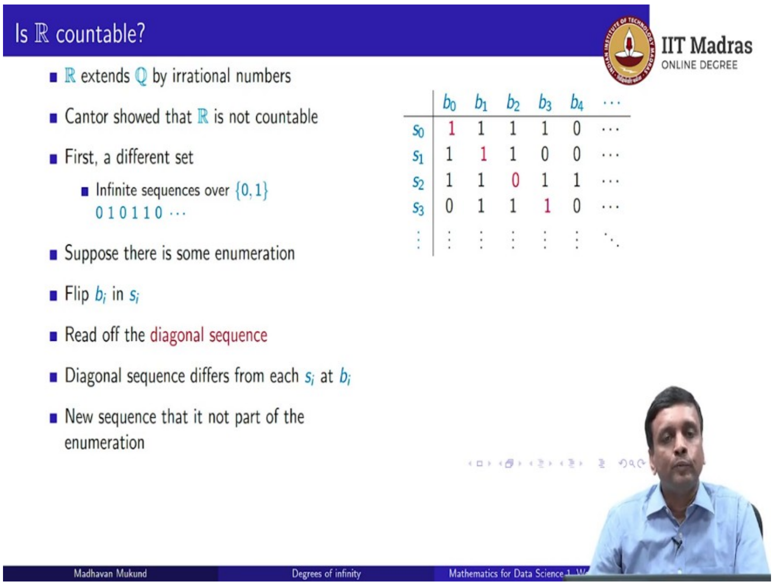
So, he said let us take each row and reverse the bit. And which bit to be reversed? Well, if we are in the ith row, then we reverse the ith bit. So, in the first row which is s 0, we reverse b_0 , in the second row. So, if you want go back, so this was 0, so we are here at 0 0 1 0. So, after flipping, it becomes 1 1 0 1 right. So, what we are doing is in 0th row, we are flipping b_0; in row s_1 we are flipping b_1; in s_2 we are flipping b_2, and so on.

(Refer Slide Time: 13:03)



So, now this gives us a new sequence which we can read off diagonally right. The sequence consists of the red numbers which we have got by flipping the number at the i-th position in the i-th sequence. What can be say about this sequence? Well, first of all it is an infinite 0, 1 sequence.

(Refer Slide Time: 13:20)



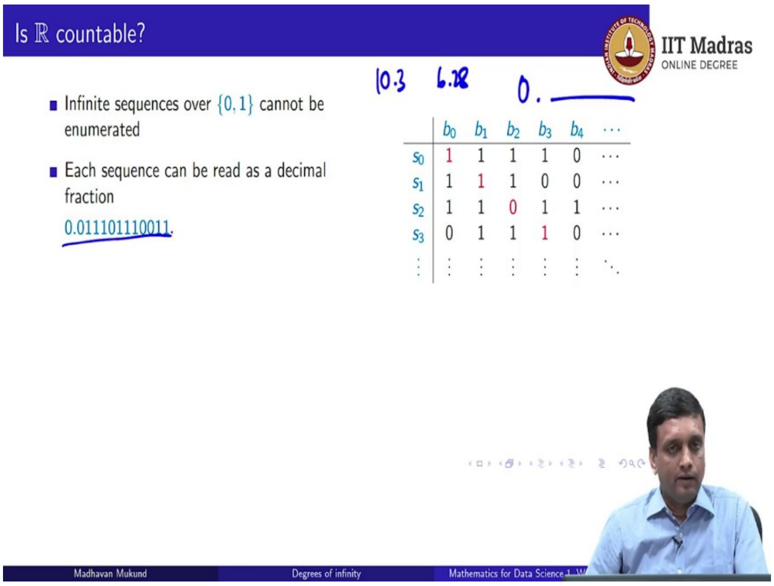
But this infinite 0, 1 sequence cannot be any of the rows in my table, because by construction if it is a row in my table it must be s_j for some j, but at position j, s_j has been flipped. So, this cannot be s_j because if I had s_j already in my table if the sequence is already in my table, the new sequence has the j-th bit flipped. So, diagonal sequence differs from each s_i at b_i, and therefore, this new sequence that I have constructed cannot be part of the enumeration.

Now, it is important that we are shown this regardless of what the enumeration looks like, we have not made any assumption about the order in which we are enumerating. We have said no matter what sequence you have in mind in terms of enumeration, you would have to be able to write down the sequences one after the other table in a sequence of rows.

However you write it down, I will be able to construct this new diagonal sequence by taking the i-th bit in the i-th row and flipping it. So, however you enumerate it, I get a new sequence which is not part of your enumeration. Therefore, there is no possible way of enumerating 0, 1 sequences.

So, as we said this is not the question we asked, the question we asked is are the real numbers enumerable, are real numbers countable? And what we have actually argued is that 0, 1 sequences, infinite 0, 1 sequences are not countable. So, from here how do we get to the real number?

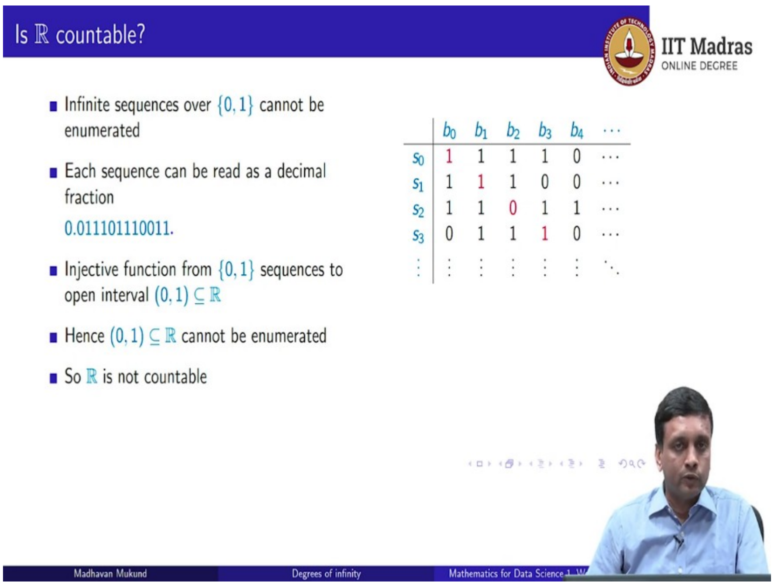
(Refer Slide Time: 14:38)



Well, it is one way to do this is to just think of these 0, 1 sequences as actually decimal fraction. Now, we know that we can write things like 10.3 and 6.28 and so on. So, now, we just restrict our self to writing in decimal fractions of the form 0 point something where everything on the right hand side of the decimal point is either a 0 or 1.

So, here is an example right. So, this is an example of a 0, 1 sequence represented as a decimal fraction. So, since each sequence is different, each such decimal fraction represents a different number.

(Refer Slide Time: 15:20)

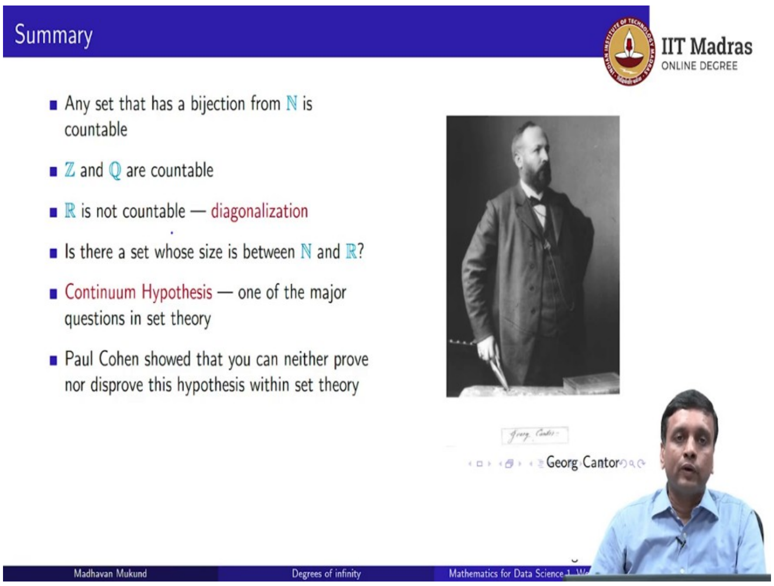


And these are all real numbers between 0 and 1, because they all have an integer part which is 0, and then we have something which is of course, we could have exactly 0 if we have all 0s ok. So, we definitely do not have, all we do not cannot get to 1, but we can think of these as numbers between 0 and 1.

So, each such sequence represents a different point in the interval 0 to 1. So, this is an injective function right. So, this is an injection that is a one-to-one function from infinite sequences 0, 1 to the interval (0, 1). Now, the interval (0, 1) is a very small fraction of the reals.

So, what this argument tells us is that in fact even this very small fraction of the reals is not countable because the set of underlying 0, 1 sequences not countable. So, if this even this small fraction of the reals is cannot be enumerated, then R itself cannot be countable right. So, this is an indirect argument saying that not saying that$R$ itself is not countable directly, but saying that there is a small part of R, which is not countable. And since R is much more than that, if the small part cannot be counted we have no hope of counting the whole thing.

(Refer Slide Time: 16:26)



So, to summarize any set that has a bijection from N is what we call a countable set. And we showed that the set of integers in the set of rationals are countable by describing a strategy to enumerate the sets. Now, this argument is due to Cantor which builds this diagonal sequence called diagonalization and has been used in many other proofs involving infinity after that. So, the proof of diagonalization by Cantor shows that the set of real numbers is not countable.

So, notice that the set of real numbers is not countable and the set of rationals is countable. What it does to the rationals to create the real numbers? We added the irrational numbers. So, actually the set of irrational numbers that we have added to the rationals must be itself uncountable, because we cannot take two countable sets and add them up and get an uncountable set. So, in other words, there are vastly more irrational numbers than there are rational numbers that is what it tells us.

Now, one question that we could ask is, is there anything in between? So, these are sets that we have been using intuitively. So, we have counted them. But can we construct something for instance which is not countable, but which is smaller than the reals right? So, is there such an infinite set?

Now, it turns out that this is a very non-trivial question. This question was actually posed when Cantor came up with this proof in the late 1800s, and it remained a very central opened question it was called the continuum hypothesis.

So, if you look at cardinal numbers in the finite sense, we have 1, 2, 3, 4, 5. So, we have a kind of small jumps between them, but we have a continuous sequence of numbers. Now, we seem to have this big jump between the real number the integers of the natural number and the real numbers, is there something in between or is it so, is there a continuum of these infinite numbers or these big jumps?

And this continuum hypothesis was a very important open question in set theory. And in the 1960s Paul Cohen actually showed that this is a question which cannot be proved or disproved. So, this is what is called independent. So, this is a fact which is independent of set theory using the axiom of the set theory, no way that you can either prove or disprove it.

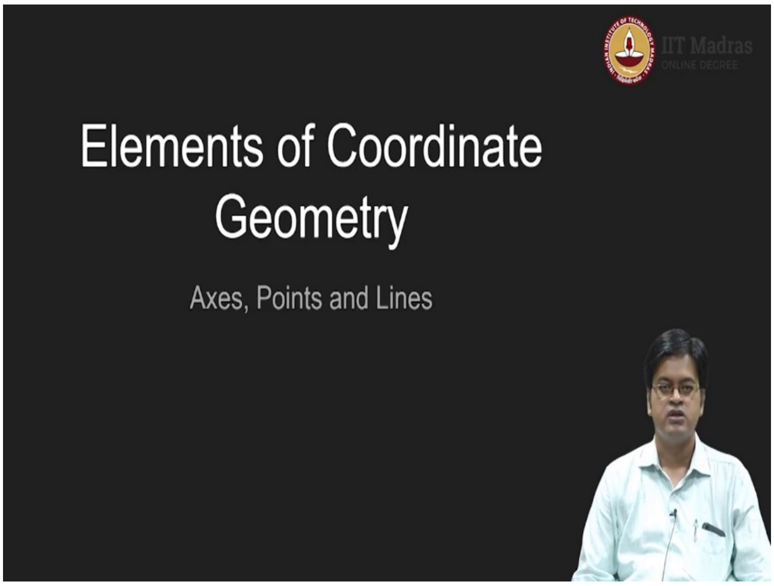
So, both the fact that there is a such a set, and there is not such a set are consistent. So, these infinite sets lead to a lot of interesting questions, some of them are quite mindboggling, and they are quite counterintuitive. But if you are interested in these things, it is well-worth looking into them.

Mathematics for Data Science 1
Prof. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras

Lecture - 12
Rectangular Coordinate system

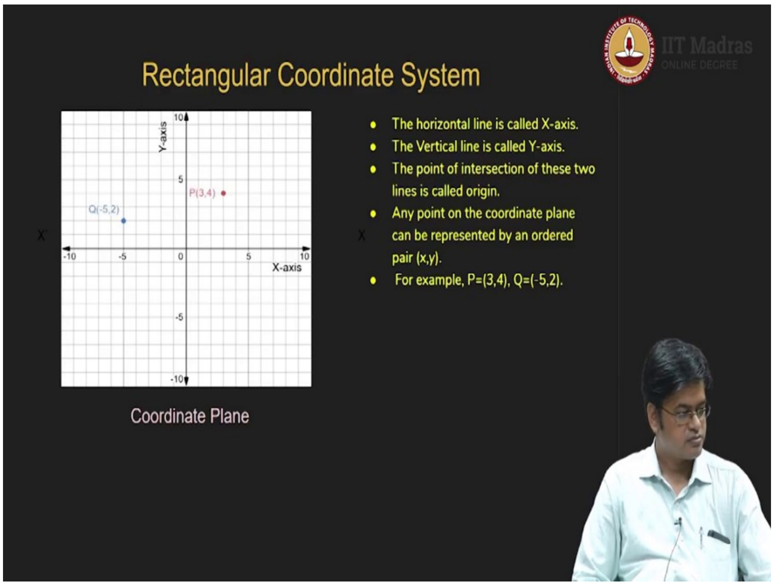
(Refer Slide Time: 00:06)

(Refer Slide Time: 00:14)



So, hello students, today we are going to see some elements of coordinate geometry. Now, let us try to identify these elements as axes, points and lines. We have already seen in basic geometry what are points, lines and planes. So, we will further study this and we will study some algebraic properties using coordinate geometry of these particular geometric objects.

(Refer Slide Time: 00:40)



So, in that context first we need to revise our Rectangular Coordinate System; why is rectangular coordinate system important and how we can study. Given a point on a plane; given a point on a plane you want to describe how this plane be how this point behaves or what is the location of this point. Now, if I want to consider this point and I want to describe the position of the point as of now I cannot say anything more than, this point is slightly towards the right top of the plane.

Now, if I introduce a horizontal line over here, then I can say the point is in the upper half of the plane. This gives a slightly better visibility to the point or slightly better description of a point. Now, if I consider a real number system associated with this line then I can say the point lies in 0 to 5, if I plot two perpendicular lines between 0 to 5 then I will get this point.

This is much better. Now, these perpendicular lines can also be replaced with one perpendicular line which is this which has a real number system associated with it. Now, when a real number system is associated with this point, then what you can actually see is if I can consider this, this particular structure or this particular square which is enclosed within 5 on the vertical line and 5 on the horizontal line; I am giving a much better description of a point.

Then I can enhance this further by putting up the grid lines. These grid lines now typically in this case locate the exact location of the point. So, what is the exact location of the point over here? If you look at this exact location of the point is on the horizontal line if you travel 3 units in one direction, horizontal direction and 4 units in the vertical direction then you will reach this point.

So, I can also name this point as in the horizontal direction I have to travel 3 units and in the vertical direction I have to travel 4 units. So, I can name this point as 3 comma 4 that will be a precise description of this point. So, in turn what we have seen just now is a reference system through which we are able to specify the location of a point in a specific manner. Let us analyze this reference system that we have introduced.

Now, in horizontal direction I have to travel 3 units and in vertical direction I have to travel 4 units; that means, I am actually specifying the coordinates in X direction and coordinates in vertical direction. So, in particular these horizontal directions and vertical directions are called X axis and Y axis respectively.

So, if you look at this horizontal direction, you can see the vertical line cuts the horizontal line into two parts; positive part of X axis and negative part of X axis. Similarly, the vertical line is cut by the horizontal line into two parts. On the upper side we have a positive part of Y axis and on the lower side we have a negative part of Y axis.

So, this is a typical structure which is called coordinate plane ok. Now, let us come to the nomenclature of this particular coordinate plane. As I mentioned if I am travelling 3 units in horizontal direction; I will call that as X coordinate and if I am travelling 4 units in vertical direction, I will call that as Y coordinate. Hence, the name coordinates.

These two lines X axis and Y axis meet each other at a 90 degrees angle; that means, both the lines are perpendicular to each other. Therefore, the name rectangular; recta means right in Latin so, rectangular means 90 degrees coordinate system; that means, a rectangular coordinate system. So, let us revise what we have studied just now in words.

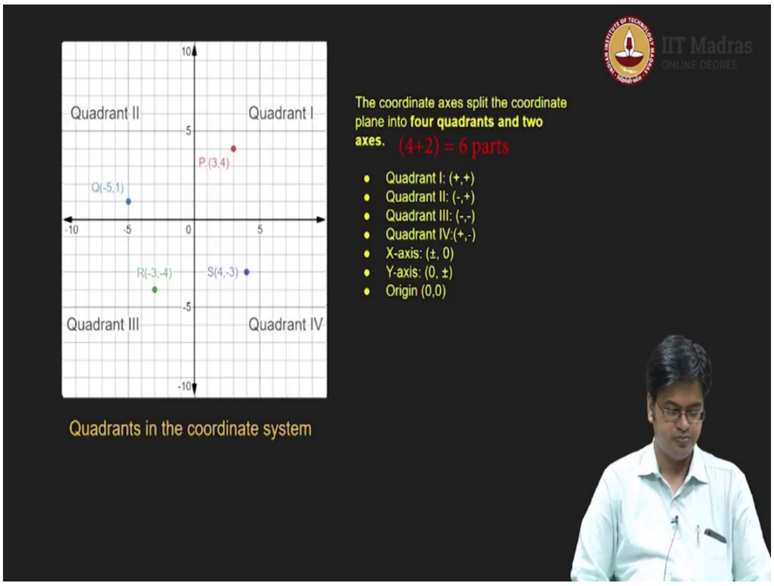
The horizontal line is called X axis, it allows you to move from left to right. The vertical line is called Y axis which allows the movement up and down, then there comes a point of intersection of these two axes which is called origin. The point of intersection of these two axes is called origin and if you look at the coordinates of these, then any point on this particular plane can be denoted by a ordered pair (x, y).

You can see one blue point is also popping up now. Now, how to describe a point using a coordinate plane? So, for example, given a point (3, 4) how will I locate this point? So, if you look at this (3, 4), we have already seen how to locate it. We have travelled 3 units in horizontal direction and 4 units in vertical direction therefore, (3, 4).

Now, suppose you are given another point which is (-5, 2), then this x coordinate corresponding x coordinate is negative; that means, I have to go to the left of the vertical line. That means, I have to travel here a 5 units distance which is - 5 and on the positive side of Y axis I have to travel that is up upper up upper half divided by X axis I have to travel 2 units which will give me the point (-5 ,2).

So, this is how we can uniquely describe points using coordinate plane. Now, when I was when we were studying these two points (3, 4) and (-5, 2), you can easily see with respect to this coordinate axes you can have 4 parts of the coordinate plane.

(Refer Slide Time: 07:24)



Let us study those parts in detail in the next slide. So, next slide is this coordinate plane. Now, I have identified 4 points in all 4 parts of the coordinate plane. So, you can see the first point P which lies in the positive side of X axis and positive side of Y axis has positive x and y coordinates which is given by quadrant I. So, any point in this plane, in this particular quarter will have positive X and positive Y axis.

Now, in general as a mathematical psychology we move in a anti-clockwise direction. So, now, I can move in a anti-clockwise direction to the next one fourth part, next quarter of the coordinate plane. And, see that my X axis has negative values and my Y axis has positive values. All points which have this form of values are called points on the second quadrant or the quadrant the one the quarter of this particular coordinate plane is called quadrant II.

Next we come in a anti-clockwise direction to the third side that is this. So, if you look at the point R which is lies in this particular quadrant is (-3, -4); that means, the x value is negative and the y value is negative. Therefore, (-3,-4) is a point which lies in quadrant III.

Remember it is easy to remember this that quadrant I and quadrant II, quadrant III that is odd quadrants have same parity of x and y coordinates. And, quadrant II and quadrant IV have opposite parity of x and y coordinates. So, let us go to quadrant IV, you can see a point S lies in quadrant IV which has coordinates 4 and -3. Now, this 4 and -3 which denotes x coordinate is positive and y coordinate is negative such a classification comes in quadrant IV.

So, this is how a coordinate plane is come split into four quadrants. Now, a question may arise in your mind; suppose I have this point which is (5, 0). Now, in which quadrant this point lie? The answer is this point does not lie in any of the quadrants. This point lies on the X axis. Similar question can be asked for a point (0, 5). The point does not lie on any quadrant, but lies on the Y axis.

So, based on this particular understanding, a coordinate plane its subdivided into first is four quadrants, two are axes. Let us try to see what are the typical features of the quadrants and these axes. Quadrant I, you will have x and y coordinates which are positive. Quadrant II, you will have x coordinate which is negative y coordinate which is positive. Quadrant III, you will have both negative values. Remember odd quadrants will have same parity that is quadrant I is positive, quadrant III is negative.

Now, quadrant IV will have positive and negative, x coordinate which is positive, y coordinate which is negative. Then comes the split into axes. So, on the X axis you will have points which can either take positive values or negative values for x coordinates and 0 for y coordinate. On the Y axis you will have points which can take positive and negative values for y coordinates, but 0 for x coordinate. Now, there remains only one point which is the point of intersection which is identified as origin ok.

So, this completes our understanding of the coordinate system. Why quadrants, quadrant system is helpful? Sometimes you have been given several points to plot. Now, those points if you look at them closely, you need not have to divide the system in a equally distance manner, like this manner. You may have many points in quadrant I, in that case you can scale this, you can bring this to the bottom right corner, bottom left corner and just focus on quadrant I.

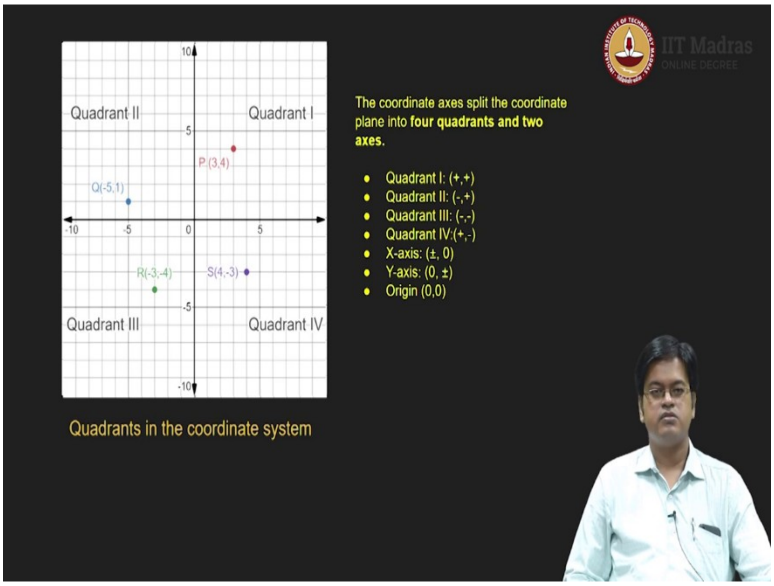
So, if you have a good understanding of quadrants, you may be able to graph the functions better, graph the points better; that is why the coordinate system is important. This ends our discussion on coordinate system.

Mathematics for Data Science-1
Prof. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras

Lecture - 13
Distance formula

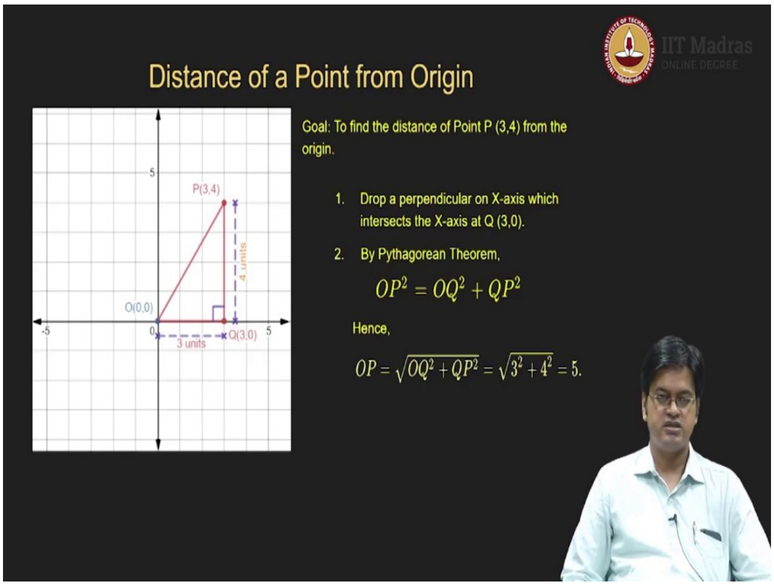
(Refer Slide Time: 00:05)

(Refer Slide Time: 00:14)



So, after coordinate system, let us try to identify one classical problem.

(Refer Slide Time: 00:20)



That is if I have a point and somebody ask me a point is located here; let us say point is (3,4). And somebody ask me what is the distance of this point from the origin? So, in this particular slide our goal is to find the distance of a point P which is (3,4) from the origin.

That essentially reduces to finding the length of this line segment which is joining points O and P. So, is there any classical tool that is of my help? Suppose, now if this point is either lying on X axis or Y axis, let us say if this point is say (3,0) ok. If this is the point that is of interest to me; do I know how to find the distance of this point? The answer is yes I know, I just need to calculate the units that are in horizontal direction.

Suppose the point is on Y axis, then do I know how to calculate the distance of this particular point from Y axis? The answer is again yes I know, I just need to calculate the number of units that I need to travel to reach this point. So, if the point lies on X axis and Y axis, I know how to calculate the distance of a point. Now, if the point is lying anywhere in the coordinate plane, how to find a distance is a question.

For that, let us try to understand the situation, that if I know if somehow I can understand this with respect to this coordinate axis. This particular position with respect to these coordinate axis then I will be able to give the answer to find the distance between the two points. So, let us try to do one thing that is let us try to get the image of this point (3, 4) on X axis.

So, how will I get the image of this point (3,4) onto the X axis? The easiest way is you drop a perpendicular on X axis, that intersects the X axis at point (3,0). Once this is done then you can actually drop a perpendicular and see that it forms a right angled triangle with X axis in place and a vertical line in place; you have a right angled triangle. Do you know any theorem in our conventional geometry that relates this particular structure?

You know Pythagoras Theorem or Pythagorean Theorem, that relates this particular structure. In a right angled triangle the hypotenuse length of the hypotenuse is given by square root of its adjacent sides; square root of squares of the lengths of the adjacent sides. So, we will try to use this for finding the distance of a point from the origin. So, by the Pythagorean Theorem, I know OP^2 is actually equal to OQ^2+QP^2.

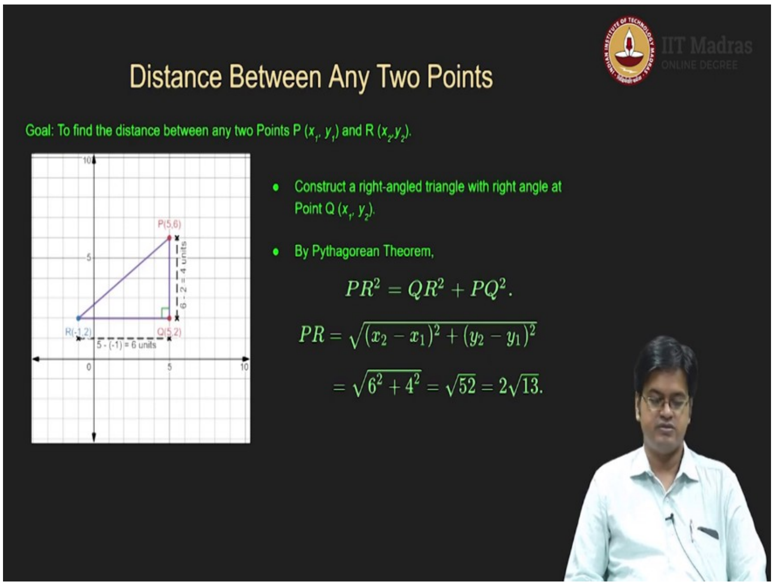
Now, the exercise that we did orally just before starting this problem will help us to understand what is OQ^2. So, what is OQ? OQ is a part of X axis, OQ is a line segment which is a part of X axis. What is the length of OQ? We have already discussed that, that length is 3 units. Similarly, if you look at QP; what is QP? QP is parallel to Y axis. So, it is as good as projection of Y axis projection onto Y axis.

So, what is the length of this particular line segment which is QP? That is 4 units; so I know the length of OQ and I know the length of QP. Therefore, by Pythagorean Theorem, I know the length of OP. So, what will be the length of OP? It will be $\sqrt{\square}$. So, 3^2 is 9, 4^2 is 16 therefore, this will give me 25; 16+9 and positive square root of it will give me number 5.

Now, has it anything special to do with point (3,4) or can I generalize this? The answer is yes, it has nothing special to do with point (3,4). I could have started with point P which is (x , y) and then projected this onto X axis or I figured out the image onto X axis which will be (x , 0). And therefore, the length of OQ will be x and the length of QP will be from 0 to y units; that means, y units.

So, length of QP will be y units and therefore, the formula OP = $\sqrt{\square}$ would have been possible. So, let us try to take this particular example and try to generalize this problem to finding the distance between any two points.

(Refer Slide Time: 06:02)



So, distance between any two points. So, again the setup is pretty common. Our goal is to find the distance between any two points P (x¿¿1, y_1)¿ and R (x¿¿2, y_2)¿. How will you find the distance between any two points? Let us see the points on the graph, then the things will be more specific. My (x¿¿1, y_1)¿ is (5, 6) and (x¿¿2, y_2)¿ is (-1 , 2).

Now, if I look at these two points, I want to find the distance between these two points. So, once easy way to find a distance between these two points is to construct a right angle triangle. But, now because the point is not located on X axis, this (-1 , 2) is not located on any of the axis; I cannot say drop a perpendicular to X axis.

So, the actual way that I should do here is I will drop a perpendicular to X axis which will intersect at (5 , 0). And, then to this line I will drop a perpendicular from the point R minus (1,2) and which will intersect this, this particular line which will be the perpendicular to X axis, at where the y coordinate will be 2 and x coordinate will be 5. So, this point will be (5,2) and then I will get a right angled triangle.

By skipping these steps, we can straight away say that you construct a right angled triangle with a right angle at point Q which is (x¿¿1, y_2)¿. Just relate this(x¿¿1, y_2)¿, if you use this terminology is the point (5,2). So, you can draw a right angle triangle using (5,2) ok.

So, this way we need not have to specify steps that you have to draw two perpendiculars and all; because the point may as well lie in the third quadrant. And, in that case dropping perpendicular to X axis may not help, you have to extend the perpendicular beyond X axis. So, it is always better to consider this kind of structure, that is construct a right angled triangle with right angle at point Q which is $(x$¿¿$1, y_2)$¿ .

Then it does not matter where the point actually lies. Now, once the right angle triangle is in place, the same theory that we used Pythagorean Theorem will come into play. And, by Pythagorean Theorem if I want to find the length of PR, I know $PR^2 = QR^2 + PQ^2$. Can I calculate the length of QR and PQ? The answer is yes I can calculate, because, the line segment QR is actually parallel to X axis and the line segment PQ is parallel to Y axis.

Therefore, this is as good as computing the length on X axis and this is as good as computing the length on Y axis; hence what we will get is. So, how to compute the length? It is basically the change in x coordinates. So, how far the x coordinates have changed? So, while computing the length parallel to X axis always remember you should go from left to right, that is when you are subtracting you should take the highest value first that is 5 – (-1).

So, the length of this will be 6 units and while subtracting or while finding the length in a vertical direction go from bottom to up. That means, you subtract the value that is highest in Y direction to the value that is lowest in Y direction. So, here 6 - 2 will give me 4 and in the X direction 5 – (-1) will give me 6 units. So, this is how we will calculate the length of these two line segments.

And therefore, I can easily find the length of PR; while calculating the length because we are in this particular case, we are considering squares. It does not matter whether you consider x_1 first or $x_2$first because, anyway we are squaring even if you get the negative value, you will be squaring it.

So, in particular in this case where the coordinates are $(x$¿¿$1, y_1)$¿and $(x$¿¿$2, y_2)$¿, I will take $(x$¿¿$2 - x_1)^2$¿; does not matter which one is bigger. And $(y$¿¿$2 - y_1)^2$¿ , does not matter which one is bigger.

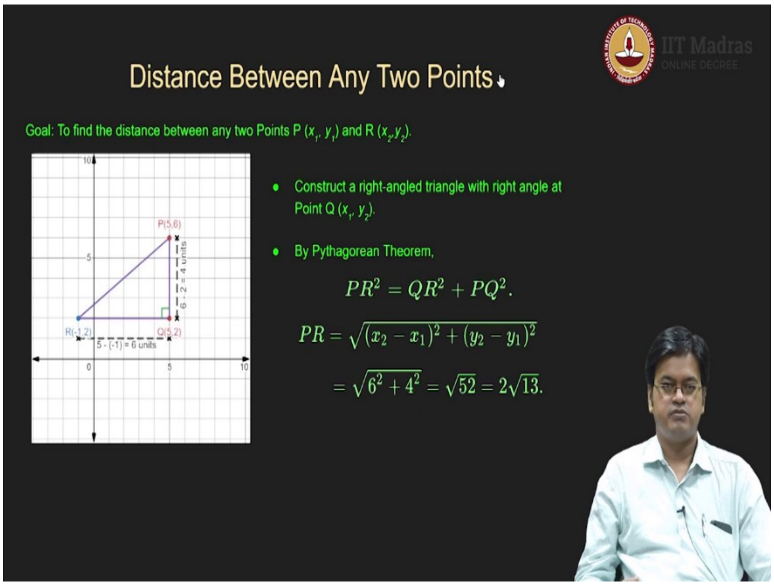
And I will take a positive square root of it. Therefore, my length PR for this particular example will be 6^2+ 4^2; $6^2 =$¿ 36, $4^2 =$¿ 16 together they will give 52 is 2√❑. So now, we have established a general formula which is called distance formula for finding the distance between any two points on a coordinate plane.

Mathematics for Data Science 1
Prof. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras

Lecture - 14
Section formula

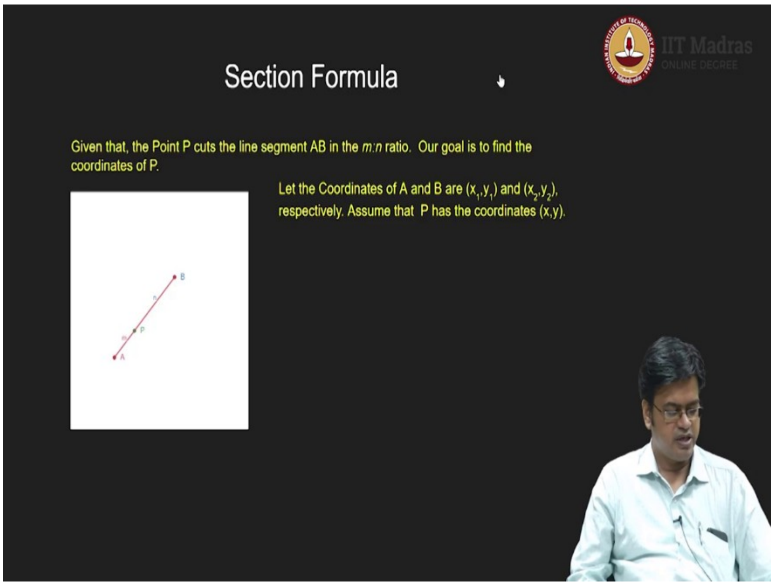
(Refer Slide Time: 00:06)

(Refer Slide Time: 00:14)



Now, let us take up the next concept. Now, we have handled two points; now let us take three points.

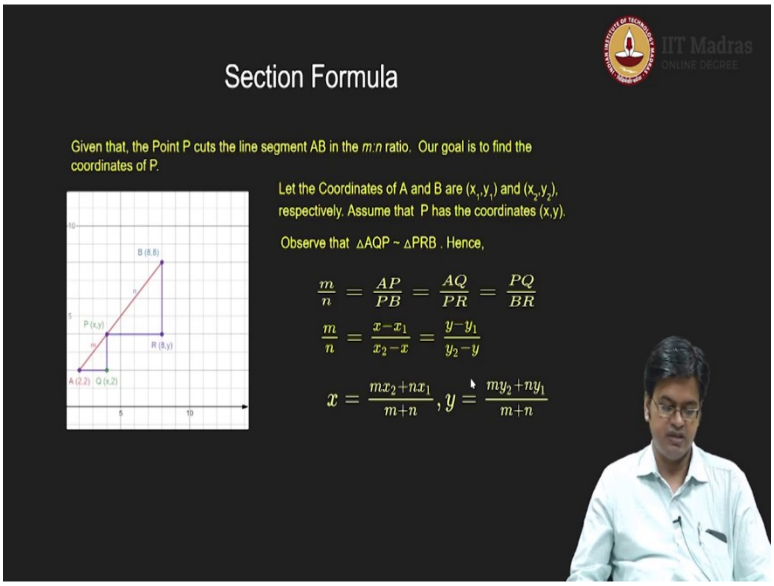
(Refer Slide Time: 00:22)



And, let us say those three points lie on a line and given that the point P cuts the line segment AB in the ratio m : n. Our goal is to know the coordinates of point P; this will give us the Section Formula. So, this is the graphical representation of the points. So, there are two, there is a line segment AB and point P cuts this line segment in the ratio m : n.

How will you find the coordinates of point P? This is the question; let us bring in our coordinate system. So, let the coordinates of A and B are ¿, y_1¿ and ¿, y_2¿, the coordinates of A are ¿, y_1¿ coordinates of B are¿, y_2¿,. I do not know what P is, let us assume it has some coordinates which are x and y ok. So, let us bring in them in the coordinate system which is this.

(Refer Slide Time: 01:29)



Let us try to understand this particular coordinate system by putting up some triangles around. So, what I have done is I have actually constructed two triangles using the same logic that I used in the distance formula. If the coordinates of point P are (x,y), then I will construct a right angled triangle in this direction; where the x coordinate will be x and y coordinate will be the coordinate of y coordinate of A that is y_1.

Similarly, I will do the same thing with respect to point B; I will drop a perpendicular which will meet at this particular point. So, basically I will drop a perpendicular which will meet the X axis and again I will draw a perpendicular here. But, let us for sake of simplicity we have constructed a right angle triangle, where the y coordinate of this point will be y and the x coordinate of this point will be the x coordinate of point B which will be x_2, (x_2, y).

With this understanding we can proceed further and see that the triangles, these two triangles are similar to each other. How? First of all let us see this line is parallel to X axis and this line is parallel to X axis as well. Therefore, these two are parallel lines and this is a transversal that is passing through these two parallel lines. Therefore, these two angles the angle A and angle P will be same or equal.

Next these two are right angles, then we know the sum of the angles in a triangle is 180 degrees, therefore this angle, angle B must be equal to angle P. Therefore, triangle AQP must be similar to triangle PRB by angle test that essentially means I have their sides in some ratio, correct. So, for simplicity I have plotted these points with some coordinate references.

So, this is A is (2,2), B is (8,8); then whatever I mentioned the coordinates of Q are (x,2) and coordinates of R are (8,y). So, now these two things will be in some ratio that is $\frac{AP}{PB}$, these are the hypotenuse of these two right angle triangles is equal to $\frac{AQ}{PR}$and this thing is $\frac{QP}{RB}$ right or you can see $\frac{AP}{PB}$ is equal to $\frac{AQ}{PR}$which is equal to $\frac{PQ}{BR}$.

Now, I already know $\frac{AP}{PB}$ have a ratio m : n. So, their ratio is m by n that is already known to us, that is given to me. Now, can I calculate the length of AQ and PR? The answer is yes, because AQ is parallel to x axis. It is just subtracting the highest x coordinate from the low.

So, it will be x -2 in the figure and in our theory it is x - x_1. Similarly, you can look at PR; it will be 8 - x or in our theory it will be x_2- x. For y axis or the lines that are parallel to y axis PQ and BR you can see you will go to the highest value that is y - 2 or y - $y_1$and the other one BR will have $y_2 - y$right.

So, together I will have a representation of this form: $\frac{m}{n} = \frac{x - x_1}{x_2 - x} = \frac{y - y_1}{y_2 - y}$. Now, take one equality at a time; that means, $\frac{m}{n}$is equal to let us say these the consider these x coordinates. So, we will just cross multiply them, rearrange them you will get what x is equal to.

In a similar manner just take $\frac{m}{n}$ is equal to this, these y coordinates ratio and then cross multiply and rearrange them. You will get the following values which are given by $x=\frac{m x_2+n x_1}{m+n}$. And, similarly $y=\frac{m y_2+n y_1}{m+n}$. This gives me the section formula, when a point divides the line in the ratio m : n.

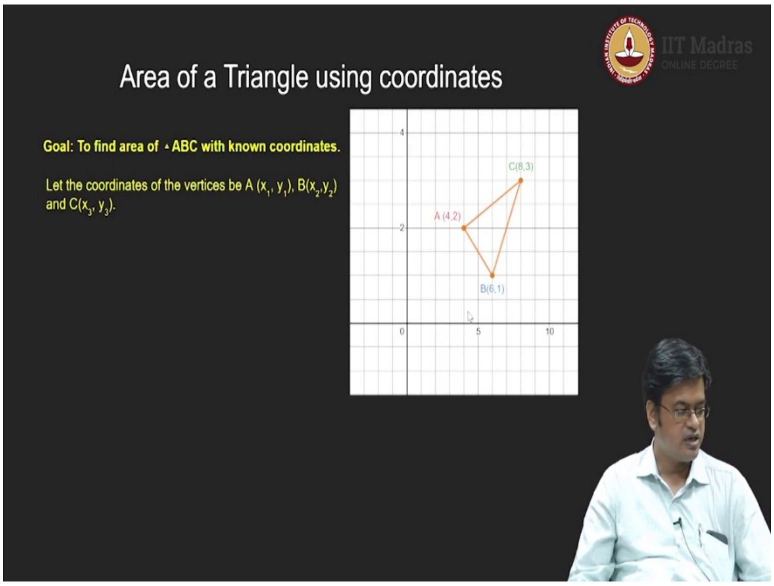
Another interesting question is suppose I know the coordinates of x and y; can I find in what ratio the line divides? Obviously, yes because you know the coordinates of the line, you just need to use this formula for finding the ratio ok; that will be more clear when you solve more problems ok.

Mathematics for Data Science 1
Prof. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras

Lecture – 15
Area of triangle

(Refer Slide Time: 00:05)

(Refer Slide Time: 00:15)

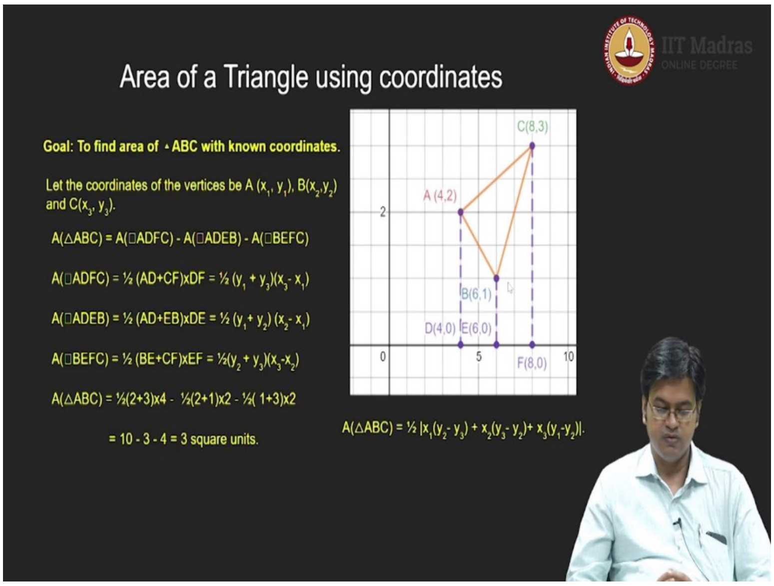


After section formula let us try to understand the three points when they are not on one line that is when they form a triangle. So, you have been given three points, and you know they are not collinear points and therefore, they will form a triangle. And the question can be how to find the area of triangle using the coordinate system.

So, let us try to see that using the coordinate system. So, there is some triangle ABC and I want to find the area of triangle ABC. Let the coordinates of that triangle be (x_1, y_1),(x_2, y_2) and (x_3, y_3). Once I have these coordinates, I can plot it here. You can see on the right there is an image of a triangle.

Now, how to find the area of this triangle? Now, whatever I discussed so far everything actually relied on dropping a perpendicular to X - axis and finding the area of the geometric object that is formed. In earlier cases, it was just a triangle. Now, if we follow that theory then you can easily see that I need to do something like dropping a perpendicular to x axis. So, I have dropped perpendiculars to X - axis.

(Refer Slide Time: 01:35)



Now, I have generated some figures. What are the figures that I have generated? In particular, I have generated 3 trapeziums, trapezium ADFC that is the biggest one encompassing everything. Then, you can look at trapezium ADEB, then you can look at the trapezium BEFC.

Now, my triangle is trapped in between these trapeziums. So, let us try to make our understanding crystal clear. If I want to find the area of triangle ABC, then I need to first consider the biggest possible quadrilateral or trapezium that is ADFC and eliminate the areas of two smaller trapeziums that is ADEB and BEFC. And whatever I am left with is the area of triangle ABC.

Now, do I know how to find the area of trapezium? Yes, I know. The formula is half times sum of parallel sides into the height of the trapezium. So, we need to quantify how will these quantities be calculated? Let us consider trapezium ADFC, if I consider a trapezium the ADFC then what are the parallel sides of this trapezium? Side AD and side FC.

So, I will take average of these two parallel sides that is half of AD plus FC. Then, what is a height? Height should have a perpendicular distance, so that is X - axis. So, I know the distance will be DF.

So, let us take the general coordinate system rather than using this coordinate system. What are the coordinates of A and D? So, A has coordinates $(x_1, y_1)$and after dropping a perpendicular on X – axis the y coordinate will vanish and therefore, the coordinate of D will be $(x_1, 0)$. So, what will be the length of AD? It will be purely in terms of y that is y_1.

Similar, thing is applicable for CF. So, it will be nothing but y_3; so, area of ADFC,

$$Area(ADFC) = \frac{1}{2}(AD+FC) \times DF = \frac{1}{2}(y_1+y_3) \times DF \cdot$$

Now, what is the length of line segment DF or FD? Highest minus the lowest. So, in this case our F is $(8, 0)$ or $(x_3, 0)$and the point D is $(x_1, 0)$. So, it is $(x_3 - x_1)$.

$$Area(ADFC) = \frac{1}{2}(AD+FC) \times DF = \frac{1}{2}(y_1+y_3) \times (x_3 - x_1) \cdot$$

In a similar manner, I can actually see a smaller trapezium that is ADEB, smaller quadrilateral that is ADEB and the height of that quadrilateral will be the length of ED which is 2 in this case or $x_2 - x_1$in the coordinate system. So, this is what our understanding of the length is. In a similar manner, the sum of lengths of parallel sides is $y_1 + y_2$.

$$Area(ADEB) = \frac{1}{2}(AD+EB) \times DE = \frac{1}{2}(y_1+y_2) \times (x_2 - x_1) \cdot$$

In a similar manner you can compute BEFC.

$$Area(BEFC)=\frac{1}{2}(BE+CF)\times EF=\frac{1}{2}(y_2+y_3)\times(x_3-x_2)\ .$$

Now, using this you can compute the area of the triangle which can be easily seen to be in this form. So, I have just taken this example and computed these values. So, the values are effectively in this particular case the length of AD was 2.

The length of CF was 3 so

$$Area(abc)=\frac{1}{2}(2+3)\times 4-\frac{1}{2}(2+1)\times 2-\frac{1}{2}(1+3)\times 2=10-3-4=3\ square\,units.$$

Now, if you look at this particular thing and rewrite this expression you will get a very nice expression. You can juggle with this expression and try to simplify it by taking a cross products and you will come up with the expression of this form.

$$A(\Delta ABC)=\frac{1}{2}\vee x_1(y_2-y_3)+x_2(y_3-y_1)+x_3(y_1-y_2)\vee ¿$$

The absolute sign is just to ensure that the area value should not be negative, but the calculation still remains same. And you need to consider one caution here that all the vertices of a triangle in an anticlockwise direction then only this formula is valid. So, I have considered area of a triangle.

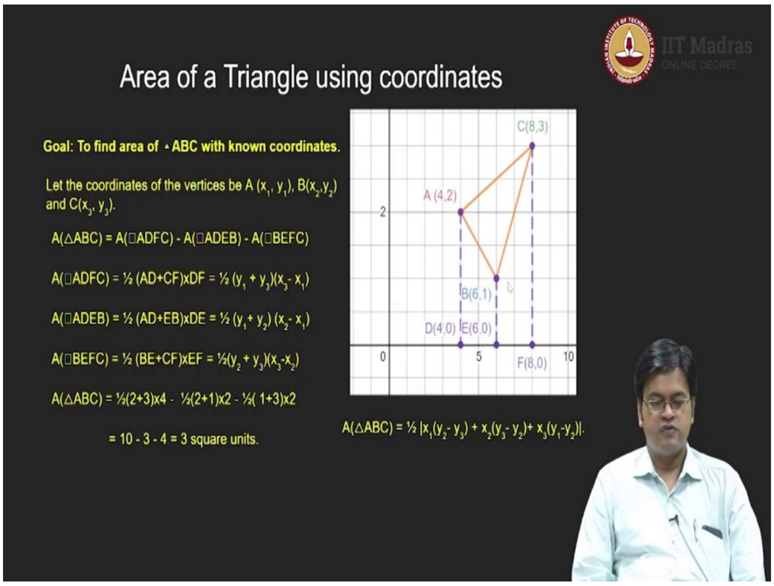
Now what we have seen so far is given two points how to find the distance between two points, given three points if they are collinear, we have found the section formula that can help us to find their ratios or the coordinates of the middle point. Now, if the points are non-collinear, we have seen how to compute the area of the triangle using the coordinate system.

Mathematics for Data Science 1
Prof. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras

Lecture – 16
Slope of a Line

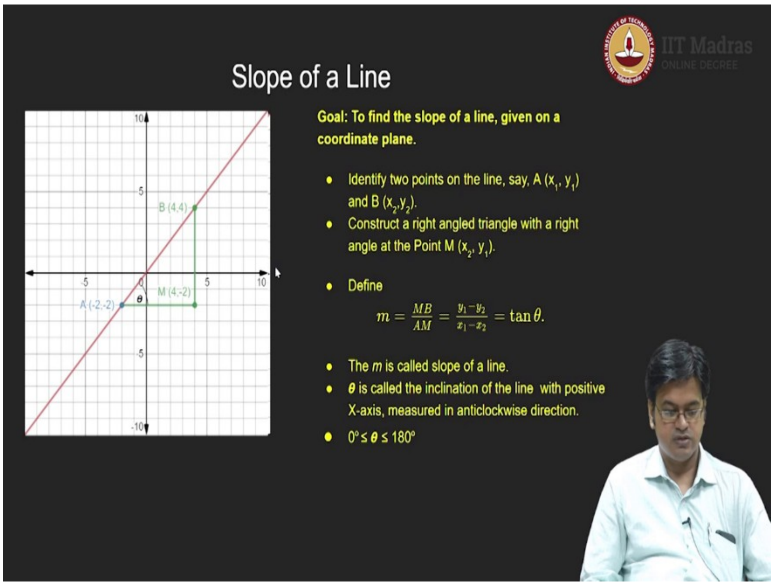
(Refer Slide Time: 00:05)

(Refer Slide Time: 00:14)



So, after looking at area of the triangle using coordinates, let us now focus our attention to again a two-point system and one-dimensional objects that is a line.

(Refer Slide Time: 00:24)



We have already seen in our basic classes that two points uniquely determine line. Now, if I want to characterize a line, and if I give you two points, I should be able to find a line passing through these two points. How is the geometric object algebraically related to the coordinate geometry? That is what we want to explore now, to explore that I need a concept of a slope of a line.

So, what essentially is the slope of a line? In a vague manner, what we understand by slope of a line? If you look at this coordinate plane which is displayed here. If I am moving some units in x directions; the question can be asked with respect to this change in x direction what is the corresponding change in y direction.

So, if I want to answer that question then I need to consider a ratio of change in y direction to change in x direction; some people call it as rise by run ratio, run is in the horizontal direction, rise is in the vertical direction. So, you can consider slope of a line as a rise by run ratio. So, let us try to make this work concept clearer by showing some examples.

So, now, here is a line with two points given onto it. Again, our standard conventional method we will construct a right-angled triangle using these two points. Now, the question that I posed is what is a rise by a run can be answered over here. For example, you look at this right-angled triangle, what is happening? This is the movement of a line in moving from one point to another point in y direction, this vertical length is the direction, is the movement of a line in moving from one point to other point in y direction and this horizontal line is a movement in x direction while moving from point A to B on a line.

So, essentially what I need to capture is the change in y direction that is from point $(4,-2)$ to point $(4,4)$, that is -6 and moving in x direction from $(-2,-2)$ to $(4,-2)$ that means, -6 here also. So, the slope of a line can be equal to 1. This we can make it more precise by giving some formal definitions.

So, if I want to find the slope of a line given the coordinate plane, I can always identify these two points as (x_1, y_1) and (x_2, y_2). I will construct a right-angled triangle which intersects the point at(x_2, y_1). And once I constructed, as I mentioned you know what is the change in x direction and what is the corresponding change in y direction, therefore, you can actually compute the ratio of this. But while computing the ratio, you can also think remember some concept from trigonometry.

For example, when I constructed this right-angle triangle there is some angle formed over here this arc denotes that angle. Let us call that angle as theta. Now, what I am saying is change in y upon change in x, but can you relate some quantity related to this trigonometric ratio that is tan of θ, right. So, what I can say is my m or the slope of a line is MB by AM which is $\frac{y_1 - y_2}{x_1 - x_2}$and which is also equal to$\tan\theta$.

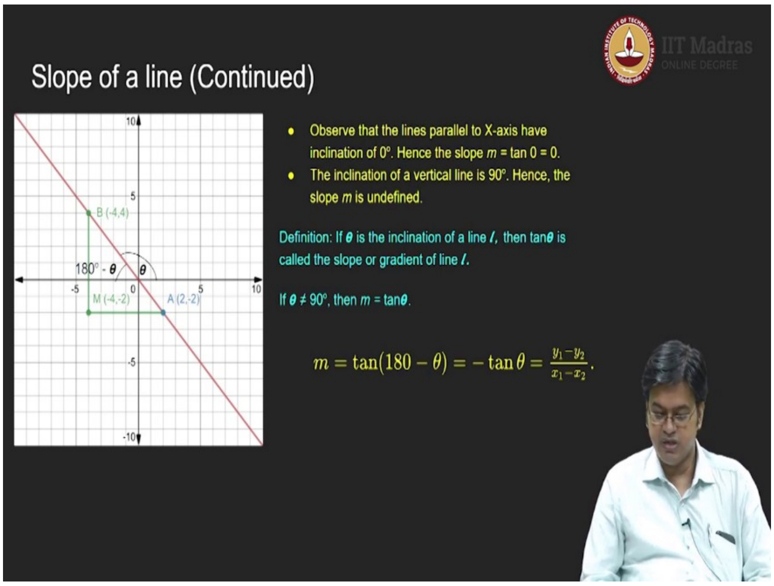
So, I have defined one thing that is m which is the ratio of these two, but which in turn turned out to be equal to$\tan\theta$. So, if it is $\tan\theta$, see here it does not matter whether I take $y_1 - y_2 \vee y_2 - y_1$whatever I am doing I should do synonymously. For example, if I have taken $y_2 - y_1$then I should take $x_2 - x_1$or if I have taken $y_1 - y_2$ then I should take $x_1 - x_2$.

So, it does not matter which order you are swapping because finally you are taking the ratio so whatever you are doing you do it asynchronously, so that there will not be any confusion. So,$m = \tan\theta$. Now, I have introduced two terminologies here m and θ. So, let us define them properly. This m is called slope of a line, which is the topic of this discussion. And then this θ is called inclination of a line with respect to positive X - axis measured in an anti-clockwise direction.

Now, somebody may say I have drawn this angle over here, but if you look at this particular line, this line is parallel to X - axis. And this line is intersecting X - axis here, that means even if I consider this angle, this angle also will be $\theta$from the basics of geometry, correct.

So, now the question can be asked how far the θ can go? So, to answer that question let us try to see if I am considering a θ then θ can be equal to 0, θ equal to 90 degrees tan is not defined. As you can see tan of 90 is not defined, but it can go up to 180 degrees. So, the variation of θ allowed is 0 to 180 degrees.

(Refer Slide Time: 07:42)



So, now let us have a look at the salient features of the slope of a line. In particular, let us see if the line is parallel to X - axis the angle of inclination is 0 degrees; therefore, the slope of a line should be 0. Now, if the angle is 90 degrees; that means, 90 degrees with respect to X - axis; that means, eventually I am on Y - axis or in fact, I am on Y - axis in such case tan 90 is undefined, right. Therefore, slope is undefined.

As you can see if I have an angle which is 90 degrees that is Y - axis; that means, x is equal to constant is the equation of the line. And you cannot have any movement in y direction or you can have infinite movement in y direction without any change in x direction. That itself creates a problem therefore, the slope is undefined for theta is equal to 90 degrees or the inclination is equal to 90 degrees.

So, with respect to inclination there is another definition of slope. If theta is the inclination of a line l then tan theta is called slope or gradient of the line. This is the second definition of our slope of a line which matches exactly with the original definition, but there will be some glitch, there may be some confusion, ambiguity.

So, let us try to resolve that ambiguity because this theta is the angle made with respect to positive X - axis. And theta not equal to 90 degrees I can define $m=\tan\theta$. That is perfectly fine and it is well-defined over there whenever it is not equal to 90 degrees. What is the ambiguity? The ambiguity can be shown in the figure. For example, now what is θover here? θ over here is actually this particular angle.

Now, if I you look at this particular angle which is θ you can see that this is an obtuse angle. Now, how to evaluate a tan of this angle? We already know some methods, but will that contradict with our definition of slope. That is the question. So, if I use the rise by run formula or the change in y to up on change in x formula, how will I figure out the slope? So, the answer is I will simply drop a perpendicular or I will construct a right-angle triangle with right angle at point M which is $(-4,-2)$.

In that case, I will be interested in this angle that is angle at A in our older definition or this angle is essentially equal to $180-\theta$. So, let us go further. This angle is equal to this angle. What is the measurement of this angle? It is$180-\theta$. That means, if I want to find a slope according to our definition that is $\frac{\delta y}{\delta x}$or change in y by change in x, then I need to consider the angle of this particular structure that is $\tan(180-\theta)$. So, $m=\tan(180-\theta)$.

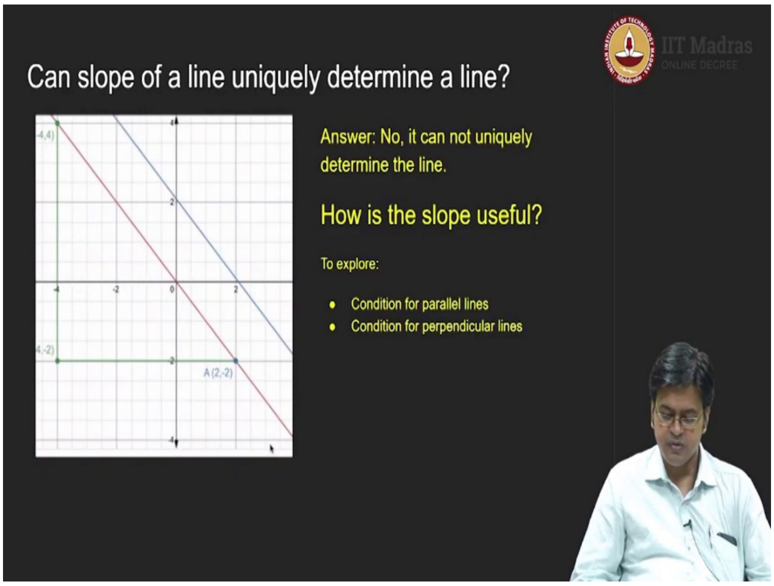
Now, what is $\tan(180-\theta)$? If you use simple trigonometric formula you will get $\tan(180-\theta)$ is nothing, but $-\tan\theta$. But what is $-\tan\theta$? You can easily see what is $-\tan\theta$which will be $\frac{y_1-y_2}{x_1-x_2}$. So, in short, our formula for slope is consistent no matter which definition we use, therefore a slope of a line is uniquely determined given a line.

Mathematics for Data Science 1
Prof. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras

Lecture – 17
Parallel and perpendicular lines

(Refer Slide Time: 00:05)

(Refer Slide Time: 00:15)



Now the question can be asked that if a line is given to me, I can uniquely determine the slope, but if a slope is given to me can I uniquely determine a line? That is the next question that I will put up. In any sense the question asks can there be many lines with same slope? The answer can be seen in this GIF image.

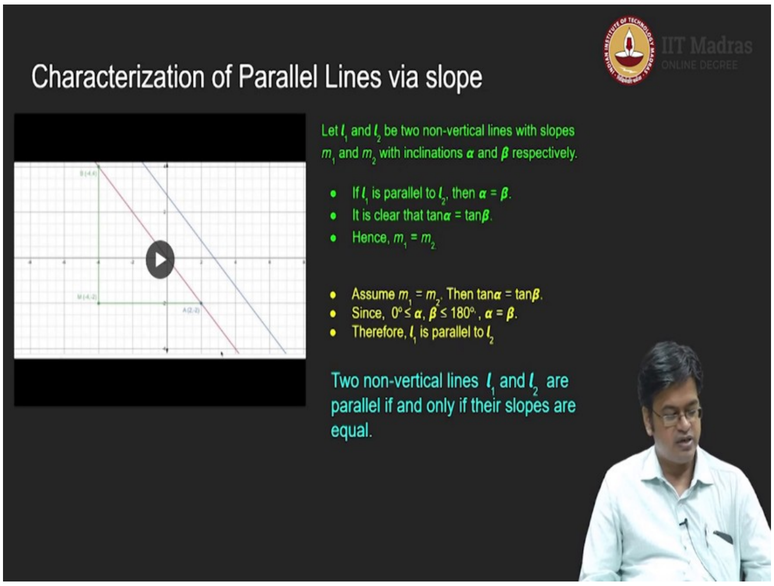
If you look at this image closely what we have done is? We have fixed one line and we know how to compute the slope of this line we have a it will be minus 1 based on the coordinates. Now, the blue line that is revolving around is actually having the same inclination as the orange line.

Now, the orange line and blue line have the same inclination; that means, tan of those inclinations will be same, will match and hence there can be infinitely many parallel lines which have a same slope. So, the answer to this question, can slope of a line uniquely determine a line? The answer is no, you cannot uniquely determine a line given the slope of a line or the inclination of the line.

Now, why do we study the concept of slope or whatever we studied how it is helpful? The helpfulness of this concept is just what we discussed in this graphical image, what we are seeing is if the inclinations are same the line better be parallel. So, for parallel lines I can use this concept and derive a condition of slope. Similarly, I can do by rotating them by 90 degrees; that means, I we can consider the perpendicular lines and I can consider general two lines intersecting each other and see what condition I can derive based on the slope.

So, I want to explore the usefulness of slope. So, to explore this I will first figure out the condition for parallel lines and I will figure out the condition for perpendicular lines, in due course we will find the relation between slopes of two lines and their intersection and their angles of intersection. This is what we will do in next few minutes.

(Refer Slide Time: 03:15)



So, let us go to the next characterization of parallel lines via slope. Now as you can see in this image there are two parallel lines, they have same inclination, but they are not unique that is what we figured out. So, if I play this video you can see again, this is similar to what we have seen in the last video.

So, I have something which is moving around and there can be infinitely many lines, what remains constant is the inclination, the inclination is same if I have parallel lines. So, let us try to see whether we you can derive something. So, let to put it in a proper context.

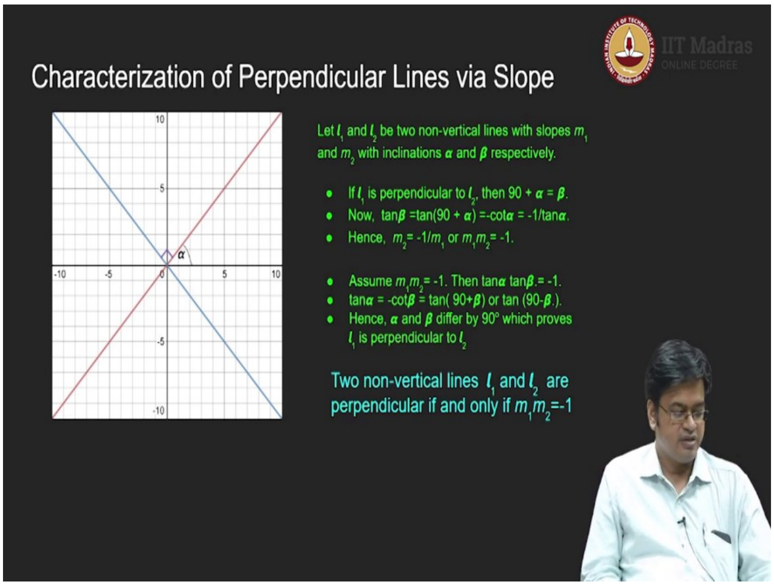
Let orange line be $l_1$and the blue line be l_2 be two non-vertical lines. Why non-vertical lines? Vertical lines have angle of 90 degrees for which the concept of slope is undefined, inclination 90 degrees for which the concept of slope is undefined. So, what I need is non-vertical lines. So, considered two non-vertical lines with slopes $m_1 \wedge m_2$given the slopes their inclinations $\alpha$and βrespectively.

Now, if you have been given that $l_1$is parallel to $l_2$then $\alpha = \beta$, inclinations are same that is what we have seen in the figure and that is what we discussed in the last slide also. So, if $\alpha = \beta$ then naturally $\tan\alpha = \tan\beta$, once $\tan\alpha = \tan\beta$; what is $\tan\alpha$? It is the slope of line1_1 that is m_1 and $\tan\beta$is the slope of line $l_2$which is m_2. Therefore, clearly the slopes are equal, $m_1 = m_2$.

The converse that is assumed that, if the slopes are equal then $\tan\alpha = \tan\beta$ by a definition. Now, $\tan\alpha = \tan\beta$ does that imply α is equal to β? In our case because we are restricting the inclinations to vary from 0 to 180 degrees the value of tan is uniquely determined. And therefore, because $\alpha \wedge \beta$ lie in 0 to 180 degrees $\alpha = \beta$ which resolves the problem; that means, their inclinations are same. That means the two lines are parallel. So, l_1 is parallel to l_2.

So, what is a characterization of parallel lines? That means, if I want to say two non-vertical lines l_1 and l_2 are parallel then it suffices to check whether their slopes are equal or not. If they are parallel then the slopes better be equal and if the slopes are equal then we have parallel lines. Now similar characterization we are searching for in perpendicular lines. So, let us go and try to figure out this characterization for perpendicular lines.

(Refer Slide Time: 06:42)



Let us try to visualize, what are the perpendicular lines? So, here are two perpendicular lines one l_1 and l_2 let us take the orange line as l_1 and blue line as l_2. So, l_1 will have slope m_1, l_2 will have slope m_2 angle of inclination of l_1 is α then inclination of β, if it is perpendicular to line l_1 is $90+\alpha$ which is β. And, then you may play with the tangent function of it and you can get something which is very interesting.

So, let us try to figure out what is that interesting thing that we are getting. So, to put it formally let l_1 and l_2 be two non-vertical lines because I cannot work with vertical lines θ equal to 90 degrees, the concept of slope is not defined which slopes m_1, m_2 inclinations α

and β respectively, no problem in this. If l_1 is perpendicular to l_2 as is the case in this a figure I have β is equal to$90+\alpha$.

So, if I want to figure out the relation between the slopes of 1 1 and l_2 then it is a good idea to take tangent of β. So, let us take that. So, $\tan\beta=\tan(90+\alpha)$, but $\tan(90+\alpha)$ if you use that simple formula that is available to you is $-\cot\alpha$ which also can be written as $\frac{-1}{\tan\alpha}$.

But what is $\tan\alpha$? $\tan\beta$ is the slope of a line l_2 which is m_2 and $\tan\alpha$is the slope of a line l_1 which is m_1. So, what we have just now derived is $m_2=\frac{-1}{m_1}$ or $m_1 m_2=-1$. That means, if you take two slopes if you take slopes of two lines take a product of them and if you get the quantity to be equal to -1; that means, you have got a perpendicular line.

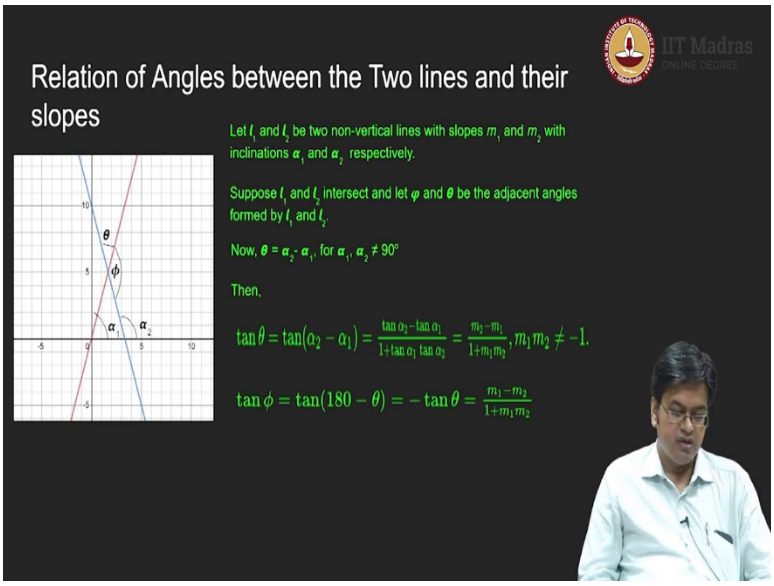
But right now, we have not proved that result, what we have proved just now is if l_1 is perpendicular to l_2 then the product of the slopes better be -1. Now I want to prove if the product of the slopes is -1 then the lines are perpendicular, how will I go about this? Exactly the way we went for parallel lines.

So, $m_1 m_2 is-1$ then I; obviously, $\tan\alpha\tan\beta=-1$; that means, $\tan\beta$will be equal to $\frac{-1}{\tan\alpha}$ or $\tan\alpha=-\cot\beta$but what is $-\cot\beta$? $\tan(90+\beta)$or either it will be this way or it will be the other way so, $\tan(90-\beta)$. So, $-\cot\beta$ is either $\tan(90+\beta)$ or $\tan(90-\beta)$, in any case the difference between α and β is 90 degrees.

Therefore, l_1 is perpendicular to l_2. Hence, we have proved a characterization that if two non-vertical lines are perpendicular to each other, the product of their slopes is equal to -1 which can be written in this form. Two non-vertical lines l_1 and l_2 are perpendicular if and only if $m_1 m_2=-1$ or you can verbally write product of their slopes is equal to - 1.

So, this is the characterization of the perpendicular lines via slope. So, what we have seen so far is the characterization of parallel lines by slope and characterization of perpendicular lines via slope, what if they are not parallel or perpendicular and they intersect just like that? If they are not parallel then they better intersect each other.

(Refer Slide Time: 11:37)



So, in general if I want to have an intersection of two lines and I know the slopes of those two lines. Can I talk about the angle of intersection of these two lines? The answer is yes. So, here is the relation of angles between the two lines and their slopes. So, what I want to say if once I show the figure it will be clear.

As of now let us understand I have two non-vertical lines with slopes m_1 and m_2, inclinations $\alpha_1$and α_2 respectively. And, l_1 and l_2 intersect each other, they are not parallel so they will intersect somehow and they are not perpendicular also. So, they intersect in angles ϕ and θ are the adjacent angles that are formed by l_1 and l_2, if they intersect in a perpendicular manner the adjacent angles will be 90 degrees each. So, that is not an interesting case because we have resolved that case.

So, now, if they intersect at any angle then this figure will look like this; let us first understand this figure. So, there are two lines l_1 and l_2. So, l_1 has angle of inclinationα_1, l_2 has inclination α_2 these two lines intersect over here near y coordinate $\phi$and they have two angles; one is θ, another one is ϕ.

So, these two angles are adjacent angles. What can you say about the angle θ that is formed? As you can see the angle α_2 is obtuse and α_1 is slight acute. So, the angle θ is actually α_2 minus α_1 provided α_1 and $\alpha_2$are not equal to 90 degrees. Why? Because I cannot consider vertical lines as simple as that. So, the angle is 90 not equal to 90 degrees, $\theta = \alpha_2 - \alpha_1$.

So, if I want to talk in terms of slopes of these lines, I better take tangent function and apply it to the angle θ. So, let me do it. So $\tan\theta = \tan(\alpha_2 - \alpha_1)$. Take a standard trigonometric formula of $\tan(\alpha_2 - \alpha_1)$, you will get $\frac{\tan\alpha_2 - \tan\alpha_1}{1 + \tan\alpha_1 \tan\alpha_2}$. But what is $\tan\alpha_2$? $\tan\alpha_2$is nothing but the slope of line l_2 which is m_2 and $\tan\alpha_1$ it is slope of line l_1 which is m_1.

Therefore, the answer to this is $\frac{m_2 - m_1}{1 + m_1 m_2}$. So, I know what is $\tan\theta$, now you can look at the angle ϕ which is $180 - \theta$. So, I can similarly derive a relationship for $\tan\phi$which is $\tan(180 - \theta)$, we have already seen, this is $-\tan\theta$.So, that $m_2 - m_1$ will be swapped to $m_1 - m_2$ denominator remains the same, the condition $m_1 m_2 \neq -1$remains the same because they should not be perpendicular.

In this case we have figured out what is the relation of tan of that angle with respect to the slopes of the lines. So, this finishes our discussion on two lines. Now another interesting question that comes is, what if the three points are collinear, then how will the slopes be interpreted? Imagine three points are collinear then what happens is their slopes must be equal because they are all lying on the same line right and there is one common point.

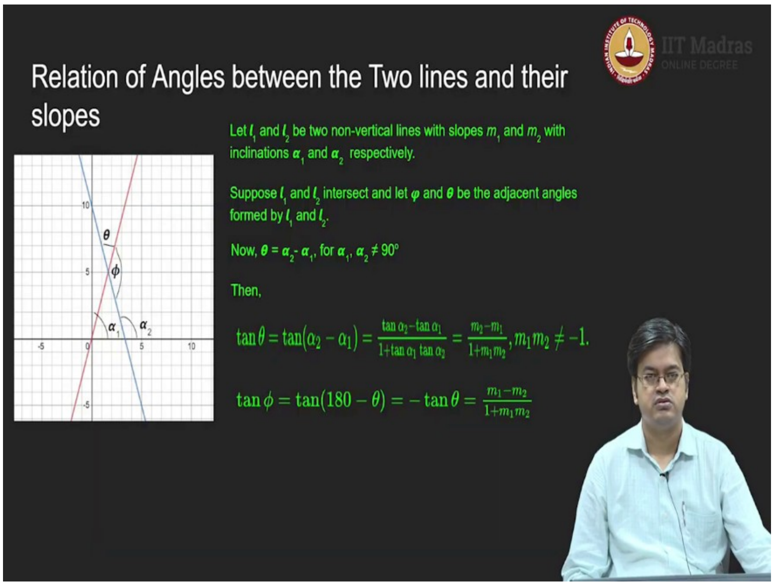
So, if A, B, C are collinear slope of AB is equal to slope of BC and therefore, all of them must be collinear. So, if there is any common point in between those three points the slopes are equal, the points are collinear, that is called the relation of collinearity using slopes.

Mathematics for Data Science 1
Prof. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras

Lecture - 18
Representation of a Line-1

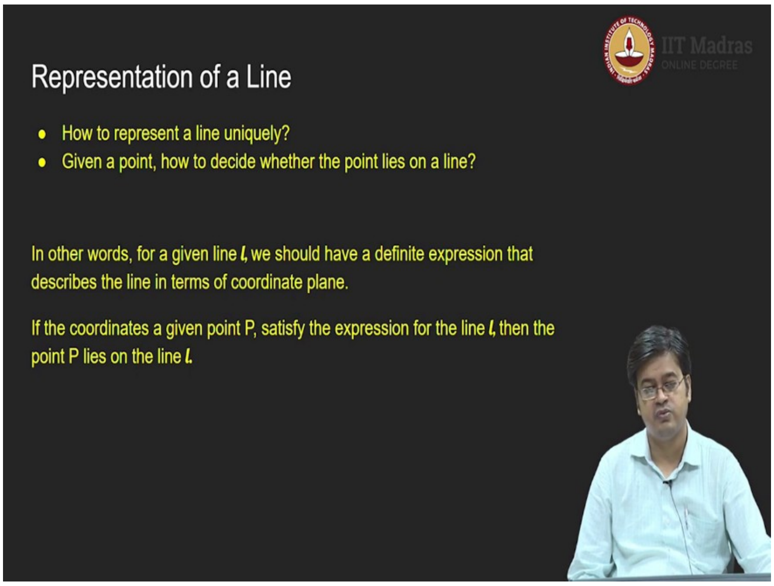
(Refer Slide Time: 00:06)

(Refer Slide Time: 00:14)



So, what we have seen so far is, what is a relation of the slope with respect to line and we have exploited certain how we can use the slope to determine whether the lines are parallel perpendicular. And, if I know the slope of the line then how will I find a slope of two non - vertical lines, then how will I find the relation between angles and other properties.

(Refer Slide Time: 00:41)



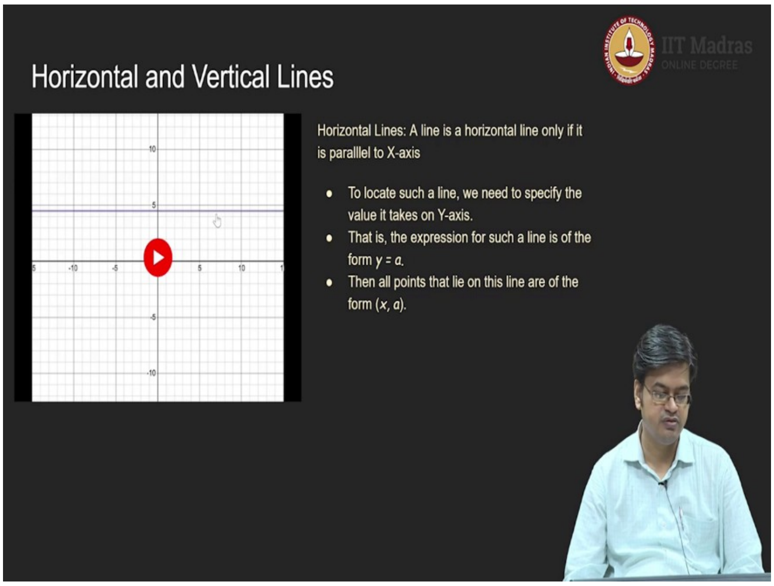
Now, we will come to the Representation of a Line, as we have already seen slope cannot represent a line uniquely. So, what is it that, that is required for representing a line uniquely? So, this raises two questions, how to represent a line uniquely? And the second question is, given any point of how will you decide whether that point lies on the line or not?

So, in order to answer these two questions, let us take the first question first and rephrase it. So, if I want to represent a line uniquely, then what I need to figure out is, I need to figure out a condition or a definite expression which will describe the line in terms of its coordinate plane. So, for a given line l I should be able to find a definite condition or expression which describes the line in terms of coordinate plane. That is in terms of the coordinates or to be more precise what should be the condition on the coordinates in order to describe the line l.

If I can understand what is this condition then the second question is automatically answered because if the coordinates of P are given to me and they satisfy the condition or expression for the line l then they must lie on the line l otherwise they do not lie on the line l, then it is just a simple job of checking whether that condition is satisfied or not. So, with this in mind we will try to answer the first question that is how to represent a line uniquely?

Now, what kind of lines we have seen so far? We have seen lines which are similar to X - axis, lines which are similar to Y - axis; those are typically horizontal and vertical lines.

(Refer Slide Time: 02:37)



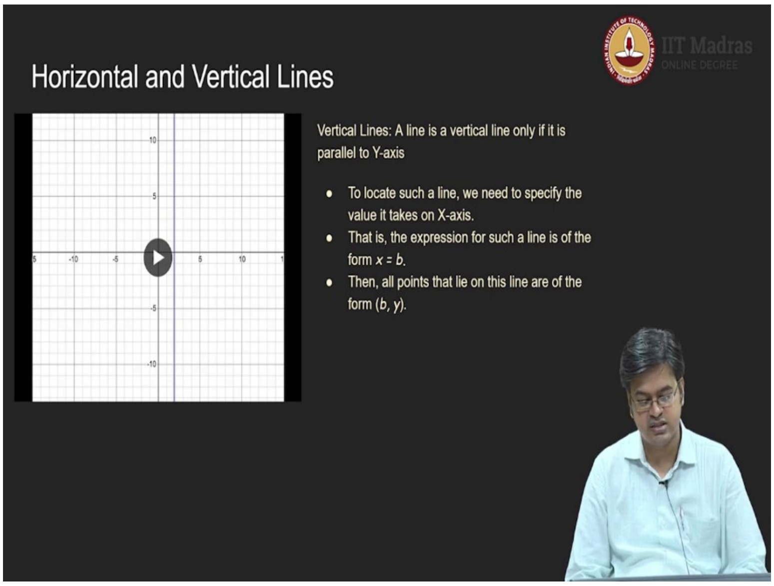
Let us first understand, what is a horizontal line. So, a line is said to be a horizontal line if it is of this form, now this line can be infinitely many. So, you can have infinitely many horizontal lines as can be seen from the video. Now, how to represent this line uniquely is my question. So, let us say I need to find this line or the condition for this line, how can I find the condition for this line?

So, let us first define this line as a horizontal line and let us say horizontal line is a horizontal line if and only if it is parallel to X - axis, this is our definition of a horizontal line. Now, if I want to specify this line uniquely what do I need to know? I need to know the distance of this line or the location of this line from X - axis, that is I need to know the y coordinate of this line you can see here. So, I want to locate this line or the value that it takes on Y - axis if I want to specify this line.

Let us say this value is given to be a then I know it is a horizontal line. So, all points will lie at a same distance from X - axis therefore, all points will satisfy the condition $y=a$. You take any point on this line it will satisfy the condition $y=a$.

So, in case of horizontal lines what I have done is I have identified the condition that is $y=a$. So, what will be the condition on points? The points will be of the form (x,a), x can be any value, but the y coordinate of that point will be fixed that is a. In a similar manner we can consider vertical lines.

(Refer Slide Time: 05:02)

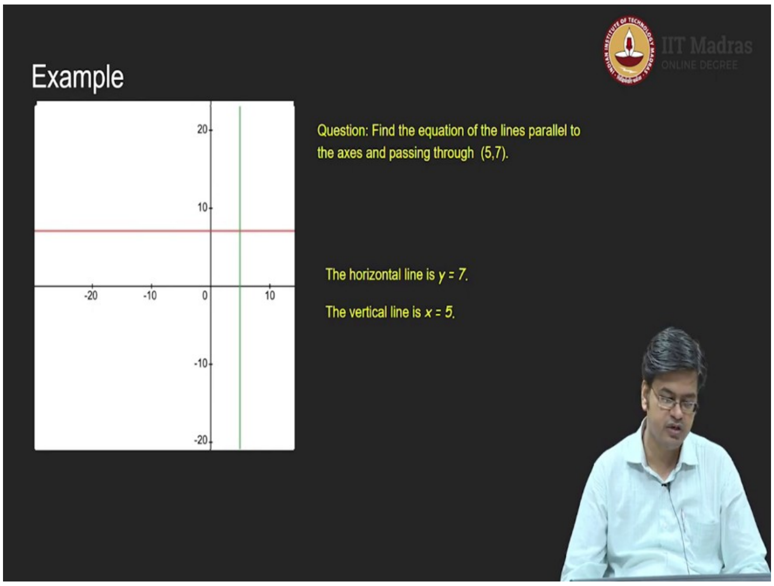


So, what is a vertical line? You can see this in this image and in this video, we can see all these kind lines are called vertical lines. So, how will I identify these vertical lines? First, I will define the vertical lines a line is a vertical line if and only if it is parallel to Y - axis. Now, to specify the location what do I need? To locate a line, I need to know the distance of this line from Y - axis ok. So, that essentially means what value it takes on x coordinate or X - axis.

So, how will you identify this? You just need to identify the one point in this particular line let us say this is the point and I need to see what is the distance of this point from X - axis, if that is b, then all points of the form (b, y) will be lying on this line; all points of the form (b, y) will be lying on this line. And therefore, the equation of the line the expression for the line will have a form $x=b$.

I mentioned the all points will be of the form (b, y). So, if I get two points where the y coordinate where the x coordinate is fixed and I know it is a line then I know it is parallel to Y - axis or it is a vertical line right. In a similar manner the other one is parallel to X - axis and it will be a horizontal line. Let us make it more crystal clear by solving one example.

(Refer Slide Time: 06:49)



So, here is an example where a question is given to you want to find the equation or expression for the lines parallel to the axis and passing through point $(5,7)$. Now, the lines are passing through point $(5,7)$ and it is also given that they are parallel to axes. So, a line which is parallel to X - axis is known as horizontal line, a line which is parallel to Y - axis is known as vertical line. So, essentially this question asks you to find one horizontal line and one vertical line.

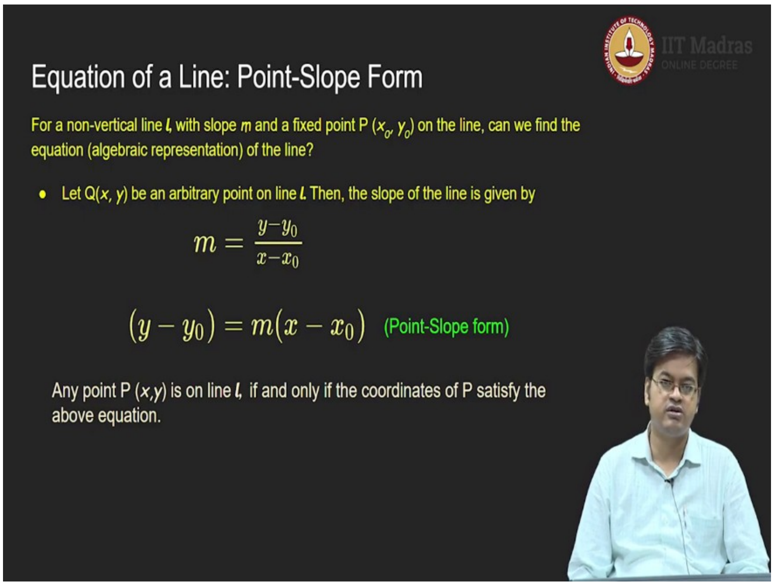
So, let us go to the coordinate plane, this is the coordinate plane let us locate the point $(5,7)$, it will be somewhere here. Now let us first focus on identifying the horizontal line. What is a horizontal line? A line which is parallel to X - axis is a horizontal line. So, a line which is parallel to X - axis, then what do I need to know? Its distance from X - axis, the distance is 7 according to this particular expression because $(5,7)$ is a point on that line.

So, the distance is 7 so, the line must appear somewhere here, now further the next question is I want to find a vertical line that passes through point $(5,7)$. So, now I need to know the distance of a line from X - axis. So, I will locate point 5 over here and all points on the line on that particular line will be of the form $(5, y)$. So, $(5,7)$ will also fall on that line. So, this is the line; so, this is how we will find the lines.

Now what are the typical equations of the line? So, the horizontal line will be $y=7$ and the vertical line will be $x=5$. This is how we will study horizontal and vertical line. So, what is a vertical line? Vertical line has inclination as at 90 degrees, and therefore, the slope of this line is not defined remember this in mind.

Another point which is horizontal line it never intersects actually X - axis, but the inclination of this line with respect to X - axis is 0 degrees therefore, it will have a slope 0. So, we have eliminated the cases where the slope does not exist or slope is 0, now we need to identify similar kind of expressions for lines which are not vertical. So, let us go further and identify such expressions.

(Refer Slide Time: 09:48)



So, here as we already know that slope cannot uniquely determine line, then the question is which slope if I give you some more information can you determine the line? So, here in this case what we are identifying is we are giving a point and giving a slope and then we are asking a question can we solve this problem or can we find a unique expression for a given line? So, the question is for a non-vertical line l vertical line we do not have to consider because the slope is not defined. So, for a non-vertical line l with slope m and a fixed-point P (x_0, y_0) on the line can we find the equation or the algebraic representation of a geometric object that is line is the question.

So, here what are the things that we know? We know slope and we know a point on the line. So, in order to answer this question, we know that two points uniquely determine a line. So, let us take another point Q (x, y). I do not know the coordinates of these points, but I assume that this point lies on line l. Now, I know from the definition of slope that I have defined change in y by change in x the slope of a line is given by. So, what are the two points now? Q and P.

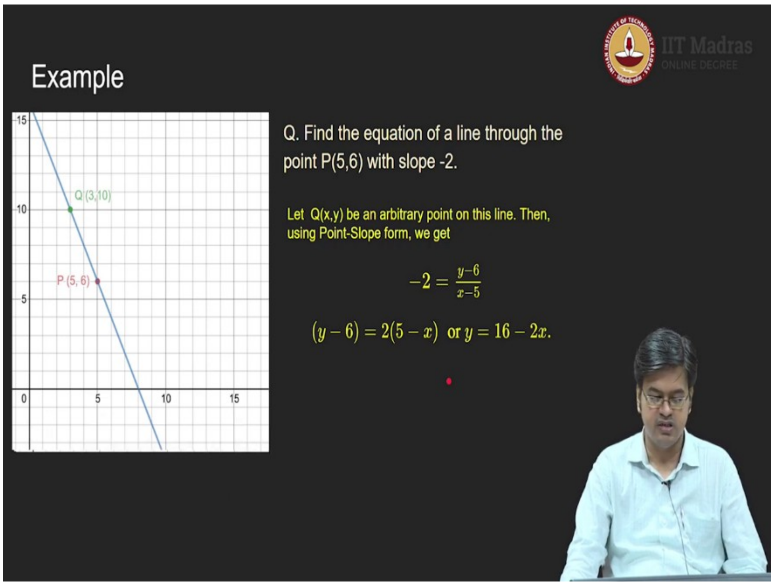
So, change in y will be $y-y_o$ and change in x will be $x-x_0$. So, I know $m=\frac{y-y_0}{x-x_0}$, this is what I know from my definition. It has nothing to do with $\tan\theta$ even if you have it you can find out what is $\tan\theta$, but since nothing is known in specific we cannot find the $\tan\theta$, but $\tan\theta$is anyway given to you in terms of slope.

So, now I have $m=\frac{y-y_0}{x-x_0}$. So, how will I find the condition on x and y? Just cross multiply this $x-x_0$, you will get an expression which is $y-y_0=m(x-x_0)$. This condition uniquely identifies my line, there cannot be any other line satisfying this condition.

So, therefore, any point that lies on this particular line that is P (x,y) that lies on this particular line, it must satisfy the condition that is given here. This form of expression is called point slope form. So, this is a point slope form of equation which essentially says that give me one point and slope of the line I will give you the equation of a line.

The beauty is the geometric object now can be represented in terms of the equation, initially when we started, we tried to represent a point which is a geometric object in terms of coordinate plane and the coordinates of the point. Now we are giving infinite set of points having certain condition that is a geometric object of line how you can represent it algebraically using the equation of a line. So, this is point slope form. Let us try to see how we can use the point slope form in our problem solving.

(Refer Slide Time: 13:40)



So, now, I have been asked to find the equation of a line which passes through point $(5,6)$ and has slope of -2. Here the interesting thing is slope is negative. So, let us identify the point $(5,6)$ on the coordinate plane and now I want to identify the line that passes through this with slope of -2. So, now I have a formula for point slope form, I can use that formula and I can straight away derive it for let us try Q (x,y) is an arbitrary point.

So, I need two points to identify a line. So, Q (x,y) with the arbitrary point on this line then using point slope formula we simply substitute $-2=\frac{y-6}{x-5}$, slightly rearrange the terms; what you will get is $y-6=2(5-x)$. If you simplify this you will get the expression $y=16-2x$.

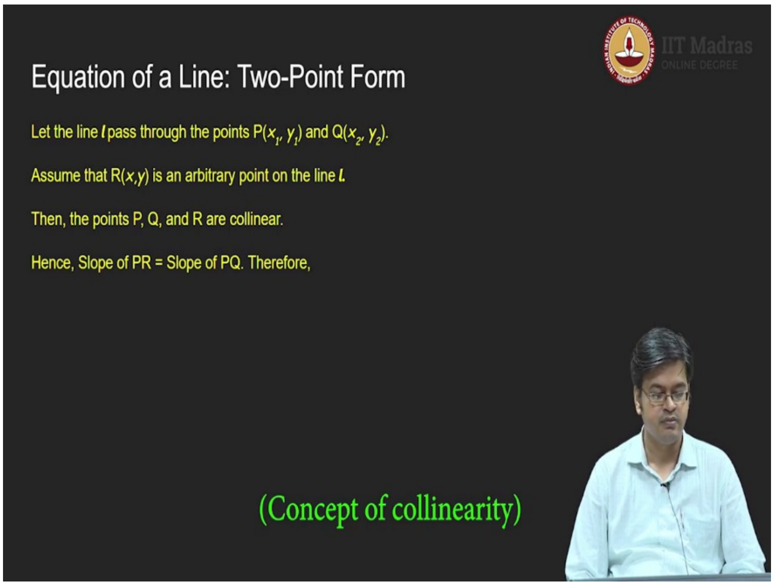
Now, let us try to see, if I want to know this value of x what point what value of y will satisfy this equation. Let us put x is equal to 3 here if I put x is equal to 3 here then I get y is equal to 10 after simplifying this I will get $y=10$.So, that means, the point $(3,10)$ should lie on this particular line. So, let us see that $(3,10)$ is here and now you know from basic geometry that two points uniquely identify a line. So, you can just draw a line using your ruler passing through these two points this is the line that we are expecting.

So, the question did not ask you to draw a graph, but drawing graph always verifies whether you have found a correct answer or not. So, it is better to cross check using graphs. So, the answer to the question is the equation of the line passing through point $(5,6)$ and slope -2 is

$y=16-2x$. Now, suppose somebody decides not to give me slope and somebody says that now you have been given only two points; can you find the equation of line?

The answer is; obviously, yes because given two points I can always determine the slope right for example, in our earlier case when we defined slope I need to figure out what is change in y and what is change in x using these two points and that will give me slope to be equal to -2. And therefore, I can always use this formulation to find the equation of the line, but you can use this knowledge and derive another form that is equation of a line two - point form.

(Refer Slide Time: 17:00)



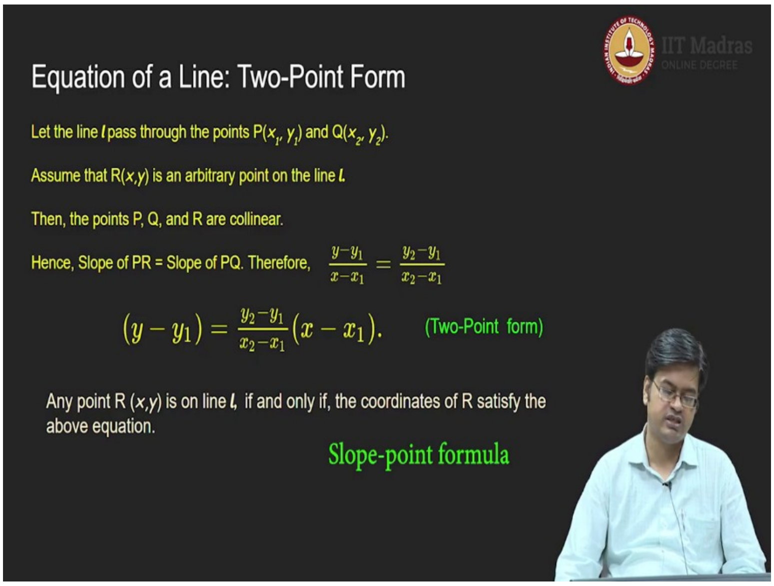
So, given two points the question is can you determine the line uniquely which should be possible and through our basic knowledge of geometry we already know that two points uniquely determine a line, now we will see that in our coordinate geometry. So, the assumption is let the line l pass through points P and Q with coordinates (x_1, y_1) and (x_2, y_2).

To start with this, I will take another point R which is arbitrary point because I want to find the condition in terms of coordinates. So, whenever I want to find the equation of line I will start with an arbitrary point. So, R (x, y) is an arbitrary point on the line l. Now, look at these three points P, Q, and R they all lie on one line therefore, the points P, Q, and R are collinear yes; so, points P, Q, and R are collinear.

Therefore, suppose I consider only these two points P and R, using these two points P and R, I can easily figure out the slope of a line. If I consider points P and Q, I also know the slope of a line; now because these points are collinear what can you say about slope of line PR and slope of line PQ, both must be same or equal? So, slope of PR is equal to slope of PQ because they are collinear.

So, if this is the case, then what is slope of PR? You can easily figure out P is this $(x_1, y_1)$and R is (x, y). So, the slope of PR first you consider change in y, $y-y_1$ upon change in x that is $x-x_1$that is slope of PR. What is slope of PQ? PQ is(x_1, y_1) and (x_2, y_2). So, y change in y is y_2-y_1 and change in x is x_2-x_1.

(Refer Slide Time: 19:21)

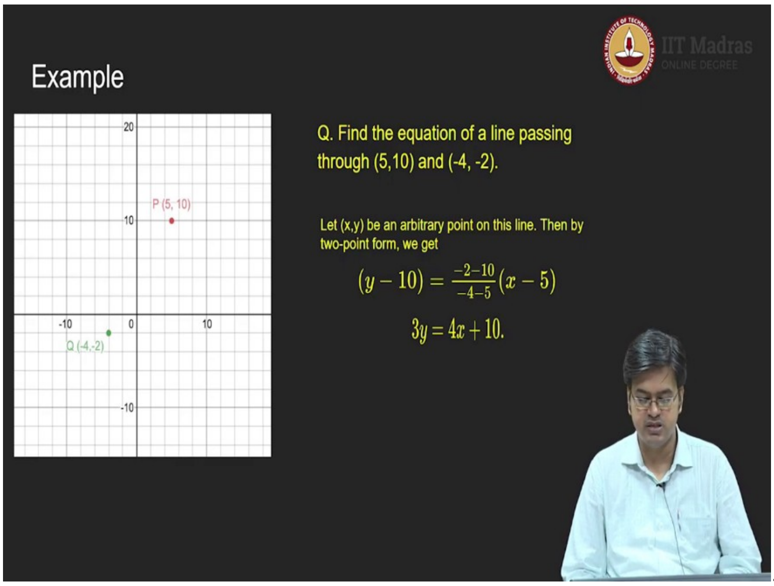


Therefore, I will get the equation of this form $\frac{y-y_1}{x-x_1}=\frac{y_2-y_1}{x_2-x_1}$.

So, again if you look at it closely this particular thing is nothing, but the slope of a line m and we are doing things which are very similar to slope point form. But instead of counting it explicitly we are counting it as a ratio and then you rearrange the term and you will get this expression because you just take this denominator on the other side and you will get this expression.

Now, this line is again uniquely characterized and therefore, any point that lies on this line must satisfy this condition. So, if your point is R lies on this line then it must satisfy this condition and this form is called two - point form. So, remember these are the formulas that we are deriving; first was slope line formula, second is two - point form.

(Refer Slide Time: 20:38)



Let us understand this formula better by solving some examples. So, let us take one example where I want to find the equation of a line that is passing through two points $(5,10) \wedge (-4,-2)$. Let us identify these two points on a coordinate plane $(5,10),(-4,-2)$. I want to find the equation of this line.

So, I will use another point Q which is an arbitrary point and it has a coordinate(x,y), I will use the two - point form. So, using two- point form what should I get? So, I am taking, this point P. So, $(y-10)=\frac{-2-10}{-4-5}(x-5)$. So, always remember this order does not matter I can always start with this as well. So, change in y is $10-(-2)$ and change in x is $5-(-4)$ in both cases my answer to this particular fraction will be $\frac{12}{9}$ which is $\frac{4}{3}$.

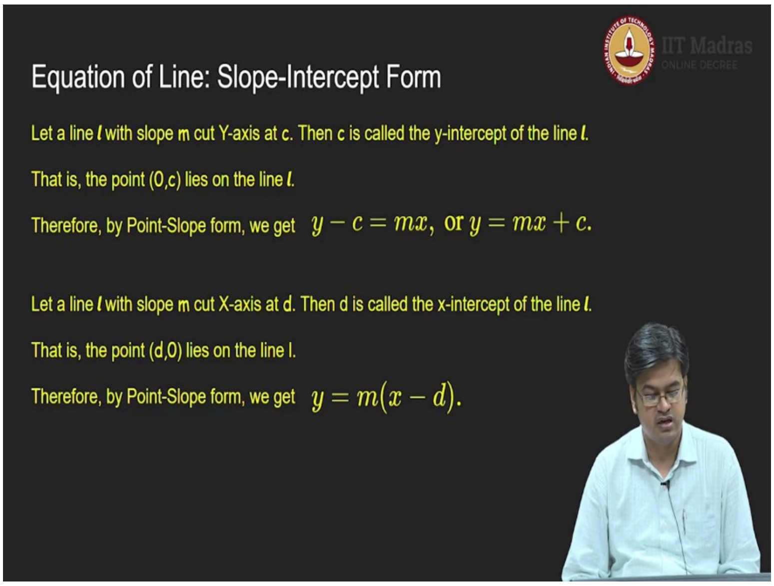
So, it does not matter whether you take this as (x_1,y_1) or you take this as (x_1,y_1), you will always get the same answer. So, if you simplify this you will get the expression of a line because as I mentioned the slope was $\frac{4}{3}$. So, you just simplify this you will get the expression of a line$3y=4x+10$. So, this will be the line that is passing through these two points.

Mathematics for Data Science 1
Prof. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras

Lecture - 19
Representation of a Line-2

(Refer Slide Time: 00:05)

(Refer Slide Time: 00:14)



Now let us go ahead and try to figure out some spatial variations where the calculations become extremely easy for. These two forms are primary two-point form and slope-point form. So, when you consider slope point form you can also consider a special case that is slope-intercept form. So, this is the methodology that we will use for considering slope-intercept form, before that let me define what is an intercept.

So, let l be the line with slope m that cuts Y - axis at point c. Then this c is called y intercept of the line l. So, what is the meaning that it cuts Y - axis at c? The y coordinate of that point is c and the x coordinate is 0; that means, any point that it cuts through Y - axis of line l will be of the form $(0,c)$ and that $(0,c)$ will lie on line l.

Now we have our slope point form instead of having any point (x,y) you have a specific point which is $(0,c)$. So, I apply the slope point form or point slope form in this expression. What you will get instead of $y-y_0$ you have $y-c$which is equal to m, m is the slope of the line m times $x-x_0$. What is x_0? Zero.

So, so we will get $y-c=mx$ and therefore, I will get a form $y=mx+c$, this is a standard form that we generally deal with when we are dealing with straight lines. So, you have got a slope-intercept form which is of the form $y=mx+c$.

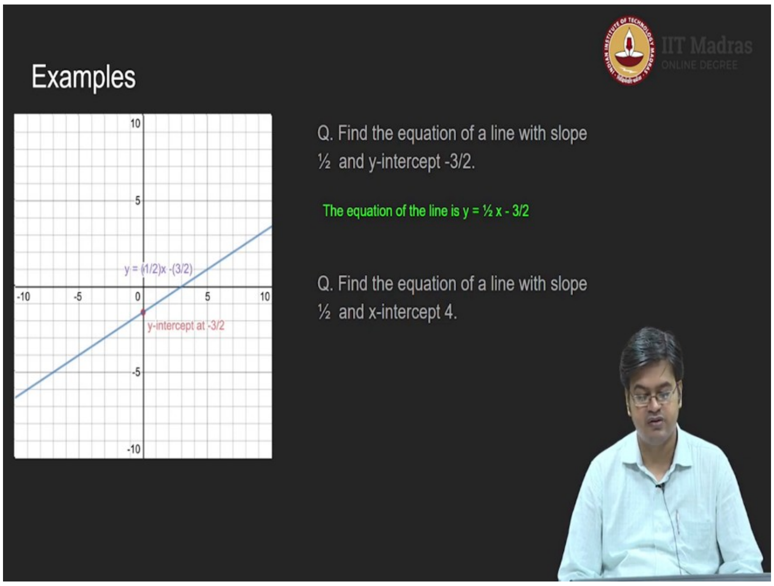
The interesting fact is the calculations are very simple whenever you are given the slope-intercept form. For example, now if you know the y intercept is at c and the slope is m you do not have to do any calculations, but straight away write this expression that is y is equal to take the slope m, take the intercept c; $y=mx+c$ will be your answer.

Therefore, the calculations simplify significantly when you are considering a slope-intercept form. If the intercept is not available then you may have to go to that point slope form and figure out what it is. Now there can be if the line cuts Y - axis the line can as well-cut X - axis. So, there can be another variation of this formulation that is if a line l with slope m cuts X - axis at point d. Then d will be called as x intercept of the line l.

If d is called as x intercept of the line l then how will this point lie on the line l or what are the coordinates of the line that intersects X - axis and line l? So, what is the point of intersection? That will be $(d,0)$ and this $(d,0)$ lies on line l. So, I will again use the point slope form of the line.

So, if I want to use point slope form $y-0=m$¿) will be the answer. So, that will be the form $y=mx-md$. So, let us try to use this and solve some problems for finding the equations of the line using slope-intercept form.

(Refer Slide Time: 04:22)



So, typically some example like this. So, I want to find the equation of a line with slope is $\frac{1}{2}$ and y intercept is $\frac{-3}{2}$. Remember here things are very easy because you just need to know $mx+c$. So, what is m? m is $\frac{1}{2}$ and c is $\frac{-3}{2}$. So, upfront I can tell you orally this, the equation of the line will be $y=\frac{1}{2}x-\frac{3}{2}$. Let us verify the result using the graphics and all other things.

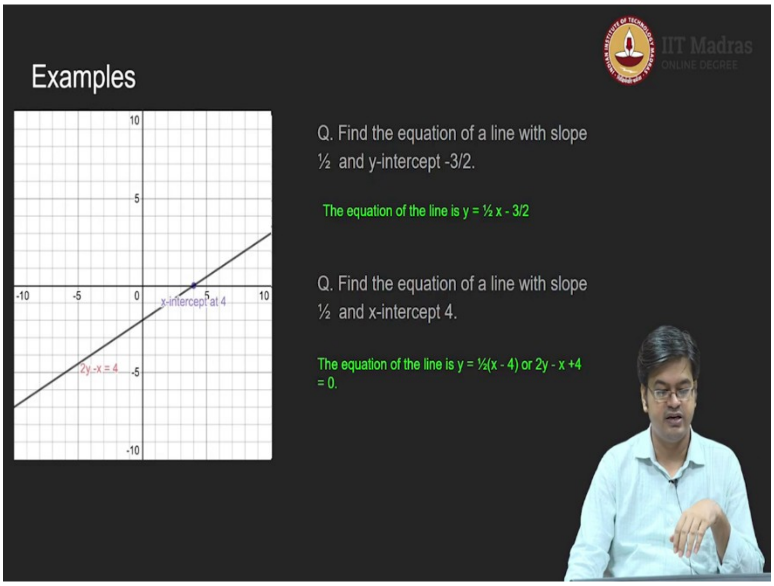
So, here is the y intercept of this particular line. So, here the y intercept is at point $\frac{-3}{2}$. Now slope is half correct. So, the equation of line you can easily see is $y=\frac{1}{2}x-\frac{3}{2}$. So, let us try to figure out what is the x intercept of this line. So, $y=\frac{1}{2}x-\frac{3}{2}$. So, the x intercept of this line is 3. So, the question could have been asked that find the equation of a line with slope half and x intercept equal to 3 that also can be a question and the answer will be same.

So, let us see what is the next question that is find the equation of a line with slope half, but x intercept is 4 it is not 3. So, it is definitely not a same line because x intercept is 4, but the slope is half. So, can you relate it to some of the concepts? The slope is half; that means, the slopes are equal, we have seen that if the slopes are equal then lines must be parallel to each other.

So, therefore, I can easily see that the line must be parallel to this line with some different intercept which is at 4 for this the intercept is 3 so, intercept is 4. So, what can be the y intercept can also be an interesting question. We will answer it later. Right now, let us see how we can answer the question that is asked here. Find the equation of line with slope half and x intercept 4.

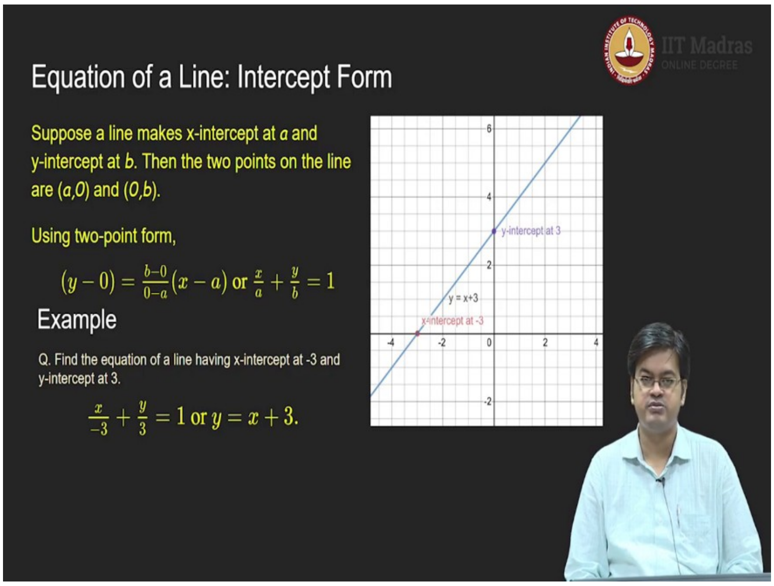
So, according to our formulation $y = mx - md$. So, where d is the intercept that is 4 so and this is half. So, $y = \frac{1}{2}x - 4$ is the equation of this line.

(Refer Slide Time: 07:00)



You can simplify this which will give you $2y - x + 4 = 0$. So, this will be the expression for the line. This is the slope-intercept form of the line, now we can go to two-point form that is suppose I have been given x intercept and y intercept how will I identify the line.

(Refer Slide Time: 07:33)



So, let us now go to the form of intercept that is intercept form, how to find equation of line when you have been given two intercepts x and y. So, let us formulate the hypothesis, suppose a line makes x intercept at a, y intercept at b, then naturally the coordinates of these two points are $(a,0)\wedge(0,b)$. So, we will use two- point form to derive the equation of line.

So, I will take this point as the first point therefore, the y coordinate is 0. So, $(y-0)=\frac{b-0}{0-a}(x-a)$.

Now, if you divide this expression throughout by b then you will get $\frac{y}{b}=\frac{-x}{a}+1$. Because this has a minus sign shift it to the left hand side and you will get this expression which is $\frac{y}{b}+\frac{x}{a}=1$, now you see how beautiful is this expression; x intercept is a so, below x you put a y intercept is b. So, below y you put b.

Therefore, there is nothing to memorize, it is just a simple trick that $\frac{x}{x-intercept}+\frac{y}{y-intercept}=1$ that is how you will get the intercept form. So, it is very easy to solve the problems if you remember this trick.

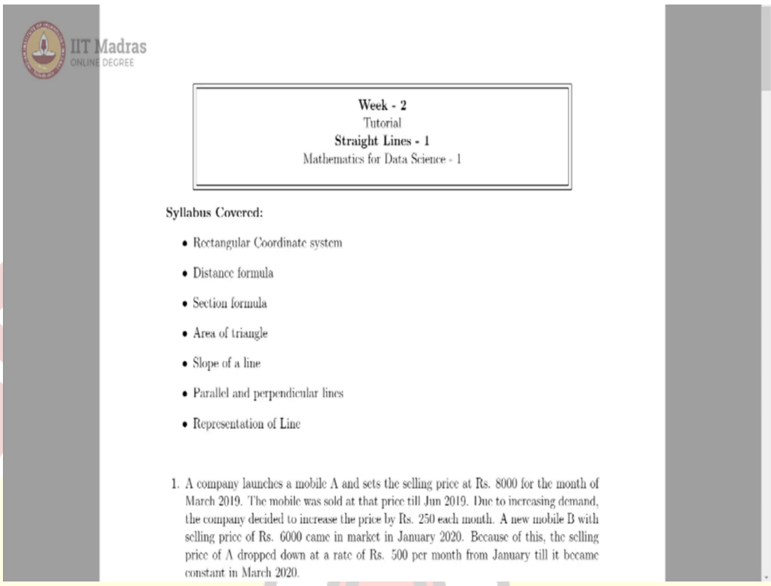
Now, let us take one example where we need to find this. So, find the equation of line having x intercept at -3 and y intercept at 3. So, you do not have to do any complicated calculations, you can simply say $\frac{x}{-3}+\frac{y}{3}=1$, multiply throughout by 3 you will get the expression $y=x+3$.

So, let us verify whether this satisfies because it is always better to verify using graph. So, x intercept is -3 y intercept is 3, the line that passes through these two points is $y=x+3$. This is what the intercept form is, it is very simple and you can practice more and more problems.

That is all for today.

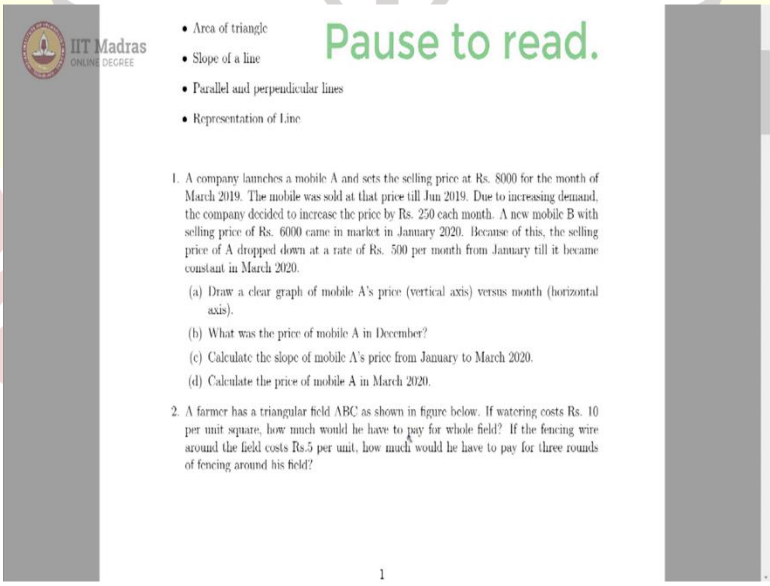
Mathematics for Data Science 1
Indian Institute of Technology, Madras
Week 02
Tutorial 01

(Refer Slide Time 00:19)



(())(0:18) In this tutorial we are going to look at the problems which are related to contents of week 2, that is to do with straight lines and all these topics here.

(Refer Slide Time 00:34)



- Area of triangle
- Slope of a line
- Parallel and perpendicular lines
- Representation of Line

1. A company launches a mobile A and sets the selling price at Rs. 8000 for the month of March 2019. The mobile was sold at that price till Jun 2019. Due to increasing demand, the company decided to increase the price by Rs. 250 each month. A new mobile B with selling price of Rs. 6000 came in market in January 2020. Because of this, the selling price of A dropped down at a rate of Rs. 500 per month from January till it became constant in March 2020.
   (a) Draw a clear graph of mobile A's price (vertical axis) versus month (horizontal axis).
   (b) What was the price of mobile A in December?
   (c) Calculate the slope of mobile A's price from January to March 2020.
   (d) Calculate the price of mobile A in March 2020.
2. A farmer has a triangular field ABC as shown in figure below. If watering costs Rs. 10 per unit square, how much would he have to pay for whole field? If the fencing wire around the field costs Rs.5 per unit, how much would he have to pay for three rounds of fencing around his field?

1

- Area of triangle
- Slope of a line
- Parallel and perpendicular lines
- Representation of Line

1. A company launches a mobile A and sets the selling price at Rs. 8000 for the month of March 2019. The mobile was sold at that price till Jun 2019. Due to increasing demand, the company decided to increase the price by Rs. 250 each month. A new mobile B with selling price of Rs. 6000 came in market in January 2020. Because of this, the selling price of A dropped down at a rate of Rs. 500 per month from January till it became constant in March 2020.
   (a) Draw a clear graph of mobile A's price (vertical axis) versus month (horizontal axis).
   (b) What was the price of mobile A in December?
   (c) Calculate the slope of mobile A's price from January to March 2020.
   (d) Calculate the price of mobile A in March 2020.
2. A farmer has a triangular field ABC as shown in figure below. If watering costs Rs. 10 per unit square, how much would he have to pay for whole field? If the fencing wire around the field costs Rs.5 per unit, how much would he have to pay for three rounds of fencing around his field?

1

- Area of triangle
- Slope of a line
- Parallel and perpendicular lines
- Representation of Line

1. A company launches a mobile A and sets the selling price at Rs. 8000 for the month of March 2019. The mobile was sold at that price till Jun 2019. Due to increasing demand, the company decided to increase the price by Rs. 250 each month. A new mobile B with selling price of Rs. 6000 came in market in January 2020. Because of this, the selling price of A dropped down at a rate of Rs. 500 per month from January till it became constant in March 2020.
   (a) Draw a clear graph of mobile A's price (vertical axis) versus month (horizontal axis).
   (b) What was the price of mobile A in December?
   (c) Calculate the slope of mobile A's price from January to March 2020.
   (d) Calculate the price of mobile A in March 2020.
2. A farmer has a triangular field ABC as shown in figure below. If watering costs Rs. 10 per unit square, how much would he have to pay for whole field? If the fencing wire around the field costs Rs.5 per unit, how much would he have to pay for three rounds of fencing around his field?

1

IIT Madras
ONLINE DEGREE

- Area of triangle
- Slope of a line
- Parallel and perpendicular lines
- Representation of Line

1. A company launches a mobile A and sets the selling price at Rs. 8000 for the month of March 2019. The mobile was sold at that price till Jun 2019. Due to increasing demand, the company decided to increase the price by Rs. 250 each month. A new mobile B with selling price of Rs. 6000 came in market in January 2020. Because of this, the selling price of A dropped down at a rate of Rs. 500 per month from January till it became constant in March 2020.
   (a) Draw a clear graph of mobile A's price (vertical axis) versus month (horizontal axis).
   (b) What was the price of mobile A in December?
   (c) Calculate the slope of mobile A's price from January to March 2020.
   (d) Calculate the price of mobile A in March 2020.
2. A farmer has a triangular field ABC as shown in figure below. If watering costs Rs. 10 per unit square, how much would he have to pay for whole field? If the fencing wire around the field costs Rs.5 per unit, how much would he have to pay for three rounds of fencing around his field?

1

IIT Madras
ONLINE DEGREE

- Area of triangle
- Slope of a line
- Parallel and perpendicular lines
- Representation of Line

1. A company launches a mobile A and sets the selling price at Rs. 8000 for the month of March 2019. The mobile was sold at that price till Jun 2019. Due to increasing demand, the company decided to increase the price by Rs. 250 each month. A new mobile B with selling price of Rs. 6000 came in market in January 2020. Because of this, the selling price of A dropped down at a rate of Rs. 500 per month from January till it became constant in March 2020.
   (a) Draw a clear graph of mobile A's price (vertical axis) versus month (horizontal axis).
   (b) What was the price of mobile A in December?
   (c) Calculate the slope of mobile A's price from January to March 2020.
   (d) Calculate the price of mobile A in March 2020.
2. A farmer has a triangular field ABC as shown in figure below. If watering costs Rs. 10 per unit square, how much would he have to pay for whole field? If the fencing wire around the field costs Rs.5 per unit, how much would he have to pay for three rounds of fencing around his field?

1

IIT Madras
ONLINE DEGREE

- Area of triangle
- Slope of a line
- Parallel and perpendicular lines
- Representation of Line

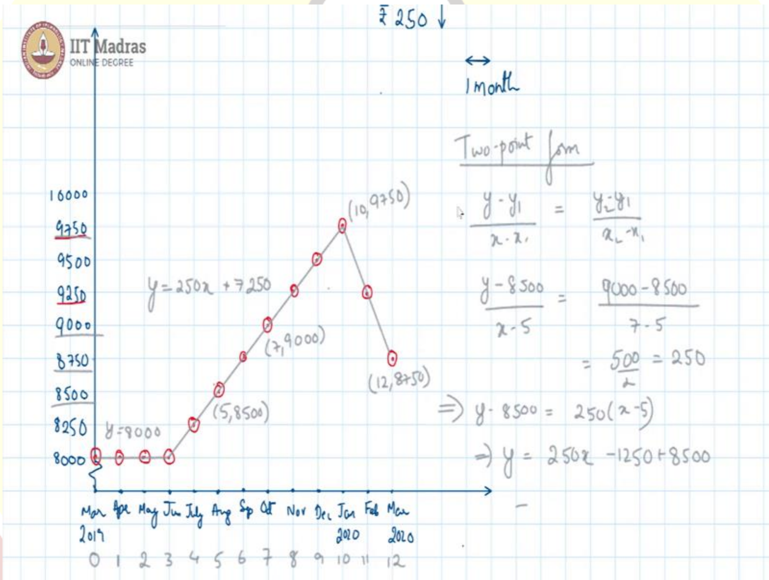
1. A company launches a mobile A and sets the selling price at Rs. 8000 for the month of March 2019. The mobile was sold at that price till Jun 2019. Due to increasing demand, the company decided to increase the price by Rs. 250 each month. A new mobile B with selling price of Rs. 6000 came in market in January 2020. Because of this, the selling price of A dropped down at a rate of Rs. 500 per month from January till it became constant in March 2020.
   (a) Draw a clear graph of mobile A's price (vertical axis) versus month (horizontal axis).
   (b) What was the price of mobile A in December?
   (c) Calculate the slope of mobile A's price from January to March 2020.
   (d) Calculate the price of mobile A in March 2020.
2. A farmer has a triangular field ABC as shown in figure below. If watering costs Rs. 10 per unit square, how much would he have to pay for whole field? If the fencing wire around the field costs Rs.5 per unit, how much would he have to pay for three rounds of fencing around his field?

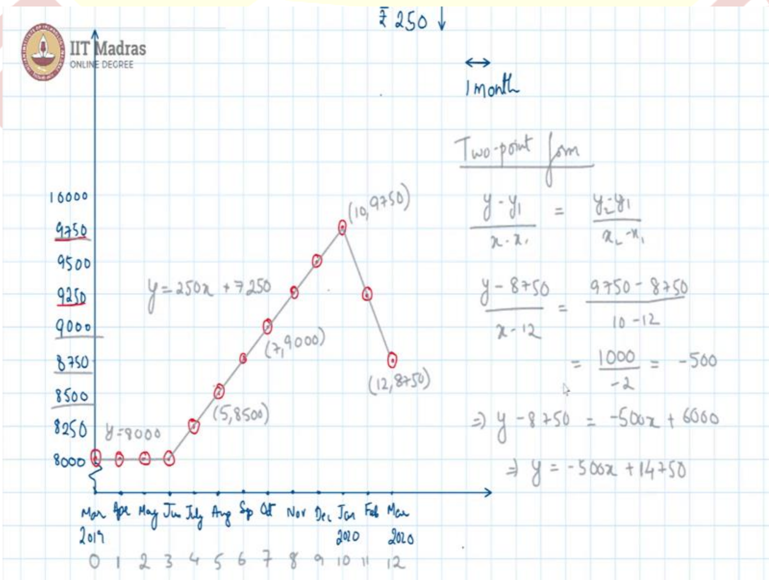
1

So, we will start with our first question. The data provided here is, there is a company which is selling mobile phones and it all begins in March 2019. In March 2019, the selling price was 8000 and it was sold at 8000 rupees, mobile A was sold at 8000 rupees from March until June. After that, due to increasing demand, the company decided to increase the price by 250 each month, so they are selling better.

So, they have decided to increase their price by 250 rupees every month. This went on until a new mobile B was launched at a lesser price, competition at a lesser price was launched in January. So, because of this the selling price of A dropped at a rate of 500 per month, from January till March 2020, so 2 months it had decreased. We are expected to demonstrate a clear graph of this. For that let us look at this graph.

(Refer Slide Time 01:58)





What we need to realize about situations like this is, the x and y axis do not necessarily represent the same units. So, we have along the x axis 1 unit is 1 month, however along the y axis 1 unit is, let us take about 250 rupees. So, 1 month and 250 rupees are not the same thing, so please remember this in situations like this. Now, because we are beginning from March let us take the starting month to be March, then this is April, May, and so on.

So, our entire problem deals with this 1 year span from March 2019 to March 2020, so this will be along our x axis. And now along the y axis, if we took each unit to be 250, then this is 250 and this is 500 and so on, the 8000 will be beyond our screen. So, to better represent our situation, we are going to introduce a zigzag here to indicate that a lot of values have been compressed into this little space. So, we are going to start from 8000 and this is going to be 8250, 8500, so on. And now we begin to mark out the points that we have, we know that in month of March the price was 8000, so this is the point for the month of March.

And then in April, May, and June the price stayed constant so it is been like this. And this portion can be represented using a horizontal line and this line is y is equal to 8000. Beyond that, the price had been increasing by 250 every month so in July we will be here, August here, September here, this will be October, this will be November, this is December, and this is January.

So, this segment can be indicated by this line, in order to find out the equation of this line we use the 2 point form, so we first write 2 points on this line segment. You could choose this one which is August, and for that let us number our months now, so March will be 0, April is 1, May is 2, June is 3, this is 4 and this is 5. So, our price point here it is $(5, 8500)$.

I am ready to take another month, so let us take October, this is the seventh month from March 2019, so this point becomes $(7, 9000)$. Using these 2 points, we can find the equation of the line by employing the 2 point form of the line equation, $\frac{y-y_1}{x-x_1} = \frac{y_2-y_1}{x_2-x_1}$, where x_1, y_1 and x_2, y_2 are two points on the line segment. So, here we can see it as $\frac{y-8500}{x-5} = \frac{9000-8500}{7-5}$. So, this would be equal to $\frac{500}{2} = 250$.

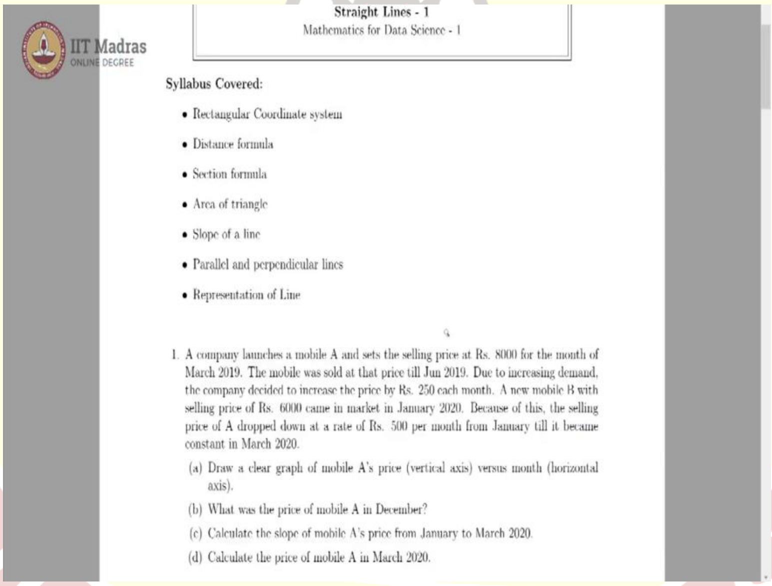
So that implies $y - 8500 = 250(x - 5)$, which finally gives us the line equation to be $y = 250x - 1250 + 8500$ plus this line is $y = 250x + 7250$. Moving on, the next 2 months, the price dropped by 500 each month. So, here we are at 9750, then for February we should be at

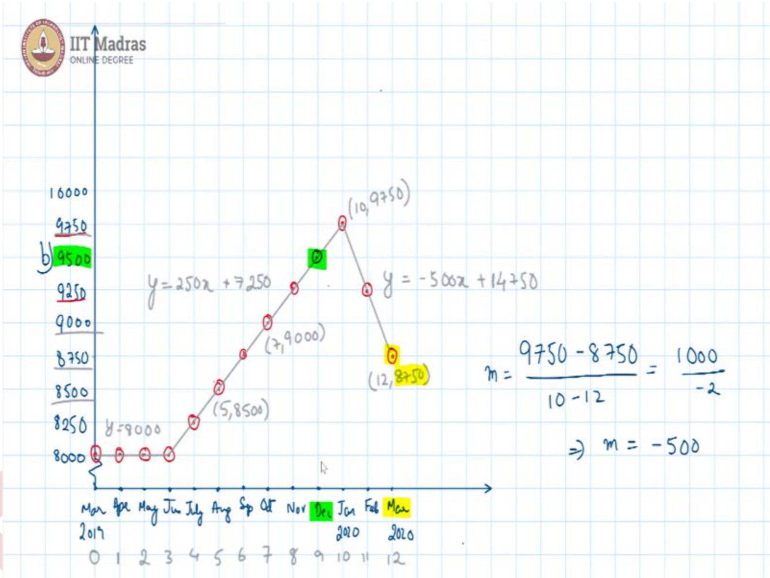
9250, so this will be our point for February and then the next month again 500 drop we will reach here, which is 8750. And this line segment also corresponds to a straight line, which also we can find using the 2 point form.

So, this point here is, let us number the months completely, this is 8, this is 9, this is 10, this is 11, this is 12. So, this point here, which is January is the tenth month, and the y axis gives us 9750 whereas this point here, this is the twelvth month, and it corresponds to 8750. And again, we would like to know the line equation for this and we use the 2 point form again.

So, this is $\frac{y-8750}{x-12}=\frac{9750-8750}{10-12}$ that gives us $\frac{1000}{-2}=-500$. Plus we have $y-8750=-500x+6000$. That gives us $y=-500x+14750$, so this is our new length. And this is a clear graph of the situation and the given question.

(Refer Slide Time 09:18)



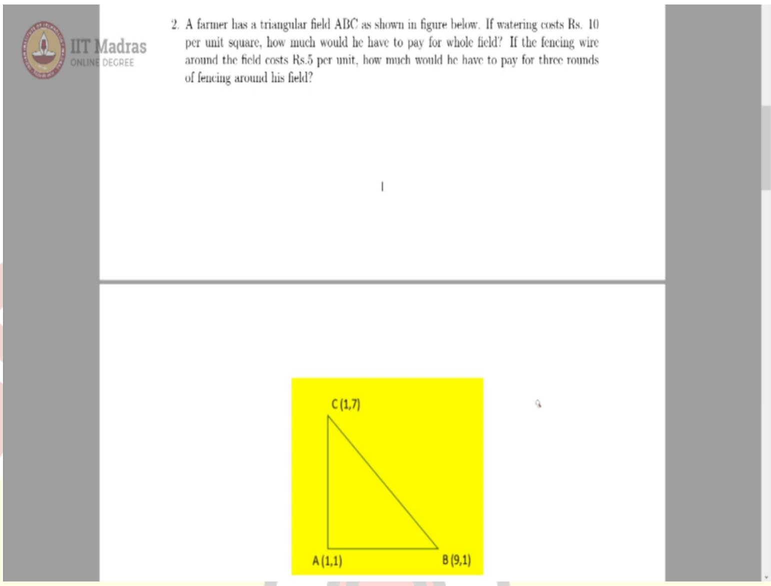


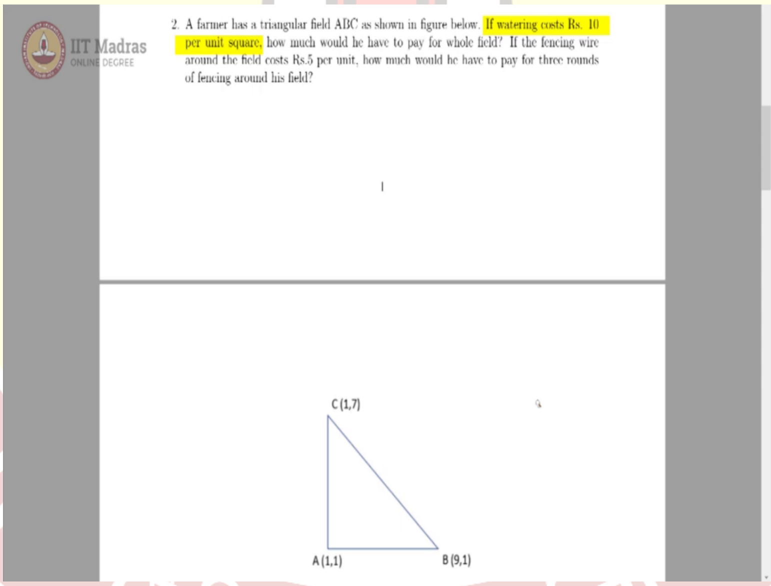
For the part B of this question, it is asked, what is the price of mobile A in December. So, this is the December month, which would be this point here which has a price of 9500, so this is our answer for B. And then in C it has asked, calculate the slope of mobile A's price from January to March 2020, so we want the slope of this segment here and this slope we had already calculated, it was $m = \frac{9750-8750}{10-12} = \frac{1000}{-2} = -500$.

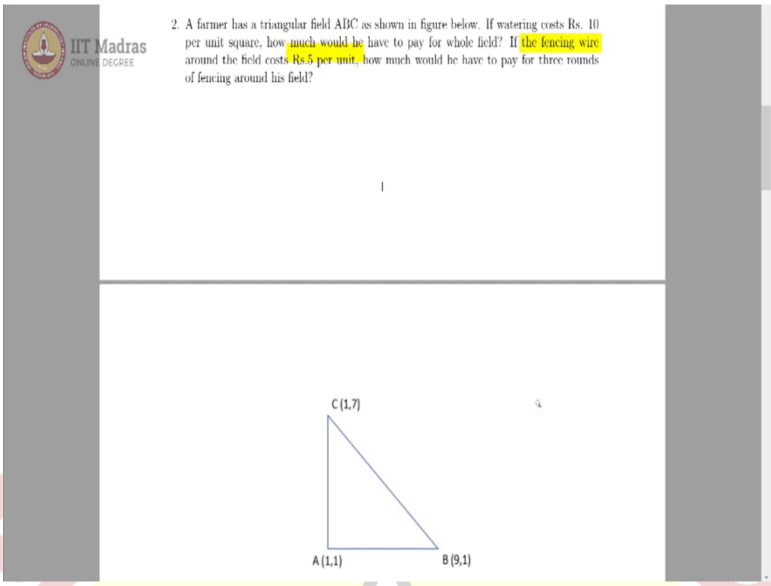
And because of the negative slope you can see that it is a decreasing function, which is what is happening, the price had fallen at 500 per month. Lastly, we have being asked, what is the price of mobile A in March 2020, so this is March 2020, this is a point and we have already found the price which is 8750 that is our part.

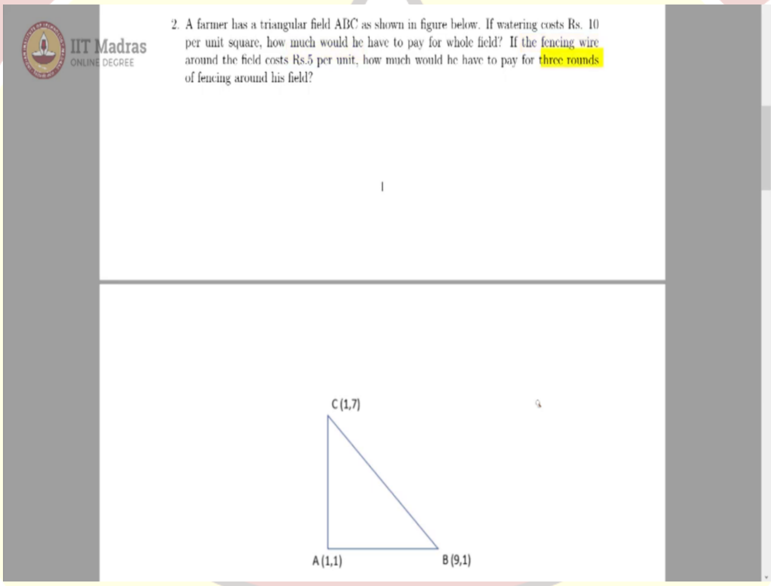
Mathematics for Data Science 1
Indian institute of Technology, Madras
Week 02
Tutorial 02

(Refer Slide Time 00:16)



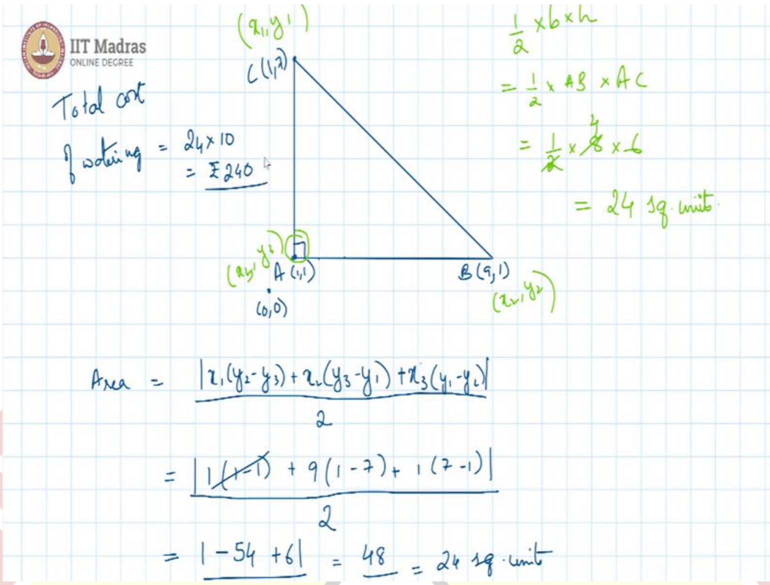






In the second question, the reserved triangular field ABC, whose coordinates are given. And if watering costs rupees 10 per unit square, so they are giving the cost of watering the field, and it is so and so amount per unit square that is area, how much would you have to pay for the whole field? So, we would like to find out the area of the field. And then if the fencing wire around the field costs rupees 5 per unit, how much should he have to pay for 3 rounds of fencing around this field that is find the perimeter, so find the area and perimeter of this particular field.

(Refer Slide Time 01:08)



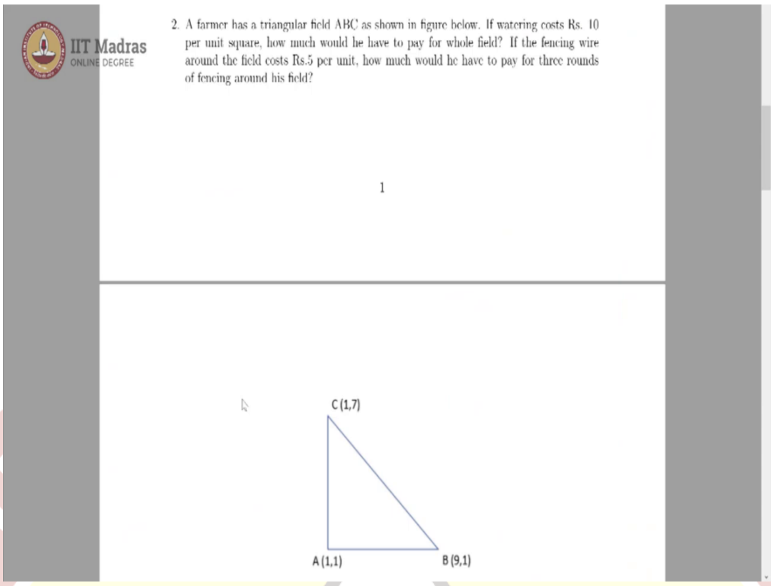
So, if we consider this to be our origin, the triangle is made up of these points, $(1, 1)$, this will be $(9, 1)$ and this is $(1, 7)$. So, this is our triangle, this is A, this is B, and this is C and you can see that AC is completely vertical, its x coordinate remains the same, it is 1, and AB is completely horizontal, its y coordinate remains the same, which is 1.

So, this is a right angled triangle. Now, we could use the area of triangle formula, the area will be $\frac{|x_1(y_2-y_3)+x_2(y_3-y_1)+x_3(y_1-y_2)|}{2}$, which in this case is going to be, you can take any of these points to be $x_1\ y_1$ and the others to, the next one to be $x_2\ y_2$ and x_1 to be $x_3\ y_3$. The, how you choose $x_1\ y_1\ \ x_2\ y_2$ and $x_3\ y_3$ does not matter, the order is what is important. Applying this formula, we get our x_1, y_2, y_3 is 1.

So, $\frac{|1(1-1)+9(1-7)+1(7-1)|}{2} = \frac{|-48|}{2} = \frac{48}{2}$, and that is 24 square unit. However, the same problem could be approached in a slightly different way which is, if I observe that this is a right angle triangle, I could just do half into base into height. And here the base would be the length AB, that is half into AB for which the height would be the length AC.

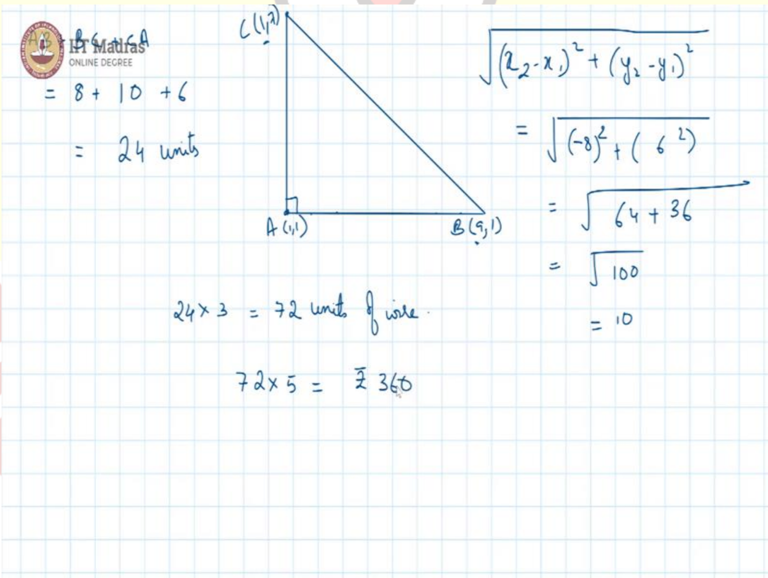
And now since AB is horizontal, you can directly take the length to be $x_2 - x_3$ which is the difference in the x coordinates, so $\frac{1}{2} \times 8 \times 6$ we have 24 square unit. So, this worked out because our triangle is a right-angled triangle. So now the cost of watering is supposed to be the area into cost of watering per square unit which is 10 rupees, so total cost of watering is equal to 24 into 10 that is rupees 240.

(Refer Slide Time 05:18)



For the second part of the question, we require the perimeter of this triangle because fencing is done along the perimeter, and they have to do 3 rounds of fencing at the rate of rupees 5 per unit. So, we first find the perimeter.

(Refer Slide Time 05:39)

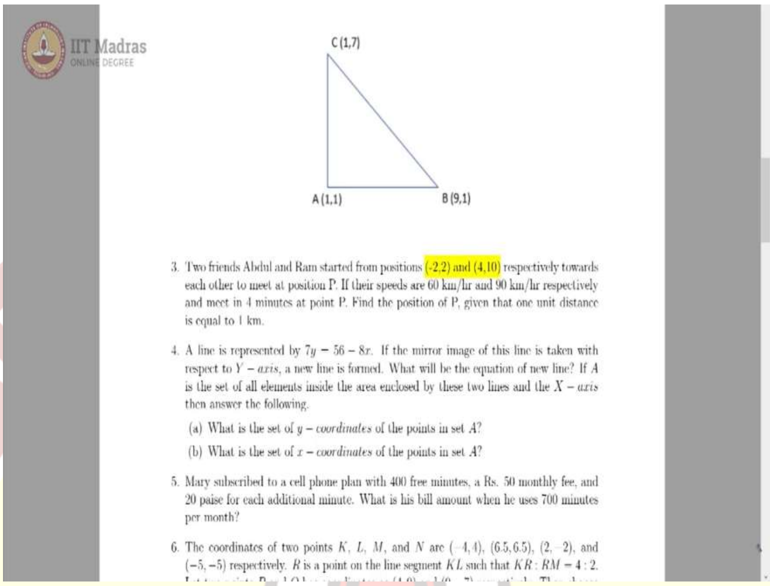


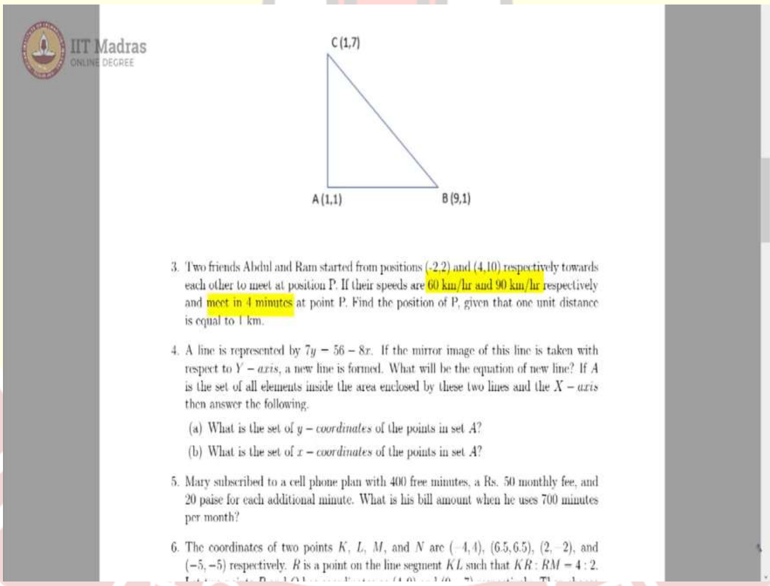
Perimeter would simply be AB plus BC plus CA, which is AB is clear it is 8 units, CA is also clear which is 6 units, BC needs to be figured out and BC we find out using the Euclidean distance, that is the $\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$ that is the square of the difference in x coordinates plus the square of the difference in y coordinates, the whole under root.

So, this gives us $\sqrt{(-8)^2 + 6^2} = \sqrt{64 + 36} = \sqrt{100} = 10$. So, we have 10. So, this quantity is 10 and thus our perimeter is also 24 units and we need wiring for fencing around 3 rounds. So, we will require 24 into 3 is equal to 72 unit of wire and then each unit has been fixed a price of 5 rupees. So, we have 72 into 5 is rupees 360 is the cost of fencing.

Mathematics for Data Science 1
Indian Institute of Technology, Madras
Week 02
Tutorial 03

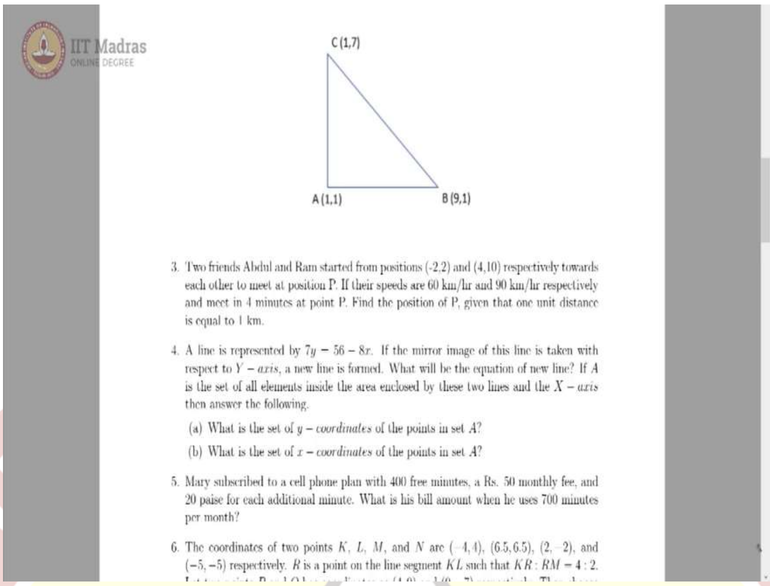
(Refer Slide Time 00:17)





In the third question, the two friends positioned at these two locations and both of them go to a position P. The speeds are given, and the time of their meeting is given, then what should be this position P given that 1unit distance is equal to 1 kilometre. So first let us look at their positions.

(Refer Slide Time 00:55)



3. Two friends Abdul and Ram started from positions (-2,2) and (4,10) respectively towards each other to meet at position P. If their speeds are 60 km/hr and 90 km/hr respectively and meet in 4 minutes at point P. Find the position of P, given that one unit distance is equal to 1 km.

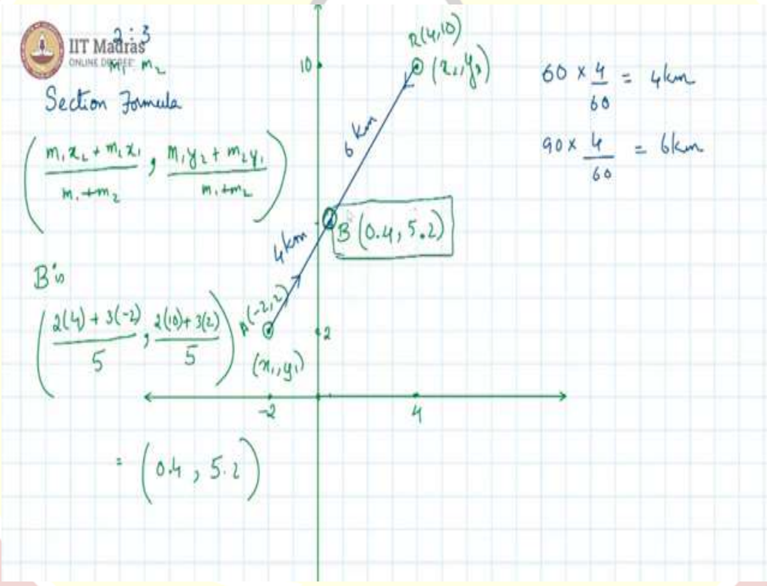
4. A line is represented by $7y - 56 - 8x$. If the mirror image of this line is taken with respect to $Y-axis$, a new line is formed. What will be the equation of new line? If A is the set of all elements inside the area enclosed by these two lines and the $X-axis$ then answer the following.

   (a) What is the set of $y-coordinates$ of the points in set A?

   (b) What is the set of $x-coordinates$ of the points in set A?

5. Mary subscribed to a cell phone plan with 400 free minutes, a Rs. 50 monthly fee, and 20 paise for each additional minute. What is his bill amount when he uses 700 minutes per month?

6. The coordinates of two points K, L, M, and N are $(-4,4)$, $(6.5,6.5)$, $(2,-2)$, and $(-5,-5)$ respectively. R is a point on the line segment KL such that $KR : RM - 4 : 2$.



So, this point is the origin, and now among the 2 friends, 1 Abdul is at (-2,2), so this is -2 here, and this is 2 here. So, Abdul is here, A (-2,2). And we have the other one Ram at (4,10), which is this is 4 on the x axis and this is 10 on the y axis, so Ram is here (4,10). It says they are moving towards each other, so this is a path they take, where Abdul is moving this way and Ram is moving this way.

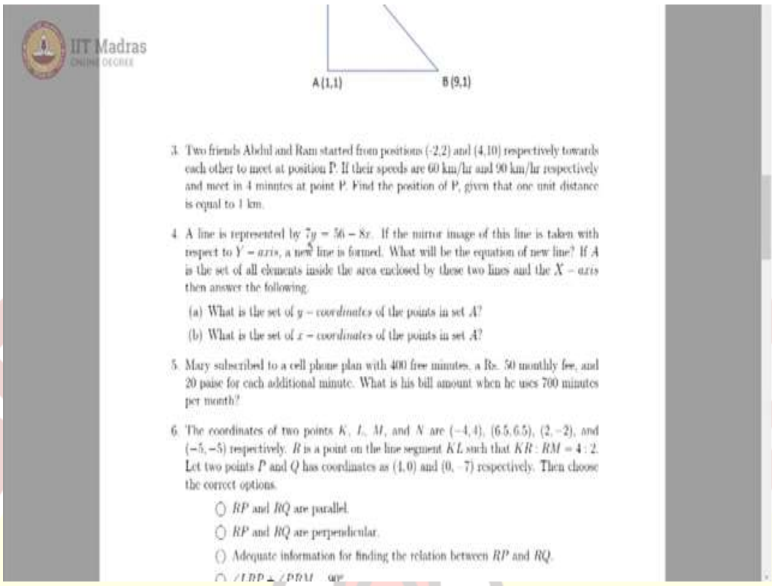
And what we know about their movement is, Abdul is moving at 60 kmph and Ram is moving at 90 kmph, so Ram is faster and they are meeting in 4 minutes. If 1 unit is a kilometre, we have 60×4/60 because it is 4 minutes and the units are in hours kilometre per hour, so we do 4/60 is equal to 4km.

So, Abdul is moving 4 km, whereas Ram is moving 90×4/60, which is 6 km, so they meet somewhere in this region and we would like to know that point. And that point we can achieve through the section formula; we do not actually need to find the distances. And for applying the section formula, what we need to know is the ratio of how this point cuts the line segment AR. And that ratio we can use it in this way.

So, we know that this length is supposed to be 4 km and this length is supposed to be 6 km which means the ratio is 4:6 that is 2:3. So, we now apply the section formula, which is $(m_1x_2+\ m_2x_1)/(m_1+m_1)$. This will be the x coordinate of that point and $(m_1\ y_2+\ m_2\ y_1)/(m_1+m_1)$ will be the y coordinate of that point. So, let us call this point B, so this is the formula for B, so we get the point B is applying m_1 is, this is the ratio $m_1: m_2$ and this is (x_1, y_1) and this is (x_2, y_2)

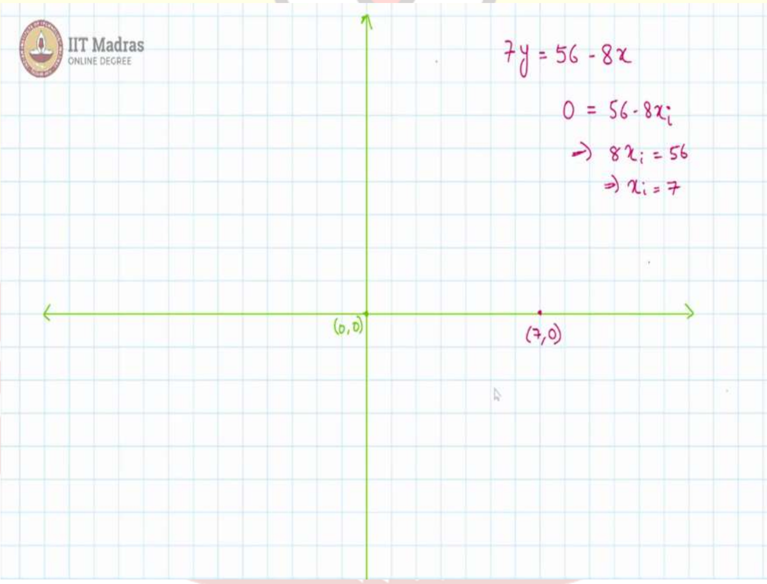
So, we have, m_1x_2 would be 2×4 + m_2x_1 would be 3×(-2) the whole by m_1+m_1 is 5 and (m_1 y_2 would be 2×10 + $m_2\ y_1$would be 3×2)/5 again. So that gives us 8-6 = 2, 2/5 is 0.4, and 2×10 = 20, 3×2 = 6 or 26 /5 = 5.2. So, B is (0.4,5.2). We can check with our intuition, this point that we marked out actually has an x coordinate between 0 and 1 and a y coordinate between 5 and 6. So the point we are looking for is (0.4,5.2).

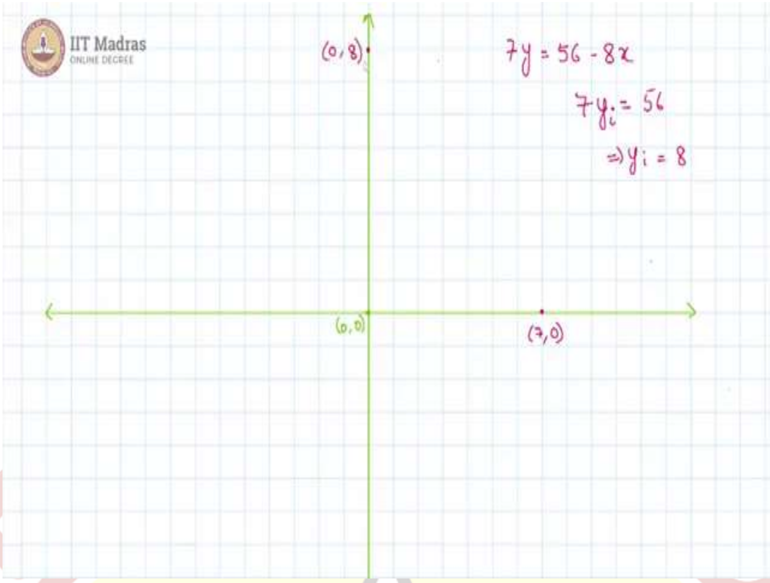
(Refer Slide Time 00:16)



Now, fourth question, there is a line which is represented by 7y = 56 - 8x.

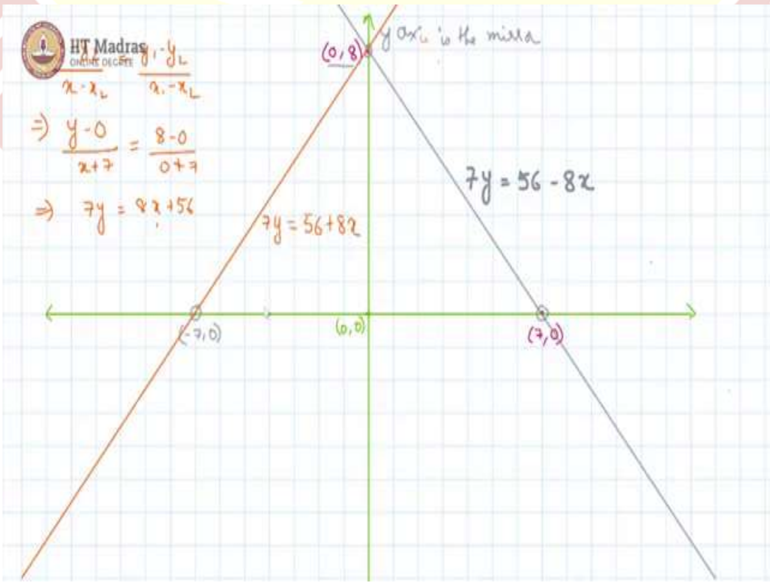
(Refer Slide Time 00:27)





Let us first draw this line, so this is our origin and our line equation is 7y = 56 - 8x. In order to draw this line, in order to find out the curve, we need two points, two points are enough. And the easiest way to find out these two points is to work with the intercepts, that is when this line cuts the X-axis and when it cuts the Y-axis. So when, it is cutting the X-axis, y will be 0, so we just take the Y-coordinate to be 0, and we write 0 = 56 - 8x and to denote that this is the intercept, I am going to call it x_i and that gives us $8x_i$ = 56 and that gives us x_i= 7. So, the x- intercept is 7 which is here. So, (7,0) is one point. And now, for the other point, we take x to be 0 and thus we can say 7y is equal to 56. Again, for the intercept, I am going to use y_i, 56 - 0, therefore y_i, , the y intercept is 8. So, this point here, which is (0,8), this is our y-intercept.

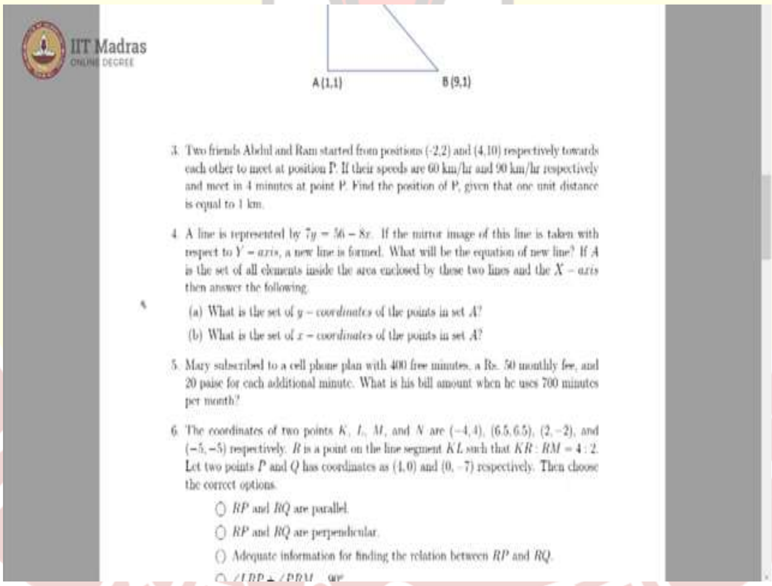
(Refer Slide Time 02:09)

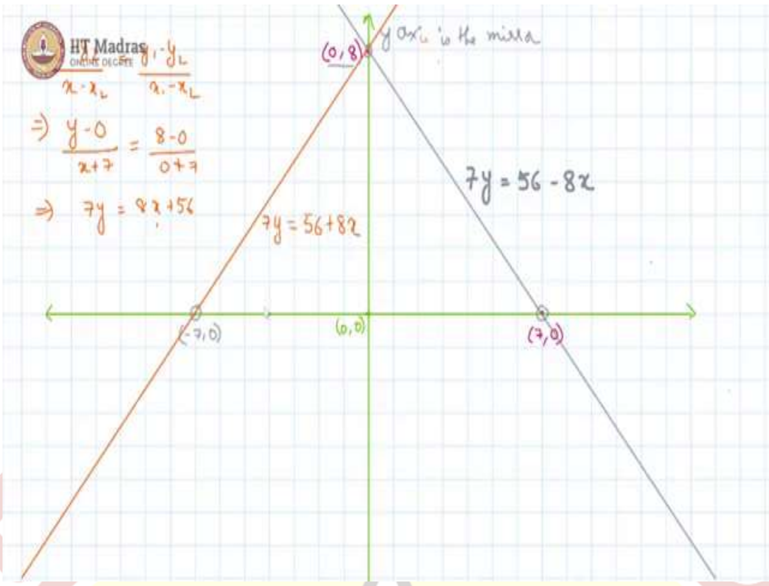


So, this is a straight line, we have been given 7 y = 56 - 8x. It passes through (7,0) and (0,8). Now, for a mirror image, what happens is, and here we are treating the Y-axis as the mirror, so Y-axis is the mirror, you are at the same distance from your mirror as your reflection. So, your reflection will be at the exact distance from the mirror on the opposite side as you, so for example, if we take our (0.7,0) on the other side, which is this point that is (-7,0), that would be the reflection of (7,0) with respect to the Y-axis as the mirror. However, (0,8), since it is already on the Y-axis, its reflection is going to coincide with itself, so this is the other point of the reflection.

And thus, the mirror image for this line is going to be this other line which passes through these two points, (-7,0) and (0,8). For finding the equation of this line, we can use the two point form. And when we apply the values, we get (y – 0) / (x +7) = (8 – 0) / (0 + 7), which gives us 7y = 8 x + 56. So, the mirror image line if you have to write it in the same form as the other one, 7y = 56 + 8 x.

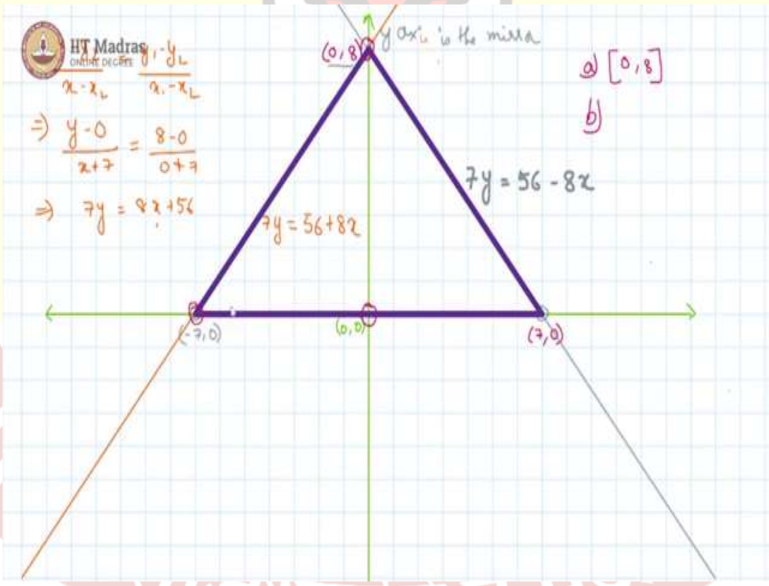
(Refer Slide Time 04:40)

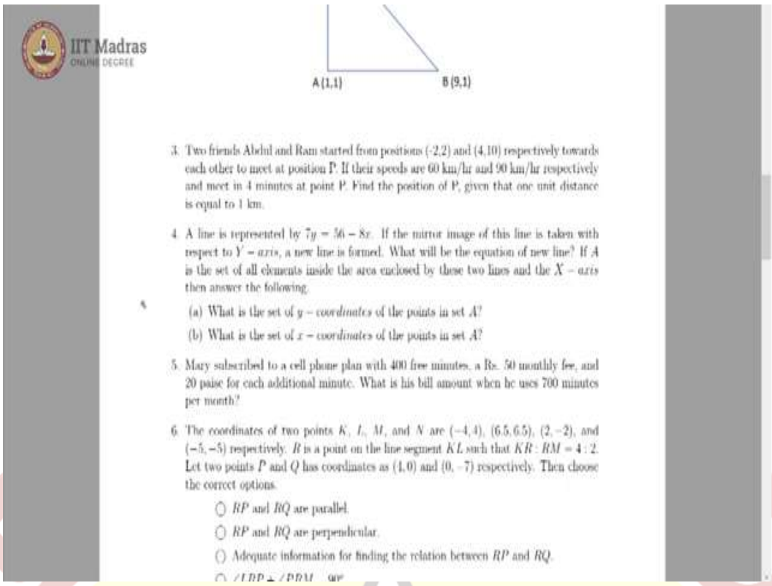




Now, in the next part of the question, they are asking if A is the set of all elements inside the area enclosed by these two lines and the X-axis. So, we are looking at this triangle, and in this triangle, we have being asked what is the set of Y coordinates of the points in set A.

(Refer Slide Time 05:03)





IIT Madras
ONLINE DEGREE

A (1,1) B (9,1)

3. Two friends Abdul and Ram started from positions (-2,2) and (4,10) respectively towards each other to meet at position P. If their speeds are 60 km/hr and 90 km/hr respectively and meet in 4 minutes at point P. Find the position of P, given that one unit distance is equal to 1 km.

4. A line is represented by $7y = 56 - 8x$. If the mirror image of this line is taken with respect to $Y - axis$, a new line is formed. What will be the equation of new line? If A is the set of all elements inside the area enclosed by these two lines and the $X - axis$ then answer the following.

   (a) What is the set of $y - coordinates$ of the points in set A?

   (b) What is the set of $x - coordinates$ of the points in set A?

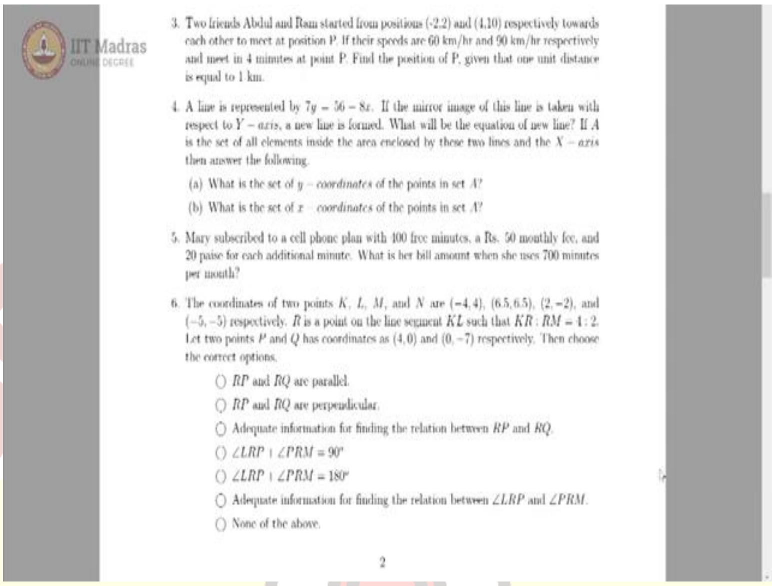
5. Mary subscribed to a cell phone plan with 400 free minutes, a Rs. 50 monthly fee, and 20 paise for each additional minute. What is his bill amount when he uses 700 minutes per month?

6. The coordinates of two points K, L, M, and N are $(-4,4)$, $(6.5,6.5)$, $(2,-2)$, and $(-5,-5)$ respectively. R is a point on the line segment KL such that $KR : RM = 4 : 2$. Let two points P and Q has coordinates as $(4,0)$ and $(0,-7)$ respectively. Then choose the correct options.

   - RP and RQ are parallel.
   - RP and RQ are perpendicular.
   - Adequate information for finding the relation between RP and RQ.

So, all possible Y coordinates in this set. So, every point within this triangle and on the triangle itself count, and as you can clearly see the least Y coordinate here is 0, and the maximum Y coordinate here is 8. So, the set of Y coordinates is going to be the closed interval [0, 8], because we are considering the triangle also to be part of this set, not just the points inside the triangle interior to the triangle, we are considering the triangle also to be part of the set. So, this is the answer for part A. And for part B we have what is the set of X coordinates of the points in set A, and again, we look for the least and the maximum here, the least is -7 and the maximum is 7. And every value in between is there so this would be again the closed interval [-7 ,7].

(Refer Slide Time 00:15)



3. Two friends Abdul and Ram started from positions (-2,2) and (4,10) respectively towards each other to meet at position P. If their speeds are 60 km/hr and 90 km/hr respectively and meet in 4 minutes at point P. Find the position of P, given that one unit distance is equal to 1 km.

4. A line is represented by $7y = 56 - 8x$. If the mirror image of this line is taken with respect to $Y-axis$, a new line is formed. What will be the equation of new line? If A is the set of all elements inside the area enclosed by these two lines and the $X-axis$ then answer the following.
   (a) What is the set of $y-coordinates$ of the points in set A?
   (b) What is the set of $x-coordinates$ of the points in set A?

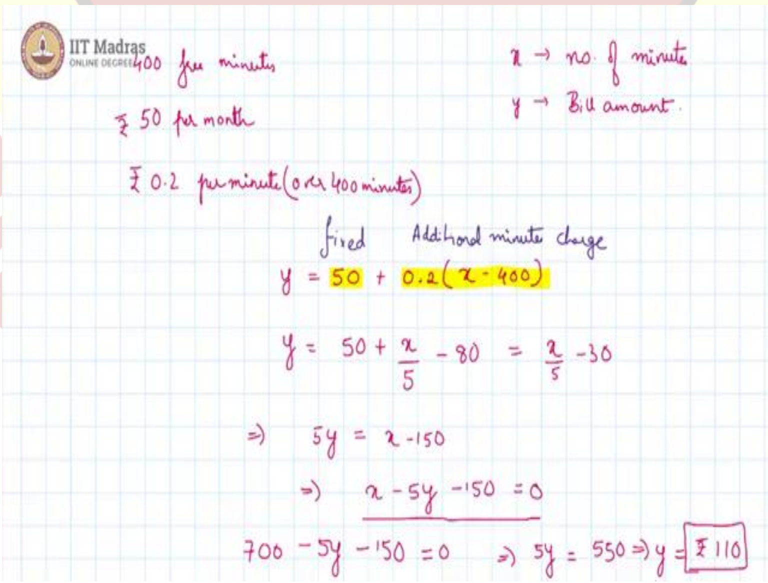
5. Mary subscribed to a cell phone plan with 400 free minutes, a Rs. 50 monthly fee, and 20 paise for each additional minute. What is her bill amount when she uses 700 minutes per month?

6. The coordinates of two points K, L, M, and N are $(-4,4)$, $(6.5,6.5)$, $(2,-2)$, and $(-5,-5)$ respectively. R is a point on the line segment KL such that $KR : RM = 1:2$. Let two points P and Q has coordinates as $(4,0)$ and $(0,-7)$ respectively. Then choose the correct options.
   - ◯ RP and RQ are parallel.
   - ◯ RP and RQ are perpendicular.
   - ◯ Adequate information for finding the relation between RP and RQ.
   - ◯ $\angle LRP + \angle PRM = 90°$
   - ◯ $\angle LRP + \angle PRM = 180°$
   - ◯ Adequate information for finding the relation between $\angle LRP$ and $\angle PRM$.
   - ◯ None of the above.

2

Now 5th problem, Mary has subscribed to a cell phone plan with 400 free minutes, a 50 rupee monthly fee and 20 paisa for every additional minute over 400. And the question is, what is her bill amount if she uses 700 minutes?

(Refer Slide Time 00:42)

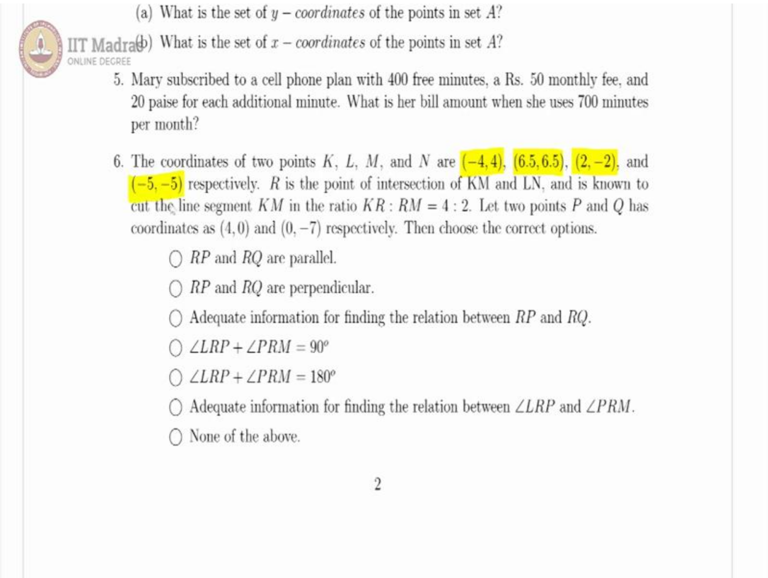


So let us put down our variables here. So there is 400 free minutes and there is a 50 rupee charge per month and we have 20 paisa that is 0.2 rupees per minute over 400 minutes. Now, our independent variable is the number of minutes, the bill is dependent on the number of minutes, so our x variable is number of minutes and the y variable is bill amount. And what we know is for every month the bill amount will always have a 50 rupee charge, and on top of that you are being charged 0.2 for every minute over 400, which means if x is the total number of minutes, then $(x – 400)(0.2)$ will be the charge for the additional minutes.

This is the fixed charge whereas this is the additional minutes charge, so we get a linear equation which is $y = 50 + x/5$(because 0.2 is 1/5) - 80 which is then $(x/5) - 30$. If we simplify it further, we get $5y = x - 150 \Rightarrow x - 5y - 150 = 0$. This is the equation that relates our bill amount to the number of minutes. So, Mary is using 700 minutes per month and we need the bill amount for that. So, if we substitute $x = 700$, we get, $700 - 5y - 150 = 0$, this gives us $5y = 550$ which implies y = Rs. 110. This is the bill amount for Mary.

Mathematics for Data Science 1
Indian Institute of Technology, Madras
Week 02 - Tutorial 06

(Refer Slide Time 00:17)



(a) What is the set of $y-coordinates$ of the points in set A?

(b) What is the set of $x-coordinates$ of the points in set A?

5. Mary subscribed to a cell phone plan with 400 free minutes, a Rs. 50 monthly fee, and 20 paise for each additional minute. What is her bill amount when she uses 700 minutes per month?

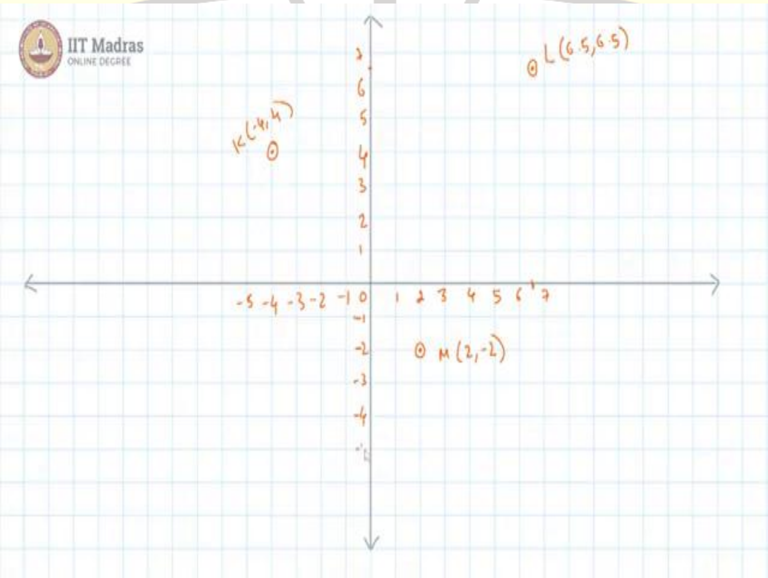
6. The coordinates of two points K, L, M, and N are $(-4, 4)$, $(6.5, 6.5)$, $(2, -2)$, and $(-5, -5)$ respectively. R is the point of intersection of KM and LN, and is known to cut the line segment KM in the ratio $KR : RM = 4 : 2$. Let two points P and Q has coordinates as $(4, 0)$ and $(0, -7)$ respectively. Then choose the correct options.

- ○ RP and RQ are parallel.
- ○ RP and RQ are perpendicular.
- ○ Adequate information for finding the relation between RP and RQ.
- ○ $\angle LRP + \angle PRM = 90°$
- ○ $\angle LRP + \angle PRM = 180°$
- ○ Adequate information for finding the relation between $\angle LRP$ and $\angle PRM$.
- ○ None of the above.

2

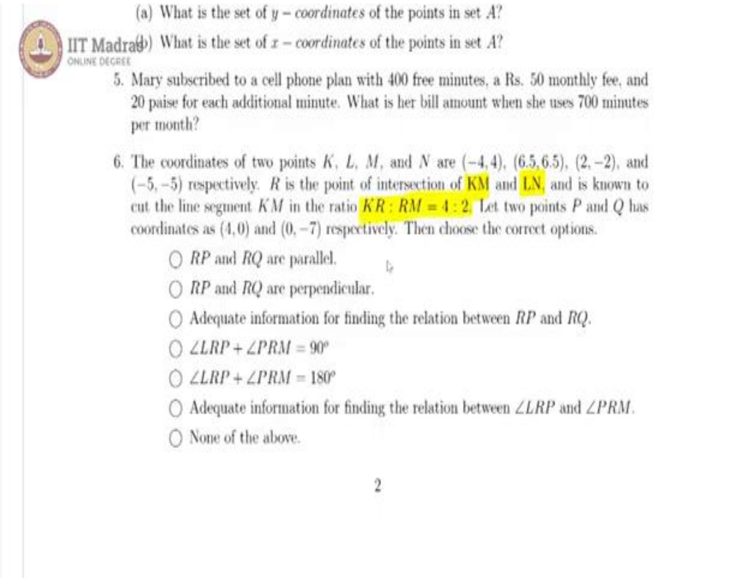
For our 6th problem we have these 4 points given to us. Let us first plot them out on a graph.

(Refer Slide Time 00:32)



And this will be 0, we have a (-4, 4), so this is -1, this is -2, this is -3, this is -4, this is 1, 2, 3 and 4, so this point here is our K (-4, 4). And then we have (6.5, 6.5), this is 1, 2, 3, 4, 5, 6, 7, this here is 6.5 and 5, 6 and 7, this here is 6.5, here we are with L (6.5, 6.5), then we have a (2, -2), -1, this is -2. So this point here is our M (2, -2). And lastly, we have (-5, -5), -4 and -5, this is our point.

(Refer Slide Time 01:48)



(a) What is the set of y − *coordinates* of the points in set A?

(b) What is the set of x − *coordinates* of the points in set A?

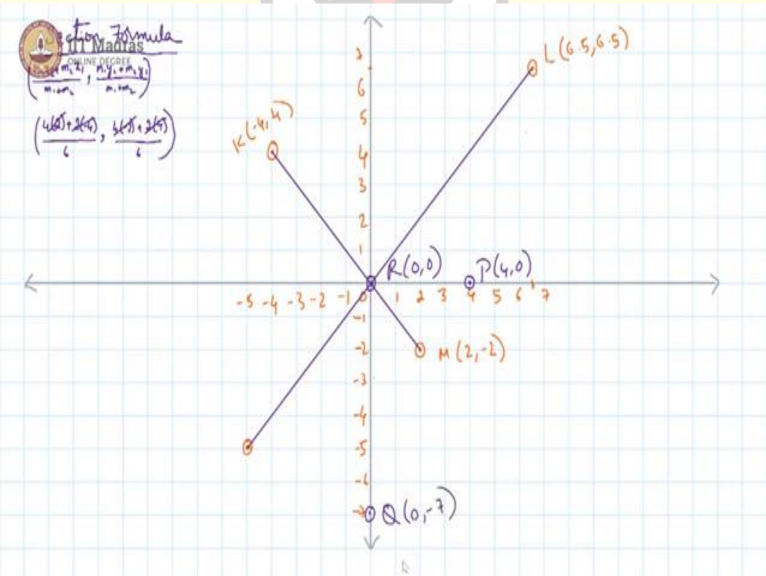
5. Mary subscribed to a cell phone plan with 400 free minutes, a Rs. 50 monthly fee, and 20 paise for each additional minute. What is her bill amount when she uses 700 minutes per month?

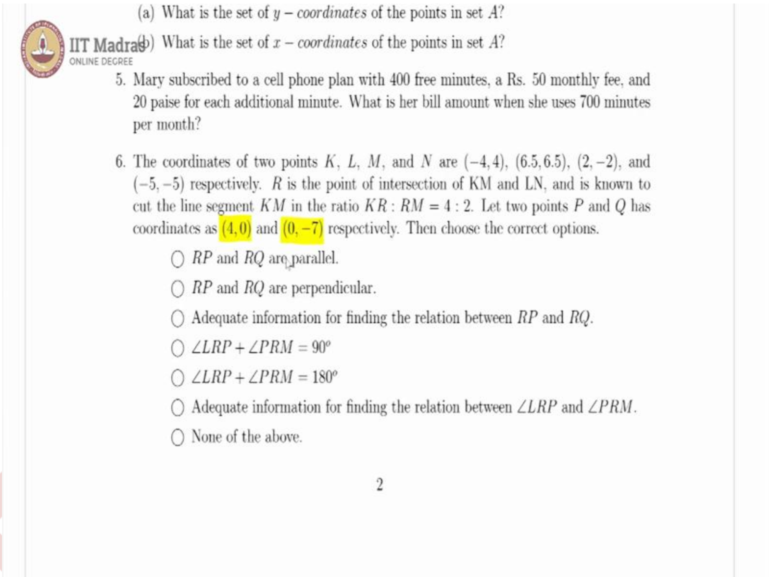
6. The coordinates of two points K, L, M, and N are $(-4, 4)$, $(6.5, 6.5)$, $(2, -2)$, and $(-5, -5)$ respectively. R is the point of intersection of KM and LN, and is known to cut the line segment KM in the ratio $KR : RM = 4 : 2$. Let two points P and Q has coordinates as $(4, 0)$ and $(0, -7)$ respectively. Then choose the correct options.
   - ○ RP and RQ are parallel.
   - ○ RP and RQ are perpendicular.
   - ○ Adequate information for finding the relation between RP and RQ.
   - ○ $\angle LRP + \angle PRM = 90°$
   - ○ $\angle LRP + \angle PRM = 180°$
   - ○ Adequate information for finding the relation between $\angle LRP$ and $\angle PRM$.
   - ○ None of the above.

2

Now we are told that R is the point of intersection of KM and LN, and it is known to cut the line segment KM in this ratio, 4 is to 2 ratio, so let us identify R.

(Refer Slide Time 02:04)





(a) What is the set of $y-coordinates$ of the points in set A?

(b) What is the set of $x-coordinates$ of the points in set A?

5. Mary subscribed to a cell phone plan with 400 free minutes, a Rs. 50 monthly fee, and 20 paise for each additional minute. What is her bill amount when she uses 700 minutes per month?

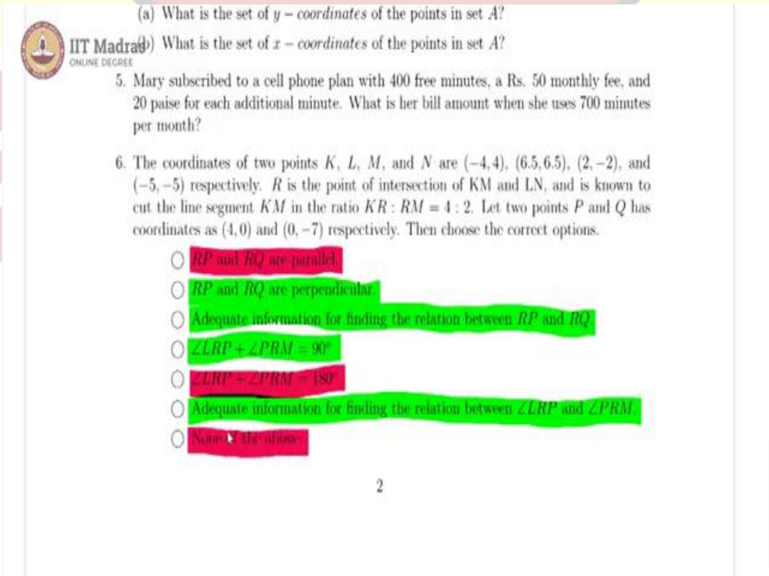
6. The coordinates of two points K, L, M, and N are $(-4,4)$, $(6.5,6.5)$, $(2,-2)$, and $(-5,-5)$ respectively. R is the point of intersection of KM and LN, and is known to cut the line segment KM in the ratio $KR : RM = 4 : 2$. Let two points P and Q has coordinates as $(4,0)$ and $(0,-7)$ respectively. Then choose the correct options.

- ○ RP and RQ are parallel.
- ○ RP and RQ are perpendicular.
- ○ Adequate information for finding the relation between RP and RQ.
- ○ $\angle LRP + \angle PRM = 90°$
- ○ $\angle LRP + \angle PRM = 180°$
- ○ Adequate information for finding the relation between $\angle LRP$ and $\angle PRM$.
- ○ None of the above.

2

So from our diagram, it appears to be the origin. Lets verify this, so we need this to be in the ratio of 4:2. So when we use the section formula, which is the coordinates of a point cutting a line segment in a ratio, $m_1 : m_2$ would be this, $(m_1x_2 + m_2x_1) /(m_1 + m_2)$. And then we have $(m_1y_2 + m_2y_1) /(m_1 + m_2)$. So in this context, R is going to be ((4(2) + 2 (-4))/ 6 , (4(-2) + 2(4))/6). And these 2 cancel out because it is 8 - 8, these two also cancel because -8 +8. So it is true, R the point is the origin. Moving on then, we have two other points, P and Q given to be (4, 0), (0, -7), so these are on the axis. So this point here is P(4, 0), and this is -6, this is -7 so this point here would become Q, which is (0, -7).

(Refer Slide Time 03:58)



(a) What is the set of $y-coordinates$ of the points in set A?

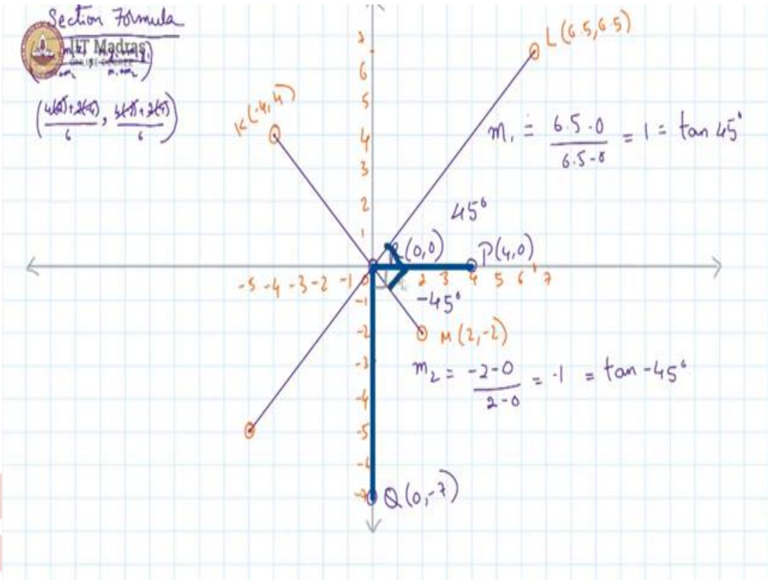
(b) What is the set of $x-coordinates$ of the points in set A?

5. Mary subscribed to a cell phone plan with 400 free minutes, a Rs. 50 monthly fee, and 20 paise for each additional minute. What is her bill amount when she uses 700 minutes per month?

6. The coordinates of two points K, L, M, and N are $(-4,4)$, $(6.5,6.5)$, $(2,-2)$, and $(-5,-5)$ respectively. R is the point of intersection of KM and LN, and is known to cut the line segment KM in the ratio $KR : RM = 4 : 2$. Let two points P and Q has coordinates as $(4,0)$ and $(0,-7)$ respectively. Then choose the correct options.

- ○ RP and RQ are parallel.
- ○ RP and RQ are perpendicular.
- ○ Adequate information for finding the relation between RP and RQ.
- ○ $\angle LRP + \angle PRM = 90°$
- ○ $\angle LRP + \angle PRM = 180°$
- ○ Adequate information for finding the relation between $\angle LRP$ and $\angle PRM$.
- ○ None of the above.

2



Lets look at the options, RP and RQ are parallel, this is one option, lets verify. Now clearly, this is 90°, PQ, PR is perpendicular to RQ and not parallel. So this is definitely wrong and this is definitely right. Is there adequate information for finding the relation between RP and RQ? Yes, we have just found the relation, so there has been adequate information.

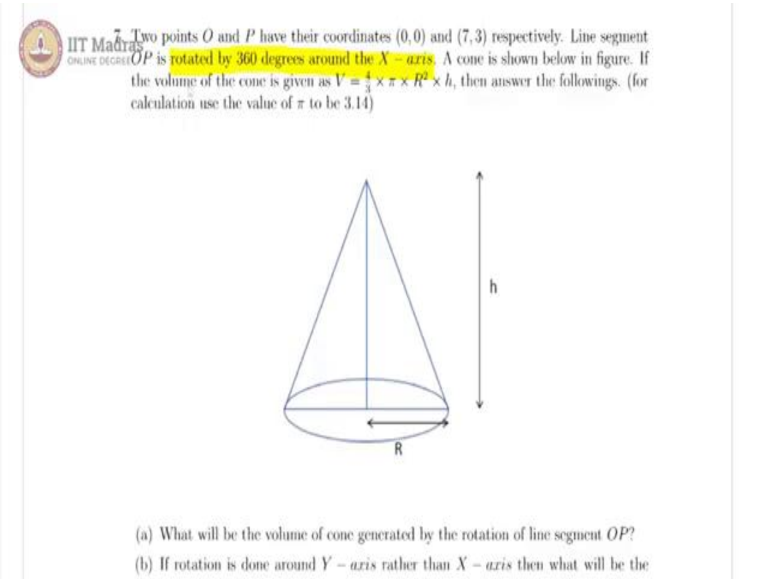
Now let us look at ∠LRP +∠PRM. So we are interested in this angle plus ∠PRM. So this sum is the total ∠LRM, so we need to know what is the angle between LR and RM. So let us look at the slope of LR. So this slope if I call it m_1, this is equal to (6.5 - 0) /(6.5 - 0), which is 1, which is basically tan 45°, so this angle here it is 45°.

And now let us look at this angle here, which is PRM. Then, if we look at the slope here, which is m_2 that is (-2 - 0)/(2 - 0), which is -1, which is equal to tan -45°, therefore this angle here is -45° because we are going clockwise from the horizontal. So in sum, we know that ∠LRM is 45°+ 45°, leading us to see that this is 90°.

So this is true, which means the following statement is false, so this is false. And here we have again adequate information for finding the relation between LRP and PRM. Clearly, we have four options correct, so none of the above is not correct.

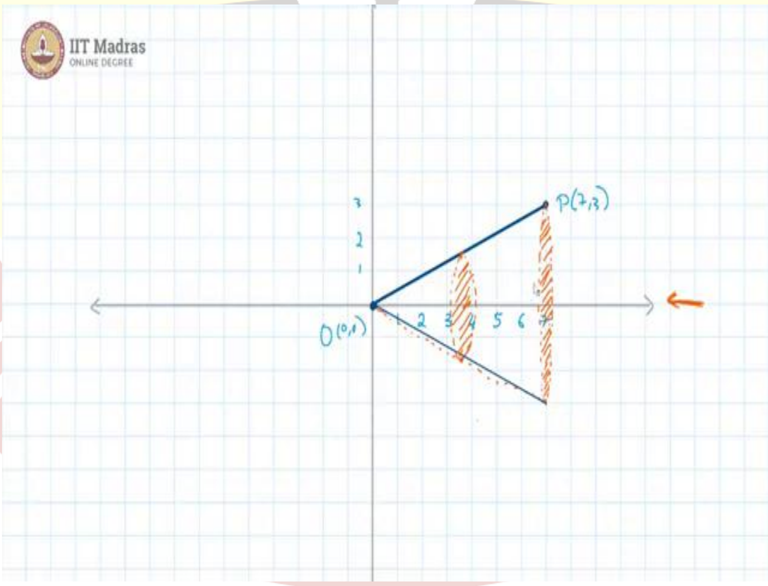
Mathematics for Data Science 1
Indian Institute of Technology, Madras
Week 02 - Tutorial 07

(Refer Slide Time 00:15)



In the 7th question we have two points, one is the origin O, and some other point P (7, 3). And this line segment OP is being rotated. So first, lets mark out these points.

(Refer Slide Time 00:32)

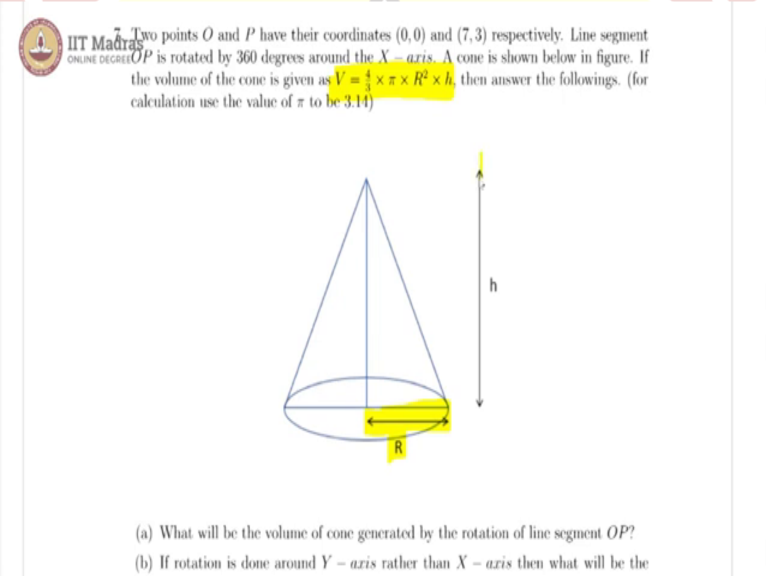


So first lets mark out these points, we have this is the origin and this is 1, 2, 3, 4, 5, 6, 7, this is 1, 2, 3. So our point P is here, this is P (7, 3) and this is the origin of course O. And we have this line segment OP given to us. Now, OP is being rotated by 360° about the x axis, lets see what that means. So every point on OP is going around the x axis in a circle, that is what rotation is, rotation is a combined circular motion of many particles, here we have this point let us take P, P goes around the x axis reaching this bottom point, then it circles back to itself.

So you would see the circle if you looked at it from the right. From the screen's perspective, this is what it will look like. And this is the case with every point on this line.

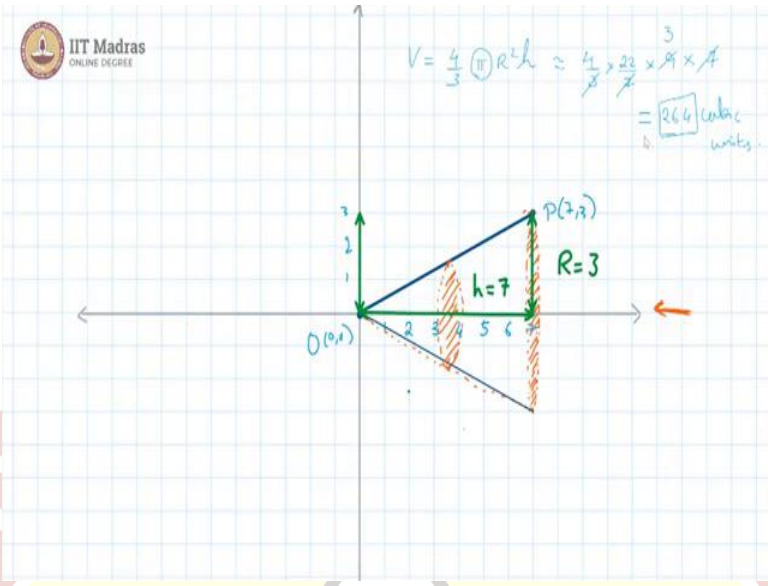
Suppose I took this point, this is just oppositely going to go till here in this circle and return back. So, every point is doing this circle, which means on this side we actually have the mirror image of OP with respect to the x axis which looks something like this. So, we have these circles being formed due to the rotation and as you can see, the final shape it appears to be a cone that is what has happened. Take a line segment and you rotate it about some central axis, you obtain a cone about that central axis.

(Refer Slide Time 02:53)



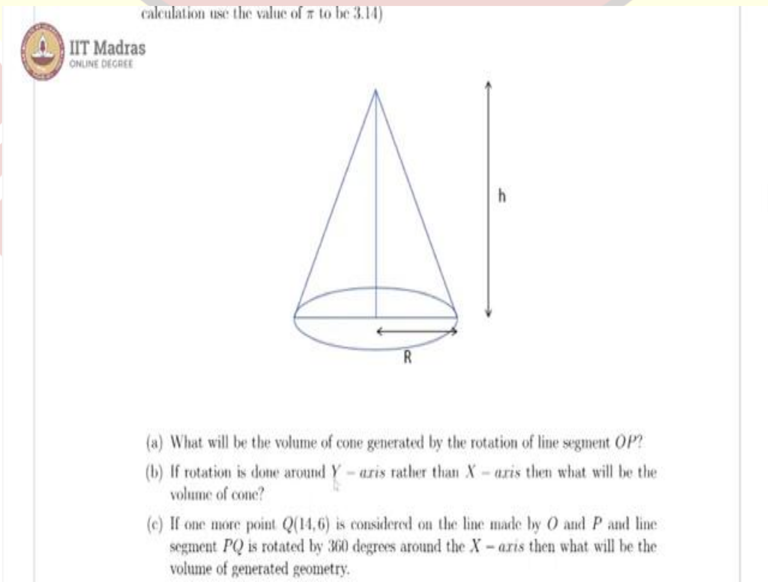
And they have given us the volume of a cone, volume of a cone is $(4/3)\pi R^2 h$, where R is the radius of the base circle and h is the height of the cone.

(Refer Slide Time 03:16)



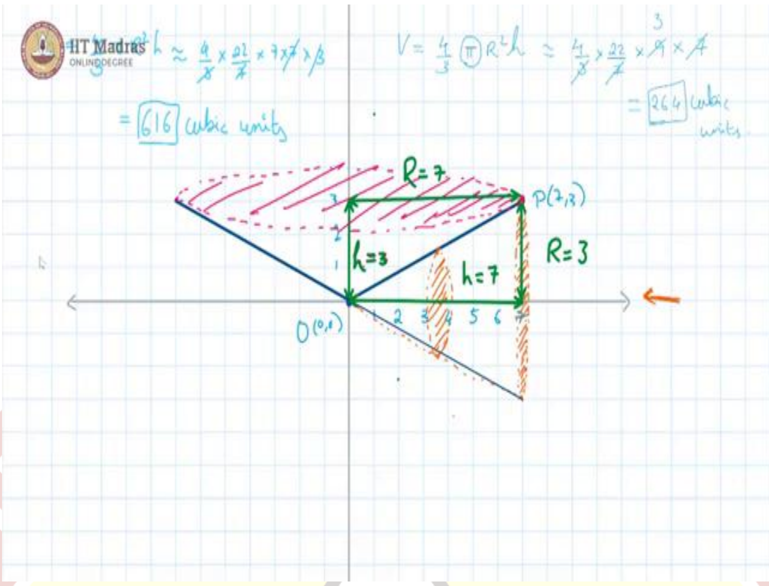
So in the cone we have obtained the radius, base radius is this quantity, which is R = 3 because it is a y coordinate of the point P that is a distance of point P from the x axis. And likewise, if you observe the height of this cone, that quantity h = 7 because the x coordinate of point. In this way, we can obtain the volume of our cone using the formula that is given V= $(4/3)\pi R^2 h$. We are going to approximate pi to be 3.14 or 22 by 7. So, this is roughly equal to (4/3)(22/7)(9)(7) so 7 and 7 cancels of, 3 and 9 gives us 3. So we get 264 cubic units, so this is our volume of the cone.

(Refer Slide Time 04:54)



Now in the second part of this question, it is being said that the rotation is done around the y axis instead of the x axis, so what will this look like?

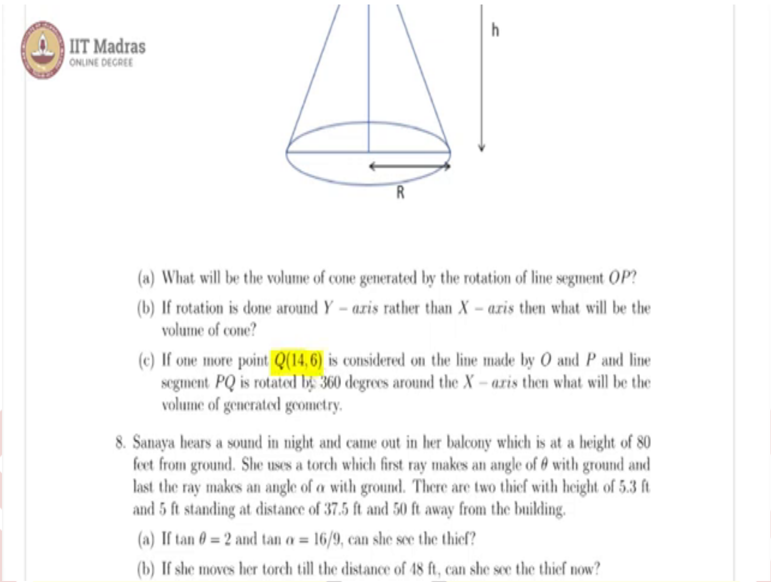
(Refer Slide Time 05:08)



So, this OP is going to have a mirror image about the y axis which is going to look like this. So that means, our point P is going around in a circle to reach this opposite point here and it is coming back to itself. So this would become the base circle for our new cone which is obtained by rotation about the y axis. Now as you can see, this value is already the height so height is 3 now, whereas the radius is basically 7. So, our height is 3 and radius is 7 so these values have changed. So, if we call this quantity, this volume to be V_2, V_2 is $(4/3)\pi R^2 h$, which will be roughly equal to (4/3)(22/7)(7)(7)(3). So, 3 and 7 cancel off here and we get this is equal to 616 cubic units. So, this is the volume if OP is rotated about the y axis 360°, so that cone's volume is 616 cubic units.

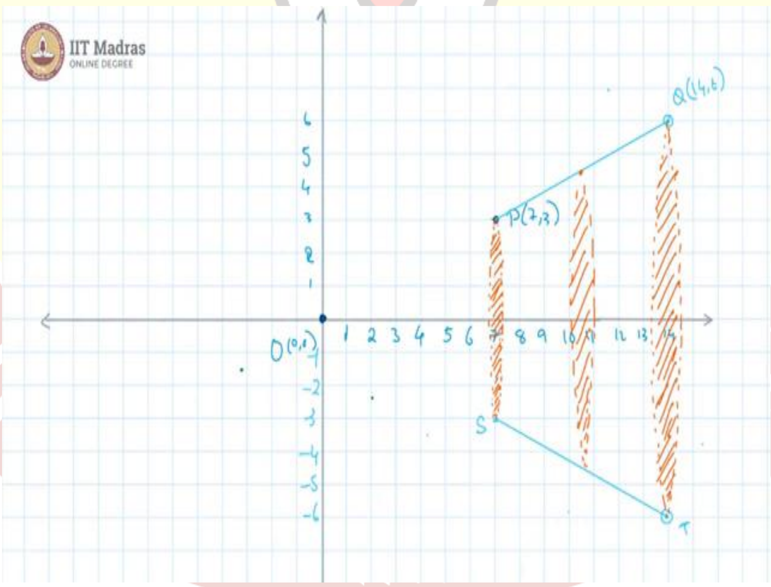
सिद्धिर्भवति कर्मजा

(Refer Slide Time 06:58)



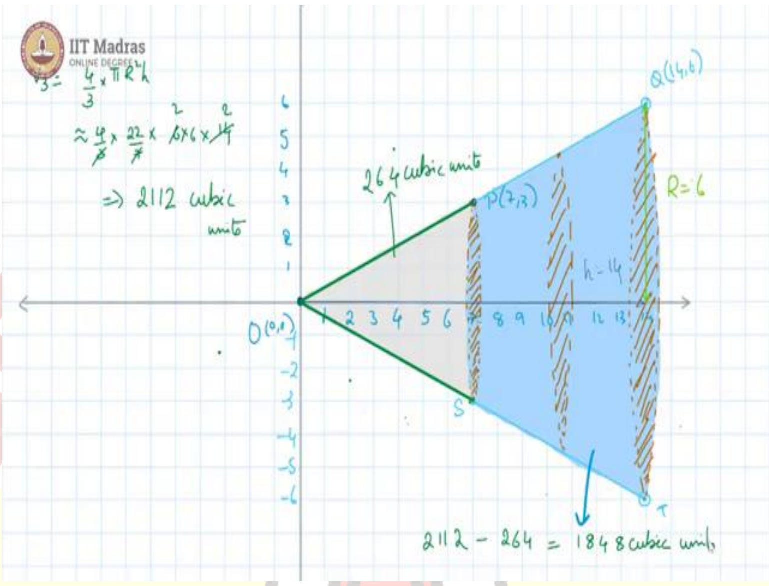
This problem gets progressively more complex, we are now adding the new point (14, 6) which is along the line segment OP, it is on the extension of OP, and then PQ is rotated about x axis by 360°

(Refer Slide Time 07:25)



So, lets see what is happening here, so our point (14, 6) Q, (14, 6) is here, which gives us PQ as this line segment. And now, they are saying that PQ is being rotated about the x axis which will result in the mirror image in this way. This is -1, -2, -3, -4, -5, and -6. So, we are here now, I think we can call this point T and this point is S. So, we have ST in this way again, so what we see here the rotated geometry. So, for reference we are going to take one more point here, which kind of moves around. So, what we are seeing here this is what is called the frustum of a cone. This is a cut-off portion of a larger cone which would be QO rotated about the x axis.

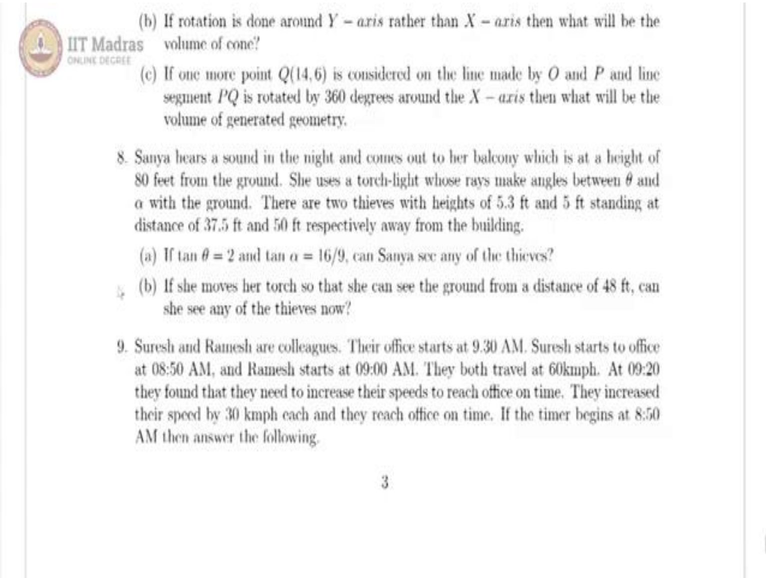
(Refer Slide Time 08:56)



So, thus as you can see, the volume we require is the frustum of the cone which is this region and this volume is the result of subtracting this volume from the total cone. So, we already know this volume OP as that cone's volume to be 264 cubic units. So, the blue shaded region that would be the volume of the large cone that is of OQ rotating about x axis and that we can calculate as $V_3 = (4/3)\pi R^2 h$, where this is approximately equal to (4/3)(22/7)(6)(6)(14). So, 3 one's 3 two's, 7 one's 7 two's, so we have 2,112 cubic units. So the volume we require is going to be 2,112 – 264 = 1,848 cubic units.

Mathematics for Data Science 1
Week 02
Tutorial 01

(Refer Slide Time 00:14)



(b) If rotation is done around $Y-axis$ rather than $X-axis$ then what will be the volume of cone?

(c) If one more point $Q(14,6)$ is considered on the line made by O and P and line segment PQ is rotated by 360 degrees around the $X-axis$ then what will be the volume of generated geometry.

8. Sanya hears a sound in the night and comes out to her balcony which is at a height of 80 feet from the ground. She uses a torch-light whose rays make angles between θ and α with the ground. There are two thieves with heights of 5.3 ft and 5 ft standing at distance of 37.5 ft and 50 ft respectively away from the building.

(a) If $\tan\theta = 2$ and $\tan\alpha = 16/9$, can Sanya see any of the thieves?

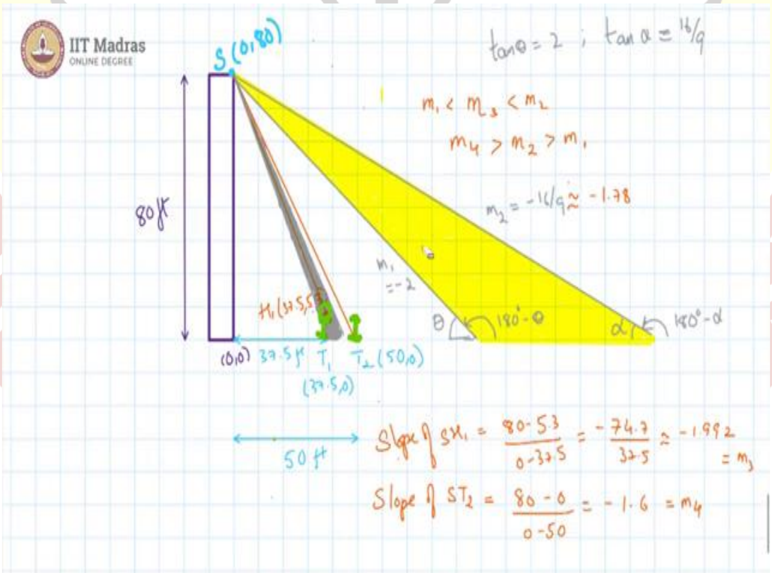
(b) If she moves her torch so that she can see the ground from a distance of 48 ft, can she see any of the thieves now?

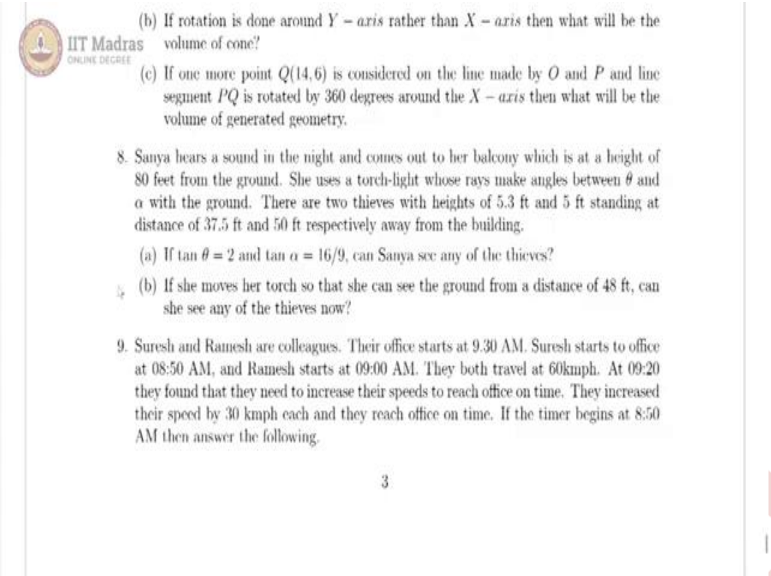
9. Suresh and Ramesh are colleagues. Their office starts at 9.30 AM. Suresh starts to office at 08:50 AM, and Ramesh starts at 09:00 AM. They both travel at 60kmph. At 09:20 they found that they need to increase their speeds to reach office on time. They increased their speed by 30 kmph each and they reach office on time. If the timer begins at 8:50 AM then answer the following.

3

The 8th problem is pretty interesting. So, we have Sania who hears a sound in a night, and she comes out to her balcony, which is at a height of 80 feet from the ground.

(Refer Slide Time 00:29)





IIT Madras
ONLINE DEGREE

(b) If rotation is done around $Y - axis$ rather than $X - axis$ then what will be the volume of cone?

(c) If one more point $Q(14, 6)$ is considered on the line made by O and P and line segment PQ is rotated by 360 degrees around the $X - axis$ then what will be the volume of generated geometry.

8. Sanya hears a sound in the night and comes out to her balcony which is at a height of 80 feet from the ground. She uses a torch-light whose rays make angles between θ and α with the ground. There are two thieves with heights of 5.3 ft and 5 ft standing at distance of 37.5 ft and 50 ft respectively away from the building.

(a) If $\tan\theta = 2$ and $\tan\alpha = 16/9$, can Sanya see any of the thieves?

(b) If she moves her torch so that she can see the ground from a distance of 48 ft, can she see any of the thieves now?

9. Suresh and Ramesh are colleagues. Their office starts at 9.30 AM. Suresh starts to office at 08:50 AM, and Ramesh starts at 09:00 AM. They both travel at 60kmph. At 09:20 they found that they need to increase their speeds to reach office on time. They increased their speed by 30 kmph each and they reach office on time. If the timer begins at 8:50 AM then answer the following.

3

So, let this be our tower, which has a height 80 feet. So, if we take this point to be origin (0, 0), Sania is here, which would be (0, 80). And she uses a torch light which makes angles θ and α with the ground, so the rays from the torch light make angles between these two.

So, this angle here, this is θ and this angle here it is α . And the two thieves, their heights are given and they are standing at these distances from the buildings. So, thief T_1 is somewhere here and T_2 is here, what is given to us is this distance is 37.5. So, T_1 is (37.5, 0), and this distance is 50 feet, so this is 37.5 feet, this is 50 feet. So, T_2 will be the point (50, 0). And we are also given to understand that T_1 is standing at a certain height, T_2 is standing at a certain height, which are roughly the same; one is 5 feet, the other is 5.3 feet.

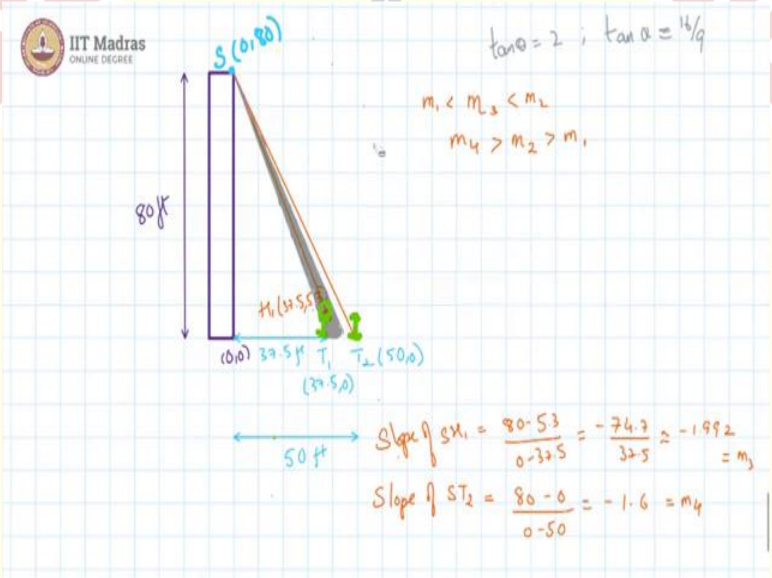
In our diagram, we have drawn the rays of light as though they are passing away from the 2 thieves, however that we need to find out. So, if $\tan\theta$ is 2, and $\tan\alpha$ is 16/9, can Sania see any of the thieves ? so it is given to us that $\tan\theta = 2$ whereas, $\tan\alpha = 16/9$.

So, that means we can find the slope of this line, which is the lowest ray of the torch, and this line which is the farthest ray from the torch, and these slopes would be m_1 =-2 and the minus is because the standard angle here, which is angle from the posture x axis is actually $180^o - \theta$. As you can see, it is clearly a line with the negative slope.

Likewise, this also is $180^o - \alpha$, thus this slope $m_2 = -16/9$. In our diagram, we have drawn it as though the 2 thieves are safe. But this is only a rough schematic diagram, we did not draw θ and α accurately. What we need to do now is to check if the line from Sania to the head of thief 1 or the line from Sanya to the foot of thief 2. If these 2 lines have slopes between m_1 and m_2, then the 2 thieves are likely to be seen. So, we need to calculate these slopes, let us call the head of thief 1 as H_1 and that point will be (37.5, 5.3). So, slope of SH_1 = (80-5.3)/(0-37.5), which is -(74.7 / 37.5), which is roughly -1.992.

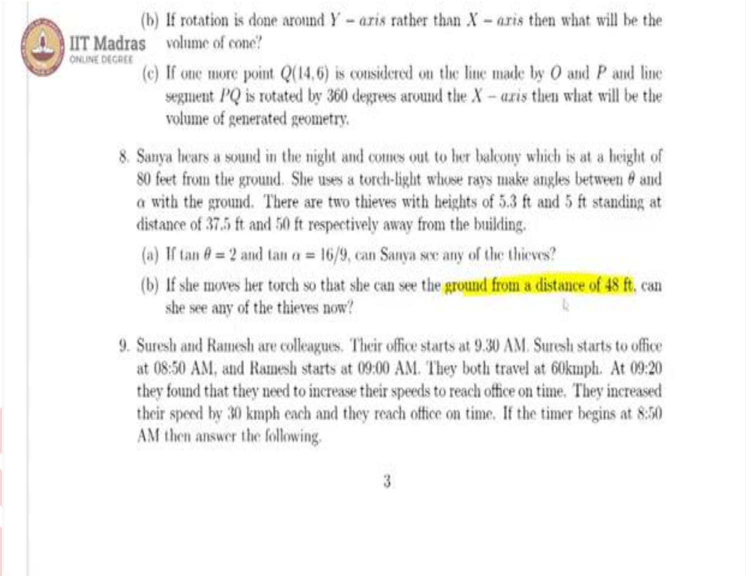
And slope of ST_2, which is to the foot of thief 2 is (80 - 0) / (0 - 15), which is equal to -1.6. So, I want to call this m_3 and this is m_4. And here m_2 is roughly equal to -1.78. So, clearly m_3 is greater than m_1 and lesser than m_2, but m_4 is greater than m_2 and also greater than m_1 which means m_4 that is the foot of thief 2 is not visible to Sania, the actual light cone looks something like this. So, thus we can say the head of thief 1 is visible whereas thief 2 is not visible.

(Refer Slide Time 7:19)



And this light cone that we have drawn earlier it is wrong.

(Refer Slide Time 7:28)


IIT Madras
ONLINE DEGREE

(b) If rotation is done around $Y-axis$ rather than $X-axis$ then what will be the volume of cone?

(c) If one more point $Q(14,6)$ is considered on the line made by O and P and line segment PQ is rotated by 360 degrees around the $X-axis$ then what will be the volume of generated geometry.

8. Sanya hears a sound in the night and comes out to her balcony which is at a height of 80 feet from the ground. She uses a torch-light whose rays make angles between θ and α with the ground. There are two thieves with heights of 5.3 ft and 5 ft standing at distance of 37.5 ft and 50 ft respectively away from the building.

(a) If $\tan\theta = 2$ and $\tan\alpha = 16/9$, can Sanya see any of the thieves?

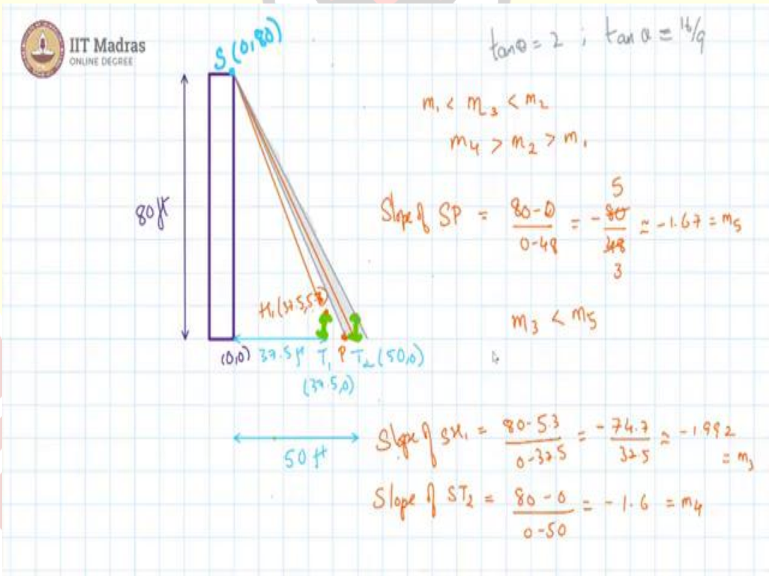
(b) If she moves her torch so that she can see the ground from a distance of 48 ft, can she see any of the thieves now?

9. Suresh and Ramesh are colleagues. Their office starts at 9.30 AM. Suresh starts to office at 08:50 AM, and Ramesh starts at 09:00 AM. They both travel at 60kmph. At 09:20 they found that they need to increase their speeds to reach office on time. They increased their speed by 30 kmph each and they reach office on time. If the timer begins at 8:50 AM then answer the following.

3

Now, she moves her torch so that she can see the ground from a distance of 48 feet. Can she see thieves or not?

(Refer Slide Time: 07:32)



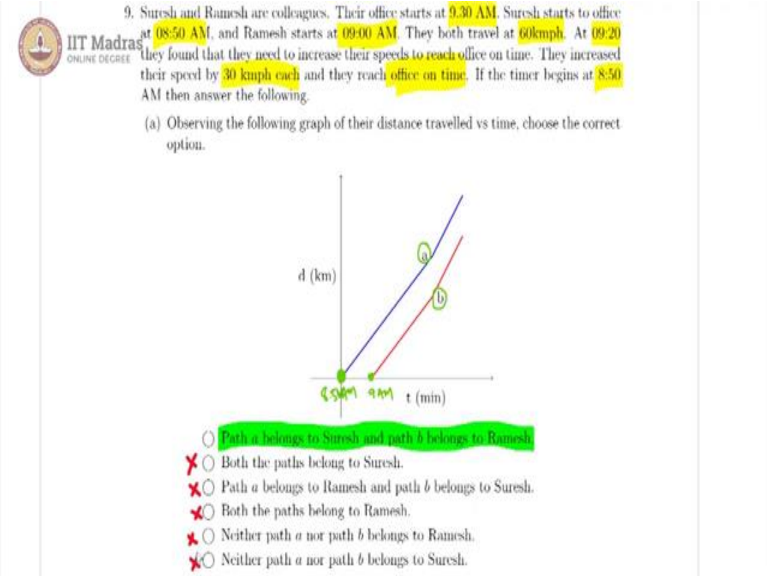
That would mean she is able to see from some point here to some point beyond. From the diagram, it is pretty clear that thief 2 is going to be visible, we do not know if thief 1 will be visible though we have to check for thief 1's head. So, this point that we are talking about, which is let us call it P gives us a slope with S as SP, the slope is equal to (80 - 0) / (0 - 48) because point P is basically (48, 0).

So, that gives us - (80 / 48) which is divisible by 16, both of them are divisible by 16. This would be 5 and this would be 3, so this is roughly -1.67. That would give us m_3 is let us call this now m_5. m_3 is lesser than m_5. And that means the head of thief 1 is not visible now, but thief 2 is visible.

INDIAN INSTITUTE OF TECHNOLOGY MADRAS · सिद्धिर्भवति कर्मजा ·

Mathematics for Data Science 1
Week-02 Tutorial-09

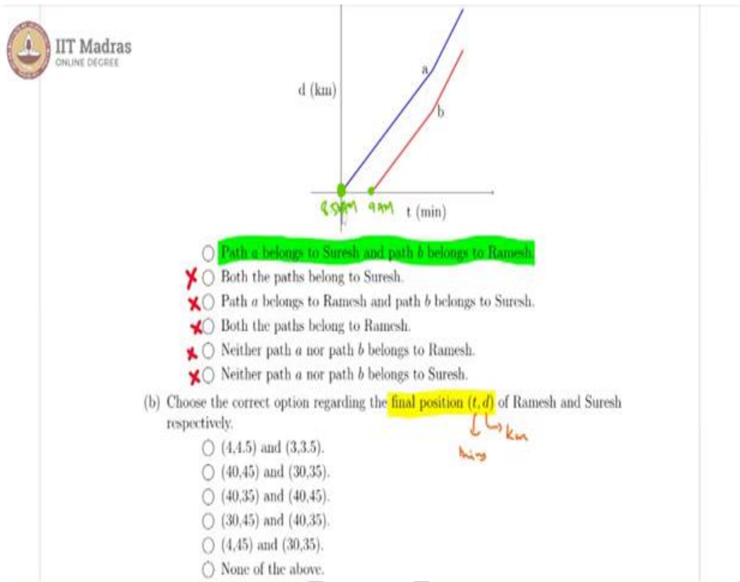
(Refer Slide Time: 0:12)



For our 9 th problem, we have 2 colleagues Ramesh and Suresh, and their office starts at 9:30 AM, Suresh starts at 8:50, Ramesh starts at 9, and they both go at equal speed. At 9:20 they decide to increase their speeds in order to reach their office on time, which is at 9:30 and this increase in speed was 30 kilometer per hour each, and they manage to reach the office on time. So, the timer begins at 8:50 AM, which means our origin is corresponding to 8:50 AM.

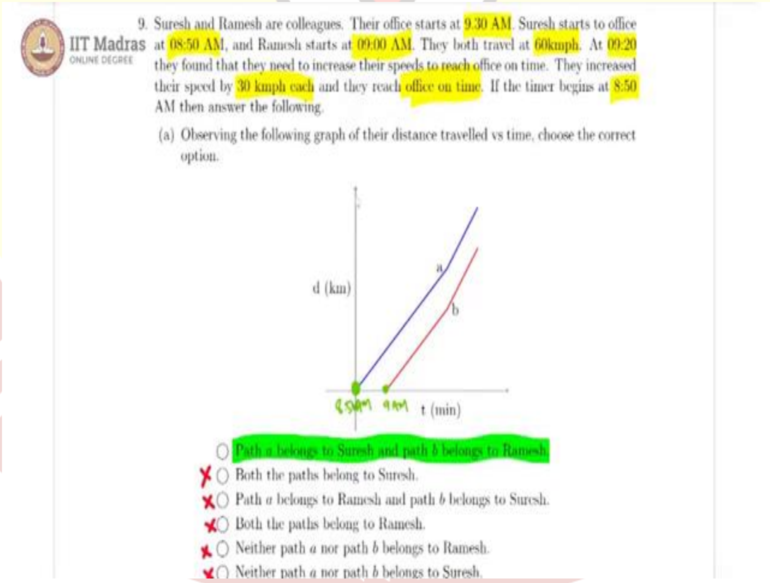
And since we know that Suresh started at 8:50 path A must belong to Suresh and Ramesh started a little late, so this here should be 9 AM. So, B, the path B corresponds to Ramesh's journey, which gives us option A is correct. Of course, this is wrong because both paths do not belong to Suresh. This is also wrong because path A does not belong to Ramesh, both paths do not belong to Ramesh and Ramesh has a path Suresh has a path so all of these options are wrong, only option A is right.

(Refer Slide Time: 1:54)



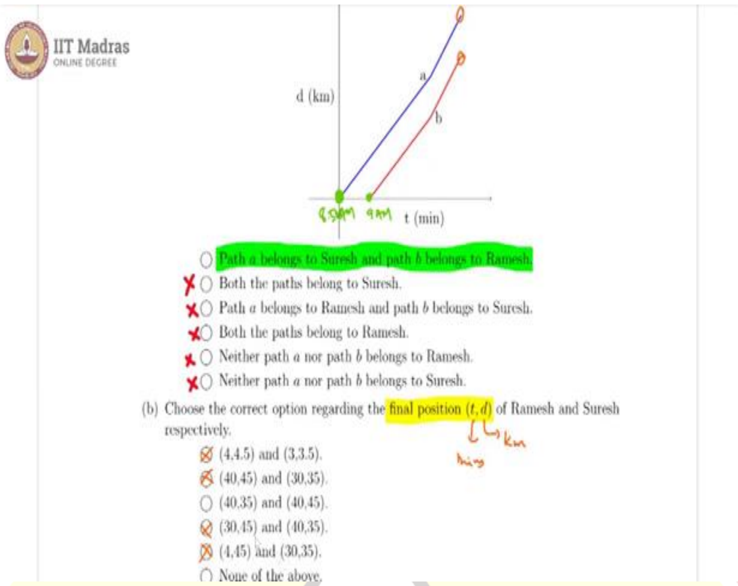
Now, in the second part, we are being asked the final position t, d, where t must be in minutes and d must be in kilometers. So, what, so this is not actually the position, is a coordinate in this particular graph regarding the final position of Ramesh and Suresh respectively.

(Refer Slide Time: 2:24)



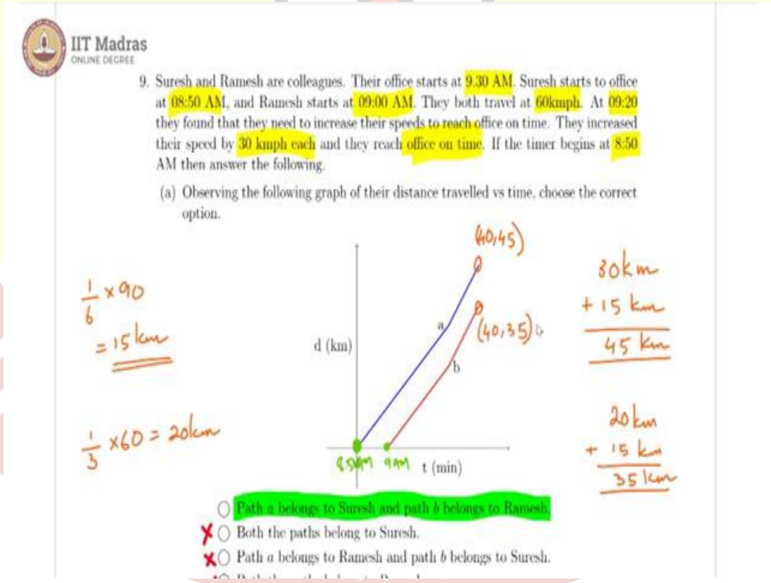
So, we know that, Suresh started at 8:50 and he traveled till 9:30. That means, Suresh traveled for 40 minutes, whereas, Ramesh started at 9 AM and reached office at 9:30 AM. So, Ramesh started for 30 minutes.

(Refer Slide Time: 2:50)



However, both of them reached at the same time, which means, this point and this point in the graph, both of them have the same x coordinate. Now, that clearly rules out this, this, this and this, because none of these have the same x coordinate.

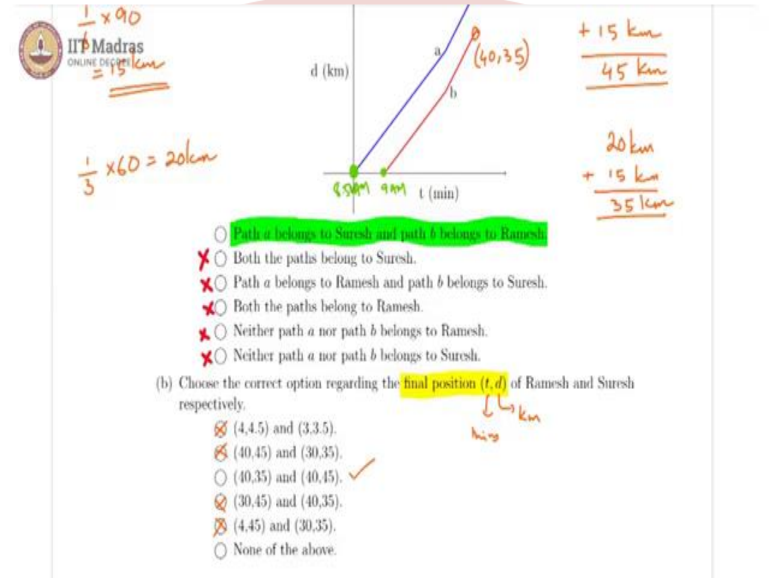
(Refer Slide Time: 3:18)



In terms of the number of kilometers traveled, Suresh goes at 60 kmph for $\frac{1}{2}$ an hour till 9:20. So, in $\frac{1}{2}$ an hour he must have covered 30 kilometers and then for 10 minutes, he goes at a speed of an additional 30 kmph, so, 90 kilometer per hour for 10 minutes. So, 10 minutes is $\frac{1}{6}$ an hour, $\frac{1}{6} \times 90$ gives us 15 km. So, overall Suresh covered 45 km, so this point it must be 40, 45.

Whereas, Ramesh also covered the same 15 km in those 10 minutes but in the initial time of established it is only 20 minutes, he did not cover 30, he instead covered 20 minutes is $\frac{1}{3}$ of an hour $\frac{1}{3} \times 60$ gives us 20 km. So, Ramesh covered 20 km + 15 km s giving us 35 kilometer overall. So, this point here it is 40, 35.

(Refer Slide Time: 4:48)



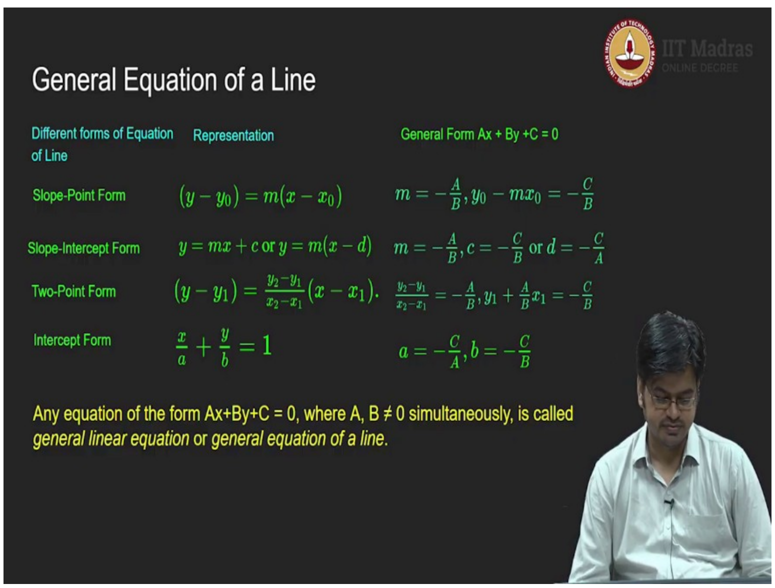
So, our correct option is this one.

Mathematics for Data Science 1
Prof. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology Madras – Chennai

Lecture-20
General Equation of Line

So, far in our journey we have studied how to represent on line which is a geometric object in algebraic manner using various forms of equations. This is a time to recollect; what are the forms of equations that we have studied and understand some common properties commonalities in that equation of line and give a general equation of line which will be helpful for further analysis. So, let us see what are the different forms of line; equations of line that we have studied.

(Refer Slide Time: 00:49)



So, in particular we had two forms one is two-point form another when its slope point form. So, first I will list the slope point form, a specialized version of this is slope intercept form where instead of a point you have been given x intercept or a y-intercept. Then we have also studied two-point form given two points how to uniquely determine a line and a specialized version of that is nothing but intercept form.

So we can quickly review these forms like slope point form we have a point (x_0, y_0) which is given to us and a slope m that is given to us. So, we come up with an equation when we give the algebraic representation of this line with slope y with slope m and point (x_0, y_0), we will come up with a representation as $(y-y_0)=m(x-x_0)$. When you come to slope intercept form suppose the x intercept is given to me if I have been given an x intercept then the y coordinate of that point will be 0.

So let us say x intercept is d, in that case my equation from slope point form as slope intercept form is a specialized version of slope in point form. My equation will become $(y-0)=m(x-d)$ if the intercept is at d. So, $y=m(x-d)$, in a similar manner so the y intercept is given to me and that intercept is at c then my y_0 will be replaced by c and x_0 will be replaced by 0 therefore I will come up with an equation $y=mx+c$ that is what is listed here, given a y-intercept and given an x intercept the equation has a form $y=m(x-d)$.

Let us come to two point form we have also seen during the course that this two point form is closely related to slope point form. We also know that given any two points on a line we can determine the slope of a line so in this particular expression m will be replaced by the ratio of the change in y upon change in x. Therefore the two point form will be just replica of this instead of m you will have the difference between y-axis difference between the coordinates of y-axis and difference between the coordinates of x-axis that will be given in this form.

$$(y-y_1)=\frac{y_2-y_1}{x_2-x_1}(x-x_1)$$

Now remember here the points given are not (x_0, y_0) so the points given are (x_1, y_1) and (x_2, y_2) therefore x_0 is replaced by x_1 and y_0 is replaced by y_1 and this thing is nothing but a replacement of m that is how these two forms are also closely related. In an intercept form you will get two intercepts x intercept let us say x intercept is a and y intercept is b then how will these forms change?

If x intercept is a that means I have a point $(a,0)$ so my (x_1, y_1) will be nothing but $x_1=a, y_1=0$ and (y_2-y_1) will be b, x_2-x_1 will be minus a, so $\frac{b}{-a}$ is equal to $\frac{-b}{a}$ x, and if you simplify that you will come up with a very simple expression of the form $\frac{x}{a}+\frac{y}{b}=1$. So, here there is a no-brainer nothing to remember below x you write x-intercept below y you write y-intercept and equate it with 1.

Now if you look at all these forms there is one common feature, let us take the slope point form given a point (x_0, y_0) this $x_0 \wedge y_0$is fixed. The slope of a line is fixed. So, now what we are identifying is we are identifying in a condition in the form of (x, y) what these coordinates should satisfy. So, the variables are x and y.

If you look at all these forms the same feature is visible, the variables are x and y and I have an expression of the form some constant times y some constant times x and added with another constant. Let us take this feature for example $y-y_0=m(x-x_0)$now I want to differentiate between variables and constant. So, I can simply write this as $y-mx=y_0-m\,x_0$. $y_0-m\,x_0$ will be the constant associated with this particular equation and y and one variable y is associated with real coefficient 1 and variable x is associated with real coefficient -m.

So in particular I can have a general form of the equation and similar story is true for all this. For example, if you come here, with variable x, $\frac{1}{a}$is a real coefficient that is associated, with variable y, $\frac{1}{b}$that is a real coefficient that is associated and the constant is c. So, I can discuss same things about all these features but one thing is common that I can have a general form of equation which will be of the form $Ax+By+C=0$.

Now let us identify this particular general form with our various expressions like slope point form, the way I discussed the slope point form we already know. In this case we have assumed that b is equal to 1 but I can as well multiply by a constant term throughout the equation and we will have the same equation. So, assuming this holds true let us discuss about this particular expression. So, in this case you can easily see if I relate this equation with this equation that is you rewrite this as $y-mx=y_0-m\,x_0$.

In that case you can have this expression which will give the value of m when you compare with respect to this expression as $\frac{-A}{B}$ and value of $y_0 - m x_0$, now remember this is a constant term because all these are constants. So, $y_0 - m x_0 = \frac{-C}{B}$. If you are able to understand this then you can easily understand the slope intercept form. Because in the slope-intercept form, $y = mx + c$you have y-intercept which is c therefore your y_0 will be replaced by c and x_0 will be replaced by 0 so if you look at this expression m will still remain $\frac{-A}{B}$, when I am identifying this equation m will still remain by minus $\frac{-A}{B}$, y_0 is identified with $C - \left(\frac{-A}{B}\right) x_0$is 0 so this becomes irrelevant so y_0 is c so $c = \frac{-C}{B}$. In a similar manner you can do for x-intercept and you will get these expressions.

So m as I mentioned $c = \frac{-C}{B}$ and for getting d you just put $x_0 = d$and $y_0 = 0$, you will get this expression. So, same exercise can be done for two point form and intercept form remember this m will be replaced by a ratio of these two differences. So, m is replaced by a ratio of these two differences there is no (x_0, y_0) there will be (x_1, y_1) therefore you will have an expression of this form.

But remember this $\frac{-C}{B}$is common everywhere the slope is $\frac{-A}{B}$ everywhere so essentially, we have got one simple general equation. Similar things you can do for $\frac{x}{a} + \frac{y}{b} = 1$ that is intercept form and you will get $a = \frac{-C}{A}, b = \frac{-C}{B}$. So, what we have seen here is an exact matching one-to-one correspondence of a general equation with respect to this equation.

Now why should I consider general equation? Remember when we figured out this representation our assumption was these are non vertical lines. For vertical lines our slope do not exist but in this case if you; and those lines are where the slope do not exist those lines are vertical lines. They are of the form x is equal to some constant. If you look at this equation which is a general form of this equation you just put B to be equal to 0 you will get $Ax+c=0$ that means x is equal to some constant x is equal to $\frac{-C}{A}$, you will get that is what our intercept form also reveals.

So all these lines are actually vertical lines, so this general equation is capable of handling vertical lines also, horizontal lines are anyway handled here because if you put m is equal to 0 the horizontal line is handled. While we were deriving these forms we were always assuming non-vertical lines. So, non-vertical lines are covered as well as vertical lines are covered therefore this equation is a general form of equation of a line.

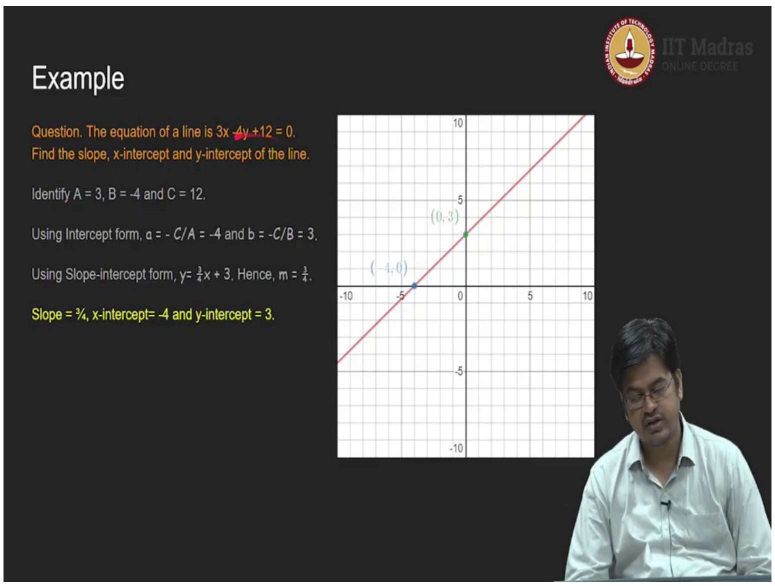
Also, in your earlier classes you might have studied this as a polynomial in without this equal to 0 Ax + By + C is a polynomial in two variables and it is a linear polynomial in two variables. Therefore you will hear a term called linear equation in two variables. So, in particular if this has to represent general form of a equation of line then A and B cannot be simultaneously equal to 0.

If A and B are simultaneously equal to 0 then I am actually equating constant with a zero which is invalid therefore the assumption will always be A and B cannot be simultaneously equal to 0. Though individually they can be equal to 0 for example you can put A is equal to 0 then you will get y is equal to some constant which is a line parallel to x axis. You can put B is equal to 0 then you will get a line x is equal to constant, x is equal to constant is parallel to y axis.

So now we will bring up a definition that any equation of the form Ax + B y + C = 0 where A and B are not equal to 0 simultaneously individually they can be 0 or they can be nonzero as well is called general linear equation because we are handling a linear polynomial which is equated to 0 so it is an equation, general linear equation or general equation of a line. So, what we are summarizing here is a polynomial in two variables or and general linear equation in two variables gives you line.

So this is the identification of a geometric object called straight line with an algebraic representation of general linear equation. So, this will give us both the strength in our analysis because now you do not have to discuss about the line. But you can as well discuss about its algebraic representation or you can start with an algebraic representation of a line and then discuss about the geometric properties of the line. How let us see in the next slide.

(Refer Slide Time: 14:30)



So, here is an example, the example gives you a question that the equation of a line is $3x-4y+12=0$. Now I do not know how this line behaves now I want to see how this linear equation represents a line. So, when I talk about a line what is the natural question we will talk about what are the two points that uniquely determine this line or you can ask what is the slope of a line and give me one point on a line because we have slope-intercept form or we have two point form any of them should be usable.

So in order to discuss about the geometric aspects we can ask a question that find the slope or x intercept or y intercept of a line. So, how will you find this the job is pretty simple let us go back and revisit the previous slide which will make the job very simple. Suppose I want to determine the x-intercept and y-intercept then I have this intercept form right which says that a is the x-intercept and b is the y-intercept.

Now you I have been given an equation in this form which is $Ax+By+C=0$ so I can image lately consider this equation and consider the values of a and b which is $-C/A$ and B is equal to $\frac{-C}{B}$. So, let us go and do the same thing on the on the our; now our problem so we have identified $Ax+By+C=0$. So, what is A, A is 3, B is minus4, C is positive 12. So, what should be my x intercept A as you have seen in the previous slide is $-C/A$.

So what is C? $\frac{12}{A}$ which is 3 so my a is 4, and a minus sign associated with it so $a=-4$. In a similar manner you can talk about y intercept which is $\frac{12}{-4}=-3$ but a minus sign because it is $\frac{-C}{B}$ so it will be 3. So, now we can readily answer the question what is on a x-intercept and y-intercept.

Now the question comes what is the slope of a line. So, for slope of a line you can use the slope intercept form $y=mx+c$. So, identify this equation in the form of $y=mx+c$ so if you look at this equation, I should push this 4y to the right hand side that gives me $y=\frac{3}{4}x+\frac{12}{4}$. So, my m should be $\frac{3}{4}$this is the answer. So, slope intercept form you have y is equal to 3 by 4 x plus 3 so the slope is naturally $\frac{3}{4}$, this easy is our calculation.

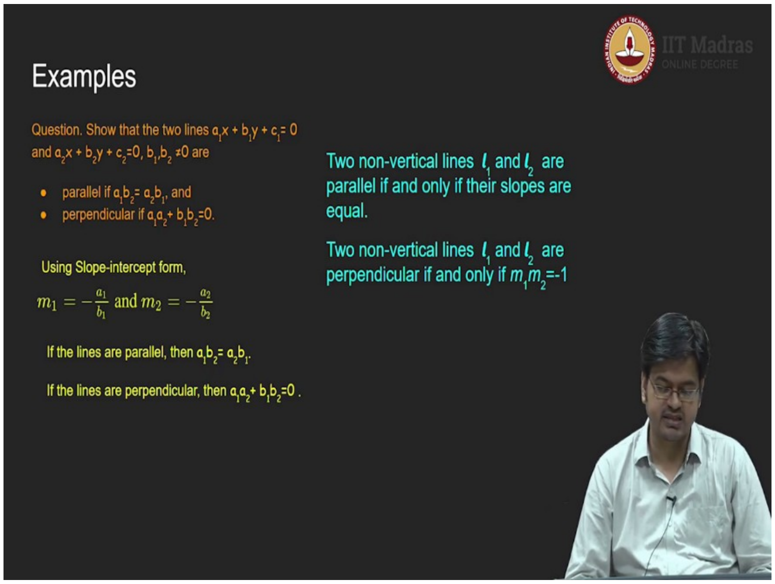
Now we have identified an algebraic object as a geometric object. Now let us see what we can do further and we can actually verify this graphically you know although it may be correct it is always better to verify it graphically. So, slope is $\frac{3}{4}$ x intercepts should be - 4 and y intercept should be 3 if you want to satisfy the equation of this line this should happen right.

So this is how we have drawn so the x-intercept is -4, y intercept is 3 and the line passes through this. Now you pick for verification purposes you can pick any point on this line and you can put the values of the coordinates into the equation of a line and verify that it will give you the value 0 that will be the identification that your answer is correct.

Mathematics for Data Science 1
Prof. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology Madras – Chennai

Lecture-21
Equation of Parallel and Perpendicular Lines in General Form

(Refer Slide Time: 00:19)



Let us look at next example which is another application of a general form of equation of a line. The example is stated in the form of a question that is if I have been given two lines $a_1x+b_1y+c_1=0$, $a_2x+b_2y+c_2=0$, $b_1, b_2 \neq 0$. What does this mean? That means the lines are non-vertical. $b_1, b_2 \neq 0$ means the lines are non-vertical you can verify for yourself.

Now two such lines are parallel if $a_1b_2=a_2b_1$and perpendicular if $a_1a_2+b_1b_2=0$. This is an interesting application of general form of equation of line. And if you recollect, we have derived some characterization of line in terms of slope. So, let us try to see this problem so let me first identify if I want to characterize parallel and perpendicular lines what should I do?

What is a parallel line, how will I identify a parallel line when I will have their slopes to be equal and how will I identify a perpendicular line, when the product of the slopes of the two lines is -1? So, if you remember this then the job reduces to finding the slopes of the two lines. Can I find a slope of these lines? Let us first consider this line $a_1 x+b_1 y+c_1=0$. You should be immediately able to identify this with slope point form which is $y=mx+c$.

So if I want to adjust this equation in the form of $y=mx+c$ then what should I do? Because b_1 is nonzero I can divide throughout by b_1 and shift this coordinate of y to their right-hand side of the equation. So, I will get $y=\frac{-a_1}{b_1}-\frac{c_1}{b_1}$. So what is the slope $\frac{-a_1}{b_1}$. A similar trick you can apply here and therefore you will get $m_2=\frac{-a_2}{b_2}$. So using slope intercept form you have got $m_1=\frac{-a_1}{b_1} \wedge m_2=\frac{-a_2}{b_2}$.

Now let us recollect the famous fact because $b_1 \wedge b_2$ are not equal to 0 we are not considering vertical lines. So, two non-vertical lines are parallel if and only if their slopes are equal. So, what you will do you will just put $m_1=m_2$because you have been given that the lines are parallel. So if you put $m_1=m_2$, minus sign will cancel each other $\frac{a_1}{b_1}=\frac{a_2}{b_2}$. Multiply both sides by $b_1 b_2$, b_1, b_2 are nonzero.

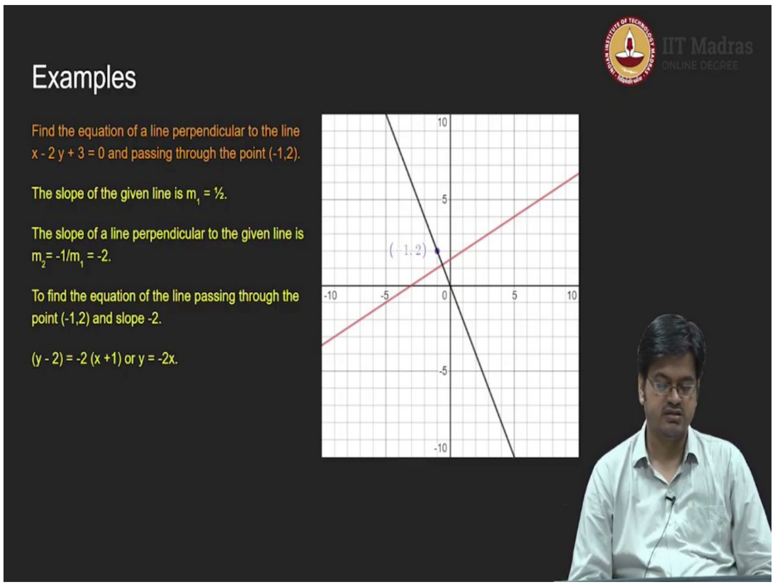
So multiply both sides by $b_1 b_2$, you will get $a_1 b_2=a_2 b_1$. Therefore, the lines are parallel then $a_1 b_2=a_2 b_1$. In a similar manner we also know something about perpendicular lines that the product of their slopes is -1, if the lines are perpendicular. So, just multiply m_1, m_2 and equated to -1. Minus sign will cancel each other so you will get $\frac{a_1}{b_1}\times\frac{a_2}{b_2}=-1$.

So take the denominator on the right hand side that is $b_1 b_2$, so $a_1 a_2=-b_1 b_2$ which essentially means $a_1 a_2+b_1 b_2=0$. Therefore, we have proved the result. So, now what we have done right now is we have related our result about the characterization of perpendicular and parallel line via slope to a general form of equation and this is the new condition that we are coming up with if the lines are parallel and you have been given to two non -vertical lines and their general forms then you just need to check that $a_1 b_2 = a_2 b_1$ for the lines to be parallel and $a_1 a_2 + b_1 b_2 = 0$ for the lines to be perpendicular. This you can consider as another characterization of parallel and perpendicular lines using a general form of the equation of lines.

Mathematics for Data Science 1
Prof. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology Madras – Chennai

Lecture-22
Equation of a Perpendicular Line Passing Through a Point

(Refer Slide Time: 00:15)



So, now you have been presented with an equation of line and a point and the question is you find the equation of a line perpendicular to the line $x-2y+3=0$ and passing through the point $(-1,2)$. So, in this case let us identify the general form of the equation that is $Ax+By+C=0$ and you can easily see that $A=1, B=-2, C=3$.

Therefore, the slope of the given line the line that is given to you will be $\frac{-A}{B}$ which is $\frac{-1}{-2}$ that will be $\frac{1}{2}$. So, the slope of the given line $m_1=\frac{1}{2}$, now if at all a line is perpendicular to it then you already know that the product of the slopes is -1. So if the product of the slopes is -1 then $m_1 m_2=-1$ that is $m_2=\frac{-1}{m_1}$. So m_1 is $\frac{1}{2}$ which will give me $m_2=-2$.

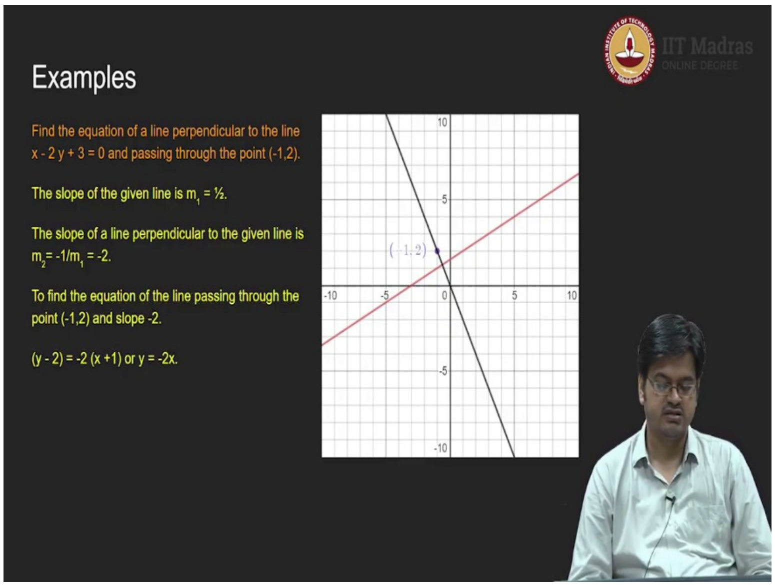
So now the problem reduces to the slope of a given line is -2 and it passes through point $(-1,2)$ and I want to find the equation of a line that is passing through point $(-1,2)$ and has slope -2. So, use the slope point form $y-y_0=m(x-x_0)$, y_0 is 2 so you can easily see $y-2=-2(x+1)$, ¿, so $(x+1)$, rearrange the terms so this 2 will get cancelled constant therefore I will get the equation of a line to be $y=-2x$ or in a general form you can write this as $-2x+y=0$.

So, let us try to figure out whether the line which we have actually found is perpendicular or not. So, the orange line is the line for which the equation is given $x-2y+3=0$ the point $(-1,2)$ is displayed in the graph and the line passing through it is also displayed and you can clearly see the angle that is made is 90 degrees therefore the lines are perpendicular and our answer is correct. So, our verification test has passed.

Mathematics for Data Science 1
Prof. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology Madras – Chennai

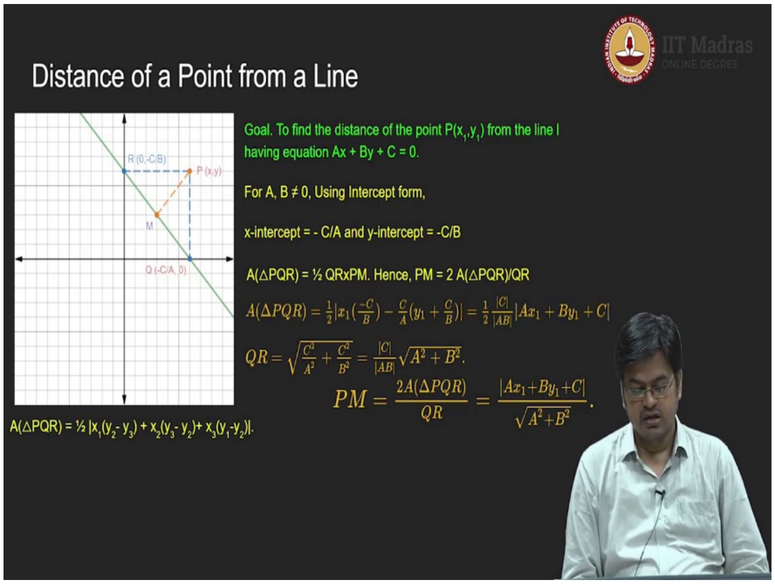
Lecture-23
Distance of a point from the given line

(Refer Slide Time: 00:16)



So, we have verified the answer now what you can see here is (-1, 2) is a point which is lying on the line which is perpendicular to the given line. An interesting question can be asked that what is the distance of this; point from the given line. Let us try to answer that question in the next slide.

(Refer Slide Time: 00:40)



So, the question is given any point I have another line the point is not collinear to the line then what is the distance of that point from a line. So, let us take that as our goal. To be precise we are interested in finding the distance of the point P which has coordinates (x_1, y_1) of this point from the line l which has equation $Ax+By+C=0$ and this is a general form of equation.

Now how will we proceed? So, if I want to understand the location of P, I need to do some analysis because $Ax+By+C=0$ is a completely geometric object. So, I need to understand this line in terms of its geometric concepts. So, what are the geometric concepts associated with this line they are slope x intercept y intercept or points on this particular line. So, let us identify those things first.

So if we assume that A and B both are not equal to 0 then I can rewrite this equation in the form of intercept form that is $\frac{x}{a}+\frac{y}{b}=1$ so in that case my a is actually $\frac{-C}{A}$ that is x intercept and my small b will be $\frac{-C}{B}$ which is y intercept. So, I have identified this line as I have identified the 2 points and these 2 points uniquely determine the line. So, I know how the line is located.

Let us try to visualize this line in terms of the graph of a function. So, as you can see I have mentioned that x intercept is $\frac{-C}{A}$ so it is mentioned as a point Q which is $\left(\frac{-C}{A},0\right)$, y intercept is $\frac{-C}{B}$ which is identified here so $\left(0,-\frac{C}{A}\right)$, the point P is located here it may be located anywhere but right now the point P is located here it has coordinates (x_1,y_1). So, now I want to identify a distance of this point from this line, the line joining the points Q and R.

So, what is the distance? It should be the shortest distance from the line, so the shortest distance in this case if you move along this line the shortest distance in this case is a point where the point is actually perpendicular to the line. So, what I want to say is the shortest distance is the one which is the perpendicular distance. So, the entire question reduces to how to find this perpendicular distance PM.

So, let us try to see what are the geometric objects associated with this. So, you can see from the dotted lines the geometric object that I can associate with this particular distance is a triangle PQR. Now if I want to find the distance PM, I can take help of this triangle PQR so that I will be able to find the distance PM. So, how will I do that? First you see if I want to compute the area of triangle PQR what do I need to know?

I need to know the base and the height and the area of a triangle is half base into height. So, half base into height means $\frac{1}{2}\times QR\times PM$. I do not know what is PM. But we have already seen in this course how to find area of a triangle where its coordinates are given. So, even though I do not know what is PM I know how to compute the area of a triangle. The next question is do I know how to compute the length QR?

Yes of course because this is x intercept this is y intercept and these are the lines which are the distances on x and y axis and all of them form a right-angled triangle. So, by Pythagorean theorem I will be able to find the length of QR. So, I can reformulate the question as

$PM = 2 \times \frac{A(\Delta PQR)}{QR}$. So, now I know how to compute the length PM if I know how to compute area of triangle and how to compute the length of line segment QR both of which I know.

So, let us go ahead and try to compute area of triangle PQR. So, here is our formula for area of triangle PQR which has coordinates $(x_1, y_1), (x_2, y_2), \wedge (x_3, y_3)$. So, let us start with (x_1, y_1), the (x_1, y_1) is the first coordinate remember you will always take this in anti-clockwise direction. So, I will start with this coordinate then I will go to R and then I will go to Q. So, this is (x_1, y_1) this is (x_2, y_2) and this is (x_3, y_3) according to the notation that is given in the formula.

So you will see $x_1(y_2 - y_3)$ so x_1 is first coordinate it will remain x_1 because P has coordinate (x_1, y_1), y_2 is $\frac{-C}{B}$, y_3 is zero. So, you will get $x_1\left(\frac{-C}{B} - 0\right)$ then the next term that is x_2, x_2 here is 0 so this entire thing vanishes then you go to x_3, what is x_3? x_3 is $\frac{-C}{A}$, into y_1, which is y_1 as it is $-\left(\frac{-C}{B}\right)$ so $\left(y_1 + \frac{C}{B}\right)$, this is how I got the formula.

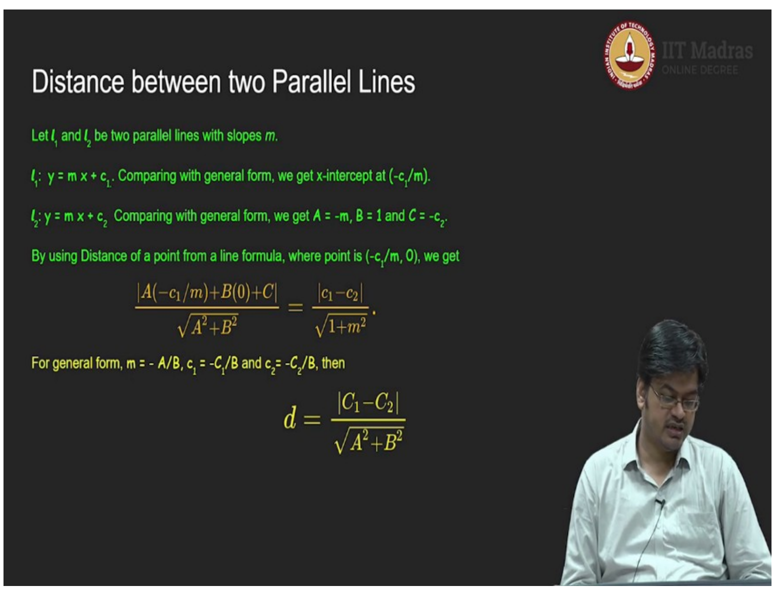
So, if you look at this formula closely you can actually take C common from all within the mod sign so you can take $\left|\frac{C}{B}\right|$, denominator has terms containing B and AB. So, if you want to take those terms out you multiply throughout by AB or you find the LCM is AB and you take AB out so you will get $\frac{1}{2} \times \frac{|C|}{|B|} \times |A x_1 + B y_1 + C|$, remember this is the term corresponding to general form of the equation.

Now we have seen how to compute area of triangle PQR. Next, we will see how to compute the length QR. But length QR is actually very easy because I have a point Q which has only x-coordinate and I have a point R which has only y-coordinate. So, it will be as if computing the distance of length QR is $\sqrt{\frac{C^2}{A^2}+\frac{C^2}{B^2}}$ these are the 2 sides of the triangle and this QR is the hypotenuse of that right-angle triangle.

So, again you can simplify this to amend to this form so you can take out C common so you will get $|C|$, you take A and B common you will get $|AB|$ and then you will get $\sqrt{A^2+B^2}$ which is which is in the numerator and now if you look at this form PM which is the length of the line segment PM is $2\times\frac{A(\Delta PQR)}{QR}$, so just now it is just a matter of feeding the values this half will get cancelled with this 2 and area of triangle PQR is this and QR is this therefore this constants also will vanish because they are same.

And you will get the formula to be equal to $\frac{|Ax_1+By_1+C|}{\sqrt{A^2+B^2}}$ this is how you will calculate a perpendicular distance of a point from a line. Now this idea can be helpful in finding one more thing that is a distance between two parallel lines.

(Refer Slide Time: 10:22)



So, the question can be asked is I have two parallel lines what is the distance between two parallel lines. So, let us take the set up because the lines are parallel l1 and l2 they have common slope or the same slope, so their slope is m. Then you can use the slope point form which is

$y = mx + c_1$. Now I want to use the previous concept that I have introduced distance of a point from a line. So, I will first identify this with x intercept. So, what will be the x intercept in this case?

If you identify this line it is very easy to see go back to the general form and figure out that x-intercept is $\frac{-c_1}{m}$, because $B=1$ here $B=-m$ and $C=-c_1$, so the intercept is this $\frac{-c_1}{m}$. Let us take another line that is l2 it has same slope identify it with our standard form $A=-m, B=1, \wedge C=-c_2$. So, now given x-intercept what are the coordinates of this x-intercept $\left(\frac{-c_1}{m}, 0\right)$.

So now the problem reduces to finding the distance of this point from this line ok. So, by using the distance of a point from a line formula where the point is $\left(\frac{-c_1}{m}, 0\right)$, you just need to substitute this point (x_1, y_1) into this formula for the distance of a line which is given as $\frac{|A x_1 + B y_1 + C|}{\sqrt{A^2 + B^2}}$.

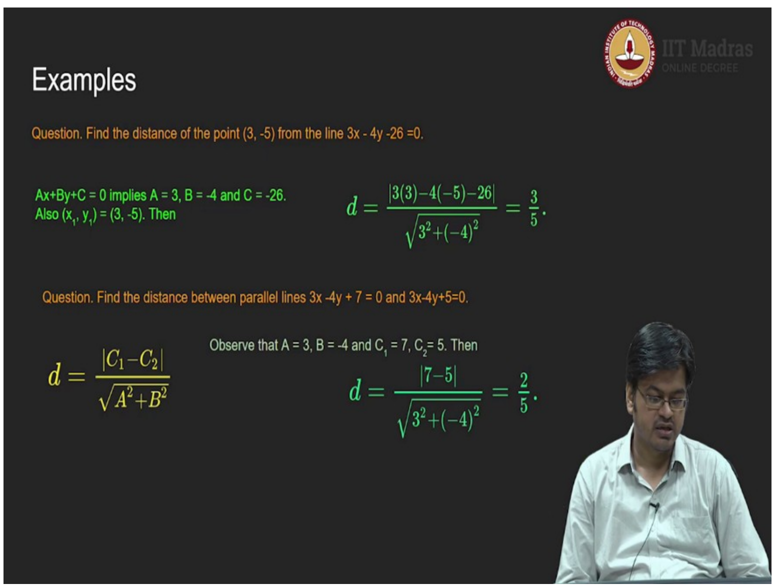
So, my point $x_1 is - \frac{c_1}{m}$ substituted here, y_1 is 0 substituted here you will get the formula to be equal to $\frac{|C_1 - C_2|}{\sqrt{A^2 + B^2}}$ so in this case $\sqrt{A^2 + B^2}$, B was 1, A is -m so it is $\sqrt{1 + m^2}$.

Now you can actually identify this formula in the general equation form also. So, in the general form instead of B = 1 we have slope which is equal to -A/B and $c_1 = \frac{-C_1}{B}$ and $c_2 = \frac{-C_2}{B}$. So, this I am matching with both equations in general form these are slope point forms but now if you match these equations with a general form you will get this description of the line where you have $Ax + By + C1$ as one line is equal to 0 as one equation of line.

$Ax+By+C_2=0$ as equation of the second line. So, in that case this is the form and therefore now you just substitute these values into this expression. So, this m will be replaced by - A/B so you will get $\sqrt{A^2+B^2}$ here and some $|AB|$ will come out common and therefore finally that will cancel off with this B and you will get the expression of the form $\frac{|C_1-C_2|}{\sqrt{A^2+B^2}}, C_1 \wedge C_2$ belong to general form of equation.

So, this gives us a clear-cut understanding of the interconnection between the slope point form and general form of equation and we have figured out what is a distance between two parallel lines using distance of a point from a line formula.

(Refer Slide Time: 14:42)



So, now we will solve some examples to concretize the concepts so here are the examples in line. So, you have been asked to find a distance of a point $(3,-5)$ from the line $3x-4y-26=0$. So, in this case you just need to apply the formula, what is a formula, $\frac{|Ax_1+By_1+C|}{\sqrt{A^2+B^2}}$. So, what is A, B, and C here is the key question? (x_1, y_1) is known to be $(3,-5)$. So, A is 3, B is -4, C is -26 then you just need to apply that formula which will give the denominator square root of 25 it will give me 5 the numerator will be 3.

In a similar manner you can ask a question what is the distance between two parallel lines $3x-4y+7=0$, $3x-4y+5=0$. So, what is a formula that we have derived it is $\frac{|C_1-C_2|}{\sqrt{A^2+B^2}}$ it is very straightforward. So, what is C_1 here? C_1 the first line the constant term is 7 the second line the constant term is 5. So, it will be modulus of $\frac{7-5}{\sqrt{3^2+(-4)^2}}$. So, you will get the answer to be 2 / 5.

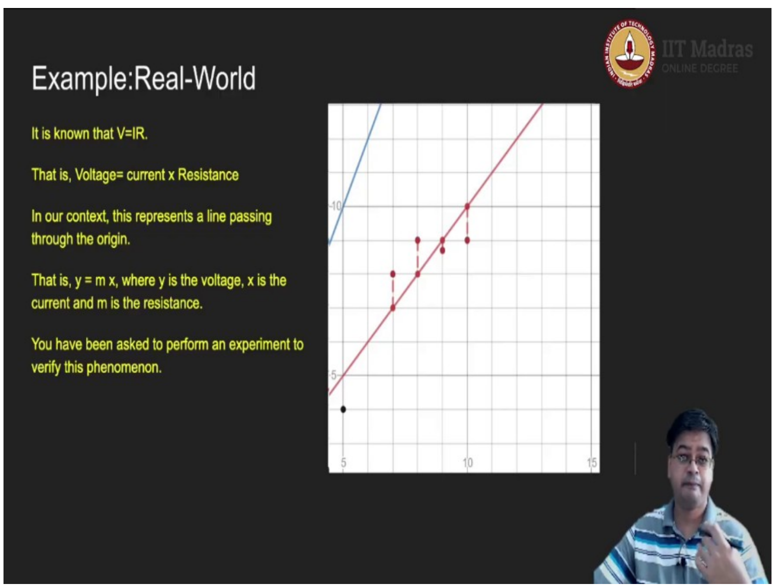
Mathematics for Data Science 1
Prof. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology Madras – Chennai

Lecture-24
Straight Line Fit

Welcome friends welcome back so far what we have seen is a distance of a line from a point distance between two parallel lines. But the question now we can ask is, is that the only distance that we can seek as a distance of a point from the line. To demonstrate this let me give you one example where the paradigm will change as we will compare several points set of points and we will compare the distance from those set of points to the line and the paradigm change that I want to say is you will think differently how the distance will change from a line.

So, let us take one simple example this example is related to a small experiment that you might have conducted in your lab.

(Refer Slide Time: 01:16)



It is a physics experiment which says $V = IR$ that is voltage is equal to current times the resistance. Voltage is equal to current times the resistance you all know this is a law this is the law of physics where voltage is measured in volts current is measured in amperes and resistance in ohms. Now the experiment that a physics teacher asked you to conduct is you have to verify this law or using this law can you compute the resistance of a particular equipment.

So now what you will do is you will actually relate this with our equation of a straight line. So, if I want to relate this with the equation of a straight line then what will happen you see V is voltage so on the right hand side you can replace this voltage by y then the current that is delivered to the circuit or the equipment you can denote it by say x and you want to determine the resistance which is an unknown so you can put it as m.

And what is the constant? The constant is 0, so you can relate this with the equation $y=mx$, where y is the voltage, x is the current and m is the resistance and the whole purpose is to determine this resistance over here m. So, the setup is ready the lab technician has arranged a set up and you just have to go and perform the experiment and verify this phenomenon. So, the catch over here is you want to determine what is a resistance.

So, the lab technician was very kind he has given you a priori information that there are only two kinds of resistors our lab has one has a resistance of 1 ohm another one has a resistance of 2 ohms. This is the information that is given to you. Also notice the fact that this line is passing through the origin that means (0, 0) is one point why (0, 0) should be a one point because there is no current then there is no voltage this is our assumption.

So (0, 0) is one point and this line is passing through the origin so if I look at a mathematical theory that I have studied so far I can safely assume if I get one reading if I get one reading from that circuit that will help me in understanding the behavior and I can safely go and tell what is the resistance of this particular equipment. Let us try to see how this assumption works out over here.

Now this is the data you have conducted some experiments you have observed some data so it is like you have passed a current of 1 ampere and you received the output of 2 volts here you can say you have passed the current of 5 amperes and you have received output which is 4 volts and so on and so forth. So, this is how it is working on. Now we want to identify what is the correct line that will fit because I know from theory that this is a line passing through the origin.

So in particular if I tell you this line which is (1, 2) and (0, 0) then I will get the equation of line using a slope point form or point - point form we also know that the intercept is (0, 0) so slope intercept form $y = mx + c$, where c is 0 you can easily see the line that passes through this point is $y = 2x$. But with the same register you also got these readings. So, let us see based on the lab technicians' knowledge if we draw two lines, they will be seen they will be visible like this, interesting.

So, if I take only one observation and stop my experiment, I will get the line $y = 2x$. But if I go for more experimentation then I am getting a line which seems to be similar to $y = x$. Now what is it that is happening here, which line is a better fit. So, I need to answer this question because this line actually passes through the point (0, 0) and $y = 2x$, this line is not passing through any of the points.

So which line is better that is a natural question that comes to our mind? So, we will try to answer this question mathematically. So, how will I answer this question mathematically? Let us zoom in and consider our notion of a perpendicular distance. What is a perpendicular distance? You will actually drop a perpendicular from this point to this point and you will compute the distance of a line. Is that distance a correct distance? Geometrically it is a correct distance that is a distance of a line.

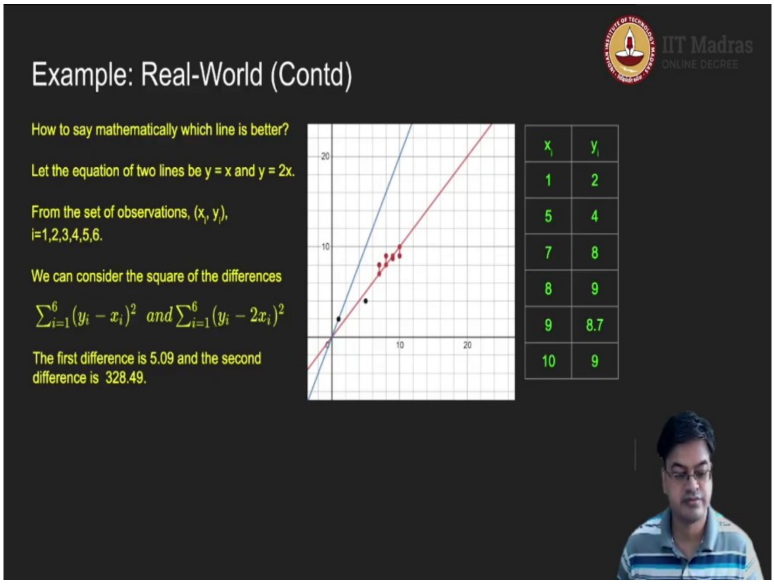
But in this context that we are taking real-world context what is happening here is if I pass a current of let us say 7 amperes this particular line is saying I should get a voltage of 7 volts but actually I got a voltage of 8 volts. So, now I may not be interested if I drop a perpendicular from this point to this point because this line is $y = x$ it may cross this line at point 7.5. I am not interested what is the value of y at point 7.5.

I should be interested in what is the value of y at point 7 because I have passed the current of 7 amperes not 7.5 amperes. So, the perspective of distance changes here because I want to find the distance for this particular value of x from the line and the point how to go about then we will not consider a perpendicular distance. This is a paradigm shift that I was talking about at the beginning of the video.

So now I will not consider this thing but I will consider this distance that is a vertical distance the distance that is parallel to y axis that is what I will consider. So, once I consider the distance that is parallel to y axis, I have to consider these distances. So, again coming back to the question which line is the best-fit line I can consider similar distances over here. And I can consider similar distances over the blue line.

So which line is the best fit? We will try to answer this question mathematically. So, mathematically we have seen that perpendicular distance will not fetch me any result directly. So, I need to consider the distances that are parallel to y axis.

(Refer Slide Time: 08:57)



So, let us formalize this in a real term, this is the data that was shown in the picture. So, for 1 ampere you have got 2 volts current. For 5 ampere you got 4 volts current, for 7 you got 8, 8 you got 9, 9 you got 8.7 and 10 you got 9. So, there is no direct relation between y and x; you cannot figure out the $y = x$ is visible over here but something is there which is making that line pass very close to all these points.

This is the this is the demonstration, so $y=2x$ is way apart and we are assuming that the hypothesis given by the lab technician is correct. So, I want to mathematically formulate this problem. There are two lines $y=x \wedge y=2x$ both pass through the origin, so current 0 voltage 0 hypothesis is correct. Now you have the set of observations xi's and yi's. I want to compute which line is better.

So, let us try to see if I consider the sum of the differences, what do I mean by some of the differences? If I consider $y=x$ is a valid equation of line then I will consider y_i-x_i, that is the distance between the line y and x because here y is equal to x if I input x_i my point that I will get is also x_i because $y=x_i$ and the actual output that I have got is y_i so I will consider y_i-x_i as one coordinate and y_i-2x_i as another difference that will be a point over here y_i-2x_i it will be a point over here.

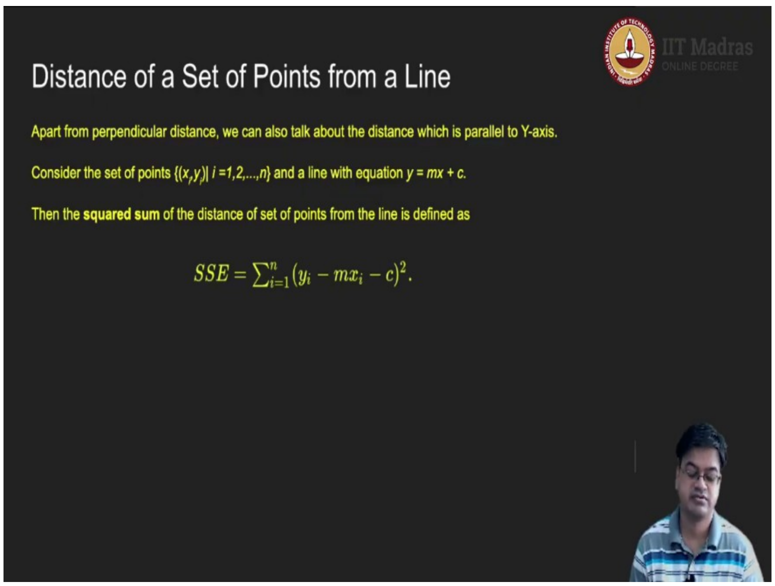
But if I just consider the differences the problem is the differences may cancel each other some differences may be positive some differences may be negative. so, I do not want those differences to cancel out each other so what I will do is I will take square of them. So, in particular we can define the sum square difference that is $\sum_{i=1}^{6} (y_i-x_i)^2 \wedge \sum_{i=1}^{6} (y_i-2x_i)^2$.

Now what this difference is calculating? It is calculating the difference between $y_i \wedge x_i$ in the first case and $y_i \wedge 2x_i$in the second case that is the error that we have made when we actually saw the output on the error the equipment has made or in our recording the error which is made in whatever way that is the error made. So, $(y_i-x_i)^2$ and $(y_i-2x_i)^2$ right. Now what do you think which one will be better the one that will be better which will have a least difference.

So, you can actually put in these values and compute these differences and square them sum over them you will get the first difference is 5.09 and the second difference is 328.49. In this situation what should be our conclusion? Our conclusion should be that the difference where the difference is least that is 5.09 this must be a better line as compared to this line that essentially reduces to a conclusion that y is equal to x is a better line as compared to y is equal to 2x which is pretty evident intuitive from the figure as well.

So, you can see this figure you can see this chunk of points that are located around $y=x$ and therefore the resistance of the equipment that is given to us must be 1 ohm that should be our conclusion. So, I want to introduce a notion of this kind to handle the real-world problems. So, let us see what is that notion? In this case you were very lucky the lab technician has given you the set of points or the resistance values there are only two resistance values.

(Refer Slide Time: 13:49)



But real life is not that lucky, so there they may not give you the set of values, and you want to find out what is the best line that is passing through these set of points. In that case this notion of a distance of a set of points from a line may help. So, what is this notion? First of all, we know one notion is perpendicular distance but that perpendicular distance may not be of much use when we are coming to the real-world perspective.

In that case we talk about the distances that are parallel to y axis from the distance of a points that are parallel to y axis. So, in particular if you have been given n points $\{(x_i, y_i) | i=1,2,\ldots,n\}$. You just plot this equation $y=mx+c$. Now remember here this equation is valid when it is not a vertical line. If it is a vertical line this equation is not valid. And if it is a vertical line you do not need such a complicated procedure to estimate it.

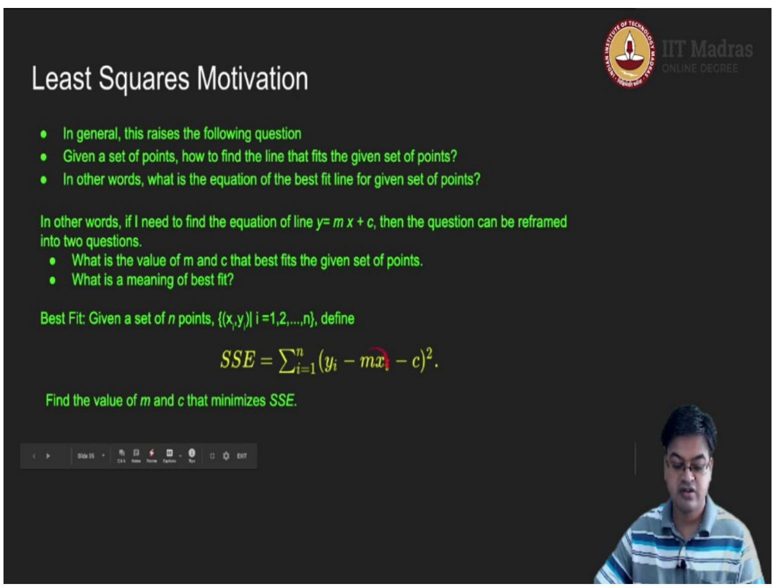
So $y = mx + c$ is our standard equation of line which is a slope point form or slope-intercept form to be precise and then as in the previous case we have defined the squared sum of the distance of the set of points from the line. So, in the previous case $y_i - x_i$ but in this case what should it be $(y_i - m x_i - c)^2$ and you have to sum over all of them. So, $\sum_{i=1}^{n} (y_i - m x_i - c)^2$.

So, we call this as sum squared error or some squared distance, sum squared error so the abbreviation is SSE.

$$SSE = \sum_{i=1}^{n} (y_i - m x_i - c)^2$$

Now the fact is when we are handling a general problem, we do not know what will be m and what will be c.

(Refer Slide Time: 16:08)



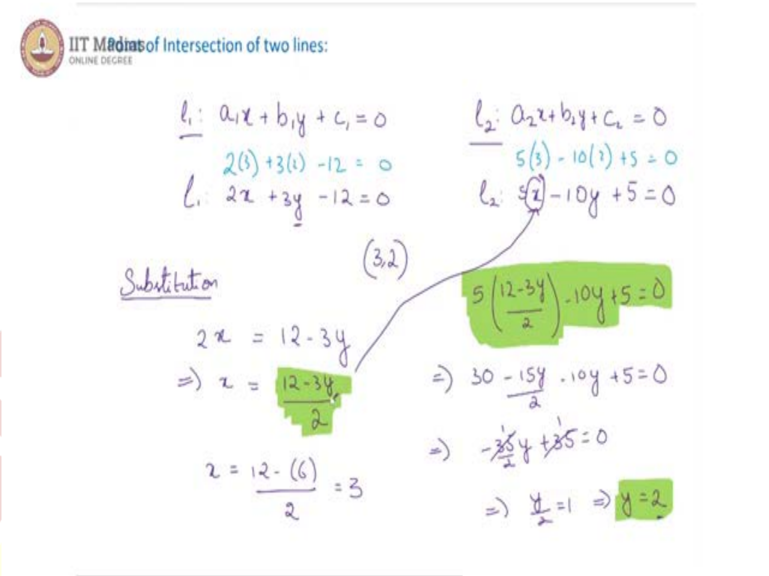
So, our goal should be if I want to find the best line, I want to find the best line passing through this point what should be my goal. So, these raises two questions if I have some square, I want to know the value of m I want to know the value of c? So, given the set of points how to find a line that fits the given set of points remember now I am not uniquely determining the line I am saying but that fits the given set of points.

The line may not pass through any of the points in this particular case in other words $y = mx + c$ so what is the equation of the line that best fits the given set of points. This will mean I need to find an equation of a line $y = mx + c$ and then the question can be reframed into two questions that is what do I mean by the value of m and c that best fits the line and then I have to define what is the best fit according to me.

Obviously, the best fit according to me will be the sum squared error minimization. And so, if I define SSE in this manner then I want to find the values of m and c that minimize SSE but this is right now beyond our scope as so far, we have handled only linear terms. But if you look at these terms, they appear to be in the form of squares of something. So, we need to divide some strategies in order to find this minimization for m and c so with that we will see in few upcoming videos of the course, thank you.

Mathematics for Data Science 1
Week-03 Tutorial - Point of Intersection of two lines

(Refer Slide Time: 0:14)



Hello mathematics students. In this tutorial, we are going to learn to find the point of intersection of two given lines. So, you have two-line equations given to you. Let us call one $a_1x + b_1y + c_1 = 0$, let this be line l1, and line l2 is a $a_2x + b_2y + c_2 = 0$. And we try to find out the point at which these two lines intersect. And that would basically be the solution the (x, y)which satisfies l1 and l2 as well. It is easier to observe this process with example. So, let us take 2 example lines and find out where they intersect.

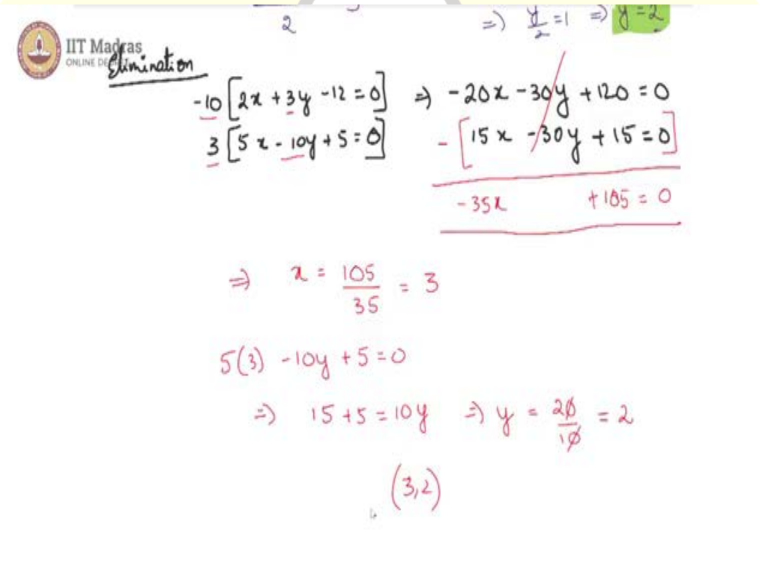
So, for our examples, let us take l1 is $2x + 3y - 12 = 0$, whereas $5x - 10y + 5 = 0$. So, when we have these 2 line equations, how do we solve for x and y. So, the best thing to do is to eliminate one variable, either x or y and get a single equation in the other variable. So, what I mean by that, and this could be done in 2 ways. One way is called substitution. In substitution, in order to remove one variable, we basically express the other in terms of it.

For example, if I wanted to eliminate the y variable, what I do is I express x in terms of y. So, I get all externs on 1 side, so $2x$ is on one side, and the other terms non x terms on the other side, which will give me $12 - 3y$. This would then indicate that x is $\frac{12-3y}{2}$, and then I take this representation of x in terms of y, and substitute it into this equation. What that gives us is, suppose I substituted it, now I will get $5\left(\frac{12-3y}{2}\right) - 10y + 5 = 0$.

So we get $30 - \frac{15y}{2} - 10y + 5 = 0$. That is essentially taking the y common I am going to get $-\frac{15y}{2} + 35 = 0$, canceling off the 35, so I get 1 here, 1 here, that would indicate $\frac{y}{2} = 1$, this implies $y = 2$. So because we eliminated the x here, we got an equation which is entirely in y, which lets us solve for y, and we get the value of y.

Now, to obtain x, we simply have to substitute this value of y in this representation of x, so we will $x = \frac{12-(6)}{2} = 3$. Which means the solution for these 2 line equations is $(3, 2)$, $x = 3$ and $x = 2$. And we can verify this quite immediately by substituting these values into the equations, I will get $2(2) + 3(1) - 12 = 0$. Likewise, $5(3) - 10(2) + 5 = 0$. So it is fairly clear that $(3, 2)$ is the solution which satisfies both linear equations.

(Refer Slide Time: 4:43)

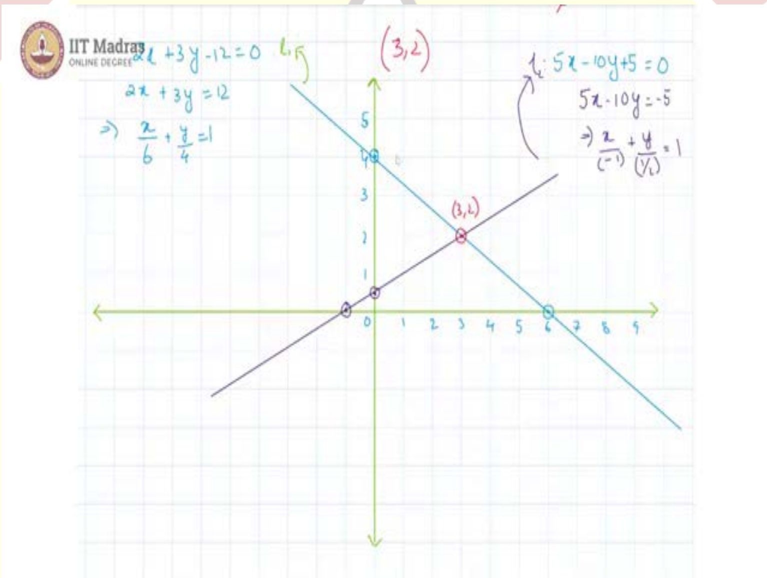


Another method of doing the same thing, which is to solve these 2 equations, we call it elimination. And in elimination, what we do is we again, take these 2 equations, which is$2x + 3y - 12 = 0$, and $5x - 10y + 5 = 0$. We again choose to eliminate either of these variables, because we earlier eliminated x and got an equation in y, now I am going to eliminate y and get an equation x. And for that, what we do is, we multiply this entire equation by the y coefficient in this equation, which is minus 10.

So, I am going to multiply this whole thing with minus 10. And we multiply this entire equation with the y coefficient here in the other equation, that is 3. What that will give us is this would give us minus $-20x - 30y + 120 = 0$. And this gives us $15x - 30y + 15 = 0$. And now what is to be observed is this is $-30y$ and this is also $-30y$, because here we multiply 3 with -10 and here we multiplied -10 with 3.

And that lets us cancel these off, if I subtracted this whole equation from the previous one now. So that will result in $-30y$ by $-30y$ getting canceled, and here, I will get $-35x + 108 = 0$. And this would indicate that $x = \frac{105}{35} = 3$. And now I can substitute $x = 3$ in either of those equations. If I substituted in the second one, I would get $5(3) - 10y + 5 = 0$, this indicates $15 + 5 = 10y$, which gives us $y = \frac{20}{10} = 2$. So, we got our value back, the point back, which is $(3, 2)$. This is the point of intersection of these 2 lines.

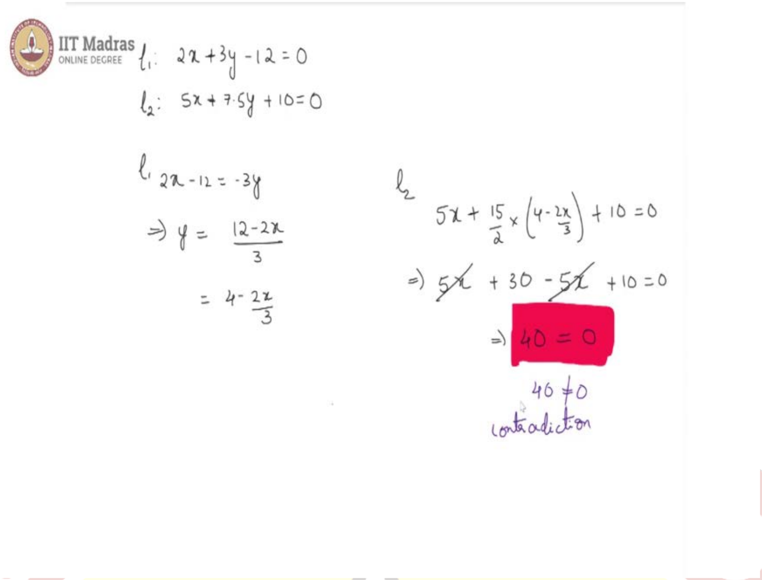
(Refer Slide Time: 7:41)



So, if we plotted these, these are our line equations, let us take the first one, I will reduce this to intercept form, which will have to be to $2x + 3y = 12$ is going to give us $\frac{x}{6} + \frac{y}{4} = 1$. So, x intercept is going to be 6, this and the y intercept is going to be 4, which is this and so our line is this is our l1. Now, if we try to plot the other equation, here, again, I will get $5x - 10y + 5 = 0$, $\frac{x}{-1} + \frac{y}{1/2} = 1$.

So, here we have this is the x intercept, whereas this is the y intercept 0.5 here. So, this is our line equation 2. And clearly the intersection is happening here at this point, which is you can see this is (3, 2). So, in this way, you can try to find the point of intersection of any 2 given lines. However, you are likely to run into a bit of trouble in 2 cases, and let us see those 2 cases.

(Refer Slide Time: 9:21)

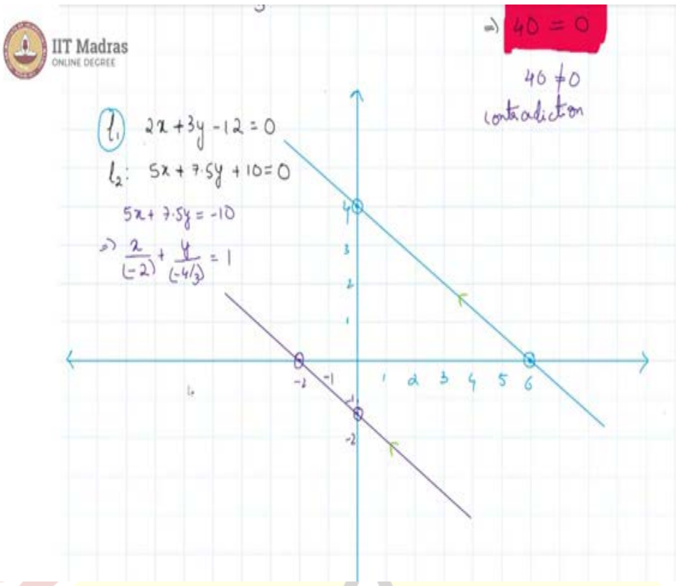


Consider these 2 line equations, 11 is still $2x + 3y - 12 = 0$, whereas $5x - 7.5y + 10 = 0$. If we try to solve this using the substitution method, for example, we would get, I would, let us say I try to eliminate the variable x in which case I should be doing to $2x - 12 = -3y$, which would indicate $y = \frac{12-2x}{3} = 4 - \frac{2x}{3}$. And substituting this in l2, I will get from l2, this is from l1.

And now in l2, if I substituted this, I would get $5x + \frac{15}{2}\left(4 - \frac{2x}{3}\right) + 10 = 0$. This gives us $5x + 30 - 5x + 10 = 0$. And you see that $5x$ and $-5x$ cancels and we come at the strange contradiction where $40 = 0$. And this is not okay right. We know that $40 \neq 0$. So, there is some contradiction we are arriving at.

And what does this contradiction indicate? It indicates that there is no point for which these 2 lines meet. So, you cannot find a point of intersection for these 2 lines. So why is that? That is because they are parallel. If we plotted these lines,

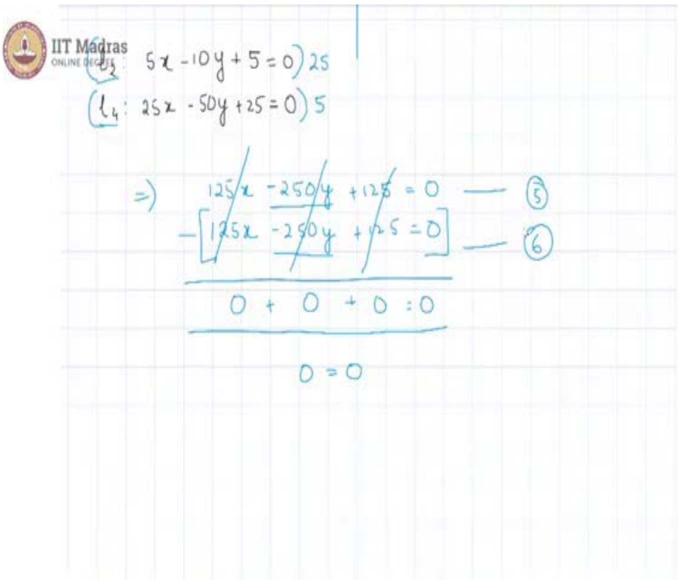
(Refer Slide Time: 11:51)



We know that for l1, the intercepts are 6 and 4, respectively. So this is 1, 2, 3, 4, 5, 6. So this is our intercept for l1, x intercept for l1 and y intercept for l1 is 4 1 ,2, 3 and 4. For l2, we have to see now for l2, we get $5x - 7.5y + 10$, which indicates $\frac{x}{(-2)} + \frac{y}{(-4/3)} = 1$.

So, in $x = -2$, so this would be our point and in $x = -\frac{4}{3}$is a little below -1, which is about one third the way from -1 and -2. So, this would be it. If we plotted these lines now we see that these are, in fact, parallel lines. They just do not meet anywhere, which is why when you try to solve for a point of intersection, you get a contradiction. So here, we can say that there is no solution for this system of linear equations.

(Refer Slide Time: 13:30)



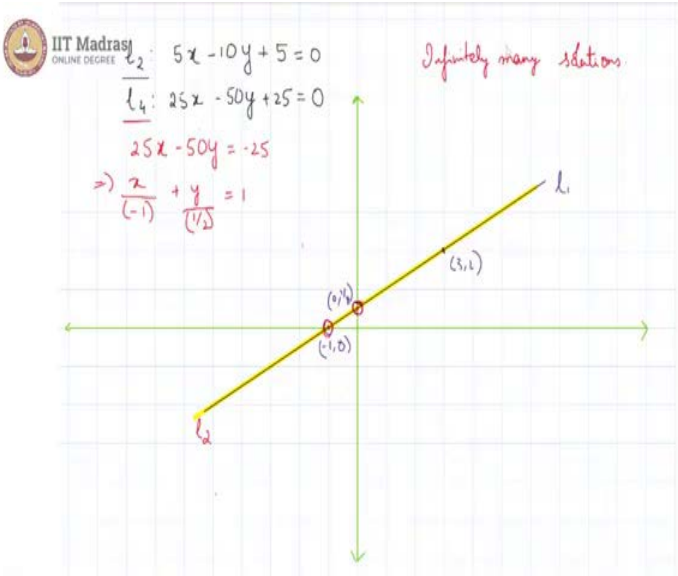
Now, in the third case, let us look at a line equation which is our l2 earlier that was $5x - 10y + 5 = 0$. And there is some other equation l4 let us call it, which is $25x - 50y + 25 = 0$. So, when we solve for these 2 equations, now let me try the elimination method. So, I am going to get 2 equations, then one is $125x - 250y + 125 = 0$. And here I am going to get another one, $125x - 250y + 125 = 0$.

We have the same coefficient for y. So if I attempted to subtract this equation entirely, I will get 0. So, I have this statement, which is always true. Unlike the previous case where it was never true, 40 was never going to be equal to 0, here I get a statement, which is always true, which is $0 = 0$, independent of the coordinates of x and y.

And this means something similar to the previous case, but not exactly the same. What is happening here is since this is always true, it means there are infinite solutions for these 2 equations. If you observe what is actually happening is l2 and l4 are the same line, which is why we got this entirely identical equations, both of these, let us call this equation 5 and let us call this equation 6. And we see that equation 5 and equation 6 are the same, there is no difference, which means our 2 original lines are coinciding.

If they are the same line, then we will get infinitely many points which satisfy both of them. So we have infinitely many solutions for these 2 lines. So whatever x you take, you are going to get a solution for that x. So in the graph, this is what is going to look like.

(Refer Slide Time: 16:28)

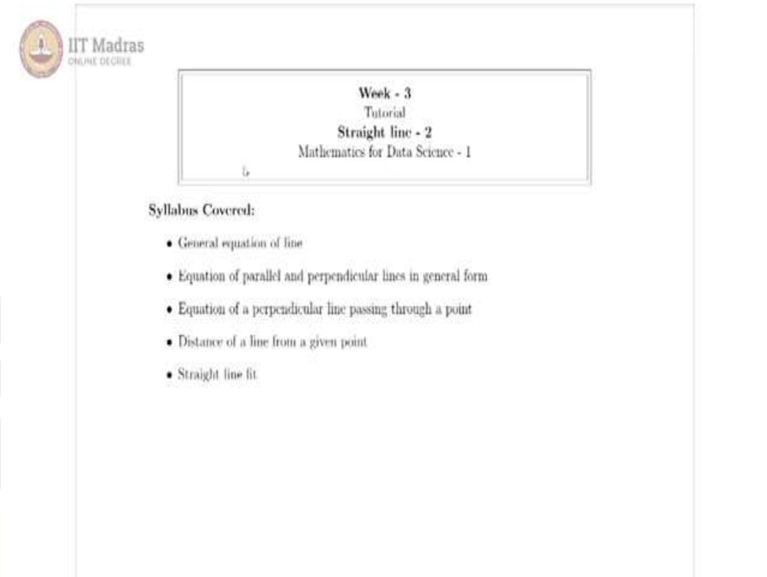


We know the intercepts of our l2, which is -1 and y intercept was half, so this would be our l1, it is passing through $(-1, 0)$, and also $(0,\ 1/2)$. And as we had found earlier, it is passing through $(3, 2)$ as well. Now let us consider the other equation. Now let us consider the other equation which is l4, and we will have $25x - 50y = -25$. This gives us $\frac{x}{(-1)} + \frac{y}{(1/2)} = 1$.

So, again we get the same intercepts. Thus, l2 will have to coincide entirely with l1. And that is what is happening, they are the same line. So, we get infinitely many solutions when we get a true statement, an always true statement independent of x and y in case of the same line, that is both line equations are representing the same line.

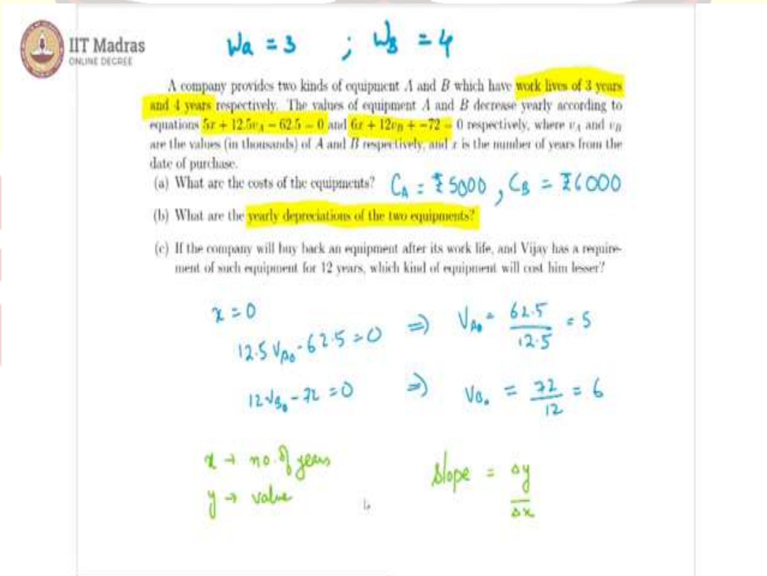
Mathematics for Data Science 1
Week-03
Tutorial-01

(Refer Slide Time: 0:16)



Hello, mathematics students. This is a tutorial for week 3, where we will be doing more straight line concepts problems. Primarily, this is the syllabus that has been covered here. Let us begin with our first question.

(Refer Slide Time: 0:31)



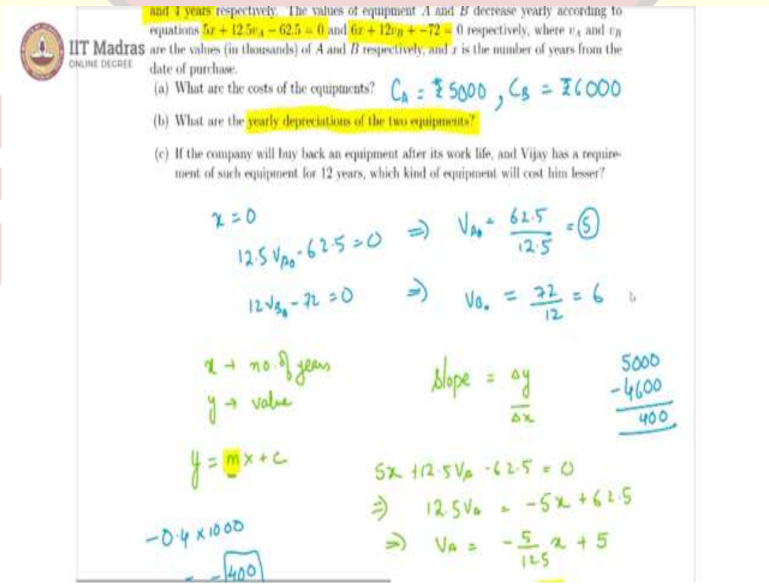
There is a company with two kinds of equipment, A and B. And they have work lives of 3 years and 4 years respectively. So, work life of A is, let us call it W_a is 3, W_B is 4 years. Further, the values of equipment A and B decrease yearly according to these equations. These are our equations, where v_A is supposed to be the value of A and v_B is supposed to be the value of B in thousands, respectively, and x is the number of years for which that value is applicable.

So, what are the costs of the equipments? So, the cost of the equipments would be v_A and v_B values when x is equal to 0, that is, when you just bought it, what is the value of the equipment. So, we just take x is equal to 0 and from this we get $0.5v_A$ - 62.5 = 0, this would give us v_A is equal to, to indicate that this is the initial time I am going to make it A_0, v_{A0} so yes, this is v_{A0} and that is 62.5 /12.5, which is equal to 5.

Therefore, the cost of A, I will call it C_A is rupees 5000. Now, let us work with B. Same thing again, we take x is equal to 0. So, we have $12v_B$ - 72 = 0, this will imply v_B again we are calling v_{B0} to indicate the initial cost that would be 72 /12 which is equal to 6. So, C_B, the cost of B is rupees 6000. Going further, we are asked what are the yearly depreciations of the two equipments.

So, yearly depreciation basically means how much value is decreasing each year. So, let us look at that. Here, in this case, x is number of years, whereas y is the value. So, what is being asked in a yearly depreciation is the change in y for a unit change in x, which is basically just a slope. Because slope is changing y, Δy by changing x. So, when Δx is equal to 1, Δy is equal to the slope.

(Refer Slide Time: 3:56)



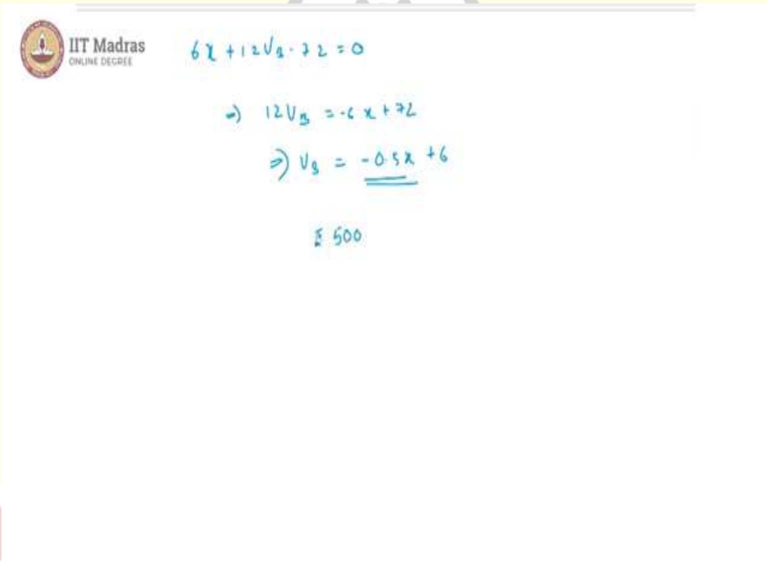
So, we can find this by just finding the slope for each of those two linear equations. And for the slope, we convert our equations to the y = mx + c form, then the m is going to be the slope.

So, one equation is 5x + 12.5v_A - 62.5 = 0. This would indicate that 12.5v_A = -5 x + 62.5. Going further then, it will have v_A = -5x / 12.5 + 62.5 /12.5, we had already seen it to be equal to 5.

So, that is equal to - 0.4 x + 5 = v_A. So here, we are, our m in the equation is basically -0.4. So, this is the reduction in one year, -0.4 ×1000 because we are taking everything in thousands, so, that is basically -400. So, this is the depreciation, 400 is the depreciation every year for the company one, we can also verify this by looking at the values of v_A for year one.

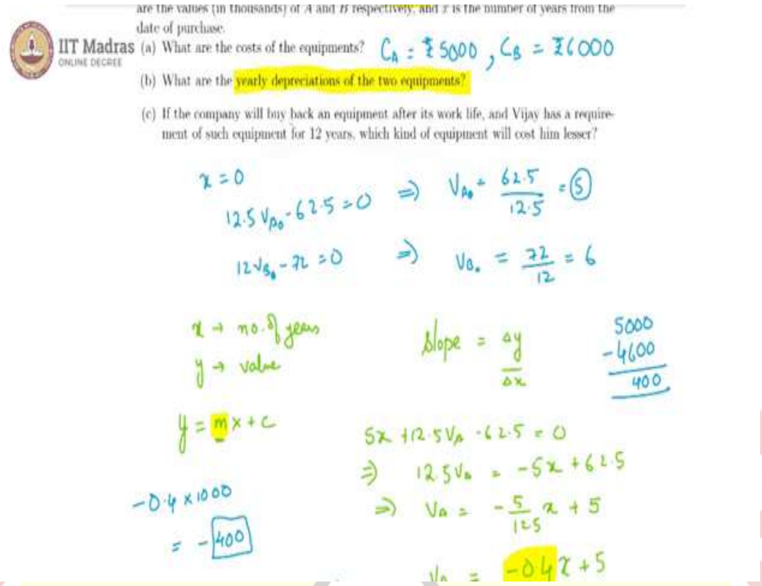
So, when x = 1 we have 5 +12.5v_A - 62.5 = 0, this gives us v_A = 57.5 / 12.5 which is equal to 4.6. So, v_A was originally 5, that means it was originally 5000 rupees and after 1 year it became 4.6 which is 4600 rupees. So, the difference is 400 rupees. So, that is the yearly depreciation for the first equipment.

(Refer Slide Time: 6:48)



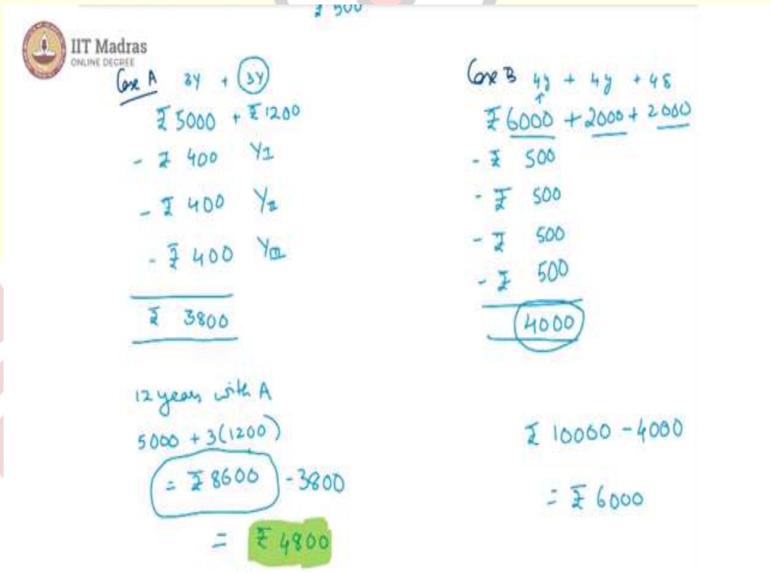
Now, let us look at the second equipment now second equipment the equation was 6x + 12v_B - 72 = 0. Again, if we put this to the y = mx+ c form, the slope intercept form we will be getting first we have to do 12v_B = -6 x + 72. This indicates v_B = -0.5 x + 6, thus -0.5 is the slope here. Which means 500 rupees is the yearly depreciation.

(Refer Slide Time: 7:42)



In the last part, they said that the company will buy back the equipment after its work life. And Vijay has a requirement of such equipment for 12 years. Which kind of equipment will cost him lesser.

(Refer Slide Time: 7:58)



So, in the case of the first equipment, let us call it case A, and here let us have case B to consider. And in case A the initial cost was 5000 rupees and each year there is a decrease of 400 rupees. So, in first year we lose this much, in the second year we lose another 400 rupees and at the end of the third years, there is a loss of another 400 rupees. And we are aware that 3 years is a worklife for A, whereas for B it is 4 years. This is to say that at the end of 3 years, the value of the machine is 3800 rupees.

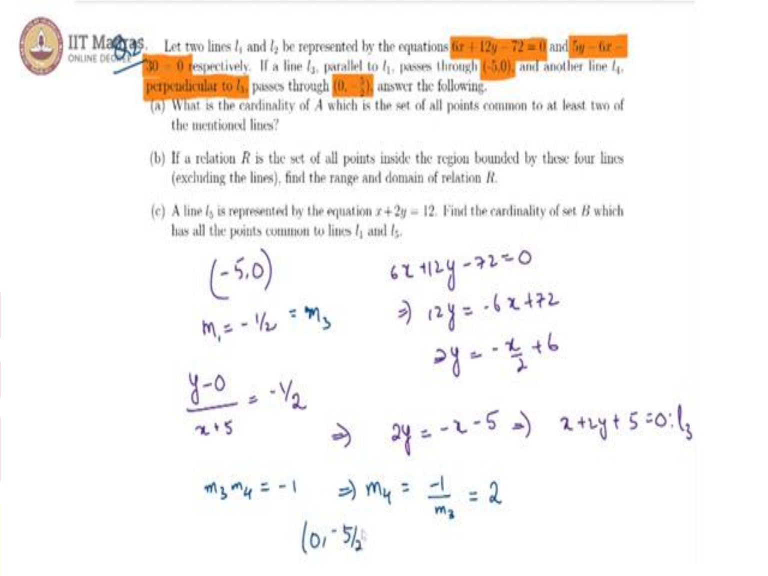
So, if now, Vijay buys the equipment afresh, then and the company is buying back this 3800. All that Vijay needs to spend now is rupees 1200 and this way he gets an additional 3 years. So, with 5000 he got 3 years and now another 3 years this way. So, in order to get 12 years with equipment A, the total money that Vijay will require to spend is 5000, which is the initial first 3 years and from then on 3 years plus 3 years plus 3 years because it is totally 12 years.

So, 3 times 1200, that is rupees 8600 in case of A, whereas in B, B is more expensive. So, we have 6000 and every year there is a loss of 500 rupees in value and this is required to be done 4 times because the work life for B is 4 times. So, we are effectively subtracting 2000 rupees from the original value. So, we have 4000 at the end of it, which means for the first 4 years there is an expenditure of 6000 but then, for the remaining 8 years, there has been only 2000 each.

This is so because the product's value is already 4000 rupees and in order to get a new version of equipment B, Vijay only has to spend 2000 rupees. So, total expenditure in this case is going to be 10,000 rupees because 6000 + 2000 + 2000. Here, we are not supposed to forget one thing though, that is the end of, after these 3 years, 3 years pass, at the end of 12 years, he can sell it off for 3800. So, we are supposed to further subtract 3800 here and likewise here, we can sell it off for 4000. So, here we get rupees 4800 whereas, here we get rupees 6000. So, the expenditure is clearly lesser for A. So, A would be the good choice for Vijay.

Mathematics for Data Science 1
Week-03
Tutorial-02

(Refer Slide Time: 0:16)

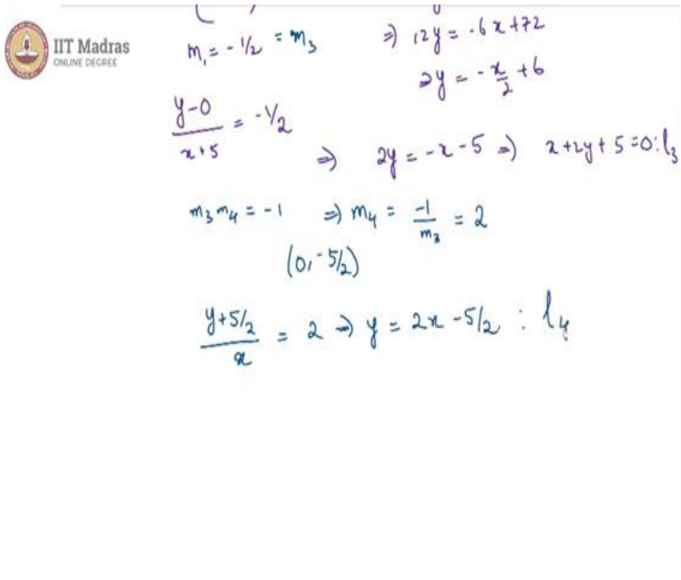


And for our second question, there are 2 lines, and these are the equations, which represent our lines. And a line l_3 is parallel to l_1, and passes through (-5, 0). Now we can find l_3, by using the point slope form, we already have the point, which is (-5, 0). And we can also find the slope from l_1 slope, we already have l_1. And we can write, so l_1 is this, $6x+12y-72=0$, which tells us that $12y=-6x+72$.

And that gives us $y=-\frac{x}{2}+6$ so, the slope here is $-\frac{1}{2}$, because $y=mx+c$. So, slope is $-\frac{1}{2}$. Now if we did point slope form on this, we would get $\frac{y-0}{x+5}=-\frac{1}{2}$ which indicates $2y=-x-5$. So therefore, $x+2y+5=0$ is basically our line l_3. And now if we look further, we have line l_4 which is passing through this point, and it is perpendicular to l_3.

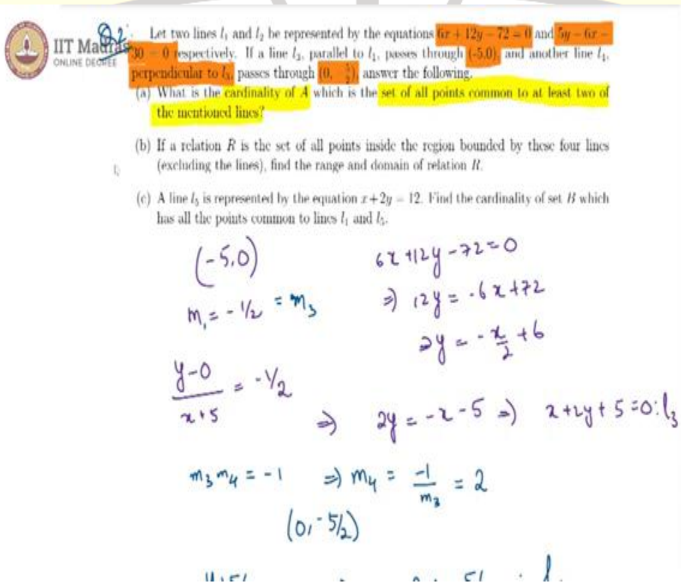
So, if we took this to be $m_1=-\frac{1}{2}=m_3$ because m_1 and m_3 are the same slope. And let us consider the slope of l_4 to be m_4, so we can say $m3 \times m4=-1$, because they are perpendicular, that would indicate $m_4=-\frac{1}{m_3}$, which is basically 2. So we now have the slope of l_4. And it also goes through this point.

(Refer Slide Time: 2:57)



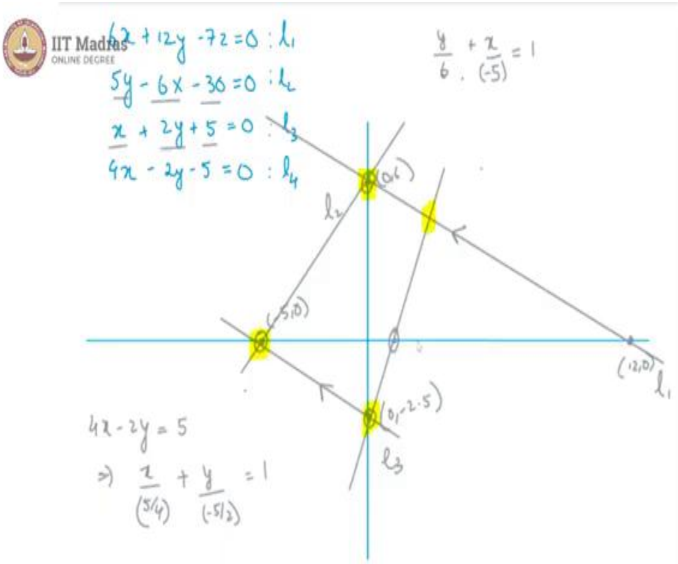
So again, using point slope form, we have $\frac{y+\frac{5}{2}}{x} = 2$, that would indicate $y = 2x - \frac{5}{2}.$ So this is our l_4.

(Refer Slide Time: 3:14)



Now, the question is being asked is, what is the cardinality of A, which is a set of all points common to at least 2 of the mentioned lines.

(Refer Slide Time: 3:28)



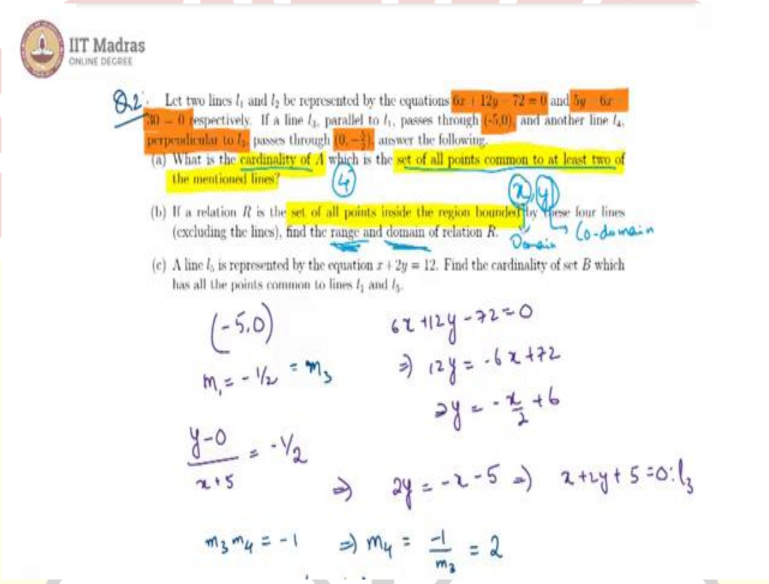
For that, let us try to draw our lines on the graph. $6x+12y-72=0$ would give us if $x=0$, it gives us $y=6$ which means some point let us call this here is (0,6), it goes through this point. And if $y=0$, you get $x=12$. So that would be some point here. So, our l_1 is this line. And now we know l_3 is parallel to this line. So l_3, if we, again did the same thing of putting $y=0$, x becomes -5, which is somewhere here.

So, as you can see, I am doing this on a rough estimate. I am not trying to be accurate, but even a rough estimate can work out here, because you might not always find graph paper when you require it. So often developing an intuition for the rough estimates is a good idea to solve problems. Now, this is one point and when $x=0$, y becomes -2.5, which is somewhere like this. So we have (0, -2.5). As you can probably see from our last rough estimate itself that these do appear to be parallel, they seem to be in the same direction.

Now, l_2 if we look into it with a similar logic, we can see that l_2 can be reduced to $\frac{y}{6}-\frac{x}{5}=1$. So in our intercept form, we can now tell that if I made this plus, this becomes -5, so the x intercept is -5, which is this point, again, and y intercept is 6, so that is this point. So, l_2, in fact, passes through these 2 points. So, this is our l_2. So, this was l_1 now, this is l_3 and this is l_2.

Lastly, let us reduce our l_4 into the intercept form, we get $4x-2y=5$, therefore, $\frac{x}{5/4}+\frac{y}{-5/2}=1$. So, when we look at this then 5/4 is a quantity just a little greater than 1, so it is probably somewhere here and 5/2 is a 2 and a half basically. So, -2.5, so this and this plus we have something like this happening. So, overall there are four points, which are common to any pair of these four lines.

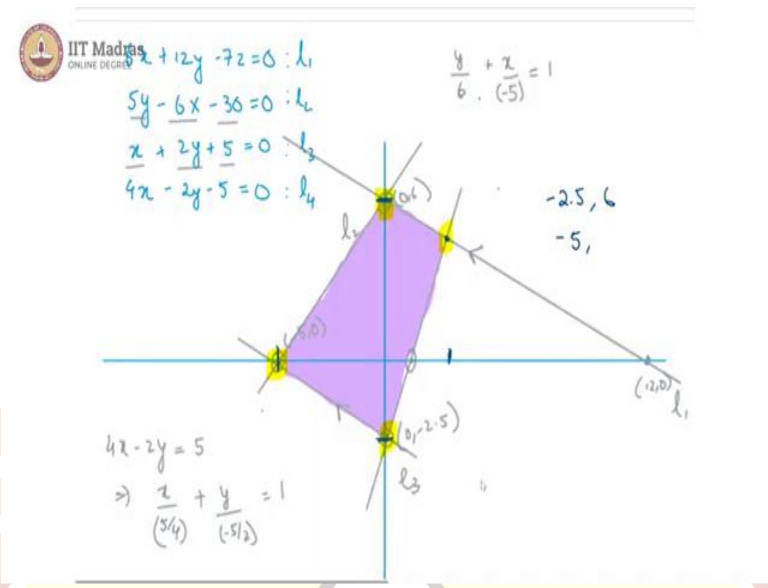
(Refer Slide Time: 7:02)



So, our question, the cardinality of A, where A is a set of all points common to at least 2 of the mentioned lines. So, that would be 4, there are 4 points of intersection here. Now, if R is a relation, and it is the set of all points inside the region bounded by these 4 lines. So, here we are, when we say relation, we are basically saying every point in the set when is taken as a ordered pair like this (x, y), then x would be from the domain of the relationship and y would be from the co-domain.

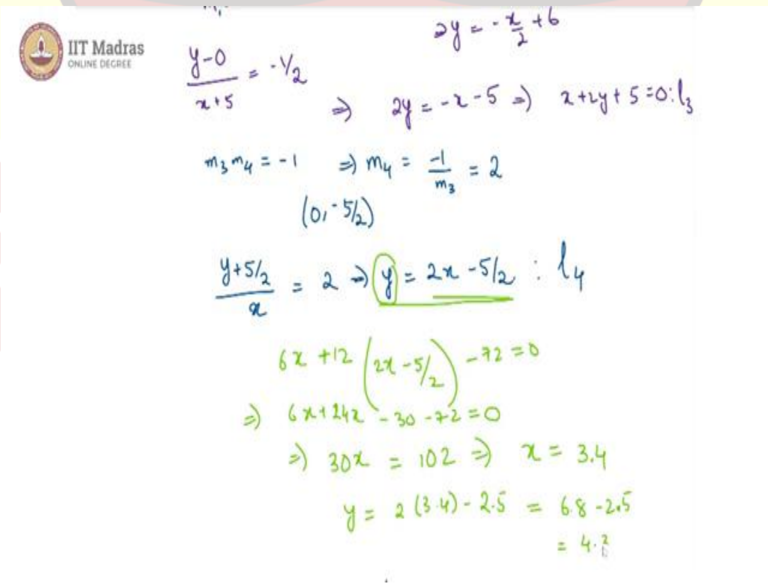
So, this is seen as a relation from the set of x values and to the set of y values. And now, we are asked to find the range and domain of relation R, which is to basically find when we say range, all the possible y values and the domain is all the possible x values.

(Refer Slide Time: 8:16)



So, here in this region that we are looking at, the possible y values would be between this value and this value. So, all possible y values are between -2.5 and 6, whereas the possible x values are between this point and this point, that is between -5 to some particular quantity, which is the x coordinate of this point. And that point is the intersection of l_1 and l_4. So, let us try to solve l_1 and l_4 to find that point of intersection.

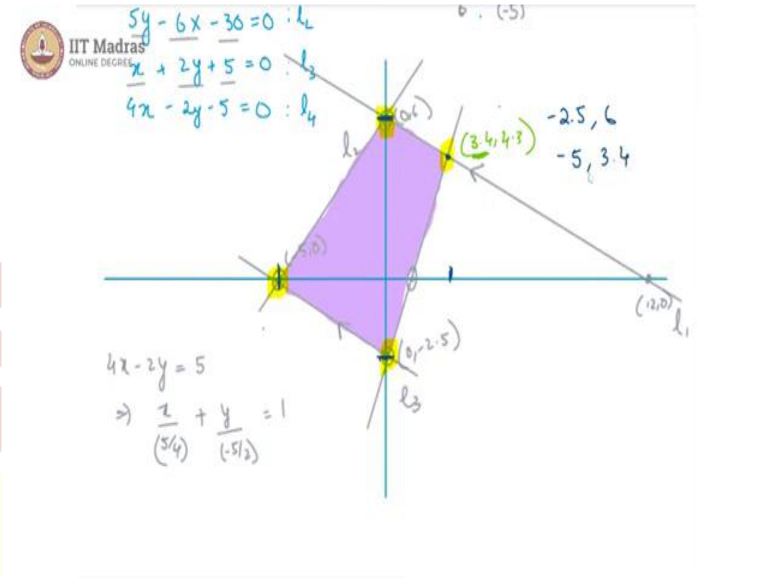
(Refer Slide Time: 9:03)



We know that this as l_1, and this is l_4 and from l_4, we know that y is basically $2x-\frac{5}{2}$. If we substituted this into l_1 we would get $6x+12(2x-5/2)-72=0$. This would give us

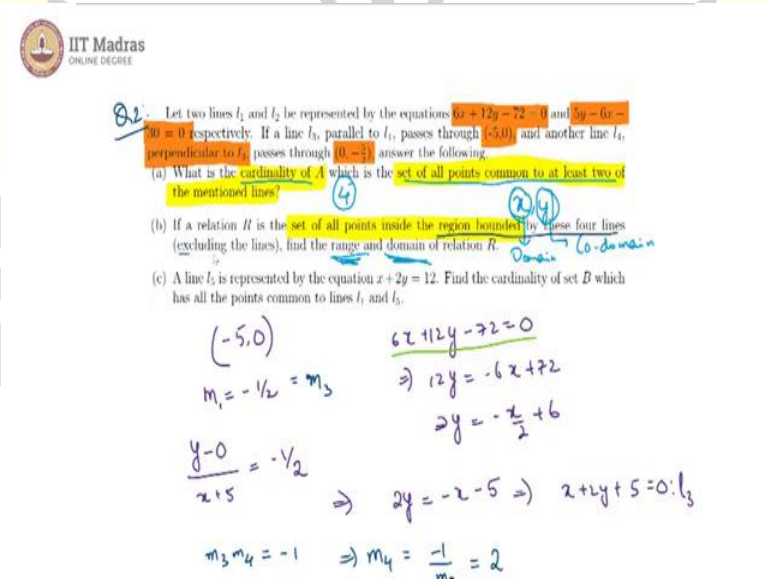
$6x + 24x - 30 - 72 = 0.$ That indicates $30x = 102$ which indicates $x = 3.4.$ Correspondingly, y would then be $2 \times 3.4 - 2.5$, because 5/2 is 2.5, which gives us 6.8 - 2.5, which is equal to 4.3.

(Refer Slide Time: 10:24)



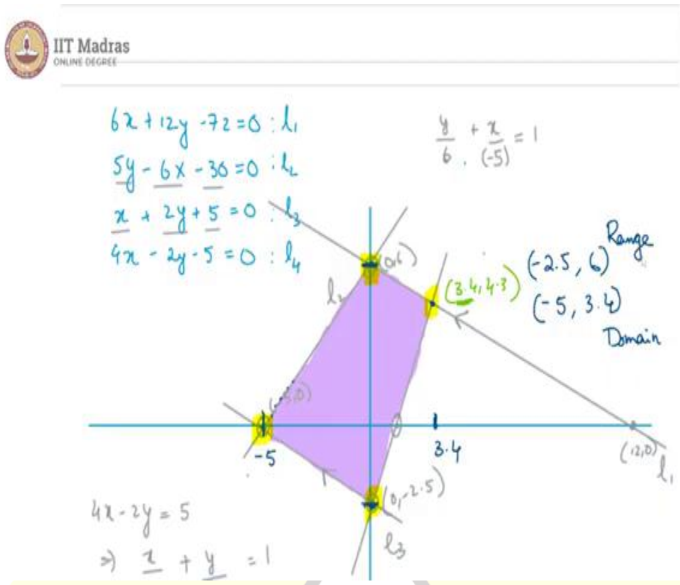
So this point here is (3.4, 4.3) and we only require the x value. So the x values range from -5 to 3.4.

(Refer Slide Time: 10:42)



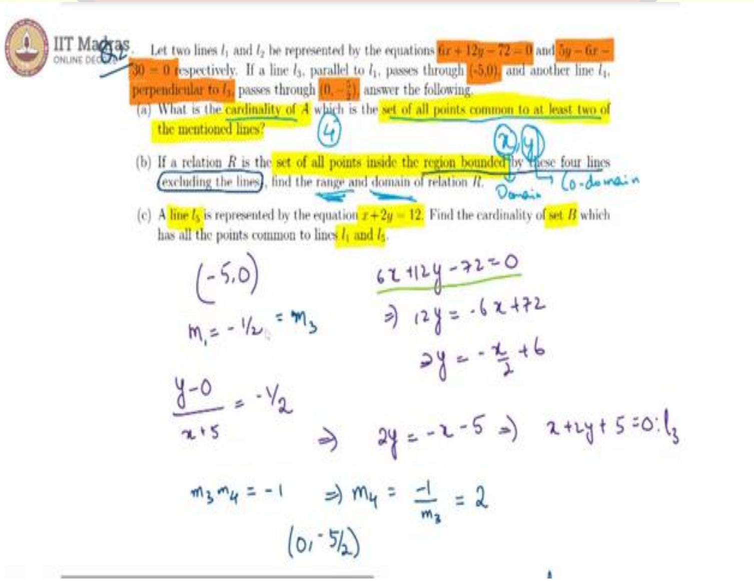
However, one important thing we need to look for here now is the region bounded by these 4 lines, but excluding the lines themselves.

(Refer Slide Time: 10:52)



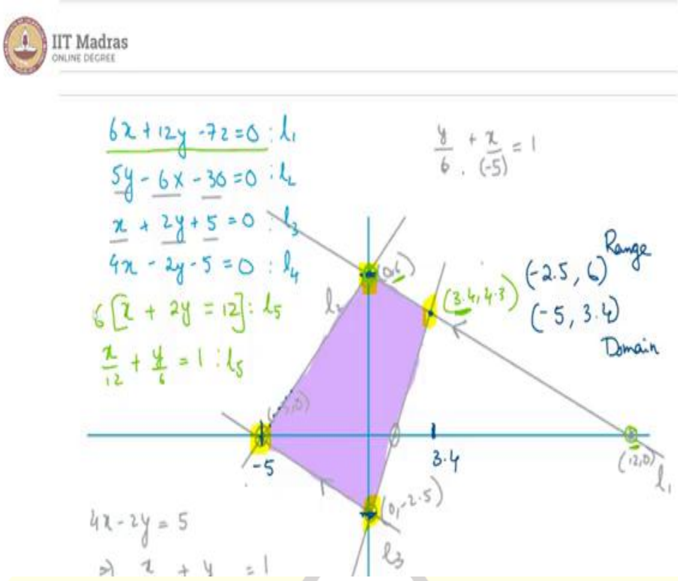
Which means -2.5 and 6 themselves do not fall into our domain because we are not interested in the points on the curve. So this point is on the curve, this point is on the curve, but it is not inside, similarly, for each of these, because they are the border points. So, -5 is not an x value inside the domain. Similarly, 3.4 is not a value inside the domain. So, our domain is the $(5,3.4)$. Likewise, -2.5 is not a y value inside the range and 6 is also not a y value inside the range, so our range is (-2.5, 6).

(Refer Slide Time: 11:50)



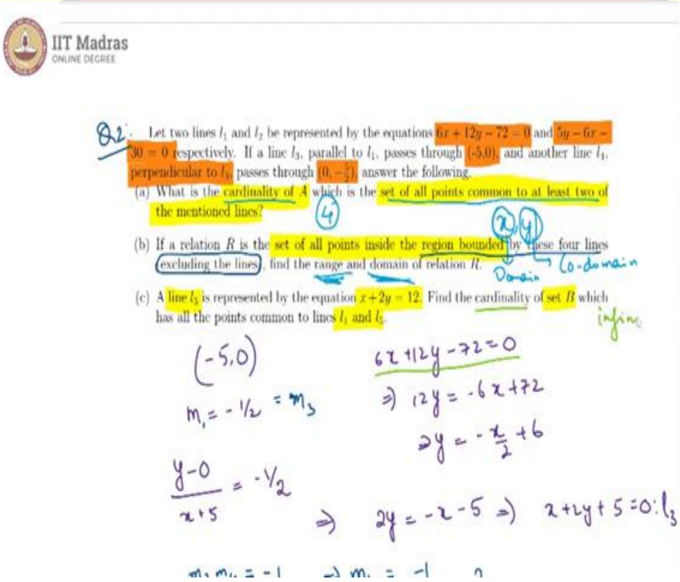
Lastly, there is a line l_5 represented by this equation given to us find the cardinality of set B, which has all the points common to l_1 and l_5.

(Refer Slide Time: 12:04)



Let us look at l_1 and l_5. l_5 is given as $x + 2y = 12$. Now, if we applied our intercept form again, we would get $x/12 + y/6 = 1$. Let us look at that $x/12$ indicates x intercept of l_2, $y/6$ indicates y intercept of 6. So, we see that l_5 is basically the same line as l_1, indeed if you multiply this whole equation with 6, you will just get the form of l_1. Therefore, l_1 and l_5 are the same lines.

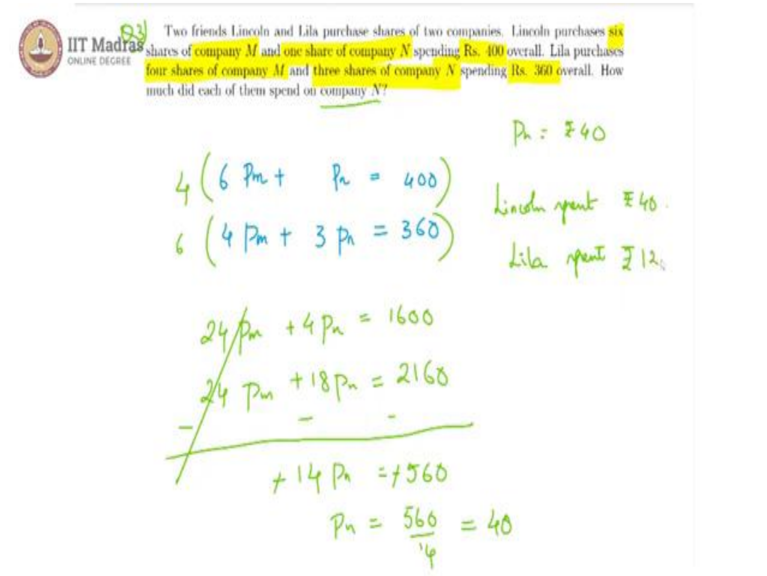
(Refer Slide Time: 12:54)



Then, the question is asking, find the cardinality of set B, which has all the points common to the lines l_1 and l_5. There are infinite points because they are the same line. So, the cardinality of set B is infinite.

Mathematics for Data Science 1
Week-03
Tutorial-03

(Refer Slide Time: 0:14)



Now, third question, you have two friends Lincoln and Lila who purchase shares of two companies. Lincoln purchases six shares of a company M and one share of company N and overall spends 400. This can be encapsulated as if the company M's share price is P_m and for n that is P_n we can say that $6P_m + P_n = 400.$ Then for Lila there is four shares of Company M and three shares of Company N coming to 360.

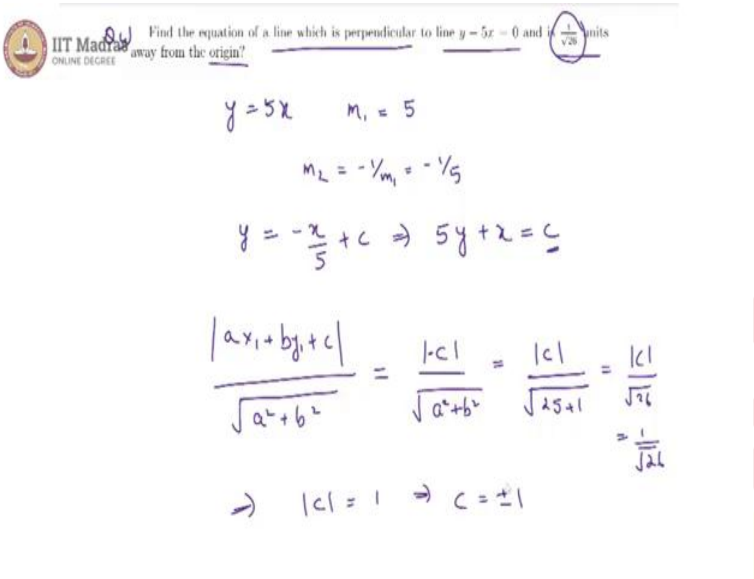
So, for Lila we have $4P_m + 3P_n = 360.$ How much did each of them spend on n? So, we need to know what is P_n and $3P_n$, that is what we are interested in. To find the values of P_m and P_n we will require to solve these two linear equations. However, we only required to find P_n because the question is only pertaining to the company N's shares. So, we can work towards eliminating the P_m variable from these two equations.

So, we can multiply this equation by 4 and this one by 6 because $4\times 6 = 24,$ $6\times 4 = 24$ and that way we should be able to subtract $24P_m$. So, we are going to get from the first equation $24P_m + 4P_n = 1600,$ whereas, from the second equation we get $24P_m + 18P_n = 2160.$ Now, if we subtract second equation from the first we get these two canceling off and here we get $-14P_n = 560.$

And this indicates that $P_n = 560/14$ because we can cancel out the plus and the plus and that is equal to 40. So, P_n is 40 rupees per share. And now since Lincoln has purchased only one share, Lincoln spent only 40 rupees on company N, whereas, Lila spent three times that which is rupees 120.

Mathematics for Data Science 1
Week-03
Tutorial-04

(Refer Slide Time: 0:15)



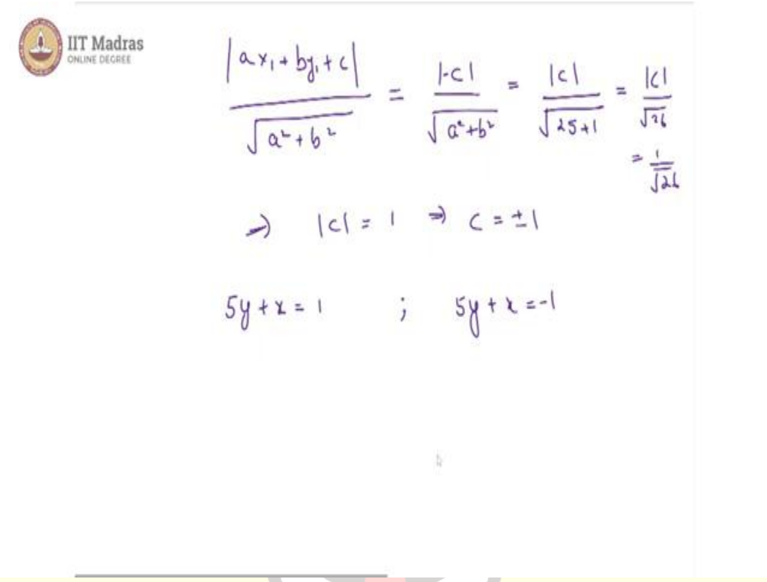
For our fourth question, we want the equation of a line which is perpendicular to this line, and is at this distance from the origin. So, from $y-5x=0$, we get $y=5x$, so therefore, the slope m_1 is 5. And if our line is perpendicular to it, then our line m_2 must be $-1/m_1$, which is equal to -1/5. So, we know that our line is some $y=-\frac{x}{5}+C.$

If we kind of simplify it, we are going to get $5y+x=C$, this C is not the same thing as the previous C, I have just used that as C because it is an arbitrary constant, which is yet to be determined, otherwise it should have been 5C. Anyway, now we have to find the value of this C in this equation. For that, we are going to use the next bit of information that is given to us, which is the distance from the origin.

So, this line has this distance from the origin. So the distance from a point formula is $\frac{|ax_1+by_1+c|}{\sqrt{a^2+b^2}}$, where (x_1, y_1) is the point from which we are measuring the distance for this line. So, in our case, (x_1, y_1) is (0, 0) because we are doing from the origin. So in our case, we get modulus of $\frac{|-c|}{\sqrt{a^2+b^2}}$, So, modulus of $|-c|$ is just the same thing as $|c|$.

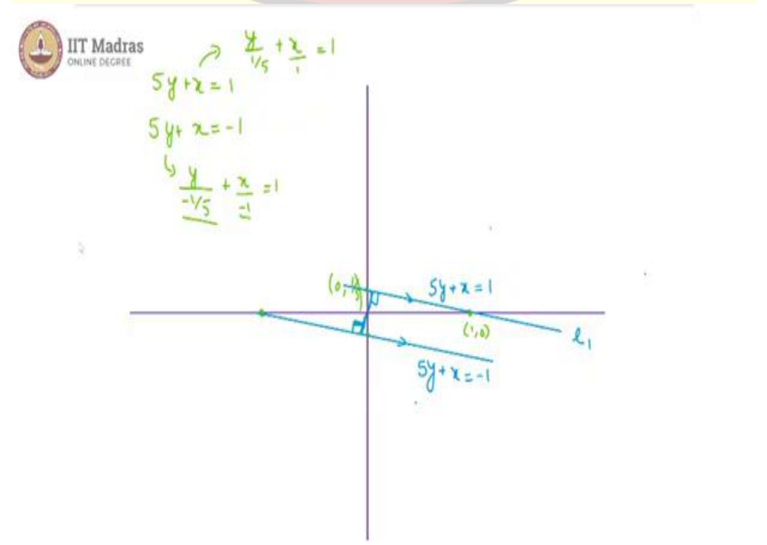
And root of $\sqrt{a^2+b^2}$, in our case comes out to be $\sqrt{25+1}$,, that is $\sqrt{26}$. So we have $\frac{|c|}{\sqrt{26}}$, this is given out to be $\frac{1}{\sqrt{26}}$, which would imply $|c|=1$, and that would imply $c \pm 1$. So, we get two answers.

(Refer Slide Time: 2:49)



What are the two answers? One is for c being $+1$, we have $5y+x=1$. And in the other case, we get $5y+x=-1$. for the other choice. So, how does this happen, what is actually happening here to try to plot our lines?

(Refer Slide Time: 3:08)



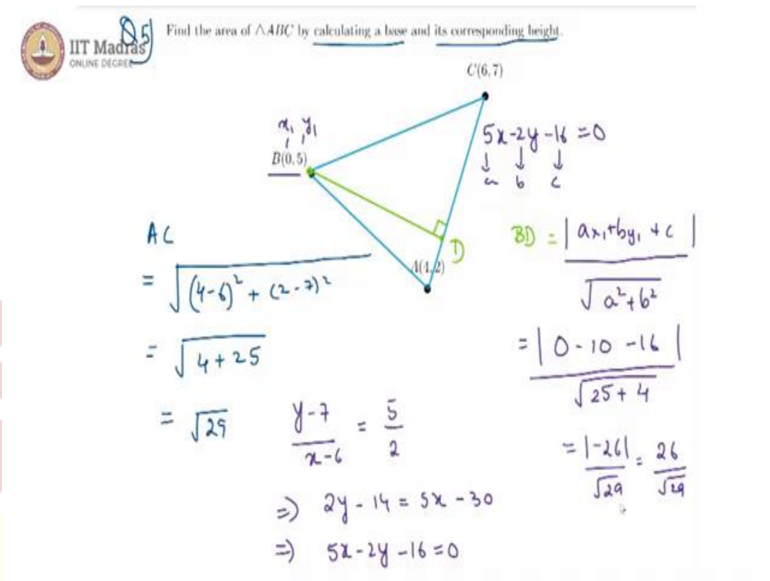
So, we have two lines, which we were looking at, which is $5y + x = 1$ and $5y + x = -1$ and from this we get the intercepts to be, the intercept form of this would be $\frac{y}{1/5} + \frac{x}{1} = 1$. And in this case, we get $\frac{y}{-1/5} + \frac{x}{-1} = 1$. So, in one case, we have a y intercept of 1/ 5. So, let us assume this is 1/5, then x intercept is 1 which is 5 times of that, so that so it must be somewhere here, so this would be (1, 0) and this is (0, 1/5).

And our line is going through these two points, giving us something like this. Let us call this l_1 and where do we get the $\frac{1}{\sqrt{26}}$ distance from the origin, we get it when we measure it perpendicularly from the origin. Now, let us look at the other equation. So -1/5, so, this should be exactly below this this way and this is -1, so this would be exactly opposite in this way at the same distance.

So now we have these two points, so we can also construct this line, which goes this way. And as you can see, they are both parallel and exactly opposite to that $\frac{1}{\sqrt{26}}$ you get this distance which is again perpendicular distance and it is also at $\frac{1}{\sqrt{26}}$. So, we have two lines which satisfy our requirements, one is $5y + x = 1$, the other is $5y + x = -1$.

Mathematics for Data Science 1
Week-03
Tutorial-05

(Refer Slide Time: 0:14)



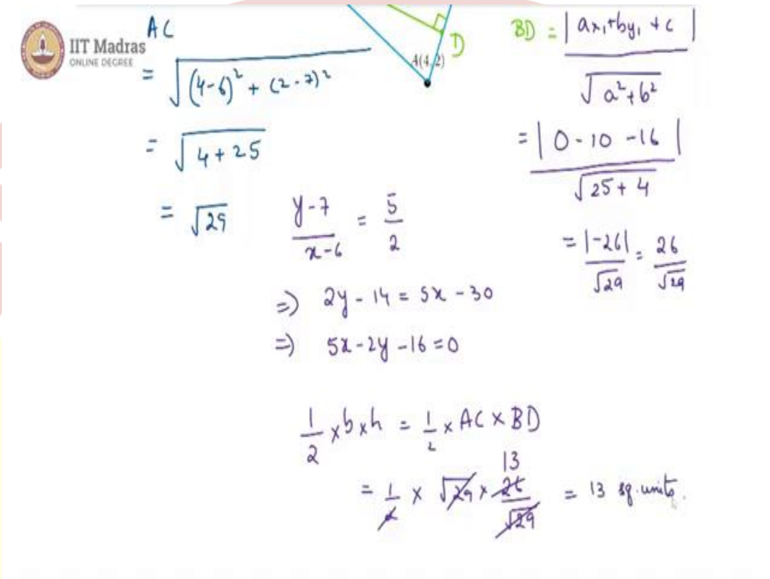
In this problem, should be our fifth question. Suppose to find the area of ABC, there are three points here. So, we need to make that triangle, our triangle would look something like this. But to find the area, we are supposed to calculate a base and its corresponding height. So, we are not supposed to use the formula which involves the three coordinates, instead, we will take any of these sides to be the base. So, let me take AC to be the base. So, we need to find the base length, which is AC, that would be by Euclidean distance formula, $\sqrt{(4-6)^2+(2-7)^2}$.

So, this comes out to be $\sqrt{4+25}$. So, that gives us $\sqrt{29}$ is the base. Now the altitude, the height from B would be something like this, let us call this point D and this is 90 degrees, so B to D that length would be the height. So, BD is going to be the distance of the point B from the line AC, the shortest distance of point B, the line AC. So, for this we can use the distance formula of a point from a straight line. However, we first need to find out the equation of AC.

For that, let us use the 2 point form because we have 2 points, we will get $\frac{y-7}{x-6}=\frac{7-2}{6-4}=\frac{5}{2}=$ 2.5 . Anyway, if we cross multiply, we get 2y-14=5x-30, which gives us the equation to be 5x-2y-16=0 and the distance of (0,5) which is our B from this particular line. So, this line is our 5x-2y-16=0 .

So, that distance can be calculated from the formula, which is the $\frac{|ax_1+by_1+c|}{\sqrt{a^2+b^2}}$. So here a is our 5, b is -2 and c is -16. So, substituting and (x_1,y_1)is our coordinates of B, this is x_1 and this is y_1, so the coordinates of B. So here we get $\frac{|0-10-16|}{25+4}$, which then gives us $\frac{|-26|}{\sqrt{29}}$, $|-26|$ is then 26. So, this would be the height.

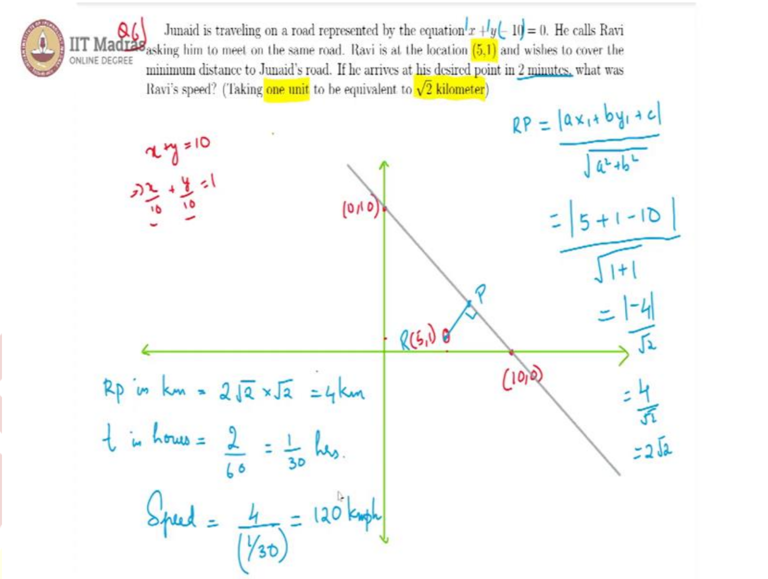
(Refer Slide Time: 3:48)



Combining these two quantities, we get our area as half into base into height, which will then be $\frac{1}{2} \times \text{AC} \times \text{BD}$, which then gives us $\frac{1}{2} \times \sqrt{29} \times \frac{26}{\sqrt{29}}$, $\sqrt{29}$ and $\sqrt{29}$ cancels off, 2 cancels with 26 giving us 13. So, we get 13 square units as the area of our triangle ABC.

Mathematics for Data Science 1
Week 03 – Tutorial 06

(Refer Slide Time: 00:14)



Sixth question. We have Junaid who is traveling on a road represented by the equation x+y-10=0. So, in the graph if we plot that, we can see that x+y=10, which gives us $\frac{x}{10}+\frac{y}{10}=1$, which means the x intercept and y intercept are both equal to 10. So, if this is 10 and this is also 10, so this would be (10,0). Whereas this is (0,10) and the line that passes through them is the road that Junaid is traveling on.

So, this is the line that Junaid is traveling on and he calls Ravi, asking him to meet on the same road. But Ravi is at this point. So, that would be 5, it would be somewhere here halfway and 1 would be somewhere here. So, this is 1 and this would become our location of Ravi, that is (5,1) and Ravi wishes to cover the minimum distance to Junaid's road. So, we know that minimum distance is achieved when you go perpendicular that is normal to the other line.

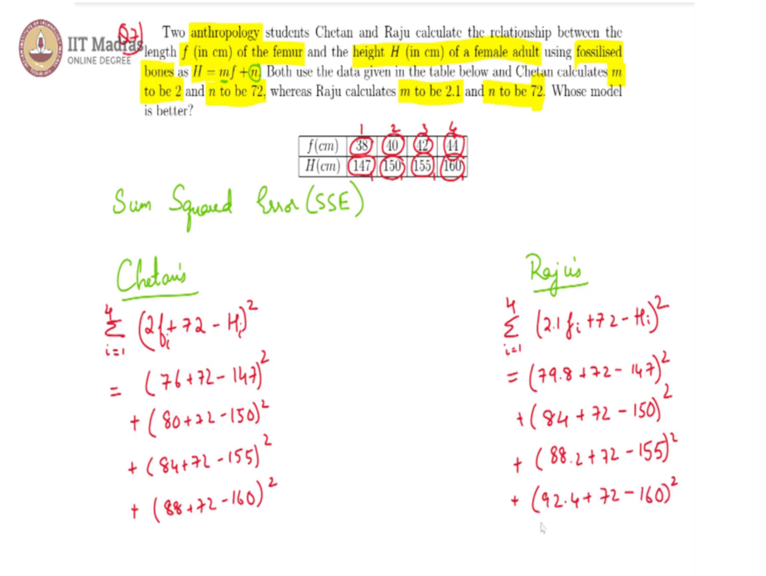
So, we can see that Ravi goes along this path and intersects, that path intersects somewhere over there, let us call this point P and he arrives at this point P in 2 minutes and we are being asked, what is Ravi's speed? So, we need to first find out, assuming Ravi's original location (5,1) is R, if we find out what RP is, then we should be able to find out the speed. So, RP is basically the shortest distance of R from this particular line. So, we can calculate it from that formula, which is $\frac{|ax_1+by_1+c|}{\sqrt{a^2+b^2}}$.

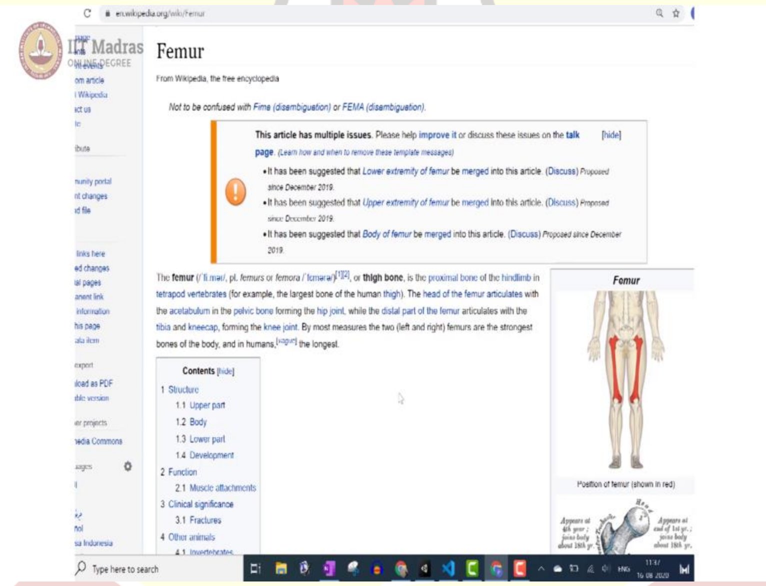
And now, here a is 1, b is also 1, c is -10. So, we have this is equal to $\frac{|5+1-10|}{\sqrt{1+1}}$, because a is 1 and b is 1, the squares are also 1. So, we get $\frac{|-4|}{\sqrt{2}} = \frac{4}{\sqrt{2}}$, which is then equal to $2\sqrt{2}$ units and now it is given that 1 unit is equivalent to $\sqrt{2}$ kilometres.

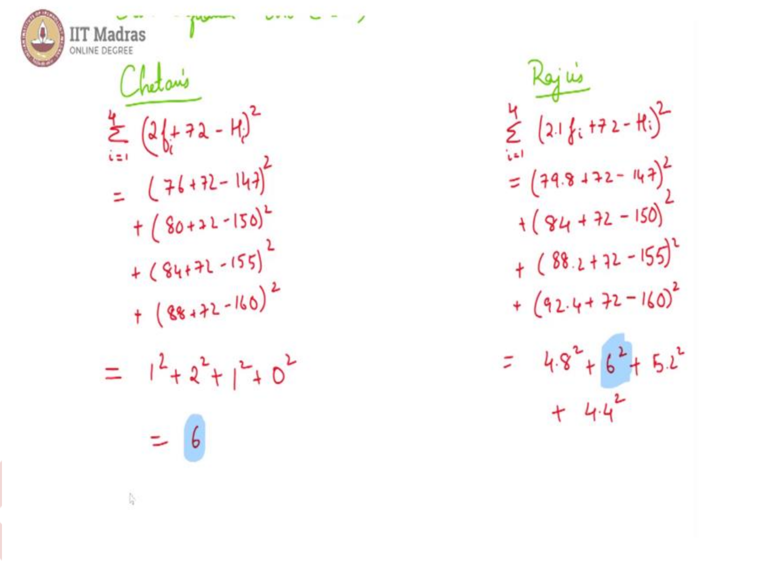
So, that means the distance in kilometre RP in km is equal to $2\sqrt{2} \times \sqrt{2}$, that is 4 km and Ravi has taken 2 minutes. If we write it in hours, t in hours is then $\frac{2}{60}$ that is $\frac{1}{30}$hours. So, the speed would be distance by time, which is $\frac{4}{\frac{1}{30}}$ which gives us 120 kmph, kilometre per hour. This is the speed that Ravi travels at.

Mathematics for Data Science 1
Week 03 – Tutorial 07

(Refer Slide Time: 0:14)







In our seventh question, we have this interesting thing, where there are two anthropology students and they are calculating the relationship between the length f of the femur and the height H of a female adult using fossilised bones. So, what is exactly happening here? What is femur?

From Wikipedia, we can see that the femur is the thigh bone, which is this particular bone. So, what is happening is, in our question, there are fossilised bones and these anthropology students, anthropologists try to study the nature of humans and their societies as they were evolving.

So, here we have fossilised bones and suppose we have the femur of what we know to be a female adult, then we are estimating the height of that female adult from the length of the femur bone, from the thigh bone. So, it is given that this relationship is linear, we have $H = mf + n$. Both use the data given below, so this is the data that is available. We have the femur length and the height of the adult female.

So, from this we are trying to develop this model and Chetan has found $m = 2, n = 72$, whereas Raju has calculated m to be 2.1 and n to be 72. So, both of them agree on n, this parameter is already fixed. It is the m that we are trying to see, whose m is better. So, in terms of linear equation, m is basically the slope of the line. So, how do we do this? We want to use the concept of Sum Squared Error, which we call SSE.

So, in both cases, we are going to look at what is being predicted in terms of height and what is the actual data. So, let s look at case one, let us look at Chetan's case here and Raju's case here. In terms of Chetan's case, we would have the $H = mf + n$, where m is 2, so we have $\sum_{i=0}^{4}(2f_i + 72 - H_i)^2$ and we sum it over how many items 1, 2, 3, 4.

So, let us call this f_i, H_i and i goes from 1 to 4 and in case of Raju's measurements, this error would be again, i goes from 1 to 4 and we have $\sum_{i=1}^{4}(2.1f_i + 72 - H_i)^2$. So, I think we just need to do the calculations now. So, let us look at this here, so this is case 1. So, this is case 2, this is case 3, this is case 4, f_1 is 38 and H_1 is 147.

So, when we put in 38 here we get 2 times 38 is 76 + 72 - 147 the whole square and then in case 2, we have 40 and 150 as f_2 and H_2. So, we will get 80 + 72 - 150 the whole square and then we have, f_3 is 42 and H_3 is 155. We have $(84 + 72 – 155)^2$ and lastly, we have f_4 is 44 and H_4 is 160. So, we have $(88 + 72 – 160)^2$.

$$\sum_{i=0}^{4}(2f_i + 72 - H_i)^2 = (76 + 72 - 147)^2 + (80 + 72 - 147)^2 + (88 + 72 - 160)^2$$

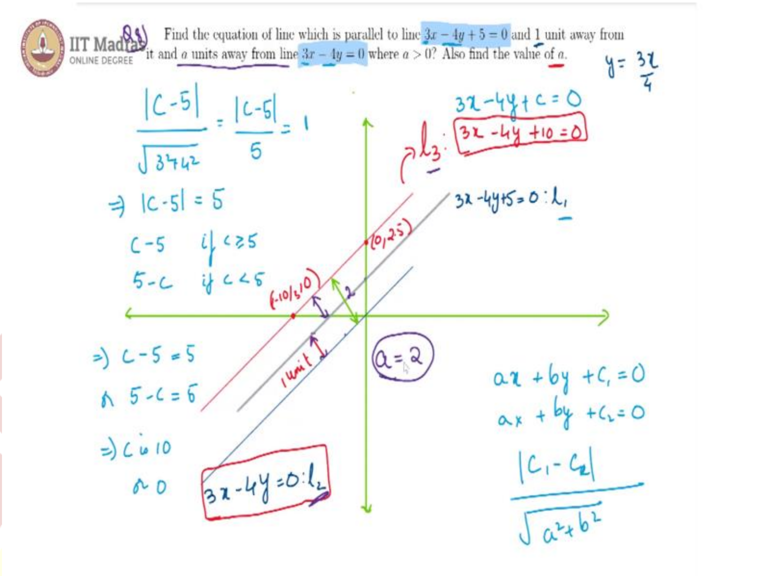
So, this is the total sum squared error for Chetan. Whereas in case of Raju, we would get 2.1 times the same thing. So, 2.1 times 38 is 79.8 + 72 - 147 the whole square and in case 2 we get 84 + 72 - 150 the whole square + in 3 we get 88.2 + 72 - 155 the whole square and lastly, in case 4, we have 92.4 + 72 - 160 the whole square. We calculate these values then we get $1^2 + 2^2 + 1^2 + 0^2$.

$$\sum_{i=1}^{4}(2.1f_i + 72 - H_i)^2 = (79.8 + 72 - 147)^2 + (84 + 72 - 150)^2 + (92.4 + 72 - 160)^2$$

So, this is 6. So, sum square error for Chetan is 6. Whereas in Raju's case we would have $4.8^2 + 6^2 + 5.2^2 + 4.4^2$. Now clearly, in this error, there is a 6^2 which has to be greater than 6, which means Raju's error is much more than Chetan's error. Therefore, Chetan's line fit is better.

Mathematics for Data Science 1
Week 03 – Tutorial 08

(Refer Slide Time: 00:14)



In our eighth question, we want the equation of a line which is parallel to this given line. So, let us first plot the line that is given, lets plot this line. If you look at that line, it is $3x - 4y = -5$, which gives us $\frac{}{-5} + \frac{}{5} = 1$. So, the x intercept and y intercepts can be marked out as 1.6s. So, if we take this to be 1 and this to be 2, so this is - 1 roughly and this is - 2, roughly. Yeah, this might be our intercept, which is (—, 0) and our y intercept, again, if we take this to be 1 and this to be 2, 1.25 is somewhere likely here.

So, this is probably our y intercept, (0,–). As you can see, we are doing a thoroughly rough plotting, we do not always have to be very accurate with our plotting. This is only for an indication. So, this would be our line, $3x - 4y + 5 = 0$ and now we have another line given to us, which is $3x - 4y = 0$.

Clearly, these two lines are parallel to each other because they have the same slope and that slope would be, we write it as y is equal to, we will get 3x / 4, so the slope is – and it is passing through the origin, because if I put x = 0 and y = 0, the line equation is satisfied, that is there is no constant term. So, this line is our and we are trying to find a line that is parallel to these 2 and it is at a distance of 1 unit from the line.

Let us name these lines as well. Let us call this and this is l_2. I am going to erase the intercepts to make it look a little clear and now, we can find our equation and for that, we will use the formula of separation between 2 parallel lines. So, that two parallel lines and we write them with the same coefficients.

So, $ax + by + c_1 = 0$ and the other one would be $ax + by + c_2 = 0$. This is how two parallel lines would look like. You can reduce them to have the same coefficients for x and y, like in this case so this is 3 and this is - 4, this is also 3 and this is also - 4. In this case, the separation between these two parallel lines would be $c_1 - c_2$ the modulus divided by $(\sqrt{a^2 + b^2})$. So, the equation we are looking for, the line we are looking for also is going to be some 3x - 4y + c = 0.

So, its separation from our l1 is going to be applying the formula modulus of $\frac{}{\sqrt{}}$.

is 25. Therefore, you have modulus of —, which is what is expected to be 1 unit. That gives us modulus of c - 5 = 5.

Now the model is, indicates that there are two possible values here, one could be c - 5 if ≥ 5,. Because then c - 5 would be positive, and the other would be 5 - c if c < 5. So, what we get is two separate solutions, one is c - 5 = 5 or 5 - c = 5, in which case we get c is 10 or 0. So, we have 2 lines, one is $3x - 4y + 10 = 0$, the other is our $3x - 4y = 0$, this is our line.

So now, because of this, we can say that this length between, the separation between these two lines is now one unit and therefore, the other line, which is our 3x - 4y + 10 = 0 is going to come on the other side of l_1, which is going to look like this. So, this line is our l_3 and it will have intercepts equal to, this should be 0,— which is 2.5 and this would be (- —,0), this is our other plan.

And now, we should also find out what the value of a is, because a would be the distance between the lines l_2 and l_3 that is what they are saying, its units away from our l_2 and we know that this is one unit and this is also one unit. So, this total length is going to be 2 units. So, a = 2.

Mathematics for Data Science 1

Professor Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras

Lecture 4.1 A

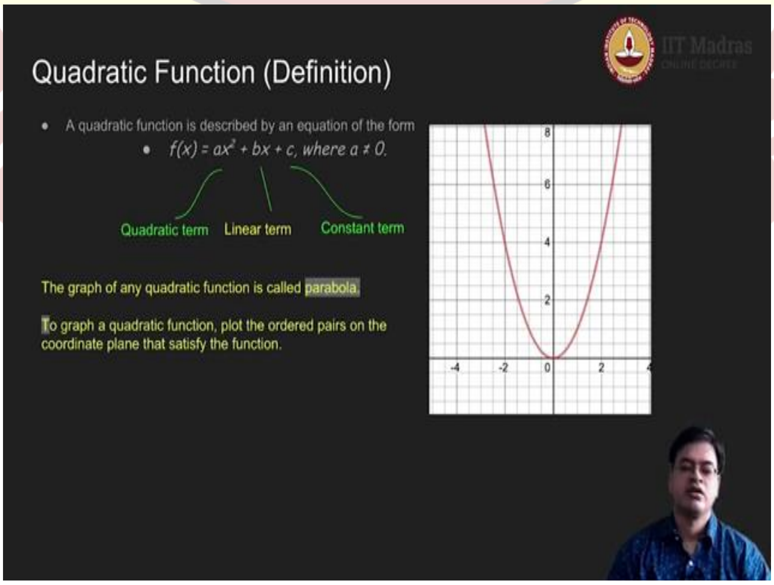
Quadratic functions

Welcome students, today we are going to start the new topic in our syllabus that is Quadratic functions. Before starting this new topic, let us revise what we have studied so far. We started with some simple geometric objects like points and lines, after studying points and lines geometrically we plotted them on coordinate plane and seen how to derive the algebraic equation of a geometric line.

When we have seen the algebraic equation of a geometric line, we got a form of the form $y = mx + c$, it is also known as linear function that we have seen in last few lectures. And if you recall recollect it from the first week where you have studied functions, this is $f(x) = mx + c$is a linear function.

Now, we want to enhance our knowledge further and add 1 more intrication or 1 more complexity in this particular function and that is why we are studying quadratic function. Here we will take an approach where we will first state the algebraic form and then derive its geometric properties as opposed to what we did in straight lines. So, let us start with quadratic functions.

(Refer Slide Time: 01:38)



The first question is, how will I define quadratic functions? The answer to this question is given in this slide. So, a quadratic function is described by an equation of the form $f(x) = ax^2 + bx + c$, where $a \neq 0$ is a crucial condition. Why? If $a = 0$ it simply reduces to a linear function. Let us talk about the name quadratic function. The name quadratic function is derived from 1 foreign language where the quadra term, actual word quadratic term means square and quadratic means related to square.

So, a quadratic function is a function that is related to square of the variable as can be seen from the definition, it has a term containing a x^2. So, if $a = 0$ then it does not have a term containing square so it no longer remains a quadratic function and it is a linear function which is equivalent to a straight line as a geometric object.

So, we will put a condition that $a \neq 0$ that means we are studying a quadratic equation. The next question is how to plot a graph of this function. So, this equation is actually composed of 3 terms, let us describe them 1 by 1 that is a x^2, this term is a quadratic term.

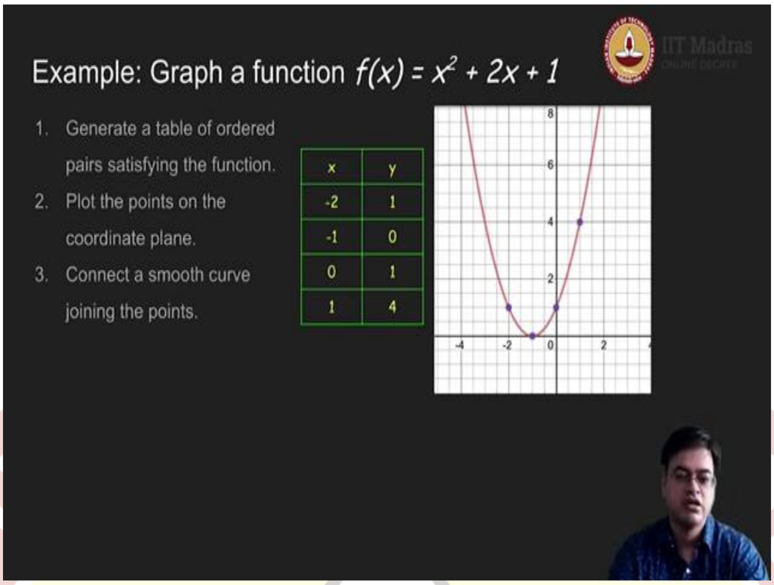
As I mentioned earlier, when $a = 0$, the term $bx + c$ survives and that term bx is a linear term. And finally, if you put $x = 0$, only term that survives is c so that is nothing but a constant term. So, a quadratic equation can be split into 3 parts. If ax^2 is not there, then I know how to handle this term on a coordinate plane, it just simply represents an equation of a non-vertical line.

So, I know how to handle these terms. So, what if the a x^2term remains that is $a \neq 0$? We can graph this particular function and graph of any quadratic function will be called as parabola. Graph of any quadratic function will be known as parabola. So, what are the important features of parabola? In order to do that we first need to plot the parabola.

So, what is the best way to do it? We have already seen to graph any function what we need to do is, we need to take the value of x, put it in the formula $f(x)$ and evaluate it and get the values of y. So, consider all ordered pairs and plot them on the coordinate plane so that they satisfy this function, is the best way to handle it. For example, let us take this let us take for example, when $b = 0$ and $c = 0$ and $a = 1$, let us take that particular function that is $y = x^2$.

In that case what I will do is, I will put $x = 0$, I will get back 0. So, I will plot a point (0,0), I will take $x = 1$ I will get 1, $x = -1$ I will get 1, and then $x = 2$ then I will get 4 and if I take $x = -2$ again I will get 4. So, $y = x^2$can be easily plotted by joining these points smoothly this is the curve $y = x^2$, this is how we plot our quadratic function.

(Refer Slide Time: 5:35)



Let us take 1 example. Let us say I want to graph a function $f(x) = x^2 + 2x + 1$. How will I graph this function? In 3 steps. First, I will generate a table of ordered pairs satisfying the given function. Second, I will plot those points on the coordinate plane. Once I plot those points on the coordinate plane, I will connect a smooth curve joining the 2 points, this is the recipe for drawing a function. Let us draw it here, for that I have computed some points you can verify by yourself, if you put the value of $x = -2$, you will get $(-2)^2 + 2(-2) + 1$ and on solving you will get 1.

You take the value $x = -1$ you will get 1, for x square you will get -2 and 1 in the constant term. So, together they will cancel and you will get 0. Similarly, you can compute for $x = 0$ it is 1 and for $x = 1$ it is 4. Now, our job is to consider a coordinate plane and plot these points so I have plotted these points.

So, these points are plotted and now I need to draw a graph, which is connecting all these points. Now, here you remember I have plotted these 3 points, how will I know the shape of this graph in this zone? That is a major question that you can ask, but this parabola is somewhat symmetric in a sense, suppose I take this point, what is the point here, the point is (1, 4).

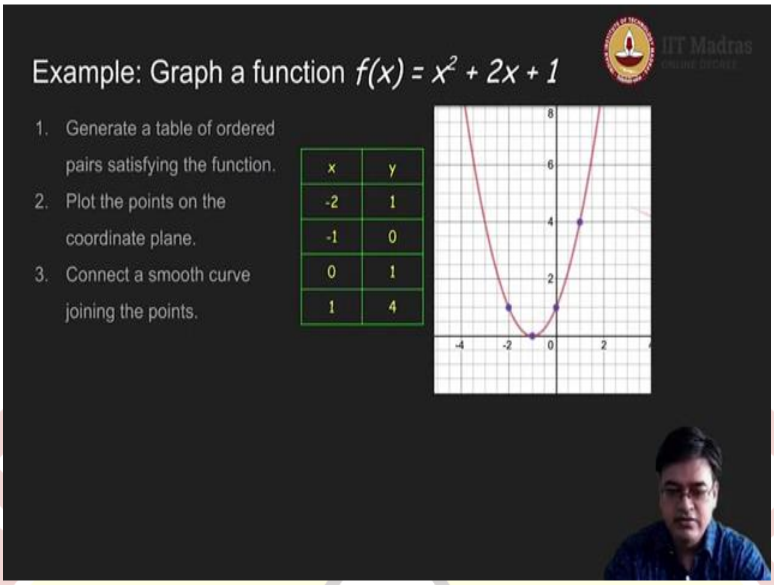
Now, if I consider this point which is -1 where it takes the value 0 and consider the point 1 it is 2 units apart. So, somewhere in this where -3 will come, which is 2 units apart from -1, the value of the parabola will be again 4. I will keep the cursor here, see. So, there is some kind of symmetry underlying this particular function, we need to understand that symmetry in a better way.

So, what essentially is happening is, if I consider this point which is the bottom of the curve, and if I draw a straight line, which is the line $x = -1$, then if you look at all these points for every point there is a similar point on $y - axis$ at the same distance from this particular point. This also can be called as a symmetry of a parabola. We will study this later in the next slide. So, right now, our job is to graph a function which we have plotted and let us explore further properties of this parabola like this symmetry, what is the meaning of the symmetry and all those things.

(Refer Slide Time: 09:00)

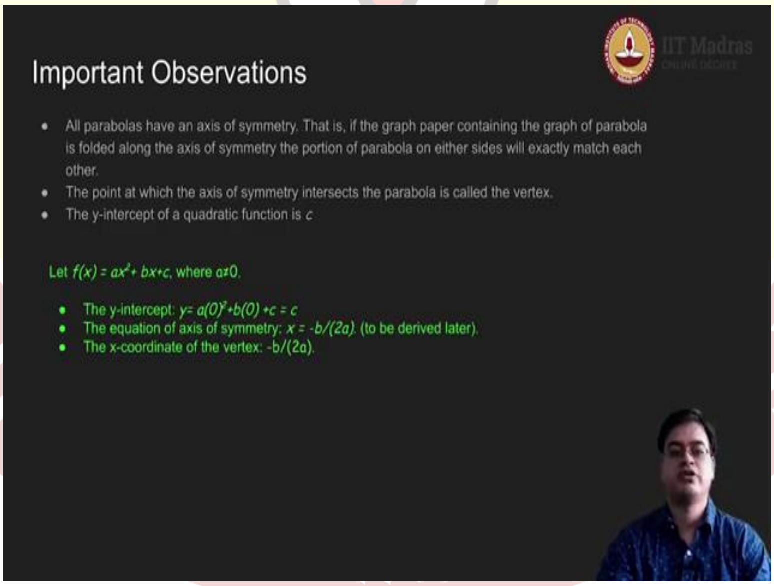
So, there are a few important observations, if you consider equation of $y = ax^2 + bx + c$, these are all parabolas, I have shown you two parabolas $y = x^2$, $y = x^2 + 2x + 1$both parabolas have axis of symmetry. Inevitably all parabolas will have an axis of symmetry that is, what is axis of symmetry. Let us go to the previous slide and see.

(Refer Slide Time: 9:28)



The axis of symmetry over here, as I mentioned was $x = -1$. If I take this graph paper and fold along $x = -1$, then the curves that we have plotted here must exactly match each other that gives us a recipe to draw a parabola.

(Refer Slide Time: 9:50)



So, all parabolas will have axis of symmetries that is if you take a graph paper containing the parabola, and if you fold it along the axis of symmetry, the portions of the parabola on both sides will exactly match with each other, this is the beauty of a parabola. So, now if I know how the parabola appears on one side, I know how the parabola appears on the other side of the axis of symmetry. It is a pure reflection of whatever is happening on one side.

Then, the point this axis of symmetry as we have seen in the previous graph, the point at which this axis of symmetry meets parabola, we will call that point as a vertex of the parabola. This is again a nomenclature we will call that point as a vertex of the parabola and the point at which x, if you put $x = 0$, then the point at which the y coordinate is taken is called the value c or you can simply refer to the equation $ax^2 + bx + c$, put $x = 0$, that will be the y intercept which will be given by c.

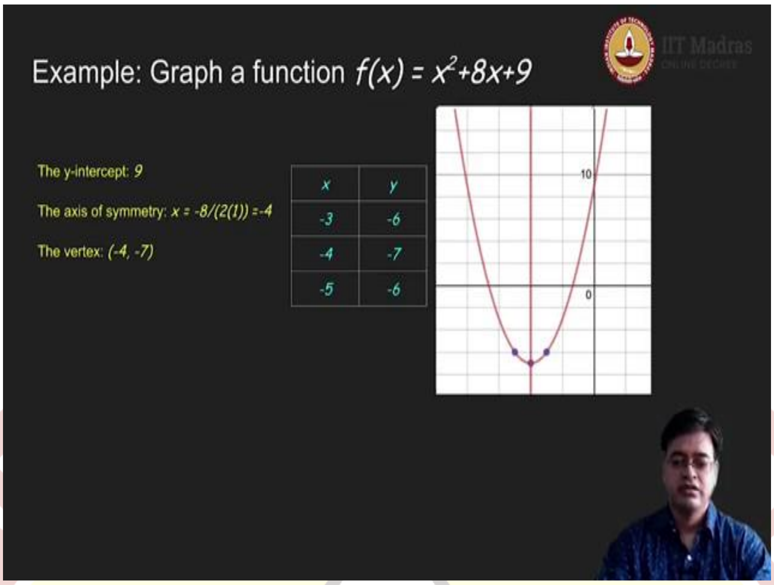
These 3 points play a crucial role in graphing the parabola. How? Let us do it one by one, Let us say, our quadratic function is $ax^2 + bx + c$ where $a \neq 0$, you can easily figure out that the y intercept of this point by putting the value 0 in c.

Now, I want to know the axis of symmetry, this plays a crucial role. So, I will derive the expression for axis of symmetry later but right now you memorize this equation as $x = \frac{-b}{2a}$ as this needs some algebraic skills which we do not have right now. So, I will derive it later. But right now, you understand that $x = \frac{-b}{2a}$.

Remember, the equation of the quadratic function is given by$ax^2 + bx + c$, and c will not play any role in this and b and a will play a role. So, it is $\frac{-b}{2a}$ is the axis of symmetry and where the graph meets this parabola, it is called vertex. So, the x coordinate of the vertex is $\frac{-b}{2a}$ obviously, because the axis of symmetry has $x = \frac{-b}{2a}$. So, the x coordinate of the vertex is $\frac{-b}{2a}$.

Let us see how this knowledge helps us in understanding how to draw a parabola. So, there are 3 steps in drawing the parabola, first you need to generate a table of values, but if you generate a table of values only on one side and you do not have a table of values on vertex, then you may not be able to draw the parabola appropriately, that is way the knowledge of these facts is important, so let us see how to draw a parabola by example.

(Refer Slide Time: 13:07)

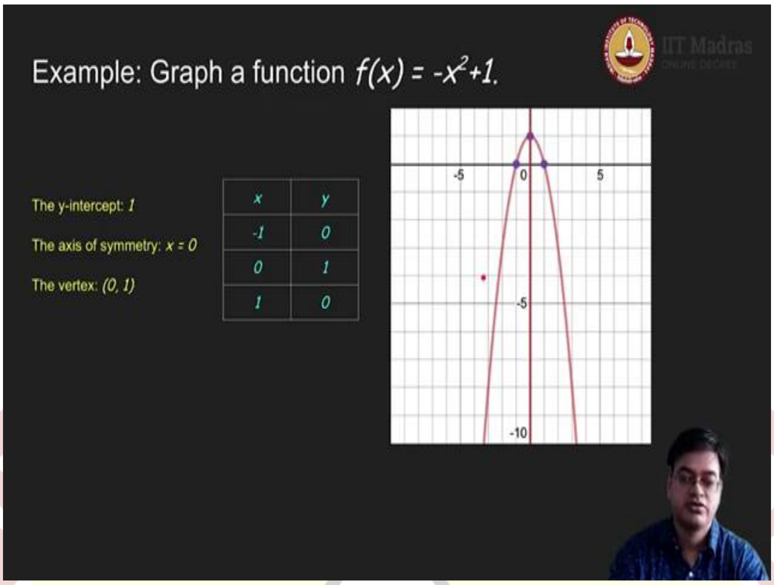


$f(x) = x^2 + 8x + 9$, I want to graph this function. So, I will reiterate on previous points. So, what is the y intercept, y intercept is 9, because if you put $x = 0$, the y intercept is 9. Next, I want to know the equation for axis of symmetry. So in this case, what is b, b is 8, a is 1, so $\frac{-b}{2a}$ is $\frac{-8}{2}$, which will give me axis of symmetry to be $x = -4$, y intercept is 9, axis of symmetry is $x = -4$ so I can evaluate the coordinate that is $(-4, -7)$ will be the vertex. How this -7 comes, you just substitute -4 over here in this expression, you will get the value to be equal to -7.

So, now with these 3 terms, how will I draw the function? So, now I know that around vertex I need to find the points. So, based on this, I will draw a table. So, around -4 I have simply taken three points, fourth point is already with me, (0, 9) is the 4th point. So, around the around the point -4 I have taken the values so -3 which is the value of -6 when you substitute in the function -4, -7 already known and -5, -6. So, I have 3 points and the point (0, 9).

So, I will plot these points on a graph paper, take a graph paper, plot the axis of symmetry because around this the curve should be symmetric, take these 3 points, these 3 points are here and I know (0, 9) is another point. So, it should be somewhere here (0, 9). Now, let us plot a graph. So, now we have plotted a graph with much ease because of the knowledge of axis of symmetry I know where the point where the minimum has occurred or the vertex point is that is the beauty of axis of symmetry. So, this is how you will be able to plot any function any quadratic function given to you, this is about the graphing of a function.

(Refer Slide Time: 15:48)



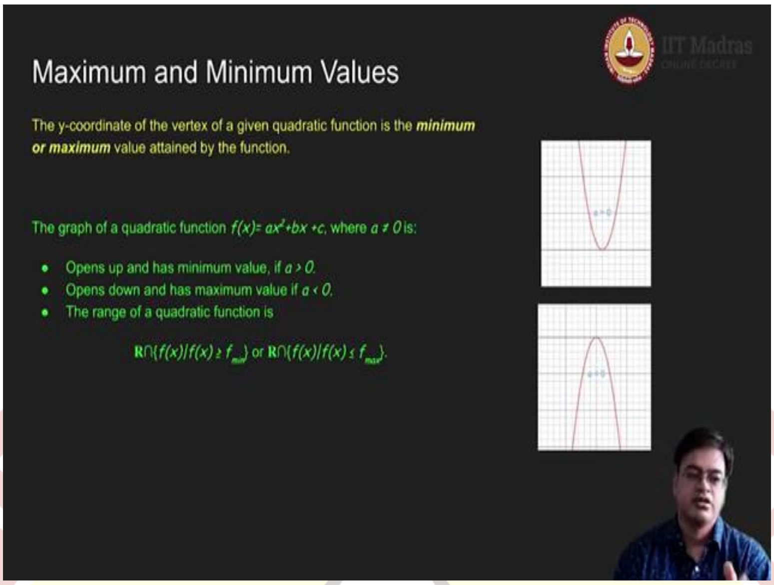
Let us try to see, is this the only shape that is possible that is the upward shape. Let us try to figure out whether is this a quadratic function first of all, $-x^2 + 1$, the answer is yes, $a = -1$, $b = 0$ and $c = 1$. Now, in this case, let us try to figure out the 3 summaries that is what will be the y intercept for this? y intercept will be 1, what will be the axis of symmetry for this because b is 0, it does not matter what is the value of a it will be 0.

So, $x = 0$ is the axis of symmetry that is y axis is the axis of symmetry for this particular function. And the vertex is (0, 1). In this case, we are not really getting much information because what this is saying is (0, 1) is the y intercept that is (0, 1) is the coordinate, axis of symmetry is 0 that means (0, 1) is the vertex as well, right?

But still this information will suffice because I know I have to find the points around 0. So, let Figure out the points around 0; -1, 0 and 1, these are the 3 points, their y coordinates respectively are 0, 1, 0. Now you see there is a change, earlier we were only dealing with positive side of y axis or the y axis where the curve is opening up, here the curve is opening down. For example, if I plot an axis of symmetry over here, which is y axis and if I plot these 3 points, these 3 points look like this that means the curve will go downward. So, the curve is opening down, why this has happened.

In earlier examples, if you look at it closely then this the form, general form of this expression $ax^2 + bx + c$, in all of them a was equal to 1, and in this particular expression $a = -1$ therefore, the curve is actually opening down instead of opening up, this point needs to be noted.

(Refer Slide Time: 18:15)



So, let us know this point and figure out what happens when this a is greater than 0 or is less than 0. So, that leads us to the next question that is maximum and minimum values. So, the y coordinate of the vertex of a given quadratic function is minimum or maximum value attained by the function.

Do you all agree with this, we have seen 3 to 4 graphs of the functions, first we have seen $y = x^2$ where it goes to bottom and 0 is the minimum value, then we have seen $y = x^2 + 2x + 1$which again gave us 0 value, then finally we have seen Third graph that we have seen the last graph that is$-x^2 + 1$, because the value of a was negative, it was going downward, and that will give me the maximum value and all the values are below that value.

So, the y coordinate of the vertex of a given quadratic function gives us the minimum or maximum value attained by the quadratic function. In particular, given any graph given any function $f(x) = ax^2 + bx + c$ where $a \neq 0$. The graph of this quadratic function if a is greater than 0 will open upwards and will have a minimum value.

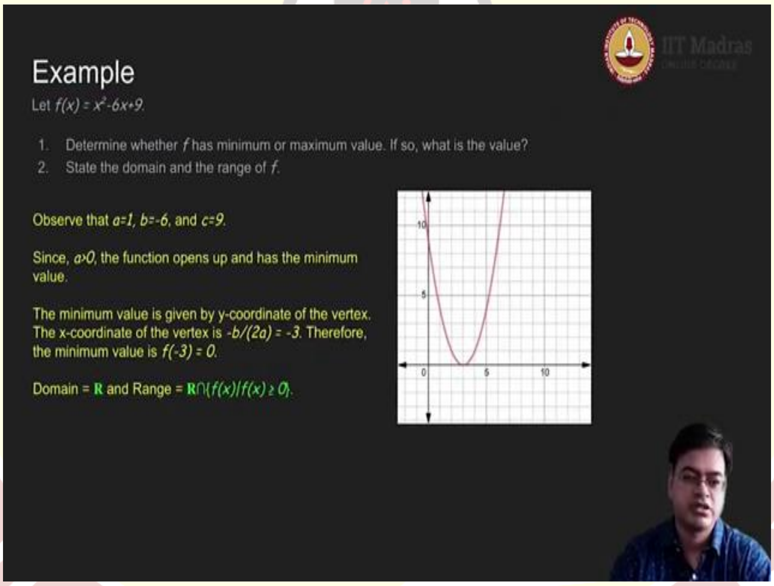
If a is less than 0, the graph will open downwards and will have the maximum value and there will be either maximum or minimum values, not both, this is the beauty of the quadratic function. Another thing that you can see is the range of the quadratic function, if you relate to your weak 1 background, where you are discussing about the domain-codomain range, so the range of this quadratic function will be.

So, let us say a is greater than 0, then it attains the minimum value then it will be the minimum value and all of the real line that is above the minimum value. And if a is less than

0, then it will be the maximum value and an entire real line which is below that particular thing, I can denote this using the set theoretic notation as it is a set of real numbers R intersected with a set of all $f(x)$, these y values such that $f(x)$ is greater than or equal to f min when $a > 0$, or if $a < 0$ it is set of real numbers intersected with $f(x)$ such that $f(x)$ is less than or equal to the maximum value that f has achieved.

So, let us try to visualize this. For example, if $a > 0$ your graph looks like this. So, in particular the range of the value, range of y values is from this point to upward. So, this is the entire real line above this value. Similarly, if $a < 0$, the range of the values that is taken by this function is this, if you relate this to domain codomain terminology, what is the domain of this quadratic function, it is an entire real name and range is restricted to some subset of real life.

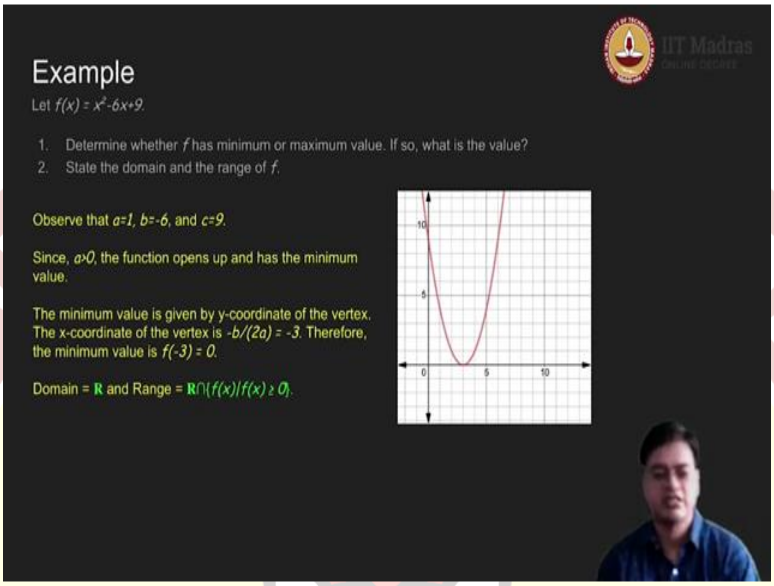
(Refer Slide Time: 21:51)



So, we will try to improve upon this concept using this example.

Mathematics for Data Science 1
Professor Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras
Lecture 4.1 B
Examples of Quadratic Functions

(Refer Slide Time: 00:16)



Let us say, this example we have been given a function $f(x) = x^2 - 6x + 9$ and we are asked to determine whether f has minimum or maximum values, if so what is the value and you need to state the domain and range of f. Let us first attempt the second question, what is a domain? Domain of f is enter real line we do not have to worry, what is the range of f?

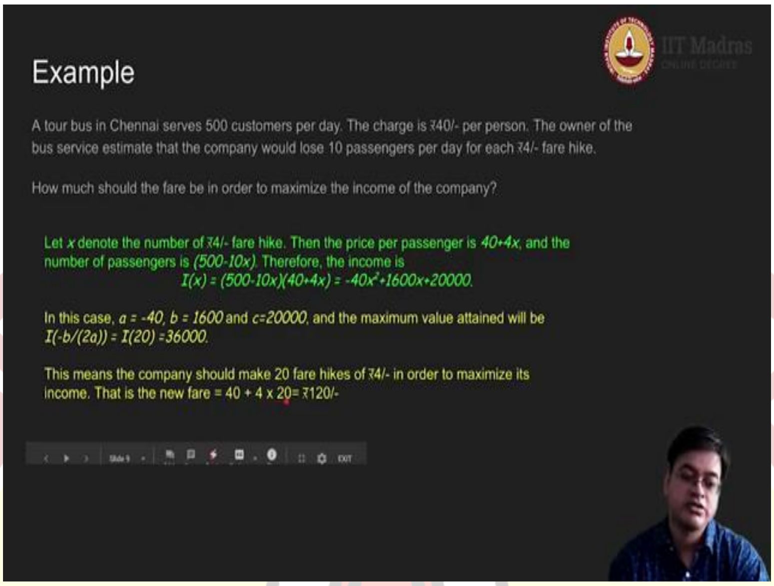
Let us take this function identify a, b, c so $a = 1, b = -6, c = 9$. Since $a > 0$, the function opens up, if the function opens up then it will have a minimum value. So, the answer to first question is whether f has minimum or maximum value, it has a minimum value, once it has a minimum value it cannot have maximum value, if so what is the value?

You need to figure out what is the vertex of this particular parabola. So, what is the formula for vertex of the parabola, —, $b = -6$, so — is —— which will give me minus 3. Sorry, this is wrong, it should give me $+3$, ——, it should give me $+3$, which is written wrong here, but the graph is correct here where we are getting is the vertex. So, if you substitute 3), what do you get? and therefore, the value of this is nothing but 0. So, this is wrong it, should be 3.

And obviously, the range if it has a minimum value, the range is minimum upwards, this is the entire real line above this minimum. So, that is R intersected with $f(x)$ such that $f($

0. So, we have understood how to find the minimum and maximum values of a function, if a is negative you can similarly find the maximum value.

(Refer Slide Time: 02:58)



So, let us try to make this example more realistic. So, let us take 1 realistic example. Where a tour bus in Chennai serves 500 customers per day, they charge rupees 40 per person. Now, they want to revamp their strategies, so the owner of the bus service estimate that the company would lose 10 passengers per day for each Rs 4 hike in the fare.

So, if they hike the fare by 4 rupees, then they will lose 10 customers per day, this is the estimate. Now, the company wants to maximize the profit, so how much should be the fair in order to maximize the income of the company is the question. So, let us try to answer this question using our knowledge of quadratic equation.

So, let us say, 1 unit of hike is 4 rupees so let x denote the number of Rs 4 fare hikes. So, what will this impact? This will impact the number of passengers because we are losing 10 passengers per fare hike. So, what will be the corresponding fair price for the passenger? It will be $40 + 4x$, 40 rupees is the fees that we are charging per person, the company charging per person and if I hike the fare it will be four times x, this will be charged per person.

Now, the number of passengers with this hike if you increase x units, that means, you will lose 10 passengers every x units increase. That means $500 - 10x$ is the passengers that still remain. So, in this case, the income of the company will be the number of passengers into the fare, they have charged so that is $(500 - 10x)(40 + 4x)$. If you open this, open the bracket and multiply them, then you will get the expression to be$-40x^2 + 1600x + 20000$. This is the income.

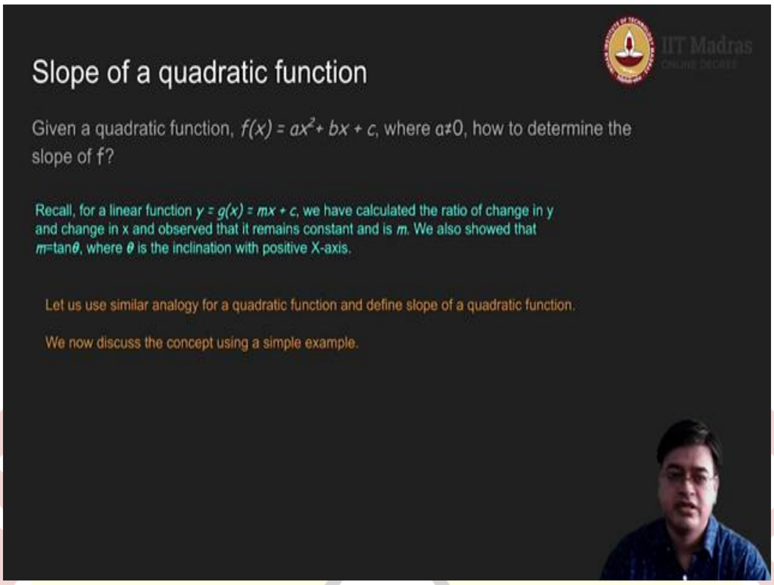
Now, the company wants to maximize the profit, first of all after getting this quadratic equation, can you tell me is the maximum possible? The answer is yes, and why the answer is yes, because it lies in the coefficients $a, b$and c. So, what is $a$here? $a = -40$, $b = 1600$ and $c = 20{,}000$. Because $a = -40$, a < 0 so, the curve will open downwards that means the maximum is possible.

And what will be the maximum value attained then that is what we have to figure out. So now, the next question is okay. So, where this maximum will be attained? The maximum is possible, maximum will be attained on the vertex, y coordinate of the vertex will give me the maximum. So, I will simply figure out what is the x coordinate of the vertex, x coordinate of the vertex is point of intersection of the axis of symmetry, what is the axis of symmetry $x = \frac{-b}{2a}$, what is b? 1600, c is 20,000 and a is -40.

So, $x = \frac{-1600}{2(-40)}$ which will give me 20, so that is what 20 is yes. And of maximize y coordinate the maximum fair that we will get is 36,000. Right now, how much we are earning, how much the company is earning, it is 500 customers they are serving, where everybody is paying 40 rupees so they are simply earning only 20,000 rupees that is when you do not increase any fare $x = 0$, you get 20,000. So, the main question is how much the fare should be?

Now, what we are suggesting here by solving this problem is there should be a 20 units of hike of rupees 4 each that means, what we are suggesting is there should be 80 rupees hike in the fare. So, the new fare for the company should be 40 plus four times x that x is 20. So, $40 + 4 \times 20 = 120$and that is what is the recommended hike by the company. So, now every person should be charged 120 rupees as opposed to 40 rupees and then the company will be profitable and you may have to serve less customers. This is how we are using real life, we are using quadratic equations to solve real life situations.

(Refer Slide Time: 8:37)

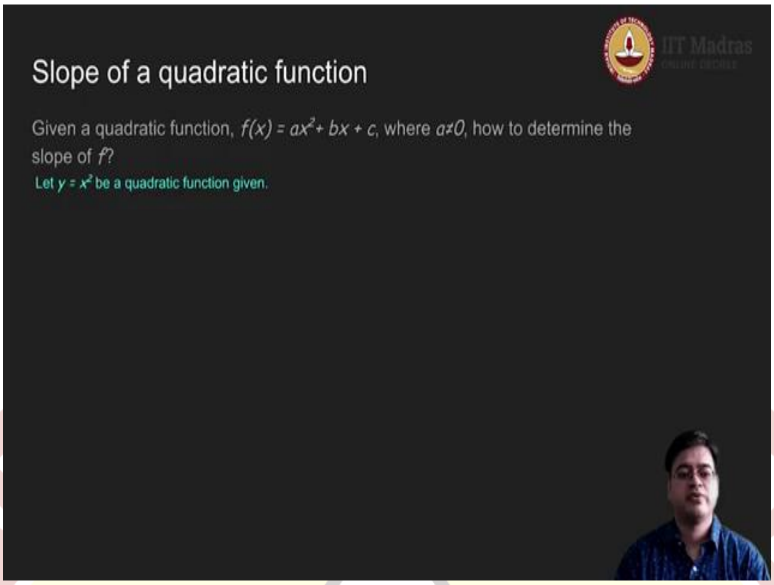


Now, let us go back to our linear functions, where we studied the slopes of the lines. What was the slope of a line? Slope of a line was change in y by change in x. Let us see what the concept of slope has to do with a quadratic function. Let us try to analyse that. So, my goal in this set of slides is given a quadratic function $f(x) = ax^2 + bx + c$ where $a \neq 0$, how to determine the slope of a function f.

So, in order to generalize this notion of slope of a function, we will first recall what we do know about linear function. So, if you look at a linear function which is y which is equal to $g(x) = mx + c$, we know that this m represents the slope and m can be calculated by considering a ratio of change in x upon change in y.

We have spent a lot of time in understanding the slope and when I consider a linear function, I also know that the slope remains constant okay. We also know that the slope is nothing but tan of some inclination and that inclination is with positive x axis. I want to relate all these concepts and try to figure out what is the slope of this quadratic function. Let us go ahead, we will use a similar analogy for a quadratic function and define the slope of a quadratic function. First let us take one example to discuss this concept of slope.

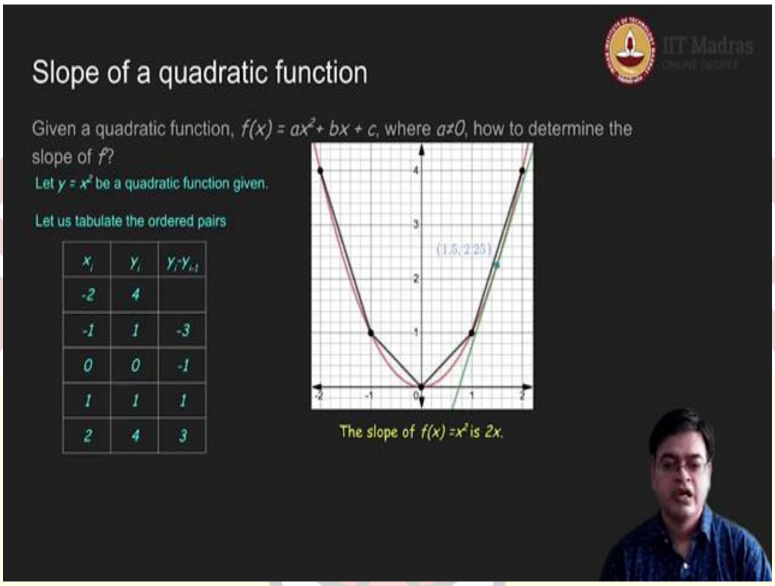
(Refer Slide Time 10:25)



Let us take our standard prototype example. We are trying to answer this question, $y = x^2$.

Mathematics for Data Science 1
Professor Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras
Lecture 4.1 C
Slope of a Quadratic Function

(Refer Slide Time: 00:15)



So, in this situation, I want to know what is the slope of this curve and is it a constant or which variable or what else? So, we want to answer this question. So, first, we need to plot this function, for plotting the function, we know what is the axis of symmetry for this b is 0. So, so $y - axis$ is the axis of symmetry and it will be symmetric about $y - axis$. Minimum value will be 0 as it can be seen.

So, I will take a symmetry about $y = 0$ that is, I have taken -2, -1, 0, 1, 2, these are the points then I have evaluated the ordered pairs, that is 4, 1, 0, 1, 4. The symmetry is clearly visible in these. Now, what is the definition of slope? It is change in y upon change in x.

So, if you look at the left-hand side, the first column, the change in x is constant. It is 1 all the time so I will use this notation, and I will go ahead and figure out what is the difference between y_i values because the denominator is always 1, it suffices to take the difference between y_i values.

So, the first value is -3, 1 -4 is -3, $y_i - y_{i-1}$. $0 - 1 = -1$, $1 - 0 = 1$,$4 - 1 = 3$, so I got the changes in y with respect to 1 unit change in x so this is the slope, but where does this slope lie or at what point is this slope? Because if it is a straight line, I know the slope is constant. So, in order to understand this let us go to a figure and try to understand.

This is a curve, $y = x^2$. Now, when I consider these 2 points $y_i - y_{i-1}$, what I am actually doing is, I am assuming a straight line connecting these 2 points and I am calculating slope for it. So, I have assumed all these straight lines and I have calculated the slope for it. Is this a slope for a curve? No, basically not because it is a slope for that straight line.

So, now how will I identify this slope? So, if at all I want to decide what is the slope of the line if you look at our old definition the change in y by change in x also associated with $tan\ \theta$, the $tan\ \theta$ plays a crucial role, what is θ? θ is the angle of inclination. So, if I consider any point over here, and if I draw the inclination of, if I draw a line passing through that point and if I measure the inclination of that point with this positive $x - axis$ then I will get a slope because the definition of $tan\ \theta$ was not dependent on the line per se, it was dependent on that line on that particular inclination.

So, tan of that is still a slope of a line. So, let us try to use this idea and see what we can get. So now, I have identified 1 point let us say this point is actually (1.5, 2.25) because I am considering a curve, which is $y = x^2$. What will be the slope of a line at this point? We can ask this question but if you look at this line, this vertical line, this vertical line and if I slide this vertical line slightly for this point, then this is nothing but a tangent to this curve, it passes through it only once.

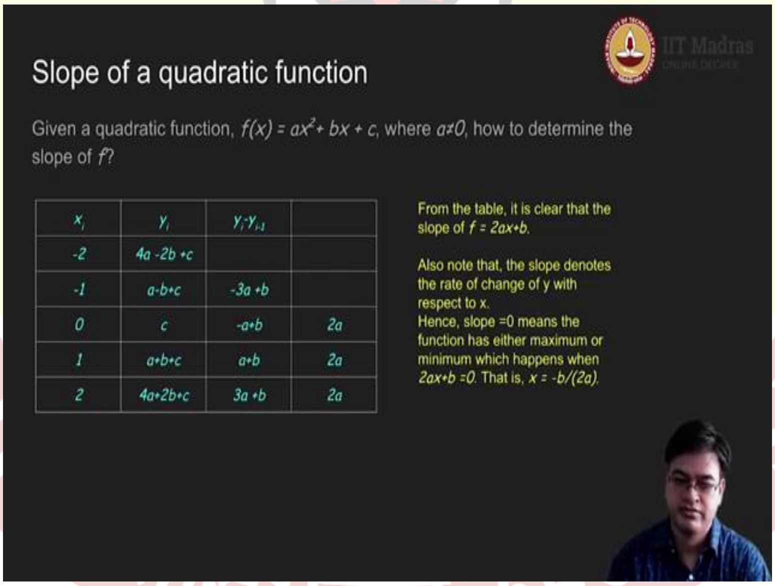
Let us try to actually plot that line. Yes so, once we have plotted this line, this is the tangent to that curve and the line is actually parallel to this line and the point is 1.5. This gives me a hint that this is something like you have -3, the point is, you have a slope between these 2 points as 3, you have a point which is 1.5 and if I divide this point, this particular difference by that point I am getting 2. Then let us look at these differences, what are these differences, the difference is $-1 - (-3) = 2, 1 - (-1) = 2$ so, all these differences are 2.

If you look at the second differences of these points, there are 2 that means there is some relation, 3 and 1.5, 1.5 times 2 is 3. So, I can safely assume that this point 1.5 is actually a midpoint of 1 and 2 on the $x - axis$ and therefore, whatever value is given to it is actually the value of the slope of a curve. And in particular, if I go here, for example, if I go here, and if I talk about the point 0 and 1, then what I will get is a point 0.5, the midpoint of this. Again, I can do a similar exercise, I can draw a line and the line again will be parallel to this line and at a point 0.5, I will get the line with a slope 1.

In a similar manner if I go here, I will get a line with a slope -1, in a similar manner here, I will get a line with a slope -3 and therefore, I can safely conclude that the slope of this particular curve is $2x$. How? I have computed it. So, let us now verify our hypothesis. So, let us take a point 0, consider any 2 points about 0, let us take symmetric points because I need a symmetry.

So, let us take the point (1, -1), what is the slope of this line? It is horizontal line, so the slope should be 0 and that is what this slope is. So, in particular, I can verify for all points if I consider a point, let us say a, a is used here. Let us say if I consider a point z then I will go z +u, z$-$u, I will consider those 2 things and I will assume their values, draw a straight line joining them and whatever is the value of the slope for that straight line will be the value of slope for my point. This is a beautiful idea that can be generalized to a general quadratic curve.

(Refer Slide Time: 8:33)



So, let us answer the general question that is, I want to find a slope of a quadratic function $ax^2 +$ $ax^2 + bx + c$ where $a \neq 0$. So, we will simply take 5 set of points, standard 5 set of points, -2, -1, 0, 1, 2, I will just substitute these values in the function. So, I will get a corresponding values of y_i's, which are here, $4a - 2b + c$, and $a + b + c$, c, $4a - 2b + c$. I will take the first differences of these two, those are given here and then I will take one more difference of these two, all these differences will turn out to be $2a$.

Now, if I look at the points which are here, and if I consider the midpoint of this midpoint of these 2 that is 1.5 so $-2a \times 1.5 + b$ will give me the answer to my question that what is the slope of that particular value, because if you look at this -3, -3 is actually 2 times 1.5. This -1 is actually 2 times -0.5, 1 is 2 times 0.5, 3 is again 2 times 1.5 so I am essentially getting the slope of all these values that means, my answer to the question that the slope of this curve quadratic function is $2ax + b$.

Now, from the table it is very clear the 2 way comes here, $ax + b$ I have derived it because this is a value containing 1.5 in the middle, so this is 2 times 1.5 So, that is ax so $2ax$ that is what this is $2ax + b$. Now, we can do some interesting observations, we have already seen around point 0 for $y = x^2$, the slope was flat it was 0. So, when will that happen? Right.

So, you can equate this $2ax + b = 0$, slope 0 means the function has reached its minimum or maximum, slope is 0. So, when will that happen? That is $x = \frac{-b}{2a}$. This is one of the reasons why $x = \frac{-b}{2a}$ is the value of the minimum or maximum, because the slope reaches the value 0. So, here what actually slope, calculates?

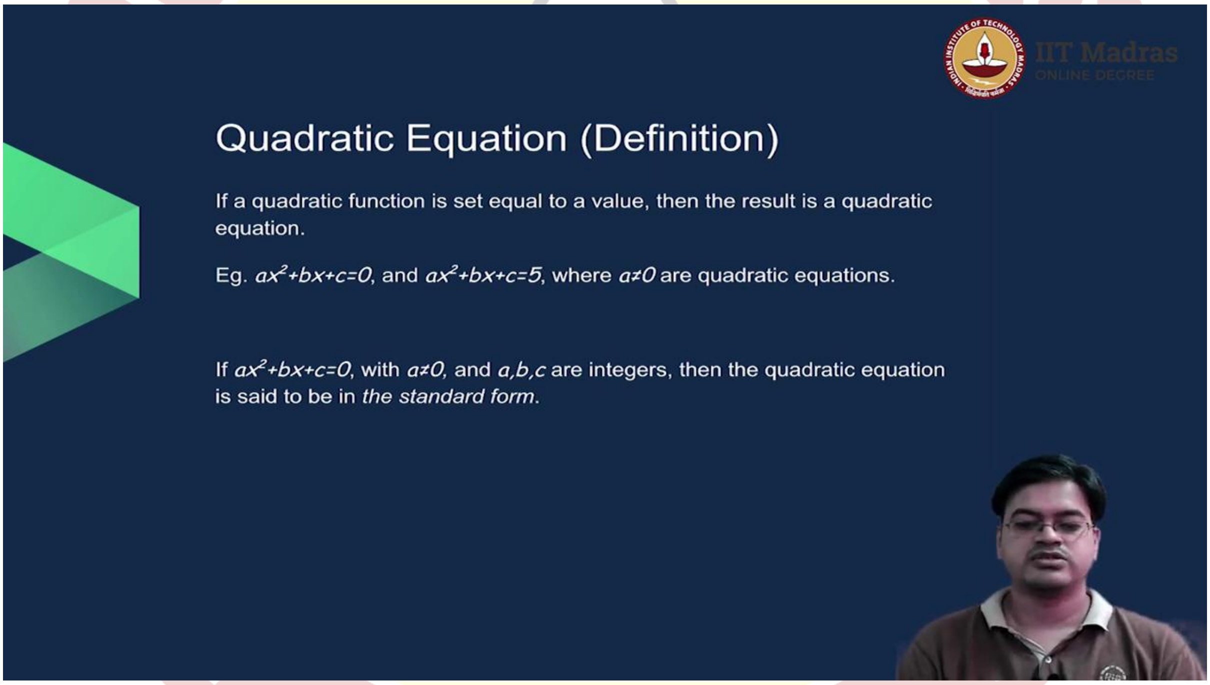
Slope actually calculate the rate of change with respect to x and a rate of change of y with respect to change in x. So, if the rate of change is becoming 0, that means the function has reached its minimum or maximum. So, this justifies the idea that why a quadratic function should have a minimum or maximum value at the point $x = \frac{-b}{2a}$.

Still, that point is pending where we want to find why the axis of symmetry is $x = \frac{-b}{2a}$ and we will come to it later. But as you can see here, the slope of a quadratic function is significantly different from slope of a line, slope of a line is constant, whereas the slope of a function quadratic function f is no longer a constant. In fact, it is variable that is $2ax + b$. It depends on a and b, not on the constant c, which is expected.

Mathematics for Data Science 1
Professor Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras
Lecture 4.4
Solution of quadratic equation using graph

In today's video, we are going to learn what are Quadratic Equations. And once we set up the Quadratic Equation, we are going to see, what are the solutions of the Quadratic Equations, that are called roots of the Quadratic Equation and how to solve these Quadratic Equations, using the technique that we have demonstrated, for quadratic functions. That is Graphing technique. So, let us start.

(Refer Slide Time: 00:42)



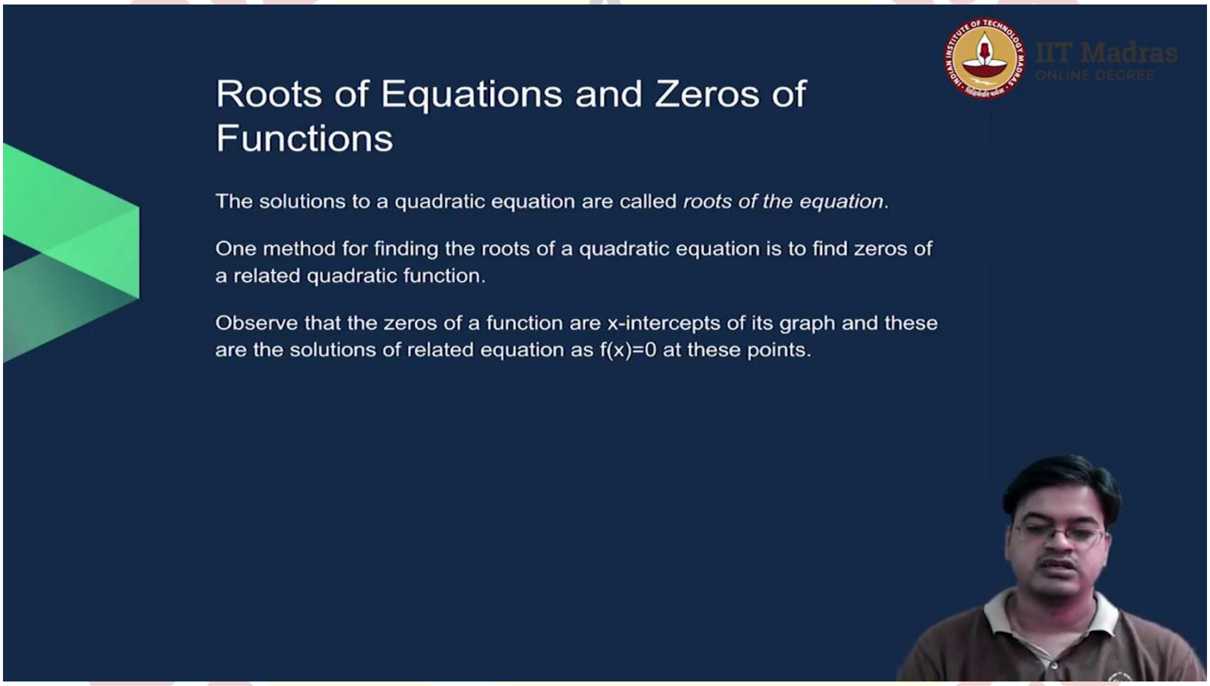
So, first of all, let us understand what is a quadratic equation and how it is related to quadratic function. So, here is a definition. If a quadratic function is set to be equal to a value, then the result is called quadratic equation. So, let us see one example. For example, $ax^2 + bx + c = 0$, is one quadratic equation, where $a \neq 0$.

In the similar manner, $ax^2 + bx + c = 5$, is another quadratic equation. Obviously a should not be equal to 0. Now, once we get the Quadratic Equation, if the coefficients, what are the coefficients. Coefficients are like a, b and c. These are called coefficients of the Quadratic Equation.

If the coefficients are from set of integers, which we have studied in week 1. So, if a, b, c, the coefficients are integers, and on the righthand side, it is equated to 0. That is, you have an equation, $ax^2 + bx + c = 0$, where a is not equal to 0 and a, b, c are integers. Then the Quadratic Equation is said to be in the standard form. So, on this slide, we have seen two definitions. One, what is Quadratic Equation. Quadratic Equation is nothing, but a quadratic function, where it is equated to some value.

And what is a standard form of Quadratic Equation? That is $ax^2 + bx + c = 0$, where $a, b, c \in \mathbb{Z}$ and a is not equal to 0. Then the Quadratic Equation is said to be in a standard form.

(Refer Slide Time: 02:37)



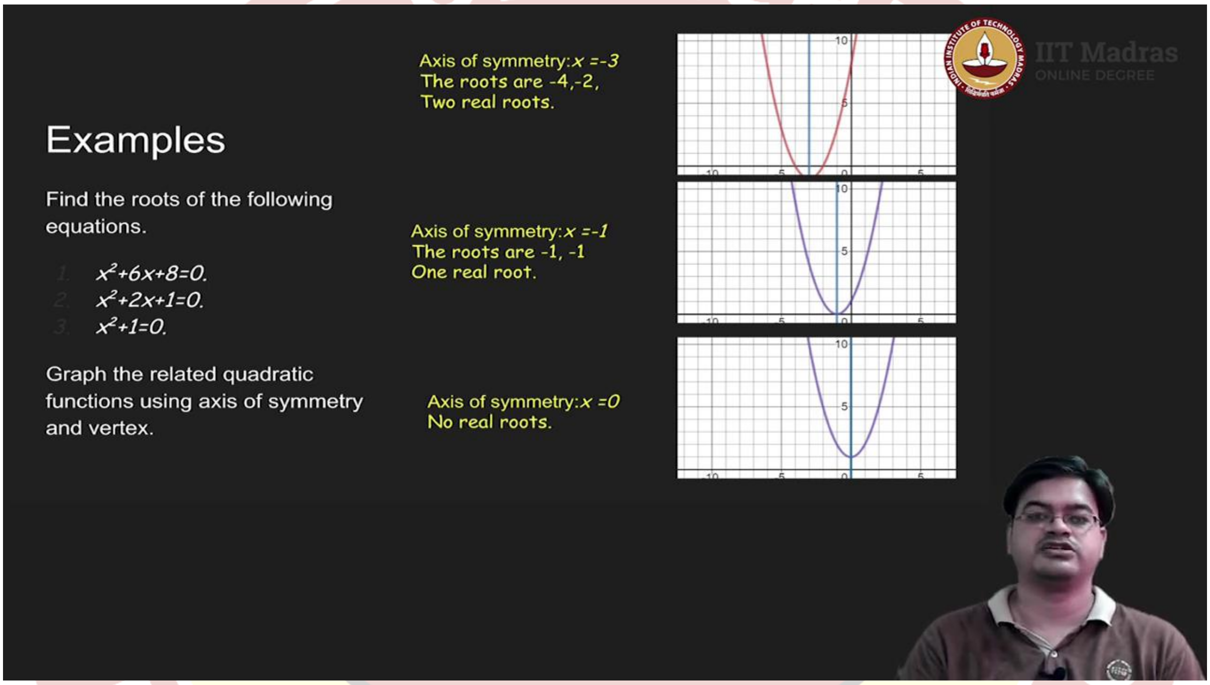
Now, once we have a Quadratic Equation in standard form, we can discuss about roots of the Quadratic Equation or zeroes of the functions. And we will see how the concept of roots of Quadratic Equation and zeroes of quadratic function are related, in this slide. So, the solutions of the Quadratic Equation are called roots of the equation.

What do I mean by that? If $ax^2 + bx + c = 0$, then what is the value of x, that gives me 0, is called the solution to the Quadratic Equation. And also, that value of x will also be known as root of the Quadratic Equation. So, this way we get the root of the Quadratic Equations.

So now, which way you can find the roots of the Quadratic Equations? One method, which is very easy. If you have a quadratic function associated with this Quadratic Equation, then you just plot the quadratic function and find its zeroes. What is a zero of a quadratic function? Zero of a quadratic function is nothing but its x intercept. So, in particular, if you observe that, zeroes of the functions are x intercepts of its graph and these are the solutions to the related equation, $f(x) = 0$, at these points?

So, if you are having a quadratic function, what you need to do is just plot it and see where it intersects x axis. If it intersects x axis, then you got the solution or the root of the Quadratic Equation. So, let us try to see this through some examples.

(Refer Slide Time: 04:32)



So, here are some examples. So, the question is to find the roots of the following equations. First equation is $x^2 + 6\,x + 8 = 0$. Second one is $x^2 + 2\,x + 1 = 0$. And third one is $x^2 + 1 = 0$. Now, we will take these equations one by one. So, essentially what we are proposing is, we want to chart these equations or we want to plot these expressions on a graph paper.

So, if you recollect from our last few videos, in order to plot a quadratic function, we need to understand the axis of symmetry of the quadratic function. So, let us take the first example, where you have $x^2 + 6x + 8$. Now, I want to understand, what is the axis of symmetry of this particular function.

Let us see. So, in this case, for our standard notation, our related quadratic function is $x^2 + 6x + 8$. So, $a = 1, b = 6$ and $c = 8$. So, y intercept is obviously 8. And axis of symmetry is $x = -\frac{b}{2a}$, which obviously means it is $-\frac{6}{2}$, which is -3. So, axis of symmetry is x = -3.

So, axis of symmetry is x = -3 and a, the value of a is positive. So, what are the things that we can conclude from our previous videos? That is, if a > 0, the curve opens up, the graph of the function opens up. It attends the minimum.

And the axis of symmetry in this particular example is x = -3. So, the simplest thing that we can do here is, put x is equal to minus 3 in this expression. And you will see that, the expression will take a negative value. That means the y value taken is negative. That means if the y value taken is negative, you can easily see, the curve opens up. That means it will intersect x axis in two points.

Now, we want to guess those two points. Without plotting, right now based on our visual interpretation of this curve, can we guess the two points? Okay. So, -3, the value is negative. That means, for -3 it is negative. Then let us check it for x is equal to -2. If you substitute x is equal to - 2, you will get $4 - 12 + 8 = 0$. So, one root I have got, which is -2. If -2 is one root, -3 is one, -3 is axis of symmetry. That means, at a distance one apart from this, there will be another root. That means - 4 will be the second root.

Wow. So, we were able to understand, that -4 and -2 will be the roots of this equation, without even drawing, just on the basis of what we have understood. So, what we have understood here is, -2 and - 4 will take the value 0 and for x is equal to -3, you will get one negative value. And based on that, you have prepared a table. And therefore, you can plot this graph easily. Right?

So, we will graph the related quadratic function, using axis of symmetry and vertex. We have already discussed this. So now, axis of symmetry x = - 3, the roots are - 4 and - 2. And therefore, the Quadratic Equation given here, $x^2 + 6x + 8$ has two real solutions, two real roots. How will the graph look like? It is very easy. We have already imagined the graph. Yes, so this is the graph, where - 4 is a point here and -2 is a point here. -4, - 2 are the roots. And here, it achieves the minimum, which is - 3.

So, you can easily plot this graph. Let us go to the second equation. Now, in this second equation, again we will consider the associated quadratic function. What is the associated quadratic function? $x^2 + 2x + 1$. What will be the axis of symmetry for this? $-\frac{b}{2a}$, that will be - 1. Because b is 2 and a is 1. So $-\frac{b}{2a} = -1$.

So, x = - 1, is the axis of symmetry for this particular quadratic function. Let us substitute the value of x = - 1, in this quadratic function. So, you will get $(-1)^2$, which is 1, 2×-1, which is – 2, + 1. So, you will get 0. Oh! so, -1 itself is a zero. Correct? But that is a point of the vertex, where it achieves the minimum. So, there using axis of symmetry, you can conclude that there cannot be any other point, other than - 1, where it will take the value 0. Because that's the point, where the vertex arises.

That means the axis of symmetry for the second equation is x is equal to -1. a is greater than 0. So, the curve opens up. And therefore, it achieves the minimum. And therefore, the roots are -1 and -1. What is the value at -1? It is 0. So, that is that itself is a root. And therefore, it has only one real root, which is repeated twice. So, in particular, the graph of a function will look like this.

Now, the next problem is very interesting.$x^2 + 1$, where if you compare this with a standard form of the equation, $ax^2 + bx + 1$, then you will get b to be equal to 0. That means this curve or this, the graph of this function will be symmetric about x = 0, that is y axis. And since a is greater than 0, the curve will open upwards. So, the curve is opened up.

Now, a > 0, it will achieve the minimum value. Where it will achieve the minimum value? At the vertex. So, what is the vertex of this particular function? Because x = 0, so, where it, you substitute x is equal to 0 here. So, that value is 1,$x^2 + 1 = 1$. So, (0, 1), so 1 is the minimum value of this function. Can this function be equal to 0 then? It cannot be. So, this will give us the answer, that axis of symmetry x is 0. There are no real roots for this particular function, because it never intersects x axis.

And the function will look like this. So, this in short summarizes, what are possible solutions in any scenario, $ax^2 + bx + c = 0$ is given to you. In particular, if $ax^2 + bx + c$, if you are able to find the vertex and the vertex takes the negative value and a is greater than 0, the curve opens up. So, it will have two roots which are real numbers. If the curve opens up, but the value at the vertex is 0, then it has only one root.

And if the curve opens up and it is above the X-axis, that is it takes a positive value on the vertex, y coordinate of the vertex is positive. Then it will never intersect x axis. In the similar manner, you it is for you to study, that when a is less than 0, what will happen. So, I can give you the rough interpretation. If a is less than 0 and it achieves the maximum on the vertex.

And if that maximum is positive, then it will have two real roots. If a is less than 0 and at the vertex, the value is 0, then it will have a single real root. And if a is less than 0 and it is below

X- axis, then it will have no real roots. So, these are the scenarios, that we can cover using this, this graphing technique.

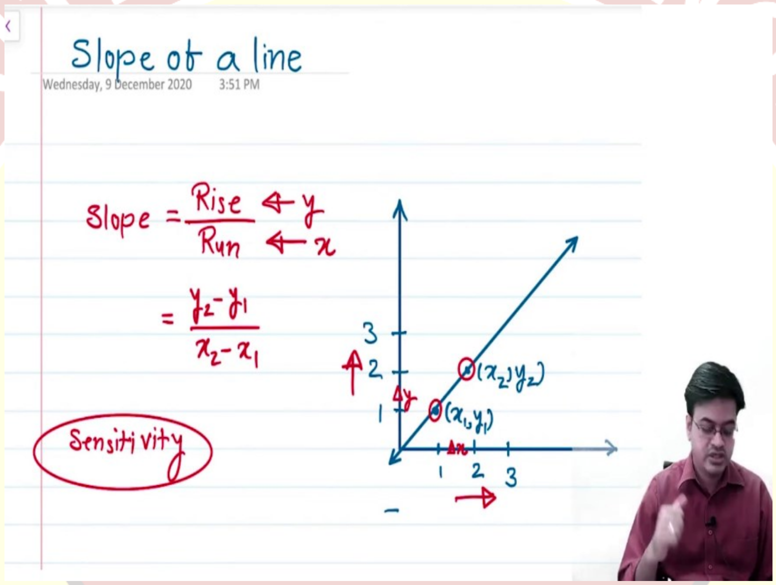
INDIAN INSTITUTE OF TECHNOLOGY MADRAS • सिद्धिर्भवति कर्मजा •

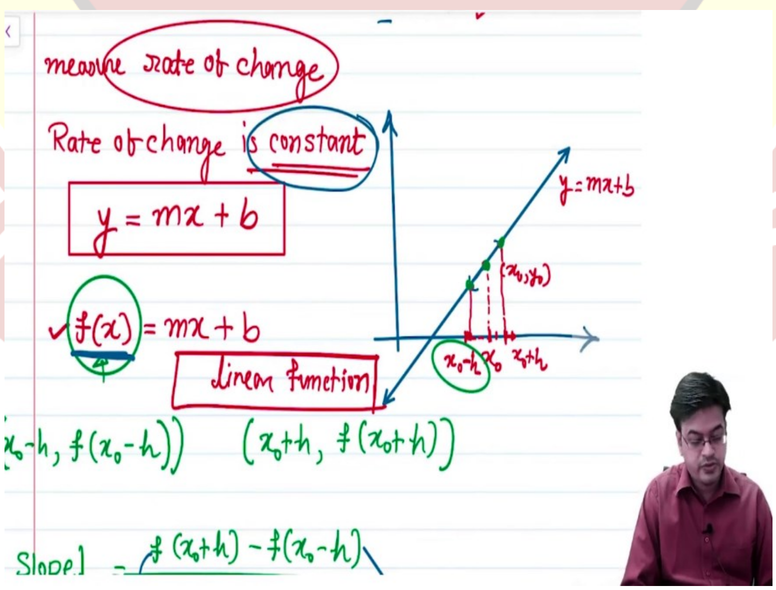
IIT Madras

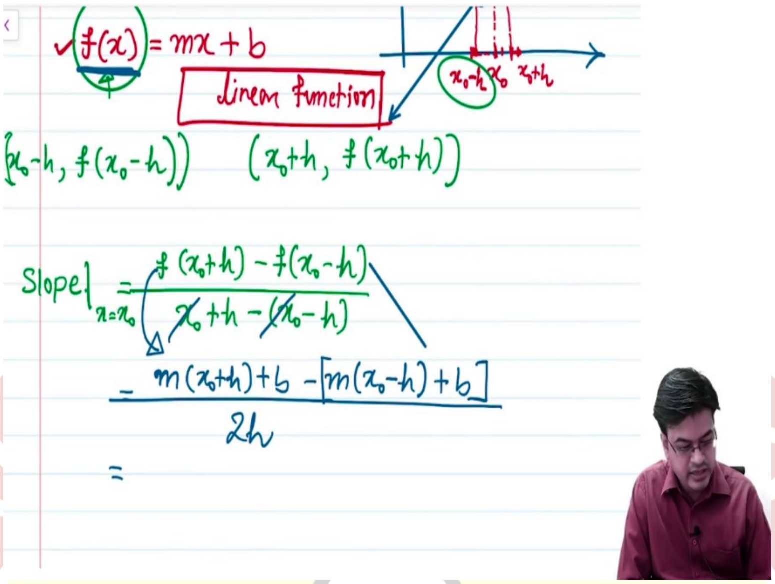
ONLINE DEGREE

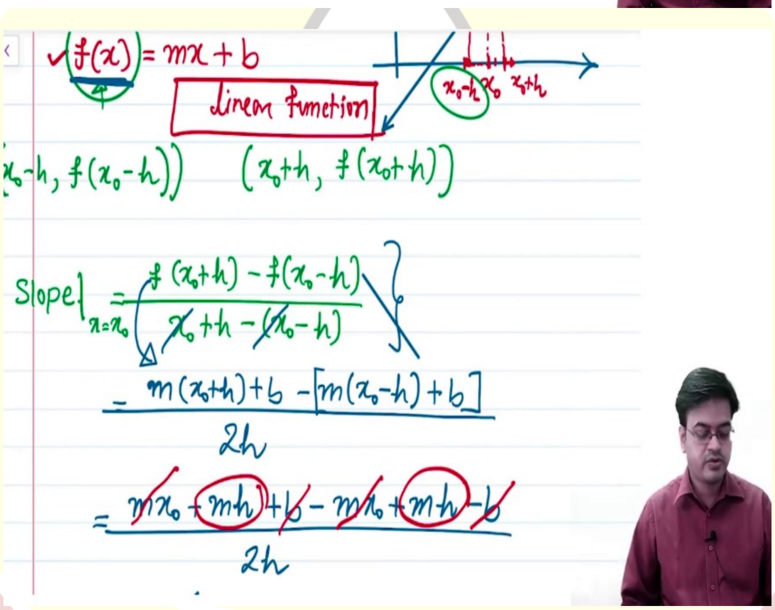
Mathematics for Data Science 1
Professor Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras
Lecture 26A
Slope: Line and Parabola

(Refer Slide Time: 00:14)









Hello students, in this video we are going to revise the concept of slope of a line and generalize this concept of slope of a line to slope of a quadratic function, let us begin. So, you are already familiar with slope of a line concept, so in particular if you want to talk about the slope of a line, then you will talk about slope of a line in terms of slope is equal to rise by run, so rise this is in the y direction and this particular thing run is in the x direction.

So, rise happens here and a run happens here, so given any two points on the line $x_1 y_1$ and $x_2 y_2$, if I want to talk about the slope of a line the formula simply reduces to $\frac{y_2-y_1}{x_2-x_1}$, we are all familiar with this, we have already seen the representation of line in terms of slope, this slope actually

INDIAN INSTITUTE OF TECHNOLOGY MADRAS
सिद्धिर्भवति कर्मजा

denotes the sensitivity of the expression with respect to change in x direction, so this slope actually denotes the sensitivity in the sense how small change in x effects the change in y.

So, if I change a small amount in x will it affect significantly in change in y, therefore this particular term actually said to measure rate of change. So, I can use the different notation, for example this thing we can write as Δy which is change in y upon Δx. So for Δx change in y there will be a Δy change in y direction, this is what it means. Now, when we measure the rate of change when I am talking about a straight line, there is no change in the rate of change, that is, the rate of change is constant if you are talking about a straight line is constant in a straight line.

And this fact in particular we used rate of change is constant on a straight line, which is the fact that we have used for deriving the equation of a line and we actually derive the equation of line for a non-vertical line where the slope is properly defined as $mx + b$ say, so this is the equation of a non-vertical line, you can also rewrite this expression in the form of functions which we have also studied. So, $f(x) = mx + b$. Now our goal so this is this in particular is called a linear function.

Now, if I want to generalize the theory of slope, why do I need to generalize the theory of slope first question, I want to generalize the theory of slope in order to be able to handle the concept of rate of change for all the functions. So, what is the rate of change of this function? That is what so in case of straight line it is a constant, so there are no brainers, but let us understand this notion in terms of functions.

So, this $f(x)$ is a linear function and this is the graph of a function $y = mx + c$ or $mx + b$. Now, I want to understand the concept of slope and suppose right now it is constant but suppose had it been a variable the slope had been a variable, how would I have encountered this? So, if I want to determine the slope at this point, which is say x_0y_0 what I would have done is I would have plotted projected this point on x axis over here this x naught then I can take a small interval a tiny interval of size h over here, so this point will be x naught plus h and I can also go behind h units which is $x_0 - h$.

And what I would have done is I would have projected this point over here, this is $x_0 - h$ and this $x_0 + h$, so slightly shaky diagram, but you understand the point. So, $x_0 + h$ and $x_0 - h$ and then what is a change from this point to this point that is what I would have calculated. So, let us try to formalize this in the language of functions.

So, this is $f(x)$. So, if the point corresponding point is $f(x_0 - h)$, then I can write the point as f of $x_0 - h$ is the point and on x axis that point is $x_0 - h$, so this is one point and then I have another point which is $x_0 - h$ and $f(x_0 - h)$. So, these are the two points which are typically denoted here in the green colour like this.

And I want to calculate the slope around this point x naught, so I have taken a symmetric distance $x_0 + h$ and $x_0 - h$. And then I will go about deriving the slope using my formula which is rise by run form. So, let us derive it. So, there are two points $x_0 - h$ and $x_0 + h$ and $f(x_0 + h)$ and $f(x_0 - h$ are the images of those points.

So, what is y? $y_2 - y_1$, so $\frac{f(x_0+h)-f(x_0-h)}{(x_0+h)-(x_0-h)}$, so this will be my slope. Now, at what point I am calculating? This is my slope because my centre is x_0 I am calculating the slope at the point $x = x_0$, so let me put it properly in the framework.

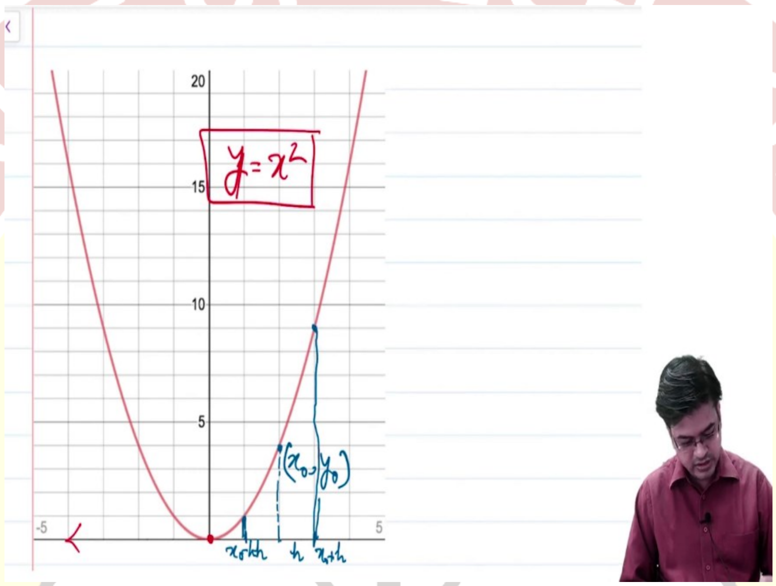
So, this is my definition of slope rise by run formula and now $f(x)$ is this, so I can simply use this definition of $f(x)$ and substitute it here, let me do it on the next line so that it will be less cluttered. So, here is the formula and let me do it here the different ink, so what is $f(x_0 + h)$ in this function $f(x)$ I will put $f(x_0 + h)$, so this will be $mx_0 + h + b - m$ times let me put a square bracket here in order to avoid confusion $x_0 - h + b$.

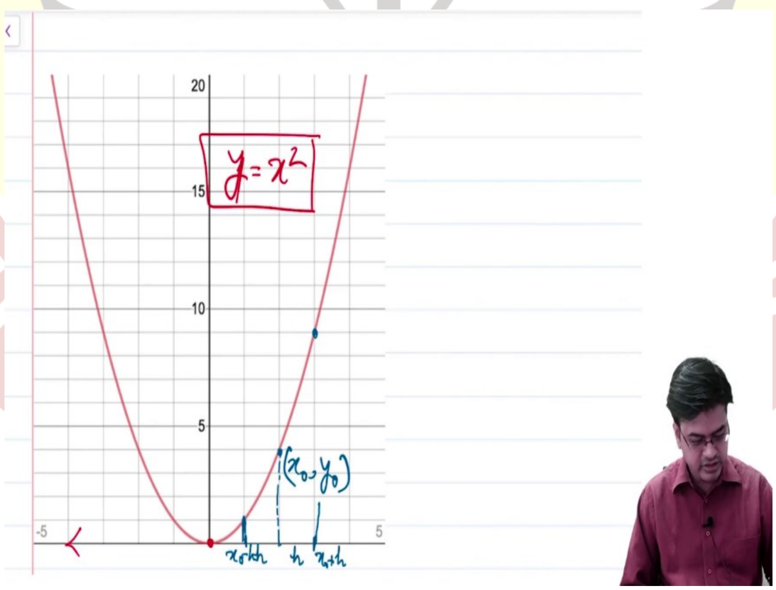
So, this is the first point $f(x_0 + h)$, this is the second point $f(x_0 - h)$, upon $x_0 - h$ will get cancelled, $h + h$ I will get, sorry this ink to shape forms so, so this will give me $2h$ in the denominator, so this is $2h$. Now, this is in fact going to be equal to so $mx_0 - mx_0$, so for clarity I will rewrite one more step, so that there will not be any confusion, this is $mx_0 + mh + b$ - will become + $mh + b$ will become $\frac{-b}{2h}$.

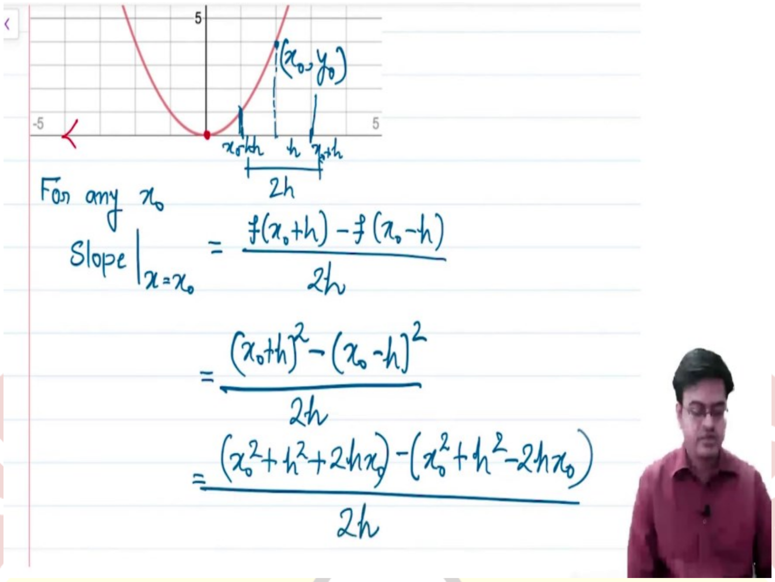
Now, as you can readily see this mx_0 cancels with this mx_0, b cancels with b and I am left with mh and mh, so together they will contribute to 2 times mh, so that will give me $2h \times \frac{m}{2h}$. And then $2h$ because $h > 0$ will cancel itself off and I will get the answer to be m, which is expected because I am dealing with a straight line.

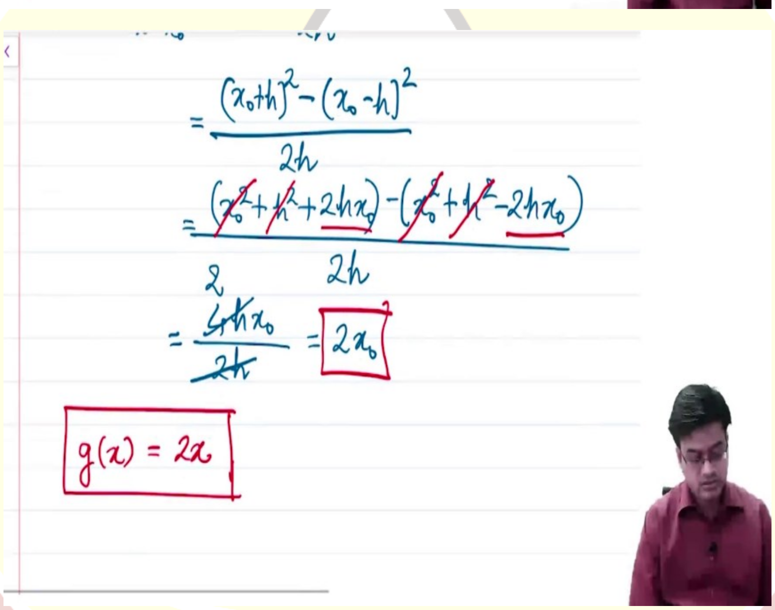
So, the slope of a line at x_0 is m, we already know and we also know that the rate of change of line is always constant. So, therefore my slope will always be m and now we have verified this using rise by run formula. So, our formula looks more sophisticated and more robust, so we can use this particular formulations to figure out the slope of a quadratic function. So, let us jump into slope of a quadratic function.

(Refer Slide Time: 10:38)









So, here is a quadratic function y is equal to x square that is given to you and I want to figure out the slope of this quadratic function at some point, let us say at point x naught y naught, so I will naturally use my same similar strategy I will project this point on x axis this is a point x naught I will go h time units h units on this side and h units on the other side, so this point will be x naught plus h and this point will be x naught minus h.

So, this is the point and then I will consider the corresponding point over here and the corresponding point over here on that line, so it will be here, let us erase this line it is not needed, so this is the point that that corresponds to $x_0 + h$ and this is the point that corresponds to $x_0 - h$ and I will use $f(x_0)$ definition.

Now, this definition is so robust that it does not distinguish between which x naught we are taking, so we can simply write for any x_0 I can write slope at x_0, because what is slope? Slope is actually calculation of rate of change, slope at the point $x = x_0$ is rise by run and that will be given by $f(x_0 + h) - f(x_0 - h)$ upon the horizontal length that you have travelled is 2h.

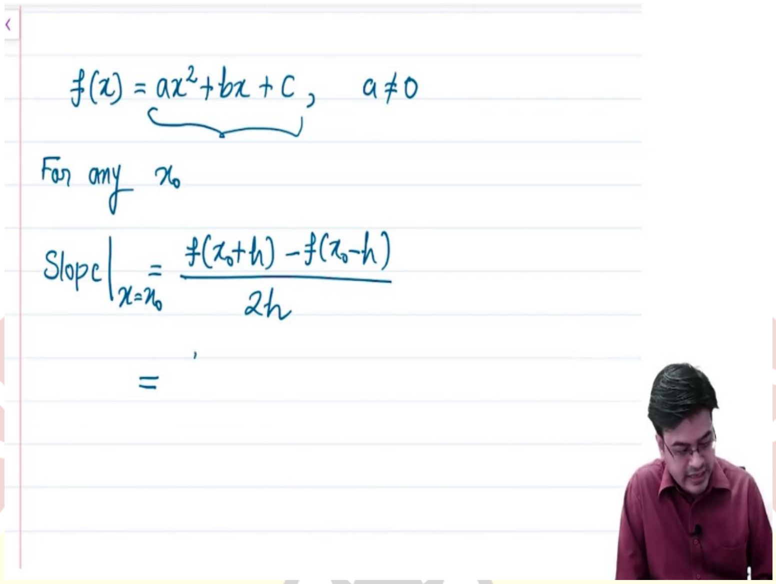
So, this will be 2h, clear. Now, let us bring in the function itself $y = x^2$ and see what is the slope of this particular function, so simply put in $y = x^2$ and therefore this will be $f(x_0)$ so it will be $\{(x_0 + h)^2 - (x_0 - h)^2\}/2h$.

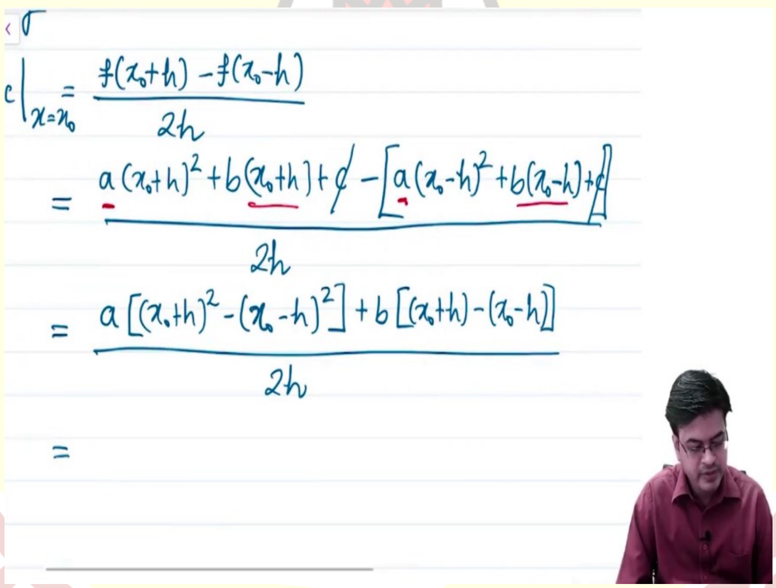
Now, you can simply do the calculation, this is again a quadratic equation, so you can simply write $x_0^2 + h^2 + 2hx_0$ minus again to be precise I will put the brackets intact $x_0^2 + h^2 - 2hx_0$, correct? Upon 2h, everything is intact now the cancellation job.

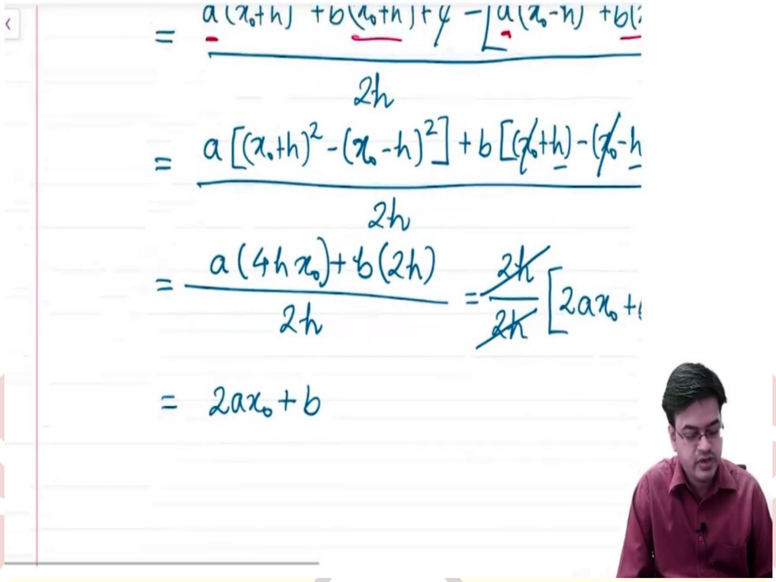
So, I will cancel x_0 from x_0, h^2 both are cancelled, I am left with $2hx_0$2hx naught minus minus plus $2hx_0$, so I will get here this is going to be equal to $\frac{4hx_0}{2h}$4 times hx naught upon $2h, h \neq 0$ and therefore I can cancel this off and I will get a factor of 2 here, so I will get the answer to be equal $2x_0$.

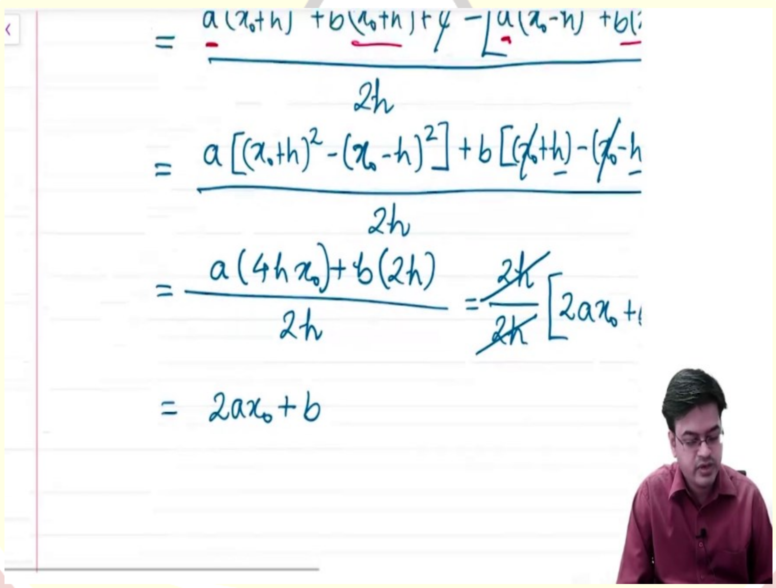
So, what is the slope at any point x naught? Slope on this curve at any point x naught this $2x_0$ the answer is $2x_0$, so in particular you can see this is nothing but another function $g(x) = 2x$ and this is the slope of this particular quadratic function. Let us, now generalize this concept to any quadratic function.

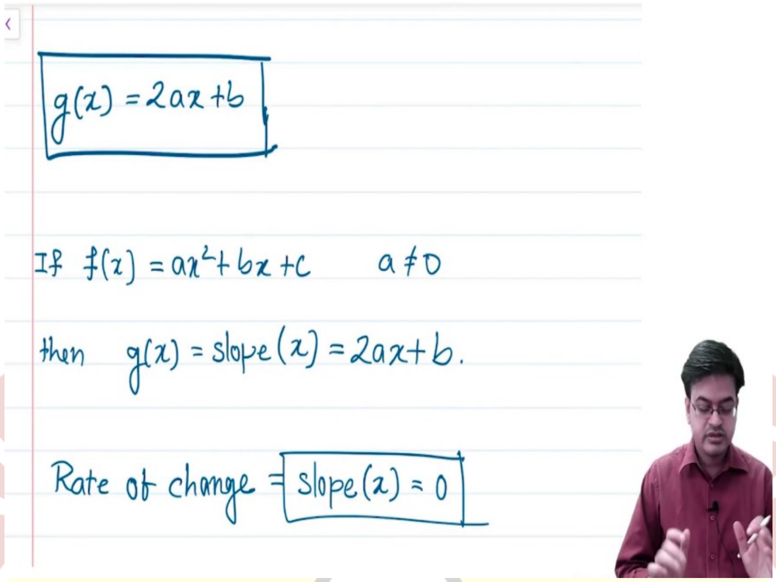
(Refer Slide Time: 14:48)

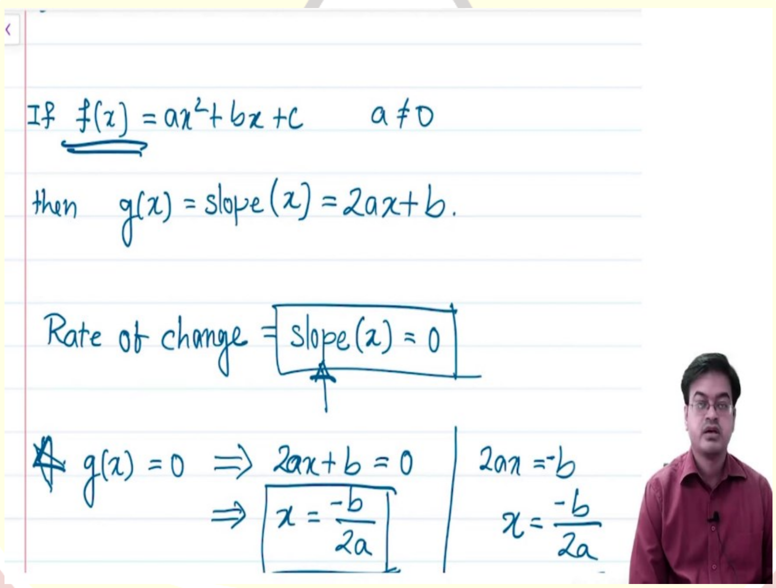


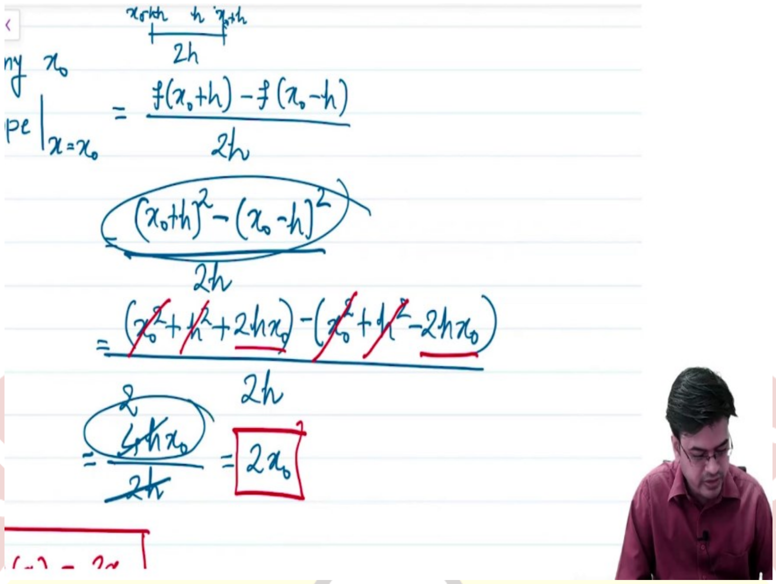



$= \frac{a(x_0+h) + b(x_0+h) + c - [a(x_0-h) + b(}{2h}$
$= \frac{a[(x_0+h)^2 - (x_0-h)^2] + b[(\cancel{x_0}+h) - (\cancel{x_0}-h}{2h}$
$= \frac{a(4hx_0) + b(2h)}{2h} = \frac{\cancel{2h}}{\cancel{2h}}[2ax_0 +$
$= 2ax_0 + b$


$= \frac{a(x_0+h) + b(x_0+h) + c - [a(x_0-h) + b(}{2h}$
$= \frac{a[(x_0+h)^2 - (x_0-h)^2] + b[(\cancel{x_0}+h) - (\cancel{x_0}-h}{2h}$
$= \frac{a(4hx_0) + b(2h)}{2h} = \frac{\cancel{2h}}{\cancel{2h}}[2ax_0 +$
$= 2ax_0 + b$


$\boxed{g(x) = 2ax + b}$
If $f(x) = ax^2 + bx + c \quad a \neq 0$
then $g(x) = \text{slope}(x) = 2ax + b.$
Rate of change = $\boxed{\text{slope}(x) = 0}$


If $f(x) = ax^2 + bx + c \quad a \neq 0$
then $g(x) = \text{slope}(x) = 2ax + b.$
Rate of change = $\boxed{\text{slope}(x) = 0}$
$g(x) = 0 \Rightarrow 2ax + b = 0$
$\Rightarrow \boxed{x = \frac{-b}{2a}}$
$2ax = -b$
$x = \frac{-b}{2a}$



So, let us all the standard assumptions kick in and let $f(x)$ be defined in this manner, $f(x) = ax^2 + bx + c$, in order to ensure that this is a quadratic function I will put the condition $a \neq 0$, because $a \neq 0$let us go ahead and calculate this. If $a = 0$ we have already calculated because it's equation of a straight line and you know the slope is constant and the slope will be b.

Now, let us see how to calculate slope of an equation slope of this particular function. Again use a same logic for any point x_0 any x_0 I have defined the slope as $\frac{f(x_0+h)-f(x_0-h)}{2h}$, this is slope at point $x = x_0$, go ahead and do a similar calculation by just substituting this function.

So, it is $\frac{[a(x_0+h)^2+b(x_0+h)+c]-[a(x_0-h)^2+b(x_0-h)+c]}{2h}$. Let us, simplify one by one slightly shrink it down so that it is visible.

So, let us first focus on these terms, a a so let us club them together and you will get something like $a[(x_0 + h)^2 - (x_0 + h)^2] + b$.. that is the second term I am clubbing in, this and this which will give me $(x_0 + h) - (x_0 - h) + c - c$ so c will cancel itself off, so I do not have to worry about that term, upon $2h$, this will again give me some idea.

Now, you can redo this calculation or you can use the calculation I have done in the previous example, so we had $[(x_0 + h)^2 - (x_0 + h)^2]$, this particular calculation and we showed that it was coming out to be $4hx_0$, so I will simply use the that calculation and substitute it here.

So, a into $4hx_0$ plus so a this is complete $+bx_0 + h$ so $x_0 - x_0$ will cancel itself, $h + h$ that is $2h$, so I will get b× $\frac{2h}{2h}$, so on further simplification you can simply take out $2h$ as a common factor upon $2h$ and here you will be left with $2ax_0 + b$.

And this $2h$ will cancel itself off and therefore my answer to that question that what is the slope of a general quadratic function is $2ax_0 + b$ for at any point $x = x_0$, so I can rewrite this as some function $g(x) = 2ax + b$. So, we have figured out one formulation, which is if $f(x) = ax^2 + bx + c$, that is a quadratic equation where $a \neq 0$ then if $f(x)$ is this then $g(x)$ which is nothing but slope at x, which is you can write this as slope at the point x is $2ax + b$.

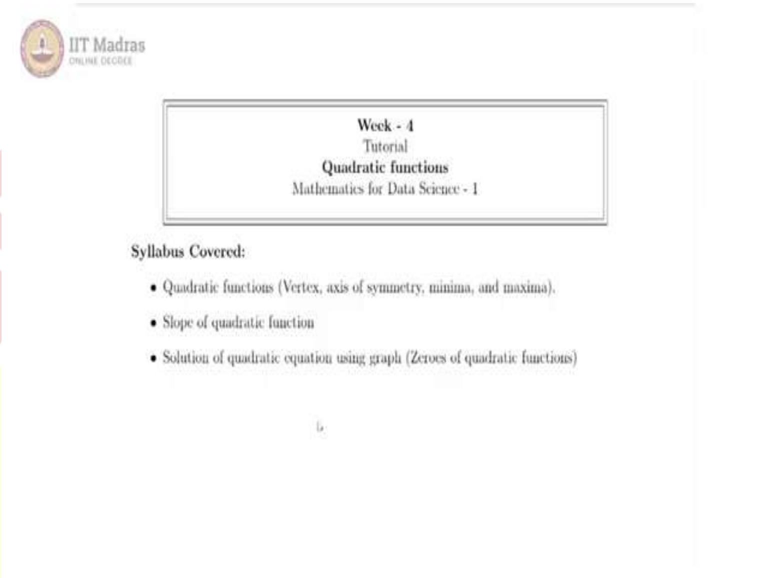
So, this is the characterization of the slope. Now, you can see the salient features of this slope. What is the slope? Slope as we mentioned in straight line case is rate of change. So, now you can ask a question rate of change is equal to slope, at any point if something happens and this slope at that point $x = 0$, then what is what is the consequence of that? So, this is an interesting question and we will try to answer this question.

So, let us say $g(x)$=0 for some x, so what is that x you can check it, that means this is 2ax plus b is equal to 0 which essentially means what you have got is $x = \frac{-b}{2a}$, if you are unsure of the calculation let me do it for you, $2ax = b$, $a \neq 0$ this is the assumption, therefore I cannot, $2ax = -b$, $x = \frac{-b}{2a}$. So, this is this is how I got this number $\frac{-b}{2a}$.

Now, you note this number is nothing but a vertex of the parabola, this number is nothing but vertex of the parabola that means this is the number at which the function the quadratic function assumes minimum or maximum, so this is an important application and you can easily you can give a physical interpretation of this as because the rate of change is coming to 0, because the rate of change is coming to 0 this rate of change prior to this might be positive and now it is going to become negative that means the rate of change is changing and therefore the slope 0 is going to be the minimum or maximum of the original function f. This is an interesting application and it comes when you study calculus, so points of maxima and minima then be easily derived using the expression for the slope, that is all about the slope of a quadratic function. Thank you.

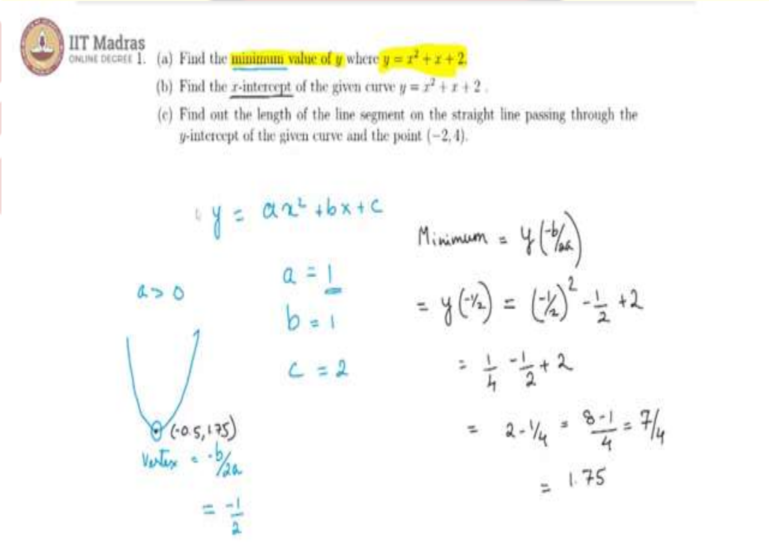
Mathematics for Data Science 1
Professor. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras
Week - 04
Tutorial - 01

(Refer Slide Time: 00:15)



Hello mathematics students in this tutorial, we are going to look at problems related to the topics covered in weak four. And so, these are the topics and we will begin with our first question.

(Refer Slide Time: 00:26)



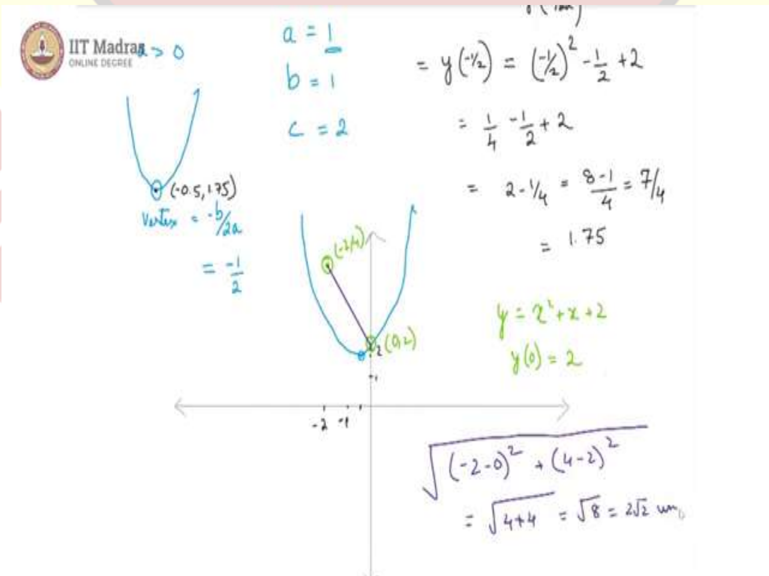
Here, we would like the minimum value of y for this particular quadratic function. And first, let us put down the quadratic function in its standard form, the standard form would be $y = ax^2 + bx + c$. In which case, our particular equation, the one that is given here would give us $a = 1, b = 1$ and $c = 2$. We are looking at the minimum value. Now, because the x square coefficient a is 1 that is a is greater than 0.

So, our parabola will be in this form, if a were lesser than 0, it would be inverted, it would be a downturned parabola, but right now it is in this form, and the minimum value is going to occur at this point, which is the vertex, which we know to be $\frac{-b}{2a}$. And so, we know our vertex for this particular equation is $\frac{-1}{2}$. And the value of y at $\frac{-1}{2}$ would be the minimum.

So, I can write the minimum is equal to $y\ (\frac{-b}{2a})$, which in this case is $y(\frac{-1}{2})$. And if I substitute that, I would get $(-\frac{1}{2})^2 - \frac{1}{2} + 2$, which is essentially $\frac{1}{4}$ - $\frac{1}{2} + 2$, which gives us 2 - $\frac{1}{4}$, which is equal to $\frac{(8-1)}{4}$, which is equal to $\frac{7}{4}$ and that is essentially 1.75. So, this point, here it is now we know it to be (-0.5, 1.75).

Now, they are asking us for the x- intercept and this is what we need to observe about the x −intercept.

(Refer Slide Time: 2:57)



Point (-0.5, 1.75) assuming this is 1 and this is 2, this is -1 of course, so this is negative side and this is -2. So, -0.5 is going to be somewhere here and on the Y- axis, this would be 1 and this would be 2, 1.75 is somewhere here so our vertex point is here.

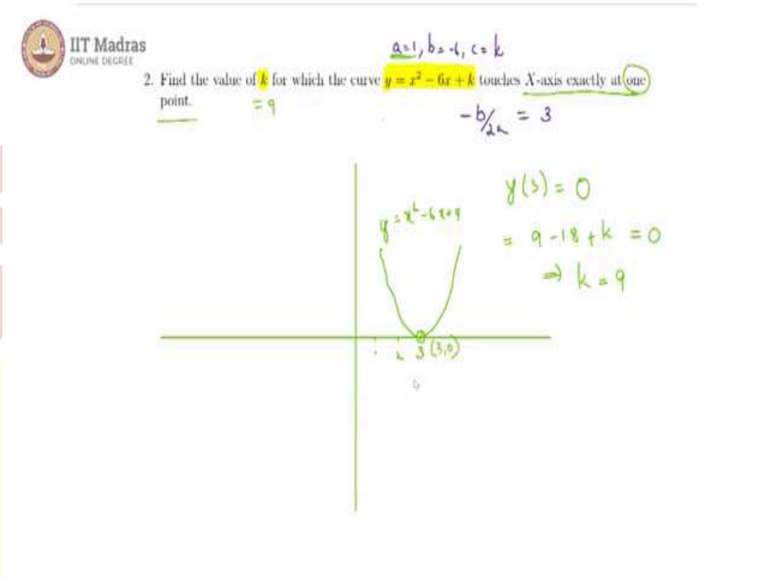
And from here, we know that this is an upward parabola, which is going to be something like this. And that means it never touches the X- axis at all. There is no x −intercept for this parabola.

And lastly, it is asked to find the length of the line segment on the straight line passing through the y- intercept of the given curve and the point (-2, 4). So, (-2, 4) is somewhere over here, and we need to find this point here the y −intercept. And the y −intercept is easy to obtain, since our curve is y = x^2 + x^2+ 2. y − intercept is obtained when the curve cuts the Y- axis that is when $x = 0$ so $y(0) = 2$, therefore, our intercept is actually (0, 2).

And the point we are looking at is (-2, 4) and it is the length of this line segment that we require. And that line segment we will get by using the Euclidean distance formula, it will be $\sqrt{(-2-0)^2 + (4-2)^2}$ which is essentially $\sqrt{4+4}$, that is root $\sqrt{8}$, which is $2\sqrt{2}$, units.

Mathematics for Data Science 1
Professor. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras
Week - 04
Tutorial - 02

(Refer Slide Time: 00:14)

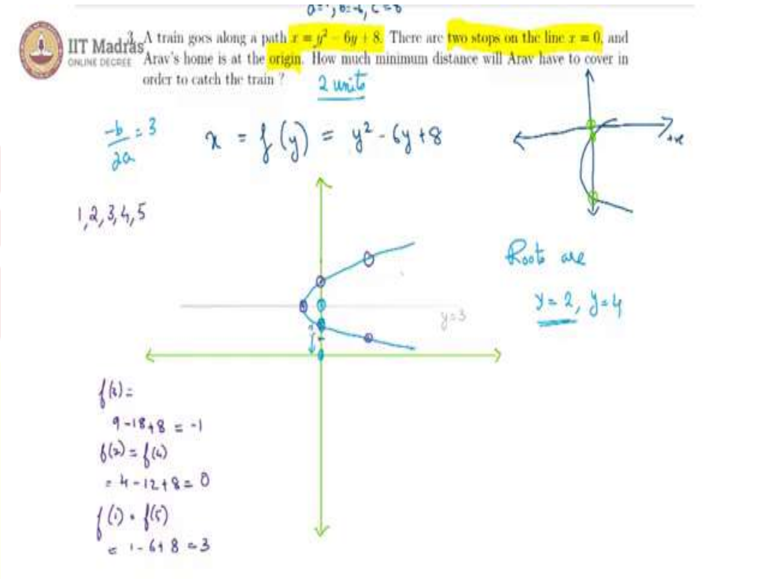


Now, second question we are going to have, this quadratic functions curve touches the X-axis exactly at 1 point. And for that what is the value of k supposed to be? First observation should be that the vertex is given to us. The vertex, which is $\frac{-b}{2a}$, here $a = 1, b = -6$, and $c = k$, thus the vertex is $\frac{-b}{2a}$, which is $\frac{6}{2}$ that is 3. So, this is 1, this is 2 and this is 3, our vertex is on this particular line that is $x = 3$. And we are told that it touches the X- axis, the parabola touches the X-axis at precisely 1 point.

We also can see that a is positive, so this is an upward turn parabola, upturned parabola. And if it touches the X-axis at exactly 1 point that is only possible when the vertex is right here on the X-axis itself, and from here, our parabola looks something like this. That means, for this condition to be satisfied at the vertex, $y = 0$ that is $y\,(3) = 0$. And that is equal to $9 - 18 + k = 0$. This gives us $k = 9$ that is it so $k = 9$. When that happens, our equation is $y = x^2 - 6x + 9$ and it has its vertex at (3, 0).

Mathematics for Data Science 1
Professor. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras
Week - 04
Tutorial - 03

(Refer Slide Time: 00:14)



In question 3, there is a path $x = y^2 - 6y + 8$. So, if we observe here, we are basically saying x is a function of y. And that function is quadratic, we have $y^2 - 6y + 8$. So, we are now switching the axis and so our parabola is expected to look something like this, or like this, because this is the X-axis and this is the Y-axis. Now, we see that the coefficient of y square, which is a is 1, and b is -6, the coefficient of y and lastly, the constant term c is 8.

Since, a is greater than 0, we expect that this is an upturned parabola. In the case of upturn in X, what we mean is it is towards the positive X-axis. So, our parabola is expected to be something like this. Of course, it could be moving about, we do not know where exactly it cuts the axis or where the point is. And for that, we will have to go further. They are saying this 2 stops on the line $x = 0$ that is on Y-axis and of course, these will be this point and this point, basically, if we looked at it in terms of our standard $y = f(x)$ these are what are the roots of our equation.

And Arav's home is at the origin, Arav lives at the origin so this is where Arav is. How much minimum distance will Arav have to cover in order to catch the train? So, the question is simple, you have two routes for your $x = f(y)$, and these routes will be on the Y-axis now, because we have switched the axis and which route is closer to Arav's home that is which route is closer to the origin. So, let us try to find out now, let us try to plot this particular graph and let us see where the tool train stops are. From the equation, we know that the vertex will be $\frac{-b}{2a}$, which is again $\frac{-(-6)}{2}$, that is 3.

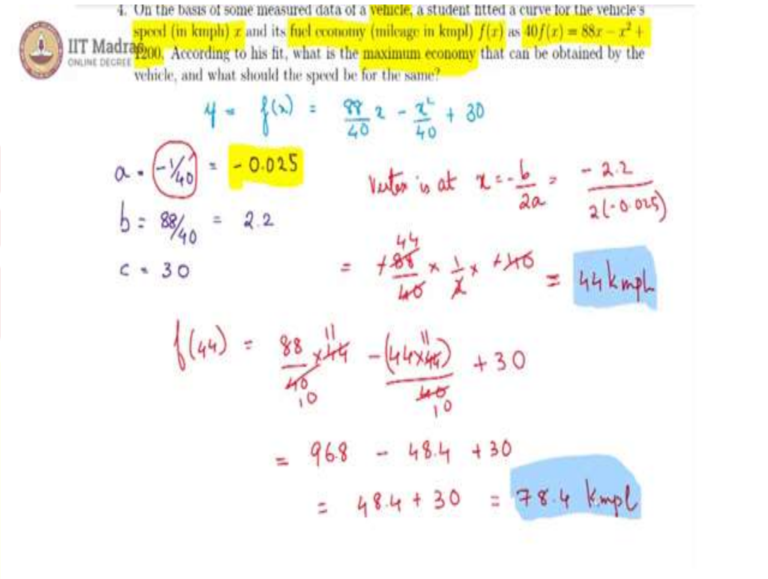
So, here we are basically saying $y = 3$ is the vertex. So, this is 1, this is 2, then this is our $y = 3$ and thus, the vertex will be along the line, $y = 3$, the axis of symmetry is $y = 3$. So, this is our axis of symmetry, $y = 3$. And for plotting the graph, we are now going to look at various points, which will be 3 and 1 to the other side of 3, 2 and 1 to this side of 3, 4 and then 5, and then 1, this should give us a reasonable idea of what the graph looks like.

So, f (3) at the vertex, what is the x-coordinate that would be $f(3) = 9 - 18 + 8 = -1$, so $x = -1$, which is going to be somewhere around here, this is our vertex. And f (2) will be equal to $f(4)$ because of symmetry. So, if I just substitute 2, I will get $4 - 12 + 8 = 0$, ok, that is good so we now have roots, we know that on 2 this point, and at 4 our curve is going to intersect the Y-axis.

So, if you want, we can further look at what is $f(1)$ which is also equal to $f(5)$, that is going to give us $1 - 6 + 8 = 3$, so we got to be somewhere over here, for these two points we are going to get somewhere here and thus our quadratic parabola looks like this. And we know for a fact that the routes are $y = 2$ and $y = 4$. Clearly $y = 2$ is closer to the origin. So, the minimum distance that Arav will have to cover is 2 units that is from the origin to this particular point, and this is the distance he will have to cover.

Mathematics for Data Science 1
Professor. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras
Week - 04
Tutorial - 04

(Refer Slide Time: 00:15)



In the fourth question, there is some data of a vehicle, and a student fitted a curve for the vehicle's speed x. So, this is our variable x and its fuel economy mileage in kilometre per litre as $f(x)$. So, it is a function of x and this function is given in this way, we are going to use it for y which means if we reduce it to the standard form, we will get $y = f(x) = \frac{88}{40}x - \frac{x^2}{40} + 30$. So, we have the situation where a the coefficient of x square is $\frac{-1}{40}$, which is equal to -0.025. And b is the coefficient of x which is $\frac{88}{40}$ and that is $\frac{22}{10}$, therefore 2.2, and lastly, $c = 30$.

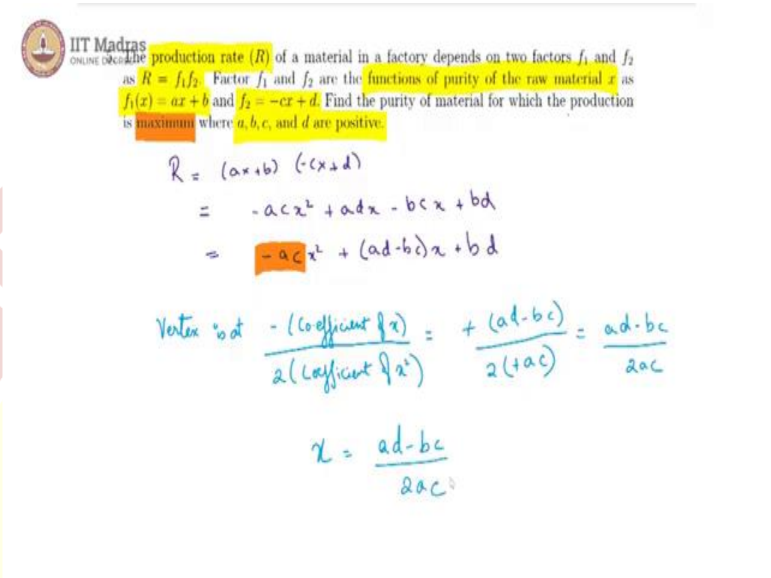
Now, we may observe that the x square coefficient is negative so this is a downturn parabola, which is why they are asking what is the maximum economy. So, at the vertex, you will get the maximum fuel economy so we need to find the vertex. And we know that the vertex is at x is equal to $\frac{-b}{2a}$, which in our case is then $\frac{-2.2}{2\times(-0.025)}$. This is probably better than in fractions.

So, if we write it down in fractions, we have $-b = \frac{88}{40}$ and this will be $\frac{1}{2}$ into $\frac{1}{2}$ and $\frac{1}{a}$ is then -40 itself, because a is $\frac{-1}{40}$. So, we have the 40 and the 40 cancelling off and minus and minus become plus 2, and 88 will give us 44. So, we have the vertex that is we get the maximum fuel economy at a speed of 44 kilometres per hour. And what is the maximum economy at this particular speed that we can calculate from our equation directly we have $f(44) = \frac{88 \times 44}{40} - \frac{44 \times 44}{40} + 30$ so this is 4, 10s a 4, 11s.

This is also 4, 10s and 4, 11s and we get $96.8 - 48.4 + 30$, which is then further equal to 96.8 is two times 48.4 so you will get $48.4 + 30$ giving us 78.4 kilometre per litre. So, we can say that this is our maximum fuel economy which is achieved at this particular speed.

Mathematics for Data Science 1
Professor. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras
Week - 04
Tutorial - 05

(Refer Slide Time: 0:15)



Our fifth problem looks a little complicated, but let us go one by one. And here we have the production rate of a material which is being made in a factory depends and it two factors f_1 and f_2 as $R = f_1 f_2$. And these two factors, they are the functions of the purity of the raw material. And that variable is x, x is the purity of the raw material. And both these functions are given to be linear $f_1(x) = ax + b$, $f_2(x) = -cx + d$. And it is given that a, b, c, d are all positive.

And it is asked find the purity of material, that is the value of x for which the production is maximum. So, let us understand what is being done here. We have two linear functions and the rate of production $R = f_1 f_2$, which will then $R = (ax + b)(-cx + d) = -acx^2 + adx - bcx + bd = -acx^2 + (ad - bc)x + bd$.

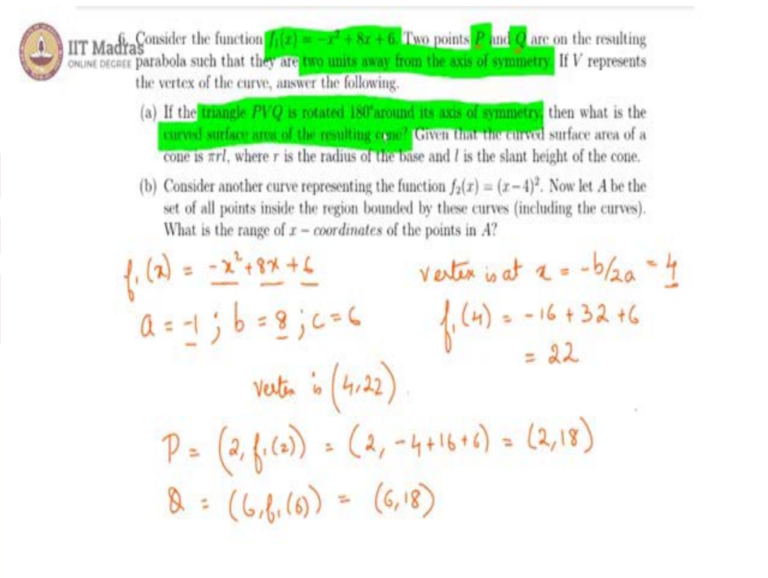
We are told that a, b, c, d are all positive, and that indicates the coefficient of x^2 is negative because the negative of ac and that means this is a quadratic function whose parabola is downturned, therefore, we will be able to get a maximum value at some point and this is going to be at the vertex, we know that this is going to be at the vertex. So, the vertex is at $\frac{-b}{2a}$, that is because here we have a, b, c, d already.

Let us write it down more carefully, that is the $\frac{-(coefficient\ of\ x)}{2(coefficient\ of\ x^2)} = -\frac{(ad-bc)}{2(-ac)} = \frac{ad-bc}{2ac}$, is where we will get the vertex. And since we know that the maximum is going to occur at this particular x, we get the $x = \frac{ad-bc}{2ac}$.

INDIAN INSTITUTE OF TECHNOLOGY MADRAS
सिद्धिर्भवति कर्मजा

Mathematics for Data Science 1
Professor. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras
Week - 04
Tutorial – 06

(Refer Slide Time: 0:14)

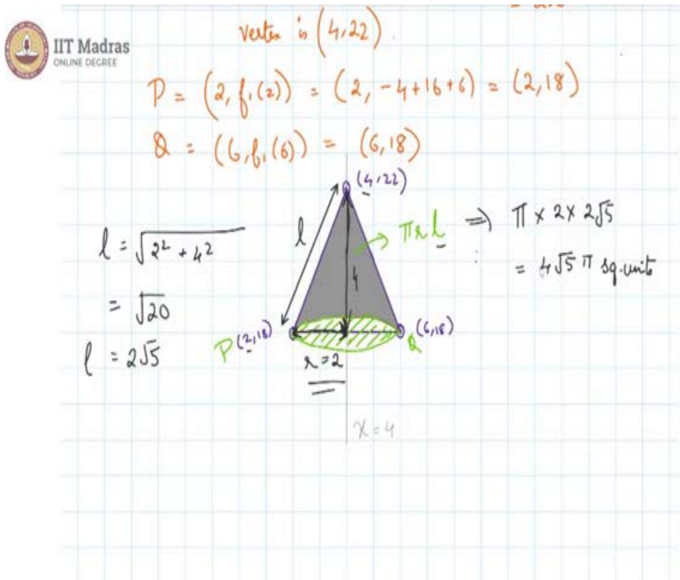


In our sixth question, we are given this particular quadratic function $f(x) = -x^2 + 8x + 6$. And we are told that two points P and Q, which are on this parabola such that they are two units away from the axis of symmetry. So, let us try to find out what the axis of symmetry is for this parabola. Our equation is $f_1(x) = -x^2 + 8x + 6$. And that would mean, in a standard form $a = -1$, $b = 8$, and $c = 6$.

And that would give us the vertex is at $x = \frac{-b}{2a}$, which in our case will then become $a = -1$, $b = 8$, so we will get 4. And the functions value at 4 is $f_1(4) = -(4)^2 + 8 \times 4 + 6 = 22$. So, the vertex is (4, 22). Further, we are told that P and Q are two units away from the axis of symmetry. So, the axis of symmetry is along $x = 4$, which means P and Q will be at $x = 2$ and $x = 6$, $4 - 2$ and $4 + 2$.

So, these points are going to be P$(2,\ f_1(2)) = (2, -4 + 16 + 6) = (2, 18)$. And the point Q is going to be $P(6,\ f_1(6))$ and from symmetry we know that this is also going to be 18, so $P(6,$ 18). And it is now told to us that the triangle PVQ is rotated 180 degrees about its axis of symmetry and we are being asked the curved surface area of the resulting cone. So, let us look at what this looks like.

(Refer Slide Time: 3:08)



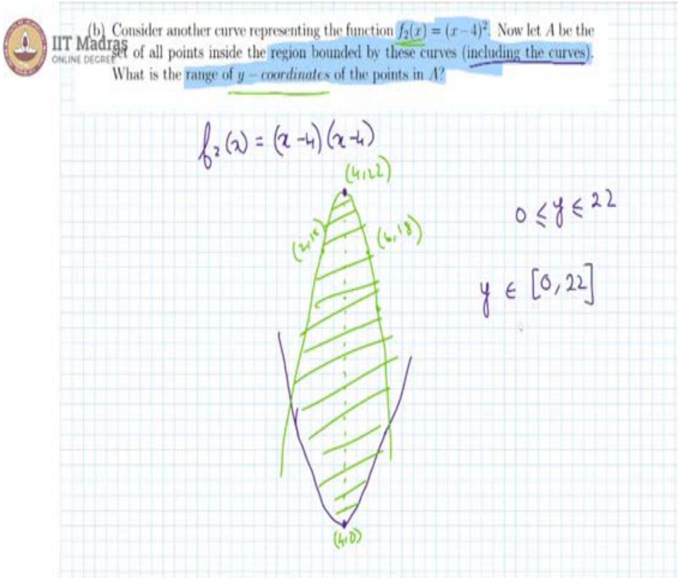
So, now let us suppose that this point here, let us call this our (4, 22), in that case 18 is 4 units below, so this will be the horizontal line passing through 18 and 2 will be here. So, (2,18) is here and this gives us (6,18) is here. This is (2,18) and this is (6,18). And that gives us a parabola which looks something like this, obviously a smoother curve than I have drawn, but something like this. And the triangle we are interested in is an isosceles triangle, which looks roughly like this.

This is the triangle that is being rotated 180 degrees about its axis of symmetry and its axis of symmetry is $x = 4$. I am erasing the parabola in order to focus on the triangle alone. If this triangle were to be rotated, this point which is our P, this is our Q, this point P basically goes around and reaches Q, whereas Q comes around and reaches P. And in this way, we have a cone in our hands and we want the curved surface area and that would be this region and the base circle is this flat surface below this is the base circle.

And we are interested in the curved region whose surface area is given to be $\pi r l$. So, what is r, r is the radius of the base circle. Which is basically then this quantity, this is r, which we can tell is $4 - 2$, so it is 2. And what is l over here, that is the slant height, which is basically this height, that height can be obtained as the hypotenuse of this base radius and height here, which is as we can see 4 units. So, $l = \sqrt{2^2 + 4^2} = \sqrt{20} = 2\sqrt{5}$.

So, we have $r = 2$ and $l = 2\sqrt{5}$, this gives us a curved surface area is $\pi \times 2 \times 2\sqrt{5} = 4\sqrt{5}\pi$ square units.

(Refer Slide Time: 6:46)



For the part B of our question we have another curve which is also quadratic and whose roots are basically 4 repeated. So, $f_2(x) = (x-4)(x-4)$. So, x being equal to 4 makes $f_2(4)$ =0. So, therefore, our root is 4 and it is repeated because coming twice here. So, let us now try to look at what they are asking. Now, let A be the set of all points inside the region bounded by these curves, including the curves. So, we are saying the region bounded by these curves and including the curves.

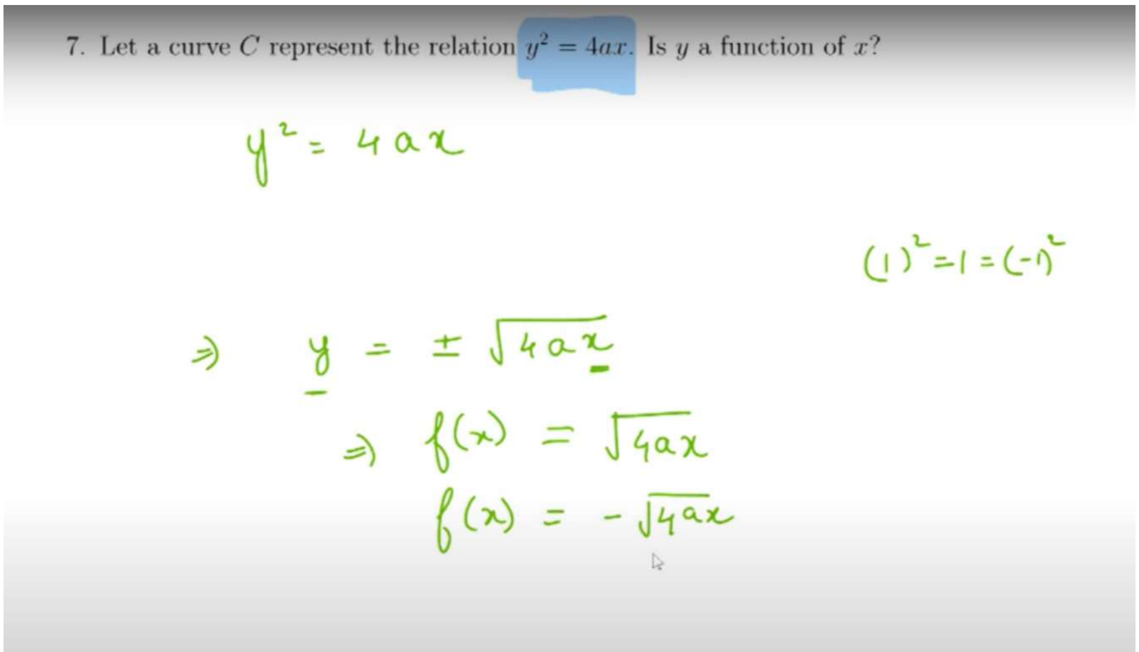
And they would like the range of y coordinates of points in it. We know already that (4, 22) is the vertex for our previous parabola. And it also passed through (2, 18) and (6, 18). And about this new parabola, the $f_2(x)$, we know that 4 is repeated root so there is only 1 root and therefore, at 4, that is 22, this would be 21, this is 20, this is 19, 18 17, 16, 15, 14, 13, 12, 1, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, and 0. So, this is going to be the repeated root and the vertex of our other parabola.

So, if one parabola is like this, f_1 had negative x^2 coefficient so it is a downturned parabola, then the other parabola $f_2(x) = (x-4)^2$ is an upturned parabola which is going to be something like this. So, these curves are going to intersect in some way this way. And we are interested in the range of $y-coordinates$. So that would be, what are all the $y-coordinates$ possible in this region.

So, if this is the region we are looking at, then clearly this is the upper bound of our $y-coordinates$ and this is a lower bound. So, $y-coordinates$ in our region range between 0 and 22. And they said including the curve, so 0 is also included, 22 is also included, so we can write the same thing as $y \epsilon [0, 22]$.

Mathematics for Data Science 1
Professor. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras
Week - 04
Tutorial - 07

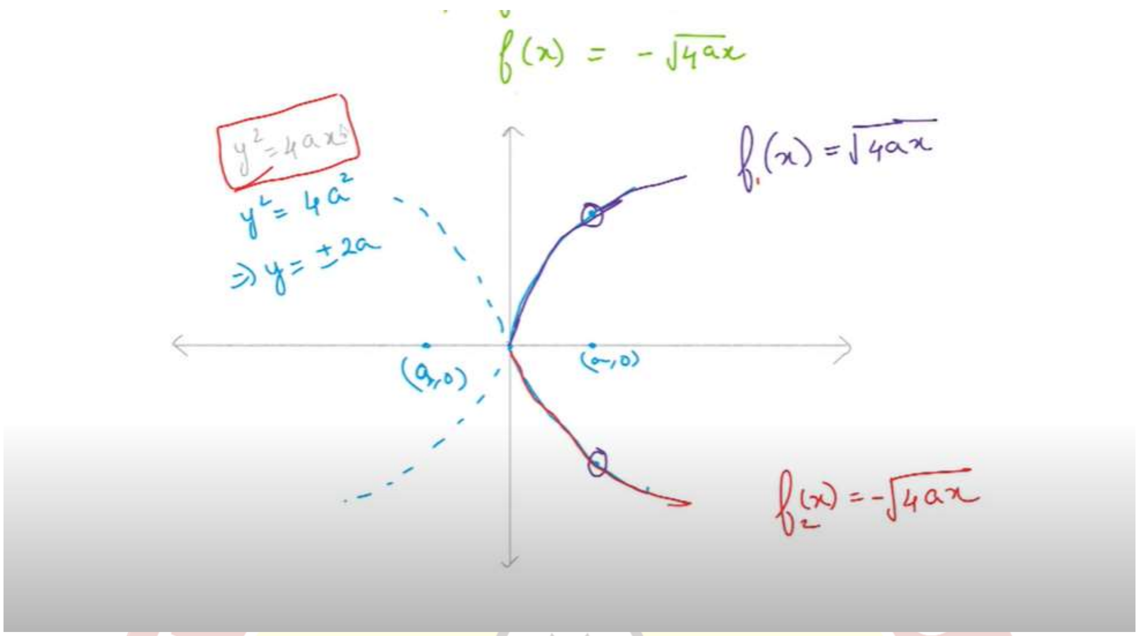
(Refer Slide Time: 0:04)



In question number 7, we have one relation given this way, $y^2 = 4ax$. And they are asking a very simple question, is y a function of x. So, we have $y^2 = 4ax$. And the interesting thing about square roots is, if I did the square root of 1, it is not just 1, it is actually ± 1. So, both $(+1)^2 = 1$, which is also equal $(-1)^2$.

So in this case, we need to consider the fact that y $= \pm\sqrt{4ax}$. Which means for the same x, I might have 2 different y's. So, put it this way, I am basically saying f(x) assuming it is a function is equal to $\sqrt{4ax}$ and f(x) is also equal to the $-\sqrt{4ax}$. And this is not allowed, for a single element in the domain, for a function, you should have only one image in the range. But here we have 2 different images for the same element in the domain. Therefore, this is not a function.

(Refer Slide Time: 1:55)



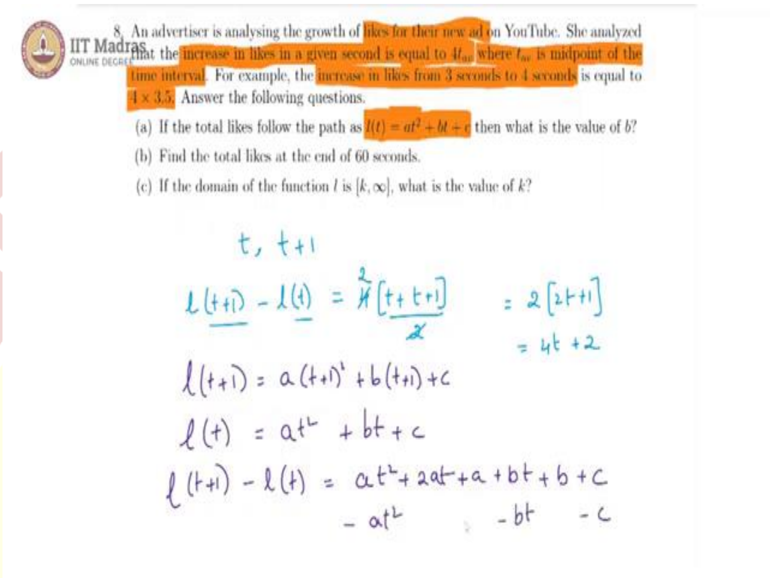
If we looked at it in terms of the plot, we have $y^2 = 4ax$, that is what we are trying to plot. And for $x = 0$, we get $y = 0$. So, this curve passes through the origin definitely. And for the next x value, I am going to take a, so therefore, $y^2 = 4a^2$, which gives $y = \pm 2a$. So, if a is positive, this is (a, 0), then 2a is going to be somewhere here like this, and -2a is going to be somewhere here like this.

And so, we have a parabola which goes something like this. And if a were to be negative, then this would have been (a, 0) and we would have a similar parabola in the negative direction. Either way, it is pretty clear that for a given value, you have two corresponding y values, for a given value of x you have two corresponding y values and that is not allowed for a function.

Independently $f(x) = \sqrt{4ax}$, which is this part of the curve, that can be treated as a function and $f(x) = -\sqrt{4ax}$, for convenience let us call this as f_1 and this is f_2. This is also possible to be treated as a function independently, but their combination which gives us this relation, that is not a function.

Mathematics for Data Science 1
Professor. Neelesh S Upadhye
Department of Mathematics
Indian Institute of Technology, Madras
Week - 04
Tutorial - 08

(Refer Slide Time: 0:15)

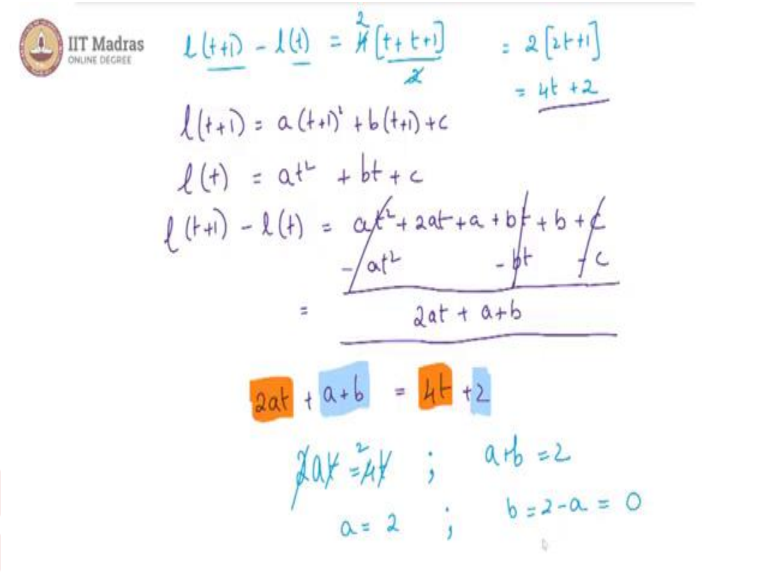


For our eighth question we have an advertiser who is analyzing the growth of likes for their new ad on YouTube. She analyzed that the increase in likes in a given second is equal to 4 times t_{av} , where t_{av} is the midpoint of the time interval, that is the average time in that time interval. And so we are given an example to explain what this is. The increase in likes from 3 seconds to 4 seconds. So, from the time $t = 3$ to the time $t = 4$, there is a number of increase in likes, which is equal to 4×3.5 and 3.5 is the midpoint of 3 and 4.

So, one way to write this is, let us look at time t seconds and the time $t + 1$ seconds. Then it is given to us that the likes at time $t + 1$, so, number of likes is a function of time. So, $l(t + 1) - l(t) = 4 \times t_{av} = 4 \times \frac{(t+t+1)}{2} = 4t + 2$, this is the difference in the likes from time t seconds to $t + 1$ seconds.

Now, it is further given to us that this particular function is a quadratic function. So, $l(t + 1) = a(t + 1)^2 + b(t + 1) + c$ and $l(t) = at^2 + bt + c$. Then $l(t + 1) - l(t) = at^2 + 2at + a + bt + b + c - (at^2 + bt + c) = 2at + a + b$

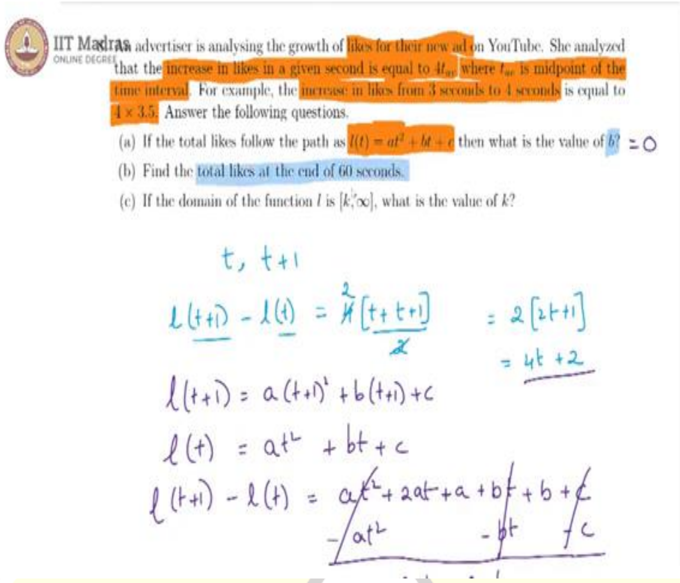
(Refer Slide Time: 3:08)



This quantity is basically equal to $2t + 4$. So, we are saying that $2at + a + b = 2t + 4$. Now, what are we supposed to acknowledge here is that the term with the t in it, that is the time dependent term is going to be same on both sides. Whereas the term which is constant is going to be same on both sides.

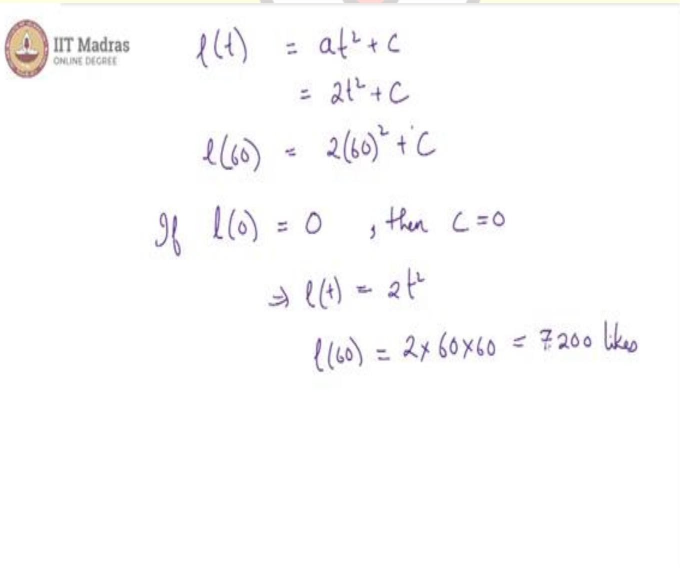
Thus, we are saying $2at = 4t$ and $a + b = 2$. This gives us 2 times t and t cancelled. So, we know $a = 2$ and that would imply $b = 2 - a = 0$.

(Refer Slide Time: 4:20)



And our question is asking us what is the value of b. So, we know this is equal to 0. Second question, the second part of the question is asking what is the total number of likes at the end of 60 seconds.

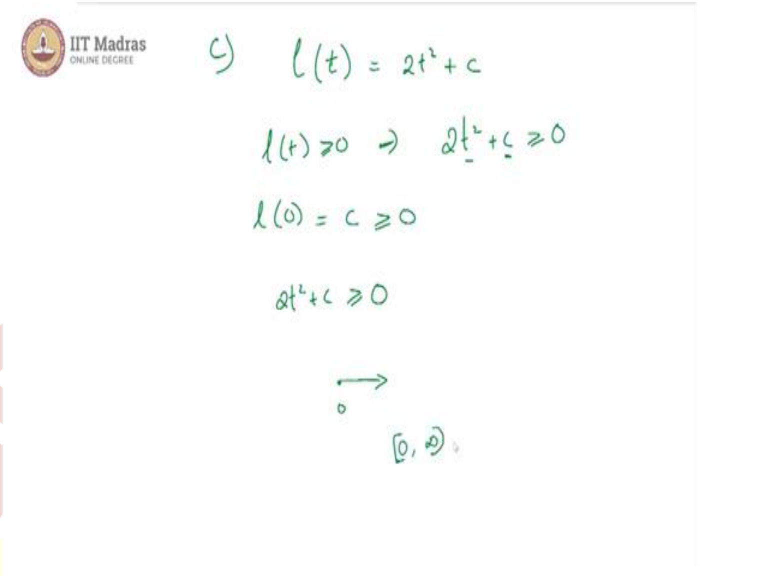
(Refer Slide Time: 4:49)



That would be impossible to calculate because we have the values of a and b, so we know that our $l(t)$, in this case we want l of 60. $l(t) = at^2 + 0t + c = 2t^2 + c$. But we do not know what c is. So, $l(60) = 2 \times 60^2 + c$. Now, if we made further interpretations that there were 0 likes at time $t = 0$. So, if $l(0) = 0$ then $c = 0$. So, this is a particular assumption we are making, we are assuming that the timer started when the likes were 0 and that would imply your $l(t) = 2t^2$.

So $l(60) = 2 \times 60 \times 60 = 7200$, that is 7200 likes at the end of 1 minute.

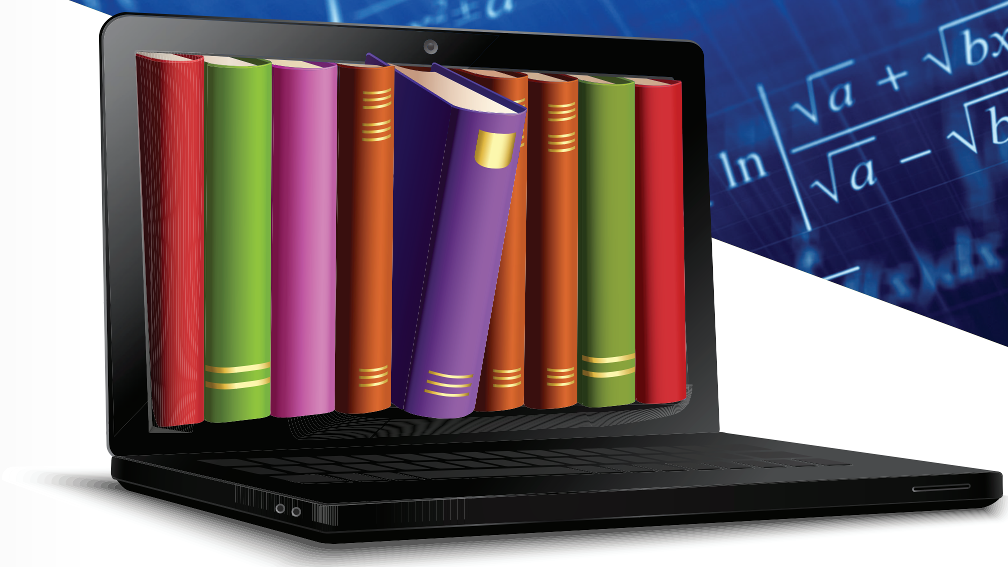
(Refer Slide Time: 6:18)



And lastly, for Part C, we are being asked the domain of the function is $[k, \infty)$, what is the value of k. We know that $l(t) = 2t^2 + c$. Now only real requirement we have is that our likes be greater than or equal to 0. So, $l(t) \geq 0 \rightarrow 2t^2 + c \geq 0$. Another thing we have is clearly that $l(0) = c \geq 0$, because at 0 time, it is not like you can have negative likes. So, $c \geq 0$.

Now we know that $t^2 \geq 0$ and now we also found that $c \geq 0$. So, $2t^2 + c \geq 0$, which means any time that is 0 or greater than 0. So, we are looking at the timer being started at a particular time and from there on, if this is 0 from there on your function is well defined and the number of likes will be greater than or equal to 0. So, the domain will be all the time from 0 seconds to ∞.

THIS BOOK IS NOT FOR SALE NOR COMMERCIAL USE



Scan QR code to chat with our WhatsApp bot

+91 9840776800

7850-999966
(Mon-Fri 9am-6pm)

Online Degree Office
3rd Floor, ICSR Building, IIT Madras, Chennai - 600036

onlinedegree.iitm.ac.in | support@onlinedegree.iitm.ac.in