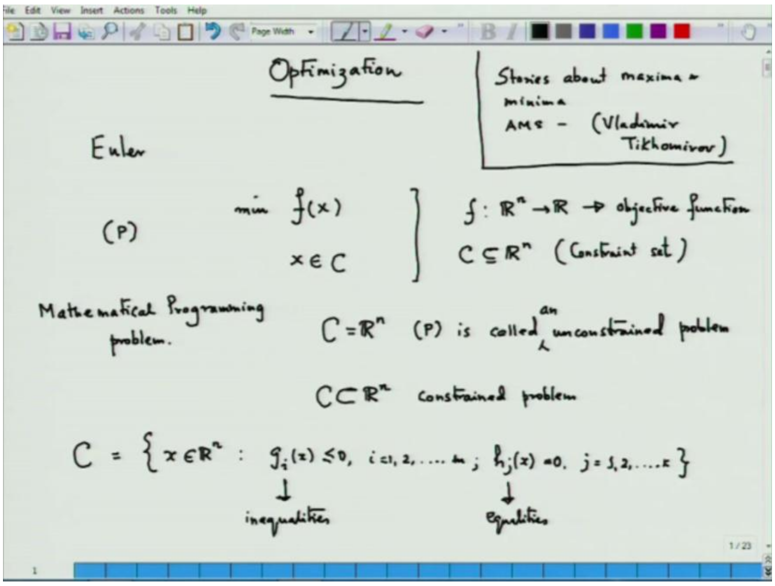
Foundation of Optimization

Prof. Dr. Joydeep Dutta
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Lecture - 1

Good evening viewers, I had already given a course on convex optimization. Those who have already seen me giving that course; compared to that course, this course would be less terse. This course is just called as optimization or may be the foundations of optimization. We are going to tell stories. This course is indented to show you that optimization is a beautiful subject. It is one of the most interesting and elegant areas of mathematics. We tend to divide mathematics into applied pure and so and so forth. But mathematics is mathematics. There is nothing applied or pure either you like, either the mathematics is beautiful or it is not. So, here we are going to show that we are indeed concerned with a very beautiful part of mathematics. We are not going to stress ourselves too much. We are going to remain in the nice world of differentiable functions. And within that nice world, from that world we would try to know how can we find the maxima or minima of a function?

(Refer Slide Time: 01:53)



So, essentially this is nothing but stories about maximum, minimum; stories about maxima a minimum is the name of a beautiful book written by a great optimizer Vladimir Tikhomirov is published by the American Mathematical Society which I write in short as A M S and the author is Vladimir Tikhomirov. Vladimir Tikhomirov is a very deep person and this book is just pure fun. So, at the end of the day we are going to have fun. So, what is optimization all about? What do we seek to do? The guiding principle in this world is that we want the best of things.

For example, we would want to have a good holiday, the best holiday with minimum expenditure. We would want to have the best chocolates. But we would like to have the prices less. We would like to have a good car with as much less cost as possible. A business man would always like to maximize his profit and would like to minimize his expenditure.

So, what I would say is that; optimization is prevalent everywhere, everywhere we see in our real social world. In the sciences, optimization is hidden and every where it comes up. So, there is a famous statement by this the famous mathematician Leonhard Euler; which says that nothing in this universe takes place without a law of maximum or minimum being satisfied. So, this is something very, very important to realize that optimization lies at the heart of scientific activity. If you and physics for example, stands on a very fundamental principle given by the French scholar Maupertuis called the least action principle. Whole of mechanics depends on least action principle. So, this word least is again pointing us that we need to minimize.

So, our problem that we will look into is quite a simple looking problem. We have to minimize f x over a set x element of C. Now of course, in our setting f would always be a function from R n to R and C is a subset of R n. Now optimization really got a growth when with the advent of calculus. So, calculus showed a very fundamentally strong way of approaching problems of maxima and minima. The derivative became a powerful tool in actually computing points where a function is maximized or minimized. This sort of finding the maxima and minima is something which you have already learnt in school, may be not so deeply but you possibly know what are why and how to do get the things?

My aim here is not to tell you just what you learnt in school, what optimization does? So, let me tell a bit about history. If you look at David Hilbert's 1900, in the year 1900 David the famous mathematician, David Hilbert gave a lecture in the international congress of mathematicians, and there one of the subjects he felt would give impetus to modern mathematics was the calculus of variations. Calculus of variations was the problem, was the subject which brought us into modern optimization and in fact a huge amount of functional analysis had been developed in order to solve the problems of calculus of variation.

Now of course, it was later realized during the World War 2, that there are many other aspects of optimization where which is not in the form of a calculus of variation problem. But, it is coming out in many social and many, many other business contexts or many other contexts like context of war fare. So, the real impetus in developing the subject for mathematical stand point whose literature is now vast has been with the introduction of simplex method in.

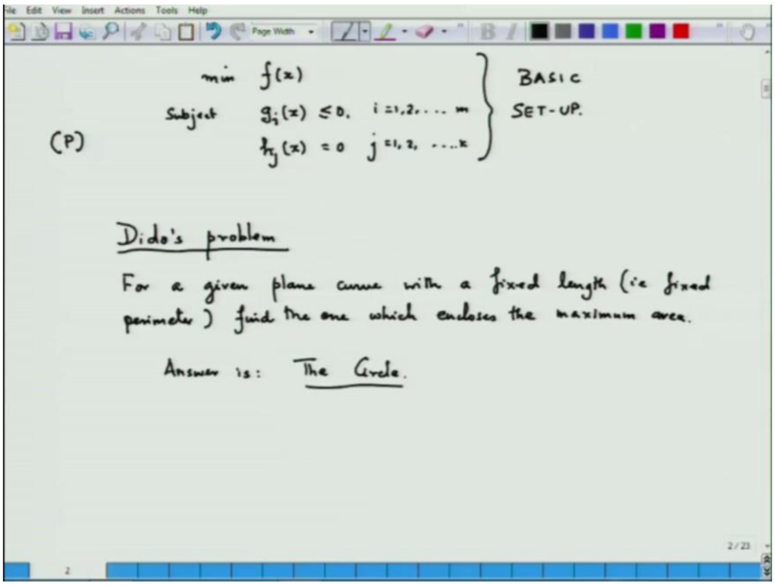
For linear programming problems, we in this course however would not discuss the simplex method. If you want to have a discussion of the simplex method, I would refer you to the lectures which I had given on convex optimization. Here we are looking into much simpler things for example, as you look into this problem which we will call a mathematical programming problem, so now I want to let you know that mathematical programming problem is fundamental different from computer programming. Now this term mathematical programming came in a very strange way when this linear programming was developed the simplex method.

Danzig was one day walking with T C Koopmans a famous economist and he was telling him that he was really working on solving problems whose objective and constraints are linear. He was trying to minimize such sort of problems where you have an objective function which is linear and constraint function which is linear. And he says that he could not find a name for it, but for all of these came from programs of the air force, he was trying to solve some issues with air force operations during the Second World War. So, Koopmans gave an idea why do not you call it linear programming. So, that name became popular and mathematical programming is optimization problems are finite dimensions in general.

So, this is now here the, if C is equal to R n which can be the case then P is called unconstraint problem. There are no constraints we will spend quite a bit of time with unconstraint problems because they are the easier ones. While if C is truly a subset of R n, and then we call it to be constraint problem. Now if you say that I just give an arbitrary C then it might be very difficult to figure it out, because at higher dimension how do I visualize the set C? Usually a set C, the set C is described by certain equality and inequality constraints. And here you see I have written down 2 sets of constraints; one described by equalities, one by inequalities. So, this is called the inequality constraints. So, all of these are inequalities and these are equalities.

As Teri-Rockefeller the greatest convex optimizer of our times had noted that the hallmark of modern optimization lies in the presence of inequalities in the constraints. Constraints means I am restricting my choice of x that is this function f which I want to minimize is usually referred to in the literature as objective function. And this set C is called the constraint set. So, what I am essentially trying to do is that; I am trying to restrict my x when I am telling that the constraint is present. That is C is a subset of R n, I am restricting my x. That is, I do not want to know what the maximum is or minimum value of the function when x is outside C. My total concentration would be on the set C itself.

(Refer Slide Time: 12:45)



So in general a mathematical programming problem would be written as follows. This is what we will be largely bothered about. So, the problem P once C is described like that and be written as. So, this is our basic set up and we would like to work in this set up. Now let me tell you that optimization problems do not just arise in business engineering economics. By the way just I want to, once I have taken the name whole of economical theory, whole of modern economics is based on optimization. You cannot have a true understanding of economic theory, modern economic theory without a true and without a good understanding of optimization.

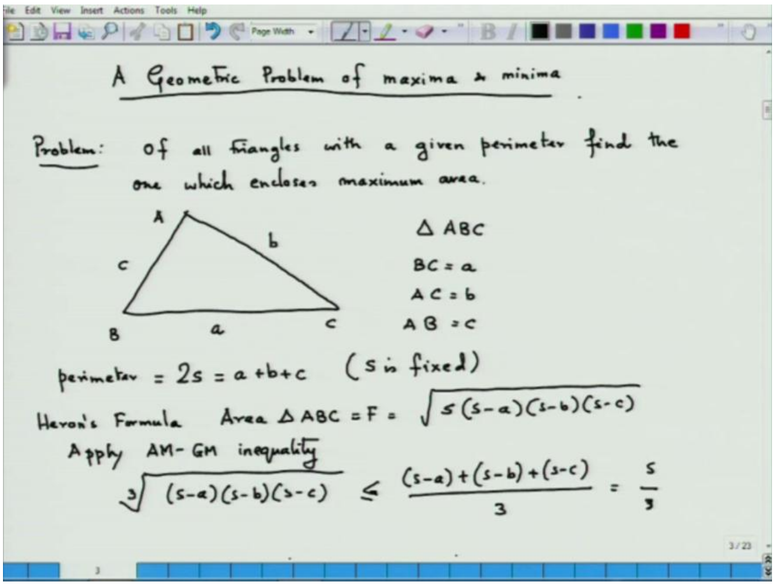
So, there are many areas which optimization permeates. But it is not that optimization problems only permeate in those things, optimization problem arises in mathematics also.

Let us go back to history a bit and I will tell you what this famous problem called the Dido’s problem the old or the ancient optimization problems where all of all geometric in nature.

So, today I will give a demonstration of how we can solve a geometric problem, geometric problem of maxima or minima. But let me tell you a little bit of history and talk about Dido’s problem. Dido was a Phoenician princess. So, she was fleeing from the prosecution of her cruel brother. She kept on kept on moving down the Mediterranean Sea and came to one place which attracted her attention. There she mates the local leader Yakub, and Yakub asked how much land do you really want? She told that you take up a bull’s hide, a bulls skin and make thin pieces out of it and now join up the pieces and now you encircle the maximum area that you can by using those pieces by joined up by joined up bulls hide. So, what she was asking was quite enormous actually.

Yakub did not realize that it would take up a huge amount of land and it is in this place a modern city of Carthage was founded. Now in our modern terms it says that Dido’s problem says, for a given plane curve with a fixed length that is fixed perimeter, find the one which encloses the maximum area. Are you trying to guess the answer? If you are trying to guess the answer, you can try it for few minutes. But, I am slightly restless I need to tell you the answer. Answer is, so answer to the Dido’s problem is the circle, the beauty of optimization has a subject lies in the fact that it is intimately tied up with geometry. And geometry specifically Euclidean geometry is not only the one of the most important parts of mathematics but it is possibly one of the most beautiful.

(Refer Slide Time: 18:06)

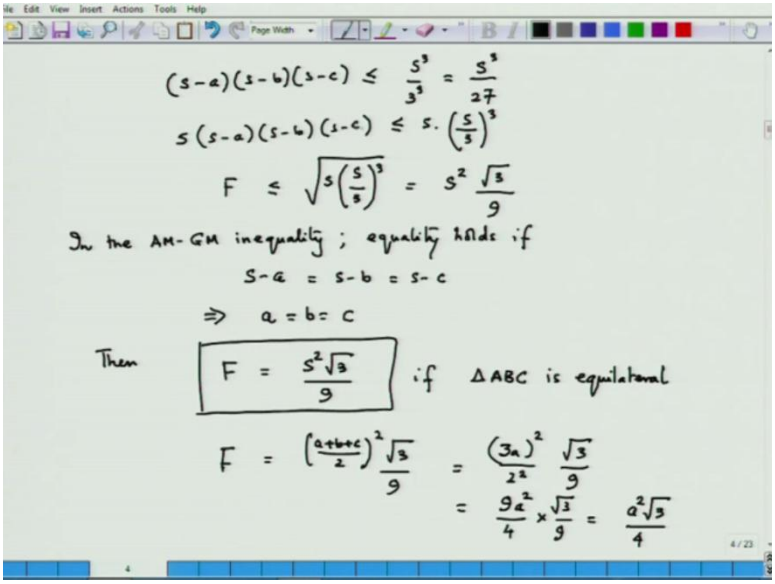


So, here we would first go in and try to talk about a geometric problem of maxima and minima and try to really solve it. So, let us talk a very simple problem. So, let us write down the problem of all triangles with a given perimeter. Find the one which encloses maximum area. Again you see we are dealing with Euclidean geometry. So, here is a triangle, say let us call it triangle a b c. So, our given triangle is triangle a b c and you know how the sides are called. The side opposite to a, that is b c is a side opposite to b c a is b and a b is c.

So, I Have b c is equal to a a c is equal to b and a b is equal to c and now how do I start usually perimeter is denoted by 2 s where s is half of the perimeter it is standard and this. So, once I know this now this for our problem S is fixed. I think you are trying to guess the answer but you will see how mathematically this answer comes out beautifully. And we know by Heron's formula, the area of triangle a b c which we denote by f is given in the following way; S into S minus a into S minus b. I hope you are remembering your school days S minus c. Now once I know this fact, what can I do?

Now we will use the arithmetic mean and geometric mean. Arithmetic mean and geometric means; this inequality is fundamental to arithmetic. It is fundamental to geometry and because large amount of this maximization minimization problem depends on this when they are of geometric nature. Now arithmetic mean is A M G M inequality. So, apply A M G M inequality that would give me the cube root of s into S minus a into so I take S minus a, S minus b and S minus c. These are my quantities, 3 quantities. So, cube root of geometric mean is less than the arithmetic mean. So, here I will use the inequality that a plus b plus c is equal to 2 S that will give me S by 3.

(Refer Slide Time: 23:19)



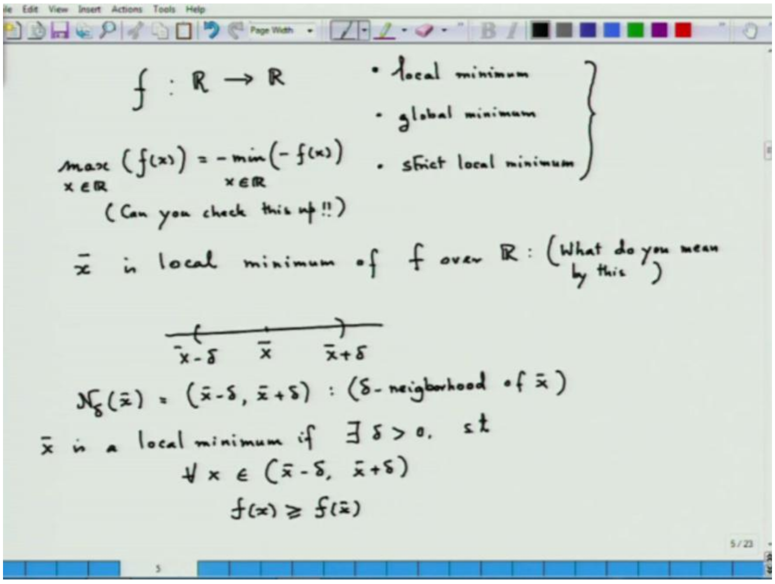
Now from here, you must have observed that f square or may be just we can look in from the A M G M inequality, we can immediately say that S minus a into S minus b into S minus c. You will cube both the sides to get because these are all positive, S minus a into S minus b into S minus c. So, the cube would be a positive quantity is less than s cube by 3 cube is 27. That is s cube by 27. Now so just let us if I multiply by s, so this is s into will see why we are writing like this S by 3 whole cube. So, f is obviously less then root over S into which comes out to be S square S cube s to the power 4 into root 3 by 9. That is 27. When you take a root of 27 3 into 9, and so here I would have S S cube by 3 cube s 4 by 3 cube S 4. If you take the square root it will become s square and by 3 cube which is 3 cube. So, I will put 3 by I will multiply up and down by 3.

So, I will get 3 square out, so which is 9 by root 3, which would be left. Now once I know this in the A M G M inequality equality holds if S minus a equals S minus b equals S minus c. So, this would imply that in our A M G M inequality, equality would hold when a is equal to b is equal to c and then f is equal to s square. So, the area is this; if triangle a b c is equilateral, so it is for a given perimeter it is equilateral triangle which encloses the maximum area and that is the value of this area. That is you can write if s is a plus b plus c by 2, so you know if a is equal to b is equal to c, it will become so if a is equal to b is equal to c you can write f has what is S? S is a plus b by, so it is a plus b plus c by 2 whole square into root 3 by 9. But now this 1 a is equal to b is equal to c. So, let me just write 3 a whole square by 2 whole square into root 3 by 9, which is 9 a square by 4 into root 3 by 9. So, it is nothing but a square into root 3 by 4.

So, this is a neat answer. So, this when I have an equilateral triangle with this every side of length a, then this is the area that it will enclose and among all such triangles with the given fixed a perimeter s that will enclose the maximum area. So, you see just by using A M G M inequality we were able to solve and in those days they were no such sophisticated tools like what we have.

Of course, you can analyze it from a more functional point of view and use derivatives and all those things which we will do quite soon. But still you see how beautifully geometric methods can be used because these are ancient problems in mathematics has not been developed to this level. There was no calculus, so you really have to use your geometry have to use basic ideas to do it. This looks very elementary but please note that when we use the word elementary it does not say, it is easy, it only says that number of tools required to analyze this problem is less. So, once you have done this, let us get going and do a little bit of more definition type study little more with more erudite type thing.

(Refer Slide Time: 29:14)

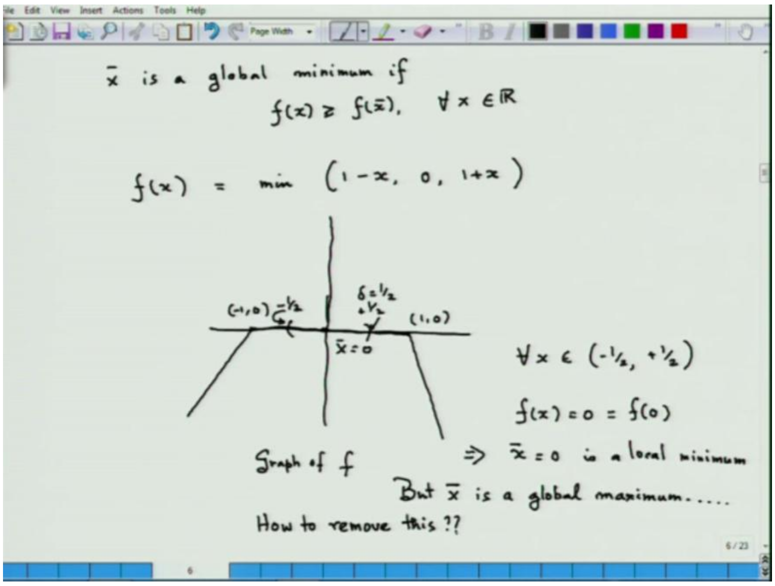


So, let me now consider a simpler situation. I have a function from R to R and now I want to recall to you few notions which I will write down one by one. I will write down things only for the notion of minimum. Because we will just minimize maximizing is just an opposite operation. And you at your leisure write down the same definitions for the maximum. Please do not neglect this, because this is a good exercise.

So, what we I want to talk about is what is a local minimum, what is a global minimum and what is a strict local minimum? I also tell you to fix up this check, this inequality max of a function f x over say x in R n. Suppose there is a maximum value, then this is nothing but so my question is to you to check this, can you check this up? Now I want to write down the definition of local minima. Once you want to talk about a definition of a local minima, so I first ask the question x bar is a local minima. What do you mean by this till this? So, my question is what do you mean by this?

Now, if I look into this very carefully, once again when I say local, I must be able to localize and in real analysis, when localize a point, you do it through the notion of the neighborhood. That is if you have a point say x bar here, we say we consider a neighborhood to be an open interval. Usually, taken to be of same length on both sides of x, but it is not necessary for our purpose. Let us take it. So, this so a neighborhood of x bar which we usually denote as usually denoted as an x bar is an open interval of this form. I will call delta because I will have kept the distance delta, it is called a delta neighborhood of x bar. Delta neighborhood of x bar. So, x bar is a local minimum if there exists a delta greater than 0, such that for all x which lies in this delta neighborhood f of x must be equal to f of x bar. A global minimum is that for whatever delta you choose this will happen. So, this is the definition of a local minima and x bar.

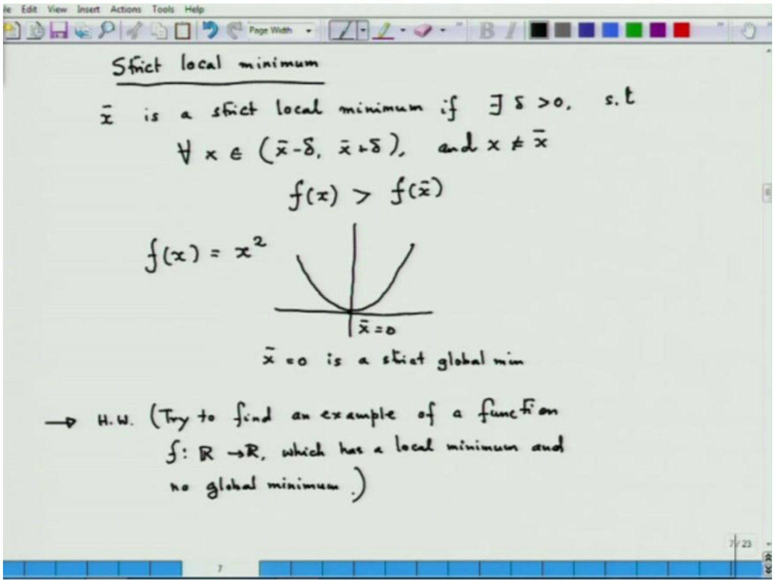
(Refer Slide Time: 33:53)



x bar is a global minimum, if f of x is bigger than equal to f of x bar for all x element of R. Let me tell you what can happen? Why we have also introduced this notion of strict global, strict local minima? I will tell you why such a notion is necessary? Such a notion comes out from this very little construction. So, consider a function like this. These are nice constructional max functions are quite important in optimization. But, we will not go to go into their details. So, basically you would have you put an x and find the maximum of these 3 numbers, and put that as your f x value, if you look into the geometric graph of that function.

So, then this is minus 1, this is 1 and this part so geometrically this is your graph. This is a graph of f. You look at a point x bar equal to 0. Now if I choose this delta is equal to half, so this will be minus half and this point would be plus half. This is mine plus half and this is minus half. So, for all x which is element of minus half plus half f of x is equal to 0 is equal to f of 0. So, this would imply by definition x bar equal to 0 is a local minimum. But, beware if you look at the function very carefully you realize that at the point x bar equal to 0 the function has a maximum value 0 because the function is negative throughout non positive throughout. So, but on the other hand, x bar is a global maximum. So, a global maximum can be a local minimum or a local minimum can be a global maximum. So, this is an anomaly which optimizers needed to remove. So, the question is how to remove this brings us to the notion of a strict local minimum.

(Refer Slide Time: 37:21)



This brings us to the notion of a strict local minimum. So, if I look into the notion of a strict local minimum, what it says? x bar is a strict local minimum, if there exits delta greater than 0 such that for all x element of and x naught equal to x bar means you forget that x bar and take anything else f of x must be strictly bigger than f of x bar. This fact has not taken place here, f of x is always equal to f x bar throughout the interval. So, this by doing this we remove the anomaly.

For example, if you take the function if you look at this function at x equal to 0. You look at the nature of the function is actually a strict global minimum in this case at x bar equal to 0 except x bar equal to 0. The function never takes the 0 value. It is always positive. So, here in this particular case, x bar equal to 0 is a strict global minimum. Of course, you can try to figure out. So, as a home work you might say home work try to find an example of a function f from R to R which has a local minimum but no global minimum.

So, tomorrow we will try to give you an example of such a function and then we will move into the more advanced case of speaking about functions from R n to R and how do we find global minima and these concepts global, local and strict local in that case. Of course, as a entertainment, we will give push on you another problem, Geometric problem of maxima and minima. And you will see how beautiful how minimal it is a geometry interacts with the notion of optimization. And with this I would like to really end my talk today.

Thank you very much for your attention.