Engineering Physics – II

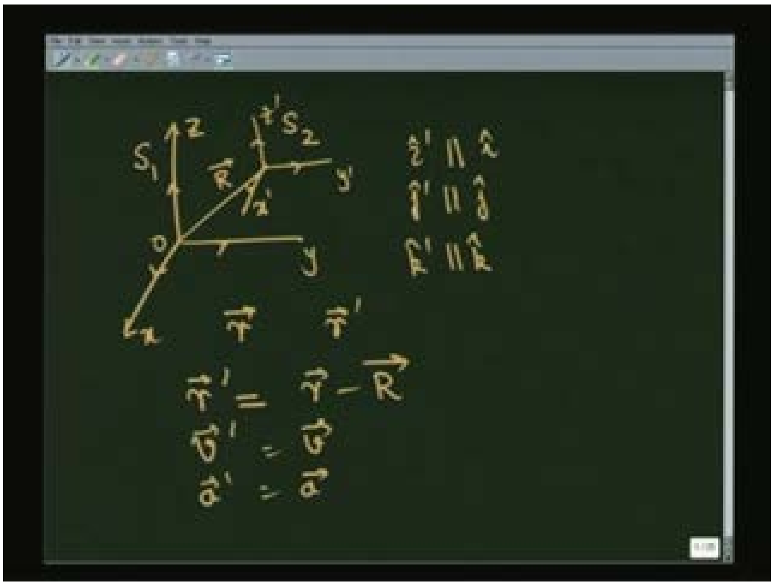
Prof. V. Ravishankar
Department of Basic Courses
Indian Institute of Technology, Kanpur

Module No. # 01
Lecture No. # 02

In the last lecture, we introduced the rectangular cartesian coordinate system, the cylindrical polar coordinate system and the spherical polar coordinate system. We were able to construct the line element, the area element and the volume element in all the three cases; but then it does not complete the preparation for us in order to launch into the study of electric and magnetic phenomena, because as we mentioned, we are still left with the freedom of choosing the origin of any of these coordinate systems or choosing the relative orientation as well.

I already discussed there at length in the previous lecture, so what I shall now do is to put those concepts on a formal footing footing, and see how we can actually write down transformation formulae from one coordinate system to another coordinate system. Thus, what we would have achieved is a dictionary that allows us to translate the observations made by one set of observers in a given coordinate system with to the observations made by another set of observers in another coordinate system.

(Refer Slide Time: 01:27)

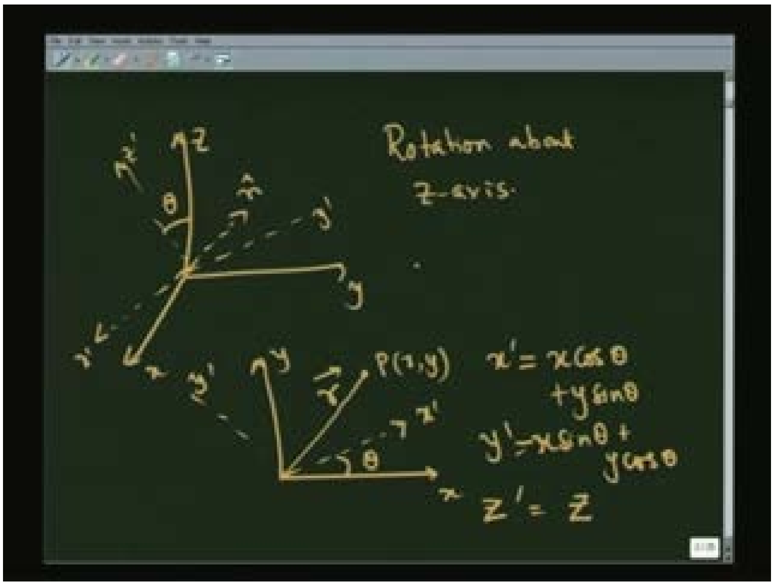


Let us start with the simplest of the situations, namely the translation of the coordinate system. So, I start with my rectangular cartesian coordinate system again; so, I have my x - axis, I have my y - axis, I have my z - axis; this is the standard configuration. Now what I do is to take the origin o of this coordinate system, and shift it by a certain distance, a certain vector R, and then erect a new coordinate system. So, if I have to denote them by x prime, y prime and z prime; I have the unit vectors or the basis vectors along these directions with our i prime, j prime and k prime; and I have i, j and k in the coordinate system - original coordinate system.

Now, when we say that we have translated, by that we mean that i prime is parallel to i; j prime is parallel to j; and k prime is parallel to k. In other words, all that we have done is to bodily move the coordinate system that we have not caused any rotation; we have not affected any rotation. This is so simple that, I do not even have to call it as a transformation formula, because let me say that my original coordinate system was labeled S 1, the new coordinate system is labeled S 2, the position vectors in my original coordinate system, if they are labeled as r; the position vector, the same position vector in the new coordinate system is labeled as r prime; then we have r prime, simply given by r minus r. So, this, is essentially the transformation from one coordinate system to another coordinate system; given this we immediately know, how to transform for example, velocity or acceleration or any another quantity, we know that v prime will be given by v, a prime will be given by a, so on and so forth.

This is only a warm up for us, in order to discuss a slightly more complicated example, and that is what we shall get into now, and that is the rotation. Of course, the more general transformation will involve both translation and rotation simultaneously; but we are not going to launch into that, although you can take it as an exercise and write it down later, we will consider rotation individually. Also for our purposes, it is very, very important to concentrate on the rotation, but that is because that is something that we will be doing routinely. So, let us see how to go on with the rotations.

(Refer Slide Time: 03:51)



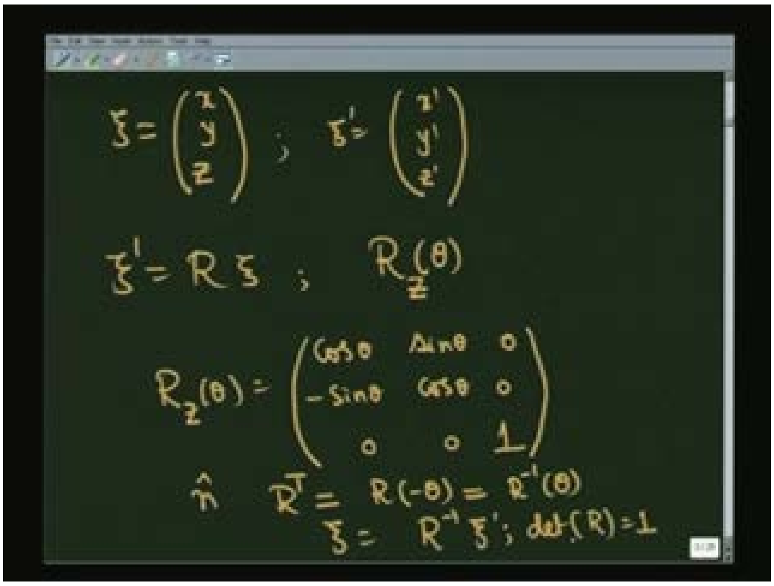
Again I have my the rectangular cartesian coordinate system x, y and z. Now I do not move my origin better rotate my original coordinate system to a new coordinate system, which I shall call as x prime, y prime and z prime. This is a rather pictorial representation, because we know that in general, a rotation is a rotation about any arbitrary axis, by any angle theta. So, there is some axis presumably let call it as n, I have rotated about this particular axis; and what I have done is to rotate by an angle theta. We do not wish to write down the most general form of the rotation that is the rather complicated form; and we won not we will not need it here. What we shall do is to consider for purposes of illustration, the simplest case namely rotation about the z - axis, rotation about z axis; equivalently, it is also rotation in the plane.

Now, our problem is considerably simplified, because I have x, I have y here, and then I rotate about the z axis. So, I get my x prime here, I get my y prime here, and I have rotated it by an angle theta, that is what I have done. So, once I rotate it, I want to know what the coordinates of a point p, which is denoted by (x,y) in the original coordinate system, you could call it x 1 if you feel like, would be into new coordinate system in terms of x prime and y prime.

I will not spend too much time, I would rather write down the answer, because we have other important things to go to; we can straightaway write down the transformation formula. And that is given by x prime equal to x cos theta plus y sin theta; y prime is equal to minus x sin theta plus y cos theta. Of course, although we have rotated at the plane, we are giving in a three-dimensional world therefore, I should complete my set of transformations by giving down what happens to z - axis also, the z coordinate also; so, we write z prime is equal to z. So, this is the transformation formula.

Of course, remember that when I speak of a point p, this is actually a vector, which I have denoted it by r. And what I am doing is to look at, a given physical vector, the position of a particle, the position of a charge for instance, in terms of two different coordinate systems. I have not altered my vector, it remains one and the same; all that I have done is to alter my coordinate system; the one and the same vector has two different coordinate descriptions; it is not as if I took the vector and rotated it by an angle theta. Now once we speak of vectors, not only should we speak of the coordinate system that we should also speak of coordinates, but we should also speak of the basis vectors. Before I get into that, let me write down this transformation in a slightly more convenient form, it is a rather compact notation; what do I do for that; what I do is to arrange the coordinates as a column vector.

(Refer Slide Time: 07:04)

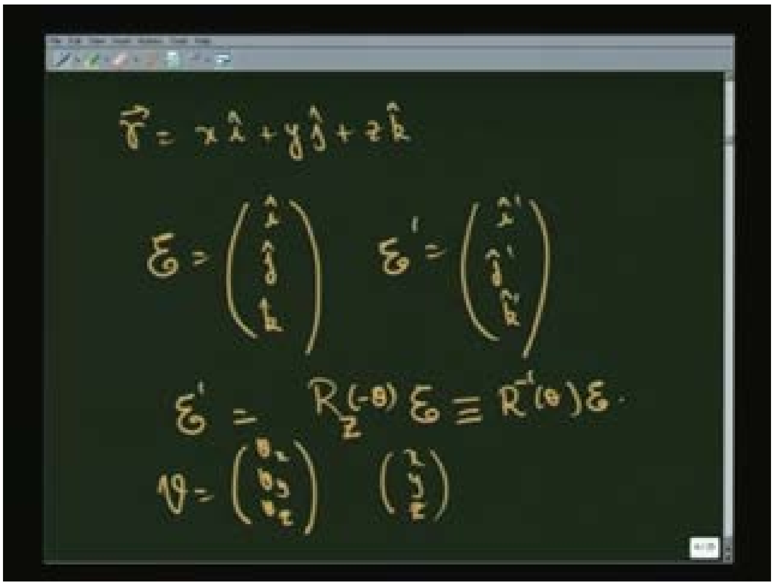


So, let me introduce a notation, let me call it as psi, and say this is your column vector with rows given by x, y and z. If I have to rotate the angle theta and if the same vector to be represented by coordinates x prime y prime and z prime; obviously, I would write psi prime is equal to x prime y prime z prime. Now the effect of rotation can be represented as a matrix operation on this column vector. So, I will now write psi prime equal to R psi; I could qualify my R even further, R standing for rotation. I am going to rotate about the z axis, while indicated by a subscript z here, I am rotating it by an angle theta, I will indicated by an argument as I have written now. And what is this given by R z theta is simply given by the 3 by 3 matrix cos theta sin theta 0 minus sin theta cos theta 0 0 0 1, this indeed is the so-called rotation matrix; when you rotate by an angle theta about the z axis, that is what we have done.

Now, even if you went to rotate about an arbitrary axis n without any loss of generality we could take that to be the z axis therefore, in principle, there should be a prescription of writing down the most general rotation matrix starting from this; we will not get into this, because that is not the purpose of our discussion here. However what we have to do is to observe certain nice properties of this rotation matrix, so what are those properties; let me list down here. If I look at the transpose of this matrix that is the same as R of minus theta; whether you interchange the rows and the column or you simply change theta to minus theta, it is one and the same, not only that this is also equal to the R inverse of theta, the inverse matrix which takes cube from x prime, y prime, z prime to x, y, z that is I can write psi is equal to R inverse psi prime.

The set of matrixes, which have this property are called as orthogonal matrixes; when the transpose is the same as the inverse, it is an orthogonal matrix. And actually what we have done is to observe a very, very important; in fact, the defining property of the rotation matrixes namely the transpose is the same as the inverse. What are the other properties that you can see about this particular matrix? Yet another property that you would see is that the determinant of this matrix is equal to 1, which I will write, determinant R equal to 1; in fact, together with orthogonality, this determinant tells you that you have ended up with a rotation matrix. What do I mean by that? Every rotation matrix satisfies the property that its inverse is the same as its transpose, not only that every rotation matrix satisfies the property determinant R equal to 1; conversely if any matrix has the property its transpose is the same as inverse and determinant is equal to 1, then it is certainly a rotation matrix; that is what we have. We are we might need it at a later step, so please remember these properties.

(Refer Slide Time: 10:42)



Now, I can actually write down, how the basis vectors transform, when I move from one coordinate system to another, where there are various ways of looking at it. One thing is to remember that the vector r, which was given by x i plus y j plus z k is in some sense quote unquote invariant, because x will go to x prime, y x will go to x prime, y will go to y prime, z will go to z prime, i will go to i prime, j will go to j prime, k will go k prime; but the position nothing has happened to the position vector or you can look at it geometrically as well in any case, what is the result that you will find.

Again what we shall do is to define two column vectors; now these column vectors do not have the coordinates as their entries, basic vectors themselves as their entries that I shall denote by curly e. This I will denote it as a column vector consisting of i j and k; obviously, if I have to rotate by theta, then I have my another column vector consisting of i prime, j prime and k prime. Now e prime will be related to e by a matrix R, but this rotation matrix is not the same as what we wrote originally that was R z of theta; in this case, it is R z of minus theta. In other words, the coordinates of a (( )) vector, and the basis vectors themselves transform in two different fashions; if that transforms according to R, this transforms according to R inverse, because we already absorbed that property R inverse of theta operating on e.

Well nobodies, no great surprise that the position vector has retained its absolute character, when I move from one coordinate system to another coordinate system. This is an important conceptual view point for us; in fact, it is a very important conceptual input for us, because normally we say a vector is a set of quantities, which transform in a particular manner under notations like what I have executed just now. When you make such a statement, we are actually referring only to the coordinates of the vector; if you simultaneously worry about the transformation of the basis factors also then of course, it has retained its absolute what nature.

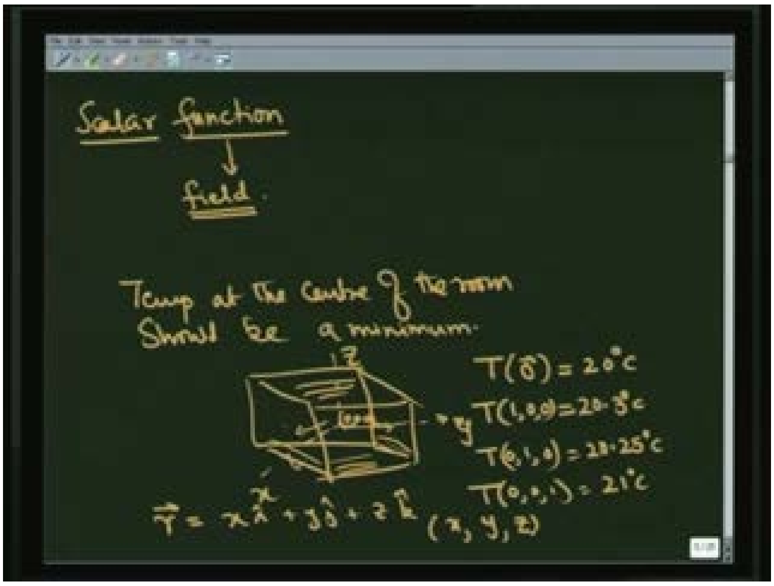
Now, whenever I speak of a vector, it is very clear that it is with respect to the set of transformations that I have in my mind, I have already introduced two different kinds of transformations; one of translation and another of rotation. Now for our purposes again, whenever I speak of a vector, I only have transformations with their, in terms of rotations, I will not worry about translations anymore. So, what is a vector now? Any set of three quantities, which transform like the position vector.

So, give me a quantity v in terms of three quantities v x, v y, v z. If you want, you can imagine them to be the components of the velocity vector. These transform exactly like the components x y z of my position vector; therefore, we say the velocity is also a vector. Similarly, acceleration is a vector, force is a vector, angular momentum is a vector, we have the electric field, which is a vector, we have the magnetic field, which is a vector. So, what do we mean by that? Every components of these fields in any coordinate system; rotate your coordinate system, then their coordinates in the new coordinate system does not depend on the nature of the quantity that you are considering, whether it is velocity or acceleration or magnetic field or electric field, it is simply dictated by this given matrix R or as universally on all of them. So, that is what we mean by a vector. So, that is something that we have to remember.

Now, why do we need this notion of a vector? The answer is simple, because as I already did just now, I have given innumerable examples of objects, which transform like vectors, except that there is a small complication; the complication is that at every point in space, there are many, many observations; for example, we can have something like energy at every point in space, energy density or we can have the momentum density; if there is a gas, there might be gas molecules in a small volume in space, which are moving in a particular direction or you can have pressure or temperature or magnetic field or electric field. Therefore we are not interested simply in three numbers v x, v y, v z, but we are actually interested in functions per set. We want to define scalars over the points in the real space; we want to define vectors over points in the real space and that is what launches on to what we call as vector algebra and vector analysis; in fact, vector analysis and that we shall do now.

Since, we have sort of endowed ourselves with all the basic properties; we are now in a position to actually make a quick review. Many of the concepts that I am going to discuss now, you are already familiar in your earlier courses for example, in your mechanics course, you already know that force can be written as the gradient of a potential and so, on and so forth. However for the sake of completeness and for the sake of setting of the notation and for the sake of causing no confusion, what we shall do is to quickly review these concepts, even if you are already familiar with them, because there is no harm in repeating these notions.

(Refer Slide Time: 16:11)



The simplest concept is of what is called as a Scalar function; it is worth spending some time on that, because once we understand this, it is very, very easy to have the notion of a vector field. In physics, we do not use the word Scalar function, but actually what we do is to replace it by the word field. So what do we mean by that? Instead of giving an abstract definition straightaway, let me start with a very, very simple example; and then go on to extract the definition out of this example, that would be a sensible way of doing things; so what shall we do now.

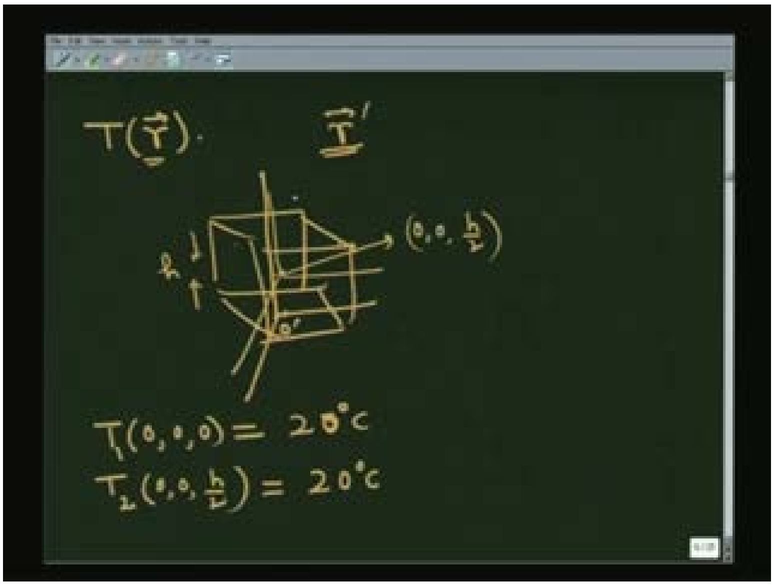
Let us look at the classroom or the room, in which you are sitting right now listening to this lecture; and let us imagine that a whole series of thermometers, which measure temperature at various points in the room. So, let us say that it is a nice bright clear sunny day; sun is shining equally on all the sides on the ceiling and on the four walls. So, what is it that you would expect the temperature to be? Well, if it is not too big a room, what we might say is that the temperature at the centre of the room, temperature at the centre of the room should be a minimum. The temperature at the centre of the room should be a minimum.

So, what do I do? Let me write a cube here; so, this is a kind of caricature of what we expect a room to be, this is the ceiling and here we have the floor, we have the ceiling here, and these are the four walls that we have. Now, if it is a real hot summer day and sun rays are shinning widely on all these objects. Let me look at the centre of the room and let me erect a coordinate system with that as the centre of the room; and then I have my x, let me call it as y; I have my x; and I have my z; and I want to measure temperature, as I move away from the centre of the room. As you approach for example, this wall or this wall or the wall, which is coming outside the plane or wall which is going inside the plane, you might expect the temperature to rise.

As to approach the ceiling, you would expect the temperature to rise, may be if you are approaching the floor the temperature would decrease; but whenever knows if there are bright windows for example, let somewhere at the lower point here, I am deliberately creating them in order to get a nice example for the purpose of our discussion; we can imagine even the floor is somewhat warmer. So, we can say, because the floor gets heated and it starts radiating. So, what we can say is that the temperature is at minimum at the centre of the room; and that is what I would call as the origin.

Now, what I would do is to take my thermometer and move around the all this all over the room; and then I would make a table; I would say T at the origin, let us say is given by 20 degree celsius. Now I want to move a point put a point nearby, so what shall I do? I will say T at the point (1,0,0); probably I am measuring the units of 1 feet 1 foot let us say or it could be even 6 inches, I would say this is given by some 20.3 degree Celsius; I would write that. There I would move in the other direction (0,1,0), and say this is 20.25 degree celsius; there I might move in the upper direction and I would say T (0,0,1) is 21 degree celsius, this is a gross exaggeration; but never mind, this is what we are doing. If you notice, I have actually smuggled in a notation, which should be clear to you, but we now let me make it explicit, I denote at the positions by (1,0,0), (0,1,0), (0,0,1) what did we mean by that; what I mean by that is that if I have a position vector r, and if I write x i plus y j plus z k and denoting it by a rho vector, not a column vector and indicate it by x, y and z; fine.

(Refer Slide Time: 20:24)



Let me again make a caricature of the room that I had, so so this is my room, ceiling floors, floor walls and so on and so forth. I had looked at the coordinate system, where the origin was centered, bank at the centre of the room. And then I had my z axis, I had my y axis, and I had my x axis. Now suppose I say no, no; I do not want to look at it from the centre of the room, but I want to look at it from the centre of the floor let us say, because that is where I am going to actually start, bringing my measuring scales, meter sticks whatever, whatever.

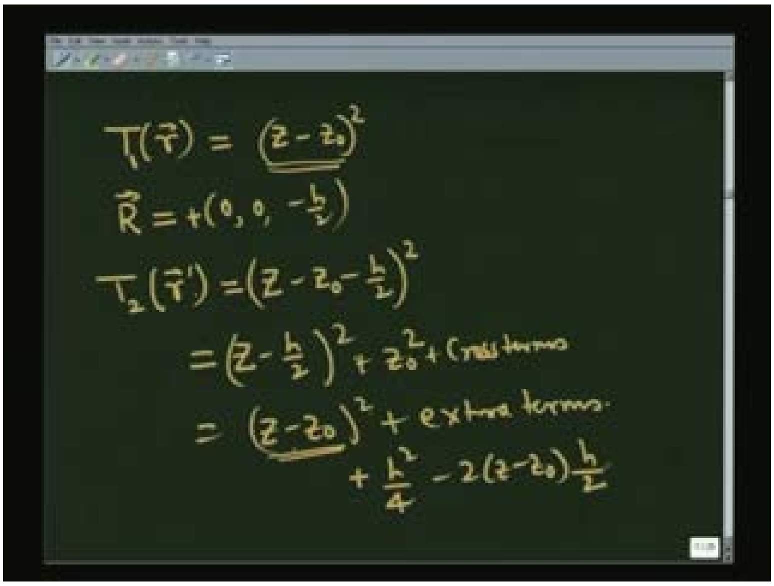
So, now what I will do is my y axis will be sitting here, my x axis will be sitting here, and my z axis is going to go apply this; this is my o prime; suppose I do that. Simply, because I chose my coordinate system, where the origin is centered at o prime on the centre of the floor, does not mean that the temperature at this point has changed. Irrespective of what you are going to do, whether you are going to choose the origin here, here, here, here, here centre of the floor at anywhere, the temperature is an invariant quantity. It is a property of the position of this point, relative to the walls, the ceilings and the floor of that room that is what we mean, when we say that there is a scalar; please remember that.

However what is it that is going to happen? Now when I do that what I called as r in my original coordinate system will be called as r prime in my new coordinate system. So, what in my original coordinate system I say, T (0,0,0) that was the origin was given by

20 degrees Celsius that is what I wrote. Now if the room has a height h, I moved on by a distance minus h by 2 in order to come to the centre of the floor; therefore, my origin, which was the original origin will have a coordinate 0, 0, h by 2. In the new coordinate system, it will have a coordinate 0,0,h by 2 my original coordinate system deeply denoted by T 1; in the new coordinate system T 2, the coordinates have changed that is 0,0, h by 2, but what is the value; it is still given by 20 degree Celsius.

Now, we say any physical absorbable, which has this property that when I shifted my coordinate system whether it is translation or rotation; for us rotation is important. Although I have illustrated it in terms of translation, the numerical value remains the same; then we say that that particular object is a scalar. Now, this scalar is defined all over the room and therefore, we say since, it is assuming continuous values as a function of the coordinates of the points in the room, we call such an object as a Scalar field. Now in order to make it more concrete, what I shall do is to give a definite functional form; this functional form has nothing much to do with whatever I have just now described, it is some functional form, which is easy for us to manipulate.

(Refer Slide Time: 23:22)

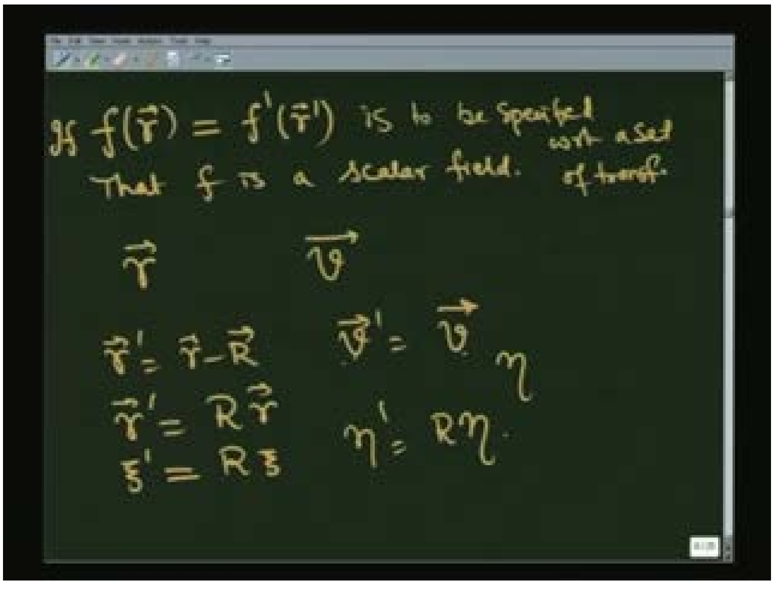


So, now let us here that my T of r in my original co-ordinate system, so I shall denote it by 1, was simply given by z minus z naught whole square, this is a bad example; but there is matter, actually it is not too bad example; what does it tell you? It tells you that the temperature is independent of x and y, it is a minimum at z equal to z naught, but z naught is actually the centre of the room, if you put z naught is equal to 0. And now what shall I do? I will shift my coordinate system, I shifted my coordinate system, how did I do that? I shifted it by a quantity R, which is simply given by 0, 0, minus h by 2, there is a plus sign here, that is what I did; if you remember the translation. Now what happens to this object in terms of the new description; well in the new coordinate system, I would write T 2 of r prime that object would be given by, now I have to be very, very careful, because z prime would be given by z minus h by 2; I will have z minus z by 2 minus h by 2 whole square, that is what I would have.

Now, suppose I were to open this up, expand this square, I would get z minus h by 2 whole square minus z naught plus z naught square plus cross terms or even better I can write it as z minus z naught whole square plus extra terms, this is a better form. Numerically speaking T 2 of r prime is exactly T 1 of r; what was being called as r is being called as r prime, but functionally speaking, the functional form has changed. In the original coordinate system, it had the functional form z minus z naught whole square; in the new coordinate system, it has the form z minus z naught whole square plus extra terms. Now I can write those extra terms if you feel like, that is simply given by plus h square by 4 minus 2 z minus z naught into h by 2, that is how it looks like.

Now, we are not going to be displayed by the difference in the functional forms, because we know both these functional forms represent one and the same functional dependence of the quantity, namely the dependence of temperature on various points in the room. So, in actually writing down this example, I have hit upon the exact definition of a scalar field. What is the definition of a Scalar field?

(Refer Slide Time: 26:08)



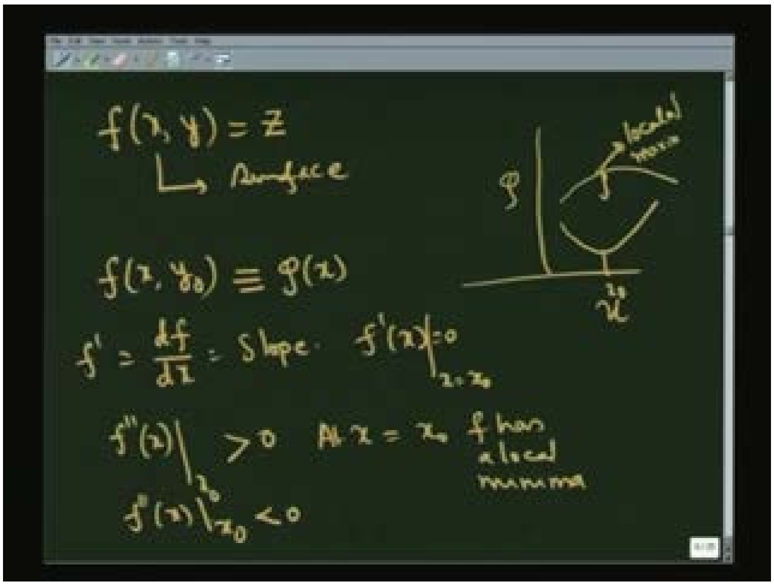
Let a function be defined over a certain region in space, if under rotations of translations it acquires the form f prime of r prime; the functional form changes, but numerically both the values are the same, then we say that f, this is scalar field. So, let me put the word f here. If f equal to f of r is equal to f prime of r prime, then f is a scalar field. It is not for us to decide what is a Scalar field, what is a vector field, what is a tensile field that is a luxury given to the mathematicians; for us, we have to demand that your physical absorbable actually has the property under these physical transformations. In order to exemplify this point, in order to emphasize this point, let me give another example; take the position vector r, and take the velocity vector v. Now under the translation we say r prime will be given by r minus R, but under translation by velocity v prime will be simply given by v.

Now, if you feel like although both of them are vectors, what is it that we had in mind when we said that both of them are vectors, we had in mind rotations as far as translations are concerned we find that the x component of v prime is the same as x component of v; y component of v prime is the same as y component of v; y z component of v prime is the same as z component of v; they actually act as scalars. So, if you remember this, then you would know that in all future lectures whenever I use the vector, I actually have rotation mind and not the translations in mind.

However if I were to rotate my coordinate system, now there is no ambiguity, because I would write r prime equal to R r, if you feel like I could be more precise and write psi prime equal to r psi. Now similarly if I were to arrange the components of velocity as a column vector, suppose I call it as a eta, then eta prime will be given by the same rotation matrix eta; under rotations, the position vector and the velocity vector had the same transformation property whereas, under translations, they do not have the same transformation properties; therefore, this statement f of r equal to f prime of R prime is to be specified specified with respect to a set of transformations.

In mechanics, in your earlier course, you had to worry not only about rotations, but also about Galilean transformation from one inertial frame to another; not only that you and I had to worry about entering rotating frames of preference, which are not inertial frames. If I were to do a complete course on electricity and magnetism, you would have worried about transformation from one inertial frame to another in terms of Lorentz transformation; but here since they are not going to bother, we can sort of rest contented by saying that we will be worried about only the rotations. So, all that discussion will be with respect to the rotations. So, now we have been able to define a scalar field, over the field, over the of real numbers; let us say over the configuration space or the physical space that is what we have done.

(Refer Slide Time: 29:48)



Now, we would like to further study the property of this field; in order to clarify what I mean by the property of the field, let me now specialize the general notion of a Scalar field to a simple case, namely my Scalar field is a function of (x,y). In fact, my temperature example was not a very good example, because it dependent only on z, but now let me make it as a function of x and y, if you feel like I can call it as a z (x,y). A very, very simple example would be is that you bring a whole lot of children let us say, and you fill them your classroom with them; each child is going to stand there, they are closely part; and you can ask what is the height of each child as I move from point to point. So, every square or every slab in your on your floor can represent a point some kind of an average.

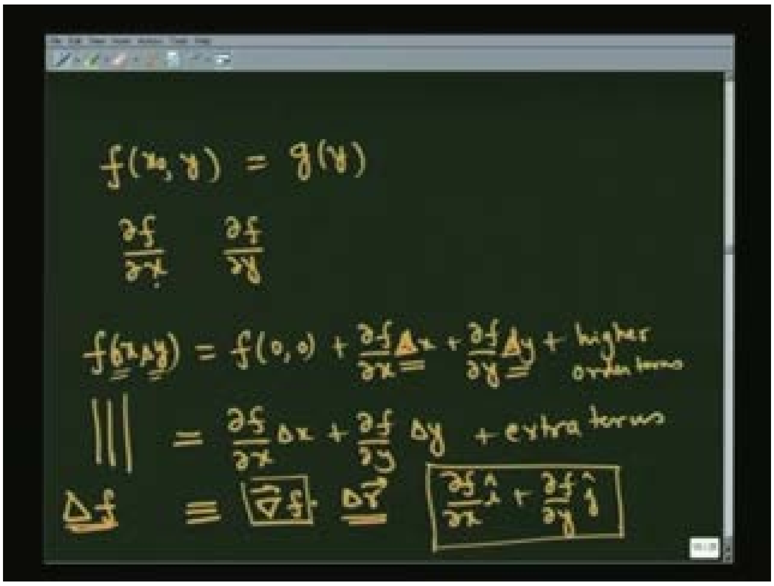
So, the height is represented as a function of (x,y); now that means, if I were to place a whole lot of children in your classroom, and if I were to look at their height, therefore form a surface and that surface is define over the x y plain. So, this is a surface; of course, if I write a Scalar form of x, y, z that would be a high per surface in a higher dimensional space, let me not worry about that at this point. Now, whenever we are given a surface of this particular kind, a very, very natural question to ask is where is the surface peak, where is the surface a minimum, where does the surface change very rapidly, where is the surface discontinuous is the surface defined everywhere, these are the question that we are going to ask; and we need some kind of a procedure, which allows us to discuss these notions.

Now, imagine that I freeze the value of y, and I look at the variation of f of x at a given value of y naught; I have held my y naught fixed, at a given value of y naught, I keep on only changing my x coordinate. Now it is as good as a function of a single variable, which I will denote it as a phi of x. Well this is a curl, defined over the x axis; once I give you a curl defined over the x axis, there is no great deal about discussing maxima minima extreme and things like that, because what you would do is to look at the quantity d f by d x, which is this slope. And you say that whenever this is 0 that is if x prime is 0 at some x naught, x equal to x naught, you say that is an external point, because at that particular point either your function is a maxima or a minimal or it could even be some kind of a inflection point.

In order to further characterize the nature of the extreme, you would set the second derivative. So, you now calculate the second derivative at the same point, and now if this function is greater than 0, you say f has a local minima at that particular point. So, I say at x equal to x naught, f has local minima; how do I illustrate that. So, I say that if this is my x axis, and this is my f; I have changed my notation to phi it may really does not matter phi at x equal to x naught, it looks like this, there is a minima sitting here.

On the other hand, suppose my f double prime x, which is same as phi double prime x; at x naught is less than 0 that is curvature is negative let me denote it here. So, this is x naught, at the while x naught if it is 0, then the curve would behave like this, and this point is what would be called as a local maxima. So, we know that the minima or the maxima depend on the curvature properties of this function; of course, if the second derivative also is 0, then we cannot say much we will have to go onto higher derivatives and that is something that we have studied from the calculus of a single variable. Now, your surface is a more complicated object, because it might so happen that if I had frozen the value of x and change the value of y let us say.

(Refer Slide Time: 33:46)



So, what I now do is I look at f of x naught y, this will be some other function, which I will denoted as g of y. This g of y may exhibit a different nature altogether; there is no reason why g of y should be following phi of x. So, what do I mean to say, what I mean to say is that in the x y plane, the function might behave differently as I keep on moving in different directions. If I move along the x direction, the function might increase; if I move along the y direction, the function might decrease; if you move on some other direction the function might not change; therefore I need a procedure to keep track of how the function changes as I move to a neighborhood point.

The procedure to do that is not calculating the derivatives that we have done, but what we should do is to calculate the partial derivatives. So, what shall we do? We shall calculate delta f by delta x and delta f by delta y. So, if you were to calculate these two points the derivatives, then this derivative will tell you how the function behaves independently, when I am changing x, keeping the value of y fixed that is what the meaning of this partial derivative is; and here I vary y keeping the value of x fixed.

Now, what is the geometric significance of the partial derivative; in order to see the geometric significance, what we shall say is that I have a good, smooth, nice differentiable continuous function; and let me say, I want to do a Taylor expansion around the point. So, what shall I do? I have a function f y f (x,y) and let us say I want to make a Taylor expansion of this function around the point x is equal to y equal to 0. So, what would I do? I would say that this is given by f (0,0) plus delta f by delta x d x plus delta f by delta y d y plus higher order terms.

Now, I have to be careful; I should not write the use the notation d. So, let me use the notation delta. So, what do I mean by x and y actually I mean, f delta x f delta y that is what I have done. Given the value of the function at x equal to 0, y equal to 0, I am giving you the value of the function at delta x and delta y. Now, stare at the right hand side of your equation, if you look at the right hand side of your equation, delta x and delta y are arbitrary displacements; they are infinitesimal displacements, but they are arbitrary otherwise, along the x and the y direction; f delta x delta y is the value of the Scalar function at the point 0 plus delta x and 0 plus delta y.

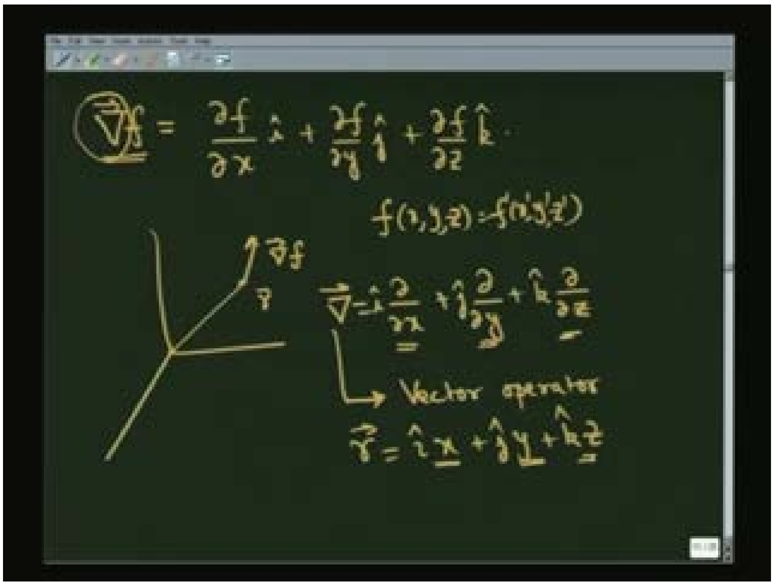
So, if I want to do, I can write it symbolically as delta f by delta x into delta x plus delta f by delta y delta y, this is what I am going to concentrate plus extra terms; let me leave them alone; this object I can write symbolically as gradient f dot delta r. In making this expansion, I may not have to specify for you, how I chose my x coordinate, how I chose my y coordinate; I can rotate my coordinate system for that matter I can even translate my coordinate system. If I want to translate what I called at 0 0 would be called as x naught and y naught; if I rotated, it will become delta x prime and delta y prime, whatever it might be. So, we have found that the change in the value of the function, that

I shall denote it by delta y f; so, because identically equal to f of delta x delta y is simply given by gradient f dot delta r.

So, what did I do? I introduced a notation. The notation that I am introducing is that I arrange these two partial derivatives as a vector, delta f by delta x i plus delta f by delta y j. When we looked at these partial derivatives, we have no clue that we are actually constructing the components of a vector. But now let me repeat the argument so that there is no confusion about it; and what is the argument; delta f is a scalar; it is a scalar defined at the point delta x delta y, but other hand delta r is an arbitrary infinitesimal displacement, this is a vector and the only way I can construct a scalar given a vector is to take the dot product or if you feel like the inner product with some other vector; therefore this object better transform like the component of a vector; and therefore, this notation is a legitimate notation. So, take the partial derivatives along individual directions, along the i direction, along the j direction, along the k direction so on and so forth.

Then we have actually constructed this vector called delta f by delta x i plus delta f by delta y j, which is a great significant improvement, you never understand it, this started with a very nice simple object like a scalar field something like an energy density. Now, we have ended up with a more complicated entity like the vector entity. And this vector entity is indeed something of some importance to us. So, we have succeeded in constructing a vector function, starting from its scalar function, which I shall call as a vector field. I will give you a quote unquote rigorous definition at a later time.

(Refer Slide Time: 39:32)



Right now, we have succeeded in constructing this function, namely the gradient, which I shall now write in its full detail with all the coordinates put namely delta f by delta x i plus delta f delta y j plus delta f by delta z k. Now obviously, it specifies a direction at every point in the space. So, I have my co-ordinate system this is my position vector r, there is a function defined, and now I am associating the direction with a certain magnitude and that is what I am denoting it by gradient f. So, we have to know what the physical significance of the magnitude of this gradient f and what the physical significance of the direction of this gradient f is; if you have done that then we know the complete geometric significance of this function.

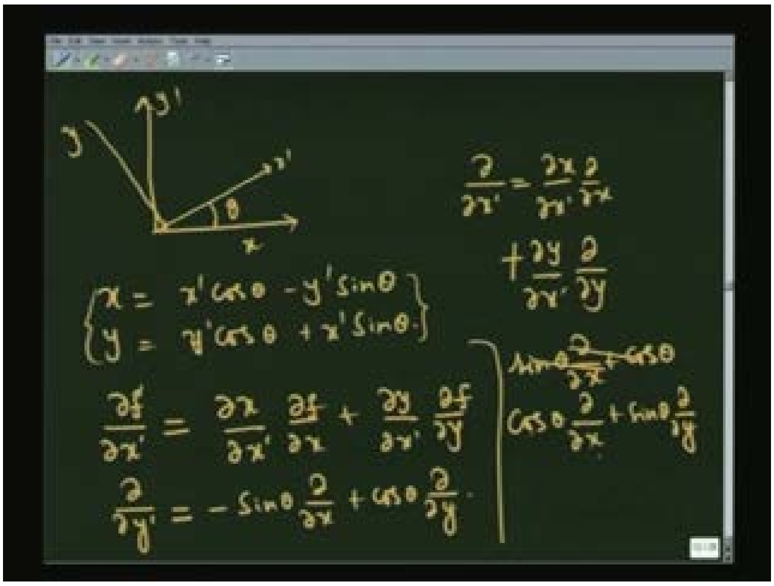
Obviously, it should have something to do with, how the function increases or decreases or changes or retains its shape as we keep on moving on different points on the space; but before I go on to do that, it is necessary for us to confirm one thing namely that this new definition, namely the gradient f confirms to our earlier notion of a vector, what is it that I said, I said that any set of three quantities, which transform like the coordinate vectors under rotations are what will be called as a vector. So, that is what I would like to erect it.

Now, what are the inputs that I have to use; the first input that I have to use is f (x,y,z) is given by f prime of x prime, y prime, z prime that is what I need to use. The second input that I need to use is the transformation property of these partial derivatives. Now, this is a very, very important notion. And let us pause to spend a little time. What is it that is happening? Look at this notation gradient f, now when I am looking at the gradient f, I am going to produce a vector on the right hand side, and the fact that the vector on the right hand side is produced is independent of the nature of f; start with energy density, start with temperature, start with pressures start with anything; the minute these derivatives operate d by d x i d by d y j d by d z k are actually land up with a vector. So, there is something that I would like to emphasize therefore, let us give an independent existence to these operators themselves these differential operators.

So, let us denote the operator by delta by delta x along the i direction, along the j direction I have delta by delta y and along the z direction, I have by delta by delta z. This is unlike anything that we have seen so far, because this is not a function. Whenever I spoke of the transformation properties what is it that I had in my mind? Set of three numbers; and these three numbers, which I have defined all over this space, they constitute a function. Right now I do not have a number, but I have a set of three operators; therefore this gradient is what one would call as a vector operator meaning if this operates on any scalar function that produces a usual vector function or a vector field.

Now, what is the great difference between this operator and the position vector that I wrote; i x plus j y plus k z. If then it is indeed a vector operator, then these derivative operators, these differential operators d by d x d by d y d by d z better have the same transformation properties as x, y and z and that is something that you should be able to verify. If you did that then the concept, the idea that this is a vector operative as shown, not only from the physical view point, but also from the view point of mathematics; we are doing something consistently. So, in order to do that, I need the transformation properties, so what do I do? Let me write down the coordinates x, y, z in terms of x prime, y prime and z prime.

(Refer Slide Time: 44:06)



As usual, I will not consider the most general rotation, but I will consider the simplest of rotations namely around this z - axis. So, I have x, y, x prime, y prime and my theta here. So, how is my x related to x prime and y prime? x is simply given by x prime cos theta minus y prime sin theta, y is given by y prime cos theta plus x prime sin theta. Given this functional dependence, I can actually express the derivatives with respect to x, y, z in terms of derivatives with respect to x prime, y prime and z prime. What is it that I want to do? Look at the partial derivative delta by delta x prime; what is it that I have in my mind? It is going to act on a function. So, how is it going to act on a function, if I had rotated my coordinate system, I say that this is the same as delta x by delta x prime delta by delta x plus delta y by delta x prime delta by delta y.

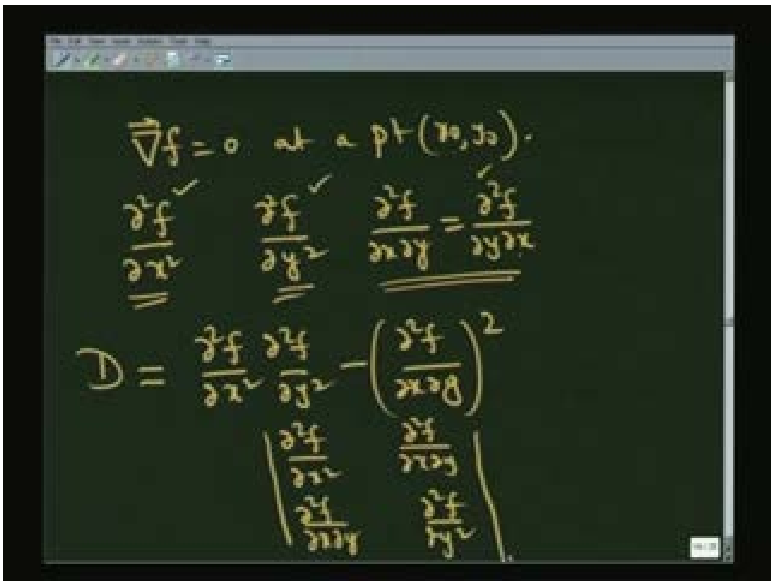
So, if there is a function f sitting here, the same function f will sit here, it has the same numerical value in both the coordinate systems, because it is a scalar field. So, what is the great difference, here I am going to write delta of f prime of x prime, y prime, z prime, but here I will write f that is understood. So, here it will be f prime, here it is f. So, once you are given that we can forget all about this f. So, let me not worry about that and let me write down the original equation, what was the original equation; delta by delta x prime is given by delta x by delta x prime d by d x plus delta y by delta x prime d by d y.

Now, I know x as a function of x prime, I know y as a function of x prime therefore, I can immediately rid it off from this formulas that I have written. So, let me write it down here, there is still lot of space; delta x by delta x prime is nothing but cos theta, delta y by delta x prime is nothing but sin theta; therefore, I have sin theta delta by delta x plus cos theta, little bit of mistake; so, I am sorry this should be cos theta delta by delta x plus sin theta delta by delta y; that is what we wrote. If you remember, we have written the transformation formula for x prime as x prime equal to cos theta x plus sin theta y. Now, my delta by delta x prime is cos theta delta by delta x plus sin theta by delta delta y.

In a similar manner, you can sit down with your paper and pencil, and verify easily that my delta by delta y prime will be given by minus sin theta delta by delta x plus cos theta delta by delta y, which is exactly the way my y coordinate transform; y prime is minus sin theta x plus cos theta y. In other words the differential operators delta x and delta y transform exactly like the coordinates x and y therefore, they qualify to be called as vectors; except that they are not vector functions, but they are what we call as vector operators.

So, we have to remember that if you are going to take the dot product of this vector operator with some other function, you have to be careful, because whether it acts from the right or the left, it will have two different meanings, we will not get into it right now, but there is a caution, which is well exercised and mentioned at this particular point. So, now let we have demonstrated the nice transformation properties of my derivative operator; in fact, you people should amuse yourself by writing down the most general rotation matrix on x, y, z and convincing yourselves that the derivative operators transform in this particular fashion.

(Refer Slide Time: 48:21)



Now, we can launch on to get all that the geometric interpretation of this gradient operator; and that is not a very difficult thing, because we already said that the change in the value of my scalar field is simply given by gradient f dot delta r; that is what we wrote; delta r are infinitesimal displacements. Now, suppose I freeze the value of my delta r, whatever the infinitesimal value is; and then change only the orientation of delta r. So, what is it that I mean, I take a point x naught, y naught and sweep all over this sphere with a radius, which is given by delta r. So, delta r is nothing but modulus of delta r, that is what I am saying; when does the right hand side become a maximum? Let us ask that question; the right hand side becomes a maximum, when that gradient f and delta r are parallel to each other, that is when it becomes a maximum; when does it attain a minimum value, it attains a minimum value, when gradient f is anti parallel to delta r.

And of course, whenever delta r is perpendicular to be gradient of f the direction, then there is no change in the value of f that is what this fellow says. In other words gradient f if I look at it as a vector, it gives you the direction, in which my function changes maximally; that is what it gives you. Now, once you give me this piece of information, module gradient of course, is the magnitude by with it changes the direction is simply given by gradient f divided by mod gradient f. So, suppose I denote it by some unit vector n, then n gives the direction, in which the function changes maximally.

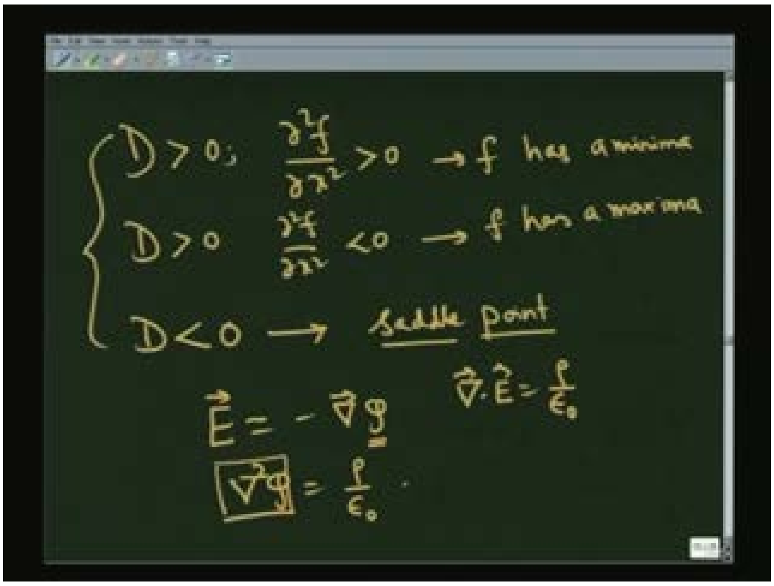
So, if I give you a hill, the hill keeps on winding up, and then you will reach the peak, and then there is a valley, there is a hill, so you imagine a range like this. If I am sitting here, the gradient here is given in this direction; the gradient at this point in given in this direction, the gradient at this point is given in this direction. Of course you also have maxima and minima, the gradient at this point is 0, the gradient at this point is 0, the gradient at this point is 0, except that what I have written here is a rather poor description of what a two-dimensional surface like a hill is, because there are more complications than what this curve shows; and that is something that we shall examine now.

In order to do that, let us look at a surface again; and let us say that gradient f is equal to 0 at a point, at a point x naught y naught. A nice question that you would like to ask is, if the function minimal or maximal at that particular point, well the answer to that depends on a higher order derivative that is what I said, when I looked at the calculus of a single variable, I said look at d square f by d x square. But now, I have two variables and there are three different kinds of second order derivative that I can construct. So, what are those objects, I can construct del square f by del x square; I can construct del square f by del y square; and I can construct del square f by del x del y. Please remember that del square f by del x del y is the same as del square f by del y del x. It does not matter whether you take the partial derivative with respect to x first and y later or y first and x later; therefore we have three species of derivatives.

So obviously, the precise behavior of the function at this point x naught y naught would depend on all the three quantities. I would like to state a result without generally proving it; that is something that you would have learnt in your mathematics course, and that is the following. In order to discuss the properties, it is convenient to define an object called D, which is called as a discriminate; and that object is devoted to be d square f by d x square d square f by d y square minus del square f over del x del y whole square.

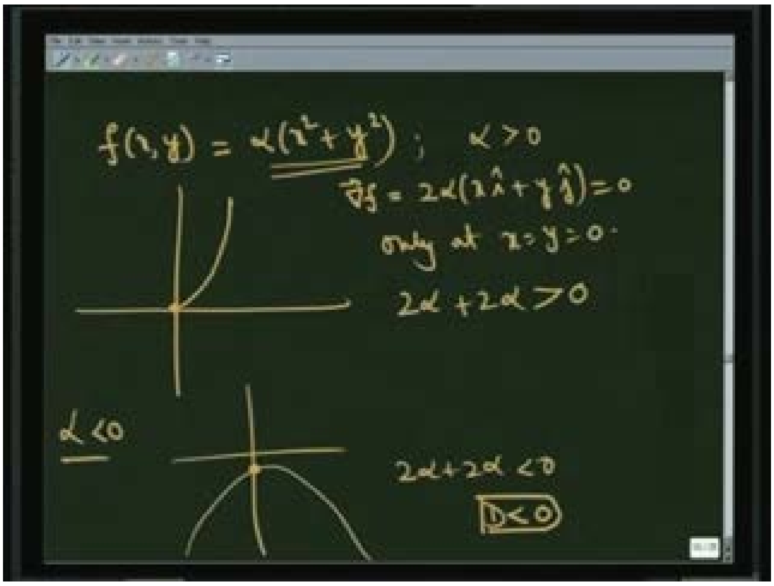
If you feel like this is actually the determinant of the matrix, which we write d square f by d x square d square f by d y square, then I write a d squared f d x d y here and d square f d x d y here. Clearly otherwise it is complicated, because of the occurrence of these so-called cross derivative terms. However it turns out that if d greater than 0, here I should go make it more carefully in the next page.

(Refer Slide Time: 53:15)



So, suppose D greater than 0, and then I have D square f by D x square greater than 0, then the function f has a minimal. Suppose D greater than 0, D square f by D x square less than 0, then f has a maxima of course, this is not evaluated at any point, that it is evaluated at the point x naught y naught where your gradient vanished. Suppose D less than 0, what is it that you are going to get? You are going to get a very new kind of a functional behavior called as the saddle point. The geometric meaning of this object is not at all tough, because the first two are very easy; so let me take some very, very simple examples and illustrate them.

(Refer Slide Time: 54:21)



The simplest example confirms a conics so, what we shall do is to consider a very, very simple function, which I shall denote as f (x,y) is equal to alpha into x square minus y square; alpha tells you how fast the function increases, essentially it is a measure of this scope, so it is not very important for us. Now, if I look at this particular surface, it is going to give you properties, which tell you what the saddle points are. So, instead what I shall do is to look at an even simpler surface, which is given by f (x,y) is equal to alpha x square plus y square.

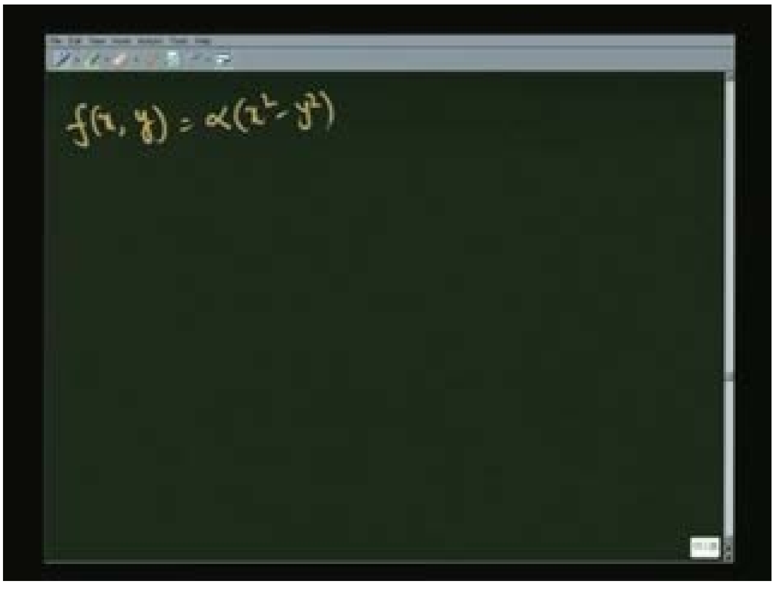
What is this function? If I hold the value of y fixed it describes a parabola along the x direction, so if y is fixed, it describes a parabola in the x direction, and then a similar manner if I were to hold the value of x fixed, it would again describe a parabola therefore, alpha of x square plus y square is a parabolic, that is what we are going to get. So, if this is the z axis, and if I imagine that this line is the section of the plane, what you do is to draw a parabola, and what I did about the z axis to get the parabola. From this geometry, it is clear that the only external point is at x equal to y equal to 0.

In order to verify that explicitly, let us calculate the gradient; you can see that gradient f is given by 2 alpha x into I hat plus y j hat and this is equal to 0 only at the origin x is equal to y equal to 0. Now, in order to further verify that the origin is indeed minima, all that we need to do is to calculate the discriminated; but however, since the cross derivative terms vanishes, I only have to calculate the second derivative of x square, and the second derivative of y square and evaluate it at the origin. The second derivative for both of them is simply given by 2 alpha, which is greater than 0, thus by discriminate at that point is greater than 0, individually both the second derivatives of the origin are greater than 0 therefore, we conclude this is indeed a maxima.

On the other hand, if I take alpha to be less than 0, I had taken alpha to be greater than 0 here, if I take alpha to be less than 0 here, then my parabola would be generated from a parabola, which is written in this fashion. Obviously, this would be a maxima no surprise about that, because although the discriminate is greater than 0, my alpha is less than 0 therefore, 2 alpha is less than 0. Thus we have been able to generate two very, very simple functions; one of each has a global minima at the origin and it is negative, which has a global maxima at the origin.

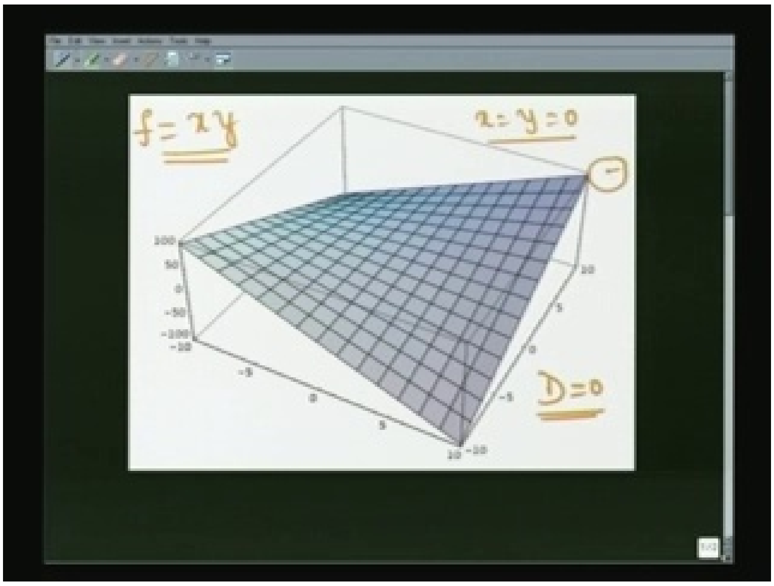
However, these are not the only examples or functions that we can have; whenever you look at a surface, it may so happen that I move along the x - axis the function increases, that I move along the y axis that the function decreases, whenever we encounter such a situation, which is what we denoted we discussed in the case D less than 0, will be called as a saddle point.

(Refer Slide Time: 57:46)



In order to illustrate a saddle point, a very, very good example is given by yet another conic, which is f (x,y) equal to x square minus y square. Now, this function is something that cannot be shown or drawn so easily or even can be picturize by rotating it therefore, what we shall do is to show you a nice computer generated picture and let us look at that.

(Refer Slide Time: 58:09)



So, this is the function, which has been generated; it is a very, very simple function as I told you, which is simply given by x square minus y square; I believe that alpha has been set to equal to 1. Now you can see that if I were to move along the x direction, the function is going to increase; in fact, it is a section of a parabola. But if I move along the perpendicular direction namely along the y direction, the function is going to decrease.

So, we say that the point x is equal to y equal to 0, because that is the only point at which the gradient is going to vanish, it is a saddle point. So, what is the definition of a saddle point, take the external point, in this case x is equal to y equal to 0, and draw a small region around that particular point. When you draw the small region around the point, however, small this region may be there will be some points at which the function will be decreasing from the original point; whereas if I move in some other direction, the value of the function will be increasing; this is not an unfamiliar surface for us, because this is the famous horse saddle, which you will see the riders use when they want to ride on a horse and this is regard the required property for this stability, that is a stability of a different kind not in the sense of mathematical stability.

In fact, yet another simple figure that shows the saddle point behavior is this function, this surface has been generated by looking at a even simpler function f equal to x y. I invite you people to convince yourself that only x equal to y equal to 0 is the external point, that is that is the point at which the gradient vanishes, and again it is a saddle point as you can see from this figure, the behavior along this direction is simply different from the behavior along this direction. And the again we find that if you draw a sufficiently small circle around the origin, depending on which direction you move, the function will either increase or decrease, but there is no circle; however, small the radius may be such that the function will only either increase or it will only decrease.

This should conclude for us, the study of the properties of the gradient, what kind of a information it gives us, when we look at it in terms of the… From the view point of maxima, minima so on and so forth; except that that I did not discuss yet another case namely, discriminates is equal to 0, D equal to 0. Whenever d equal to 0, unfortunately there is no way for us to do, except that we have to start computing higher alternate derivatives, and see how the function behaves. But there is something that is not very important to us, what we shall do is to now take the q from the fact that gradient f is a vector field or a vector function, generalize the notion to discuss any arbitrary vector field and do not go on to discuss other operations such as curl and the divergence, which we shall take up in the next lecture.