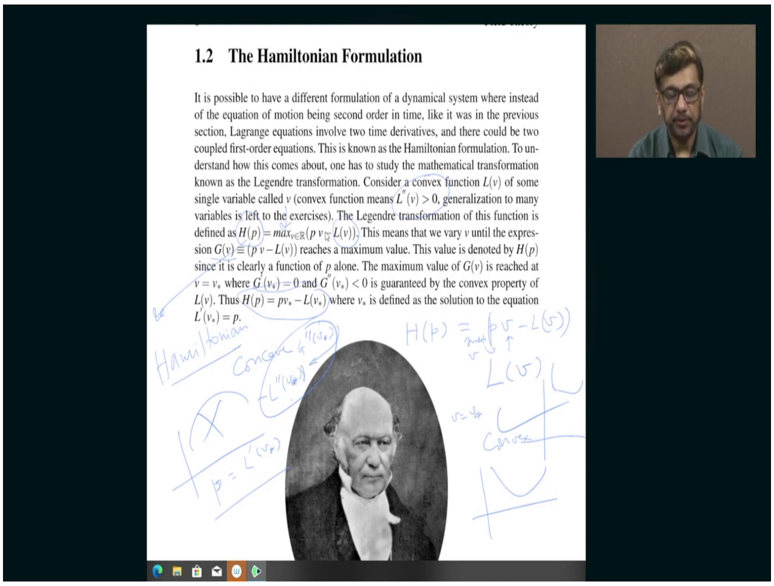
Dynamics of Classical and Quantum Fields: An Introduction
Prof. Girish S. Setlur
Department of Physics
Indian Institute of Technology, Guwahati

Review of point particle mechanics
Lecture - 03
Hamiltonian Mechanics

(Refer Slide Time: 00:32)



So, in today’s lecture I am going to discuss an alternative to the Lagrangian approach which we discussed in the earlier lecture. So, if you recall the Lagrangian approach to classical mechanics involves rewriting the vectors form of the Newton’s second law in terms of quantities which are purely scalars and these are called Lagrangian’s and the Lagrange the Lagrangian obeys a certain set of equations which are called the Euler Lagrange equations.

Now, the main advantage of the Lagrange equations over Newton’s laws is that the constraint forces in Lagrange formalism do not have to be explicitly specified because that is typically how most interesting problems in classical mechanics are formulated by just specifying that the system obeys certain constraints or it is a constraint to move in a certain way rather than specifying the forces that compel those particles to move in that fashion.

So, whereas, Newton's second law forces you to know what those forces are that compel those particles to move in that fashion. So, that is the main advantage of using the Lagrange formalism. And it so happens that there is an equivalent formalism which is typically more useful for especially for generalizing to quantum mechanics, but also within classical mechanics itself it has many advantages in the sense that it allows you to study symmetries in a more convincing and transparent manner.

So, that is the formulation of the Hamiltonian mechanics the Hamilton's formulation of classical mechanics which I am going to discuss. So, superficially the distinction between the two can be captured by the following assertion namely that the Lagrange equations the Euler-Lagrange equations of classical mechanics are second order in other words that if you write them down the it will involve 2 time derivatives of the generalized coordinates just like Newton's second law does.

So, if you recall Newton's second law is mass times acceleration is force. So, what is acceleration? Acceleration is nothing but the second time derivative of the position coordinate, but basically Euler–Lagrange equation pretty much says the same thing except it says that something analogous to mass times the second derivative of the generalized coordinate equals generalized force.

So, it is basically a curvilinear analog of Newton's second law, except that those generalized coordinates already obey constraints so that the generalized forces do not involve forces of constraint. So, that is the main advantage. Otherwise the analogy between Newton's second law and the Euler–Lagrange equations is pretty much one to one, ok.

However, in contrast the Hamilton's formulation of classical mechanics recasts these equations in terms of two first order equations. So, in other words instead of having one second order equation you write them as two first order equations which of course, effectively is especially if the coupled which they in fact, r in this case these two first order equations is basically equivalent to one second order equation.

So, I am going to describe to you the relation between Hamilton's formulation of classical mechanics and the Lagrangian mechanics and they are related by this beautiful mathematical notion called the Legendre transformation. So, I am going to describe what it is and I will tell you that it has a very beautiful and intuitively appealing geometrical meaning. So, the Legendre transformation can be visualized in a very geometrical way and that is quite appealing and interesting to know ok.

So, in order to develop the Hamilton's approach to classical mechanics I am going to introduce to you the concept of Legendre transformation. So, I hope you can see my slides. So, if you can make out that I have written down that imagine that there is a function L(v) is a function of some variable called v_n. So, the idea is that I am going to introduce a notion called convex function.

So, a convex function is basically something that looks like this. I mean basically anything that looks like this ah. So, this would be convex ok. So, rather than this would be concave. So, specifically the mathematically what it means is that you are saying the second derivative of the function is positive at all points. So, the slope continuously increases as v increases the slope increases rather than decreases ok. So, then you say it is convex.

So, if you have such a convex function one can define formally another function which is basically the maximum value of, so, the right now this definition is not very intuitive, but I am going to tell you that it has a very well defined geometrical meaning, but analytically you can define H(p) as the maximum value of p times p. So, p is a fixed quantity and v is your variable ok. So, v is your variable and I maximize with respect to v.

So, I change v until this becomes maximum and the reason why that is guaranteed to exist is because if you look at the first derivative. So, when does an extremum exist? The extremum exists at $v = v^*$ where the derivative of this vanishes. So, in other words that is what I have I have told you that. So, if you think of G of v as this difference, then the place at which the function becomes an extremum so, extremum recall for a single variable is either maximum or minimum.

So, at this stage equating the first derivative to 0 just guarantees that point is an extremum it does not guarantee that it is a minimum, but right now it guarantees that it is an extremum. So, in other words it is either a minimum or a maximum. So, now, what is the value of v and clearly that value is obtained by the solution of this equation. So, by inverting this correspondence you will be successful in finding the extremum.

But, then I want to convince you that extremum is in fact, the maximum, right. So, because it is supposed to be the maximum value so, why is that the maximum? Because if you take the second derivative. So, if you take the second derivative of G it is simply $-L''(v_*)$ is not it because p times v second derivative is 0 because p times v is linear in b.

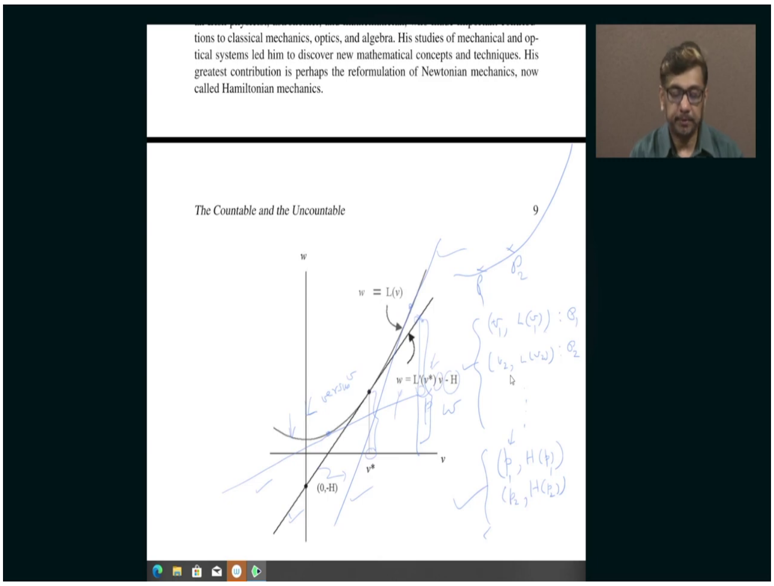
So, the first derivative is constant the second derivative is 0. So, the only term that contributes to the second derivative is the one with the Lagrangian. So, in other words $G''(v_*)$ is basically equal to this ok. So, the bottom line is that because the function is convex because it is convex. So, you know it is not this rather it is this. So, because it is convex L'' is positive ok.

So, because L'' is positive, G'' is negative and you know when second derivative is negative the function is a maximum, right. If it is an extremum and the second derivative is negative, the function is a maximum. So, that makes sense. So, the H(p)'s definition therefore, makes sense because there does indeed exist a maximum. So, we are not making a mistake we are not assuming more than we know.

So, we have to assume that there is a maximum only when we know beforehand that there in fact, is a maximum we just convinced ourselves that there is a maximum ok um. So, now, that is the definition therefore, of H(p) is that. So, $H(p) = pv_* - L(v_*)$ and v_* is the place where G(v) becomes a maximum ok. So, now, so, now, this is the analytical description. So, this is called the Hamiltonian ok.

So, this is called the Hamiltonian just like this is called the Lagrangian. So, they are named after different scientists, but the question is that you know the claim that I am going to make is the following.

(Refer Slide Time: 10:58)



That imagine think of L(v) as this sort of curve. So, this is L versus v. So, it is some convex curve is not it. So, it is a convex curve. So, it means it points upward like that. So, and then the thing is that in order to specify this convex curve you can do the traditional obvious thing namely you specify the x coordinate of some point that is the horizontal coordinate of some point and its vertical coordinate.

So, by specifying the horizontal and the vertical coordinates that is the abscissa and the ordinates you will be successful in locating all the points on the curve L(v). So, that is the usual traditional way of drawing a curve you just locate the horizontal and the vertical coordinates of each point and then you draw points at those places and then you join them all using a smooth curve.

So, that is how you would think of L(v) in the traditional way, but then there is another way of thinking about a convex curve like this and the way to think about it is not by specifying the points themselves like we did just now, but by specifying the set of all straight lines that are tangent to this curve. So, I am going to convince you that specifying a whole bunch of straight lines each of which are tangent to the curve at different different points is equivalent to specifying the curve itself.

So, why is that? So, the reason why that is the case is because so, just imagine that this is one such straight line which is clearly tangent to the curve at this point. Now, you can imagine another straight line like this which is tangent to the curve at this point. So, then you can imagine another straight line which is tangent to the curve at this point.

So, if I specify so, if I collect all these straight lines, right I collect all these straight lines and then I put them in one place and then I just collect is. So, just like in order to draw the curve I am collecting all the bunch of points, right. So, basically in order to specify the curve the usual way of doing it is to collect all those bunch of points and put them together in some place and then deal with them.

But, instead of doing that what I am doing is I am going to collect a bunch of straight lines and then I am going to put them all together in a certain way. And, the way to do that is, so, first draw those straight lines on a piece of paper and then ask yourself which is that single curve which I can draw that is tangent to all these straight lines. So, if you are able to draw such a curve then basically that curve is exactly what your you are going to describe using those points.

So, instead of so, what you have done is you have been successful in replacing that point description of the curve using the tangent description. So, in other words rather than specifying the points on the curve what you have done is you have replaced those points by a bunch of tangents and what you are saying now is that the curve that I am looking for is that curve which is simultaneously tangent to all these straight lines at whatever points they want to be tangent at.

So, now, I am going to convince you that these two ways of looking at the curve are in fact, equivalent ok. So, how do I convince you about that? So, that is exactly what the that is the geometrical interpretation of the Hamiltonian ok. So, now, imagine that there is some v^*. So, imagine this is my v^* and then there is a at this point there is a straight line which is tangent to this curve.

So, now, I am going to ask myself what are the slope so obviously, a straight line is described by two numbers – one is the slope of the straight line and the other is the y intercept; that means, the vertical intercept and the slope. So, that is what if you recall from our high school days that is how we have defined or choose to describe straight lines, is not it? So, if we if I wanted to analytically describe a straight line I would describe it by specifying the slope and y intercept.

So, now, I am going to ask myself what is the slope of this straight line that is tangent to this L(v) clearly that slope is nothing but $L^{'}(v)$ which happens to be exactly p. So, remember that $L^{'}(v*)$ is nothing but p. So, this is so, if I specify the slope of the straight line and that is bound to be equal to by definition the derivative of this L(v) at v* .

So, now, I am going to ask myself what is the y intercept? So, I am going to call the y intercept the point at which the straight line intersects the vertical axis at (0,-H) . So, when I do that clearly the equation for the straight line now becomes. So, this is the w is the some point on the straight line that is this is my w. So, w is nothing but the slope $y = mx + c$. So, this is y and this is my slope which is m which happens I have called it p, but.

So, if you used to $y = mx + c$ that this would be x, this would be m, this would be c and this would be y. So, that is what that is. So, it is a straight line whose slope is nothing but $L^{'}(v) = p$ and the y intercept is - H. So, that is my straight line. So, the claim is that rather than specifying v and L, v what I am going to do is ah.

So, if I specify a whole bunch of these things, right so, then I am of course, going to specify the curve. So, this is a certain point this is a certain different point and so on and so forth. So, by specifying all these points I can clearly draw a smooth curve through that and that should be my L of v, but the claim here is that alternatively instead of doing this instead of specifying (v,L(v)) what I can do is I can specify (p,H(p)) which is equivalent.

And, what is p? It is the slope of the straight line that is tangent to this curve that I am eventually going to generate. So, there is a slope. So, rather than specifying v and L(v) so, I am going to specify the slope of a whole bunch of straight lines and those straight lines are all going to be simultaneously tangent to the curve that I have in mind. So, this is the slope and then -H is pretty much the y intercept of that straight line. So, I am specifying slope and intercept.

So, for there are there is one straight line whose slope is p_1 and whose y intercept is $-H(p_1)$. So, by specifying both that straight line now becomes unique because there is only one straight line with slope p_1 and y intercept $-H(p_1)$. So, similarly I can generate different straight line whose slope is p_2 and the y intercept is $-H(p_2)$. So, now, I have in this way of doing things I am generating a whole bunch of straight lines.

Now, I am going to ask myself which is that curve that is tangent to all these straight lines at the same time in other words, but they are going to be tangent at different points, but then they are simultaneously tangent to all these straight lines and the answer is precisely this the curve that you generate through this procedure. So, that is because the Hs and the ps have been generated through the Ls if you recall. So, this is the geometrical meaning of Legendre transformation.

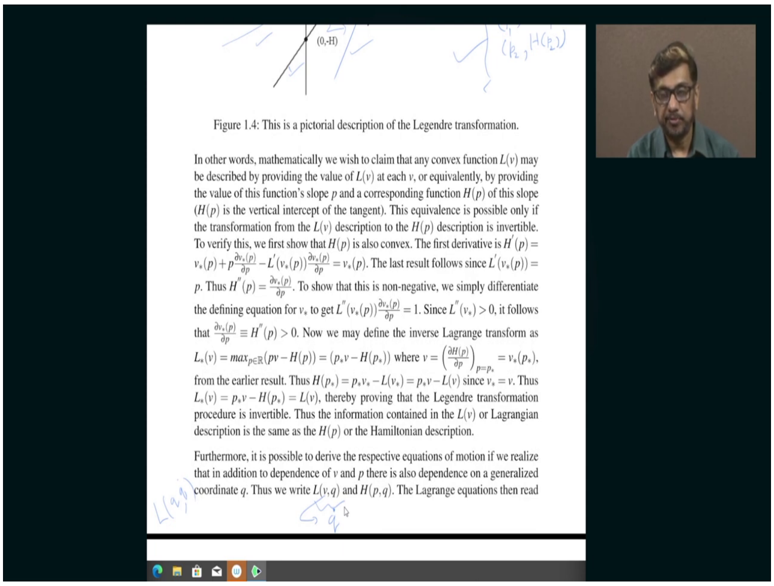
So, the transformation that we are talking about now is the transformation from the (v, L (v)) language to the (p,H(p)) language. So, from the language of specifying points on the curve in the usual way to the language of specifying the curve indirectly by specifying the slope and the y intercept of the vertical intercept of the tangents to the curve that those straight lines eventually imply.

So, in other words, there is going to be a curve which is finally, going to be tangent to all those straight lines at the same time. So, you either specify the curve directly by specifying v and L(v) or you specify this whole bunch of straight lines and then there is always going to be a unique curve that is going to be tangent to all these straight lines at various points, ok. So, this is the Hamiltonian formulation of.

So, of course, you might think that where is physics here? So, the physics comes if I interpret L (v), L as my Lagrangian and v as my generalized velocity or the so, if v is L (v) then clearly $L^{'}(v)$ is basically the usual way of defining canonical momentum and so, H(p) will then become the Hamiltonian. So, that is called the Hamiltonian.

So, this is how Hamiltonians are formally introduced into classical mechanics and this is the geometrical interpretation ok.

(Refer Slide Time: 22:30)



So, in my textbook I have made some other technical statements about invertibility that you have to convince yourself that this correspondence is invertible. So, that so, in other words, if you go from the (v,L(v)) language to (p,H(p)) language and then you should be able to do the reverse. So, if I start from the (p, H(p)) language I should be able to get back the (v, L(v)) description as well.

So, I have intuitively convinced you that is in fact, possible, but there is also a mathematical rigorous way of doing that which is described in my book here which I will encourage you to read, because the textbook that I am using right now which is being displayed in front of you is the prescribed textbook.

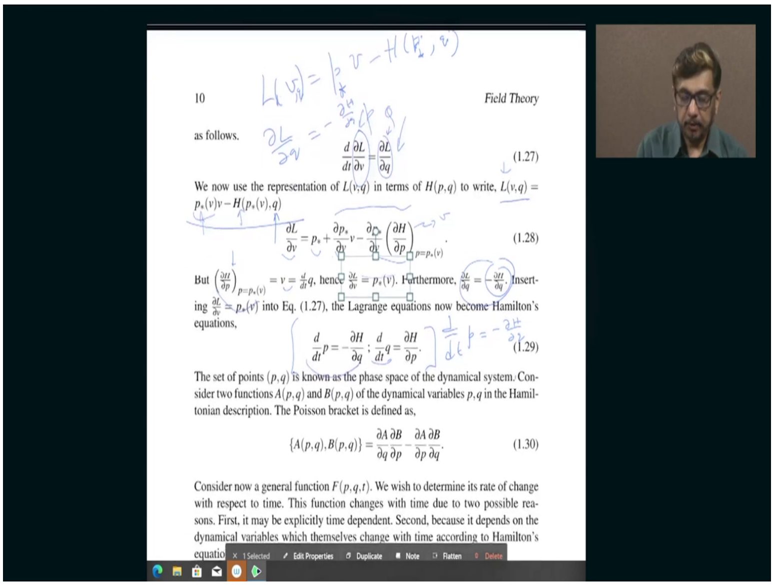
So, you please consult the relevant page number 9 of my textbook has this mathematical proof which tells you why it is that it is this correspondence is invertible ok. So, remember that L(v) implies that the Lagrangian is a function of generalized velocity. I have purposely suppressed another independent variable namely the generalized rather the position itself.

So, you see the so, you know that the Lagrangian is actually not just a function of $\dot{q}$ which is your generalized velocity it is also a function of q. So, but because q does not play any role in this geometrical description of the Legendre transformation I purposely suppress that.

But, now I can go ahead and bring it back. So, if I bring it back my Lagrangian is going to be not only a function of the generalized velocity which is v which is of course, $\dot{q}$, but it is also a function of yeah. So, it is also a function of the generalized coordinate itself which is q.

So, as a result of course, that same dependence carries over into the Hamiltonian as well because after all that q was anyway suppressed earlier. So, it should once you explicitly display it should also be displayed also inside the Hamiltonian bracket. So, the Hamiltonian is not only a function of p which is the slope of L versus v, but it is also of course a function of the generalized coordinates. So, I write H(p,q) to emphasize this point.

(Refer Slide Time: 25:19)



So, now, recall that the Lagrange equations were basically this. So, this is like I told you know this has a very close resemblance to Newton's second law because what this is generalized momentum and this is generalized force. So, rate of change of generalized $\frac{d}{dt}(p) = F$. So, $\frac{d}{dt}$ of generalized momentum is generalized force.

So, its pretty much Newton second law in disguise and, but of course, the disguise is a very efficient disguise in the sense that it gets rid of aspects that you are anyway not privy to namely the forces of constraints, alright. So, now, what I am going to do is that I know how H is expressed in terms of L. So, I am going to invert that and express L in terms of H.

So, that is easy to do because all I have to do is invert this and if I do that. So, I am going to use this inversion to see what it is if I can rewrite the Lagrange equations in terms of my newly generated Hamiltonian. So, how would I do that? I define L(v) like this and then the derivative of L(v) with respect to v which is the v is if you remember is the generalized velocity which is $\dot{q}$ and that is going to be nothing but . So, remember that this p^* ok.

So, what is p^*? p^* is basically ok I think I should not have skipped this, but bottom line is that just like you can define the Hamiltonian as the maximum value of p (v - L). Similarly, you can define the Lagrangian as the maximum value by varying p as p (v - H(p). So, its completely invertible. These relations are completely invertible and the maximum takes place exactly at some $p = p^*$ which is basically given by the solution to the equation $\frac{d}{dp}(H(p))$ ok.

So, yeah so, this is this page number 9 is not fully cosmetic I mean its not some pedantic discussion of invertibility because I am going to use this idea in the very next page. So, I feel its worthwhile for you to go through this carefully. So, bottom line is that just like there was a v_* which was basically the value of v at which that function G became an extremum here too there is a p_* which makes the analogous function an extremum that is p(v - H(p)) now becomes an extremum as you vary p.

So, if you keep that in mind then obviously, this p_* is now going to be a function of the v that because so, clearly what is how do you define p_*, p_* is determined indirectly through this. So, if you invert this whatever I have circled here. So, you will get p_* in terms of v so, which is why I have written p_* as a function of v there.

So, now, I take the derivative of L with respect to v and I end up getting p_* and then I should not forget to differentiate p_* because now that also is a function of v and I end up getting this relation. But then keep in mind that $\frac{d}{dH}(p) = v$. So, because this is nothing but v these two will cancel out and then it will give me this ok.

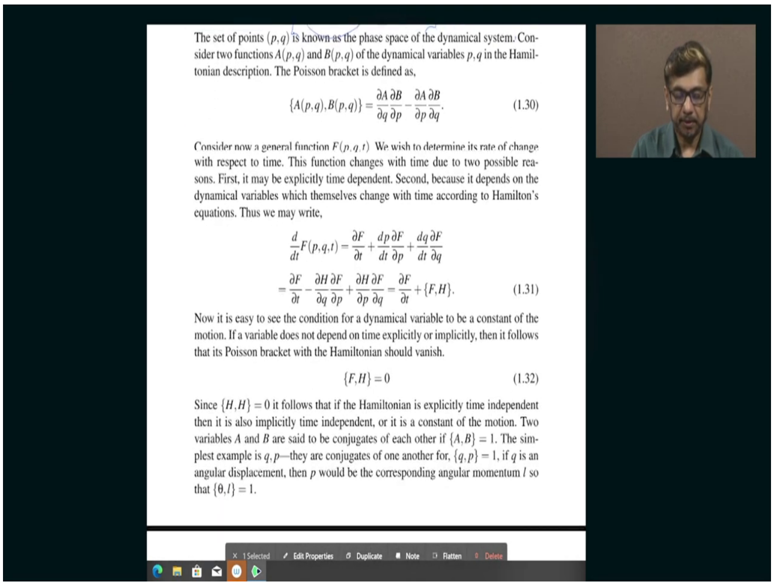
So, $\frac{dL}{dv} = p_*$ ok and furthermore it is obvious that $\frac{dL}{dq} = -\frac{dH}{dq}$ and why is that is fairly obvious because you see it is L is defined L is defined as p_*v. So, H is a function of q, right. So, the q dependence now is here ok. Sorry, I am messing up. So, I am just going to this one.

So, L of the $L(v) = p_*v - H(p_*)$ which happens to depend on q. So, now, if I take the derivative with respect to q, so, $\frac{dL}{dq} = -\frac{dH}{dq}$ ok yeah. So, just take dL by dq here you will see that q is only sitting here nowhere else. So, then that is what you get here ok.

So, as a result what will be successful in doing is that so, this equation is nothing but $\frac{d}{dt}(p) = \frac{dL}{dq}$, but then what is $\frac{dL}{dq} = -\frac{dH}{dq}$. So, this is the first equation first of the Hamiltonian equations. Then of course, dH by dp itself is nothing but v right. So, but then what is v it is nothing but q dot.

So, what is happening now is that you have instead of one second order equation, you have two first order equations involving p and q. So, in the Lagrangian formalism there was one second order equation involving only q which is the generalized coordinate. Now, you have two first order equations involving generalized momentum as well as generalized coordinate ok.

(Refer Slide Time: 32:38)



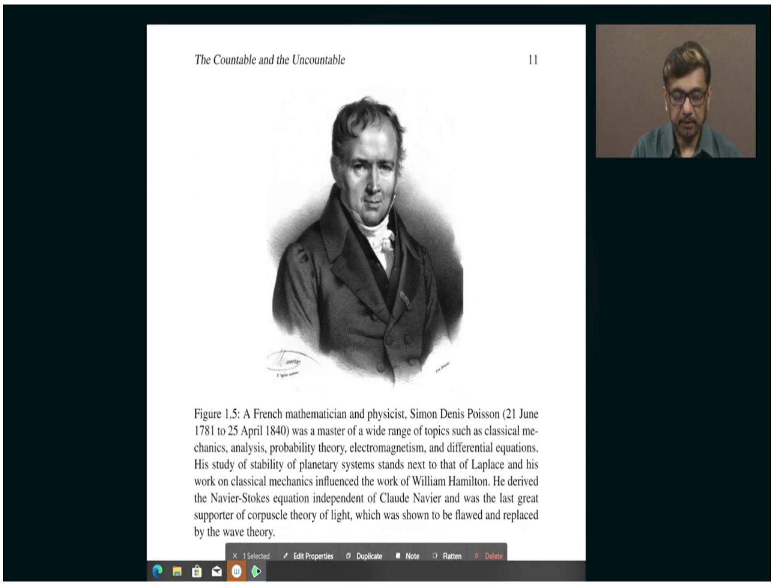
So, this is the so called Hamilton's approach to classical mechanics, but notice that in both the Lagrangian approach to classical mechanics as well as the Hamilton's approach to classical mechanics the forces of constraints are explicitly omitted. So, they are superior to Newton's second law of for studying classical systems for this reason. So, namely that you do not have to know beforehand what are the forces of constraints acting on the system.

So, even without knowing that you will be able to find the trajectory of all the particles in the system which is basically the fundamental question that is of interest in classical mechanics is to just explicitly work out the trajectory of each of the particles knowing the initial state of the system.

So, the answer to that question is facilitated by both the Lagrangian as well as the Hamiltonian approach because both these approaches do not explicitly require you to know the forces of constraints whereas, Newton's second law requires you to know what the forces of constraint are ok. So, I am going to stop here and in the next class I am going to discuss what are called flows and symmetries.

So, the Hamilton's description of classical systems enables a very elegant description of symmetries and in fact, the Hamilton's equations themselves describe a kind of flow, but the flow is with respect to time. So, the so, these ps and qs are flowing with respect to time, but then the independent variable can be something other than time which enables us to study certain kinds of symmetries called dynamical symmetries in a very elegant way.

(Refer Slide Time: 34:58)



So, I am going to describe that in the next class and I hope you will join me for that.

Thank you.