EULER-LAGRANGE EQUATIONS FOR PARTICLE AND FIELD THEORIES

Link to: physicspages home page.

To leave a comment or report an error, please use the auxiliary blog.

Post date: 29 Dec 2020.

It's important to understand the distinction between a *particle* theory and a *field* theory. To see how this works, we'll start by looking at classical theories of particles and fields using the Lagrangian formalism.

1. Classical particle theory

We'll look first at the Euler-Lagrange equations for a system of classical particles. Suppose we have N particles in 3-d space, for a total of $3N$ degrees of freedom. We define the Lagrangian as

$$L(q_i,\dot{q}_i) \equiv T(\dot{q}_i) - V(q_i) \tag{1}$$

where T is the kinetic energy (that depends only on velocities $\dot{q}_i$) and V is the potential energy (that depends only on positions q_i). We want to find the path followed by the system between times t_1 and t_2, that is, we want to find $q_i(t)$ between those times, subject to the constraint that $q_i(t_1)$ and $q_i(t_2)$ are fixed at some known values. In general, there are an infinite number of paths the system *could* take between these two times, and each path is specified by choosing the functions $q_i(t)$ (which in turn determines $\dot{q}_i(t)$). Each choice of path gives a different form for the Lagrangian.

The principle of least action states that the action S, defined as a functional of the paths that can be followed, is an extremum (in practice, almost always a minimum, hence the principle of *least* action). The action is defined as

$$S \equiv \int_{t_1}^{t_2} L(q_i,\dot{q}_i)\;dt \tag{2}$$

The condition that S be an extremum is specified by requiring $\delta S = 0$, which means that if the particular Lagrangian L_0 gives a minimum (we'll assume the extremum is always a minimum from here on to avoid confusion), then any slight variation of the paths that make up L_0 increases S. Thus the condition $\delta S = 0$ is just an extension of the usual condition that the first derivative of an ordinary function be zero in order for that function to have a minimum.

To calculate δS, we need to vary the paths slightly. Using the chain rule (actually, we should justify that the chain rule works when calculating variations in functions, but we'll trust the mathematicians on this point) we get

$$\begin{aligned}
\delta S &= \delta\left[\int_{t_1}^{t_2} L\,dt\right] && (3)\\
&= \int_{t_1}^{t_2} \delta L\,dt && (4)\\
&= \int_{t_1}^{t_2}\left(\frac{\partial L}{\partial q_i}\delta q_i+\frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i\right)dt && (5)\\
&= \int_{t_1}^{t_2}\left(\frac{\partial L}{\partial q_i}\delta q_i+\frac{\partial L}{\partial \dot{q}_i}\frac{d\left(\delta q_i\right)}{dt}\right)dt && (6)
\end{aligned}$$

where the repeated index i is summed.

We can now integrate the second term on the RHS of 6 by parts to get

$$\delta S=\int_{t_1}^{t_2}\frac{\partial L}{\partial q_i}\delta q_i dt+\left.\frac{\partial L}{\partial \dot{q}_i}\delta q_i\right|_{t_1}^{t_2}-\int_{t_1}^{t_2}\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\delta q_i dt \qquad (7)$$

The requirement that $q_i\left(t_1\right)$ and $q_i\left(t_2\right)$ are fixed means that $\delta q_i=0$ at the limits of integration, so the middle term is zero. We're then left with

$$\delta S=\int_{t_1}^{t_2}\left[\frac{\partial L}{\partial q_i}-\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\right]\delta q_i dt=0 \qquad (8)$$

This must be true for all possible variations δq_i so the quantity in brackets must be zero, which gives us the Euler-Lagrange equations:

$$\boxed{\frac{\partial L}{\partial q_i}-\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)=0} \qquad (9)$$

It's worth digressing at this point to explain why we can take q_i and $\dot{q}_i$ as independent variables. It would seem that they are *not* independent since once you've specified q_i you can get $\dot{q}_i$ by just taking the derivative. The point is that when we specify the Lagrangian L, we don't know what $q_i\left(t\right)$ is; all we have is the function L which depends on both q_i and $\dot{q}_i$. The goal of minimizing the action is to find the curves $q_i\left(t\right)$ such that these curves together with their derivatives minimize the integral of $L\left(q_i,\dot{q}_i\right)$. The physics comes in specifying the Lagrangian; the mathematics then allows us to determine the paths $q_i\left(t\right)$ followed by the particles. In principle, we can specify L to be any old function of q_i and $\dot{q}_i$, but once we've done this, the form of L is fixed and we can then solve the Euler-Lagrange equations to find the particle paths. In other words, the Euler-Lagrange equations specify the q_i so that the q_i together with their derivatives $\dot{q}_i$ minimize the action S.

In an alternative universe, we could conceive of a Lagrangian that depended on q_i, $\dot{q}_i$ and $\ddot{q}_i$, say. In that case all of q_i, $\dot{q}_i$ and $\ddot{q}_i$ would be independent variables in the derivation above, and we'd end up with a different form of the Euler-Lagrange equations. The fact that physical Lagrangians depend only on q_i, and $\dot{q}_i$ is a consequence of Newton's second law $F = ma$, since this allows only the positions and velocities to be specified as independent variables. The law $F = ma$ is, when written out, a second order differential equation, since the acceleration is the second derivative of the position with respect to time. The general solution of a second order differential equation allows initial or boundary conditions to be specified for the solution function and its first derivative.

All this is fine, except how do we know that these equations, when solved for $q_i(t)$, actually do give the path followed by the system? The key is to look back at the definition of L in 1. Then

$$\begin{aligned}
\frac{\partial L}{\partial q_i} &= -\frac{\partial V}{\partial q_i} && (10)\\
\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) &= \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right) && (11)\\
&= \frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_i}\sum_j \frac{1}{2} m_j \dot{q}_j^2\right) && (12)\\
&= \frac{d}{dt}\left(m_i \dot{q}_i\right) \quad \text{(no sum over } i\text{)} && (13)\\
&= \dot{p}_i && (14)
\end{aligned}$$

where p_i is the momentum of degree of freedom i. Therefore, the Euler-Lagrange equations are equivalent to

$$\dot{p}_i = -\frac{\partial V}{\partial q_i} = F_i \tag{15}$$

where F_i is the force acting on degree of freedom i. This is just Newton's second law, so the Euler-Lagrange formulation is indeed equivalent to Newton's laws.

2. CLASSICAL FIELD THEORY

The main difference between particle theory and field theory is that the variables q_i no longer describe the motion of anything, that is, they are no longer functions of time. Rather, they become fixed labels for points in space. The position variables q_i become independent variables in the same way that the time t is independent. Taken together, they label points in spacetime.

A *field* is some quantity that has a value for each point in spacetime, and it is this quantity that can change as we move from place to place or forward in time. For a scalar field such as temperature or density of a substance, the field consists of a single value $\phi(q^\mu)$ attached to each point in spacetime, where we now use the notation q^μ to represent the space components together with time. (That is, q^μ is a four-vector in special relativity, with $q^0 = t$, $q^1 = x$ and so on.) A vector field, such as the electric field **E**, is actually composed of three separate fields, one for each spatial coordinate. Each of these fields again has a single value for each point q^μ.

To work out the Euler-Lagrange equations for classical field theory, we need to think about what is meant by a 'path' that the system follows. Because the spacetime coordinates q^μ are no longer dynamical variables, it doesn't make sense to ask how q^μ changes with time. What *does* change is the value of the field ϕ (or ϕ^r if we have more than one field, as with the electric field, in which case the index r ranges over all the fields), so it is the field ϕ that is the dynamical variable. As such, the path followed is determined by a function of the field values. By analogy with the Lagrangian in the particle case, we define the *Lagrangian density* $\mathcal{L}\left(\phi^r, \phi^r_{,\mu}, q^\mu\right)$. The notation $\phi^r_{,\mu}$ is defined as

$$\phi^r_{,\mu} \equiv \frac{\partial \phi^r}{\partial q^\mu} \tag{16}$$

The Lagrangian density is the Lagrangian per unit volume, and each infinitesimal volume element $d^3x = dq^1 dq^2 dq^3$ follows a path through time, so the action element of this volume element between times t_1 and t_2 is

$$dS = \int_{t_1}^{t_2} \mathcal{L}\left(\phi^r, \phi^r_{,\mu}, q^\mu\right) dt \tag{17}$$

Note that this is the action dS for only the volume element centred at q^μ. The total action of the entire system is the integral of this over some spacetime volume Ω that encloses the entire system spatially during the time interval, so

$$S = \int_\Omega \mathcal{L}\left(\phi^r, \phi^r_{,\mu}, q^\mu\right) d^4q \tag{18}$$

The idea now is to apply the calculus of variations to this integral and require $\delta S = 0$ as in the particle case. Remember that we're varying the *fields* at each spacetime point and not the coordinates q^μ. Therefore (I'll drop the superscript r to avoid confusion with the summation convention, so the following should be taken to apply to each field ϕ^r separately. A summation over μ is implied):

$$\delta S=\int_{\Omega}\left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi+\frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}\delta\phi_{,\mu}\right]d^4q \tag{19}$$

To work out the second term, we write out the derivative explicitly:

$$\frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}\delta\phi_{,\mu}=\frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}\frac{\partial(\delta\phi)}{\partial q^{\mu}} \tag{20}$$

$$=\frac{\partial}{\partial q^{\mu}}\left[\frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}\delta\phi\right]-\frac{\partial}{\partial q^{\mu}}\left[\frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}\right]\delta\phi \tag{21}$$

where the last line follows from the product rule. We therefore get

$$\delta S=\int_{\Omega}\left[\frac{\partial\mathcal{L}}{\partial\phi}-\frac{\partial}{\partial q^{\mu}}\left(\frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}\right)\right]\delta\phi d^4q+\int_{\Omega}\frac{\partial}{\partial q^{\mu}}\left[\frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}\delta\phi\right]d^4q \tag{22}$$

The last term is the integral of a 4-d divergence over a 4-d volume and (trusting the mathematicians again) we can use a 4-d analog of Gauss's theorem to convert this to a surface integral over a 3-d surface Σ that bounds the 4-d volume Ω. Making the usual assumption that this surface can be removed to infinity and that our system is finite so that $\mathcal{L}\to 0$ at infinity, this integral goes to zero. We're left with

$$\delta S=\int_{\Omega}\left[\frac{\partial\mathcal{L}}{\partial\phi}-\frac{\partial}{\partial q^{\mu}}\left(\frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}\right)\right]\delta\phi d^4q=0 \tag{23}$$

The requirement that this is valid for all variations $\delta\phi$ in the field gives us the field theory version of the Euler-Lagrange equations (where I've restored the index r indicating which field we're talking about; note that μ is still summed):

$$\boxed{\frac{\partial\mathcal{L}}{\partial\phi^r}-\frac{\partial}{\partial q^{\mu}}\left(\frac{\partial\mathcal{L}}{\partial\phi^r_{,\mu}}\right)=0} \tag{24}$$

PINGBACKS

Pingback: Hamilton's equations and Poisson brackets
Pingback: Hamilton's equations for relativistic fields