Physics H190
Spring 2013
Homework 3
Due Wednesday, February 20, 2013

Reading Assignment: Please read the lecture notes for Wednesday, February 13, and make sure you understand the Euler-Lagrange equations for classical fields.

At a later point (probably tomorrow) I will add a section to this homework assignment in which I will discuss some of the questions that I have received.

1. A region of three-dimensional space is occupied by a medium with index of refraction $n(\mathbf{x})$. Let $\mathbf{x}_0$ and $\mathbf{x}_1$ be two points of space. Consider the space of paths $\mathbf{x}(\lambda)$, where λ is a parameter, $0 \le \lambda \le 1$,

$$\mathcal{P} = \{\mathbf{x}(\lambda) | \mathbf{x}(0) = \mathbf{x}_0, \mathbf{x}(1) = \mathbf{x}_1\}. \tag{1}$$

Consider the functional defined on paths in $\mathcal{P}$,

$$T[\mathbf{x}(\lambda)] = \frac{1}{c} \int_0^1 n(\mathbf{x}) \, \frac{ds}{d\lambda} \, d\lambda, \tag{2}$$

where c is the speed of light, and where s is the arc length along the path. Note that

$$\frac{ds}{d\lambda} = \left| \frac{d\mathbf{x}}{d\lambda} \right|. \tag{3}$$

Compute the functional derivative,

$$\frac{\delta T(\mathbf{x}(\cdot)]}{\delta \mathbf{x}(\lambda)}, \tag{4}$$

and by setting it to zero, obtain equations for the light ray connecting $\mathbf{x}_0$ and $\mathbf{x}_1$. Express the equations of the path as a differential equation in which s is the independent variable, not λ.

2. In class and in the notes we worked out the equations of motion for $y(x,t)$, which gives the shape of the vibrating string. We assumed the string was undergoing only transverse motions, in the y-direction.

Suppose the y-direction is vertical. Add a term to the Lagrangian to take into account the gravitational potential of the string, assuming that g is the acceleration of gravity. What is the new Lagrangian density? Work out the new equations of motion.

3. I will not ask for a question this week, but if you have one anyway you would like to send to me, please do. It is not obligatory.

Answers to some questions.

Q1. My question for this week is about action extrema. You mentioned that the physical path is an extremum of the action, and not necessarily the minimum. Could you give an example of a classical system whose physical path is the maximum of the action? I initially thought a ball resting on a hill (unstable equilibrium) would work, but the physical path action in that case is 0, while any other path falling off the hill would have a positive action ($L = T - V$, $T > 0$, $V = -T < 0$, so integral of $L > 0$ would be > 0). Also, I can't think of a quantum system with a path of most action off of the top of my head, but I'm sure there's a good example for that as well.

A: The physical paths in the space of all paths connecting the given endpoints and end-times are the ones that cause the action to be *stationary*, not necessarily extremum (maximum or minimum). But many books state this principle incorrectly, saying that the physical path minimizes the action.

I think the student who asked this question must be thinking of physical paths only, since the condition $V = -T$ is conservation of energy when $E = T + V = 0$. But nonphysical paths need not conserve energy. It is important to realize that most of the paths in the path space over which the action functional is defined are not physical (that is, they do not satisfy Newton's laws).

With that understood, we may ask when, if ever, the action actually is maximum or minimum, instead of a saddle point or something else. Insofar as classical mechanics is concerned, it does not matter, since the condition that the path be physical is merely that the first functional derivative of the action should vanish.

But the question is of academic interest, and it turns out to be important in a semiclassical analysis of the path integral. For Lagrangians of the kinetic-minus-potential type, it turns out that the action along the physical path is actually a minimum if the elapsed time is not too long, but after a certain amount of time it becomes a saddle. The action is never maximum for these types of Lagrangians.

Q2. In the notes from Classical Mechanics, section 6, Hamilton's principle defines the quantity $A[q(t)]$ as the action associated with the path $q(t)$. What sort of physical significance (if any) does the action hold? The physical path is a critical point of the action functional, but does the action manifest itself physically?

A: The action appears as the phase associated with the path in the Feynman path integral. This is too elaborate to explain here, but I can post a set of notes from Physics 221A for those who are interested. You may also want to look at the book by Feynman and Hibbs on path integrals.

In addition, the action $A[q(t)]$ when evaluated along a physical path is the generating function for the time evolution of the classical system. Again, it would take too much space to explain this, but this fact was known to Hamilton and generating functions are covered in advanced courses in classical mechanics.

The short answer is that the action $A[q(t)]$, for both physical and nonphysical paths (that is, nonphysical from a classical standpoint), does have a physical significance, partly in classical mechanics and partly in quantum mechanics.

Q3. I was wondering if you could briefly discuss the usefulness of functional and Lagrangian mechanics for our purposes moving forward in the course. I think I would have an easier time wrapping my head around the concepts given a sense of where they will become useful.

A: We are going to be looking at some relativistic quantum field theories, as is necessary for the topic of quantum effects in gravity. The description of quantum fields begins with the Lagrangian for the classical field. We are building up the pieces we need, from classical Lagrangian field theory, to relativistic quantum field theory. Phenomena such as the Casimir effect and the Unruh effect require this background.

Q4. After reading your notes I feel like I have a better understanding of quantization of fields and zero field energy. However, I have trouble understanding what is the Langrangian. Is it equivalent to the 2D or 3D version of the velocity or equation of motion? I am taking Physics 105 at the moment and we have not gone over the Lagrangian just yet. I feel like the concept of the Lagrangian and the action principle is somewhat confusing. In addition, I have trouble understanding why we would have to minimize something. I assumed that minimization was only for photons and light paths. If the Lagrangian is the equation of motion then why must there be minimization if any motion is possible?

A: The Lagrangian (or Lagrangian density, in field theory) is a scalar that implicitly contains within it all the equations of motion. This is explained in the notes on classical mechanics that I posted, and, in the case of field theory, in my lecture notes from last Wednesday.

From a purely classical standpoint it is a mystery why the equations of motion (Newton's laws) should have a variational formulation. (That is, the physical paths are the ones that cause the action to be stationary.) In spite of this, the variational formulation of classical mechanics is very useful. For example, Lagrangians provide the easiest way of transforming the equations of motion from one coordinate system to another. The real significance of the variational formulation of mechanics, however, is to be found in the Feynman path integral in quantum mechanics.

As mentioned, we do not minimize the action to get the physical paths, we only require that it be stationary (the first functional derivative must vanish). Variational formulations were first discovered for light rays (photons), but it was later realized that they also exist for mechanical problems. The Lagrangian is not the equation of motion, but the equations of motion (the Euler-Lagrange equations) can be obtained from the Lagrangian.

Q5. Is it possible to arrive at the Lagrangian formulation of classical mechanics without using Newton's laws? Though the equations of motion can be derived form Hamilton's principle, how is the form of the Lagrangian (i.e. $L = T - V$) determined? In Physics105, the Lagrangian formulation was presented mostly as just a different way to solve problems, and L is essentially defined as $T-V$. Is there a more fundamental reason for the form of the Lagrangian?

A: The Lagrangian is not always $T-V$. There is an extra term if magnetic forces are present, and, in any case, this form is only valid for nonrelativistic problems. In relativistic problems, T is replaced by a term which is not the kinetic energy. This is explained in the notes I posted on classical mechanics, although not in the sections I asked you to read. One can say that the Lagrangian is whatever makes Hamilton's principle work, for a given set of equations of motion.

A more fundamental approach relies on relativistic covariance. According to relativity theory, there is no privileged Lorentz frame, and the equations of motion should be the same in all Lorentz frames. This is most easily achieved by making the Lagrangian (more generally, the action) a Lorentz scalar. Requiring the Lagrangian to be a Lorentz scalar is a severe restriction on the form of the Lagrangian, which in many cases determines it almost uniquely. This is the case for particle motion in electromagnetic fields, for example. Again, this is discussed in the notes on classical mechanics, although not in the sections I asked you to read.

Relativistic covariance determines the Lagrangian for a charged particle in an electromagnetic field almost uniquely. If we then assume that there is no magnetic field, and the electric field is static, and if we assume the particle is slowly moving and take the low velocity limit of the relativistic expression, then we get the Lagrangian $L=T-V$, where $V=q\Phi$ (q is the charge of the particle, Φ is the electrostatic potential).

Q6. I had a question about last week's lecture – how do you create a Lagrangian or a Hamiltonian for a magnetic field, since there is not a scalar potential that can create it?

A: See the previous question and answer. In the low velocity limit, the Lagrangian for a particle of charge q and mass m in an electromagnetic field described by scalar and vector potentials Φ and $\mathbf{A}$ is

$$L=\frac{m}{2}|\dot{\mathbf{x}}|^2-q\Phi+\frac{e}{c}\dot{\mathbf{x}}\cdot\mathbf{A}. \tag{5}$$

It was shown in lecture how this Lagrangian leads to the Hamiltonian,

$$H=\frac{1}{2m}\Big(\mathbf{p}-\frac{q}{c}\mathbf{A}\Big)^2+q\Phi. \tag{6}$$

Q7. I was wondering if minimizing functionals are always unique. In the examples of path of light in different media that we went over in class, there was only one path that minimized the time traveled, and it seemed to me that functionals would be very practical if they had unique minimizing solutions. Also, similarly, do functionals always have solutions?

A: Just like when you look for extrema of ordinary functions (even functions of one variable, like $y=f(x)$), there may be more than one maximum or minimum (more generally, critical points), or there may not be any. It is true that the optics problems we considered in lecture had only one solution.

Many books state state this incorrectly, saying that given endpoints and end-times, there is a unique classical trajectory connecting the two. That this is not so is easily seen in some simple examples, such as the particle in the box (for which there are an infinite number of trajectories, given endpoints and end-times).

In the case of the harmonic oscillator, the trajectory is unique for most times, but for special times (multiples of a half-period) there are either an infinite number of solutions, or none at all, depending on the endpoints.

Q8. Can we generalize the notion of a derivative of higher order functions? We have a hierarchy: parameter, function, functional, function of a functional, etc.

This is more like a math question than physics, but I was just curious.

A: Yes, there is the functional derivative, which was discussed in class and in the posted notes. It's important for this course. One can also define second, third, etc functional derivatives. For example, the second functional derivative of a functional $F[x(t)]$ can be denoted,

$$\frac{\delta^2 F[x(\cdot)]}{\delta x(t)\,\delta x(t')}. \tag{7}$$

It is the generalization of the second partial derivative of an ordinary function of several variables,

$$\frac{\partial^2 f}{\partial x_i \partial x_j}, \tag{8}$$

where $f = f(x_1, \ldots, x_n)$, in which the discrete indices i and j are replaced by continuous "indicices" t and t'. The second functional derivative is discussed in the book.

Q9. In the lecture last week I was a little bit confused when we were dealing with the matrix that represents $\delta^2(f)/\delta(x_i)\delta(x_j)$. We said when some of the eigenvalues of this matrix are positive and some of them are negative, then we have a saddle. But whenever I've dealt with saddles in the past, I've used the relation that the double derivative of the function we're looking at is equal to zero. So wouldn't a saddle mean that we'd have to in fact look at the 3rd order term? Or is this not the case because we're looking at partial derivatives with respect to two different variables?

A: You must mean the matrix in i and j shown in (8). I think you're confusing a saddle (which applies to functions of many variables, here $f(x_1, \ldots, x_n)$) with an inflection point (a point of a function of one variable, $f(x)$, where $f'' = 0$). A saddle is where some of the eigenvalues of the matrix (8) are positive, and some negative. I didn't attempt to draw a saddle in lecture, because it's hard, but a mountain pass gives the right idea. The generalization of an inflection point to many variables is a point at which some of the eigenvalues of the matrix (8) are zero.

Q10. As I was reading your notes on the form of the Lagrangian for relativistic applications I was a bit confused as to why we require a covarient formulation? What are the benefits of such a formulation over a contravarient formulation?

A: Confusing point. The word "covariant" is used in two different senses in tensor analysis. It is used to describe a certain transformation law for vectors (called covariant vectors), which is in contrast to the contravariant transformation law (a different kind of vector, the contravariant vectors). But the word "covariant" is also used to describe a set of equations that have the same form in all coordinate systems (of some given class). For example, Maxwell's equations are covariant (have the same form) in all Lorentz frames.

Maxwell did not know this, he thought his equations were valid in only one frame, the frame of the "ether."

Q10. Why are some Lagrangians that we encounter in classical mechanics manifestly Galilean-variant, whereas those in field theory are manifestly Lorentz invariant?

For example, consider the simple Lagrangian:

$$L = \frac{1}{2}mv^2, \tag{9}$$

where v is just $\dot{x}$. under the Galilean boost $v \to v + a$ (where a is some constant velocity), the Lagrangian picks up additional terms:

$$L' = \frac{1}{2}m(v^2 + a^2 + 2va), \tag{10}$$

which breaks the symmetry; that is, $L \neq L'$. However, all the Lagrangians in field theory are manifestly Lorentz invariant – one of the ways we can motivate the free field Lagrangian is to write down the Lorentz invariant terms associated with the field. Shouldn't we try to ensure the Galilean invariance of our Lagrangians in classical mechanics?

A: I think this student must have meant, "Why are some Lagrangians that we encounter in classical mechanics *not* manifestly Galilean-variant, whereas those in field theory are manifestly Lorentz invariant?"

Actually, what we want is that the equations of motion should be the same after doing the transformation (Galillean or Lorentz). If the Lagrangian is invariant, then the equations of motion are the same, but there are changes to the Lagrangian that will not affect the equations of motion. First, if a constant is added to the Lagrangian, that will not affect the equations of motion, because adding a constant to the Lagrangian just adds a constant to the action, and doesn't affect the stationary paths. So the term $(1/2)ma^2$ in (10) doesn't affect the equations of motion, and can be dropped.

Second, we can add the total time derivative of any function of q and possibly t to the Lagrangian, without affecting the equations of motion, that is, we can make the replacement,

$$L \to L + \frac{df(q,t)}{dt}, \tag{11}$$

where

$$\frac{df(q,t)}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial q}\,\dot{q}. \tag{12}$$

This is because the extra term causes the action to change by

$$\int_{t_0}^{t_1} dt\,\frac{df(q,t)}{dt} = f(q_1,t_1) - f(q_0,t_0). \tag{13}$$

This is constant (the same for all the paths in path space), and so does not affect the paths that cause the action to be stationary. The term $(1/2)(2mva)$ in (10) is of this form,

$$mva = \frac{d}{dt}(mvx). \tag{14}$$

Altogether, we see that certain changes in the Lagrangian do not affect the equations of motion; these are allowed under transformations when the equations of motion are supposed to be form-invariant.