M393C NOTES: TOPICS IN MATHEMATICAL PHYSICS

ARUN DEBRAY
AUGUST 31, 2017

These notes were taken in UT Austin's M393C (Topics in Mathematical Physics) class in Fall 2017, taught by Thomas Chen. I live-TEXed them using `vim`, so there may be typos; please send questions, comments, complaints, and corrections to a.debray@math.utexas.edu.

Contents

Lecture 1.

The Lagrangian formalism for classical mechanics: 8/31/17

The audience in this class has a very mixed background, so this course cannot and will not assume any physics background. We'll first discuss classical and Lagrangian mechanics. Quantum mechanics is, of course, more fundamental, and though historically people obtained quantum mechanical mechanics from classical mechanics, it should be possible to go in the other direction.

We'll start, though, with classical and Lagrangian mechanics. This involves understanding symplectic and Poisson structures, and the principle of least action, the beautiful insight that classical mechanics can be formulated variationally; there is a Lagrangian L and an action functional

$$S = \int_{t_0}^{t_1} L \, \mathrm{d}t,$$

and the system evolves through paths that extremize the action functional.

The history of the transition from classical mechanics to quantum mechanics to quantum field theory happened extremely quickly in the historical sense, all fitting into one lifetime. JJ Thompson discovered the electron in 1897, and in 1925, GP Thompson, CJ Dawson, and LH Germer discovered that it had mass. This led people to discover some inconsistencies with classical physics on small scales, ushering in quantum mechanics, with all of the famous names: Einstein, Schrödinger, Heisenberg, and more. The basic equations of quantum mechanics fall in linear dispersive PDE for functions living in the Hilbert space, typically L^2 or the Sobolev space H^1 (since energy involves a derivative).

One of the key new constants in quantum mechanics is *Planck's constant* $\hbar := h/2\pi$. It has the same units as the classical action S, and therefore they are comparable. There is a sense in which quantum mechanics is the regime in which $S/\hbar \approx 1$, and classical mechanics is the regime in which $S/\hbar \gg 1$. In this sense, quantum mechanics is the physics of very small scales. Sometimes people take a "semiclassical limit," and say they're letting $\hbar \to 0$, but this makes no sense: $\hbar$ is a physical quantity. Instead, it's more accurate to say taking a semiclassical limit lets $(S/\hbar)^{-1} \to 0$.

If you want to analyze a fixed number of electrons, life is good. They will always be there, and so on. But this is a problem for photons, as there are physical processes which create photons, and processes which destroy photons. Thus imposing a fixed number of quantum particles is a constraint — and the theory which describes the quantum physics of arbitrary numbers of quantum particles, quantum field theory, was worked out a little later. In this case, the Hilbert space is a direct sum over the Hilbert subspace of 1-particle states, 2-particle states, etc., and is called *Fock space*. The symplectic and Poisson structures of classical mechanics, transformed into commutation relations of operators in quantum mechanics, is again interpreted as commutation relations of creation and annihilation operators.

The mathematics of quantum field theory is rich and diverse, drawing in more PDE as well as large amounts of geometry and topology. But there's a problem — many important integrals and power series don't converge. And this is not a formal series problem: it's too central. Physicists have used renormalization as a formal trick to solve these divergences; it feels like a dirty trick that produces incredibly accurate results agreeing with experiment. But again there are problems: renormalization expresses Fock space and the commutation relations in terms of the noninteracting case, and the results you get don't necessarily agree with what you did *a priori*.

For example, quantum field theory contains a Hamiltonian H whose spectrum is of interest. One can imagine starting with the noninteracting Hamiltonian H_0 and perturbing it by some small operator W: $H \coloneqq H_0 + W$. You're often interested in the resolvent

$$\begin{aligned} R(z) &= (H - z)^{-1} \\ &= (H_0 - z)^{-1} \sum_{\ell=0}^{\infty} \Big(W(H_0 - z)^{-1} \Big)^{\ell}. \end{aligned}$$

The issue is that adding W does not do nice things to the spectrum, and this is part of the complexity of quantum field theory.

Let λ denote the interaction, and N denote the number of particles, and suppose $\lambda \sim 1/N$ as we let $N \to \infty$. Then, the equations describing the mean field theory for this system are complicated, typically nonlinear PDEs. Typical examples include the nonlinear Schrödinger equation, the nonlinear Hartree equation, the Vlasov equation, or the Boltzmann equation. We'll hopefully see some of these equations in this class.

This is a lot of stuff that's tied together in complicated and potentially confusing ways, and hopefully in this class we'll learn how to make sense of it.

Classical mechanics and symplectic geometry In classical mechanics, we think of objects in idealized ways, e.g. thinking of a stone as a point mass at its center of mass. Thus, we're studying the motion of idealized point masses (or particles, in the strictly classical sense). We do this by letting time be $t \in \mathbb{R}$; at a time t, the particles $x_1, \dots, x_N$ have positions $\mathbf{q}(t) \coloneqq (q_1(t), \dots, q_N(t))$, with $q_i(t) \in \mathbb{R}^d$; these are called "generalized coordinates."

Classical mechanics says that the kinematics of particles can be completely described by their position and velocity. Thus the motion of a system is completely determined by $\mathbf{q}(t)$ and $\dot{\mathbf{q}}(t) \coloneqq \frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t}$.

The next question: what determines the motion? The answer is the Newtonian equations of motion: $\ddot{\mathbf{q}}$ is expressed as a function of $\dot{\mathbf{q}}$ and $\mathbf{q}$ using *Hamilton's principle*, also known as the *principle of least action*.

(1) Let $\mathbf{q} \in C^2([t_0, t_1], \mathbb{R}^{Nd})$ be a curve in $\mathbb{R}^{Nd}$. We associate to $\mathbf{q}$ a weight function $L(\mathbf{q}, \dot{\mathbf{q}})$ called the *Lagrangian*.
(2) Given $\mathbf{q}$ as above, define the *action functional*

$$S[\mathbf{q}] \coloneqq \int_{t_0}^{t_1} L(\mathbf{q}(t), \dot{\mathbf{q}}(t))\, \mathrm{d}t.$$

(3) Then, among all C^2 curves with $\mathbf{q}(t_0)$ and $\mathbf{q}(t_1)$ fixed, the curve that minimizes S is the one that satisfies the equations of motion.

Now let $\mathbf{q}_\bullet(t)$ be a C^2 family of curves $[t_0, t_1] \times \mathbb{R} \to \mathbb{R}^{Nd}$ and that $\mathbf{q}_0$ minimizes S. Then,

$$\partial_s|_{s=0}\, S[\mathbf{q}_s] = 0.$$

We can apply this to the Lagrangian to derive the equations of motion.

$$\begin{aligned} \partial_s|_{s=0}\, S[\mathbf{q}_s] &= \int_{t_0}^{t_1} \left((\nabla_{\mathbf{q}_s} L) \cdot \partial_s \mathbf{q}_s(t) + (\nabla_{\dot{\mathbf{q}}_s} L) \cdot \partial_s \dot{\mathbf{q}}_s(t) \right) \mathrm{d}t \Big|_{s=0} \\ &= \int_{t_0}^{t_1} \left(\nabla_{\mathbf{q}_s} L - (\nabla_{\dot{\mathbf{q}}_s} L)^{\bullet} \right)\big|_{s=0} \cdot \underbrace{\partial_s|_{s=0}\, \mathbf{q}_s(t)}_{\delta\mathbf{q}(t)}\, \mathrm{d}t + (\nabla_{\dot{\mathbf{q}}_0} L) \cdot \underbrace{(\partial_s|_{s=0} \mathbf{q}(t))}_{=0} \Big|_{t_0}^{t}, \end{aligned}$$

where $\delta\mathbf{q}(t)$ is the variation. For all variations, this is nonzero. Thus, minimizers of S satisfy the *Euler-Lagrange equations*

$$\nabla_{\mathbf{q}} L - (\nabla_{\dot{\mathbf{q}}} L)^{\bullet} = 0. \tag{1.1}$$

We'll now impose some conditions on L that come from reasonable physical principles.

Additivity: if we analyze a system $A \cup B$ which is a union of two subsystems A and B that don't interact, then

$$L_{A\cup B} = L_A + L_B.$$

Uniqueness: Assume L_1 and L_2 differ only by a total time derivative of a function $f(\mathbf{q}(t), t)$; then, they should give rise to the same equations of motion:

$$\begin{aligned} S_2 &= S_1 + \int_{t_0}^{t_1} \partial_t f(\mathbf{q}(t), t)\, \mathrm{d}t \\ &= S_1 + f(\mathbf{q}(t_1), t_1) - f(\mathbf{q}(t_0), t_0), \end{aligned}$$

so the minimizers for S_1 and S_2 are the same.

Galilei relativity principle: The physical laws of a closed system are invariant under the symmetries of the *Galilei group* parameterized by $a, v \in \mathbb{R}^d$, $t \in \mathbb{R}$, and $R \in \mathrm{SO}(d)$, the group element $g_{a,v,R,b}$ acts by

$$\begin{aligned} \mathbf{q} &\longmapsto a + vt + Rq \\ t &\longmapsto t + b. \end{aligned}$$

That is, in each component j, $q_j \mapsto a + vt + Rq_j$.

This actually determines L for a system consisting of a single particle. By homogeneity of space (by the Galilei group contains translations), L can only depend on $V = \dot{q}$. Since space is isotropic (because the Galilei group contains rotations), L should depend on v^2. Next, the Euler-Lagrange equations imply

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial v} - \frac{\partial L}{\partial q} = 0,$$

and since L does not depend on q, $\frac{\partial L}{\partial q} = 0$, so $\frac{\partial L}{\partial v}$ must be a constant.

Now we consider Galilei invariance of v If $v \mapsto v + \varepsilon$, the equations of motion must be invariant, so

$$L[(v')^2] = L[(v+\varepsilon)^2] = L(v^2) + \frac{\partial L}{\partial v^2} 2v \cdot e + O(\varepsilon),$$

and this should only differ by a total time derivative $\dot{q}$:

$$F(\dot{q}) \cdot \dot{q} = \partial_t G,$$

where $F(\dot{q})$ is a constant, and $\frac{\partial L}{\partial v^2}$ is also constant. This latter constant is denoted m, and called the *mass*, and the Lagrangian expresses its kinetic energy:

$$L(v) = \frac{1}{2} m v^2.$$

Now imagine adding N particles, which we assume don't interact. Then additivity tells us they have masses $m_1, \ldots, m_N$, and the Lagrangian is

$$L = \frac{1}{2} \sum_{j=1}^{N} m_j v_j^2.$$

If the particles are interacting, there's some potential function $U(q_1, \ldots, q_N)$, and the Lagrangian is instead

$$L = \frac{1}{2} \sum_{j=1}^{N} m_j v_j^2 - U(q_1, \ldots, q_N).$$

Now, by (1.1),

$$m_j \ddot{q}_j = -\partial_{q_j} U = F,$$

and this is called the *force*. This is Newton's second law $F = ma$.

Symmetries and conservation laws There's a general result called Noether's theorem which shows that any symmetry of a physical system leads to a conserved quantity. We'll see the presence of symmetry in classical mechanics and then how it changes in quantum mechanics.

For example, the systems we saw above had symmetries under time translation invariance $t \mapsto t + b$, so the Lagrangian doesn't depend on t, just on $\mathbf{q}$ and $\dot{\mathbf{q}}$. Therefore

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}L &= \sum_j \left(\frac{\partial L}{\partial q_j}\dot{q}_j + \frac{\partial L}{\partial \dot{q}_j}\ddot{q}_j\right) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\sum_{j=1}^{N}\left(\frac{\partial L}{\partial \dot{q}_j}\right)\cdot \dot{q}_j,\end{aligned}$$

and therefore

$$\underbrace{\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{j=1}^{N}\frac{\partial L}{\partial \dot{q}_j}\cdot\dot{q}_j - L\right)}_{E} = 0.$$

The quantity E is the *energy* of the system, and time translation invariance tells is that energy is conserved. The component $p_j \coloneqq \frac{\partial L}{\partial \dot{q}_j}$ is called the j^{th} *canonical momentum*.

The homogeneity of space, told to us by invariance under the Galilei translations $q_j \mapsto q_j + \varepsilon$, tells us that

$$\begin{aligned}\delta L &= \sum_i \frac{\partial L}{\partial q_j}\cdot\varepsilon \\ &= \varepsilon\frac{\mathrm{d}}{\mathrm{d}t}\sum \partial L q_j = 0.\end{aligned}$$

Thus, the quantity

$$\mathbf{p} \coloneqq \sum_{j=1}^{N}\frac{\partial L}{\partial q_j}$$

is conserved, and is constant. This is called the *total momentum*, so translation-invariance gives you conservation of momentum. In the same way, rotation-invariance around any center gives you conservation of angular momentum around any center.

Hamiltonian dynamics The Euler-Lagrange equations express $\ddot{\mathbf{q}}$ as a second-order ODE. One might want to reformulate this into a first-order ODE; there are many ways to do this. There's one that's particularly important. Since

$$p_j = \frac{\partial L}{\partial \dot{q}_j}(\mathbf{q}, \dot{\mathbf{q}}),$$

then it looks like one could solve for $\dot{\mathbf{q}}$ in terms of $\mathbf{p}$ and $\mathbf{q}$.

Lemma 1.2. *Let $f \in C^2(\mathbb{R}^n, \mathbb{R})$ be such that its Hessian $D^2 f$ is uniformly positive definite, i.e. there's an $\alpha > 0$ such that*

$$D^2 f(x)(h,h) = \sum_{i,j}\frac{\partial^2 f}{\partial x_j \partial x_\ell}h_j h_\ell \geq \alpha\|h\|^2$$

uniformly in $x \in \mathbb{R}^n$, then there is a unique solution to

$$Df(x) = y$$

for every $y \in \mathbb{R}^n$.

Proof. Let $g(x,y) \coloneqq f(x) - \langle x, y\rangle$. Then, $\nabla_x g(x,y) = \nabla f - y$, and $D^2 g = D^2 f$. Hence it suffices to check for $y = 0$.

The positive definite assumption on $D^2 f$ means f is strictly convex, and hence has at most a single critical point, at which $\nabla f = 0$. Thus it remains to check that there's at least one solution.

If you Taylor-expand, you get that

$$f(x) = f(0) + \langle Df(0), x\rangle + \frac{1}{2}D^2 f(sx)(x,x) + \cdots,$$

so for all x,

$$f(x) \geq f(0) - |\nabla f(0)||x| + \frac{\alpha}{2}|x|^2.$$

Thus, there's an $R > 0$ such that if $|x| \geq R$, then $f(x) \geq f(0)$, so f has at most one minimum in the ball $\overline{B_R(0)}$, so by compactness, it has a minimum x_0, which must be the global minimum, so $Df(x_0) = 0$. ⊠

Definition 1.3. Suppose f is continuous on $\mathbb{R}^n$. Then, its *Legendre transform* or *Legendre-Fenchel transform* is

$$f^*(y) \coloneqq \sup_{x \in \mathbb{R}^n} (\langle y, x \rangle - f(x)).$$

You can think of this as measuring the distance from the graph of f to the line cut out by $\langle y, x \rangle$ (i.e. between the two points with minimum distance).