Advanced Mechanics

Alberto Nicolis

Transcribed by Ron Wu

This is an advanced undergraduate course, offered in spring 2013 at Columbia University. Required Course textbook: Landau, Lifshitz, *Mechanics.* Recommended books: Goldstein, *Classical Mechanics.* Arnold, *Mathematical Methods of Classical Mechanics.*

Contents

Course Overview

Lecture 1 (1/22/13)

This course mainly focuses on three things: rewriting, generalizing, and quantizing.

We will rewrite Newton's mechanics into two formalism: Lagrangian, Hamiltonian.

After that (near the end of semester), we will generalize these two formalism to (classical field theory) give the universal structure of other branches of modern physics. We will also modify the rules and introduce new rules (quantization) to get interpretation QM.

$$\begin{array}{ccccc}
\text{Newton Mechanics} & \rightarrow & \text{Lagrangian, Hamiltonian} & \rightarrow & \text{QM} \\
 & & \downarrow & & \downarrow \\
\begin{array}{c}\text{Maxwell eq} \\ \text{Einstein eq}\end{array} & \rightarrow & \text{classical field theory} & \rightarrow & \text{QFT} \\
 & & \downarrow & & \downarrow \\
 & & \text{classical string theory} & \rightarrow & \text{Q string theory}
\end{array}$$

From QM, we can add relativity to get QFT, then replace particles by strings, to get string theory. QFT is so far the most precious theory we know, and it agrees experiments spectacularly. Surprisingly that string theory at classical level is nothing but classical mechanics.

1 Lagrangian Mechanics

This is of course equivalent to Newton's mechanics, but we will see it has two advantages.

1.1 Euler-Lagrange Equation

Consider system of N particles,

$$m_a \ddot{\vec{r}}_a = \vec{F}_a$$

we use a, b, $c = 1, ..., N$ to label particles, $\vec{F}_a$ total force acting on particle a.

We only consider *conservative force,* i.e. $\exists V(\vec{r}_1, ..., \vec{r}_N)$ s.t. $\vec{F}_a = -\vec{\nabla}_a V$. The symbol means

$$F_a^i = -\frac{\partial}{\partial r_a^i} V$$

we use i,j,$k = 1, 2, 3$ to label Cartesian directions. So

$$m_a \ddot{\vec{r}}_a + \vec{\nabla}_a V = 0$$

or

$$\frac{d}{dt}(m_a \dot{\vec{r}}_a) + \vec{\nabla}_a V = 0 \tag{1.1}$$

We define kinetic energy:

$$T \equiv \sum_{a=1}^{N} \frac{1}{2} m_a |\dot{\vec{r}}_a|^2 = T(\dot{\vec{r}}_1, ..., \dot{\vec{r}}_N)$$

then (1.1) becomes

$$\frac{d}{dt}(\frac{\partial T}{\partial \dot{\vec{r}}_a}) + \vec{\nabla}_a V = 0 \tag{1.2}$$

We define Lagrangian of the system

$$\mathcal{L} = T - V = \mathcal{L}(\vec{r}_1, ..., \vec{r}_N, \dot{\vec{r}}_1, ..., \dot{\vec{r}}_N)$$

the minus is crucial, so that $\mathcal{L}$ is not the energy.

Therefore (1.2) becomes

$$\frac{d}{dt}(\frac{\partial \mathcal{L}}{\partial \dot{\vec{r}}_a}) - \frac{\partial \mathcal{L}}{\partial \vec{r}_a} = 0 \tag{1.3}$$

This form has two advantages over Newton's mechanics.

1. Independent of coordinates. (while in Newton change coordinate may result inertial force)

2. Constraints come in naturally. (while in Newton that becomes adding extra force too)

1.2 Generic Coordinates

We will show this now and leave constraints to next lecture.

Let us consider two sets of trajectories

$$\{\vec{r}_a(t)\}_{a=1,\dots,N}, \quad \{\vec{r}_a(t)+\delta\vec{r}_a(t)\}_{a=1,\dots,N}$$

where $\vec{r}_a(t)$ is the trajectory of particle a, i.e. solution of the system. $\delta\vec{r}_a(t)$ infinitesimal arbitrary derivation from the path.

Consider $\delta\vec{r}_a(t)$ dot (1.3)

$$\begin{aligned} 0 &= \sum_{a=1}^{N}\delta\vec{r}_a(t)\cdot\left[\frac{d}{dt}(\frac{\partial\mathcal{L}}{\partial\dot{\vec{r}}_a})-\frac{\partial\mathcal{L}}{\partial\vec{r}_a}\right] \\ &= \sum_{a=1}^{N}\left[\frac{d}{dt}(\frac{\partial\mathcal{L}}{\partial\dot{\vec{r}}_a}\cdot\delta\vec{r}_a(t))\right]-\sum_{a}^{N}\left(\frac{\partial\mathcal{L}}{\partial\dot{\vec{r}}_a}\cdot\delta\dot{\vec{r}}_a(t)+\frac{\partial\mathcal{L}}{\partial\vec{r}_a}\cdot\delta\vec{r}_a(t)\right) \end{aligned} \tag{1.4}$$

The second term on the right is the first order Taylor expansion of $(\delta\mathcal{L})_a \equiv \mathcal{L}(\vec{r}_a + \delta\vec{r}_a, \dot{\vec{r}}_a+\delta\dot{\vec{r}}_a) - \mathcal{L}(\vec{r}_a,\dot{\vec{r}}_a)$, $(\delta\mathcal{L})_a$ is variation of path of particle a only.

Recall

$$f(x+\Delta x, y+\Delta y) = f(x,y)+\frac{\partial f}{\partial x}\Delta x+\frac{\partial f}{\partial y}\Delta y+O(x,y)^2$$

Now we use a generic coordinate. It can be curvilinear, non-inertial, or whatever

$$q_1,\dots,q_{3N}\to q_\alpha$$

where $3N$ is the number of degree of freedom, and we will use q_α to denote the new coordinate, we use α, β, $\gamma ... = 1,\dots,3N$ to label the order of degree of freedom.

Consider the same two trajectories in new coordinate

$$\{q_\alpha(t)\}_\alpha, \qquad \{q_\alpha(t)+\delta q_\alpha(t)\}_\alpha$$

and the coordinate transformation

$$\vec{r}_a = \vec{r}_a(q_1,\dots,q_{3N};t) \text{ for each } a \tag{1.5}$$

Note: (1) t here is due to wrt $\vec{r}_a$ coordinate, q_α coordinate maybe is in motion, so q_α coordinate changes with time. One can think this way, if q_α coordinate is moving with the point particle, then q_α will be constant, then if $\vec{r}_a = \vec{r}_a(q_1, ..., q_{3N})$ had no explicit time dependence, then $\vec{r}_a$ would be constant too.

(2) the order of $q_1, ..., q_{3N}$ is not necessary q_1, q_2, q_3 for particle 1 and etc, the order does not matter at all.

We then define

$$\mathcal{L}(q, \dot{q}, t) = \mathcal{L}(\vec{r}(q; t), \dot{\vec{r}}(q, \dot{q}; t))$$

Note: (1) Now $\mathcal{L}$ has explicit time dependance too, for the same reasoning given above.

(2) Here $\dot{\vec{r}}_a(q, \dot{q}; t)$ has $\dot{q}$ dependence. That is due to (1.5), so

$$\dot{\vec{r}}_a = \sum_\beta \frac{\partial \vec{r}_a}{\partial q_\beta} \dot{q}_\beta + \frac{\partial \vec{r}_a}{\partial t} \tag{1.6}$$

(3) Scalar function $\mathcal{L}$ defined above in terms of new coordinate has the same value as in the old coordinate. Therefore $(\delta \mathcal{L})_a$ is independent of coordinate.

So only need to show the first term on the right of (1.4) is independent of coordinate.

Is

$$\sum_{a=1}^{N} \left[\frac{d}{dt} (\frac{\partial \mathcal{L}}{\partial \dot{\vec{r}}_a} \cdot \delta \vec{r}_a(t)) \right] = \sum_\beta^{3N} \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\beta} \delta q_\beta \right) ? \tag{1.7}$$

Notice the left hand side has dot product, and the right hand side doesn't.

So we do

$$\begin{aligned}
\sum_\beta^{3N} \frac{\partial \mathcal{L}}{\partial \dot{q}_\beta} \delta q_\beta &= \sum_{\beta, a} \frac{\partial \mathcal{L}}{\partial \dot{\vec{r}}_a} \frac{\partial \dot{\vec{r}}_a}{\partial \dot{q}_\beta} \delta q_\beta \text{ by chain rule} \\
&= \sum_{\beta, a} \frac{\partial \mathcal{L}}{\partial \dot{\vec{r}}_a} \frac{\partial \vec{r}_a}{\partial q_\beta} \delta q_\beta \text{ by (1.6)} \\
&= \sum_a \frac{\partial \mathcal{L}}{\partial \dot{\vec{r}}_a} \partial \vec{r}_a
\end{aligned}$$

This proves (1.7). Hence (1.4) becomes

$$0 = \sum_{\beta=1}^{3N} \left[\frac{d}{dt} (\frac{\partial \mathcal{L}}{\partial \dot{\vec{q}}_\beta} \delta q_\beta(t)) \right] - \sum_\beta \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\beta} \delta \dot{q}_\beta(t) + \frac{\partial \mathcal{L}}{\partial q_\beta} \delta q_\beta(t) \right)$$

Or

$$\sum \delta q_\beta \left[\frac{d}{dt} (\frac{\partial \mathcal{L}}{\partial \dot{\vec{q}}_\beta}) - \frac{\partial \mathcal{L}}{\partial q_\beta} \right] = 0$$

$\delta \vec{r}_a$ is arbitrary, so is δq_β. That implies

$$\frac{d}{dt} (\frac{\partial \mathcal{L}}{\partial \dot{\vec{q}}_\beta}) - \frac{\partial \mathcal{L}}{\partial q_\beta} = 0 \, \forall \beta \tag{1.8}$$

Just like Newton mechanics, the $3N$ components are independent, but here coordinate is completely free to choose.

(1.8) is Euler-Lagrange equation, or called "equation of motion" (EOM).

Example 1. free particle in 1D

$$\mathcal{L}(x, \dot{x}) = T - V = \frac{1}{2} m \dot{x}^2 = \mathcal{L}(\dot{x})$$

Apple E-L

$$\frac{d}{dt}(m\dot{x}) = 0 \implies m\ddot{x} = 0 \tag{1.9}$$

Let now choose a crazy coordinate

$$q = e^x$$

so $x = \ln q$, $\dot{x} = \dot{q}/q$, and $\ddot{x} = \ddot{q}/q - \dot{q}^2/q^2$

$$\mathcal{L}(q, \dot{q}) = \frac{1}{2} m \frac{\dot{q}^2}{q^2}$$

Apple E-L, we get

$$\frac{1}{q}(m\frac{\ddot{q}}{q} - m\frac{\dot{q}^2}{q^2}) = 0$$

agrees with (1.9).

Lecture 2 (1/24/13)

Example 2. Particle in 2D central potential

Lagrangian in Cartesian

$$\mathcal{L} = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - V(\sqrt{x^2 + y^2})$$

in term of polar

$$\mathcal{L} = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2) - V(r)$$

Then equation of motion in r component

$$m\ddot{r} = mr\dot{\phi}^2 - V' \tag{1.10}$$

notice $mr\dot{\phi}^2$ is the usual centripetal force and $-V'$ is the radial force associated with V.

In ϕ component

$$\frac{d}{dt}(mr^2\dot{\phi}) = 0$$

$mr^2\dot{\phi}$ is constant. This corresponds to conservation law for angular momentum.

This leads to interpret q, $\dot{q}$ in $\mathcal{L}(q, \dot{q}; t)$ as generalized coordinate and generalized velocity.

Definition. We also define

$$p_\alpha \equiv \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \tag{1.11}$$

to be momentum conjugate to q_α.

In the old way

$$\vec{p} = m\dot{\vec{r}} = \frac{\partial}{\partial \dot{\vec{r}}}(\frac{1}{2}m\dot{\vec{r}}) = \frac{\partial \mathcal{L}}{\partial \dot{\vec{r}}}$$

Notice in example 1,

$$\frac{\partial \mathcal{L}}{\partial x} = 0$$

By EOM,

$$p_x \equiv \frac{\partial \mathcal{L}}{\partial \dot{x}} = \text{const}$$

In example 2,

$$\frac{\partial \mathcal{L}}{\partial \phi} = 0$$

so

$$p_\phi \equiv \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \text{const}$$

Notice in this example, $p_\phi = mr^2\dot{\phi}$, so p_ϕ is not just depended on $\dot{\phi}$. This is a big difference between Cartesian and generalized coordinate, i.e. $p_\alpha \not\propto \dot{q}_\alpha$.

Theorem. *If $\mathcal{L}$ does not depend explicitly on one of the q's, then associated conjugate momentum is conserved.*

This is a special case of Noether theorem, saying symmetry implies conservation.

Proof. EOM for q_α

$$\frac{d}{dt}p_\alpha = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} = \frac{\partial \mathcal{L}}{\partial q_\alpha} = 0$$

$\square$

One may ask why p_α goes with $\dot{q}_\alpha$ not q_α. This is a very deep question. It relates to the fact that not all coordinate are dynamical variables. (EM for example coulomb gauge introduce scalar potential, which added more degree of freedom than needed.) So only the one has $\dot{q}_\alpha$ meaning should give p_α.

1.3 Constraints

Constraints relate among different degrees of freedom. This can be because not all degrees are independent, not all EOM are independent, or sometimes some forces are not interesting.

For examples:

(1) Mass moving on sphere;

(2) fly flying inside a room;

(3) sphere rolling (no slide) on the floor

ex1 is *holonomic* constraint. It is expressed through equalities

$$f_A(q_1, ..., q_{3N}; t) = 0 \quad A = 1, 2, ..., M \le 3N$$

here t is due to q coordinate is moving or constraints is changing.

ex2 is constraint that is given by inequality

$$f_A(q_1, ..., q_{3N}; t) \leq 0$$

ex3 is constraint that has $\dot{q}$ dependence

$$f_A(q_1, ..., q_{3N}, \dot{q}_1, ..., \dot{q}_{3N}; t) = 0$$

because for rolling ball, rotation speed, rolling direction and rolling distance are all correlated.

In our class, we will only consider holonomic constraints. (2), (3) types are hard, and can only be solved cases by cases.

We will follow Arnold section 17.

Although in ex1, mass moves on a sphere, then we should have $x^2 + y^2 = R^2$ for all time. Or choose polar coordinate, then $dr = 0$ for all time, but in reality, they are tiny displacement that violate the constraints (which may be result of very steep potential in those direction).

So the argument is that let us keep all $q_1, ..., q_{3N}$ *"as"* independent, and choose arrange them that the first n of q's be unconstrained, and the rest q's to be constraint.

Then EOM is still valid, so one can solve the $3N$ pdes for q_α, and the solutions of q_α for $\alpha = (n+1), ..., 3N$ should be $\cong$ constants (choose to be 0). Then one will use these q_α for $\alpha = n+1, ..., 3N$, plugging into the solutions for q_α, $\alpha = 1, ..., n$ to get a simplified answer.

But we claim there is a simpler way

Theorem. *The EOM one gets for $\alpha = 1, ..., n$ are the same as the EOM associated with*

$$\mathcal{L}_{eff}(q_1, ..., q_n, \dot{q}_1, ..., \dot{q}_n; t)$$

$$= \mathcal{L}(q_1, ..., q_n, q_{n+1} = 0, ...q_{3N} = 0, \dot{q}_1, ..., \dot{q}_n, \dot{q}_{n+1} = 0, ..., \dot{q}_{3N} = 0; t)$$

What the theorem is saying is that one can solve the $3N$ pdes of

$$\frac{d}{dt}(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}) - \frac{\partial \mathcal{L}}{\partial q_\alpha} = 0$$

in the usual way (so the constraint is handled by potential that is constant if the constraint $q_\alpha = 0$ and change steeply if q_α becomes non-zero just a tiny bit), or one can just use $\mathcal{L}_{eff}$ which is to plug into $q_\alpha = 0$, $\dot{q}_\alpha = 0$ before taking the derivatives and solve the pdes using $\mathcal{L}_{eff}$. Later we will show this method has serious limitations. Must only be used for holonomic constraints, i.e. both q_α, $\dot{q}_\alpha$ are known to be zero no matter what other variables do.

Here are some steps for solving problems involving holonomic constraints

(1) Parametrize the system with generalized unconstrained variables $q_1 ..., q_n$

(2) write $\mathcal{L} = T - V = \mathcal{L}(q_1 ..., q_n, \dot{q}_1 ..., \dot{q}_n; t)$

(3) solve EOM for $\alpha = 1, ..., n$

Example. 2 masses, M_1, M_2 connected by a pulley, with rope l , under gravity

Without the rope, the two masses would move vertically independently so the degree of freedom is 2.

Let $q_1 = x_1$ the vertical distance from M_1 to the pulley. x_2 the vertical distance from M_2 to the pulley. Then clearly $x_2 = l - x_1$, if we choose $q_2 = x_2 - (l - x_1)$ to represent the change of length of the rope, then we better have $q_2 \cong 0$. Then $\mathcal{L}$ in the old way

$$T = \frac{1}{2}M_1\dot{x}_1^2 + \frac{1}{2}M_2\dot{x}_2^2 = \frac{1}{2}(M_1 + M_2)\dot{x}_1^2$$

$$V = -M_1gx_1 - M_2gx_2 = g(M_2 - M_1)x_1 - M_2gl$$

notice (1) we can ignore $-M_2gl$, for constant in V makes no physical meaning.

(2) the two minus in $-M_1gx_1$ and $-M_2gx_2$ are important. They give the direction of gravity, which is down. That is consistent with the direction of $x_{1,2}$ axes chosen.

So

$$\mathcal{L} = \frac{1}{2}(M_1 + M_2)\dot{x}_1^2 - g(M_2 - M_1)x_1$$

EOM gives

$$\ddot{x}_1 = g\frac{M_1 - M_2}{M_1 + M_2}$$

Notice in the calculation, tension is not considered. Unless one is interested in things like whether the rope is going to break.

In the following example, let's analyze constraints more explicitly.

Lecture 3 (1/29/13)

Example. Consider mass M is hung from a rigid rod l_0.

Let us imagine the rod is replaced by some very stiff spring with $k \to \infty$, so there are two degrees of freedom $q_1 = \theta$, $q_2 = r - l_0$. Then we compute

$$T = \frac{1}{2}M(\dot{r}^2 + r^2\dot{\theta}^2) = \frac{1}{2}M\dot{q}_2^2 + \frac{1}{2}M(l_0 + q_2)^2\dot{q}_1^2$$

$$V = -Mg(l_0 + r)\cos q_1 + \frac{1}{2}kq_2^2$$

In q_2 component EOM gives

$$M\ddot{q}_2 - M(l_0 + q_2)\dot{q}_1^2 - Mg\cos q_1 + kq_2 = 0$$

$k \to \infty$, ignore other terms

$$M\ddot{q}_2 + kq_2 = 0$$

so q_2 oscillates rapidly.

In q_1 component EOM gives

$$M\frac{d}{dt}(l_0 + q_2)^2\dot{\theta} + Mg(l_0 + q_2)\sin\theta = 0$$

if we take $q_2 = 0$, we have the usual

$$Ml_0^2\ddot{\theta} + Mgl_0\sin\theta = 0$$

This will finish our discussion of constraints. Notice we skip Lagrangian multiplier, because it is used for constraints that have velocity dependence.

1.4 The Action Principle

It is also called "least action principle" or "variational principle".

Reference: Goldstein chapter 2, Landau Lipschitz chapter 1.

Background: Lagrangian system with n unconstrained degree of freedom

$$q = \{q_1, ..., q_n\}$$

then Euler-Lagrange equations are equivalent to a minimization problems.

Let's introduce Calculus of Variations

A *functional* is a function of a function. Recall function is to send a number to a number; functional is to send a function to a number. e.g. length of a curve is a functional of the function defining curves.

Calculus of variation deals with minimizing functional over the space of all functions

Example. "Brachistochrone" (shortest time)

Given two points A, B, find a curve that particle releases from A with 0 speed, moves along under the influence of gravity, gives the shortest time to B.

Let $y(x)$ to be a curve connecting A, B, then the functional

$$\begin{aligned} T_{A\to B}[y] &= \int dt \\ &= \int_A^B \frac{dl}{v} \\ &= \int_{x_A}^{x_B} \frac{\sqrt{1+y'^2(x)}dx}{\sqrt{2(y_A - y(x))g}} \end{aligned}$$

(We use square bracket for the argument of a functional.)

This problem has two fixed boundary conditions, which is important to recognize for later variation computation.

In general, the problem goes down to solve

$$S[y] = \int_{x_A}^{x_B} dx F(y(x), y'(x))$$

where F is a function. We want to minimize $S[y]$ over all possible functions $y(x)$ with fixed boundary conditions

$$y(x_A) = y_A \quad y(x_B) = y_B$$

The *necessary condition* for $y(x)$ to be a solution:
$S[y]$ should be stationary, i.e. its value should not change under

$$y(x) \to y(x) + \delta y(x) \tag{1.12}$$

at 1st order in $\delta y(x)$.

Note: (1) This of course makes sense. Recall for ordinary function $f(x)$, the necessary condition for $f(x)$ to have an extreme at x_0 is $f'(x_0) = 0$, which is equivalent to

$$f(x_0 + \delta x) = f(x_0) + f'(x_0)\delta x = f(x_0)$$

at 1st order in δx.

(2) in (1.12), δy is arbitrary but with $\delta y(x_A) = 0$, $\delta y(x_B) = 0$, so no variance at the boundaries.

Now let's do variation of S under (1.12)

$$\begin{aligned}
0 = \delta S &\equiv S[y+\delta y] - S[y] \\
&= \int dx[F(y(x)+\delta y(x), y'(x)+\delta y'(x)) - F(y(x), y'(x))] \\
&= \int dx[F(y,y') + \frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y'}\delta y' - F(y,y')] \text{ (1st order)} \\
&= \int dx \frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y'}\delta y\Big|_{x_A}^{x_B} - \int dx \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\delta y \text{ (integration by parts)} \\
&= \int dx \left(\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right)\delta y \text{ (vanishing at boundary)}
\end{aligned}$$

δy is arbitrary,

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) - \frac{\partial F}{\partial y} = 0 \tag{1.13}$$

so it looks just like E-L, if we are brave enough to think

Lecture 4
(1/31/13)

$$x \leftrightarrow t$$

$$y(x) \leftrightarrow q(t)$$

$$y'(x) \leftrightarrow \dot{q}(t)$$

$$F \leftrightarrow \mathcal{L}$$

Let utilize the discussion of last lecture. Modify S from last lecture, define "Action" of 1 degree of freedom

$$S[q] = \int_{t_A}^{t_B} \mathcal{L}(q(t), \dot{q}(t))dt$$

therefore minimizing S (or more properly extremizing S) gives E-L equations with fixed boundary conditions

$$\begin{cases} q(t_1) = q_1 \\ q(t_2) = q_2 \end{cases}$$

Theorem. *(Action Principle) a mechanical system evolves in time in such a way as to minimize S.*

Use action and variation is the more economical way of describing mechanical system. In string theory, one has similar definition, action of a string.

Notice: (1) we can define "Action" of n degree of freedom

$$S[q_1, ..., q_n] = \int_{t_A}^{t_B} \mathcal{L}(q(t), \dot{q}(t))dt \tag{1.14}$$

we require δq_α is arbitrary (like before) and independent (so we can fix other δq_α to be constant 0 and only vary one δq_α, this will give n independent E-L equations) But this in turns requires that $q_1, ..., q_n$ are the unconstrained coordinates and $\mathcal{L}$ is essentially the effective Lagrangian we studied before, basically treating constrained coordinates to be 0.

(2) We didn't put explicit t dependence in $\mathcal{L}$ in (1.14), that does not mean $\mathcal{L}$ has no explicit t dependence. It just uses as a convention. What it means is that

we want t not to affect the action principle, i.e.

$$\begin{cases} q_\alpha(t) \to q_\alpha(t) + \delta q_\alpha(t) \\ t \to t \end{cases}$$

meaning we deviate path, but not to deviate time, so no partial in t in the Taylor expansion.

(3) The fixed boundaries used here can be replaced by

$$\begin{cases} q(t_1) = q_1 \\ \dot{q}(t_1) = \dot{q}_1 \end{cases}$$

In fact the latter one gives unique solution, by the theory of existence and uniqueness of ode, while as the formal boundary condition doesn't give unique solutions. Notice the solution $q(t)$ of stationary point is almost always a saddle point rather than a minimum. E.g. in Quantum Tunneling.

Theorem. *For any $f(q,t)$, $\mathcal{L}(q,\dot{q};t)$, one can add*

$$\mathcal{L}'(q,\dot{q};t) = \mathcal{L}(q,\dot{q};t) + \frac{d}{dt}f$$

both $\mathcal{L}$, $\mathcal{L}'$ yield the same EL equations.

Hence $\mathcal{L}$, $\mathcal{L}'$ are physically equivalent.

Proof.

$$S' = \int_{t_1}^{t_2} dt\mathcal{L}' = \int_{t_1}^{t_2} dt\mathcal{L} + f(q(t),t)|_{t_1}^{t_2} = S + \text{const}$$

($f(q(t),t)|_{t_1}^{t_2}$ is a constant because we fix end points, but we don't specify the velocities at end points, so f must not be a function of $\dot{q}$.) Then it is clear that to minimize S' is same as to minimize S. □

Another Proof.

Proof.

$$\begin{aligned}
\mathrm{EL}(\mathcal{L}') - \mathrm{EL}(\mathcal{L}) &= \frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}}\frac{d}{dt}f\right) - \frac{\partial}{\partial q}\frac{d}{dt}f \\
&= \frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}}\left(\frac{\partial f}{\partial q}\dot{q} + \frac{\partial f}{\partial t}\right)\right) - \frac{\partial}{\partial q}\left(\frac{\partial f}{\partial q}\dot{q} + \frac{\partial f}{\partial t}\right) \\
&= \frac{d}{dt}\frac{\partial f}{\partial q} - \frac{\partial^2 f}{\partial q^2}\dot{q} - \frac{\partial^2 f}{\partial q \partial t} \\
&= \frac{\partial^2 f}{\partial q^2}\dot{q} + \frac{\partial^2 f}{\partial q \partial t} - \frac{\partial^2 f}{\partial q^2}\dot{q} - \frac{\partial^2 f}{\partial q \partial t} \\
&= 0
\end{aligned}$$

Lagrangian formalism freeze out intuition of physical meaning, so is the extra term added to $\mathcal{L}$. □

Example. Suppose we look at particle in 1D with potential $V(x)$

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2 - V(x)$$

we can add $f = \frac{1}{3}x^3$ to it

$$\mathcal{L}' = \mathcal{L} + \dot{x}x^2$$

both $\mathcal{L}$, $\mathcal{L}'$ should give the same EOM, but the physical meaning of $\dot{x}x^2$ is unclear.

Example. Another example free particle in 1D

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2$$

if we add $f = mxv_0 + \frac{1}{2}mv_0^2 t$, v_0 is some constant,

$$\begin{aligned}
\mathcal{L}' &= \mathcal{L} + m\dot{x}v_0 + \frac{1}{2}mv_0^2 \\
&= \frac{1}{2}m(\dot{x} + v_0)^2
\end{aligned}$$

This also leaves EOM invariant, and this has physical meaning: the extra term added corresponds to *Galilean Boost.*

1.5 Symmetry and Conservation Laws

Definition. A *symmetry* is a transformation of the q's, of the $\dot{q}$'s, and of t, that leaves the value of $\mathcal{L}$ unaltered.

Recall q_α is *cyclic* if $\partial\mathcal{L}/\partial q_\alpha = 0$.
If q is cyclic, of course the following transformation won't change $\mathcal{L}$

$$\begin{cases} q \to q + \epsilon \\ \dot{q} \to \dot{q} \\ t \to t \end{cases}$$

where we consider ϵ to be infinitesimal small (although not necessary here, will be useful later.) This transformation is a symmetry. Indeed

$$\begin{aligned} \delta\mathcal{L} &= \mathcal{L}(q+\epsilon, \dot{q}; t) - \mathcal{L}(q, \dot{q}; t) \\ &\approx \frac{\partial\mathcal{L}}{\partial q}\epsilon \\ &= 0 \end{aligned}$$

in the 1st order.

Continuous Symmetry

Lecture 5 (2/5/13)

Let's make the idea clear. Consider the following continuous transformation

$$\begin{cases} q_\alpha(t) \to q_\alpha(t) + \epsilon\gamma_\alpha(t) \\ \dot{q}_\alpha(t) \to \dot{q}_\alpha(t) + \epsilon\dot{\gamma}_\alpha(t) \\ t \to t \end{cases} \tag{1.15}$$

under which $\mathcal{L}$ is unchanged in the first order of ϵ, i.e.

$$\mathcal{L}(q + \epsilon\gamma, \dot{q} + \epsilon\dot{q}; t) = \mathcal{L}(q, \dot{q}; t) + O(\epsilon^2) \tag{1.16}$$

where ϵ some continuous infinitesimal parameter (not dependent on t), γ_α is some deformed path, and $\alpha = 1, 2, ..., n$ degrees of freedom.

Note: (1) for the following theorem (Noether) to hold, it suffices to talk about infinitesimal. Although we are more used to finite transformation, finite transformation and infinitesimal are not the same thing. One does not imply the other. (2) Not all transformations are continuous. E.g. parity, time reversal. Called discrete symmetries. One can either do it or not do it. (3) In (1.15) if $\epsilon = 0$, we have identity transformation.

Theorem. *(Noether) If* (1.15) *is a symmetry for some* $\{\gamma_\alpha\}_{\alpha=1,\dots,n}$*, then*

$$\sum_{\alpha=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \gamma_\alpha = const$$

Proof.

$$\mathcal{L}(q + \epsilon\gamma, \dot{q} + \epsilon\dot{\gamma}; t) = \mathcal{L}(q, \dot{q}; t) + \sum \epsilon(\frac{\partial \mathcal{L}}{\partial q_\alpha}\gamma_\alpha + \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\dot{\gamma}_\alpha)$$

compare with (1.16). Since $O(\epsilon^1) \gg O(\epsilon^2)$, we get

$$\epsilon \sum(\frac{\partial \mathcal{L}}{\partial q_\alpha}\gamma_\alpha + \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\dot{\gamma}_\alpha) = 0$$

that follows by Euler-Lagrange

$$\sum(\frac{d}{dt}(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha})\gamma_\alpha + \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\dot{\gamma}_\alpha) = 0$$

This is of course

$$\frac{d}{dt}\sum_{\alpha=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\gamma_\alpha = 0$$

□

Example. Euclidean translation

N particles in 3D interacting via a potential V only depends on the positions difference $(\vec{r}_a - \vec{r}_b)$ notice we did not say it depends on $|\vec{r}_a - \vec{r}_b|$. This is the case when there is no external field.

$$\mathcal{L} = \sum_{a=1}^{N} \frac{1}{2} m_a \dot{\vec{r}}_a^2 - V$$

so the overall translations of the system is a symmetry.

Take an arbitrary $\hat{n}$, $\mathcal{L}$ is unchanged under

$$\begin{cases} \vec{r}_a \to \vec{r}_a + \epsilon \hat{n} \\ \dot{\vec{r}}_a \to \dot{\vec{r}}_a \\ t \to t \end{cases}$$

so by Noether

$$\sum_{a=1}^{N} \sum_{i=1}^{3} \frac{\partial \mathcal{L}}{\partial \dot{r}_{a,i}} n_i = \sum_{i=1}^{3} n_i \sum_{a=1}^{N} m_a \dot{\vec{r}}_{a,i} = \hat{n} \cdot \vec{P} = \text{const} \tag{1.17}$$

where $\vec{P}$ is total momentum. Since $\hat{n}$ is arbitrary, $\vec{P}$ is constant. Hence in the absence of external field, total momentum is conserved.

If $\mathcal{L}$ is unchanged only under some specific direction, then we need to take $\hat{n}$ to be in that direction. So in other directions $\hat{n}$ are 0, hence in (1.17), i is not from 1 to 3, but we can still get that $\vec{P}$ is constant in the direction of $\hat{n}$.

The example is an infinitesimal translation. In case of electrons in a solid (Bloch model) there is a finite translation symmetry i.e. translated from one lattice unit to another. But that is not an infinitesimal translation, so we don't have momentum conversation for Bloch.

Note: there is a summation in the Noether theorem. This comes from Taylor expansion and chain rule. This makes big difference with Euler-Lagrange equation. There is no summation in E-L. E-L is correct for each component of α. But for Noether, only the sum is constant, hence each individual particle's momentum is not necessarily conserved but the total momentum is conserved.

Example. Euclidean Rotation

If $V = V(|\vec{r}_a - \vec{r}_b|, |\vec{r}_a - \vec{r}_c|, ...)$, we have rotational symmetry. Note the difference in the argument of V with the previous example, so that translation symmetry is not a rotational symmetry.

Suppose the reference has been chosen, and take $\hat{n}$ to be the axis of rotation that goes through the origin. Then the infinitesimal displacement

$$\epsilon\vec{\gamma}_a = \epsilon\hat{n} \times \vec{r}_a$$

where ϵ is the angle of rotation of the perpendicular component of $\vec{r}_a$ with respect to $\hat{n}$. That is because

$$|\epsilon\vec{\gamma}_a| = \epsilon r \sin \angle(\hat{n}, \vec{r}_a)$$

Then the transformation is

$$\begin{cases} \vec{r}_a \to \vec{r}_a + \epsilon\hat{n} \times \vec{r}_a \\ \dot{\vec{r}}_a \to \dot{\vec{r}}_a + \epsilon\hat{n} \times \dot{\vec{r}}_a \\ t \to t \end{cases}$$

(the velocity part of the transformation is actually taken from the definition of velocity. The definition of velocity says velocity vector is a transformation of the position vector under rotational transformation.)

By Noether

$$\begin{aligned} \sum_{a=1}^{N}\sum_{i=1}^{3} \frac{\partial \mathcal{L}}{\partial \dot{r}_{a,i}} (\hat{n} \times \vec{r}_a)_i &= \sum_{a=1}^{N} m_a \dot{\vec{r}}_a \cdot (\hat{n} \times \vec{r}_a) \\ &= \hat{n} \cdot \sum_{a=1}^{N} (\vec{r}_a \times m_a \dot{\vec{r}}_a) \\ &= \hat{n} \cdot \vec{L} = \text{const} \end{aligned}$$

So use the same argument before $\vec{L}$ is conserved, or if $\mathcal{L}$ is only invariant under certain direction, then let $\hat{n}$ be in that direction, then $\vec{L}$ is conserved in that direction.

Similarly if one has finite rotation symmetry, then one doesn't get conservation of angular momentum.

Example. Time translation

The case can be realized as doing experiment today yields the same result as yesterday. Because of the universe expansion, time translation is not exactly possible, but in order of ten billion years so we will ignore this, but if one studies early universe, one will see there is no well-defined energy concept.

Consider the time translation

$$\begin{cases} q_\alpha(t) \to q_\alpha(t) \\ \dot{q}_\alpha(t) \to \dot{q}_\alpha(t) \\ t \to t + \epsilon \end{cases}$$

then

$$\begin{aligned} \delta\mathcal{L} &= \mathcal{L}(q_\alpha(t), \dot{q}_\alpha(t); t + \epsilon) - \mathcal{L}(q_\alpha(t), \dot{q}_\alpha(t); t) \\ &= \epsilon\frac{\partial \mathcal{L}}{\partial t} + O(\epsilon^2) \end{aligned}$$

This is a continuous symmetry iff

$$\frac{\partial \mathcal{L}}{\partial t} = 0$$

Consider the total time derivative

$$\frac{d\mathcal{L}}{dt} = \frac{\partial \mathcal{L}}{\partial t} + \sum_\alpha (\frac{\partial \mathcal{L}}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\ddot{q}_\alpha)$$

Using E-L, one gets

$$\frac{d\mathcal{L}}{dt} = \frac{\partial \mathcal{L}}{\partial t} + \frac{d}{dt}\sum_\alpha \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\dot{q}_\alpha \tag{1.18}$$

We now define the *Hamiltonian* of the system

$$H \equiv \sum_\alpha \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\dot{q}_\alpha - \mathcal{L} \tag{1.19}$$

then (1.18) says

$$\frac{dH}{dt} = -\frac{\partial \mathcal{L}}{\partial t}$$

so if $\mathcal{L}$ doesn't depend on t explicitly, H is conserved.

Note: although it is natural to think time translation invariant as

$$\begin{cases} q_\alpha(t) \to q_\alpha(t+\epsilon) \\ \dot{q}_\alpha(t) \to \dot{q}_\alpha(t+\epsilon) \\ t \to t+\epsilon \end{cases}$$

with $\mathcal{L}$ unchanged, then

$$\begin{aligned} \delta\mathcal{L} &= \mathcal{L}(q_\alpha(t+\epsilon), \dot{q}_\alpha(t+\epsilon); t+\epsilon) - \mathcal{L}(q_\alpha(t), \dot{q}_\alpha(t); t) \\ &= \epsilon\frac{\partial\mathcal{L}}{\partial t} + \epsilon\sum_\alpha(\frac{\partial\mathcal{L}}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial\mathcal{L}}{\partial\dot{q}_\alpha}\ddot{q}_\alpha) + O(\epsilon^2) \end{aligned}$$

This would say time translation is a continuous symmetry iff

$$\frac{\partial\mathcal{L}}{\partial t} + \sum_\alpha(\frac{\partial\mathcal{L}}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial\mathcal{L}}{\partial\dot{q}_\alpha}\ddot{q}_\alpha) = 0$$

but this would just be $d\mathcal{L}/dt = 0$. This would not give conservation of H. Recall from fluid dynamics, total time derivative means following the particles, so it is not the same as simply time translation: compare same experiment was done today and yesterday.

So in accordance with our definition, we will just require $\partial\mathcal{L}/\partial t = 0$. But where does the explicit time dependence come from? There are two places: from potential that depends explicitly on t; and from the generalized coordinate q_α that is a moving coordinate wrt usual coordinates $\vec{r}$. So we eliminate these two cases. Soon we will see that we need another assumption that V only depends on $\vec{r}$. They are the most common situations.

We now claim if $\partial\mathcal{L}/\partial t = 0$, then H is the total energy, and so energy is conserved. Proof we start from $\vec{r} = \vec{r}(q)$ (note no t in $\vec{r}$),

$$\frac{d}{dt}\vec{r}_a = \sum_\beta \frac{\partial\vec{r}_a}{\partial q_\beta}\dot{q}_\beta \tag{1.20}$$

So

$$T = \sum_a^N \frac{1}{2}m_a\dot{\vec{r}}_a^2 = \sum_{\beta\gamma}\frac{1}{2}M_{\beta\gamma}(q)\dot{q}_\beta\dot{q}_\gamma \tag{1.21}$$

where $M_{\beta\gamma}(q)$ is symmetric mass tensor given by (1.20). Why is $M_{\beta\gamma}(q)$ symmetric? Suppose not, then we can write $M_{\beta\gamma}$ as sum of symmetric matrix and antisymmetric matrix. Since antisymmetric matrix makes no contribution to the sum in Kinetic energy, so we can assume $M_{\beta\gamma}$ is symmetric.

So by the definition (1.19) and assume $V = V(\vec{r}) = V(q)$

$$\begin{aligned}
H &= \sum_{\alpha} \frac{\partial T}{\partial \dot{q}_\alpha} \dot{q}_\alpha - T + V \\
&= \sum_{\alpha} \dot{q}_\alpha \frac{\partial}{\partial \dot{q}_\alpha} \sum_{\beta\gamma} \frac{1}{2} M_{\beta\gamma}(q) \dot{q}_\beta \dot{q}_\gamma - T + V \\
&= \sum_{\alpha} \sum_{\beta} M_{\beta\alpha} \dot{q}_\beta \dot{q}_\alpha - T + V \\
&= 2T - T + V = E
\end{aligned}$$

Example. 2D particle in central potential

$$\mathcal{L} = \frac{1}{2} m \dot{r}^2 + \frac{1}{2} m r^2 \dot{\phi}^2 - V(r)$$

So by the definition (1.19),

$$H = \frac{\partial \mathcal{L}}{\partial \dot{r}} \dot{r} + \frac{\partial \mathcal{L}}{\partial \dot{\phi}} \dot{\phi} - \mathcal{L} = m\dot{r}^2 + m r^2 \dot{\phi}^2 - \mathcal{L} = T + V$$

Generalized Noether

Lecture 6 (2/7/13)

Last time we showed translation, rotation, time translation symmetries give three conservation laws. Later when we do Hamilton formulation, we will show conservation laws give symmetries. For now if one wants to include Galilean transformation, one needs to work with generalized Noether.

Suppose we have a transformation like (1.15), such that

$$\mathcal{L}(q, \dot{q}; t) \to \mathcal{L}(q + \epsilon\gamma, \dot{q} + \epsilon\dot{\gamma}; t) = \mathcal{L}(q, \dot{q}; t) + \epsilon \frac{d}{dt} F(q; t) + O(\epsilon^2) \tag{1.22}$$

This doesn't leave $\mathcal{L}$ unchanged but this gives same equations of motion.

Theorem. *(general Noether) If* (1.15) *satisfies* (1.22), *then*

$$\sum_\alpha \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\gamma_\alpha - F(q;t) = const$$

Proof. Same steps as last time

$$\mathcal{L}(q+\epsilon\gamma, \dot{q}+\epsilon\dot{\gamma}; t) = \mathcal{L}(q,\dot{q};t) + \epsilon\sum_\alpha\left(\frac{\partial \mathcal{L}}{\partial q_\alpha}\gamma_\alpha + \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\dot{\gamma}_\alpha\right) + O(\epsilon^2)$$

so

$$\sum_\alpha\left(\frac{\partial \mathcal{L}}{\partial q_\alpha}\gamma_\alpha + \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\dot{\gamma}_\alpha\right) = \frac{d}{dt}F(q;t)$$

By E-L

$$\frac{d}{dt}\left(\sum_\alpha \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\gamma_\alpha - F(q;t)\right) = 0$$

□

Example. Galilean transformation

Consider

$$\mathcal{L} = \frac{1}{2}\sum_a^N m_a\dot{\vec{r}}_a^2 - V(\vec{r}_a - \vec{r}_b)$$

(this of course implies $\mathcal{L}$ is translational invariant, so $\vec{P}$ is conserved) and in addition the transformation

$$\begin{cases} \vec{r}_a \to \vec{r}_a + \epsilon\hat{n}t \\ \dot{\vec{r}}_a \to \dot{\vec{r}}_a + \epsilon\hat{n} \\ t \to t \end{cases}$$

Then

$$\begin{aligned} \delta\mathcal{L} &= \frac{1}{2}\sum_a^N m_a(\dot{\vec{r}}_a + \epsilon\hat{n})^2 - \frac{1}{2}\sum_a^N m_a\dot{\vec{r}}_a^2 \\ &= \epsilon\sum_a^N m_a\dot{\vec{r}}_a\cdot\hat{n} + O(\epsilon^2) \\ &= \epsilon\frac{d}{dt}F(\vec{r}) + O(\epsilon^2) \end{aligned}$$

$F(\vec{r}) = \sum_a^N m_a \vec{r}_a \cdot \hat{n} = M_{tot}\vec{X}_{cm} \cdot \hat{n}$

Noether says

$$(\sum_a^N \frac{\partial \mathcal{L}}{\partial \dot{\vec{r}}} t - M_{tot}\vec{X}_{cm}) \cdot \hat{n} = \text{const}$$

Since $\sum_a^N \frac{\partial \mathcal{L}}{\partial \dot{\vec{r}}} = \vec{P}$ is a constant, and $\hat{n}$ is arbitrary, we get

$$\vec{X}_{cm}(t) = \vec{X}_0 + \frac{\vec{P}}{M_{tot}} t$$

hence the center of mass is moving at constant velocity.

Note: although in our derivation, we get Galilean symmetry from $\mathcal{L}$ that has translational symmetry. But there are cases where translation symmetry exists but not Galilean. E.g. in fluid mechanics, one cannot defined center mass position, and there is actually a preferred reference frame.

2 Kepler's Problem

2.1 Single Particle Motion in 1D

$$\mathcal{L} = \frac{1}{2} m \dot{x}^2 - V(x)$$

Clearly E is conserved ($\because$ $\mathcal{L}$ satisfies the three assumptions last lecture).

$$H = \frac{\partial \mathcal{L}}{\partial \dot{x}} \dot{x} - \mathcal{L} = m\dot{x}^2 - \mathcal{L} = \frac{1}{2} m \dot{x}^2 + V(x) = E$$

So

$$\dot{x} = \sqrt{\frac{2(E - V(x))}{m}} \tag{2.1}$$

(note: we only keep + square root, the − square root gives same $x(t)$ solution but with time reversal, or one moves counterclockwise, the other clockwise.)

From (2.1), we get

$$t = \int \frac{dx}{\sqrt{\frac{2(E-V(x))}{m}}} \tag{2.2}$$

Now if we are able to integrate it, we will get $t(x)$, and if we are able to invert it, we will get $x(t)$.

Example. Harmonic Oscillator

$$V = \frac{1}{2}kx^2$$

then by (2.2), we get

$$t = \sqrt{\frac{m}{k}} \arcsin \sqrt{\frac{k}{2E}} x + t_0$$

so

$$x(t) = \sqrt{\frac{2E}{k}} \sin w_0(t - t_0)$$

where $w_0 = \sqrt{k/m}$. One can easily verify that E in the equation gives correct energy. Compute

$$\frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2 = \frac{1}{2}m\frac{2E}{k}w_0^2 \cos^2 + \frac{1}{2}k\frac{2E}{k} \sin^2 = E$$

An arbitrary $V(x)$ may cause (2.2) unsolvable. It has solution iff $E - V(x) \geq 0$. If $E = V(x_0) = V_{local\,min}$, particle stuck at x_0. There are also bound solution (related to bound orbits we will see later), and unbound solutions if $E \geq V(x)$ for $x \to \infty$ or $-\infty$.

2.2 2-Body Problem in Central Potential

2 particle in 3D, $V = V(\vec{r}_2 - \vec{r}_1)$,

$$\mathcal{L} = \frac{1}{2}m_1\dot{\vec{r}}_1^2 + \frac{1}{2}m_2\dot{\vec{r}}_2^2 - V(\vec{r}_2 - \vec{r}_1)$$

we know from last lecture, $\vec{R}_{cm}$ moves in simple fashion, so this motivates to use CM frame, let

$$\vec{R} = \frac{\vec{r}_1 m_1 + \vec{r}_2 m_2}{m_1 + m_2} \quad \vec{r} = \vec{r}_2 - \vec{r}_1$$

$\vec{r}$ is the relative position, it gives vector from m_1 to m_2. Later we will consider a specific example where m_1 is the sum and m_2 is the earth, so making $\vec{r}_2 - \vec{r}_1$ to be

$\vec{r}$ is very natural. Inverting, we get

$$\begin{cases} \vec{r}_1 = \frac{-m_2}{m_1+m_2}\vec{r} + \vec{R} \\ \vec{r}_2 = \frac{m_1}{m_1+m_2}\vec{r} + \vec{R} \end{cases}$$

$$\therefore \mathcal{L} = \underbrace{\frac{1}{2}(m_1 + m_2)\dot{\vec{R}}^2}_{\mathcal{L}_{cm}} + \underbrace{\frac{1}{2}\frac{m_1 m_2}{m_1 + m_2}\dot{\vec{r}}^2 - V(\vec{r})}_{\mathcal{L}_{rel}} \tag{2.3}$$

Set $\mu \equiv \frac{m_1 m_2}{m_1+m_2}$ reduced mass. It is rightly named, because it has unit of mass, and it is smaller than both masses, and when $m_1 \gg m_2$, $\mu = m_2$.

Lecture 7 (2/12/13)

It is clear to see from that (1) $\vec{R}$ is cyclic, so $\vec{P}_{cm}$ is constant; (2) $\mathcal{L}$ is the sum of two Lagrangian that depend on two separated sets of variables. So we can solve two parts separated, then add the solutions together. This technique used heavily in particle physics. Since $\mathcal{L}_{cm}$ is very simple, let's concentrate on $\mathcal{L}_{rel}$ and just write $\mathcal{L}$ for $\mathcal{L}_{rel}$.

If further more we have central potential $V = V(|\vec{r}|)$, it gives rotational symmetry, so $\vec{L}$ is constant. Since

$$\vec{r} \cdot \vec{L} = \vec{r} \cdot (m\vec{r} \times \vec{v}) = 0$$

so $\vec{r}$ is always in a fixed plane perpendicular to $\vec{L}$. Therefore the motion is in 2D.

Choose $\hat{z} \parallel \vec{L}$ (so $L > 0$), in spherical coordinates $\theta = \pi/2$, $\dot{\theta} = 0$.

$$\mathcal{L} = \frac{1}{2}\mu\dot{r}^2 + \frac{1}{2}\mu r^2\dot{\phi}^2 - V(r)$$

Equations of motion

$r:\quad \mu\ddot{r} - \mu r\dot{\phi}^2 + V'(r) = 0$, where $-\mu r\dot{\phi}^2$ is called centrifugal force (fictitious force).

$\phi:\quad P_\phi = \frac{\partial\mathcal{L}}{\partial\dot{\phi}} = \mu r^2\dot{\phi} \equiv L \implies$

$$\dot{\phi} = \frac{L}{\mu r^2} \tag{2.4}$$

This says $\dot{\phi} > 0 \ \forall t$. Hence rotating in same direction. Plugging (2.4) into r component of equation of motion, yield

$$\mu\ddot{r} + \frac{d}{dr}\left(V(r) + \frac{1}{2}\frac{L^2}{\mu r^2}\right) = 0$$

if one can treat

$$V_{eff} = V(r) + \frac{1}{2}\frac{L^2}{\mu r^2} \tag{2.5}$$

where $\frac{1}{2}\frac{L^2}{\mu r^2}$ is centrifugal potential, we have a 1D problem

$$\mu\ddot{r} + \frac{d}{dr}V_{eff} = 0 \tag{2.6}$$

If we define "energy" as

$$E = \frac{1}{2}\mu\dot{r}^2 + V_{eff}$$

we have energy conservation. Later we will see that the two initial conditions: E will tell us which orbit we are on; L will tell us the sharp of V_{eff}.

Example. Gravitational potential or attractive Coulomb potential

Suppose

$$V(r) = -\frac{c}{r}$$

$c > 0$

$$V_{eff} = -\frac{c}{r} + \frac{1}{2}\frac{L^2}{\mu r^2}$$

Because r^{-2} term dies quicker and blows faster, near 0, $V_{eff} > 0$, and near ∞, $V_{eff} < 0$. And there is a minimum E_0, below which has no solution.

So if $E = E_0$, $r = r_0 \ \forall t$, then from (2.4)

$$\dot{\phi} = \frac{L}{\mu r_0^2}$$

so uniform circular orbits.

If $E_0 < E < 0$, we have bound orbits, but is the orbit closed or not? We will discuss this soon. If $E > 0$, unbound orbits

Example. 3D harmonic oscillator

Consider two masses connected by a spring.

$$V(r) = \frac{1}{2}kr^2$$

Here we have assume that the rest length of the spring is 0, so that we will have a nice central potential problem. We can too draw

$$V_{eff} = \frac{1}{2}kr^2 + \frac{1}{2}\frac{L^2}{\mu r^2}$$

which has a U shape. This means all orbits are bounded. This makes sense. This is guaranteed by the fact for $k > 0$ the spring will not extend to ∞.

2.3 Necessary Condition for Closed Orbits

It turns out the two examples given above are the only two possible potential $V(r)$ that bound orbits are closed for all L, E suitable. It is proved in Goldstein 3.6. We won't do that in class, but we will show a necessary condition for bound orbits that are very closed to the minimal orbit to be closed.

Let $r = r_0$ be at which $V(r_0) = E_0$ a local minimum of $V(r)$. Consider $r(t)$ deviates from r_0 slightly, so

$$\begin{cases} r(t) = r_0 + \delta r(t) \\ \phi(t) = \phi_0(t) + \delta\phi(t) \end{cases}$$

where $\dot{\phi}_0 = L/\mu r_0^2 \equiv w_0$

Lecture 8 (2/14/13)

This tells us immediately $\phi_0(t) = w_0 t$ periodic motion at frequency w_0. Suppose $\delta r(t)$ oscillates at some frequency w (yet to be determined). Then

$$\begin{aligned} \dot{\phi}(t) &= \dot{\phi}(r_0 + \delta r(t)) = \frac{L}{\mu(r_0 + \delta r(t))^2} \\ &= \frac{L}{\mu r_0^2} - 2\frac{L}{\mu r_0^3}\delta r(t) \end{aligned}$$

So

$$\phi(t) = w_0 t - 2\frac{L}{\mu r_0^3}\delta r(t) t + \text{const}$$

so clearly $\phi(t)$ has two parts: one constant rotation, the other is oscillation which has the same frequency w.

So after $r(t)$ finishes a couple of cycles, the oscillation part of $\phi(t)$ cancels out too, so only the constant rotation part makes actual motion, which has frequency w_0. In sum, it is clear now that orbit is closed iff w, w_0 are commensurate

$$\frac{w}{w_0} = \frac{m}{n} \quad m, n \in \mathbb{N} \tag{2.7}$$

Substitute $r(t) = r_0 + \delta r(t)$ into (2.6), and use Taylor

$$\mu \delta \ddot{r} + V'_{eff}(r_0) + V''_{eff}(r_0)\delta r \cong 0$$

Since $V'_{eff}(r_0) = 0$, we get SHO, and use (2.5),

$$\begin{aligned} w^2 &= \frac{V''_{eff}(r_0)}{\mu} = \frac{V''(r_0)}{\mu} + \frac{d^2}{dr^2} \frac{L^2}{2\mu^2 r^2}\bigg|_{r_0} \\ &= \frac{V''(r_0)}{\mu} + 3\frac{L^2}{\mu^2 r_0^4} = \frac{V''(r_0)}{\mu} + 3w_0^2 \end{aligned} \tag{2.8}$$

so combining (2.7), we gives the necessary condition

$$w^2 = \frac{V''(r_0)}{\mu} + 3w_0^2$$

2.4 Power Law Potential

$$V(r) = Ar^\alpha$$

$V'_{eff}(r_0) = 0 \implies V'(r_0) = \frac{L^2}{\mu r_0^3}$, and $V''(r) = \frac{\alpha-1}{r} V'(r)$ give

$$V''(r_0) = \frac{\alpha - 1}{r_0} \frac{L^2}{\mu r_0^3} = (\alpha - 1)\mu w_0^2$$

To satisfy (2.8), we need

$$w^2 = (\alpha + 2) w_0^2$$

or

$$\alpha = \frac{m^2}{n^2} - 2 \qquad m, n \in \mathbb{N}$$

They closed after $lcm(m, n) = mn$ cycles, for m, n are relative primes.

If $m = n = 1$, $V(r) \propto 1/r$; if $m/n = 2$, $V(r) \propto r^2$. They both give elliptic orbits. Question if they both give elliptic orbits why the ratio of m, n are different?

That is because for the gravitational potential, one mass is at the focus, the other is on the elliptic. So r is from focus to the ellipse. So after one cycle, both radial and angle complete the period in one cycle. For the 3D harmonic potential one mass is at the center of the ellipse, and the other is on the ellipse, so after angle part finishes one cycle, the radial part finishes two cycles. We will see these at the end of this lecture and next lecture.

We are now going to show the two potentials mentioned above have closed elliptic orbits for all bound orbits, not just the ones that are closed (within linear order) to the minimum. What we are not going to show is that these are the only potentials having this property. (see proof in Goldstein)

Proof for gravitational potential leaves to next lecture.

3D harmonic potential

We know for central potential 2-body problem is essentially 1D in r and 1D in ϕ problem. But now we want to use Cartesian coordinate. So

$$V = \frac{1}{2}k(x^2 + y^2)$$

$$\mathcal{L} = (\frac{1}{2}\mu\dot{x}^2 - \frac{1}{2}kx^2) + (\frac{1}{2}\mu\dot{y}^2 - \frac{1}{2}ky^2) \tag{2.9}$$

Hence we have 2 independent 1D harmonic oscillator.

Let $a = r_{max}$, hence $V(a) = E$ so at a no radial velocity. Choose $t = t_0$, $r(t_0) = a$. and align the x axis with the point $(r(t_0), \phi(t_0))$, so in our xy plane, $\phi(t_0) = 0$.

Since

$$\begin{cases} x(t) = r(t)\cos\phi(t) \\ y(t) = r(t)\sin\phi(t) \end{cases} \implies \begin{cases} \dot{x} = \dot{r}\cos\phi - r\sin\phi\dot{\phi} \\ \dot{y} = \dot{r}\sin\phi + r\cos\phi\dot{\phi} \end{cases}$$

use the initial conditions

$$\begin{cases} x(t_0) = a \\ y(t_0) = 0 \\ \dot{x}(t_0) = 0 \\ \dot{y}(t_0) = a\dot{\phi}(t_0) = \frac{L}{\mu a} \end{cases}$$

and the ansatz to (2.9)

$$\begin{cases} x(t) = A\cos(wt - \psi) \\ y(t) = B\cos(wt - \psi') \end{cases}$$

where $w = \sqrt{k/\mu}$, we obtain

$$\begin{cases} x(t) = a\cos w(t - t_0) \\ y(t) = \frac{L}{\mu a}\frac{1}{w}\sin w(t - t_0) \end{cases}$$

and it is easy to check $b = L/\mu a w$ gives r_{min}, so in this xy plane, $\vec{r}$ moves on an ellipse and the origin is the center of ellipse.

2.5 Kepler's Problem

$$V(r) = -\frac{k}{r} \quad k = Gm_1m_2 = G\mu M > 0$$

$$V_{eff}(r) = -\frac{k}{r} + \frac{L^2}{2\mu r^2}$$

From (2.2) 1D single particle, change V to V_{eff}, m to μ, we get

$$t(r) = \int \frac{\sqrt{\mu}dr}{\sqrt{2(E - V_{eff}(r))}} \tag{2.10}$$

or

$$dt = \frac{\sqrt{\mu}dr}{\sqrt{2(E - V_{eff}(r))}}$$

And use (2.4),

$$\phi = \int dt\dot{\phi} = \int \frac{\sqrt{\mu}dr}{\sqrt{2(E - V_{eff}(r))}}\frac{L}{\mu r^2} \tag{2.11}$$

Lecture 9 (2/19/13)

With changing of variables first set $\xi = 1/r$, then

$$\phi(\xi) \quad = \quad -\frac{L}{\sqrt{2\mu}} \int \frac{d\xi}{\sqrt{E + k\xi - \frac{L^2}{2\mu}\xi^2}}$$

Write $k\xi - \frac{L^2}{2\mu}\xi^2 = k^2\frac{\mu}{2L^2} - (\xi - \frac{k\mu}{L^2})^2\frac{L^2}{2\mu}$, and set $\xi' = \xi - \frac{k\mu}{L^2}$, then

$$\phi(\xi') = -\frac{L}{\sqrt{2\mu}} \int \frac{d\xi'}{\sqrt{(E + k^2\frac{\mu}{2L^2}) - \xi'^2\frac{L^2}{2\mu}}} = -\frac{L}{\sqrt{2\mu(E + k^2\frac{\mu}{2L^2})}} \int \frac{d\xi'}{\sqrt{1 - \frac{\frac{L^2}{2\mu}\xi'^2}{E+k^2\frac{\mu}{2L^2}}}}$$

So if we put $\xi'' = \sqrt{\frac{\frac{L^2}{2\mu}}{E+k^2\frac{\mu}{2L^2}}}\xi'$, (here ξ'' is guaranteed to be $(-1,1)$, because orbits are bounded), we get

$$\phi = -\int \frac{d\xi''}{\sqrt{1 - \xi''^2}} = -\arcsin\left(\sqrt{\frac{\frac{L^2}{2\mu}}{E + k^2\frac{\mu}{2L^2}}}(\frac{1}{r} - \frac{k\mu}{L^2})\right) + \text{const}$$

Choosing constant to be $\pi/2$, we can obtain

$$\cos\phi = \sqrt{\frac{\frac{L^2}{2\mu}}{E + k^2\frac{\mu}{2L^2}}}(\frac{1}{r} - \frac{k\mu}{L^2})$$

yielding

$$\frac{P}{r} = 1 + e\cos\phi \tag{2.12}$$

where

$$P = \frac{L^2}{k\mu}, \qquad e = \sqrt{\frac{2L^2E}{k^2\mu} + 1}$$

(2.12) describes a conic section of eccentricity e and of Latus rectum P. Because by definition a curve is a conic section iff the distance of the points on the curve to a fixed focus is r, and the distance of the points on the curve to a fixed line, called directrix, is r/e.

If we consider the point at which the line connecting the point on the curve to the focus is parallel to the directrix, and the connecting line has length P (because

$\phi = \pi/2$), then clearly the distance from the focus to the directrix, denote line x, will be P/e, then suppose any other point on the curve that is r distance from the focus and r/e distance from the directrix, and makes angle ϕ between line r and x. Then we must have

$$\frac{r}{e} + r\cos\phi = \frac{P}{e}$$

which is (2.12).

There are four cases: (for other E, gives no solutions)

1) $e = 0$ then $E_0 = -\frac{k^2\mu}{2L^2}$. Circular Orbit

We check it agrees with the argument given before. Circular orbit means $E_0 = \min V_{eff}$, hence

$$V'_{eff} = \frac{k}{r^2} - \frac{L^2}{\mu r^3} = 0 \implies r_0 = \frac{L^2}{\mu k}$$

$$E_0 = V_{eff}(r_0) = -\frac{k}{r_0} + \frac{L^2}{2\mu r_0^2} = -\frac{k^2\mu}{2L^2} \tag{2.13}$$

2) $0 < e < 1$ when $E_0 < E < 0$, elliptic orbit, i.e. orbits are closed.

Easy see from (2.12)

$$r_{min,max} = \frac{P}{1 \pm e}$$

3) $e = 1$ when $E = 0$. Parabolic orbits.

$r_{min} = P/2 = L^2/2k\mu$. This is barely unbound orbit, meaning that it goes to ∞ with 0 speed. $r_{max} \to \infty$ and the asymptotic direction, given by ϕ

$$1 + \cos\phi = 0$$

so $\phi = \pi$, hence the two asymptotic lines become parallel to x.

4) $e > 1$ when $E > 0$. Hyperbolic orbits. This is an interesting case of attracted scattering. From (2.12), we know the two asymptotic directions are given by

$$e\cos\phi_a = -1$$

where ϕ_a as usual is angle between the line from $r \to \infty$ to the focus and x axis (i.e. the line passing the focus and perpendicular to the directrix.). Since here

$r \to \infty$, we can also assume this is the same angle that the asymptote makes with the x axis. Then the angle between the two asymptotes is

$$\beta = \phi_a - (\pi - \phi_a) = 2 \arccos -\frac{1}{e} - \pi = 2 \arcsin \frac{1}{e} \tag{2.14}$$

β is call the deflection angle, because say a particle is coming in along one of the asymptotes then it will deflect by the mass sitting at the focus and coming out along the other asymptote, so β is the angle of change the direction of the particle.

We can simplify (2.14), if we consider a very energetic orbit, coming in particle with very high energy

$$E \gg |E_0|$$

where E_0 is in (2.13). So

$$e = \sqrt{1 + \frac{E}{|E_0|}} \gg 1$$

$$\beta \approx \frac{2}{e} \ll 1$$

Hence very energetic particle doesn't get deflected much. We can further parametrize this in term of r_{min}, called "impact parameter" (commonly denoted by b). $L = \mu r^2 \dot{\phi} = \mu r_{min} v_{at\, the\, min\, position}$, and $E = \frac{1}{2} \mu v_\infty^2$. Suppose

$$v_{at\, the\, min\, position} \approx v_\infty \tag{2.15}$$

then

$$e \approx \sqrt{\frac{E}{|E_0|}} = \frac{\mu r_{min} v_\infty^2}{k}$$

For gravity $k = GM\mu = G(m_1 + m_2)\frac{m_1 m_2}{m_1 + m_2} = Gm_1 m_2$,

$$\beta = \frac{2GM}{r_{min} v_\infty^2}$$

Since β doesn't depend on μ, we can safely consider light ($\mu = 0$) deflected by the sum, and the assumption (2.15) works very well for this case too, taking $v_\infty = c$.

Therefore

$$\beta = \frac{2GM_{sum}}{r_{min}c^2} \tag{2.16}$$

One can do the experiment, when a distance star, sum, moon, earth are collinear. The star is behind the sum, and we are in front of the sum, so we are not supposed to observe the star. Because of light deflect, we see the star with deflection angle β. We use radius of the sum to be r_{min} in (2.16). The present of the moon is just to block light directed from the sum, which is the reason why astronomical observation are made at nights.

So one can then verify (2.16). But experiments tell us that the factor 2 in (2.16) is not correct. It should be 4, but this requires general relativity to explain.

Lecture 10 (2/21/13)

We now consider the last case before we turn to summarize Kepler's three laws. Consider

$$V(r) = \frac{k}{r}$$

with $k > 0$. This is the repulsive case. If one draw V_{eff}, one will see there that $V_{eff} > 0$ and there is no minimum, so only unbounded orbits. One can goes through the same algebra before with replacing $k \to -k$, $P \to -P$, and $e \to e$. Therefore we have

$$-\frac{P}{r} = 1 + e\cos\phi \tag{2.17}$$

where P, e are defined last time, and $e > 1$, because E has to be > 0 otherwise no solution.

Let's see what (2.17) looks like in Cartesian coordinate, $r = \sqrt{x^2+y^2}$, $x = r\cos\phi$, hence choose the heavy object to be the origin, then (2.17) becomes

$$-P = \sqrt{x^2+y^2} + ex$$

We can write it as, but only keep the one of the two branches, i.e. entirely $x < 0$,

$$\frac{(x-x_0)^2}{a^2} - \frac{y^2}{b^2} = 1 \tag{2.18}$$

where $x_0 = -\frac{eP}{e^2-1}$, $a^2 = \frac{P^2}{(e^2-1)^2}$, and $b^2 = \frac{P^2}{e^2-1}$ (here $e > 1$ so $b^2 > 0$. If we have $0 < e < 1$, $b^2 = \frac{P^2}{1-e^2}$ and we get ellipse.)

Recall for hyperbolic

$$c = \sqrt{a^2 + b^2} = \frac{eP}{e^2 - 1} = |x_0|$$

This shows that to obtain (2.18), we shift the usual hyperbola to the left until the right focus meets the origin, and keep the left branch only. We can also get r_{min} for the graph

$$r_{min} = a + c = \frac{P}{e - 1}$$

which can also be found by setting $\phi = \pi$ in (2.17).

Similarly the asymptote for the repulsive case is the same as $r \to \infty$

$$\cos \phi \to -\frac{1}{e}$$

2.6 Kepler Laws

Now we are back to attractive case $E < 0$ elliptic orbits. We already studied the first law. We can easily show Kepler's 2nd law, area sweeps by the line joining a planet and the sun in dt, using (2.4)

$$dA = \frac{1}{2} r^2 d\phi = \frac{1}{2} \frac{L}{\mu} dt$$

So

$$\frac{dA}{dt} = \text{const}$$

We now study the 3rd law, the time dependency. Ideally we want to find $r(t)$, $\phi(t)$, but practically it is easier to find

$$t(\psi),\ r(\psi),\ \phi(\psi)$$

From (2.10), we have

$$t = \int \frac{\sqrt{\mu} dr}{\sqrt{2(E + \frac{k}{r} - \frac{L^2}{2\mu r^2})}} = \sqrt{\frac{\mu}{2|E|}} \int \frac{r dr}{\sqrt{-r^2 + \frac{k}{|E|} r - \frac{L^2}{2\mu |E|}}}$$

Using $a = \frac{P}{1-e^2} = \frac{k}{2|E|}$,

$$t = \sqrt{\frac{\mu}{k}a} \int \frac{rdr}{\sqrt{-r^2 + 2ar + a^2(e^2-1)}}$$

Let $r - a = -ae\xi$, so $dr = -aed\xi$,

$$t = \sqrt{\frac{\mu}{k}a} \int \frac{a(e\xi - 1)d\xi}{\sqrt{1-\xi^2}} = \sqrt{\frac{\mu a^3}{k}}(-e\sqrt{1-\xi^2} - \arcsin\xi + \text{const})$$

choosing constant to be $\pi/2$, and let $\xi = \cos\psi$, we get

$$t = \sqrt{\frac{\mu a^3}{k}}(\psi - e\sin\psi)$$

showing t is an increasing function, because $t' \sim 1 - e\cos\psi > 0$, for $0 < e < 1$

And we find

$$r = a - ae\xi = a(1 - e\cos\psi)$$

So we start from the initial positive $t = 0$, $\psi = 0$, perihelia

$$r = a(1-e) = r_{min}$$

After $\psi = 2\pi$, $r = r_{min}$, showing ϕ and ψ have the same periodicity 2π. And the period of the motion

$$T = t(\psi = 2\pi) - t(\psi = 0) = 2\pi\sqrt{\frac{\mu a^3}{k}}$$

Or

$$T^2 = 4\pi^2 \frac{\mu}{k} a^3$$

which is Kepler's third law.

For the completion, let's find $\phi(\psi)$. From (2.12)

$$\cos\phi = \frac{1}{e}\left(\frac{a(1-e^2)}{a(1-e\cos\psi)} - 1\right) = \frac{\cos\psi - e}{1 - e\cos\psi}$$

so indeed ψ, ϕ have the same period 2π.

We have now completely solved 2 body problems in $\vec{R}$ and $\vec{r}$ coordinates. We want to come back to $\vec{R}_1$, $\vec{R}_2$ coordinates.

Choose reference frame where $\vec{R} = 0$ for all t, i.e. center mass frame, so

$$\vec{R}_1 = \vec{R} - \frac{m_2}{m_1 + m_2}\vec{r} = -\frac{\mu}{m_1}\vec{r}$$

$$\vec{R}_2 = \vec{R} + \frac{m_1}{m_1 + m_2}\vec{r} = \frac{\mu}{m_2}\vec{r}$$

Hence in the center mass frame, $\vec{R}_1$, $\vec{R}_2$ are orbiting in exactly the same way as $\vec{r}$ apart from an overall scaling about the center mass.

If $m_1 \gg m_2$, e.g. ratio of mass of earth to the sum is 10^{-1000}

$$\vec{R}_1 \approx \vec{R}, \qquad \vec{R}_2 \approx \vec{R} + \vec{r}$$

3 General Theory of Small Oscillations

Lecture 11 (2/26/13)

Reference Goldstein ch6, L.L. ch5.

3.1 General Solutions

In what follows we consider $\mathcal{L}$ has no explicit time dependence, i.e. constant energy. Later we will do adiabatic invariance, where we will allow E to change.

For system with n degrees of freedom, we say the system is in equilibrium position $q^{(0)} = (q_1^{(0)}, ..., q_n^{(0)})$ if

$$\left.\frac{\partial \mathcal{L}}{\partial q_\alpha}\right|_{q^{(0)}} = 0 \tag{3.1}$$

for all $\alpha = 1, ..., n$. Hence there is no force.

Apply EL to (1.21), for each α,

$$\frac{d}{dt}\left(\sum_\beta M_{\alpha\beta}(q)\dot{q}_\beta\right) - \sum_{\beta\gamma} \frac{1}{2}\frac{\partial M_{\beta\gamma}(q)}{\partial q_\alpha}\dot{q}_\beta\dot{q}_\gamma - \frac{\partial V}{\partial q} = 0$$

If the system starts out at $q^{(0)}$ which obeying (3.1) and with no initial velocity, then

$$\sum_\beta M_{\alpha\beta}(q)\dot{q}_\beta = \text{const} \implies \dot{q}^{(0)} = 0$$

so the system will stays at $q^{(0)}$.

Now consider some small departures from equilibrium

$$\begin{cases} q_\alpha = q_\alpha^{(0)} + \eta_\alpha(t) \\ \dot{q}_\alpha = \dot{\eta}_\alpha(t) \end{cases}$$

small departures means η_α, $\dot{\eta}_\alpha$ are small so that Taylor expansion is kept up to quadratic order. Why 2nd order? Because HO is quadratic, hence we can interpret our system as harmonic like system.

$$T(q^{(0)} + \eta) = \frac{1}{2}\sum_{\beta\gamma} M_{\beta\gamma}(q^{(0)} + \eta)\dot{\eta}_\beta\dot{\eta}_\gamma = \frac{1}{2}\sum_{\beta\gamma} M_{\beta\gamma}(q^{(0)})\dot{\eta}_\beta\dot{\eta}_\gamma + O(\eta^3)$$

Here because we only keep 2nd order, the mass tenor is treated as unperturbed. If we use $M_{\beta\gamma}(q^{(0)} + \eta)$ instead, we get into anharmonic system, which we will not do.

$$V(q^{(0)} + \eta) = V(q^{(0)}) + \sum_\beta \left.\frac{\partial V}{\partial q_\beta}\right|_{q^{(0)}} \eta_\beta + \frac{1}{2}\sum_{\beta\gamma} \left.\frac{\partial^2 V}{\partial q_\beta \partial q_\gamma}\right|_{q^{(0)}} \eta_\beta\eta_\gamma + O(\eta^3) \tag{3.2}$$

We drop the first term on the right because it has no physical significance, and drop the second term because it is 0.

Defining $\frac{\partial^2 V}{\partial q_\beta \partial q_\gamma} = K_{\beta\gamma}$, we get Lagrangian up to 2nd order

$$\mathcal{L} = \frac{1}{2}\sum M_{\beta\gamma}\dot{\eta}_\beta\dot{\eta}_\gamma - \frac{1}{2}\sum K_{\beta\gamma}\eta_\beta\eta_\gamma$$

Or written in vector form

$$\mathcal{L} = \frac{1}{2}\dot{\vec{\eta}}^T \hat{M}\dot{\vec{\eta}} - \frac{1}{2}\vec{\eta}^T \hat{K}\vec{\eta} \tag{3.3}$$

Apply EL with respect to generalized coordinate η, for each α

$$\frac{d}{dt}(\sum_\beta M_{\alpha\beta}\dot{\eta}_\beta) + \sum_\beta K_{\alpha\beta}\eta_\beta = 0$$

Since $M_{\alpha\beta}$, $K_{\alpha\beta}$ are functions of unperturbed $q^{(0)}$ only, we treat them as constants here, so

$$\sum_\beta (M_{\alpha\beta}\ddot{\eta}_\beta + K_{\alpha\beta}\eta_\beta) = 0$$

or

$$\hat{M} \cdot \ddot{\vec{\eta}} + \hat{K} \cdot \vec{\eta} = 0 \tag{3.4}$$

We look for oscillating solution

$$\vec{\eta} = \vec{a}e^{-iwt} \tag{3.5}$$

with constant vector $\vec{a}$, and we will take the real part at the end of calculation. (3.5) is in fact a good solutions, although it says the same w for each component α of $\vec{\eta}$. As we will see there are many possible w's, so the final solutions are linear combinations of them.

Plugging (3.5) into (3.4) gives

$$(-w^2\hat{M} + \hat{K})\vec{a} = 0 \tag{3.6}$$

hence we have an eigenvalue problem. One trivial solution is $\vec{a} = 0$, i.e. $\vec{\eta} = 0$, so system stays at equilibrium forever. For more interesting solutions, we want

$$\det(-w^2\hat{M} + \hat{K}) = 0$$

which gives nth order algebraic equations for w^2. There are no explicit formula for solving polynomials with $n > 3$, nevertheless we know there are n solutions, not necessary distinct.

Claim: all $w^2 \geq 0$, so w are real.

Proof: First show w^2 are real, multiply hermitian of $\vec{a}$ to (3.6)

$$\vec{a}^\dagger(\hat{K} - w^2\hat{M})\vec{a} = 0$$

then

$$w^2 = \frac{\vec{a}^\dagger\hat{K}\vec{a}}{\vec{a}^\dagger\hat{M}\vec{a}}$$

showing $(w^2)^\dagger = w^2 \implies w^2$ is real.

To show $w^2 \geq 0$, we recognize that $\hat{M}$, $\hat{K}$ are positive definite matrices. Since kinetic energy is positive for all classical allowed paths, and all small departures from minimum potential paths are classical allowed,

$$T = \sum_{\beta\gamma}\frac{1}{2}M_{\beta\gamma}\dot{\eta}_\beta\dot{\eta}_\gamma \geq 0 \quad \forall\vec{\eta}$$

showing $\hat{M}(q^{(0)})$ is positive definite.

The statement that $\hat{K}(q^{(0)})$ is positive definite is just to say $q^{(0)}$ is a local minimum for V. Because from (3.6),

$$V(q^{(0)} + \vec{\eta}) \geq V(q^{(0)})\ \forall\vec{\eta} \iff \hat{K} \text{ is positive definite}$$

3.2 Examples

Example. 1D suppose $V(x)$ has a peak at x_0, so x_0 is an unstable equilibrium, $k = V''(x_0) < 0$. So

$$w = \pm i\sqrt{\frac{|k|}{m}}$$

and

$$x(t) = x_0 + ae^{-iwt} = x_0 + Ae^{\sqrt{\frac{|k|}{m}}t} + Be^{-\sqrt{\frac{|k|}{m}}t}$$

Our model works only for small η, hence the exponential solution is not appropriate solutions. The B term is bad too, because for unstable equilibrium, we don't expect the system to come to equilibrium x_0 eventually. Therefore our discussion should restrict to $w \in \mathbb{R}$. In case if $V''(q^{(0)}) = 0$, we will have $k = 0$ so $w = 0$. We will study them later.

Now back to (3.6). Since $(\hat{K} - w^2\hat{M})$ is real, we can choose eigenvector $\vec{a}$ to be real as well. Moreover we will see that for different values of w, say w_1, w_2, and each has eigenvector $\vec{a}_1$, $\vec{a}_2$, then

$$0 = \vec{a}_1^T \cdot \hat{K} \cdot \vec{a}_2 - \vec{a}_1^T \cdot \hat{K} \cdot \vec{a}_2 = (w_1^2 - w_2^2)\vec{a}_1^T \cdot \hat{M} \cdot \vec{a}_2 \implies \vec{a}_1^T \cdot \hat{M} \cdot \vec{a}_2 = 0$$

Mathematically speaking, one can define a quadratic form with respect to the positive definite symmetric matrix $\hat{M}$, so that in this form the dot product of $\vec{a}_1$, $\vec{a}_2$ is zero, i.e. $\vec{a}_1$, $\vec{a}_2$ are orthogonal in this form.

Lecture 12 (2/28/13)

Example. Spring Pendulum (spring constant k, natural length l, mass m)

2 degrees of freedom r, ϕ. First find equilibrium, i.e. extreme of potential $V = \frac{1}{2}k(r-l)^2 - mgr\cos\phi$

$$\begin{cases} \frac{\partial V}{\partial r} = 0 \\ \frac{\partial V}{\partial \phi} = 0 \end{cases} \implies \begin{cases} \phi_0 = 0 \\ r_0 = l + \frac{mg}{k} \end{cases}$$

Hence the perturbation coordinates (this is related to what we will study later: normal coordinates) are

$$\eta_1 = \phi - \phi_0 = \phi \qquad \eta_2 = r - r_0 = \delta r$$

and find

$$\hat{K}(r_0, \phi_0) = \frac{\partial^2 V}{\partial q_\beta \partial q_\gamma} = \begin{pmatrix} \frac{\partial^2 V}{\partial \phi^2} & \frac{\partial^2 V}{\partial \phi \partial \gamma} \\ \frac{\partial^2 V}{\partial r \partial \phi} & \frac{\partial^2 V}{\partial r^2} \end{pmatrix} = \begin{pmatrix} mgr_0 & 0 \\ 0 & k \end{pmatrix}$$

Then find $\hat{M}$, since

$$\begin{aligned} T &= \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\phi}^2 \\ &= \frac{1}{2}m\delta\dot{r}^2 + \frac{1}{2}m(r_0 + \delta r)^2\dot{\phi}^2 \\ &= \frac{1}{2}m\delta\dot{r}^2 + \frac{1}{2}mr_0^2\dot{\phi}^2 \end{aligned}$$

In the last step we ignored δr as we did last time, because it only contributes to higher orders. Therefore

$$\hat{M}(r_0, \phi_0) = \begin{pmatrix} mr_0^2 & 0 \\ 0 & m \end{pmatrix}$$

then

$$\det(\hat{K} - w^2\hat{M}) = (mgr_0 - mr_0^2 w^2)(k - mw^2) = 0$$

so we get two modes: pendulum frequency

$$w^2 = \frac{g}{r_0}$$

and spring frequency

$$w^2 = \frac{k}{m}$$

3.3 Non-Degenerated Normal Modes

We now formally define quadratic form

$$(\vec{v}, \vec{w}) = \vec{v}^T \cdot \hat{M} \cdot \vec{w}$$

this has all properties of usual inner product, because $\hat{M}$ is real, symmetric, and positive definite. In the formulation $\vec{v}$, $\vec{w}$ are also given as real vectors.

We can also define the length of vector $\vec{v}$ using this quadratic form

$$\|\vec{v}\|^2 = (\vec{v}, \vec{v})$$

We can of course normalize the vector $\vec{a}$, in (3.6), namely

$$\|\vec{a}\| = 1$$

(as we will see this is important, because later $\mathcal{L}$ turns out to be $\frac{1}{2}\dot{Q}_\alpha^2 - \frac{1}{2}w_\alpha Q_\alpha^2$, called canonical normalization, which is also used in field theory). Of course after the realization and normalization, $\vec{a}$ is no long the same $\vec{a}$ in (3.5). In (3.5) $\vec{a}$ agrees the initial conditions, here $\vec{a}$ doesn't, so we need to impose initial condition through other means.

If we have all w^2 different as in the non-degenerated cases, then as we showed last time, eigenvectors associated different w^2 are orthogonal, i.e.

$$(\vec{a}_\alpha, \vec{a}_\beta) = \delta_{\alpha\beta}$$

Let us define a $n \times n$ matrix A,

$$\hat{A} = \begin{pmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & ... & \vec{a}_n \\ | & | & & | \end{pmatrix}$$

so $\hat{A}$ is orthogonal matrix, in the sense that

$$\hat{A}^T \hat{M} \hat{A} = \hat{I} \tag{3.7}$$

And we define

$$(\hat{w}^2) = \begin{pmatrix} w_1^2 & & \\ & \ddots & \\ & & w_n^2 \end{pmatrix}$$

then (3.6) becomes

$$\hat{K} \cdot \hat{A} = \hat{M} \cdot \hat{A} \cdot \hat{w}^2 \tag{3.8}$$

that is because for each α, the β component of LHS of (3.6) is $\sum_\gamma K_{\beta\gamma} A_{\gamma\alpha}$, and the β component of the RHS of (3.6) is $w_\alpha^2 \sum_\gamma M_{\beta\gamma} A_{\gamma\alpha} = \sum_{\gamma\mu} M_{\beta\gamma} A_{\gamma\mu} (\hat{w}^2)_{\mu\alpha}$.

Apply $\hat{A}^T$ to (3.8), and using (3.7), we get

$$\hat{A}^T \cdot \hat{K} \cdot \hat{A} = \hat{w}^2 \tag{3.9}$$

Mathematically this is saying $\hat{K}$, $\hat{w}^2$ can be diagonalized simultaneously by a congruence transformation.

Let us compare congruence transformations with similarity transformations. Let M be the similarity or congruence transformation and let $\hat{M}$ be the corresponding matrices for the reference basis. Let C be invertible basis change matrix.

Similarity Transformations $C^{-1}\hat{M}C$	Congruence Transformations $C^T\hat{M}C$
M is linear map	M is quadratic form
$M : \vec{v} \mapsto \hat{M}\vec{v}$	$M : \vec{v} \mapsto \vec{v}^T\hat{M}\vec{v}$
Change basis $\vec{v} \mapsto \vec{v}'$, $\vec{v} = C\vec{v}'$, then	
$M : \vec{v}' \mapsto C^{-1}\hat{M}C\vec{v}'$	$M : \vec{v}' \mapsto \vec{v}'^T C^T\hat{M}C\vec{v}'$

Clearly $C^{-1}\hat{M}C$, $C^T\hat{M}C$ are the new matrices wrt the new base for these two transformations.

Now let $\vec{\eta} = \hat{A}\vec{Q}$, then $\dot{\vec{\eta}} = \hat{A}\dot{\vec{Q}}$,

$$\vec{Q} = \begin{pmatrix} Q_1 \\ \vdots \\ Q_n \end{pmatrix}$$

It is commonly called "normal coordinates", because by (3.7) and (3.9), (3.3) becomes

$$\mathcal{L} = \frac{1}{2}\dot{\vec{Q}}^T\dot{\vec{Q}} - \frac{1}{2}\vec{Q}^T(\hat{w}^2)\vec{Q}$$

which only involves diagonal matrices. Lagrangian is now separable

$$\mathcal{L}(\vec{Q}, \dot{\vec{Q}}) = \sum_\alpha (\frac{1}{2}\dot{Q}_\alpha^2 - \frac{1}{2}w_\alpha^2 Q_\alpha^2) = \sum_\alpha \mathcal{L}_\alpha(Q_\alpha, \dot{Q}_\alpha) \tag{3.10}$$

showing there are n independent harmonic oscillations. The solution to each $\mathcal{L}_\alpha$ as given by EL is

$$Q_\alpha = C_\alpha e^{-iw_\alpha t}$$

where $C_\alpha = \rho_\alpha e^{i\phi_\alpha}$. Both amplitude ρ_α, and phase ϕ_α in C_α are real and they are determined by the initial conditions, i.e.

$$\hat{A}^T\hat{M}\begin{pmatrix} \eta_1(0) \\ \vdots \\ \eta_n(0) \end{pmatrix} = \begin{pmatrix} \rho_1\cos\phi_1 \\ \vdots \\ \rho_n\cos\phi_n \end{pmatrix} \text{ and } \hat{A}^T\hat{M}\begin{pmatrix} \dot{\eta}_1(0) \\ \vdots \\ \dot{\eta}_n(0) \end{pmatrix} = \begin{pmatrix} \rho_1 w_1\sin\phi_1 \\ \vdots \\ \rho_n w_n\sin\phi_n \end{pmatrix}$$

Assume no $w = 0$, so we can fully solve $2n$ unknowns. We will discuss $w = 0$ later.

So the general solution in perturbation coordinates is that the β component of $\vec{\eta}$ is

$$(\vec{\eta})_\beta = \sum_\alpha A_{\beta\alpha} C_\alpha e^{-iw_\alpha t}$$

or

$$\vec{\eta}(t) = \sum_\alpha C_\alpha \vec{a}_\alpha e^{-iw_\alpha t}$$

Therefore we can start the system with $C_\alpha = \begin{cases} 1 & \alpha = \gamma \\ 0 & \alpha \neq \gamma \end{cases}$ (i.e. $\begin{cases} \rho_\alpha = 1, & \phi_\alpha = 0 \\ \rho_\alpha = 0, & \phi_\alpha = 0 \end{cases}$), or $\vec{\eta}(t) = \vec{a}_\gamma e^{-iw_\gamma t}$ hence the system is undergoing a simple oscillatory motion with one specific frequency w_γ. Because $\vec{a}_\gamma$ is a real vector, the components of $\vec{\eta}(t)$ of normal mode solution are either in phase or 180^0 out of phase.

Lecture 13 (3/5/13)

Example. Spring Pendulum (revisit)

Last time we found two eigenmodes. The associated normalized (wrt quadratic form $\hat{M}$) eigenvectors are

$$\vec{a}_1 = \begin{pmatrix} \frac{1}{\sqrt{mr_0}} \\ 0 \end{pmatrix} \quad \vec{a}_2 = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{m}} \end{pmatrix}$$

so

$$\hat{A} = \begin{pmatrix} \frac{1}{\sqrt{mr_0}} & 0 \\ 0 & \frac{1}{\sqrt{m}} \end{pmatrix}$$

or

$$\vec{\eta} = \begin{pmatrix} \phi \\ \delta r \end{pmatrix} = \frac{1}{\sqrt{m}} \begin{pmatrix} \frac{1}{r_0} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix}$$

Normal modes:

1) $Q_1 \neq 0$, $Q_2 = 0$

$$Q_1 = C_1 e^{-i\sqrt{g/r_0}t} \text{ or } \phi = \frac{1}{\sqrt{mr_0}} Q_1,\ \delta r = 0$$

2) $Q_2 \neq 0$, $Q_1 = 0$

$$Q_2 = C_2 e^{-i\sqrt{k/m}t} \text{ or } \phi = 0,\ \delta r = \frac{1}{\sqrt{m}} Q_2$$

3.4 Vanishing Eigenfrequencies

Suppose some $w^2 = 0$. It usually comes from the symmetry of the problem, as in the following example.

Example. Spring model of CO_2. Suppose three atoms O, C, and O are collinear with C in the middle. Each is connected with springs k with rest length a. let the mass of O be M, and mass of C be m. The reason we pretend the interaction between O and C is spring because the Lagrangian of oscillation up to 2nd order is spring. Moreover although analyzing inter atomic motion is really a quantum mechanics problem, our discussion is still valid. That is because QM HO is exact solvable, and they have classical counterpart.

Because the overall positioning is irrelevant, let us choose the equilibrium positions $x_1^{(0)} = -a$, $x_2^{(0)} = 0$, and $x_3^{(0)} = a$. Before we said that η must be small, so we keep the 2nd order term in V, but we do require η small now, because here V has no higher order term.

$$V = \frac{1}{2}k(x_2 - x_1 - a)^2 + \frac{1}{2}k(x_3 - x_2 - a)^2 \implies \hat{K} = \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix}$$

and

$$T = \frac{1}{2}M\dot{x}_1^2 + \frac{1}{2}m\dot{x}_2^2 + \frac{1}{2}M\dot{\eta}_3^2 = \frac{1}{2}M\dot{\eta}_1^2 + \frac{1}{2}m\dot{\eta}_2^2 + \frac{1}{2}M\dot{\eta}_3^2 \implies \hat{M} = \begin{pmatrix} M & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & M \end{pmatrix}$$

Now solve for w^2

$$\det \begin{pmatrix} k - w^2M & -k & 0 \\ -k & 2k - w^2m & -k \\ 0 & -k & k - w^2M \end{pmatrix} = 0 \implies \begin{cases} w_1^2 = 0 \\ w_2^2 = \frac{k}{M} \\ w_3^2 = \frac{k}{M}(1 + \frac{2M}{m}) \end{cases}$$

w_2^2, w_3^2 are usual normal modes, with eigenvectors

$$\vec{a}_2 = \frac{1}{\sqrt{2M}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \qquad \vec{a}_3 = \frac{1}{\sqrt{2M + \frac{4M^2}{m}}} \begin{pmatrix} 1 \\ -\frac{2M}{m} \\ 1 \end{pmatrix}$$

What is w_1^2? From the separated $\mathcal{L}$ (3.10), apply EL to Q_1, we get the usual SHO equation

$$\ddot{Q}_1 + w_1^2 Q_1 = 0 \implies Q_1 = c + dt$$

we get uniform translational motion in Q_1, and eigenvector

$$\begin{pmatrix} 1 & -1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \vec{a}_1 = 0 \implies \vec{a}_1 = \frac{1}{\sqrt{2M+m}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

then

$$\vec{\eta} = \vec{a}_1(c + dt) + \vec{a}_2(C_2 e^{-iw_2 t}) + \vec{a}_3(C_3 e^{-iw_3 t}) \tag{3.11}$$

Now we show how to setup the system in normal modes.

$$\hat{Q} = \hat{A}^T \hat{M} \vec{\eta} = \begin{pmatrix} \frac{1}{\sqrt{2M+m}}(M\eta_1 + m\eta_2 + M\eta_3) \\ \frac{\sqrt{M}}{\sqrt{2}}(\eta_1 - \eta_3) \\ \frac{\sqrt{M}}{\sqrt{2+\frac{4M}{m}}}(\eta_1 - 2\eta_2 + \eta_3) \end{pmatrix} \tag{3.12}$$

To incorporate the initial conditions, we look at

$$\begin{pmatrix} c \\ \rho_2 \cos\phi_2 \\ \rho_3 \cos\phi_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2M+m}}(M\eta_1(0) + m\eta_2(0) + M\eta_3(0)) \\ \frac{\sqrt{M}}{\sqrt{2}}(\eta_1(0) - \eta_3(0)) \\ \frac{\sqrt{M}}{\sqrt{2+\frac{4M}{m}}}(\eta_1(0) - 2\eta_2(0) + \eta_3(0)) \end{pmatrix}$$

and

$$\begin{pmatrix} d \\ \rho_2 w_2 \sin\phi_2 \\ \rho_3 w_3 \sin\phi_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2M+m}}(M\dot{\eta}_1(0) + m\dot{\eta}_2(0) + M\dot{\eta}_3(0)) \\ \frac{\sqrt{M}}{\sqrt{2}}(\dot{\eta}_1(0) - \dot{\eta}_3(0)) \\ \frac{\sqrt{M}}{\sqrt{2+\frac{4M}{m}}}(\dot{\eta}_1(0) - 2\dot{\eta}_2(0) + \dot{\eta}_3(0)) \end{pmatrix}$$

1) If we start the system, e.g. with

$$\eta_1(0) = \eta_2(0) = \eta_3(0) = 0 \text{ and } \dot{\eta}_1(0) = \dot{\eta}_2(0) = \dot{\eta}_3(0) = 1$$

so

$$c = \rho_{2,3} = 0,\ d \neq 0$$

so

$$\vec{\eta} = \vec{a}_1 dt = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} t$$

constant motion.

Note: That $Q_1(t) = \frac{1}{\sqrt{2M+m}}(M\eta_1(t) + m\eta_2(t) + M\eta_3(t))$ doesn't contradict the fact that we can setup the system into a mode that all three $\eta_{1,2,3}$ move at the constant speed and at the same time $Q_1(t)$ moves at the same speed.

2) If we start the system, e.g. with

$$\eta_1(0) = -\eta_3(0) = 1,\ \eta_2(0) = 0 \text{ and } \dot{\eta}_1(0) = \dot{\eta}_2(0) = \dot{\eta}_3(0) = 0$$

so

$$c = d = \phi_2 = \rho_3 = 0,\ \rho_2 \neq 0$$

so

$$\vec{\eta} = \vec{a}_2 \rho_2 e^{-iw_2 t} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} e^{-iw_2 t}$$

3) If we start the system, e.g. with

$$\eta_1(0) = \eta_3(0) = 1,\ \eta_2(0) = -\frac{2M}{m} \text{ and } \dot{\eta}_1(0) = \dot{\eta}_2(0) = \dot{\eta}_3(0) = 0$$

so

$$c = d = \phi_3 = \rho_2 = 0,\ \rho_3 \neq 0$$

so

$$\vec{\eta} = \vec{a}_3 \rho_3 e^{-iw_3 t} = \begin{pmatrix} 1 \\ -\frac{2M}{m} \\ 1 \end{pmatrix} e^{-iw_3 t}$$

There are many other ways to setup the system in normal modes. But of course in the setup of modes 2 and 3 we need to make sure $|\eta_{1,2,3}(t)| < a$, $\forall t$ so no collisions between masses, since our Lagrangian doesn't consider this.

Also notice (3.11), (3.12) shows that the normalization in $\vec{a}$ doesn't matter much, it will come from initial conditions anyway, regardless whether $\hat{M}$ is already diagonal or not.

The above example shows $w^2 = 0$ may come from symmetry. The next simple example shows $w^2 = 0$ may come from accidental feature of quadratic approximation.

Example. Let

$$\mathcal{L} = \frac{1}{2} m \dot{x}^2 - \lambda x^4$$

equilibrium $x^{(0)} = 0$, so $\eta = x - x^{(0)} = x$, hence in quadratic approximation

$$\mathcal{L} = \frac{1}{2} m \dot{x}^2$$

o $w^2 = 0$. But in fact we have second-order nonlinear differential equation, setting $w^2 = 6\lambda/m$

$$\ddot{x} + w^2 x^3 = 0$$

showing this is still an oscillatory motion with acceleration is smaller near the equilibrium and larger when displacement is large.

3.5 Degenerated Eigenfrequencies

Lecture 14 (3/7/13)

This is when some $w_\alpha^2 = w_\beta^2$. Spectral theory tells us since $(\hat{K} - w_\alpha^2 \hat{M})$ is real symmetric, there are enough eigenvectors. Say the dimension of the null space $(\hat{K} - w_\alpha^2 \hat{M})$ is N, then there are N independent eigenvectors, so we use Gram-Schmidt to find $\vec{a}_\alpha$, ...

Hence (3.10) is still true. But Q_α is not unique.

For simplicity suppose we have double root as in the following example.

2D oscillator

$$\mathcal{L} = \mathcal{L}_x + \mathcal{L}_y$$

$$\mathcal{L}_x = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2$$

we can rotate the xy axes, and choose a different set of eigenvectors x', y'.

$$\mathcal{L} = \mathcal{L}_{x'} + \mathcal{L}_{y'}$$

$$\mathcal{L}_{x'} = \frac{1}{2}m\dot{x}'^2 - \frac{1}{2}kx'^2$$

4 Hamiltonian Formalism of Mechanics

Reference Goldstein ch8; LL ch7.

Consider a system with n degrees of freedom

$$\mathcal{L}(q, \dot{q}, t)$$

and the conjugate momentum (1.11),

$$p_\alpha = \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}$$

so p_α is a function of q, $\dot{q}$, and t. If we are able to invert the above system of equations, we can get

$$\dot{q}_\beta = \dot{q}_\beta(q, p, t)$$

The simplest such inversion happens for 1D particles

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2 - V(x)$$

then

$$p_x = \frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x} \text{ and } \dot{x} = \frac{p_x}{m}$$

Later we will see the advantages of using $(q_\alpha, p_\alpha)_{\alpha=1,\ldots,n}$ as set of the variables. They are called "canonical variables" and they parametrize a $2n$ dimensional space, called "phase space".

4.1 Hamiltonian Equations

What would the Euler Lagrangian equation look like in (q_α, p_α) variables?

EL

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\right) - \frac{\partial \mathcal{L}}{\partial q_\alpha} = 0$$

is a second order pde with some initial conditions $q_\alpha(0)$, $\dot{q}_\alpha(0)$, but in (q_α, p_α) variables, it becomes two 1st order equations

$$\begin{cases} \frac{d}{dt}p_\alpha - \frac{\partial \mathcal{L}}{\partial q_\alpha} = 0 & (1) \\ \frac{d}{dt}q_\alpha = \dot{q}_\alpha(q,p,t) & (2) \end{cases} \tag{4.1}$$

with some initial conditions $q_\alpha(0)$, $p_\alpha(0)$.

In (4.1)(2), we already assume the inversion is possible, i.e. the explicit form of $\dot{q}_\alpha(q,p,t)$ is achievable. But there is still problem in (4.1)(1), $\frac{\partial \mathcal{L}}{\partial q_\alpha}$ says one fixes $\dot{q}$, t constant while taking the partial, but $\dot{q}$ is not independent variable now. So the operation $\frac{\partial \mathcal{L}}{\partial q_\alpha}$ is not legitimate.

Recall the Hamilton of the system (1.19)

$$\begin{aligned} H(q,p,t) &= \sum_\beta \frac{\partial \mathcal{L}}{\partial \dot{q}_\beta}\dot{q}_\beta - \mathcal{L} \\ &= \sum_\beta p_\beta \dot{q}_\beta(q,p,t) - \mathcal{L}(q, \dot{q}(q,p,t), t) \end{aligned}$$

Consider

$$\frac{\partial H}{\partial p_\alpha} = \dot{q}_\alpha + \sum_\beta p_\beta \frac{\partial \dot{q}_\beta}{\partial p_\alpha} - \sum \frac{\partial \mathcal{L}}{\partial \dot{q}_\beta}\frac{\partial \dot{q}_\beta}{\partial p_\alpha} = \dot{q}_\alpha$$

$$\frac{\partial H}{\partial q_\alpha} = \sum_\beta p_\beta \frac{\partial \dot{q}_\beta}{\partial q_\alpha} - \frac{\partial \mathcal{L}}{\partial q_\alpha} - \sum \frac{\partial \mathcal{L}}{\partial \dot{q}_\beta}\frac{\partial \dot{q}_\beta}{\partial q_\alpha} = -\frac{\partial \mathcal{L}}{\partial q_\alpha}$$

Both cancellations are due to $\frac{\partial \mathcal{L}}{\partial \dot{q}_\beta} = p_\beta$. Now using them, we can rewrite EOM and obtain Hamilton equations

$$\begin{cases} \frac{d}{dt} q_\alpha = \frac{\partial H}{\partial p_\alpha} \\ \frac{d}{dt} p_\alpha = -\frac{\partial H}{\partial q_\alpha} \end{cases}$$

The minus is crucial.

4.2 Legendre Transformation

The transformation above can be better expressed in turns as mathematical language belonged to the name of Legendre transformation.

We want to transform from $\mathcal{L}(q, \dot{q})$ to $H(q, p)$.

Think a generic function $f(x)$, [our prototype of $\mathcal{L}(\dot{q})$, ignoring q, t, because they are independent of p, $\dot{q}$, they don't make much difference], let

$$y = f'(x)$$

[which should gives us p,]

To invert

$$x = x(y)$$

[giving $\dot{q}(p)$], we put

$$g(y) \equiv yx(y) - f(x(y))$$

[this is our $H = p\dot{q}(p) - \mathcal{L}(\dot{q}(p))$], and try

$$g'(y) = x(y) + yx'(y) - f'(x(y))x'(y) = x(y)$$

Hence

$$y = f'(x) \leftrightarrow x = g'(y)$$

4.3 $\dot{q}$ and p Inversion

Lecture 15 (3/12/13)

How likely we can invert $\dot{q}$ and p? It turns out to be not too hard if $\partial\mathcal{L}/\partial t = 0$.

Couple weeks ago we showed if $\partial\mathcal{L}/\partial t = 0$, then H is the total energy, and total energy is conserved. cf (1.20) and follow. We now deduce from

$$T = \sum_{\beta\gamma} \frac{1}{2} M_{\beta\gamma}(q)\dot{q}_\beta \dot{q}_\gamma \text{ and } M_{\beta\gamma} = M_{\gamma\beta}$$

that

$$p_\alpha = \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} = \frac{\partial T}{\partial \dot{q}_\alpha} = \sum_\beta M_{\alpha\beta}(q)\dot{q}_\beta$$

If we employ vector form for $\vec{q}$, $\dot{\vec{q}}$, and $\vec{p}$ and matrices form for $\hat{M}(q)$,

$$\vec{p} = \hat{M}(q) \cdot \dot{\vec{q}}$$

$$\dot{\vec{q}} = \hat{M}^{-1}(q) \cdot \vec{p}$$

Consequentially

$$T = \frac{1}{2}\dot{\vec{q}}^T \hat{M}(q)\dot{\vec{q}} = \frac{1}{2}\vec{p}^T (\hat{M}^{-1})^T \hat{M}\hat{M}^{-1}\vec{p} = \frac{1}{2}\vec{p}^T \hat{M}^{-1}\vec{p}$$

looking very much to $p^2/2m$ in 1D case.

Example. 2D particle with central potential in polar

$$\mathcal{L} = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\phi}^2 - V(r)$$

$$\hat{M}(r,\phi) = \begin{pmatrix} m & \\ & mr^2 \end{pmatrix},\ p_r = \frac{\partial \mathcal{L}}{\partial \dot{r}} = m\dot{r},\ p_\phi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = mr^2\dot{\phi}$$

and

$$H = \frac{1}{2}\begin{pmatrix} p_r \\ p_\phi \end{pmatrix}^T \begin{pmatrix} \frac{1}{m} & \\ & \frac{1}{mr^2} \end{pmatrix} \begin{pmatrix} p_r \\ p_\phi \end{pmatrix} + V = \frac{p_r^2}{2m} + \frac{p_\phi^2}{2mr^2} + V(r)$$

Equation of Motion

$$r: \begin{cases} \dot{r} = \frac{\partial H}{\partial p_r} = \frac{p_r}{m} & (1) \\ \dot{p}_r = -\frac{\partial H}{\partial r} = -V'(r) + \frac{p_\phi^2}{mr^3} & (2) \end{cases}$$

$$\phi : \begin{cases} \dot{\phi} = \frac{\partial H}{\partial p_\phi} = \frac{p_\phi}{mr^2} & (3) \\ \dot{p}_\phi = -\frac{\partial H}{\partial \phi} = 0 & (4) \end{cases}$$

Working with Hamilton's mechanics reveal some interesting mathematical structures that are not available in Lagrangian mechanics. First of them is Poisson's bracket.

4.4 Poisson Bracket

Reference Goldstein 9.5, 9.6; LL section 42

Definition. For any two functions of phase space

$$f(q,p;t) \quad g(q,p;t)$$

we define

$$[f,g] = \sum_{\alpha=i}^{n} \left(\frac{\partial f}{\partial q_\alpha} \frac{\partial g}{\partial p_\alpha} - \frac{\partial f}{\partial p_\alpha} \frac{\partial g}{\partial q_\alpha} \right)$$

note: in LL, $[f,g]$ the order is reversed.

Properties of Poisson Bracket

1. Antisymmetric

$$[f,g] = -[g,f] \tag{4.2}$$

hence $[f,f] = 0$

2.

$$[f, \text{const}] = 0 \tag{4.3}$$

3. Linearity

$$[\alpha f_1 + \beta f_2, g] = \alpha[f_1, g] + \beta[f_2, g] \tag{4.4}$$

4.

$$[fg, h] = f[g,h] + [f,h]g \tag{4.5}$$

5. Jacob's Identity

$$[f,[g,h]] + [h,[f,g]] + [g,[h,f]] = 0 \tag{4.6}$$

similar identity holds if f, g, h are permuted to the left. Both are not so obvious to prove.

6. Some interesting Poisson brackets

$$[q_\beta, f] = \sum_{\alpha=i}^{n} \left(\frac{\partial q_\beta}{\partial q_\alpha} \frac{\partial f}{\partial p_\alpha} - \frac{\partial q_\beta}{\partial p_\alpha} \frac{\partial f}{\partial q_\alpha} \right) = \sum_{\alpha=i}^{n} \left(\delta_{\alpha\beta} \frac{\partial f}{\partial p_\alpha} - 0 \frac{\partial f}{\partial q_\alpha} \right) = \frac{\partial f}{\partial p_\beta} \tag{4.7}$$

Similar to show

$$[p_\beta, f] = -\frac{\partial f}{\partial q_\beta} \tag{4.8}$$

7. Canonical conditions

$$[q_\alpha, q_\beta] = 0,\ [p_\alpha, p_\beta] = 0,\ [q_\alpha, p_\beta] = \delta_{\alpha\beta}$$

We can use Poisson bracket to write Hamilton's equations

$$\begin{cases} \dot{q}_\alpha = \frac{\partial H}{\partial p_\alpha} = [q_\alpha, H] \\ \dot{p}_\alpha = -\frac{\partial H}{\partial q_\alpha} = [p_\alpha, H] \end{cases}$$

For any $f(q, p; t)$

$$\begin{aligned} \frac{d}{dt} f &= \frac{\partial f}{\partial t} + \sum_\alpha \frac{\partial f}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial f}{\partial p_\alpha} \dot{p}_\alpha \\ &= \frac{\partial f}{\partial t} + \sum_\alpha \left(\frac{\partial f}{\partial q_\alpha} \frac{\partial H}{\partial p_\alpha} - \frac{\partial f}{\partial p_\alpha} \frac{\partial H}{\partial q_\alpha} \right) = \frac{\partial f}{\partial t} + [f, H] \end{aligned} \tag{4.9}$$

Moreover if $\partial f / \partial t = 0$,

$$f \text{ is conserved} \iff [f, H] = 0 \tag{4.10}$$

Very soon we will apply this to angular momentum.

Claim: if f, g have no explicit time dependence,

$$f, g \text{ both conserved} \implies [f, g] \text{ conserved}$$

Proof. $[[f, g], H] = -[H, [f, g]] = [f, [g, H]] + [g, [H, f]] = 0$ □

In the next example, we will see that Poisson bracket can help to generate conservative law. Of course the process will not go on forever; after some point, Poisson bracket will be 0, and no new conservative law will be produced.

Example. Suppose we know L_x, L_y are conserved

By above $[L_x, L_y]$ is conserved. We also claim $[L_x, L_y] = L_z$. Recall $L = \vec{r} \times \vec{p}$

$$L_x = yp_z - zp_y \quad L_y = zp_x - xp_z$$

$$[L_x, L_y] = \sum_{i=1}^{3} \left(\frac{\partial L_x}{\partial r_i} \frac{\partial L_y}{\partial p_i} - \frac{\partial L_x}{\partial p_i} \frac{\partial L_y}{\partial r_i} \right)$$

$i = 1, 2$ give 0, only $i = 3$ survives

$$[L_x, L_y] = (-p_y)(-x) - (y)(p_x) = L_z$$

Conclusion if L_x,L_y are conserved, then L_z has to be conserved. This is quite deep fact and not easy to prove in group theory.

Lecture 16 (3/26/13)

From (4.10), take $f = H$

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} \tag{4.11}$$

Hence if H doesn't depend explicitly, then energy is conserved.

4.5 Canonical Transformation

LL section 45.

We won't discuss generating functions, for they are confusing.

Recall in lecture one we showed that Lagrangian formulation allowed to use any generic coordinates

$$q_\alpha = q_\alpha(\tilde{q}; t)$$

and

$$\tilde{\mathcal{L}}(\tilde{q}, \dot{\tilde{q}}; t) = \mathcal{L}(q, \dot{q}; t)$$

and EOM has same form
$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\right) - \frac{\partial \mathcal{L}}{\partial q_\alpha} = 0$$
$$\frac{d}{dt}\left(\frac{\partial \tilde{\mathcal{L}}}{\partial \dot{\tilde{q}}_\alpha}\right) - \frac{\partial \tilde{\mathcal{L}}}{\partial \tilde{q}_\alpha} = 0$$

For Hamilton it's not that easy. Suppose we do

$$q_\alpha = q_\alpha(Q; t)$$

How to find p_α? One can first write down the new $\mathcal{L}$ then compute the conjugate momentum. But there is a more robust way, without going through Lagrangian, and it works for more general transformation.

Suppose we want to do

$$\begin{cases} q_\alpha = q_\alpha(Q_\beta, P_\beta) \\ p_\alpha = p_\alpha(Q_\beta, P_\beta) \end{cases} \tag{4.12}$$

henceforth let's assume no explicit t dependence in the transformation. (later we will show why we require no t.)

Theorem. (4.12) *is good transformation, i.e. leaving the form of Hamilton equations unaffected, iff* (4.12) *is canonical transformation, i.e. satisfying canonical Poisson bracket.*

Let's prove the forward direction, and leave the converse to the next lecture.

Proof. From (4.9),
$$\dot{Q}_\alpha = [Q_\alpha, H] \qquad \dot{P}_\alpha = [P_\alpha, H]$$

[notice: here the Poisson brackets are derivatives wrt to old q, p. Later we will do Poisson bracket derivatives wrt to Q, P, and we will show they are equivalent]

Since Q, P are good,

$$\frac{\partial H}{\partial P_\alpha} = [Q_\alpha, H] \qquad -\frac{\partial H}{\partial Q_\alpha} = [P_\alpha, H]$$

Compare to

$$\frac{\partial H}{\partial P_\alpha} = \sum_\beta \left(\frac{\partial H}{\partial q_\beta}\frac{\partial q_\beta}{\partial P_\alpha} + \frac{\partial H}{\partial p_\beta}\frac{\partial p_\beta}{\partial P_\alpha} \right)$$

and

$$[Q_\alpha, H] = \sum_\beta \left(\frac{\partial Q_\alpha}{\partial q_\beta}\frac{\partial H}{\partial p_\beta} - \frac{\partial Q_\alpha}{\partial p_\beta}\frac{\partial H}{\partial q_\beta} \right)$$

Since H is arbitrary, we obtain

$$\frac{\partial q_\beta}{\partial P_\alpha} = -\frac{\partial Q_\alpha}{\partial p_\beta} \qquad \frac{\partial p_\beta}{\partial P_\alpha} = \frac{\partial Q_\alpha}{\partial q_\beta} \tag{4.13}$$

Similarly, we obtain

$$\frac{\partial q_\beta}{\partial Q_\alpha} = \frac{\partial P_\alpha}{\partial p_\beta} \qquad \frac{\partial p_\beta}{\partial Q_\alpha} = -\frac{\partial P_\alpha}{\partial q_\beta} \tag{4.14}$$

Using (4.13) we compute

$$\begin{aligned}
[Q_\alpha, Q_\gamma] &= \sum_\beta \left(\frac{\partial Q_\alpha}{\partial q_\beta}\frac{\partial Q_\gamma}{\partial p_\beta} - \frac{\partial Q_\alpha}{\partial p_\beta}\frac{\partial Q_\gamma}{\partial q_\beta} \right) \\
&= \sum_\beta \left(\frac{\partial Q_\alpha}{\partial q_\beta}\frac{-\partial q_\beta}{\partial P_\gamma} - \frac{\partial Q_\alpha}{\partial p_\beta}\frac{\partial p_\beta}{\partial P_\gamma} \right) = -\frac{\partial Q_\alpha}{\partial P_\gamma} = 0
\end{aligned}$$

Similarly using (4.13), (4.14),

$$[P_\alpha, P_\gamma] = 0$$

$$[Q_\alpha, P_\gamma] = \delta_{\alpha\gamma}$$

$\square$

The following two examples exemplify that canonical transformation is much wilder than coordinate transformation in Lagrangian mechanics. As a result of this, often its physical realization looses completely.

Example. Consider

$$Q = p$$

$$P = -q$$

check this is canonical

$$[Q,Q] = [P,P] = 0$$

$$[Q,P] = [p,-q] = 1$$

From Lagrangian $\mathcal{L} = \frac{1}{2}m\dot{x}^2 - V(x)$ viewpoint such transformation is crazy

$$q = x = -P$$

$$p = \frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x} = Q$$

Example. 1D harmonic oscillator (Goldstein 9.3)

$$H = T + V = \frac{p^2}{2m} + \frac{1}{2}kq^2$$

Suppose transformation

$$\begin{cases} p = f(P)\cos Q \\ q = \frac{f(P)}{\sqrt{km}}\sin Q \end{cases}$$

is canonical and $f(P)$ is to be determined. Then

$$H = \frac{f^2(P)}{2m} \tag{4.15}$$

then by Hamilton's equations

$$\dot{P} = -\frac{\partial H}{\partial Q} = 0 \implies P = P_0$$

$$\dot{Q} = \frac{\partial H}{\partial P} = \frac{f(P_0)f'(P_0)}{m} \implies Q = Q_0 + \dot{Q}_0 t$$

i.e. P is constant of motion and Q is constant speed motion.

To check Q, P are canonical wrt $[.,.]_{qp}$, we check their inverse q, p are canonical wrt $[.,.]_{QP}$. Next lecture we will prove that Q, P are canonical wrt $[.,.]_{qp}$ is equivalent to q, p are canonical wrt $[.,.]_{QP}$.

$$[q,q]_{QP} = [p,p]_{QP} = 0$$

$$[q,p]_{QP} = \frac{ff'}{\sqrt{km}}$$

which has to be 1, so

$$\frac{1}{2}\frac{df^2}{dP} = \sqrt{km}$$

hence

$$f^2(P) = 2\sqrt{km}P$$

plus some constant and we choose it to be 0.

$$f(P) = \sqrt{2P}(km)^{1/4} = \sqrt{2P_0 mw}$$

Therefore by (4.15)

$$H = P_0 w$$

so

$$\dot{Q} = \dot{Q}_0 = \frac{\partial H}{\partial P} = w$$

Hence in usual q, p coordinates

$$q(t) = A\sin(wt + \phi)$$

$$p(t) = Amw\cos(wt + \phi)$$

where $w = \dot{Q}_0$, $\phi = Q_0$, $A = \sqrt{2P_0/m\dot{Q}_0}$.

Lecture 17 (3/28/13)

We can draw the trajectory in the phase plane of Q, p and in the usual phase plane of q, p.

In Q, P, for given w, ϕ, A, the motion is a straight line parallel to the Q axis. While in q, p/mw phase plane after rescaling the p axis, the motion is a circle centered at the origin.

If we consider a rectangle region $[P_1, P_1 + \Delta P] \times [Q_1, Q_1 + \Delta Q]$ in the Q, P phase plane, we find it corresponds to a fan region in the $q, p/mw$ phase plane bounded by two arcs of radii $\sqrt{2P_1/mw}$ and $\sqrt{2(P_1 + \Delta P)/mw}$, and bounded by two rays of polar angles Q_1, $Q_1 + \Delta Q$.

If we compute the area of the rectangle in Q, P plane, we get

$$(Q_1 + \Delta Q - Q_1)(P_1 + \Delta P - P_1) = \Delta Q \Delta P$$

If we then compute the corresponding region in q, p plane, we get

$$\int dp \int dq = mw \int dq \int d\left(\frac{p}{mw}\right) = mw \int_{\sqrt{2P_1/mw}}^{\sqrt{2(P_1+\Delta P)/mw}} rdr \int_{Q_1}^{Q_0+\Delta Q} d\phi = \Delta Q \Delta P$$

Hence the two areas are equal. This is a instance of a very general fact: Liouville theorem.

Recall (4.7), and (4.8)

$$[q_\alpha, f(q,p)]_{qp} = \frac{\partial f}{\partial p_\alpha} \quad [p_\alpha, f(q,p)]_{qp} = -\frac{\partial f}{\partial q_\alpha}$$

What is $[Q_\alpha, g(Q,P)]_{qp}$, if Q, P are canonical?

$$\begin{aligned}
[Q_\alpha, g(Q,P)]_{qp} &= \sum_\beta \left(\frac{\partial Q_\alpha}{\partial q_\beta}\frac{\partial g}{\partial p_\beta} - \frac{\partial Q_\alpha}{\partial p_\beta}\frac{\partial g}{\partial q_\beta}\right) \\
&= \sum_{\beta\gamma} \left(\frac{\partial Q_\alpha}{\partial q_\beta}\left(\frac{\partial g}{\partial Q_\gamma}\frac{\partial Q_\gamma}{\partial p_\beta} + \frac{\partial g}{\partial P_\gamma}\frac{\partial P_\gamma}{\partial p_\beta}\right) - \frac{\partial Q_\alpha}{\partial p_\beta}\left(\frac{\partial g}{\partial Q_\gamma}\frac{\partial Q_\gamma}{\partial q_\beta} + \frac{\partial g}{\partial P_\gamma}\frac{\partial P_\gamma}{\partial q_\beta}\right)\right) \\
&= \sum_\gamma \left(\frac{\partial g}{\partial Q_\gamma}[Q_\alpha, Q_\gamma]_{qp} + \frac{\partial g}{\partial P_\gamma}[Q_\alpha, P_\gamma]_{qp}\right) \\
&= \sum_\gamma \left(\frac{\partial g}{\partial Q_\gamma}0 + \frac{\partial g}{\partial P_\gamma}\delta_{\alpha\gamma}\right) = \frac{\partial g}{\partial P_\alpha}
\end{aligned}$$

Similarly one can show

$$[P_\alpha, g(Q,P)]_{qp} = -\frac{\partial g}{\partial Q_\alpha}$$

What about $[g(Q,P), f(Q,P)]_{qp}$?

$$\begin{aligned}
[g(Q,P), f(Q,P)]_{qp} &= \sum_{\beta\gamma}\left(\left(\frac{\partial g}{\partial Q_\gamma}\frac{\partial Q_\gamma}{\partial q_\beta} + \frac{\partial g}{\partial P_\gamma}\frac{\partial P_\gamma}{\partial q_\beta}\right)\frac{\partial f}{\partial p_\beta} - \left(\frac{\partial g}{\partial Q_\gamma}\frac{\partial Q_\gamma}{\partial p_\beta} + \frac{\partial g}{\partial P_\gamma}\frac{\partial P_\gamma}{\partial p_\beta}\right)\frac{\partial f}{\partial q_\beta}\right) \\
&= \sum_{\gamma}\left(\frac{\partial g}{\partial Q_\gamma}[Q_\gamma, f]_{qp} + \frac{\partial g}{\partial P_\gamma}[P_\gamma, f]_{qp}\right) = [g(Q,P), f(Q,P)]_{QP}
\end{aligned}$$

This allows we to define Poisson bracket wrt to derivatives of QP.

In en route we have proven the converse direction of the theorem stated last lecture: (4.12) is canonical transformation $\implies$ the form of Hamilton equations will be the same.

Proof. Suppose QP are canonical, then by (4.9)

$$\frac{dQ_\alpha}{dt} = \frac{\partial Q_\alpha}{\partial t} + [Q_\alpha, H]_{pq} = [Q_\alpha, H]_{pq} = \frac{\partial H}{\partial P_\alpha}$$

and

$$\frac{dP_\alpha}{dt} = [P_\alpha, H]_{pq} = -\frac{\partial H}{\partial Q_\alpha}$$

This also shows the reason why in the canonical transformation we require no explicit t dependence. □

Lastly we briefly prove the fact mentioned in the last lecture: Q, P are canonical wrt to $[.,.]_{qp}$ iff q, p are canonical wrt to $[.,.]_{QP}$.

Proof. Suppose Q, P are canonical wrt to $[...]_{qp}$, then use above

$$[q(Q,P), p(Q,P)]_{QP} = [q,p]_{qp}$$

Exchanging the role of $QP \to qp$ proves the converse. □

In summary, we have seen Q, P has the same standing as q, p, not one set of canonical variables is more prestigious than others. All formulas involving Q, P or q, p are exactly the same.

4.6 Liouville Theorem

Theorem. *(Liouville)*

1) Time evolution is a canonical transformation;

2) Canonical transformation preserve the volume in phase space;

Notice in 1) we say time evolution, not time translation. Time translation means simply shift the time scale and leaving other dynamical variables constant. Time evolution means all variables evolve with time.

1)+2) implies time evolution preserve volume in phase space.

Proof. 1) Let $t \to t + dt$, treat $q(t) = Q$, $p(t) = P$ as new variables. By Hamilton equations

$$\begin{cases} Q_\alpha = q_\alpha + \frac{\partial H(q,p)}{\partial p_\alpha} dt \\ P_\alpha = p_\alpha - \frac{\partial H(q,p)}{\partial q_\alpha} dt \end{cases}$$

check

$$\begin{aligned}
[Q_\alpha, Q_\beta]_{qp} &= [q_\alpha + \frac{\partial H}{\partial p_\alpha} dt, q_\beta + \frac{\partial H}{\partial p_\beta} dt] \\
&= [q_\alpha, q_\beta] + [q_\alpha, \frac{\partial H}{\partial p_\beta}] dt + [\frac{\partial H}{\partial p_\beta}, q_\beta] dt + O(dt)^2 \\
&= 0 + \frac{\partial^2 H}{\partial p_\alpha \partial p_\beta} - \frac{\partial^2 H}{\partial p_\alpha \partial p_\beta} \\
&= 0
\end{aligned}$$

$$[P_\alpha, P_\beta]_{qp} = 0$$

$$\begin{aligned}
[Q_\alpha, P_\beta]_{qp} &= [q_\alpha + \frac{\partial H}{\partial p_\alpha} dt, p_\beta - \frac{\partial H}{\partial q_\beta} dt] \\
&= [q_\alpha, p_\beta] + [q_\alpha, -\frac{\partial H}{\partial q_\beta}] dt + [\frac{\partial H}{\partial p_\alpha}, p_\beta] dt + O(dt)^2 \\
&= \delta_{\alpha\beta} - \frac{\partial^2 H}{\partial p_\alpha \partial q_\beta} + \frac{\partial^2 H}{\partial q_\beta \partial p_\alpha} \\
&= \delta_{\alpha\beta}
\end{aligned}$$

2) To show

$$\int_A dq_1 ... dq_n dp_1 ... dp_n = \int_{A'} dQ_1 ... dQ_n dP_1 ... dP_n$$

Since $dQ_1...dQ_ndP_1...dP_n = |D|dq_1...dq_ndp_1...dp_n$, where

$$D = \left(\frac{\partial \Xi}{\partial \xi}\right), \; \Xi = \begin{pmatrix} Q_1 \\ \vdots \\ Q_n \\ P_1 \\ \vdots \\ P_n \end{pmatrix}, \; \xi = \begin{pmatrix} q_1 \\ \vdots \\ q_n \\ p_1 \\ \vdots \\ p_n \end{pmatrix}$$

put indices on rows and columns $A, B = 1, ..., 2n$, $\alpha, \beta = 1, ..., n$

$$D_{BA} = \frac{\partial \Xi^A}{\partial \xi^B} = \begin{pmatrix} \frac{\partial Q^\alpha}{\partial q^\beta} & \frac{\partial P^\alpha}{\partial q^\beta} \\ \frac{\partial Q^\alpha}{\partial p^\beta} & \frac{\partial P^\alpha}{\partial p^\beta} \end{pmatrix}$$

Claim: $(Q, P) \longleftrightarrow (q, p)$ is a canonical transformation iff

$$D^T J D = J$$

where $J = \begin{pmatrix} 0 & I_{n\times n} \\ -I_{n\times n} & 0 \end{pmatrix}$.

Lecture 18 (4/2/13)

Indeed we can check for $n = 2$. The upper right matrix of $D^T JD$ is

$$\begin{aligned} \left(\frac{\partial Q^\alpha}{\partial q^\beta}\right)^T \frac{\partial Q^\alpha}{\partial p^\beta} - \left(\frac{\partial Q^\alpha}{\partial q^\beta}\right)^T \frac{\partial Q^\alpha}{\partial p^\beta} \quad &= \begin{pmatrix} \frac{\partial Q^1}{\partial q^1} & \frac{\partial Q^2}{\partial q^1} \\ \frac{\partial Q^1}{\partial q^2} & \frac{\partial Q^2}{\partial q^2} \end{pmatrix}^T \begin{pmatrix} \frac{\partial Q^1}{\partial p^1} & \frac{\partial Q^2}{\partial p^1} \\ \frac{\partial Q^1}{\partial p^2} & \frac{\partial Q^2}{\partial p^2} \end{pmatrix} - \text{switch } p, q \\ &= \begin{pmatrix} [Q_1, Q_1] & [Q_1, Q_2] \\ [Q_2, Q_1] & [Q_2, Q_2] \end{pmatrix} \end{aligned}$$

so we believe in general

$$D^T JD = \begin{pmatrix} [Q_\beta, Q_\alpha] & [Q_\beta, P_\alpha] \\ [P_\beta, Q_\alpha] & [P_\beta, P_\alpha] \end{pmatrix} = J$$

proving the claim.

$$\det J = (-1)^n \det \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$$

by permuting the columns n times, then

$$\det \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} = (-1)^n \det \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$$

Hence

$$(\det D)^2 \det J = \det J = 1 \implies |D| = 1$$

$\square$

Liouville says consider all points (different possible initial conditions) that constitute a phase volume V at time t_0, then check at another time t_1 of the all these points (initial conditions) evolve to make up volume V', and $V = V'$. Poincare recurrence suggests to fix a U in phase, then follow a point $a \in U$, some time a will be in U again.

Corollary. *(Poincare recurrence theorem, reference Arnold 3.16) If for a mechanical system motion is bounded in phase space, then in any neighborhood, U, in space space, there is a point $a \in U$ that after some time will go back to U.*

A more stronger statement which we won't prove says:

Corollary. *Almost all points of U, (apart from a set of measure 0) will go back to U.*

The assumption that motion is bounded is important, as we will use in the proof. Typically motion is bounded in coordinate space if we have normal potential that don't allow for long excursions. That is $V(\pm\infty) \not\to$some constant, so no scattering states.

It is not a prior fact that momentum is bounded. Although speed is bounded by c,

$$p = \frac{mv}{\sqrt{1 - v^2/c^2}}$$

is not necessarily bounded. However it is bounded if V is bounded below, i.e.

$$\min_{\{q\}} V > -\infty$$

Because $T = E - V \leq E - \min V < \infty$ for finite energy, and $T = \frac{1}{2}\vec{p}^T M \vec{p}$ where mass matrix M is positive definite, so momentum is bounded.

Proof. Let $g_t =$ the time evolution map (aka phase flow). For given initial (q_0, p_0) at $t = 0$

$$g_t(q_0, p_0) = (q(t), p(t))$$

where $q(t), p(t)$ is a solution of Hamilton equations. By $g_t(U)$ we mean the region in phase space generated by time evolving of all points in U.

Take any $U \subset A$ bounded in A, consider a sequence of time laps $g_1(U)$, $g_2(U)$, ... By Liouville, they all have same volume, so after some time, they must overlap, namely

$$\exists\, k > l \geq 0 \text{ s.t. } g_k \cap g_l \neq \varnothing$$

Reset the counter, set $l = 0$, or evolve them back in time by $-l$, because the map is bijective,

$$g_{-l}\left(g_k \cap g_l\right) = g_{k-l} \cap g_0 = g_{k-l}(U) \cap U \neq \varnothing$$

□

Note: If we start out an experiment with releasing some gas at the corner of the room, Poincare says at later time all gas molecules will come back to the corner at the same time. (not just one of them will come back to the corner) But in our cases, the phase space is $\sim 6 \times 10^{23}$ dimension, and U is tiny, (if the corner is about 1/10 of the room, then $U/A \sim 10^{-10^{23}}$) so it may take trillion and trillion of years to see $g_t(U)$ and U overlapping.

4.7 Effective Hamilton

Lecture 19 (4/4/13)

Recall when we studied holonomic constraints, we introduced $\mathcal{L}_{eff}$. We now introduce H_{eff}. Suppose q_1 is cyclic, i.e. $\partial H/\partial q_1 = 0 \implies p_1 = \alpha$ constant, hence p_1

is not really a variable, even it is presented in the form of H. We may just define

$$H_{eff}(q_2, ..., q_n; p_2, ..., p_n; t) = H(q_2, ..., q_n; \alpha, p_2, ..., p_n; t)$$

So we forget about p_1. What happens to the Hamilton equations? e.g.

$$\dot{q}_2 = \left.\frac{\partial H}{\partial p_2}\right|_{p_1=\alpha} = \frac{\partial H_{eff}}{\partial p_2}$$

hence no change at all, because evaluate p_1 at α before or after taking ∂p_2 should be the same by the definition of partial derivatives.

This method doesn't work for Lagrangian. Suppose $\partial L/\partial q_1 = 0$, by Euler Lagrange, $p_1 = \partial L/\partial \dot{q}_1 = \alpha$ constant. Does it imply anything about $\dot{q}_1$? The relation between p_1 and $\dot{q}_1$ is complicated,

$$p_1 = p_1(q_2, ..., q_n; \dot{q}_1, ..., \dot{q}_n; t)$$

so inversely

$$\dot{q}_1 = \dot{q}_1(p_1 = \alpha, q_2, ..., q_n; \dot{q}_2, ..., \dot{q}_n; t)$$

Pretend we want to do the same L_{eff} thing to eliminate both q_1 and $\dot{q}_1$, by setting

$$L_{eff}(q_2, ..., q_n; \dot{q}_2, ..., \dot{q}_n; t) = L(q_2, ..., q_n; \dot{q}_1(\alpha, q_2, ..., q_n; \dot{q}_2, ..., \dot{q}_n; t), ..., \dot{q}_n; t)$$

Does this replicate the same EL equation? ANS: No

Consider the following example.

Example. 2D motion in central potential

$$\mathcal{L} = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\phi}^2 - V(r)$$

so $\partial\mathcal{L}/\partial\phi = 0 \implies p_\phi = \partial\mathcal{L}/\partial\dot{\phi} = mr^2\dot{\phi} = L$ constant, so we get $\dot{\phi} = L/mr^2$.

If we do

$$'L'_{eff}(r, \dot{r}) = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}\frac{L^2}{mr^2} - V(r)$$

apply EL

$$\frac{d}{dt}\frac{\partial L_{eff}}{\partial \dot{r}} - \frac{\partial L_{eff}}{\partial r} = 0$$

get

$$m\ddot{r} + \frac{L}{mr^3} + V'(r) = 0$$

But if apply EL to the original $\mathcal{L}$, we get

$$m\ddot{r} - \frac{L}{mr^3} + V'(r) = 0$$

This gives another advantage of Hamilton over Lagrange.

4.8 Adiabatic Invariants

Reference L.L section 49.

Consider 1D system, bounded motion, and there must be some natural time scale, a "period" T of the system.

Suppose L or H depend explicitly on time vary slowly via some parameter $\lambda(t)$, meaning

$$\frac{1}{\lambda}\frac{d\lambda}{dt} \cdot T \ll 1 \iff \frac{d\lambda}{dt} \ll \frac{\lambda}{T} \tag{4.16}$$

and because the motion has to stay bounded, we ask the varying not to take too long that H changes too much or our observation doesn't take too long, normally $t = O(T)$.

Why do we impose slowly varying? If the change is too abrupt, say V suddenly changes at some places not near the particle, then the particles may know nothing about the change. If change takes place over long time, particles will certainly adjust the paths accordingly.

We'll have to give up energy conservation, by (4.11),

$$\frac{dH(q,p;\lambda(t))}{dt} = \frac{\partial H}{\partial t} = \frac{\partial H}{\partial \lambda} \cdot \dot{\lambda} \tag{4.17}$$

we know $\dot{\lambda}$ is small but $\frac{\partial H}{\partial \lambda}$ is not necessarily small, nor slowly varying.

Example. Let λ be spring constant k, and $k = k(t)$

$$H(q, p, k) = \frac{p^2}{2m} + \frac{1}{2}kq^2$$

k vary slowly. Differentiate equation, like

$$\ddot{q}(t) + w^2(t)q(t) = 0, \quad w = \sqrt{k/m}$$

is always solved in a small neighborhood of t. Within a small neighborhood, comparing to the change in q, w is almost constant, so we say the solution is still

$$q(t) = q_0 \cos wt$$

Let's see how correct the above solution is.

$$\dot{q}(t) = -q_0(w't + w) \sin wt$$

$$\ddot{q}(t) = -q_0(w't + w)^2 \cos wt - q_0(w''t + 2w') \sin wt$$

so we would like

$$w't \ll w, \quad (w''t + 2w') \ll w^2$$

They are met by (4.16), indeed $t \sim T$

$$\frac{w'}{w}t = \frac{k'}{k}T \ll 1,$$

i.e. $w' \ll w/T$, then take derivative,$w'' \ll \frac{w'}{T}$, so

$$(w''t + 2w') \sim w' \ll w \cdot \frac{1}{T} \sim w^2$$

We can also check that

$$\frac{\partial H}{\partial k} = \frac{1}{2}q^2 = \frac{1}{2}q_0^2 \cos^2 wt \tag{4.18}$$

is not too small, nor it varies slowly.

Averaging Over one period

Part of (4.17) is not small varying, so we would like to compute

$$\left\langle \frac{dH}{dt} \right\rangle = \frac{1}{T} \int_0^T \frac{dH}{dt} dt$$

the average rate of steady variation of the energy, which can be simplified

$$\left\langle \frac{dH}{dt} \right\rangle \doteq \left\langle \frac{\partial H}{\partial \lambda} \right\rangle \dot{\lambda}$$

$\dot{\lambda}$ changes very little over one period, so take it out of integral,

The average $\left\langle \frac{\partial H}{\partial \lambda} \right\rangle$ is computed on 1st order approximation, as if $\lambda(t) =$ constant. E.g. in (4.18), we'll treat w as a constant, for its change is small.

$$\left\langle \frac{\partial H}{\partial \lambda} \right\rangle = \frac{1}{T} \int_0^T \frac{\partial H}{\partial \lambda} dt$$

Since $T = \int_0^T dt$ and

$$dt = dq/\dot{q} = \frac{dq}{\frac{\partial H}{\partial p}}$$

$$\left\langle \frac{\partial H}{\partial \lambda} \right\rangle = \frac{\int_0^T \frac{\partial H}{\partial \lambda} \frac{dq}{\frac{\partial H}{\partial p}}}{\int_0^T \frac{dq}{\frac{\partial H}{\partial p}}}$$

For $\lambda(t) =$ constant,

$$H(q, p; \lambda) = E(t) = \text{constant} \tag{4.19}$$

Then solve for p,

$$p = p(q; \lambda, E)$$

treating $(q; \lambda, E)$ as the set of independent variables instead of $(q, p; \lambda)$, take $\partial\lambda$ of (4.19),

$$\frac{\partial H}{\partial \lambda} + \frac{\partial H}{\partial p} \frac{\partial p}{\partial \lambda} = 0 \implies \frac{\partial H/\partial \lambda}{\partial H/\partial p} = -\frac{\partial p}{\partial \lambda}$$

take ∂E of (4.19),

$$\frac{\partial H}{\partial p}\frac{\partial p}{\partial E} = 1 \implies \frac{\partial H}{\partial p} = \frac{1}{\partial p/\partial E}$$

Hence

$$\left\langle \frac{\partial H}{\partial \lambda} \right\rangle = -\frac{\int_0^T \frac{\partial p}{\partial \lambda} dq}{\int_0^T \frac{\partial p}{\partial E} dq} = -\frac{\frac{\partial}{\partial \lambda}\int_0^T p(q;\lambda,E)dq}{\frac{\partial}{\partial E}\int_0^T p(q;\lambda,E)dq} \tag{4.20}$$

Moreover, since it varies slowly

$$\left\langle \frac{dE}{dt} \right\rangle = \left\langle \frac{dH}{dt} \right\rangle = \frac{d\lambda}{dt}\left\langle \frac{\partial H}{\partial \lambda} \right\rangle \doteq \left\langle \frac{d\lambda}{dt} \right\rangle \left\langle \frac{\partial H}{\partial \lambda} \right\rangle$$

Combine everything

$$\left\langle \frac{dE}{dt} \right\rangle \frac{\partial}{\partial E}\int_0^T pdq + \left\langle \frac{d\lambda}{dt} \right\rangle \frac{\partial}{\partial \lambda}\int_0^T pdq = 0 \tag{4.21}$$

$p(q;E,\lambda)$ may well change rapidly in response to q, but after taking the integral

$$I(E,\lambda) \equiv \frac{1}{2\pi}\int_0^T p(q;E,\lambda)dq$$

this quantity depends on E, λ, which implies I changes slowly over the period. If we assume $\frac{\partial}{\partial E}\int_0^T pdq$, $\frac{\partial}{\partial \lambda}\int_0^T pdq$ change little over a period, we can move them into the average, then left hand side of (4.21) becomes

$$\left\langle \frac{d}{dt} I \right\rangle = 0 \tag{4.22}$$

reconfirming the fact that I is approximately constant over one period. I is an adiabatic invariant.

(4.20) implies

$$2\pi\frac{\partial I}{\partial E} = \frac{\partial}{\partial E}\int_0^T pdq = T$$

and

$$\frac{\partial}{\partial \lambda}\int_0^T pdq = \left\langle \frac{dE}{dt} \right\rangle T$$

Next example shows many of the slowing varying conditions are indeed reasonably accurate.

Example. Back to the spring example

$$H = \frac{p^2}{2m} + \frac{1}{2}kq^2$$

λ can be k or m, or combination of the two. For k, m are constant, q, p diagram is an ellipse, because

$$q = q_0 \cos wt \qquad p = m\dot{q} = -mwq_0 \sin wt$$

so the area of the ellipse

$$I = \frac{\text{area}}{2\pi} = \frac{1}{2}mwq_0^2$$

which is indeed a slow varying quantity. Total energy

$$E = H = \frac{1}{2}mw^2q_0^2$$

is not slow varying, but

$$\left\langle \frac{dE}{dt} \right\rangle T \approx mww'q_0^2 T \approx mw'q_0^2$$

is small and small varying.

Another example

Lecture 20 (4/9/13)

Example. Keplerian orbits

$$H = \frac{p_r^2}{2\mu} + \frac{p_\psi^2}{2\mu r^2} - \frac{k}{r}$$

$k = k(t) > 0$ may due to change in black holes. ψ is cyclic, $p_\psi = L$ is constant (this holds regardless whether k is constant or not, because this is consequence of Hamilton equation), so put

$$H_{eff}(r, p_r; k) = \frac{p_r^2}{2\mu} + V_{eff}(r, k)$$

where $V_{eff} = \frac{L^2}{2\mu r^2} - \frac{k}{r} \sim -\frac{1}{r}$ for long r, then

$$E(t) = \frac{p_r^2}{2\mu} + \frac{L^2}{2\mu r^2} - \frac{k}{r}$$

Look for bounded motion $E < 0$, using variables $(r; E, k)$, solve for p_r,

$$p_r = \sqrt{2\mu\left(E - \frac{L^2}{2\mu r^2} + \frac{k}{r}\right)}$$

Compute

$$I = \frac{1}{2\pi}\oint p_r dr = \frac{1}{\pi}\int_{r_{\min}}^{r_{\max}} \sqrt{2\mu\left(E - \frac{L^2}{2\mu r^2} + \frac{k}{r}\right)} dr$$

Recall $r_{\max/\min} = \frac{P}{1 \mp e}$, eccentricity e and of Latus rectum P, cf (2.12) and follow.

$$P = \frac{L^2}{k\mu}, \quad e = \sqrt{\frac{2PE}{k} + 1} \implies E = \frac{k}{2P}(e^2 - 1)$$

so

$$I = \frac{1}{\pi}\int_{P/1+e}^{P/1-e} \sqrt{2\mu\left(\frac{k}{2P}(e^2-1) - \frac{Pk}{2r^2} + \frac{k}{r}\right)} dr$$

define dimensionless term $\xi = r/P$,

$$I = \frac{1}{\pi}\int_{1/1+e}^{1/1-e} P\sqrt{\frac{2\mu k}{P}}\sqrt{\frac{1}{2}(e^2-1) - \frac{1}{2\xi^2} + \frac{1}{\xi}} d\xi = \frac{\sqrt{2\mu k P}}{\pi} f(e) = \frac{\sqrt{2}}{\pi} L f(e)$$

where $f(e) = \int_{1/1+e}^{1/1-e} \sqrt{\frac{1}{2}(e^2-1) - \frac{1}{2\xi^2} + \frac{1}{\xi}} d\xi$.

That L, I are constants implies $f(e)$ is constant, so as k changes slowly, $e(t)$ is constant. Therefore I constant says the size of phase orbit stays the same. The size of the actual coordinate orbit does change, as seen from numerical calculation[1], but the sharp of the orbits will stay the same.

[1] See Mathematica simulation: http://phys.columbia.edu/~nicolis/G4003.html

5 Field Theory

We now study the mechanics of continuous systems. Reference Goldstein ch13.

5.1 Wave Equation

Consider a 1D infinite chain (i.e. no worry of boundaries) of ∞ many particles connected by springs of k and rest length a apart. Use the same notation in oscillation system,

$$q_\alpha = q_\alpha^{(0)} + \eta_\alpha$$

$\alpha = ..., -1, 0, 1, ...$ labels individual particles, and equilibrium position

$$q_\alpha^{(0)} = \alpha a$$

Later we will take limit $a \to 0$ to make the chain uniform. It turns out that the Lagrangian will survive the continuous limit.

$$T = \sum_\alpha \frac{1}{2} m \dot{q}_\alpha^2 = \sum_\alpha \frac{1}{2} m \dot{\eta}_\alpha^2$$

Potential energy is stored in the spring, the stretch of spring between particle α and $\alpha + 1$ is $(\eta_{\alpha+1} - \eta_\alpha)$

$$V = \sum_\alpha \frac{1}{2} k (\eta_{\alpha+1} - \eta_\alpha)^2$$

so

$$L = \sum_\alpha \frac{1}{2} m \dot{\eta}_\alpha^2 - \frac{1}{2} k (\eta_{\alpha+1} - \eta_\alpha)^2 \tag{5.1}$$

EL of β component:

$$m \ddot{\eta}_\beta + k(\eta_\beta - \eta_{\beta-1}) - k(\eta_{\beta+1} - \eta_\beta) = 0 \tag{5.2}$$

There are two methods to solve above ∞ coupled ODEs:

1) Diagonalize the matrix: as we did before, get normal nodes.
2) field theory method: take continuous limit. This is what we will do.

Put mass density

$$\mu = \frac{m}{a}$$

Multiply $\frac{1}{a}$ to (5.2)

$$\mu\ddot{\eta}_\beta - k[\frac{\eta_{\beta+1} - \eta_\beta}{a} - \frac{\eta_\beta - \eta_{\beta-1}}{a}] = 0 \tag{5.3}$$

Set η_β to $\eta(x)$, which is called displacement field, then

$$\eta_{\beta+1} - \eta_\beta = \eta(x+a) - \eta(x) \approx a\frac{d\eta(x+a)}{dx}$$

so (5.3) becomes

$$\mu\ddot{\eta}_\beta - k[\frac{d}{dx}\eta(x) - \frac{d}{dx}\eta(x-a)] = 0$$

Set $Y \equiv ka$, young's modulus, notice Y is constant as $a \to 0$, because recall the effective spring constant of two springs in series is

$$\frac{1}{k_{eff}} = \frac{1}{k_1} + \frac{1}{k_2} \implies k_1 = 2k_{eff}$$

if $k_1 = k_2$. This shows if reduce a by half, k will be double, so ka is constant. Finally

$$\mu\frac{\partial^2}{\partial t^2}\eta(x,t) - Y\frac{\partial^2}{\partial x^2}\eta(x,t) = 0$$

we obtain 2nd order PDE, wave equation. The solutions are linear superposition of left and right moving waves

$$\eta(x,t) = e^{ikx-iwt}$$

with speed $v = \sqrt{Y/\mu}$, and $w = vk = v2\pi/\lambda$.

This analysis is valid as long as wave length $\lambda \gg a$.

5.2 Field Theory: Lagrangian Viewpoint

Lecture 21 (4/11/13)

Last time, we started from a discrete Lagrangian (5.1), than took limit to get continuous wave equation. Can we get a continuous Lagrangian that lead to wave equation without taking $a \to 0$?

Let us try to convert the summation of the discrete Lagrangian into a integral, by interplay

$$\alpha \longleftrightarrow \frac{x}{a}, \text{ and } \sum_\alpha \longleftrightarrow \int \frac{dx}{a}$$

so

$$L \to \int dx \frac{1}{2}\frac{m}{a}(\partial_t \eta)^2 - \frac{1}{2}k(a\partial_x \eta)^2 = \int dx \left[\frac{1}{2}\mu(\partial_t \eta)^2 - \frac{1}{2}Y(\partial_x \eta)^2\right]$$

Since the total summation is not likely convergent, we make $\mathcal{L}$ Lagrangian density

$$\mathcal{L}(\partial_t \eta, \partial_x \eta) = \frac{1}{2}\mu(\partial_t \eta)^2 - \frac{1}{2}Y(\partial_x \eta)^2 \tag{5.4}$$

so that $L = \int \mathcal{L} dx$.

For higher dimension

$$\mathcal{L}(\partial_t \eta_A, \vec{\nabla}\eta_A, \eta_A), \quad L = \int \underbrace{dxdy...}_{N}\mathcal{L}$$

Label $A = 1, 2, .., N$, $N = 3$ for three displacement field in 3D. $N = 4$ for electromagnetic field.

Now we want to get EOM for the continuous $\mathcal{L}$.

First in 1D.

Use least action principle. When we did action integral before, t was the only ultimately variable. Now both x and t are variables

$$S = \int_{t_1}^{t_2} dt \int_{-\infty}^{\infty} dx \mathcal{L}$$

Here we should let $t_2, t_1 \to \pm\infty$, because EOM should hold for any $t_{2,1}$ and from relativistic view (we'll do EM soon) spatial and temporal are the time.

Having these in mind, we consider an arbitrary deviation

$$\eta(x, t) \to \eta(x, t) + \delta\eta(x, t)$$

such that

$$\begin{cases} \delta\eta(x, \pm\infty) = 0 & \forall x \\ \delta\eta(\pm\infty, t) = 0 & \forall t \end{cases}$$

Therefore the density action of the deviation is

$$\begin{aligned} S[\eta + \delta\eta] &= \int dxdt \mathcal{L}(\partial_t(\eta + \delta\eta), \partial_x(\eta + \delta\eta)) \\ &= \int dxdt \mathcal{L}(\partial_t\eta + \partial_t\delta\eta, \partial_x\eta + \partial_x\delta\eta) \\ &\approx \int dxdt \left[\mathcal{L}(\partial_t\eta, \partial_x\eta) + \frac{\partial\mathcal{L}}{\partial(\partial_t\eta)}\partial_t\delta\eta + \frac{\partial\mathcal{L}}{\partial(\partial_x\eta)}\partial_x\delta\eta\right] \end{aligned}$$

expending 1st order. Then use integration by parts

$$\delta S = \left(\int dx \frac{\partial\mathcal{L}}{\partial(\partial_t\eta)}\delta\eta\right)_{t=-\infty}^{t=\infty} + \left(\int dt \frac{\partial\mathcal{L}}{\partial(\partial_x\eta)}\delta\eta\right)_{x=-\infty}^{x=\infty} - \int dxdt \left(\partial_t \frac{\partial\mathcal{L}}{\partial(\partial_t\eta)} + \partial_x \frac{\partial\mathcal{L}}{\partial(\partial_x\eta)}\right)\delta\eta$$

So EOM

$$\partial_t \frac{\partial\mathcal{L}}{\partial(\partial_t\eta)} + \partial_x \frac{\partial\mathcal{L}}{\partial(\partial_x\eta)} = 0$$

or use label $\mu = t$ or x and Einstein summation, EOM

$$\partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\eta)} = 0$$

To check the validity of above, we use (5.4), and indeed we get 1D wave equation.

For more general 3D N coupled fields

$$\mathcal{L} = \mathcal{L}(\eta_A, \partial_t\eta_A, \partial_x\eta_A)$$

$A = 1, ..., N$, EOM

$$\partial_t \frac{\partial\mathcal{L}}{\partial(\partial_t\eta_A)} + \vec{\nabla} \cdot \frac{\partial\mathcal{L}}{\partial(\vec{\nabla}\eta_A)} - \frac{\partial\mathcal{L}}{\partial\eta_A} = 0 \quad \forall A \tag{5.5}$$

which looks more and less like the EL. The ∂_t, $\vec{\nabla}$ in front look like partial derivative, but they actually behave just like total derivative in EL, because now t, x are independent.

5.3 Electromagnetic Field Theory

When we got the continuous $\mathcal{L}$ before, we started from a discrete model. Now we have to start everything from continuous things.

We define EM $\mathcal{L}$ to be very special

$$\mathcal{L} = (\text{density of } T)\text{-}(\text{density of } V)$$

where

$$\text{density of } T\text{=part of energy density involves time derivatives}$$

$$\text{density of } V\text{=the rest of the energy density not included in } T$$

Caveat: as we will see that is not equal to "part of energy density involves spatial derivative".

We now do EM field in vacuum ($\rho = 0$, $\vec{j} = 0$) in cgs units.

Let ϕ be electrostatic potential, $\vec{A}$ be vector potential. Hence ($\vec{E}$, $\vec{B}$) field is characterized by

$$\eta_A = (\phi, \vec{A}) = (\phi, A_x, A_y, A_z)$$

$A = 0, 1, 2, 3$.

Recall from EM

$$\epsilon = \frac{1}{8\pi}(E^2 + B^2)$$

and

$$\vec{B} = \nabla \times \vec{A} \quad \vec{E} = -\nabla\phi - \frac{1}{c}\partial_t\vec{A} \tag{5.6}$$

So only $\vec{E}$ depends on ∂_t, but it is wrong to decompose $E^2 \sim (\partial_t A)^2 + (\partial_t A)(\nabla\phi) + (\nabla\phi)^2$ and take only the time part to be T, because $\phi, \vec{A}$ are not physical, manifest gauge invariance should not be violated,

$$\begin{cases} \vec{A} \to \vec{A} + \nabla\Lambda \\ \phi \to \phi - \frac{1}{c}\partial_t\Lambda \end{cases}$$

So we do still keep terms in $\vec{E}$ together,

$$T = \frac{1}{8\pi}E^2 \quad V = \frac{1}{8\pi}B^2$$

and

$$\mathcal{L} = \frac{1}{8\pi}(E^2 - B^2)$$

Lecture 22 (4/16/13)

We now verify above $\mathcal{L}$ is indeed the right choice, by showing it gives the right EOM: Maxwell's equations.

Two of the Maxwell are free

$$(5.6) \implies \begin{cases} \vec{\nabla} \cdot \vec{B} = 0 \\ \vec{\nabla} \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t} \end{cases}$$

so

$$\mathcal{L} = \frac{1}{8\pi}\left[(\nabla\phi)^2 + \frac{2}{c}(\partial_t \vec{A})(\nabla\phi) + \frac{1}{c^2}(\partial_t \vec{A})^2 - \sum_i (\epsilon^{ijk}\partial_j A_k)^2\right] \tag{5.7}$$

Recall $\epsilon^{ijk}\epsilon^{ilm} = \delta^{jl}\delta^{km} - \delta^{jm}\delta^{kl}$,

$$\begin{aligned} (\epsilon^{ijk}\partial_j A_k)^2 &= \epsilon^{ijk}\epsilon^{ilm}\partial_j A_k \partial_l A_m \\ &= \partial_j A_k \partial_j A_k - \partial_j A_k \partial_k A_j \\ &= (\partial_j A_k)^2 - \partial_j A_k \partial_k A_j \end{aligned}$$

Apply (5.5) for $A = 0$ to (5.7)

$$\vec{\nabla} \cdot \frac{1}{8\pi}\underbrace{2\nabla\phi + \frac{2}{c}\partial_t \vec{A}}_{\vec{E}} = 0 \implies -\frac{1}{4\pi}\nabla \cdot \vec{E} = 0$$

Gauss law for $\rho = 0$. Here once again we see that gauge invariance is manifested, as we group $\vec{E}$ terms together.

Apply (5.5) for $A = i = 1, 2$ or 3 to (5.7)

$$\frac{\partial \mathcal{L}}{\partial(\partial_t A_i)} = \frac{1}{8\pi}\left(\frac{2}{c}\partial_i\phi + \frac{2}{c^2}\partial_t A_i\right) = -\frac{1}{4\pi c}E_i$$

$$\frac{\partial \mathcal{L}}{\partial(\partial_l A_i)} = -\frac{1}{8\pi}\left(2\partial_l A_i - 2\partial_i A_l\right)$$

so we get

$$\partial_t\left(-\frac{1}{4\pi}E_i\right) + \sum_l \partial_l \frac{-1}{4\pi}\left(\partial_l A_i - \partial_i A_l\right) = 0$$

which is

$$-\frac{1}{c}\partial_t \vec{E} + \underbrace{\left(-\nabla^2 \vec{A} + \vec{\nabla}(\vec{\nabla}\cdot\vec{A})\right)}_{\nabla\times(\nabla\times\vec{A})} = 0$$

So we recovered Ampere law

$$\nabla \times \vec{B} = \frac{1}{c}\frac{\partial \vec{E}}{\partial t}$$

Relativistic EM

Our previous $(\phi, \vec{A})$ changes to $A_\mu = (-\phi, \vec{A})$, 4-vector potential. Then define field strength

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

which is the first object that is both gauge invariant and Lorentz invariant. But to get that we need to have a correct definition of derivatives.

Then

$$\mathcal{L} = \frac{1}{8\pi}(E^2 - B^2) = -\frac{1}{16\pi}F_{\mu\nu}F^{\mu\nu}$$

which is also gauge invariant and Lorentz invariant.

5.4 Field Theory: Hamilton Viewpoint

We want to find Hamiltonian for the field. Similarly starting from (5.1), we get conjugate momentum of the displace η_β is

$$p_\beta = \frac{\partial L}{\partial \dot{\eta}_\beta} = m\dot{\eta}_\beta = (\mu a)\dot{\eta}_\beta$$

so

$$H = \sum_\beta p_\beta \dot{\eta}_\beta - L$$

In the continuous limit, $a \to 0$, $\sum_\beta \to \int \frac{dx}{a}$, we get

$$\begin{aligned} H &\to \int \frac{dx}{a}(\mu a)(\partial_t \eta)^2 - \int dx \left(\frac{1}{2}\mu(\partial_t\eta)^2 - \frac{1}{2}Y(\partial_x\eta)^2\right) \\ &= \int dx \left(\frac{1}{2}\mu(\partial_t\eta)^2 + \frac{1}{2}Y(\partial_x\eta)^2\right) \end{aligned}$$

suggesting to define Hamilton density

$$\mathcal{H} = \frac{1}{2}\mu(\partial_t\eta)^2 + \frac{1}{2}Y(\partial_x\eta)^2$$

But we don't want $\mathcal{H}$ to be function of $\partial_t\eta$, $\partial_x\eta$, normal $\mathcal{H}$ should be function of $\partial_x\eta$ and some sort of momentum. In fact we want momentum density, defined as

$$\Pi(x,t) \equiv \lim_{a\to 0}\frac{p_\beta}{a} = \mu\partial_t\eta$$

notice $\mu = \mu(x,t)$ is not necessary constant. We also see Π has a similar relationship to $\mathcal{L}$ as p_β to L

$$\Pi = \frac{\partial\mathcal{L}}{\partial(\partial_t\eta)} \qquad p_\beta = \frac{\partial L}{\partial\dot{\eta}_\beta}$$

Using Π, we can rewrite

$$H = \int dx\mathcal{H} = \int dx\left(\frac{1}{2\mu}\Pi^2 + \frac{Y}{2}(\partial_t\eta)^2\right)$$

and

$$\mathcal{H} = \Pi\partial_t\eta_\alpha - \mathcal{L}$$

Energy of EM field

Lecture 23 (4/18/13)

As a back check, we want to see that Hamilton gives the right energy of EM field in vacuum. $\eta_A = (\phi, \vec{A})$, using calculation we had before

$$\begin{aligned}
\mathcal{H} &= \sum_A \Pi^A \partial_t \eta_A - \mathcal{L} \\
&= \sum_A \frac{\partial \mathcal{L}}{\partial(\partial_t \eta_A)} \partial_t \eta_A - \mathcal{L} \\
&= -\frac{1}{4\pi c} \vec{E}(\partial_t \vec{A}) - \mathcal{L} \\
&= \frac{1}{4\pi} \vec{E}(\vec{E} + \nabla\phi) - \frac{1}{8\pi}(E^2 - B^2) \\
&= \frac{1}{8\pi}(E^2 + B^2) + \frac{1}{4\pi} \vec{E} \cdot \vec{\nabla}\phi
\end{aligned}$$

The second doesn't matter if we integrate $\mathcal{H}$ over all spaces, because

$$\iiint \vec{E} \cdot \vec{\nabla}\phi d^3x = \oint\!\!\oint_S (\vec{E}\phi) d\vec{a} - \iiint \phi \underbrace{\nabla \cdot \vec{E}}_{0} d^3x$$

The first term on the right is too 0, because Maxwell must be supplemented by consistent boundary conditions, i.e. both $\vec{E}$, ϕ must be 0 at ∞.

5.5 Quantum Mechanical Field Theory

Hamilton Formulation: Commutator

Recall properties of Poisson bracket (4.2) − (4.6). They work exactly the same for commutator of matrices A, B, putting

$$[A, B] = AB - BA$$

So let us promote q, p to matrices $\hat{q}$, $\hat{p}$. For function $f(q, p)$, first Taylor expand

$$f(q,p) = f(0,0) + \frac{\partial f}{\partial q}(0,0)q + \frac{\partial f}{\partial p}(0,0)p + \frac{1}{2}\frac{\partial^2 f}{\partial q^2}(0,0)q^2 + ...$$

then substitute in $\hat{q}$, $\hat{p}$, to get

$$f(\hat{q},\hat{p}) = f(0,0)I + \frac{\partial f}{\partial q}(0,0)\hat{q} + \frac{\partial f}{\partial p}(0,0)\hat{p} + \frac{1}{2}\frac{\partial^2 f}{\partial q^2}(0,0)\hat{q}^2 + ..$$

In terms of measurements (i.e. observables), we require further that $\hat{q}^\dagger = \hat{q}$ and $\hat{p}^\dagger = \hat{p}$, hence measurable f has to be hermitian.

Canonical Poisson

$$\begin{cases} [q,q] = 0 \\ [p,p] = 0 \\ [q,p] = 1 \end{cases}$$

becomes commutator relation

$$\begin{cases} [\hat{q},\hat{q}] = 0 \\ [\hat{p},\hat{p}] = 0 \\ [\hat{q},\hat{p}] = i\hbar I \end{cases}$$

The $i, \hbar$ factors added to the commutator because

1) $[\hat{q},\hat{p}]$ is antihermitian. Indeed

$$[\hat{q},\hat{p}]^\dagger = (\hat{q}\hat{p} - \hat{p}\hat{q})^\dagger = (\hat{p}^\dagger\hat{q}^\dagger - \hat{q}^\dagger\hat{p}^\dagger) = -[\hat{q},\hat{p}]$$

2) adding $\hbar$ because Poisson bracket is much bigger than commutator.

Therefore we have shown that one may equate the two

$$i\hbar[\ .\ ]_{pb} = [\ ,\ ]_c \tag{5.8}$$

Let $f(\hat{q},\hat{p})$ be state of the system. $f \in H$ called Hilbert space: all possible states of the system. H is a vector space. It has ∞ dimensions, and it has positive definite Hermitian scalar product: For all states ψ, ϕ

$$(\psi,\phi) = (\phi,\psi)^* \text{ and } (\psi,\psi) > 0$$

What have we done?

$$\text{classical states} \to \text{vectors}$$

$$\text{phase space} \to \text{Hilbert space}$$

$$\text{experimental outcome for observable} \to \text{eigenvalues of } f(\hat{q},\hat{p})$$

If state ϕ is an eigenvector of $f(\hat{q},\hat{p})$ with eigenvalue λ, then measurement returns λ.

If ϕ is not an eigenvector, measurement returns eigenvalues, $\lambda_1, ..., \lambda_n, ..$ with probabilities, $p_1, ..., p_n, ..$, i.e. we get certain probability distributions after multiple trails, and the expectation value of the measurement

$$\sum \lambda_i p_i = \langle f(\hat{q},\hat{p})\rangle_\phi = (\phi, f(\hat{q},\hat{p})\phi)$$

where ϕ is normalized $(\phi,\phi) = 1$.

Uncertainty Principle

One interesting feature of commutator is to do with uncertainty principle, which only assume certain mathematical properties of the inner product.

Consider the system is in state ψ, measure $\hat{q}$, $\hat{p}$. The expectation value of

$$\langle\hat{q}\rangle = (\psi, \hat{q}\psi), \qquad \langle\hat{p}\rangle = (\psi, \hat{p}\psi)$$

Put operator deviation (they are hermitian)

$$\Delta\hat{q} = \hat{q} - \langle\hat{q}\rangle I \qquad \Delta\hat{p} = \hat{p} - \langle\hat{p}\rangle I \tag{5.9}$$

and mean of the deviation square

$$\Delta q^2 \equiv \left\langle(\Delta\hat{q})^2\right\rangle = (\psi, (\Delta\hat{q})^2\psi) = |\Delta\hat{q}\psi|^2$$

similar for Δp^2.

Applying Cauchy Schwartz on vectors $\Delta\hat{q}\psi$, $\Delta\hat{p}\psi$

$$\begin{aligned}
\Delta q^2 \Delta p^2 = \|\Delta\hat{q}\psi\|^{\,2} \, \|\Delta\hat{p}\psi\|^{\,2} &\geq |(\Delta\hat{q}\psi)\cdot(\Delta\hat{p}\psi)|^2 \\
&= |(\psi, \Delta\hat{q}\Delta\hat{p}\psi)|^2 \\
&= \Re^2(\psi, \Delta\hat{q}\Delta\hat{p}\psi) + \Im^2(\psi, \Delta\hat{q}\Delta\hat{p}\psi) \\
&\geq \Im^2(\psi, \Delta\hat{q}\Delta\hat{p}\psi) \\
&= \left(\frac{(\psi, \Delta\hat{q}\Delta\hat{p}\psi) - (\psi, \Delta\hat{p}^\dagger\Delta\hat{q}^\dagger\psi)}{2i}\right)^2 \\
&= \left(\frac{(\psi, [\Delta\hat{q}, \Delta\hat{p}]\psi)}{2i}\right)^2 \\
&= \left(\frac{(\psi, [\hat{q}, \hat{p}]\psi)}{2i}\right)^2 \\
&= \left(\frac{\hbar}{2}\right)^2
\end{aligned}$$

Hence

$$\Delta q \Delta p \geq \frac{\hbar}{2}$$

where Δq, Δp are the root mean square $\sqrt{\langle(\Delta\hat{q})^2\rangle}$, not (5.9).

Time Evolution

Use (5.8), we may write down some quantum mechanical version of Hamilton equations

$$\begin{aligned}
\frac{d}{dt}\hat{q}(t) &= \frac{1}{i\hbar}[\hat{q}, \hat{H}]_c \\
\frac{d}{dt}\hat{p}(t) &= \frac{1}{i\hbar}[\hat{p}, \hat{H}]_c \\
\hat{H} &= \frac{\hat{p}^2}{2m} + V(\hat{q})
\end{aligned}$$

Let's see what they imply.

One example is to look at operators that commutes with $\hat{H}$. E.g. $V(q) = 0$ free particles, then

$$[\hat{p}, \hat{H}] = \frac{1}{2m}[\hat{p}, \hat{p}^2] = 0$$

then $\frac{d}{dt}\hat{p}(t) = 0$, so $\hat{p}$ is constant. What does this mean? What we have is in Heisenberg picture: operator evolve in time, states are fixed.

While in Schrodinger picture: operator are fixed, states evolves. The two pictures are equivalent, by

$$i\hbar \frac{d}{dt}\psi = \hat{H} \cdot \hat{\psi}$$

Lagrange Formulation: Path Integral

Suppose we know at $t = t_i$, the system is in one of eigenstate of $\hat{q}$ with eigenvalue q_i. At later time t_f, what is the probability that measurement returns eigenvalue q_f,

$$P(q_i, t_i \to q_f, t_f) = |(\psi_{q_f}(t_f), \psi_{q_i}(t_i))|^2$$

What inside of module is called transition amplitude. Feynman derived that one can compute this transition amplitude without all the terminologies of operators, commutators, etc.

$$(\psi_{q_f}(t_f), \psi_{q_i}(t_i)) = \int_{q(t_i)=q_i}^{q(t_f)=q_f} \mathcal{D}q(t) e^{\frac{i}{\hbar}S_{cl}} \tag{5.10}$$

$$S_{cl}[q(t)] = \int_{t_i}^{t_f} dt \mathcal{L}(q, \dot{q})$$

for example in Cartesian coordinate

$$S_{cl}[q(t)] = \int dt \left[\frac{1}{2}m\dot{q}^2 - V(q)\right]$$

Lecture 24 (4/23/13)

This is ill-defined in string theory, but good for QM. Some things are easier to show by path integral.

We can examine (5.10) a little bit closely.

$$\mathcal{D}q(t) = (dq_1 \cdots dq_N)_{N\to\infty}$$

∞ dimensional integral. $e^{\frac{i}{\hbar}S_{cl}}$ is complex, so it is not simply a probabilistic weight. One can show that technically $e^{\frac{i}{\hbar}S_{cl}}$ is not very different from $e^{-\frac{1}{\hbar}S_{cl}}$, for $S_{cl} \gg \hbar$.

6 Rigid Body

We study dynamics of "zero indices". Reference L.L ch 6, Goldstein ch 4,5.

6.1 Euler Angles

Rigid body of N particles iff

$$|\vec{r}_a - \vec{r}_b| = \text{constant in time} \quad \forall a, b = 1, ..., N$$

We know that for N particles there are 3N degrees of freedom if the particles are allowed to move wrt each other. What about now?

Claim: there are 6 degrees of freedom. only motion (overall translations and rotations)

Proof: Once we know three of the N particle's trajectories: $\vec{r}_1$, $\vec{r}_2$, and $\vec{r}_3(t)$, we know all other trajectories. That is because the intersections of three spheres are at most two separated points, and since the motion is continuous, only one of the two is the destination. So now degrees of freedom is 9, but

$$|\vec{r}_1 - \vec{r}_2|, \ |\vec{r}_2 - \vec{r}_3|, \ |\vec{r}_1 - \vec{r}_3|$$

are also fixed, so degree of freedom is

$$9 - 3 = 6 = 3 + 3$$

So we would like label them as 3 make up translation (x, y, z) and 3 make up Euler angles (ϕ, ψ, θ)

Determine Euler angles as following steps

0) start the body with xyz axes fixed in space.

1) rotate the body around z by ϕ, the axes are rotated too, $xyz \to \xi\eta z$

2) rotate around ξ by θ, $\xi\eta z \to \xi\eta' z'$

3) rotate around z' by ψ, $\xi\eta' z' \to x'y'z'$

Therefore $\vec{R}$, center mass point and (ϕ, ψ, θ) specifies the new position of the body.

6.2 Infinitesimal Displacement

Infinitesimal displacement, particular infinitesimal rotation is what we study now. This is much easier than finite rotation.

Take a point in the body, $\vec{R}$ position of CM

$$\vec{r} = \vec{R} + \vec{r}' \tag{6.1}$$

$$d\vec{r} = d\vec{R} + d\vec{r}'$$

$\vec{r}'$ is the displacement from the center mass point, so it can only be rotational. Put

$$d\vec{r}' = (\hat{n} \times \vec{r}')d\phi$$

where $\hat{n}$ is the rotational axis going through CM point, and $d\phi$ is rotational angle, the angle swept by the perpendicular component of $\vec{r}'$, then

$$\frac{d\vec{r}}{dt} = \frac{d\vec{R}}{dt} + (\hat{n} \times \vec{r}')\frac{d\phi}{dt} \tag{6.2}$$

so infinitesimal rotation means the rotational axis itself is fixed, so we only consider the first step in Euler angle.

From (6.2), the instantaneous velocity of any point p of rigid body is

$$\vec{v} = \vec{V} + \vec{\Omega} \times \vec{r}' \tag{6.3}$$

where

$$\vec{\Omega} \equiv \hat{n}\frac{d\phi}{dt}$$

called angular velocity vector.

The derivation of (6.3) seems to say that $\vec{\Omega}$ has to go through the CM point, $\vec{V}$ is the velocity of the center mass (whether CM itself is rotating or not doesn't matter, it is still $\vec{V}$), $\vec{r}'$ is the vector pointing from CM to the point p.

Suppose the rotational axis passes other point O' instead of CM. Let $\vec{\Delta}$ be the vector pointing from CM to O', and $\vec{r}''$ be the vector pointing from O' to p, so

$$\vec{r}' = \vec{r}'' + \vec{\Delta}$$

compute velocity of point p

$$\begin{aligned}\vec{v} &= \vec{V} + \vec{\Omega} \times (\vec{r}'' + \vec{\Delta}) \\ &= (\vec{V} + \vec{\Omega} \times \vec{\Delta}) + \vec{\Omega} \times \vec{r}''\end{aligned}$$

By (6.3) $\vec{V}_{O'} = \vec{V} + \vec{\Omega} \times \vec{\Delta}$ is the velocity of point O', and $\vec{r}''$ is now from O' to p, so the formula (6.3) is the same for any other rotational axes.

Because the formula is the same, we can interpret the motion of the points on a rigid body either as the sum of the motion of the CM plus rotation about the CM, or as the sum of the motion of any other point O' plus rotation about O', as long as $\vec{\Omega}$ is the same, i.e. the direction of the rotational axis $\hat{n}$ and angular speed $d\phi/dt$ are the same.

6.3 Kinetic Energy

of N particles of the rigid body

$$\begin{aligned}T = \sum_{a=1}^{N} \frac{1}{2} m_a \dot{\vec{r}}_a^2 &= \sum \frac{1}{2} m_a (\vec{V} + \vec{\Omega} \times \vec{r}'_a)^2 \\ &= \sum \frac{1}{2} m_a V^2 + m_a \vec{V} \cdot (\vec{\Omega} \times \vec{r}'_a) + \frac{1}{2} m_a (\vec{\Omega} \times \vec{r}'_a)^2\end{aligned}$$

Let us evaluate the three terms on the right

1. gives

$$\frac{1}{2} M V^2$$

2. become

$$\vec{V} \cdot \sum (\vec{\Omega} \times m_a \vec{r}'_a) = \vec{V} \cdot (\vec{\Omega} \times \underbrace{\sum m_a \vec{r}'_a}_{0}) = 0$$

3. we simplify

$$\begin{aligned}(\vec{\Omega}\times\vec{r}'_a)^2 &= \Omega^2 r_a'^2 \sin^2\theta \\ &= \Omega^2 r_a'^2 - \Omega^2 r_a'^2\cos^2\theta \\ &= \Omega^2 r_a'^2 - (\vec{\Omega}\cdot\vec{r}'_a)^2\end{aligned}$$

Therefore

$$T = \frac{1}{2}M\dot{\vec{R}}^2 + \frac{1}{2}\sum_a m_a\left[\Omega^2 r_a'^2 - (\vec{\Omega}\cdot\vec{r}'_a)^2\right]$$

6.4 Inertia Tensor

Use Eisenstein convention, write

$$\begin{aligned}\Omega^2 r_a'^2 - (\vec{\Omega}\cdot\vec{r}'_a)^2 &= \Omega_i\Omega_i r_a'^2 - \Omega_i r_a'^i \Omega_j r_a'^j \\ &= \Omega_i\Omega_j(\delta_{ij}r_a'^2 - r_a'^i r_a'^j)\end{aligned}$$

Define inertia tensor (a matrix generalization of inertia moments)

$$I_{ij} = \sum m_a(\delta_{ij}r_a'^2 - r_a'^i r_a'^j)$$

Since $\vec{\Omega}$ is the same for all points on the body, we can carefully pull out $\vec{\Omega}$ out of summation. Hence

$$T = \frac{1}{2}M\dot{\vec{R}}^2 + \frac{1}{2}I_{ij}\Omega_i\Omega_j$$

summing over i, j is of course assumed.

The continuous version

$$I_{ij} = \int d^3\vec{r}'\rho(\vec{r}')(\delta_{ij}r_a'^2 - r_a'^i r_a'^j)$$

$\vec{r}'$ is referred to origin at $\vec{R}$.

Properties of I_{ij}

Lecture 25 (4/25/13)

1) Symmetric $I_{ij} = I_{ji}$

2) Discrete version looks like

$$\hat{I} = \sum_{a=1}^{N} m_a \begin{pmatrix} y_a'^2 + z_a'^2 & -x_a' y_a' & -x_a' z_a' \\ & x_a'^2 + z_a'^2 & -y_a' z_a' \\ & & x_a'^2 + y_a'^2 \end{pmatrix}$$

3) Positive definite for $m_a \geq 0$ or $\rho \geq 0$. Pf

$$\vec{w}^T \hat{I} \vec{w} = \int d^3\vec{r}' \rho(\vec{r}') \left[r'^2 w^2 - (\vec{w} \cdot \vec{r}')^2 \right] \geq 0$$

because Cauchy Schwartz $r'^2 w^2 \geq (\vec{w} \cdot \vec{r}')^2$.

4) From 1)+3), $\hat{I}$ has 3 real non-negative eigenvalue $I_{1,2,3}$ called principal moments of inertia.

5) By spectral theorem, $\hat{I}$ can be diagonalized via a rotation ($SO(3)$)

If $I_{1,2,3}$ are all different, then there are 3 real orthogonal eigenvectors $\vec{w}_{1,2,3}$.

Proof: $\hat{I}\vec{w}_A = I_A \vec{w}_A$, $\vec{w}_B^T \hat{I} = \vec{w}_B^T I_B \implies 0 = (I_B - I_A)\vec{w}_B^T \cdot \vec{w}_A \implies \vec{w}_B^T \cdot \vec{w}_A = \delta_{AB}$.

So $\{\vec{w}_{1,2,3}\}$ is an orthonormal basis, put

$$\hat{R} = \begin{pmatrix} | & | & | \\ w_1 & w_2 & w_3 \\ | & | & | \end{pmatrix}$$

check $\hat{R}^T \hat{I} \hat{R}$ diagonalize $\hat{I}$, indeed

$$\begin{aligned} (\hat{R}^T \hat{I} \hat{R})_{il} &= R_{ji} I_{jk} R_{kl} = (w_i)_j I_{jk} (w_l)_k \\ &= (w_i)_j (\hat{I}\vec{w}_l)_j = (w_i)_j I_l (\vec{w}_l)_j = I_l (w_i)_j (w_l)_j = I_l \delta_{il} \end{aligned}$$

After choosing principal axes

$$I = \sum m_a \begin{pmatrix} y_a 2 + {z_a}^2 & & \\ & x_a^2 + z_a^2 & \\ & & x_a^2 + y_a^2 \end{pmatrix}$$

Since principal axes are calculated at very instant of the body, their positions change with the body. Later we will show however they are not "fixed" with the body.

If not all eigenvalue distinct, use Gram-Schmidt, to get an orthonormal basis.

E.g. Homogenous cylinder has $I_2 = I_3$, and cube has all three I the same, so cube is indistinguishable from a sphere from the viewpoint of rotation.

We call all three $I_{1,2,3}$ are distinct as asymmetric top. $I_{1,2}$ same as symmetric top. $I_{1,2,3}$ same as spherical top.

Exercise. Compute I for a thin rod of length L, diameter ϵ, setting on z axis. Choose three principal axes

$$I_z = \int d^3 r \rho(\vec{r})(x^2 + y^2) \sim M\epsilon^2$$

$$I_x = I_y \;\; = \;\; \int d^3 r \rho(r)(y^2 + z^2)$$

assume uniform density, and ignore y^2 integral because it's order of ϵ^2.

$$I_x = \int_{-L/2}^{L/2} dz \lambda z^2 = \frac{1}{3}\lambda \frac{L^3}{4} = \frac{1}{12} M L^2$$

6) Triangle inequality, any $I_{1,2,3}$ is less or equal than the sum of the other two.

$$I_3 \leq I_1 + I_2 \text{ etc}$$

Proof:

$$I_1 + I_2 = \int d^3 r \rho(\vec{r}) \left[(y^2 + z^2) + (x^2 + z^2) \right] \geq I_3 + 2 \int d^3 r \rho z^2 \geq I_3$$

7) We showed the formula (6.3) is the same for any rotation axis, but in the derivation of kinetic energy we assume $\vec{r}_a^{\,\prime}$ is with respect to CM.

What happen to I_{ij}, if we parallel shift $\vec{\Omega}$ from passing through CM to passing through O'?

Similarly as before let $\vec{\Delta}$ be the vector pointing from CM to O', and $\vec{r}''$ be the vector pointing from O' to particle a at point p, so

$$\vec{r}'_a = \vec{r}''_a + \vec{\Delta}$$

We have

$$\begin{aligned} T &= \sum \frac{1}{2} m_a (\vec{V}_{O'} + \vec{\Omega} \times \vec{r}''_a)^2 \\ &= \sum \frac{1}{2} m_a V_{O'}^2 + m_a \vec{V}_{O'} \cdot (\vec{\Omega} \times \vec{r}''_a) + \frac{1}{2} m_a (\vec{\Omega} \times \vec{r}''_a)^2 \end{aligned}$$

If $\vec{V}_{O'}$ is zero, which is the most common situation, then we can still define

$$\begin{aligned} I'_{ij} &= \sum m_a (\delta_{ij} r''^2_a - r''^i_a r''^j_a) \\ &= \sum m_a \left[(r'^2_a + \Delta^2 - 2\vec{\Delta} \cdot \vec{r}'_a)\delta_{ij} - (r'^i_a - \Delta^i)(r'^j_a - \Delta^j) \right] \\ &= I_{ij} - (2\vec{\Delta} \cdot \underbrace{\sum m_a \vec{r}'_a}_{0})\delta_{ij} + \Delta^i \underbrace{\sum m_a r'^j_a}_{0} + \Delta^j \underbrace{\sum m_a r'^i_a}_{0} + \sum m_a (\Delta^2 \delta_{ij} - \Delta^i \Delta^j) \\ &= I_{ij} + M(\Delta^2 \delta_{ij} - \Delta^i \Delta^j) \end{aligned}$$

where I_{ij} is the inertia tensor with respect to CM, and the second term is the inertia moment of point O' wrt to CM.

Notice inertia tensor contains no information of $\vec{\Omega}$ so the direction of the rotational axis doesn't matter, only the point CM or O' matters, i.e. the assumption of parallel shift of $\vec{\Omega}$ from CM to O' is just to make derivation simple, and the resulting I'_{ij} needs not to rotate under the same direction of the rotational axis of I_{ij}.

6.5 Angular Momentum

Choose CM as rotational center, find

$$\vec{L} = \sum m_a \vec{r}_a \times \vec{v}_a$$

and $\vec{r}_a$, $\vec{v}_a$ are given in (6.1), (6.3). We compute

$$\vec{L} = \sum m_a \left[\vec{R} \times \vec{V} + \vec{R} \times (\vec{\Omega} \times \vec{r}_a^{\,\prime}) + \vec{r}_a^{\,\prime} \times \vec{V} + \vec{r}_a^{\,\prime} \times (\vec{\Omega} \times \vec{r}_a^{\,\prime})\right]$$

Two terms in the middle are 0, for $\sum m_a \vec{r}_a^{\,\prime} = 0$. The last term is dealt with BAC-CAB

$$\sum m_a \left(\vec{r}_a^{\,\prime} \times (\vec{\Omega} \times \vec{r}_a^{\,\prime})\right) = \sum m_a \left(\vec{\Omega} r_a^{\prime 2} - \vec{r}_a^{\,\prime}(\vec{\Omega} \cdot \vec{r}_a^{\,\prime})\right) = \hat{I} \cdot \vec{\Omega}$$

Indeed

$$\begin{aligned} (\hat{I} \cdot \vec{\Omega})_i &= I_{ij}\Omega_j = \sum m_a(\delta_{ij} r_a^{\prime 2} \Omega_j - r_a^{\prime i} r_a^{\prime j} \Omega_j) \\ &= \sum m_a \left(\Omega_i r_a^{\prime 2} - r_a^{\prime i}(\vec{\Omega} \cdot \vec{r}_a^{\,\prime})\right) \end{aligned}$$

Hence

$$\vec{L} = M\vec{R} \times \vec{V} + \hat{I} \cdot \vec{\Omega}$$

Even if $\vec{R} = 0$, in general $\vec{L} \nparallel \vec{\Omega}$, because of the tensor $\hat{I}$. But if we use principal axes then

$$L_{1,2,3} = I_{1,2,3}\Omega_{1,2,3}$$

So for asymmetric top, $\vec{L} \parallel \vec{\Omega}$ iff $\vec{\Omega}$ is parallel to one of the principal axes.

For symmetric top ($I_1 = I_2 \neq I_3$), $\vec{L} \parallel \vec{\Omega}$ iff $\Omega_3 = 0$ or $\Omega_1 = \Omega_2 = 0$. E.g. a rotator (thin rod we study last time $I_3 \sim 0$, $I_1 = I_2$), $\vec{L} \parallel \vec{\Omega}$ iff $\Omega_3 = 0$, because $\Omega_1 = \Omega_2 = 0$, i.e. only rotate along z is unphysical.

For spherical top, $\vec{L} \parallel \vec{\Omega}$ always.

Lecture 26 (4/30/13)

6.6 Euler Equation

There are some technical similarities between Euler equation for rigid bodies and Euler equation for incompressible fluid in hydrodynamics.

From Newton's equation for momentum and angular momentum

$$\begin{cases} \frac{d\vec{P}}{dt} = \vec{F} = \sum \vec{F}_a^{ext} & 1 \\ \frac{d\vec{L}}{dt} = \vec{N} = \sum \vec{r}_a \times \vec{F}_a^{ext} & 2 \end{cases} \tag{6.4}$$

Internal forces cancel because they are contract forces, i.e. same $\vec{r}_a$ for both force and reaction. We have 6 components equations match degree of freedom, so they suffice to solve the system.

(6.4)-1 tells how the CM moves

$$\vec{P} = M\vec{V}, \text{ or } \vec{F} = M\dot{\vec{V}}$$

(6.4)-2 tells how the body rotates about the CM. That is what we study. So we study rotation only and the origin is CM. It is simpler to use principal axes $\vec{w}_{1,2,3}$, so we need to convert infinitesimal rotation between lab frame and the principal axes frame. This will give us Euler equations. Just like in fluid dynamics when we switch to the moving coordinate, the coordinate axes are moving with the fluid element, but the velocity of fluid element is still measured respect to the lab frame. So the moving coordinate is not a definite moving reference frame. If we had literally fixed the principal axes with the body, then in the principal axes frame there would be no motion and the principal axes frame would not be inertial either.

What we do is at very instant to look at the position of the body, calculate $\vec{w}_{1,2,3}$ at that position, and fix $\vec{w}_{1,2,3}$ in the space, then look at the instantaneous motion of the body wrt the $\vec{w}_{1,2,3}$ frame, i.e. find angular velocity $\Omega_{1,2,3}$, whose scalar values are wrt $\vec{w}_{1,2,3}$, and at the same time $\Omega_{1,2,3}$ are eigenvectors of $\hat{I}$, so they are parallel to $\vec{L}$. This way allows us to work in an semi inertial frame and have very simple geometric pictures.

For any vector $\vec{A}$, in particular $\vec{L}$, under time change $t \to t + dt$, say vector $\vec{A}$ change to $\vec{A} + d\vec{A}$. We will decompose the change, $d\vec{A}$, into two parts

$$(d\vec{A})_{Lab} = (d\vec{A})_\alpha + (d\vec{A})_\beta \tag{6.5}$$

$(d\vec{A})_\alpha$ is the motion wrt the principal axes, $(d\vec{A})_2$ is the motion moving with the principal axes. Later when we do rotational top to exemplify the combination of these two motions.

Put $\vec{A} = \vec{L}$, and take t derivative

$$(\frac{d\vec{L}}{dt})_{Lab} = \vec{N}$$

$$(\frac{d\vec{L}}{dt})_\alpha = (\frac{dL_1}{dt})_\alpha + (\frac{dL_2}{dt})_\alpha + (\frac{dL_3}{dt})_\alpha = I_1\dot{\vec{\Omega}}_1 + I_2\dot{\vec{\Omega}}_2 + I_3\dot{\vec{\Omega}}_3$$

where $L_{1,2,3}$ is the component of $\vec{L}$ wrt the principal axes.

$$(\frac{d\vec{L}}{dt})_\beta = (\vec{n} \times \vec{L})\frac{d\phi}{dt} = \vec{\Omega} \times \vec{L}$$

Here use the same $\vec{\Omega}$ for both $(\frac{d\vec{L}}{dt})_\alpha$ and $(\frac{d\vec{L}}{dt})_\beta$. This is the idea discussed above. $\vec{w}_{1,2,3}$ rotates with the body so $\vec{\Omega}$ goes into $(\frac{d\vec{L}}{dt})_\beta$ and at very instant $\vec{w}_{1,2,3}$ is fixed in the space, so same $\dot{\vec{\Omega}}$ goes into $(\frac{d\vec{L}}{dt})_\alpha$. Later we will do completely non-initial frame and we will have different velocity goes into $(\frac{d\vec{L}}{dt})_\alpha$ cf (7.2).

Therefore in $\vec{w}_{1,2,3}$ frame, we get Euler equations

$$\begin{cases} I_1\dot{\Omega}_1 = N_1 - \Omega_2\Omega_3(I_3 - I_2) & 1 \\ I_2\dot{\Omega}_2 = N_2 - \Omega_1\Omega_3(I_1 - I_3) & 2 \\ I_3\dot{\Omega}_3 = N_3 - \Omega_2\Omega_2(I_2 - I_1) & 3 \end{cases} \tag{6.6}$$

We can now solve some simple system, and convert back to lab frame.

Example. Free symmetric top $\vec{N} = \vec{F} = 0$, $I_1 = I_2 \neq I_3$, $\vec{P}$ is constant, take $\vec{R} = 0$, and $\vec{L}$ is constant but this doesn't imply $\vec{\Omega}$ is constant. We will show $\vec{\Omega}$ is constant unless we start out the motion with $\Omega_3 \parallel z$ and $\Omega_1 = \Omega_2 = 0$. Otherwise we will see precision.

(6.6)-3 says

$$\Omega_3 = \text{constant}$$

Let

$$\omega = \frac{\Omega_3(I_3 - I_1)}{I_1}$$

then

$$\begin{cases} \dot{\Omega}_1 = -\Omega_2\omega \\ \dot{\Omega}_2 = \Omega_1\omega \end{cases} \implies \begin{cases} \Omega_1 = A\cos(\omega t+\psi) \\ \Omega_2 = A\sin(\omega t+\psi) \end{cases}$$

Now we want back to lab frame, we see $\Omega_{1,2}$ are not oscillating with infrequence ω but $\vec{\Omega}_\perp = \vec{\Omega}_1 + \vec{\Omega}_2$ rotates with frequency ω about$\vec{w}_3$, that is the principal axes $\vec{w}_{1,2}$ rotate about $\vec{w}_3$ with frequency ω, and $\vec{\Omega}$, or $\vec{\Omega}_\perp$ is in the plane of $\vec{L}$, $\vec{w}_3$.

In $\vec{w}_{1,2,3}$ frame

$$\begin{cases} L_1 = I_1\Omega_1 = I_1 A\cos(\omega t+\psi) \\ L_2 = I_1\Omega_2 = I_1 A\sin(\omega t+\psi) \\ L_3 = I_3\Omega_3 = \text{constant} \end{cases} \tag{6.7}$$

$\vec{L}_\perp = \vec{L}_1 + \vec{L}_2$ rotates in $\vec{w}_{1,2}$ plane at angular speed ω with constant magnitude

$$|\vec{L}_\perp|^2 = I_1^2|\vec{\Omega}_\perp|^2 = I_1^2A^2$$

so $\vec{L}$ and $\vec{\Omega}_\perp$ rotate about $\vec{w}_3$ at constant angular speed ω. But in the lab frame $\vec{L}$ is constant, so it is that $\vec{w}_3$ rotates about $\vec{L}$ at speed, but at what speed? ω? No

Use (6.5),

$$\begin{aligned} (\frac{d\vec{w}_3}{dt})_{lab} &= (\frac{d\vec{w}_3}{dt})_\alpha + (\frac{d\vec{w}_3}{dt})_\beta \\ &= 0 + \vec{\Omega}\times\vec{w}_3 \\ &= \vec{\Omega}_\perp \times \vec{w}_3 \\ &= \frac{\vec{L}_\perp}{I_1}\times\vec{w}_3 \\ &= \frac{\vec{L}}{I_1}\times\vec{w}_3 \end{aligned} \tag{6.8}$$

Hence $\vec{w}_3$ rotates (precession) about $\vec{L}$ with angular velocity

$$\vec{\omega}_{pr} = \frac{\vec{L}}{I_1} \tag{6.9}$$

The reason the angular velocities are not the same, because each one is determined from setting one vector ($\vec{L}$ or $\vec{w}_3$) stand still and do the perpendicular projection of the other vector onto itself, then see the angle swapped. For example ω is determined by the change of $\vec{L}$ in view of $\vec{w}$,

$$\frac{d\vec{L}}{dt} = (\hat{n} \times \vec{L})\frac{d\phi}{dt}$$

where $\hat{n} \parallel \vec{w}_3$ and then we find $\frac{d\phi}{dt} = w$, because in $\vec{w}_{1,2,3}$ frame by (6.7)

$$\frac{d\vec{L}}{dt} = w(-L_2 + L_1)$$

$$\hat{n} \times \vec{L} = L_2 - L_1$$

the minus is due to active v.s. passive views of coordinate transformations.

Similarly find ω_{pr}, put

$$\frac{d\vec{w}}{dt} = (\hat{n} \times \vec{w})\frac{d\phi}{dt}$$

where $\hat{n} \parallel \vec{L}$ and by (6.8), we confirm (6.9).

So the ratio of the two

$$\frac{\omega}{\omega_{pr}} = \frac{L_3/I_3}{L/I_1} = \frac{L\cos\theta/I_3}{L/I_1} = \frac{I_1\cos\theta}{I_3}$$

where θ is the angle between $\vec{L}$ and $\vec{w}_3$.

For asymmetric top there will be oscillation (mutation) of $\vec{w}_3$ around the orbit around $\vec{L}$.

6.7 Lagrangian Formation of Rigid Body

$$\begin{aligned}\mathcal{L} &= T - V \\ &= \frac{1}{2}M\dot{\vec{R}}^2 + \frac{1}{2}I_{ij}\Omega_i\Omega_j - V(\vec{R};\phi,\theta,\psi)\end{aligned}$$

where $\vec{\Omega} = \vec{\Omega}(\phi,\theta,\psi,\dot{\phi},\dot{\theta},\dot{\psi})$, ϕ,θ,ψ Euler angles.

As one can find in L.L., Lagrangian formation is way more complicate than

Newton's way via Euler equations. So we won't do much about it, unless the motion is only 2D.

Example. Physical pendulum, 3 Euler angle get 1 θ.

Suppose a 2D object is nailed at point O which is a distance from CM, let θ be the angle between the line connecting O and CM and the vertical. Then

$$V = -Mga\cos\theta$$

$$\begin{aligned} T &= \frac{1}{2}M\dot{\vec{R}}^2 + \frac{1}{2}I_{ij}\Omega_i\Omega_j \\ &= \frac{1}{2}a^2M\dot{\theta}^2 + \frac{1}{2}I_{zz}\dot{\theta}^2 \end{aligned}$$

Hence

$$\mathcal{L} = \frac{1}{2}\left(I_{zz} + a^2M\right)\dot{\theta}^2 - Mga\cos\theta$$

$I_{zz} + aM^2$ agrees parallel axis theorem. Compare this to simple pendulum

$$\mathcal{L} = \frac{1}{2}ml^2\dot{\theta}^2 - Mgl\cos\theta$$

7 Miscellaneous Topics

7.1 Adiabatic in Thermodynamics

Lecture 27 -Last Lec- (5/2/13)

Consider a single particle thermodynamics, trapped in ∞ square well $[a(t), b(t)]$, elastic bounces with the wall

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2 \text{ and } H = \frac{p^2}{2m}$$

also assume the initial momentum of the particle is P_0, then in the phase plane, the initial condition gives a path of square because momentum is constant within the interval and become negative on its way back. The square has phase volume

$$I = 2p_0(b-a)$$

Suppose $a(t)$, $b(t)$ slow vary

Under what condition is $E =$ constant? Clearly it is constant if $b(t) - a(t) =$ constant. So the work done by the wall to the particle is 0.

This tells us E is a function of $(a-b)(t)$, from Adiabatic invariance we know if a, b slow vary, I is constant.

We now study such particle as a thermodynamical object. We need to find the meaning of pressure, volume, temperature, etc.

Is there a thermodynamical interpretation of

$$PV = nk_BT \quad n = 1 \tag{7.1}$$

for one such particle?

Here

$$V = (b-a)$$

Here pressure acts on the wall=Force/area of end point of interval, so we guess it may= force

$$P = \frac{1}{\tau}\int_0^\tau \frac{dP}{dt}dt$$

τ is time over many cycles,

$$P \approx \frac{2p_0}{T}$$

T round trip period

$$T = \frac{2(b-a)}{p_0/m}$$

so we put

$$P = \frac{p_0^2}{m(b-a)}$$

Then

$$PV = \frac{p_0^2}{m} = 2E = 2\frac{1}{2}k_BT$$

hence the ideal gas law (7.1) reappears.

We know adiabatic transfer in thermodynamic conserve entropy.

Is our I entropy?

Start from thermodynamics identify

$$dE + PdV = Tds$$

and write thermodynamics variables in mechanical variables

$$dE = \frac{1}{2m}d(p_0^2) = \frac{p_0}{m}dp_0 \quad T = \frac{p_0^2}{mk_B}$$

so we get

$$\begin{aligned} ds &= \frac{dE}{T} + P\frac{dV}{T} \\ &= \frac{dp_0}{p_0^2}p_0 k_B + \frac{p_0^2}{mV}\frac{dV}{p_0^2}mk_B \\ &= k_B\left(\frac{dp_0}{p_0} + \frac{dV}{V}\right) \end{aligned}$$

From $I = 2p_0V$, we get

$$dI = 2Vdp_0 + 2p_0dV \implies \frac{dI}{I} = \frac{dp_0}{p_0} + \frac{dV}{V}$$

Hence

$$ds = k_B\frac{dI}{I}$$

or

$$S = k_B \ln I + \text{const}$$

we find I is indeed related to entropy S.

For $n > 1$

$$S = nk_B \ln \Omega + \text{const}$$

and we need to QM and quantize the phase space, dividing $1/2\pi\hbar$. But we can only do $n = 1$, because in 1D only $n = 1$ happens to have periodic motion, which is required by Adiabatic invariance in mechanics.

7.2 Non-Inertial Frame

Lastly we study Lagrangian for a point particle in 3D in a non-inertial reference frame. Reference L.L section 39.

Consider a total generic frame O' such as accelerating, rotating etc

In generalized coordinate

$$\vec{r'} = (x', y', z')$$

and let $\vec{r}_{O'}$ be the position of O' wrt to lab frame, then

$$\vec{r} = \vec{r}_{O'} + \vec{r'}$$

then

$$\frac{d\vec{r}}{dt} = \frac{d\vec{r}_{O'}}{dt} + \frac{d\vec{r'}}{dt}$$

and the $\frac{d\vec{r'}}{dt}$ is the familiar term that its position wrt to O' frame and t derivative is still in lab frame, cf (6.5), so

$$\frac{d\vec{r'}}{dt} = \vec{v'} + \vec{w} \times \vec{r'} \tag{7.2}$$

where now $\vec{v'}$ is the velocity wrt to O' frame and $\vec{w}$ is the angular velocity of O' frame wrt the lab.

We know EL equation works for all coordinates, but we don't have much faith in

$$\mathcal{L} = T - V$$

working for all coordinates. So we'd better plug $\vec{r}$ into above, then find the form of Lagrangian in O' frame

$$\begin{aligned} \mathcal{L} &= \frac{1}{2}m(\frac{d\vec{r}_{O'}}{dt} + \frac{d\vec{r'}}{dt})^2 - V(\vec{r}_{O'} + \vec{r'}) \\ &= \frac{1}{2}m\left(V_{O'}^2 + (\frac{d\vec{r'}}{dt})^2 + 2\vec{V}_{O'} \cdot \frac{d\vec{r'}}{dt}\right) - V(\vec{r}_{O'} + \vec{r'}) \end{aligned} \tag{7.3}$$

The $\frac{1}{2}mV_{O'}^2$ term doesn't depend on the particle, so no implication in EOM, we drop it. Therefore by (7.2) we get

$$\mathcal{L} = \frac{1}{2}m[v'^2 + (\vec{w} \times \vec{r}')^2 + 2\underbrace{\vec{v}' \cdot (\vec{w} \times \vec{r}')}_{\vec{r}' \cdot (\vec{v}' \times \vec{w})}]^2 + m\vec{V}_{O'} \cdot \vec{v}' - V(\vec{r}_{O'} + \vec{r}')$$

Later we will show how to get a cleaner form of $\mathcal{L}$.

Now apply EL

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{d\vec{v}'}\right) - \frac{\partial \mathcal{L}}{\partial \vec{r}'} = 0$$

We obtain

$$m\frac{d}{dt}\left(\vec{v}' + \vec{w} \times \vec{r}' + \vec{V}_{O'}\right) + \frac{\partial V}{\partial \vec{r}'} - \frac{1}{2}m\frac{\partial (\vec{w} \times \vec{r}')^2}{\partial \vec{r}'} - m(\vec{v}' \times \vec{w}) = 0$$

First term

$$\frac{d}{dt}\left(\vec{v}' + \vec{w} \times \vec{r}' + \vec{V}_{O'}\right) = \ddot{\vec{r}}' + \dot{\vec{w}} \times \vec{v}' - \vec{v}' \times \vec{w} + \vec{a}_{O'}$$

Next

$$\frac{\partial (\vec{w} \times \vec{r}')^2}{\partial \vec{r}'} = \frac{\partial}{\partial \vec{r}'}\left(w^2 r'^2 - (\vec{w} \cdot \vec{r}')^2\right) = 2\left(w^2 \vec{r}' - (\vec{w} \cdot \vec{r}')\vec{w}\right) = 2\vec{w} \times (\vec{r}' \times \vec{w})$$

Collecting everything, write them in decreasing familiarity

$$m\ddot{\vec{r}}' = \underbrace{-\frac{\partial V}{\partial \vec{r}'}}_{I} + \underbrace{-m\vec{a}_{O'}}_{II} + \underbrace{m\vec{w} \times (\vec{r}' \times \vec{w})}_{III} + \underbrace{2m(\vec{v}' \times \vec{w})}_{IV} - \underbrace{m\dot{\vec{w}} \times \vec{v}'}_{V}]$$

where I: usual force as if in the inertia; II: translation fictitious force; III: centrifugal force; IV: Coriolis; V: no name (usually small because angular momentum usually conserved e.g. O' is the earth).

Consequence of Coriolis: one side of river is washed out more than the other.

When we did $\vec{V}_{O'} \cdot \frac{d\vec{r}'}{dt}$ in (7.3), we could have recognized that

$$\vec{V}_{O'} \cdot \frac{d\vec{r}'}{dt} = \frac{d}{dt}(\vec{V}_{O'} \cdot \vec{r}') - \vec{a}_0 \cdot \vec{r}'$$

and the first term $\frac{d}{dt}(\vec{V}_{O'} \cdot \vec{r}')$ on the right didn't eventually contribute to EOM, so we could drop it a way ago. Therefore we get

$$\mathcal{L} = \frac{1}{2}m(\frac{d\vec{r'}}{dt})^2 - m\vec{a}_{O'} \cdot \vec{r'} - V(\vec{r}_{O'} + \vec{r'})$$

This looks like our old Lagrangian if we put

$$\tilde{V}(\vec{r'}) = m\vec{a}_{O'} \cdot \vec{r'} + V(\vec{r}_{O'} + \vec{r'})$$

Simplify a little, we get the perfect form $\mathcal{L}$ in O' frame

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}m\left(v'^2 + (\vec{w} \times \vec{r'})^2 + 2\vec{v'} \cdot (\vec{w} \times \vec{r'})\right)^2 - \tilde{V}(\vec{r'}) \\ \mathcal{L} &= T' - V'\end{aligned}$$

where

$$T'(\dot{\vec{r'}}) = \frac{1}{2}mv'^2$$

$$V'(\vec{r'}) = \tilde{V}(\vec{r'}) - \frac{1}{2}m(\vec{w} \times \vec{r'})^2 - m\vec{r'} \cdot (\vec{v'} \times \vec{w})$$

One can figure out the corresponding meaning of each term in $\mathcal{L}$ to the inertial force by looking up the derivation above.