Advanced Mechanics I. Phys 302

Artem G. Abanov

This work is licensed under the Creative Commons Attribution 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by/3.0/ or send a letter to Creative Commons, 444 Castro Street, Suite 900, Mountain View, California, 94041, USA.

Contents

LECTURE 1
Introduction.

Preliminaries.

- Contact info.
- Syllabus. Homework 40%, First exam 30%, Final 30%.
- Grading scheme: 90-100 A; 75-89.99 B; 60-74.99 C; 50-59.99 D; 0-49.99 F. Grades may be curved a bit.
- Attendance policy.
- Lecture, feedback. Going too fast, etc.
- Office hours (Mondays 10:30-11:30 am on Zoom or in-person)
- Canvas.
- Homework submissions through Canvas. PDF SINGLE FILE.
- Homeworks are due by the start of Wednesday's lectures. (Except the first one.) No late submissions!
- Homework session. **WHEN???**.
- Homeworks: To cheat or not to cheat? collaborations!!!!! make study groups, mistakes, procrastination, etc. Help from the internet.
- Honors problems. Indicate if you are an Honors student on top of your homework.
- Grading. Every assignment is 100 pt. The points split equally between the problems in a given assignment.
- Exams.
  - The same point system as in homework assignments.
  - First exam is take home. Most of the problems are taken from the problem bank.
  - The bank is on the web `http://people.tamu.edu/~abanov/QE/`.
- Book. Lecture notes. Lectures and lecture notes. Problems, problems, and problems!!!!
- Language.
- Course content and philosophy.
- Questions: profound vs. stupid.
- Lecture is a conversation.

LECTURE 2
Coordinates. Scalars. Vectors. Einstein notations.

2.1. Coordinates.

- Coordinate systems. You chose a coordinate system to describe a process (positions, motion, fields, etc)
- The physical process does not depend on the system of coordinates you use to describe it!
- This obvious statement leads to the requirement, that all physics laws were formulated in a system-of-coordinate independent way.
- We need to define mathematical objects which allow for manipulation in a coordinate independent way. Such objects are: scalars, vectors, tensors, etc.

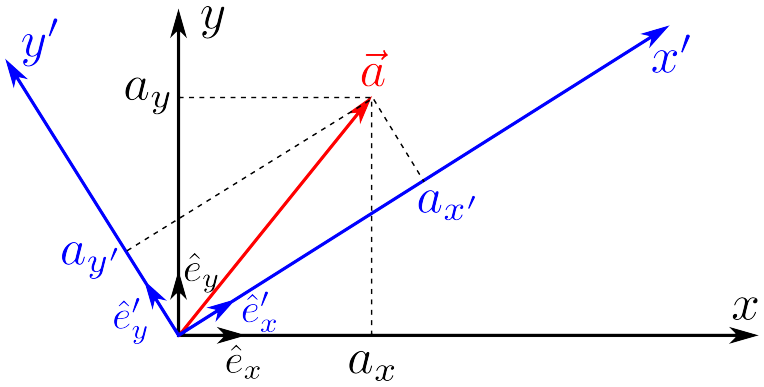
2.2. Scalars

- Scalars do not depend on the coordinate system.
- Representation of the scalar does not depend on the coordinate system.
- Examples are:
  - Energy, charge, mass, etc.
  - Example of not a scalar: a component of a vector.
- You operate with scalars as with normal numbers.

2.3. Vectors

Vectors have DIRECTION. The magnitude and the direction of a vector are EQUALLY important!

- Vectors do not depend on the coordinate system.
- Representation of a vector through components does depend on the coordinate systems!
- Examples are:
  - Forces, velocity, acceleration, etc.
- If you or I say that something is a vector, it is NOT an empty statement.
  - This statement demands, requires, and allows certain operations with them. The same way as if something is a number, then we know that we can add, subtract, multiply, and divide (unless the number is zero).



  - With vectors the operations are different than with numbers/scalars! For example, one cannot divide by a vector! Such operation does not exist. The expression $\frac{1}{\vec{a}}$ makes no sense and MUST never appear.
- Vector components: Given a coordinate system one can express/represent a vector by its components in this particular coordinate system.
- In order to operate with vector components you MUST establish the coordinate system first!
- If one changes the coordinate system the vector DOES NOT change, but its components in the new coordinate system will be different from the components in the old one.
- All operations with vector can be formulated in the coordinate independent way. They also can be formulated through components in a GIVEN coordinate system.
- What can be done with vectors? What kind of operations can we define that will not depend on the coordinate system?
- Linearity
  - One can multiply a vector by a number.
  - One can add two vectors.
  - Using these properties, in a coordinate system given by three orthonormal vectors $\hat{e}_x$, $\hat{e}_y$, $\hat{e}_z$, we can represent a vector $\vec{a}$ through its components in THE SAME coordinate system

(2.1) $$\vec{a} = a_x\hat{e}_x + a_y\hat{e}_y + a_z\hat{e}_z.$$

    Again, vector $\vec{a}$ is independent of the coordinate system. The unit orthogonal vectors $\hat{e}_x$, $\hat{e}_y$, $\hat{e}_z$ define/describe the coordinate system. In that coordinate system the vector $\vec{a}$ has components a_x, a_y, and a_z. The vector $\vec{a}$ can be represented in the coordinate system given by vectors $\hat{e}_x$, $\hat{e}_y$, $\hat{e}_z$ by the above expression.
  - If we change the coordinate system $\hat{e}_x, \hat{e}_y, \hat{e}_z \to \hat{e}'_x, \hat{e}'_y, \hat{e}'_z$, the coordinates of a vector $\vec{a}$ will also change $a_x, a_y, a_z \to a_{x'}, a_{y'}, a_{z'}$. The REPRESENTATION of the vector $\vec{a}$ will change

$$\vec{a} = a_{x'}\hat{e}'_x + a_{y'}\hat{e}'_y + a_{z'}\hat{e}'_z.$$

    but the vector $\vec{a}$ is the same!
- Scalar (dot) product.
  - Coordinate independent definition.

$$\vec{a}\cdot\vec{b} = |\vec{a}||\vec{b}|\cos(\phi)$$

– Definition through components in a given coordinate system:

$$\vec{a}\cdot\vec{b}=\sum_{i=1}^{3}a^ib^i\equiv a^ib^i \text{ — Einstein notations.}$$

I denoted the vector's components by a^1, a^2, and a^3, instead of a_x, a_y, and a_z. You should check, that the two definitions are the same. It means that for ANY two vectors $\vec{a}$ and $\vec{b}$ and any coordinate system $|\vec{a}||\vec{b}|\cos(\phi)=a^ib^i$.

– Bilinear.

– Symmetric.

– The magnitude of a vector $\vec{a}$ is given by

$$|\vec{a}|=\sqrt{\vec{a}\cdot\vec{a}}.$$

– Using these properties, in a coordinate system given by three orthonormal vectors $\hat{e}_x$, $\hat{e}_y$, $\hat{e}_z$, we can find the components of a vector $\vec{a}$ by

$$a_x=\hat{e}_x\cdot\vec{a},\qquad a_y=\hat{e}_y\cdot\vec{a},\qquad a_z=\hat{e}_z\cdot\vec{a}.$$

All you need to do is to take the vector representation (2.1) and take the dot product of this expression with $\hat{e}_x$, $\hat{e}_y$, and $\hat{e}_z$

– In particular any vector can be written as

$$\vec{a}=\hat{e}_x(\hat{e}_x\cdot\vec{a})+\hat{e}_y(\hat{e}_y\cdot\vec{a})+\hat{e}_z(\hat{e}_z\cdot\vec{a})$$

- Vector (cross) product.
  – Coordinate independent definition.

$$\vec{c}=\vec{a}\times\vec{b},\qquad |\vec{c}|=|\vec{a}||\vec{b}|\sin(\phi),\qquad \text{Direction — right hand rule.}$$

– In a coordinate system given by three orthonormal vectors $\hat{e}_x$, $\hat{e}_y$, $\hat{e}_z$, we have

$$\vec{a}\times\vec{b}=\begin{vmatrix}\hat{e}_x & \hat{e}_y & \hat{e}_z\\ a_x & a_y & a_z\\ b_x & b_y & b_z\end{vmatrix}.$$

– Bilinear.

– Antisymmetric – RHR, this is why it is $\sin(\phi)$ and not $\cos(\phi)$. Determinant.

2.4. Kronecker symbol. Symbol Levi-Chivita. Einstein notations.

- Kronecker symbol: $\delta_{i,j}$. (In Euclidean space there is no need to distinguish between upper and lower indexes. In non-Euclidean space they are different!)
  – Definition

$$\delta_{i,j}=\begin{cases}1, & \text{if } i=j\\ 0, & \text{if } i\neq j\end{cases}$$

– Einstein notations

$$a_i\delta_{i,j}=a_j,\qquad \vec{a}\cdot\vec{b}=a^ib^i=a^i\delta_{i,j}b^j$$

- Symbol Levi-Chivita.

- Definition

$$\epsilon^{i,j,k} = 0, \quad \text{if any of the indexes equal each other.}$$
$$\epsilon^{1,2,3} = \epsilon^{2,3,1} = \epsilon^{3,1,2} = 1$$
$$\epsilon^{1,3,2} = \epsilon^{2,1,3} = \epsilon^{3,2,1} = -1$$

- Useful formulas:

$$\epsilon^{ijk}\epsilon^{ijl} = 2\delta^{kl}, \qquad \epsilon^{ijk}\epsilon^{ilm} = \delta^{jl}\delta^{km} - \delta^{jm}\delta^{kl}.$$

Notice the use of Einstein notations: In the first formula we sum over all values of two indexes i and j; In the second we sum over all the values of only one index i.

- Examples:
  - Vector product $\vec{c} = \vec{a} \times \vec{b}$:

$$c^i = [\vec{a} \times \vec{b}]^i = \epsilon^{ijk} a^j b^k$$
$$c^x = [\vec{a} \times \vec{b}]^x = \epsilon^{xyz} a^y b^z + \epsilon^{xzy} a^z b^y = a^y b^z - a^z b^y, \quad \text{etc.}$$

Importance of the order of indexes.

  - Scalar product of two vector products:

$$[\vec{a} \times \vec{b}] \cdot [\vec{c} \times \vec{d}] = [\vec{a} \times \vec{b}]^i [\vec{c} \times \vec{d}]^i = \epsilon^{ijk}\epsilon^{ilm} a_j b_k c_l d_m = \left(\delta^{jl}\delta^{km} - \delta^{jm}\delta^{kl}\right) a_j b_k c_l d_m =$$
$$a_j c_j b_k d_k - a_j d_j b_k c_k = (\vec{a} \cdot \vec{c})(\vec{b} \cdot \vec{d}) - (\vec{a} \cdot \vec{d})(\vec{b} \cdot \vec{c})$$

  - Triple vector product:

$$[\vec{a} \times [\vec{b} \times \vec{c}]]^i = \epsilon^{ijk}\epsilon^{klm} a_j b_l c_m = \epsilon^{kij}\epsilon^{klm} a_j b_l c_m =$$
$$\left(\delta^{il}\delta^{jm} - \delta^{im}\delta^{jl}\right) a_j b_l c_m = b_i a_j c_j - c_i b_j a_j = \left[\vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b})\right]^i$$

so

$$[\vec{a} \times [\vec{b} \times \vec{c}]] = \vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b})$$

- Bilinearity.
- Differentiation of scalar and vector products.
  - Example: Consider a unit vector $\vec{n}(t)$ which depends on time t (or any other parameter). As $\vec{n}$ is a unit vector we have $\vec{n} \cdot \vec{n} = 1$. Differentiating with respect to time gives $\dot{\vec{n}} \cdot \vec{n} = 0$ — the derivative is orthogonal to the vector $\vec{n}$ at all times.
  - Notations:

$$\dot{f} \equiv \frac{df}{dt}.$$

- Differentiation of $|\vec{r}|$. We start with $|\vec{r}| = \sqrt{\vec{r} \cdot \vec{r}}$, then using the chain rule

$$\frac{d}{dt}|\vec{r}| = \frac{d}{dt}\sqrt{\vec{r} \cdot \vec{r}} = \frac{d}{dt}\sqrt{r_i r_i} = \frac{dr_j}{dt}\frac{\partial\sqrt{r_i r_i}}{\partial r_j} = \frac{dr_j}{dt}\frac{\partial r_i}{\partial r_j}\frac{r_i}{|\vec{r}|} = \frac{dr_j}{dt}\delta_{i,j}\frac{r_i}{|\vec{r}|} = \frac{\vec{r} \cdot \dot{\vec{r}}}{|\vec{r}|}.$$

LECTURE 3
Newton's laws.

- Notations:

$$\dot{f} \equiv \frac{df}{dt}.$$

3.1. Frames of reference.

Notations:

If $\vec{r}$ is a position vector, then

$\dot{\vec{r}} \equiv \vec{v}$ — velocity, the rate of change of the position,

$\dot{\vec{v}} = \ddot{\vec{r}} \equiv \vec{a}$ — acceleration, the rate of change of the velocity.

All three: the position $\vec{r}$, the velocity $\vec{v}$, and the acceleration $\vec{a}$ are vectors!!! ALL three vectors generally depend on time!

Last lecture we discussed the concept that the physical processes do not depend on the system of coordinate one uses to describe them. Another observation is that the physical processes do not depend on observer.

If we have two observers observing the SAM process, their description of the process will be different. So in order to formulate the laws of physics we must find the formulation which does not depend on observer!

Moving frame of reference:

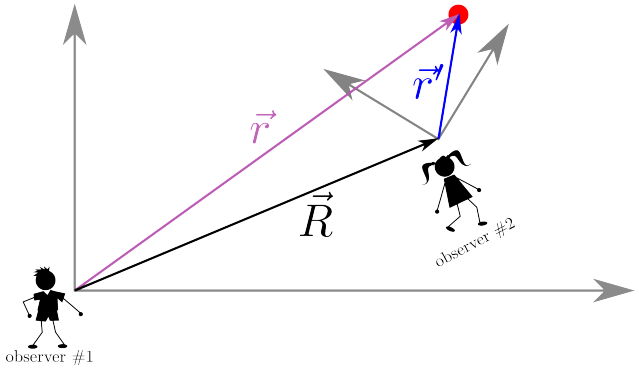
- Let's consider a process/object (red dot on the figure). This object observed by two observers: observer #1 and observer #2. The object is moving with some velocity, but at this particular moment of time t it is observed by the observer #1 to be at the position $\vec{r}$. Observer #2 observes the object at the position $\vec{r}'$ at the SAME moment of time t.
- The observer #1 also observes the observer #2. At the SAME moment of time the observer #1 sees the observer #2 at the position $\vec{R}$.
- From the figure we see

$$\vec{r} = \vec{R} + \vec{r}'$$

(3.1)

- Differentiating this relation with respect to time we get

$$\dot{\vec{r}} = \dot{\vec{R}} + \dot{\vec{r}}', \qquad \vec{v} = \vec{V} + \vec{v}'$$



- A not on the assumptions. Different meaning of dt and $d\vec{r}$. It is not guaranteed, that dt is the same in all frames of reference, but it is in classical (non-relativistic) mechanics.
- So the velocity of the object as measured by the observer #1 equals the velocity of the SAME object measured at the SAME time by the observer #2 plus the velocity of the observer #2 measured by the observer #1 at the SAME time.
- Velocity of the same object is different in the different frames of references! So the fundamental laws cannot be formulated in terms of the velocities!
- If $\vec{V}$ is constant, then $\dot{\vec{v}} = \dot{\vec{v}}'$.
- Galileo: The laws of physics must be the same in all inertial frames of reference.
- The laws then must be formulated in terms of acceleration.
- Initial conditions: initial position and initial velocity – we need to set up the motion.
- First Newton's law. If there is no force a body will move with constant velocity.
  - What is force? Interaction. Is there a way to exclude the interaction?
  - The existence of a special class of frames of reference – the inertial frames of reference.
- Force, as a vector measure of interaction.
- Point particle and mass.
- The requirement that the laws of physics be the same in all inertial frames of references. The second Newton's law: $\vec{F} = m\vec{a}$.

3.2. Second Newton's law.

The second Newton's law:

In order to apply it you must:

- Identify the object!!!!! Object is whatever! you! want! it to be! The only thing is that it must have a finite mass.
- Identify ALL the forces acting on THAT object. Remember, the forces are vectors. Remember forces may and in most cases will depend on the object's position, velocity etc. They may also depend on time. They may also depend on the positions and velocities of other objects that interact with a given one.
- Compute the net force. Superposition. The net force is simply the vector sum of all forces acting on the object. As a result you will have the net force $\vec{F}$ which depends on time as well as on the object's position, velocity, etc: $\vec{F}(t, \vec{r}, \dot{\vec{r}}, \dots)$.

- Write the equation of motion

$$m\ddot{\vec{r}} = \vec{F}(t, \vec{r}, \dot{\vec{r}}, \dots).$$

This is a system of three second order non-linear coupled differential equations!!!!

- In order to find a solution $\vec{r}(t)$ this system must be supplied with six initial conditions! (two conditions (second order=two conditions) for each of the three equations.)

3.3. Third Newton's law.

- A force is the result of INTERACTION!
- Interaction involves TWO objects.
- If object 1 interacts with the object 2, then this interaction results in TWO forces: one $\vec{F}_{\text{on 1 from 2}}$ acts on the object 1 and the other $\vec{F}_{\text{on 2 from 1}}$ acts on the object 2.
- The third Newton's law states that:

$$\vec{F}_{\text{on 1 from 2}} = -\vec{F}_{\text{on 2 from 1}}.$$

3.4. Examples.

In the following I give very simple examples of the use of the Newton's Laws.

$\vec{F} = m\vec{a}$ works both ways.

- Given the motion we can find the total force.
  - An object of mass m is sitting on a table. The coefficient of friction is μ. We apply a force F to the object parallel to the table. The object is not moving. What is the friction force?
  - An object of mass m is going around a circle of radius R with constant angular velocity ω. The position of the object at time t is given by $\vec{r}(t) = \hat{e}_x R\cos(\omega t) + \hat{e}_y R\sin(\omega t)$ — this is our motion. We find the acceleration $\vec{a} = \ddot{\vec{r}} = -\omega^2\vec{r}$. There must be a force acting on the particle: $\vec{F} = m\vec{a} = -m\omega^2\vec{r}$.
  - Archimedes law.
- Given the force we can find the motion.
  - Vertical motion. $\vec{F} = -mg\hat{e}_y$. So if $\vec{r} = x\hat{e}_x + y\hat{e}_y$, then $m\ddot{\vec{r}} = m\ddot{x}\hat{e}_x + m\ddot{y}e_y = -mg\hat{e}_y$. Or in components

    $$\ddot{x} = 0, \qquad \ddot{y} = -g$$

    Two second order differential equations: four initial conditions

    $$x(t=0) = x_0, \quad \dot{x}(t=0) = v_{x0}, \qquad y(t=0) = y_0, \quad \dot{y}(t=0) = v_{y0}$$

    The solution

    $$x(t) = x_0 + v_{x0}t, \qquad y(t) = y_0 + v_{y0}t - \frac{gt^2}{2}.$$

  - Wedge. Using $\vec{F} = m\vec{a}$ both ways.
  - Wedge with friction.
  - Pulley.

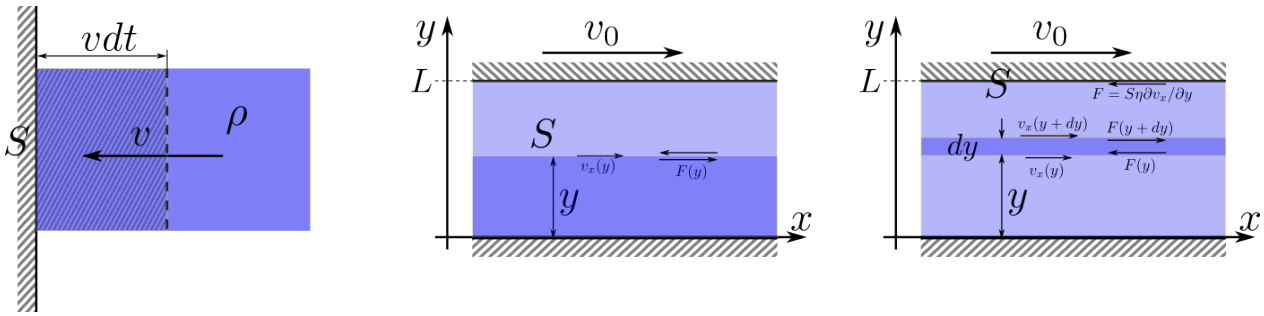
LECTURE 4
Air resistance.

4.1. Another formulation.

- The formulation
$$\vec{F} = m\vec{a}$$
is good for point-like particles. However, for the extended objects (an object does not have to consist of one piece) it can be challenging to apply.
- Many laws of physics are formulated for point-like objects. For example the Newton's gravity $\vec{F} = -\frac{Gm_1m_2}{r^3}\vec{r}$ can only be applied for point-like masses m_1 and m_2, as otherwise, for an extended object, it is impossible to define what $\vec{r}$ is. The same holds for Coulomb's law: $\vec{F} = \frac{kq_1q_2}{r^3}\vec{r}$ which is also valid only for point-like charges q_1 and q_2, for the same reason.
- The second Newton's law formulated as $\vec{F} = m\vec{a}$ is not very convenient is the mass of the object is changing in time. However, starting from the the constant mass, we can reformulate it in the following way
$$\vec{F} = m\vec{a} = m\frac{d\vec{v}}{dt} = \frac{d(m\vec{v})}{dt} = \frac{d\vec{p}}{dt}.$$
where $\vec{p}$ is the vector of momentum. In short
$$\vec{F} = \dot{\vec{p}}.$$
- Another benefit of this formulation is that momentum is additive! the momentum of the object is just the sum (vector sum) of the momenta of all the pieces of the object. In this case in the above equation the force $\vec{F}$ must be understood as the total force acting on the object (the vector sum of all the forces acting on all the pieces of the object) and momentum $\vec{p}$ as the total momentum of the object.

4.2. Air resistance.

- Water hose. Water stream hits a wall perpendicular to the wall. After hitting the wall the water stops.
  - The force acting on the wall is obviously proportional to the cross-section area S of the water stream. So we want to find the force per unit area (pressure) of the stream on the wall.



  - At time t consider a piece of the stream between the wall and the stream's cross-section the distance vdt from the wall, where dt is small interval of time and v is the velocity of the water in the stream, see left panel of the figure.
  - The volume of this piece is $Svdt$, its mass is $dm = Svdt\rho$, where ρ is the mass density of the water.
  - The momentum of this piece at time t is $dp = vdm = Sv^2\rho dt$.
  - The momentum of this piece at time $t + dt$ is zero, as water stops after hitting the wall.
  - So the change of the momentum of the stream during time dt is $dp = Sv^2\rho dt$.
  - So the magnitude of the force acting on the stream from wall is

$$F = \frac{dp}{dt} = \frac{Sv^2\rho dt}{dt} = S\rho v^2.$$

  - The force acting on the wall must have the same magnitude. So the force per unit area acting on the wall is

$$f = \frac{F}{S} = \rho v^2$$

  - Force is proportional to the velocity squared.
- Force of viscous flow.
  - Setup: Two infinite parallel plates at distance L from each other. The top plate is moving with velocity v_0 in the direction parallel to the plates, which we will take as $\hat{x}$ direction. The bottom plate is at rest. There is a viscous liquid in between the plates. What force is acting on the plates?
  - The force per area of a viscous flow is proportional to the velocity difference, or derivative $F/S = f \sim -\partial v_x/\partial y$. This follows simply from the fact, that if the velocity of liquid is the same for all y, then there is no force. And if the area doubles, the force obviously also doubles. The coefficient of proportionality depends on the liquid, we will call it η, so

$$F(y) = -\eta S\frac{\partial v_x}{\partial y}.$$

    See the middle panel on the figure.
  - This is the force that acts from the lower part on the upper part. The force that acts from the upper part on the lower part is, by Newton's third law, the same, but opposite in direction.
  - It is not enough, as we do not know how v_x depends on y in our setup. We need to find it.

– In order to find $v_x(y)$ we must use the equations of motion. In order to use these equations we MUST first define what the object is for which we apply the equations of motion.
– Consider a slab of liquid of thickness dy, the total force which acts on a liquid of area S of this slab is $\eta S\left(-\frac{\partial v_x}{\partial y}\Big|_y + \frac{\partial v_x}{\partial y}\Big|_{y+dy}\right) = \eta S dy \frac{\partial^2 v_x}{\partial y^2}$. See right panel on the figure.
– The mass of the slab of the area S is $\rho S dy$. So if the slab has acceleration a, then we must have

$$\eta S dy \frac{\partial^2 v_x}{\partial y^2} = a\rho S dy, \qquad \eta S \frac{\partial^2 v_x}{\partial y^2} = a\rho S$$

– In the steady state the acceleration $a = 0$, so

$$\frac{\partial^2 v_x}{\partial y^2} = 0, \qquad v_x(y=0) = 0, \qquad v_x(y=L) = v_0.$$

This is a differential equation with boundary conditions. The boundary conditions tell us that the water close to walls has the same velocity as the walls. (This is not the only possibility. The boundary conditions depend on what the liquid/gas is and what material the walls are made of. The real physics is, in fact, in these boundary conditions.)
– The solution of this equation which satisfies the boundary conditions is

$$v_x(y) = v_0 \frac{y}{L}.$$

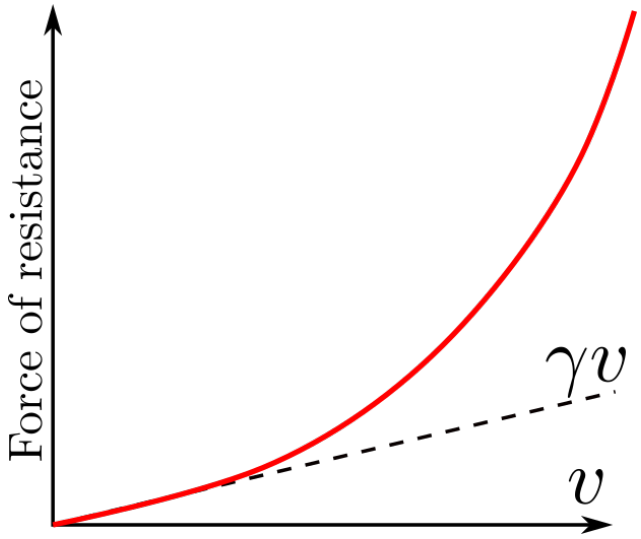
– The force per area then is proportional to

$$f \sim -\frac{\partial v_x}{\partial y} = -v_0/L.$$

– So the force is linear in velocity.

LECTURE 5
Air resistance.

5.1. Air resistance.



We consider two model cases: the air resistance is proportional to v, or to v^2 – linear, or quadratic. These forms of the air resistance should not be taken literary. These two cases are just models we will use to learn how the motion depends on the forms of the air resistance.

In reality one can think of the two regimes: for small velocity the resistance is mainly proportional to v, for large velocity it is mainly proportional to v^2. And for even larger velocities the air flow will become turbulent and the resistance force will become a lot more complicated (not even a function of velocity).

In fact one will have both cases. If we start with large velocity and resistance proportional to v^2 and the velocity is decreasing, then it will eventually become small enough to switch to the regime where the resistance is proportional to v.

What is the boundary, or more precisely what does it mean small or large velocity (small or large in comparison to what?) depends on the body, air, etc.

The main point of this lecture is to show that the motion is very different for these two cases. And the difference comes exactly from different powers of v, not from the prefactors.

In this lecture I will only consider $1D$ cases. The particle moves along a straight line.

5.2. No gravity.

5.2.1. Linear in v case

$$F = -\gamma_1 v$$

The direction of the force is opposite to the direction of velocity.

The parameter γ_1 is some number which depends on many factors: shape of the body, properties of its surface, air composition etc. This number is not dimensionless, the units of γ_1 can be computed: the units of force $[F] = kg \cdot m/s^2$, the units of velocity $[v] = m/s$, so the units of γ_1 are

$$[\gamma_1] = \frac{[F]}{[v]} = \frac{kg \cdot m/s^2}{m/s} = kg/s$$

The main result is that the linear in v resistance force is strong enough that the body will travel only a finite distance.

We denote the coordinate of the body by x. The velocity is $v = \frac{dx}{dt}$, the acceleration $a = \frac{dv}{dt} = \frac{d^2x}{dt^2}$. As the force in our case depends only on the velocity, the equation of motion $ma = F$ takes the form:

$$m\dot{v} = -\gamma_1 v, \qquad v(t=0) = v_0, \qquad x(t=0) = 0,$$

where I placed the origin of the coordinate at the initial point of the motion and the initial velocity of the motion is v_0.

The differential equation is the standard one, so the solution is

$$v(t) = v_0 e^{-\frac{\gamma_1}{m}t}.$$

(Check that the initial condition $v(t=0) = v_0$ is satisfied!!)

In order to find the coordinate of the particle as the function of time $x(t)$ we use the definition $v(t) = \frac{dx}{dt}$, rewrite it in the form $dx = v(t)dt$ and integrate $\int_0^{x(t)} dx' = \int_0^t v(t')dt'$

$$x(t) = \int_0^t v(t')dt' = \frac{mv_0}{\gamma_1}(1 - e^{-\frac{\gamma_1}{m}t}).$$

(Check, that the initial condition $x(t=0) = 0$ is automatically satisfied.)

5.2.1.1. *Analyze the result!!!!!!! Always!!!!!.*

- CHECK THE UNITS!!!

  The first thing to check is that units match. Anything which is under the exponent (or under sin or cos or log) MUST be unitless/dimensionless. We have $\gamma_1 t/m$ in the exponent. The units for this expression are $[\gamma_1 t/m] = \frac{kg}{s} s \frac{1}{kg} = 1$, so it is indeed unitless/dimensionless.

  The expression for the distance $x(t)$ has the units of $\frac{mv_0}{\gamma_1}$. Its units are $\left[\frac{mv_0}{\gamma_1}\right] = kg\frac{m}{s}\frac{s}{kg} = m$. So indeed we have the units of length.
- SEE WHAT THE ANSWER TELL US.
  - The velocity decaying exponentially the typical time of the exponential decay is $\tau = m/\gamma_1$ (check units!).
  - From the equation for $x(t)$ one sees, that the total distance traveled by the body is finite.

$$x(t \to \infty) = \frac{mv_0}{\gamma_1}.$$

Notice, that the total distance
- ∗ increases, if initial velocity is increases.
- ∗ decreases, if the resistance coefficient γ_1 increases.
- ∗ increases, if the mass increases.

All three statements make sense on the intuitive level!

- CHECK THE LIMITING CASES

  Next, we consider the limiting case which we know: if γ_1 is very small we should (almost) recover the known result, that the velocity stays constant (almost), and that the coordinate $x(t)$ is (almost) $v_0 t$. The word almost means that there are corrections which are getting smaller and smaller as γ_1 decreases.

  Taking $\gamma_1 t/m$ to be small, we use the Taylor expansion to the first non-zero order on this small parameter.

  So, if $\frac{\gamma_1}{m} t \ll 1$, then

$$v(t) \approx v_0 - v_0 \frac{\gamma_1 t}{m},$$
$$x(t) \approx v_0 t - \frac{1}{2} v_0 t \frac{\gamma_1 t}{m}.$$

  Notice an important lesson, that the expansion parameter is $\gamma_1 t/m$. No matter how small γ_1 is this parameter will become large at large enough time. At these times the Taylor expansion will no longer be valid. So the Taylor expansion is only valid for times $t \ll m/\gamma_1 = \tau$.

5.2.2. Quadratic in v case.

The set up is the same as in the previous case. The motion is along one line. The coordinate of the body is x. The motion starts from $x = 0$ with the initial velocity v_0 in the positive direction.

The resistance force is now quadratic in velocity v

$$F = -\gamma_2 |v| v.$$

The parameter γ_2 (this parameter is very different from γ_1 used in the previous section) is some number which depends on many factors: shape of the body, properties of its surface, air composition etc. The units of γ_2 are

$$[\gamma_2] = kg/m$$

(see, even units of this γ_2 are different from the previous one.)

The main result is that the body will travel infinite distance, no matter how small the initial velocity is.

$$m\dot{v} = -\gamma_2 v^2, \qquad v(t=0) = v_0, \qquad x(t=0) = 0,$$

The solution of this equation gives:

$$\frac{m}{v} = \gamma_2 t + \frac{m}{v_0}, \qquad v(t) = \frac{v_0}{1 + \frac{v_0 \gamma_2}{m} t}, \qquad x(t) = \int_0^t v(t')dt' = \frac{m}{\gamma_2} \log\left(1 + \frac{v_0 \gamma_2}{m} t\right).$$

5.2.2.1. *Analyze the result!!!!!!! Always!!!!!.*

- CHECK THE UNITS!!!
- SEE WHAT THE ANSWER TELL US.
  - At small times the velocity is almost constant v_0.
  - At larger times $t \gg \frac{m}{v_0\gamma_2}$ the velocity decaying as $v \sim \frac{m}{\gamma_2}\frac{1}{t}$ – inversely proportional to time with coefficient which is independent(!!!!) of the initial velocity.
  - Because of the $1/t$ decay, the total traveled distance is logarithmic in time. So the total distance traveled by the body is infinite — the distance diverges logarithmically with time.
  - This may not be intuitively obvious, but this is correct if the resistance force stays proportional to v^2 for all velocities.
  - However, the velocity is decreasing. Eventually it will decrease enough for the resistance force to become linear in v. Then the total traveled distance will be finite.
- CHECK THE LIMITING CASES

  Consider the limiting case of small γ_2. If $\frac{v_0\gamma_2}{m}t \ll 1$ (again, notice what plays the role of the expansion parameter!), then

$$v(t) \approx v_0 - v_0\frac{v_0\gamma_2}{m}t,$$
$$x(t) \approx v_0 t - v_0 t\frac{1}{2}\frac{v_0\gamma_2}{m}t.$$

5.3. Air resistance and gravity. Linear case.

The gravity points down. The x axis points up. We start the motion at $x = 0$ with the positive velocity v_0. The equation of motion is

$$m\dot{v} = -mg - \gamma_1 v, \qquad v(t=0) = v_0, \qquad x(t=0) = 0.$$

so rewriting the equation as

$$\frac{m}{\gamma_1}\frac{dv}{v + mg/\gamma_1} = -dt$$

and integrating this equation

$$\frac{m}{\gamma_1}\int \frac{dv}{v + mg/\gamma_1} = -\int dt$$

we find the general solution

$$\frac{m}{\gamma_1}\log(v + mg/\gamma_1) = -t + C.$$

Using the initial condition $v(t=0) = v_0$ we find from above

$$\frac{m}{\gamma_1}\log(v_0 + mg/\gamma_1) = C.$$

So

$$\frac{m}{\gamma_1}\log\frac{v + mg/\gamma_1}{v_0 + mg/\gamma_1} = -t.$$

and

$$v(t) = v_0 e^{-\frac{\gamma_1}{m}t} + \frac{mg}{\gamma_1}\left(e^{-\frac{\gamma_1}{m}t} - 1\right).$$

$$x(t) = \int_0^t v(t')dt' = v_0\frac{m}{\gamma_1}\left(1 - e^{-\frac{\gamma_1}{m}t}\right) - \frac{mg}{\gamma_1}\left(\frac{m}{\gamma_1}\left(e^{-\frac{\gamma_1}{m}t} - 1\right) + t\right)$$

5.3.0.1. *Analyze the result!!!!!!! Always!!!!!.*

- CHECK THE UNITS!!!
- SEE WHAT THE ANSWER TELL US, AND CHECK THE LIMITING CASES.
  - If $g = 0$ we recover the case of no gravity considered before.
  - Limit of $\gamma_1 t/m \ll 1$:

    $$v \approx v_0 - \left(g + \frac{v_0\gamma_1}{m}\right)t = v_0 - \tilde{g}t, \qquad \tilde{g} = g + \frac{v_0\gamma_1}{m}$$

    $$x(t) \approx v_0 t - \frac{\tilde{g}t^2}{2}$$

    Notice, that the motion is as if the acceleration of free fall is a bit different than g. It is $\tilde{g}$, and it does depend on mass!!! Although, one must be careful. This is correct only for small times $t \ll m/\gamma_1$. At these times the particle is still going up. If we set up the motion such, that the particle is thrown down the acceleration would be $\tilde{g} = g - \frac{v_0\gamma_1}{m}$
- What else can we learn from the solution?
  - Time to the top. At the top the velocity is zero, so we need to find such T that $v(t = T) = 0$,

    $$T = \frac{m}{\gamma_1}\log\left(1 + \frac{\gamma_1 v_0}{mg}\right),$$

    for $\frac{\gamma_1 v_0}{mg} \ll 1$

    $$T \approx \frac{m}{\gamma_1}\frac{\gamma_1 v_0}{mg} - \frac{1}{2}\frac{m}{\gamma_1}\left(\frac{\gamma_1 v_0}{mg}\right)^2 = \frac{v_0}{g} - \frac{1}{2}\frac{v_0}{g}\frac{\gamma_1 v_0}{mg}.$$

  - Height. Computing $x(T)$ we find

    $$x(T) = \frac{v_0^2}{g}\frac{mg}{\gamma_1 v_0}\left[1 - \frac{mg}{\gamma_1 v_0}\log\left(1 + \frac{\gamma_1 v_0}{mg}\right)\right]$$

    for $\frac{\gamma_1 v_0}{mg} \ll 1$

    $$x(T) \approx \frac{1}{2}\frac{v_0^2}{g} - \frac{1}{3}\frac{\gamma_1 v_0^3}{mg^2}.$$

    However, for $\frac{\gamma_1 v_0}{mg} \gg 1$ we can neglect $\frac{mg}{\gamma_1 v_0}\log\left(1 + \frac{\gamma_1 v_0}{mg}\right)$ in comparison to 1 and obtain

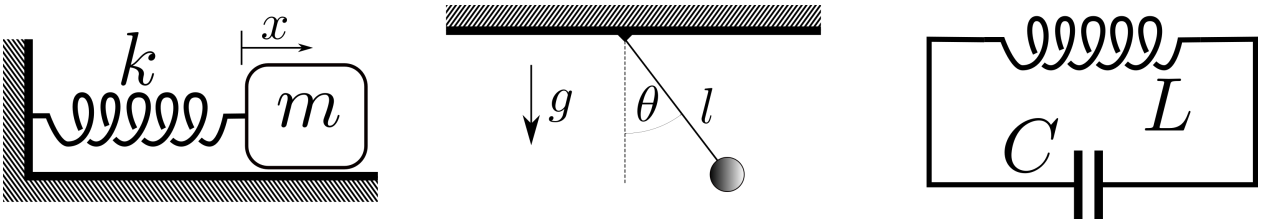
    $$x(T) \approx \frac{mv_0}{\gamma_1},$$

    the same result as for the case of no gravity.
  - Terminal velocity.

    $$t \to \infty, \qquad v_\infty = -\frac{mg}{\gamma_1}, \qquad mg = -v_\infty\gamma_1$$

LECTURE 6
Harmonic oscillations.

6.1. Harmonic oscillations.



- Equation:

$$m\ddot{x} = -kx, \qquad ml\ddot{\phi} = -mg\sin\phi \approx -mg\phi, \qquad -L\ddot{Q} = \frac{Q}{C},$$

All of these equation have the same form

$$\ddot{x} = -\omega_0^2 x, \qquad \omega_0^2 = \begin{cases} k/m \\ g/l \\ 1/LC \end{cases}, \qquad x(t=0) = x_0, \quad v(t=0) = v_0.$$

  - Second order differential equation $\longrightarrow$ two initial conditions.
  - Units of ω_0 are $[\omega_0] = s^{-1}$.
  - Notice the minus sign! This is a very important minus sign!!! It quarantines that the oscillator returns back – oscillates, instead of running away.
  - Notice the simple structure of the equation — the second derivative of a function is LINEARLY proportional to the function itself.
  - Notice, that you can read the frequency of the oscillations directly from the equation — the frequency of oscillation is simply a square root of the proportionality coefficient (without the minus sign). This is a unique property of the harmonic oscillator. It does NOT depend on initial conditions! For any other oscillations the frequency DOES depend on the initial conditions!
- The general solution is

$$x(t) = A\sin(\omega_0 t) + B\cos(\omega_0 t) = C\cos(\omega_0 t + \phi) = \Re\left(\mathcal{C}e^{i\omega_0 t}\right),$$

where A and B are arbitrary constants. $A = C\cos\phi$, $B = -C\sin\phi$, $\mathcal{C} = Ce^{i\phi}$.
- $C = \sqrt{A^2 + B^2}$ — amplitude; $\phi = -\tan^{-1}(B/A)$ — phase.

- Second order differential equation $\longrightarrow$ two arbitrary constants A and B, or C and ϕ.
- The velocity as a function of time is
$$v(t) = \dot{x} = \omega_0 A \cos(\omega_0 t) - \omega_0 B \sin(\omega_0 t).$$
- Our initial conditions give
$$x(t=0) = B = x_0, \qquad v(t=0) = A\omega_0 = v_0,$$
so the arbitrary constants are given by
$$B = x_0, \qquad A = \frac{v_0}{\omega_0}.$$
(check units)
- The solution for the given initial conditions is
$$x(t) = \frac{v_0}{\omega_0}\sin(\omega_0 t) + x_0 \cos(\omega_0 t)$$
- Oscillates forever. The frequency of oscillations ω_0 does not depend on the initial conditions and can be read straight from the equation of motion. This is the property of harmonic oscillations. It also means, that the frequency ω_0 is the property of the system itself, not of the way we set up the motion.

6.1.1. Energy conservation.

Energy. Conserved quantity: $E = \frac{m\dot{x}^2}{2} + \frac{m\omega_0^2 x^2}{2}$. It stays constant on a trajectory!
$$\frac{dE}{dt} = m\dot{x}\left(\ddot{x} + \omega_0^2 x\right) = 0.$$
So E is a constant — it does not depend on time during the motion. So the value of this constant during the motion is the same as at the initial moment of time. So the value of this constant can be obtained from the initial conditions $E = \frac{mv_0^2}{2} + \frac{m\omega_0^2 x_0^2}{2}$.

6.1.2. Limiting case $\omega_0 \to 0$.

Let's consider what happens to the oscillations in the limit $\omega_0 \to 0$.
In this case the equation is
$$\ddot{x} = -\omega_0^2 x \to 0$$
Se we expect to have $\ddot{x} = 0$, or $x(t) = v_0 t + x_0$ — the motion with the constant velocity.

Intuition: we know that for the "spring an mass" $\omega_0 = \sqrt{k/m}$, so the limit $\omega_0 \to 0$ means $k \to 0$, this means that there is no spring. Then we indeed expect the mass to be moving forever with the initial velocity.

However, we have an exact solution for the oscillator
$$x(t) = A\sin(\omega_0 t) + B\cos(\omega_0 t).$$
We want to take the limit $\omega_0 \to 0$ in this solution. We must recover the motion with constant velocity.

If we naively take $\omega_0 \to 0$ in the solution, we will get $x(t) = B$, which is incorrect. The fact that it is incorrect, can be seen from mathematics. The Newton's equation is a second order differential equation. So it requires TWO initial conditions and a solution MUST

depend on TWO arbitrary constants. The limit which we obtained depends on only ONE constant B. So we will not be able to accommodate TWO initial conditions.

What we need to do is to first impose the initial conditions: $x(t = 0) = x_0$ and $v(t = 0) = v_0$. Then we get

$$x(t) = \frac{v_0}{\omega_0} \sin(\omega_0 t) + x_0 \cos(\omega_0 t).$$

Now the limit $\omega_0 \to 0$ is not so trivial, as in the first term zero is divided by zero. So we need to use the Taylor expansion $\sin(\omega_0 t) \approx \omega_0 t$. Then we get

$$x(t) = v_0 t + x_0.$$

A lesson to remember.

- Oscillations with zero frequency is just a linear motion with constant velocity.
- Such oscillations are called zero modes.
- In an interacting system, such modes are a consequence of a symmetry.

LECTURE 7
Oscillations with dissipation.

- Equation of motion.

$$m\ddot{x} = -kx - \beta\dot{x}, \qquad -L\ddot{Q} = \frac{Q}{C} + R\dot{Q},$$

- Dissipation

$$\frac{dE}{dt} = \frac{d}{dt}\left(\frac{m\dot{x}^2}{2} + \frac{kx^2}{2}\right) = \dot{x}\left(m\ddot{x} + kx\right) = -\beta\dot{x}^2 < 0.$$

  If $\beta > 0$, the energy is decreasing! – dissipation!
- We simplify the equation a bit introducing $\omega_0^2 = k/m$ and $2\gamma = \beta/m$

$$\ddot{x} = -\omega_0^2 x - 2\gamma\dot{x}, \qquad x(t=0) = x_0, \quad v(t=0) = v_0.$$

- Units of γ are s^{-1} – the same as for ω_0.
- Solution: This is a linear equation with constant real coefficients. We look for the solution in the form $x = \Re C e^{-i\omega t}$, where ω and C are complex constants.

$$\omega^2 + 2i\gamma\omega - \omega_0^2 = 0, \qquad \omega = -i\gamma \pm \sqrt{\omega_0^2 - \gamma^2}$$

- Two solutions, two independent constants.
- Two cases: $\gamma < \omega_0$ and $\gamma > \omega_0$.

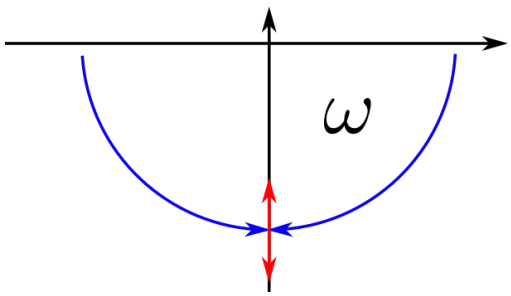


Figure 1. Evolution of the complex ω as γ increases.

7.0.1. The case $\gamma < \omega_0$ (underdamping):

- For $\gamma < \omega_0$ we may write
$$\omega = -i\gamma \pm \Omega, \qquad \Omega = \sqrt{\omega_0^2 - \gamma^2}.$$
- The solution in this case is
$$x = e^{-\gamma t}\Re\left[\mathcal{C}_1 e^{i\Omega t} + \mathcal{C}_2 e^{-i\Omega t}\right] = |C|e^{-\gamma t}\cos\left(\Omega t + \phi\right).$$
- Second order differential equation $\to$ two arbitrary constants $|C|$ and ϕ. The two constants are to be obtained from the initial conditions.
- Decaying oscillations. Shifted frequency.
- The rate of decay of oscillations is γ. The "life-time" of the oscillations $\sim 1/\gamma$.
- Notice, that γ is the negative of the imaginary part of the complex ω.
- For $\gamma \ll \omega_0$ we can use the Taylor expansion
$$\Omega \approx \omega_0 - \frac{1}{2}\frac{\gamma^2}{\omega_0}.$$
The frequency shift with respect to the undamped case ($\gamma = 0$) is proportional to γ^2.

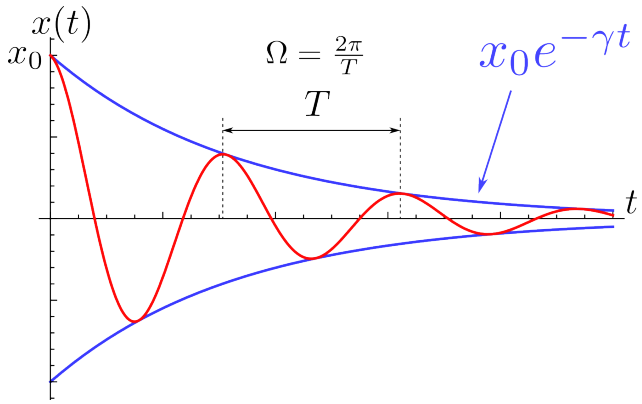


Figure 2. $x(t)$ for underdamped oscillations for the initial condition $x(t=0) = x_0$, $v(t=0) = v_0$.

7.0.2. The case $\gamma > \omega_0$ (overdamping):

- In this case the solution is
$$x = A_- e^{-\Gamma_- t} + A_+ e^{-\Gamma_+ t}, \qquad \Gamma_\pm = \gamma \pm \sqrt{\gamma^2 - \omega_0^2} > 0, \qquad \Gamma_+ > \Gamma_-.$$
- NO OSCILLATIONS!!!
- For the initial conditions $x(t=0) = x_0$ and $v(t=0) = 0$ we find
$$A_- = x_0\frac{\Gamma_+}{\Gamma_+ - \Gamma_-}, \qquad A_+ = -x_0\frac{\Gamma_-}{\Gamma_+ - \Gamma_-}$$
For $t \to \infty$ the A_+ term can be dropped as $\Gamma_+ > \Gamma_-$, then $x(t) \approx x_0\frac{\Gamma_+}{\Gamma_+ - \Gamma_-}e^{-\Gamma_- t}$.

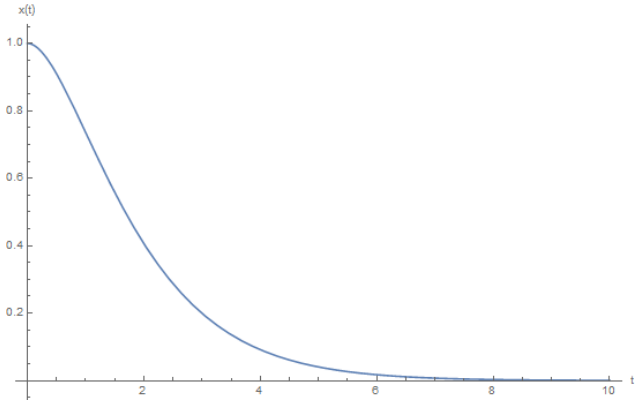


Figure 3. $x(t)$ for overdamped oscillations for the initial condition $x(t=0)=1$, $v(t=0)=v_0$.

7.1. The case of very strong damping.

This case corresponds to $\gamma \gg \omega_0$.
We found before that in the overdamped case:

$$x = A_- e^{-\Gamma_- t} + A_+ e^{-\Gamma_+ t}, \qquad \Gamma_\pm = \gamma \pm \sqrt{\gamma^2 - \omega_0^2} > 0$$

Consider a limit $\gamma \to \infty$. Then we have

$$\Gamma_+ \approx 2\gamma, \quad \text{– large}, \qquad \Gamma_- \approx \frac{\omega_0^2}{\gamma}, \quad \text{– small}$$

$$x_+(t) \approx A_+ e^{-2\gamma t}, \qquad x_-(t) \approx A_- e^{-\frac{\omega_0^2}{2\gamma} t}.$$

The first solution decays over time $\sim 1/2\gamma \to 0$ – almost immediately. The secons one almost does not decay!

Let's see where these solutions came from. In the equation

$$\ddot{x} = -\omega_0^2 x - 2\gamma \dot{x}$$

in the limit $\gamma \to \infty$ the last term is huge. It must be compensated by one of the others terms. Let's see what will happen if we drop the $\omega_0^2 x$ term. Then we get the equation $\ddot{x} = -2\gamma\dot{x}$. Its solution is $\dot{x} = Be^{-2\gamma t}$. After one more integration we see, that we will get the $x_+(t)$ solution.

Now let's see what will happen if we drop the $\ddot{x}$ term. We get the equation $\dot{x} = -\frac{\omega_0^2}{2\gamma} x$. Its solution is $x = Ae^{-\frac{\omega_0^2}{2\gamma} t}$ – this is our $x_-(t)$ solution.

LECTURE 8
External force. Resonance. Response.

- Why complex ω is always in the lower half plane?

8.1. Why harmonic oscillators are so important?

- The potential energy of the spring is $\frac{kx^2}{2}$ — a parabola as function of x.
- This parabolic potential energy is what defines the harmonic oscillator. Any system with such dependence of the potential energy will behave as a harmonic oscillator.
- Typically, the starting point to study a system is the system in equilibrium.
- In equilibrium a system is at a minimum of its potential energy.
- Any function close to its minimum can be well approximated by a parabola.
- A small disturbance of a system will leave the system close to its the minimum of the potential energy, where the potential energy can be approximated by a parabola.
- So any system close to equilibrium will behave as a collection of harmonic oscillators.

8.2. Response.

- In order to observe a system in equilibrium one must disturb it.
- Say you shine light on an object. The light interacts with the molecules, they start to vibrate and emit light back.
- You observe that back emitted light and analyze its properties: brightness, color, etc
- In more abstract terms: You have a system in equilibrium, you disturb/perturb it, you observe the response.
- So a property (say, color) of an object is a property of its response function!

The two points mean that we need to study how an oscillator responses to an external force.

8.3. External force.

In equilibrium everything is at the minimum of the potential energy, so we have the harmonic oscillator with dissipation. All we measure are the response functions, so we need to know how the harmonic oscillator behaves under external force.

- Let's add an external force:

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = f(t), \qquad x(t=0) = x_0, \quad v(t=0) = v_0.$$

- The full solution is the sum of the solution of the homogeneous equation with any solution of the inhomogeneous one. This full solution will depend on two arbitrary constants. These constants are determined by the initial conditions.
- Let's assume, that $f(t)$ is not decaying with time. Any solution of the homogeneous equation will decay in time. There is, however, a solution of the inhomogeneous equation which will not decay in time. So in a long time $t \gg 1/\gamma$ the solution of the homogeneous equation can be neglected. In particular this means that the asymptotic of the solution does not depend on the initial conditions.
- We can say, that the system "forgets" about the initial conditions after about the time $1/\gamma$. Notice, that this forgetfulness is the consequence of dissipation.
- Let's now assume that the force $f(t)$ is periodic with some period. It then can be represented by a Fourier series.

  $$f(t) = \sum_i f_i \sin(\Omega_i t + \phi_i).$$

  As the equation is linear the solution will also be a series,

  $$x(t) = \sum_i x_i(t)$$

  where each term corresponds to a force with a single frequency.

  $$\ddot{x}_i + 2\gamma\dot{x}_i + \omega_0^2 x_i = f_i \sin(\Omega_i t + \phi_i).$$

- So we can consider just one of these equations and shift time to compensate for the phase. So we need to solve

  $$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = f \sin(\Omega_f t),$$

  where f is the force's amplitude, and Ω_f is the force's frequency.

8.4. Resonance.

— *Resonance:*

- We need to solve:

  $$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = f \sin(\Omega_f t),$$

  where f is the force's amplitude, and Ω_f is the force's frequency.
- Let's look at the solution in the form $x = -f\Im C e^{-i\Omega_f t}$, and use $\sin(\Omega_f t) = -\Im e^{-i\Omega_f t}$.
- Substituting this guess into the equation we get

  $$C = \frac{1}{\omega_0^2 - \Omega_f^2 - 2i\gamma\Omega_f} = |C|e^{i\phi},$$

  $$|C| = \frac{1}{\left[(\Omega_f^2 - \omega_0^2)^2 + 4\gamma^2\Omega_f^2\right]^{1/2}}, \quad \tan\phi = \frac{2\gamma\Omega_f}{\omega_0^2 - \Omega_f^2}$$

  $$x(t) = -f\Im|C|e^{-i\Omega_f t + i\phi} = f|C|\sin\left(\Omega_f t - \phi\right),$$

- Resonance frequency — the frequency Ω_f at which amplitude $|C|$ — is at maximum for the position x measurement

  $$\Omega_f^r = \sqrt{\omega_0^2 - 2\gamma^2}.$$

- Phase changes sign at $\Omega_f^\phi = \omega_0$.

- Role of the phase: delay in response. The force is zero at $t = 0$, the response $x(t)$ is zero at $t = \phi/\Omega_f > 0$, so if $\phi > 0$ the response is "delayed" in comparison to the force.

— *Resonance in velocity measurement*

- The velocity is given by
$$v(t) = \dot{x}(t) = f\Im i\Omega_f C e^{-i\Omega_f t}.$$
- The velocity amplitude is given by
$$f\Omega_f|C| = f\frac{\Omega_f}{\left[(\Omega_f^2 - \omega_0^2)^2 + 4\gamma^2\Omega_f^2\right]^{1/2}} = f\frac{1}{[(\Omega_f - \omega_0^2/\Omega_f)^2 + 4\gamma^2]^{1/2}}$$
- The maximum is when $\Omega_f - \omega_0^2/\Omega_f = 0$, so the resonance frequency for the velocity is ω_0 — without the damping shift.
- Current is velocity.

8.5. Experiment.

We want to study an oscillator. We apply a "force" with some amplitude f and some frequency Ω_f. Without changing the amplitude f of the force slowly change the frequency Ω_f and we measure the amplitude of oscillation, either in "position" or in "velocity". (What the words "force", "position", and "velocity" mean depends on the oscillator.)

The measured amplitude will depend on the force's frequency and in case $\gamma \ll \omega_0$ it will have a very sharp peak. The position and the width of this peak tells you everything about the oscillator.

8.5.1. Analysis for small γ.

- To analyze resonant response we analyze $|C|^2$.
- The most interesting case $\gamma \ll \omega_0$, then the response $|C|^2$ has a very sharp peak at $\Omega_f \approx \omega_0$:

$$|C|^2 = \frac{1}{(\Omega_f^2 - \omega_0^2)^2 + 4\gamma^2\Omega_f^2} \approx \frac{1}{4\omega_0^2}\frac{1}{(\Omega_f - \omega_0)^2 + \gamma^2},$$

  so that the peak is very symmetric.
- $|C|^2_{\max} \approx \frac{1}{4\gamma^2\omega_0^2}$.
- to find HWHM we need to solve $(\Omega_f - \omega_0)^2 + \gamma^2 = 2\gamma^2$, so HWHM $= \gamma$, and FWHM $= 2\gamma$.
- Q factor (quality factor). The good measure of the quality of an oscillator is
$$Q = \omega_0/\text{FWHM} = \omega_0/2\gamma.$$
(decay time) $= 1/\gamma$, period $= 2\pi/\omega_0$, so $Q = \pi\frac{\text{decay time}}{\text{period}}$.

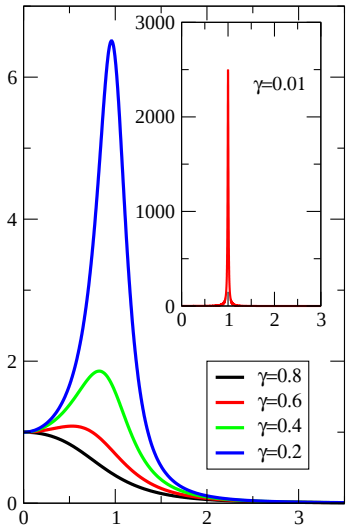


Figure: Resonant response. For insert $Q = 50$.

- Quality factor Q is the property of the resonator.
- Typical Q factors:
  - A grandfather's clock $Q \approx 100$.

– A quartz watch $Q \sim 10^4$.
– An atomic clocks $Q \sim 10^{11} - 10^{16}$.
– The insert in the Fig. shows the resonance for $Q = 50$.

8.6. Useful points.

- The *complex* response function

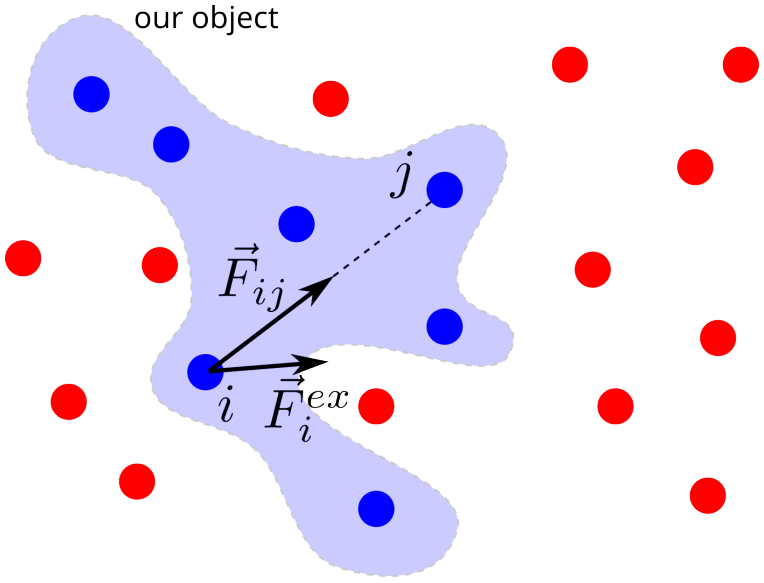
$$C(\Omega_f) = \frac{1}{\omega_0^2 - \Omega_f^2 - 2i\gamma\Omega_f}$$

as a function of *complex* frequency Ω_f has simple poles at $\Omega_f^p = -i\gamma \pm \sqrt{\omega_0^2 - \gamma^2}$. Both poles are in the lower half plane of the complex Ω_f plane. This is always so for any linear response function. It is the consequence of causality!
- The resonator with a high Q is a filter. One can tune this filter by changing the parameters of the resonator.
- By measuring the response function and its HWHM we can measure γ. By changing the parameters such as temperature, fields, etc. we can measure the dependence of γ on these parameters. γ comes from the coupling of the resonator to other degrees of freedom (which are typically not directly observable) so this way we learn something about those other degrees of freedom.

LECTURE 9
Momentum Conservation. Rocket motion. Charged particle in magnetic field.

9.1. Momentum Conservation.

It turns out that the mechanics formulated by Newton implies certain conservation laws. These laws allows us to find answers to many problems/questions without solving equations of motion. Moreover, they are very useful even when it is impossible to solve the equations of motion, as happens, for example, in Stat. Mech. But the most important aspect of the conservation laws is that they are more fundamental than the Newtonian mechanics itself. In Quantum mechanics or Relativity, or quantum field theory the very same conservation laws still hold, while the Newtonian mechanics fails.



- Consider a system of many point-like interacting particles.
- We select N of them and call it our object.
- We number the particles in our object with indexes $i = 1, \dots N$.
- All particles interact with each other and with all other particles outside of our object.

- At ANY moment of time t a particle j in the object acts on a particle i in the object with a force $\vec{F}_{ij}$.
- At the same moment of time the particle i also experiences a force $\vec{F}_i^{ex}$ from its interaction with everything outside of our object.
- Then for a particle i we can write the equation of motion, the second Newton's law

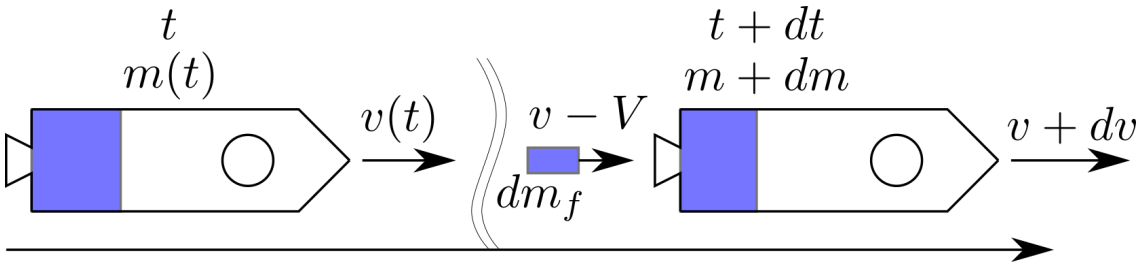
$$\dot{\vec{p}}_i = \vec{F}_i = \vec{F}_i^{ex} + \sum_j \vec{F}_{ij}.$$

  We take $F_{ii} = 0$ — no self action.
- According to the Newton's third law If a particle j acts on particle i with the force $\vec{F}_{ij}$, then the particle i acts on the particle j with the force $\vec{F}_{ji}$ and $\vec{F}_{ij} = -\vec{F}_{ji}$.
- Consider the total momentum of the whole object $\vec{P} = \sum_i \vec{p}_i$ — this is a simple vector sum of the momenta of individual particles of our object, then

$$\dot{\vec{P}} = \sum_i \dot{p}_i = \sum_i \vec{F}_i^{ex} + \sum_{i,j} \vec{F}_{ij} = \sum_i \vec{F}_i^{ex}.$$

  because $\sum_{i,j} \vec{F}_{ij} = 0$ as in this sum for every term $\vec{F}_{ij}$ there is a term $\vec{F}_{ji}$.
- So internal forces in an object do not contribute to the change of the total momentum.
- The total external force $\sum_i \vec{F}_i^{ex}$ is a simple vector sum of ALL external forces acting on All the particles of the object.
- The momentum of any closed object/system (when there is no interaction with outside $\vec{F}_i^{ex} = 0$) is conserved $\dot{\vec{P}} = 0$.
- Important points:
  - It is of paramount importance to clearly define what your object/system is and what the "outside" is.
  - The statement is only about the *total* momentum of the object/system.
  - The nature of the forces does not matter. They can be dissipative, or non-dissipative it will still work.
  - It is THE SUM of all outside forces that leads to the change of the total momentum. The points to which the forces are applied do not matter.
  - The momentum is a vector! there are three conservation laws — one for each component.
  - If only some components of the total external force are zero, then only the corresponding components of the total momentum will be conserved.
- Examples of the momentum conservation law.
  - A bullet hits a wooden block.

9.2. Rocket motion.



9.2.1. Statement of the problem:

- A rocket is a shell with the engine and the fuel.
- A rocket/engine burns fuel. The spent fuel is ejected with velocity V in the **frame of reference of the rocket**. The velocity V is the property of the engine.
- Both the mass of the rocket $m(t)$ (this is the total mass: the shell, the engine and the fuel) and its velocity $v(t)$ are functions of time t. The function $m(t)$ is in our hands – this is how we burn the fuel – how hard we press on the gas pedal.
- We want to find the function $v(t)$ — the rocket velocity as a function of time.
- The initial mass of the rocket is m_{initial}. The initial velocity of the rocket is v_{initial}.

9.2.2. Solution.

An important point is: when engine fires the rocket accelerates. So the rocket itself is NOT an inertial frame of reference. We need to work in some external inertial frame of reference. If a rocket has velocity v the velocity of the ejected fuel is $v - V$.

- At some time t the velocity of the rocket is $v(t)$ and its mass is $m(t)$.
- From this moment on our object is the rocket and the fuel which is on the rocket at time t. We discard all the fuel which was burnt before the moment t.
- The momentum of the system rocket+fuel at time t is

$$P_{\text{r+f}}(t) = mv.$$

  (I dropped the argument t in $m(t)$ and $v(t)$ for clarity, but it is still there.)
- Let's compute the total momentum of the system rocket+fuel at time $t+dt$, $P_{\text{r+f}}(t+dt)$.
  - The engine fires constantly. At time $t + dt$ the mass of the rocket changes and becomes $m + dm$ (where dm is negative), its velocity becomes $v + dv$. The momentum of the rocket is

$$P_{\text{r}}(t + dt) = (m + dm)(v + dv) \approx mv + mdv + vdm$$

    As dt is infinitesimally small, so are dm and dv, so keeping only the no more than linear in infinitesimal increments terms, we have

$$P_{\text{r}}(t + dt) \approx mv + mdv + vdm$$

  - The spent fuel has a mass dm_f and has velocity $v - V$, so its momentum is

$$P_{\text{f}} = (v - V)dm_f.$$

  - As the total mass of a rocket with the fuel does not change $dm + dm_f = 0$. So the momentum of the burned fuel is

$$P_{\text{f}} = -(v - V)dm.$$

- Thus the momentum of the system rocket+fuel at time $t + dt$ is

$$P_{\text{r+f}}(t + dt) = mv + mdv + vdm - (v - V)dm$$

- So the change of the total momentum of the system rocket+fuel during time dt is

$$dP_{\text{r+f}} = P_{\text{r+f}}(t + dt) - P_{\text{r+f}}(t) = mdv + vdm - (v - V)dm = mdv + Vdm.$$

- As there is no external forces acting on the system rocket+fuel the total momentum of this system must be conserved — must not change — so $dP_{\text{r+f}} = 0$

$$mdv = -V dm,$$
$$dv = -V\frac{dm}{m},$$
$$v_{\text{final}} = v_{\text{initial}} + V \log \frac{m_{\text{initial}}}{m_{\text{final}}}.$$

- Notice, that the answer does not depend on the exact form of the function $m(t)$. It depends only on the ratio of the initial mass to the final mass.
- As final moment of time is arbitrary we can write

$$v(t) = v_{\text{initial}} + V \log \frac{m_{\text{initial}}}{m(t)}$$

- Consider now that there is an external force F_{ex} acting on the rocket. Then we must have

$$\frac{dP_{\text{r+f}}}{dt} = F_{ex}.$$

or

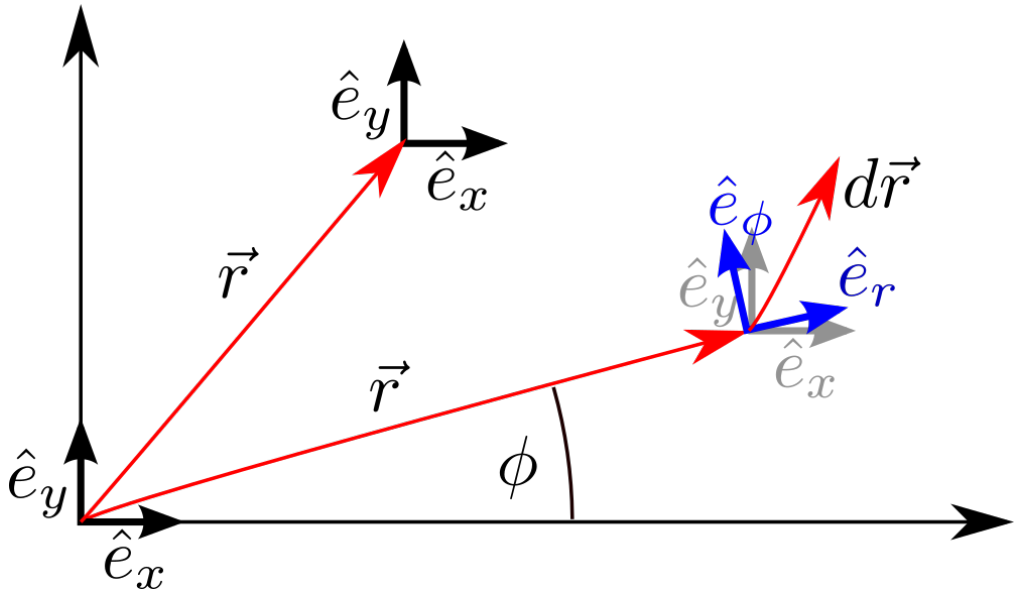
$$dP_{\text{r+f}} = F_{ex}dt, \qquad mdv = -V dm + F_{ex}dt, \qquad m\frac{dv}{dt} = F_{ex} - V\frac{dm}{dt}.$$

- This equation looks like the second Newton law if we say that there is a new force "thrust"$= -V\frac{dm}{dt}$, which acts on the rocket. Notice, that $\frac{dm}{dt} < 0$, so this force is positive.

9.3. Charged particle in magnetic field.

- Lorentz force: $\vec{F} = q\vec{v} \times \vec{B} + q\vec{E}$.
- No electric field — $\vec{F} \perp \vec{v}$, so there is no component of the force $\vec{F}$ along the vector of velocity $\vec{v}$, so $|\vec{v}| = \text{const}$.. Trajectories. $gvB = m\omega^2 R = m\omega v$, I used $\omega R = v$. Cyclotron frequency $\omega_c = \frac{qB}{m}$. Cyclotron radius $R_c = \frac{mv}{qB}$.
- Boundary effect.

LECTURE 10
Kinematics in cylindrical/polar coordinates.



We used to think of a space as a flat space. However, in many cases this is not correct. It is certainly not correct in general relativity, but also if restrict our motion to the Earth surface, then the motion happens on a (almost) sphere, which is not flat.

In a flat space we have a luxury to define the vectors globally. What it means is that the points of the space can be defined as position vectors of these points. It works because the sum of any two such vectors will also be a position vector of some point of our space. It is not so in a not flat, curved, space.

In a curved space we still can defined the coordinates, say latitude and longitude on the surface of Earth. But we cannot operate with these numbers as with vectors. However, at each (non-singular) point of a curved space we can consider an infinitesimally small neighborhood which is flat. So at any point we can define "its own" flat space, introduce the system of coordinates in each neighborhood and operate with the infinitesimal vectors in that neighborhood.

To restore the whole motion we ill need the infinitesimal displacement vector at each point and a way to translate the coordinates from one point to the next.

This is a general idea. Here we will consider how it works for a flat space. The trick is to watch that we never operate with global position vectors.

In this lecture we will consider different coordinate systems in flat $2D$ space.

10.1. Cartesian coordinates.

- The Cartesian coordinates are given by the origin and two unit vectors $\hat{e}_x$ and $\hat{e}_y$.
- These vectors have the following properties.
$$\hat{e}_x^2 = \hat{e}_y^2 = 1, \qquad \hat{e}_x \cdot \hat{e}_y = 0.$$
- These two vectors $\hat{e}_x$ and $\hat{e}_y$ are the same in any point of space. (It is possible to define such vectors only because the space is flat.)
- Any vector can be represented as
$$\vec{r} = x\hat{e}_x + y\hat{e}_y.$$
- We used to describing a position vector this way. However, the position vector is only possible in flat space.
- Instead of using the position vector, we will be using an infenitesimal vector of displacement: Say a particle moves from point (x, y) to point $(x + dx, y + dy)$, then the vector of displacement $d\vec{r}$ is
$$d\vec{r} = dx\hat{e}_x + dy\hat{e}_y$$
If we know the displacement vector $d\vec{r}$ at every moment of time AND we know how to translate the unit vectors $\hat{e}_x$ and $\hat{e}_y$ from one point to the next, then we will be able to restore the whole path! So to describe motion we do NOT need to define vectors globally!
- For a moving particle dividing $d\vec{r}$ by dt we find its vector of velocity
$$\vec{v} = \dot{x}\hat{e}_x + \dot{y}\hat{e}_y, \qquad v_x = \dot{x} \quad v_y = \dot{y}$$
- Differentiating the *vector* of the velocity we find the vector of acceleration
$$\vec{a} = \dot{\vec{v}} = \ddot{x}\hat{e}_x + \ddot{y}\hat{e}_y, \qquad a_x = \ddot{x} \quad a_y = \ddot{y}$$
- A trajectory is given by $x(t)$ and $y(t)$, where t is a parameter – usually time. If we are not interested on the time dependence, then we can give the trajectory as a function $y(x)$.

10.2. Polar coordinates

- In $2D$ we can use r and ϕ as coordinates.
- The polar coordinates are given by the origin and two vectors $\hat{e}_r$ and $\hat{e}_\phi$.
- Both $\hat{e}_r$ and $\hat{e}_\phi$ are different in different points of space. These vectors are not defined at the origin, they are defined in every point of space and are different from point to point.
- These vectors have the following properties *at every point of space*

(10.1)
$$\hat{e}_r^2 = \hat{e}_\phi^2 = 1, \qquad \hat{e}_r \cdot \hat{e}_\phi = 0.$$

- Our unit vectors $\hat{e}_r$ and $\hat{e}_\phi$ can be represented through the Cartesian vectors $\hat{e}_x$ and $\hat{e}_y$ at every point.
$$\begin{array}{l} \hat{e}_r = \hat{e}_x \cos\phi + \hat{e}_y \sin\phi \\ \hat{e}_\phi = -\hat{e}_x \sin\phi + \hat{e}_y \cos\phi \end{array}; \qquad \begin{array}{l} \hat{e}_x = \hat{e}_r \cos\phi - \hat{e}_\phi \sin\phi \\ \hat{e}_y = \hat{e}_r \sin\phi + \hat{e}_\phi \cos\phi \end{array}$$

Notice, that the transformation rules $(\hat{e}_r, \hat{e}_\phi) \leftrightarrow (\hat{e}_x, \hat{e}_r)$ do not depend on distance r — this is a peculiarity of this particular system, it is not a general property.

- Let's move from one point of space (ϕ, r) to another nearby $(\phi + d\phi, r + dr)$. As $(\hat{e}_x, \hat{e}_y)$ are the same in every point of space we have

$$d\hat{e}_r = -d\phi\hat{e}_x \sin\phi + d\phi\hat{e}_y \cos\phi = d\phi\left(-\hat{e}_x \sin\phi + \hat{e}_y \cos\phi\right) = d\phi\hat{e}_\phi$$
$$d\hat{e}_\phi = -d\phi\hat{e}_x \cos\phi - d\phi\hat{e}_y \sin\phi = -d\phi\left(\hat{e}_x \cos\phi + \hat{e}_y \sin\phi\right) = -d\phi\hat{e}_r$$

- Notice the following properties of this result

$$\hat{e}_r \cdot d\hat{e}_r = 0, \qquad \hat{e}_\phi \cdot d\hat{e}_\phi = 0, \qquad \hat{e}_r \cdot d\hat{e}_\phi + \hat{e}_\phi \cdot d\hat{e}_r = 0.$$

- From (10.1) it is easy to see, that first two of these relations are the consequence of $\hat{e}_r \cdot \hat{e}_r = \hat{e}_\phi \cdot \hat{e}_\phi = 1$ and the last one the consequence of $\hat{e}_r \cdot \hat{e}_\phi = 0$. In this sense these relations are more general than the particular form of $d\hat{e}_r$ and $d\hat{e}_\phi$.
- The first two relations show that a an infinitesimal increment of a unit vector must be orthogonal to that vector (its length must not change) The second relation shows how the orthogonal unit vectors must change in order to keep their orthogonality.
- If a point is moving as a function of a parameter, say time t, then dividing the above expression for $d\hat{e}_r$ and $d\hat{e}_\phi$ by dt we get:

$$\dot{\hat{e}}_r = \dot{\phi}\hat{e}_\phi, \qquad \dot{\hat{e}}_\phi = -\dot{\phi}\hat{e}_r$$

- If we simply differentiate the relationships (10.1), then we get

$$\hat{e}_r \cdot \dot{\hat{e}}_r = \hat{e}_\phi \cdot \dot{\hat{e}}_\phi = 0, \qquad \hat{e}_r \cdot \dot{\hat{e}}_\phi = -\hat{e}_\phi \cdot \dot{\hat{e}}_r$$

10.3. Motion in polar coordinates.

- If a particle moves from point (r, ϕ) to point $(r + dr, \phi + d\phi)$, then the vector of displacement is

$$d\vec{r} = dr\hat{e}_r + rd\phi\hat{e}_\phi$$

Notice, that the vectors $\hat{e}_r$ and $\hat{e}_\phi$ are taken at point (r, ϕ).

- The vector of velocity is simply $\vec{v} = \frac{d\vec{r}}{dt}$.

$$\vec{v} = \frac{d\vec{r}}{dt} = \dot{r}\hat{e}_r + r\dot{\phi}\hat{e}_\phi.$$

We can get the same result differently, simply writing, that $\vec{r} = r\hat{e}_r$ (this is only possible in a flat space!), and using $\dot{\hat{e}}_r = \dot{\phi}\hat{e}_\phi$

$$\vec{v} = \dot{\vec{r}} = \dot{r}\hat{e}_r + r\dot{\hat{e}}_r = \dot{r}\hat{e}_r + r\dot{\phi}\hat{e}_\phi.$$

- We see that the components of the velocity in the polar coordinates are given by

$$v_r = \dot{r}$$
$$v_\phi = r\dot{\phi}$$

- Acceleration – we *must* differentiate the *vector* of the velocity! NOT ITS COMPONENTS!!!

$$\vec{a} = \dot{\vec{v}} = \ddot{r}\hat{e}_r + \dot{r}\dot{\hat{e}}_r + \dot{r}\dot{\phi}\hat{e}_\phi + r\ddot{\phi}\hat{e}_\phi + r\dot{\phi}\dot{\hat{e}}_\phi = \left(\ddot{r} - r\dot{\phi}^2\right)\hat{e}_r + \left(r\ddot{\phi} + 2\dot{r}\dot{\phi}\right)\hat{e}_\phi.$$

from where we can read the components of the acceleration vector $\vec{a}$

$$a_r = \ddot{r} - r\dot{\phi}^2$$
$$a_\phi = r\ddot{\phi} + 2\dot{r}\dot{\phi}$$

- In the case $r = \text{const}$, $\dot{\phi} = \omega$, we have $\vec{a} = -r\omega^2\hat{e}_r + r\dot{\omega}\hat{e}_\phi$.
- Notice, if $\dot{\phi} = \omega = \text{const}$, then $a_\phi = 2\dot{r}\omega$ – this is the origin of the Coriolis force.
- In the polar coordinates we use r and ϕ to describe the position. As our space is flat we can write the position vector $\vec{r}$ as

  $$\vec{r} = r\hat{e}_r(r, \phi)$$

  Notice, that differentiating this relation with respect to time we will recover all previous formulas. But such construction is possible only in flat space, and is not needed!

10.4. Free motion.

Free motion means that there are no forces, so $\vec{a} = 0$.

10.4.1. Cartesian coordinates.

- In Cartesian coordinates $\vec{a} = 0$ gives

  $$\ddot{x} = 0, \quad \ddot{y} = 0, \qquad x(t) = v_{x,0}t + x_0, \quad y(t) = v_{y,0}t + y_0.$$

  where constants $v_{x,0}$, $v_{y,0}$, x_0, and y_0 are obtained from the initial conditions.
- Or the trajectory

  $$y = y_0 + \frac{v_{y,0}}{v_{x,0}}(x - x_0).$$

  This is the equation for a straight line in the Cartesian coordinates.

10.4.2. Polar coordinates.

- In the polar coordinates: $\vec{a} = 0$, so both components of $\vec{a}$ must be zero

$$\begin{array}{lllllll} a_\phi = 0 & \Longrightarrow & r\ddot{\phi} + 2\dot{r}\dot{\phi} = 0 & \Longrightarrow & \frac{d(r^2\dot{\phi})}{dt} = 0 & \Longrightarrow & r^2\dot{\phi} = \text{const} = A \quad \Longrightarrow \quad \dot{\phi} = \frac{A}{r^2} \\ a_r = 0 & \Longrightarrow & \ddot{r} - r\dot{\phi}^2 = 0 & \Longrightarrow & \ddot{r} - \frac{A^2}{r^3} = 0 & & \end{array}$$

  The constant A must be obtained from the initial conditions.
- Notation

  $$\frac{\partial}{\partial x} \equiv \partial_x$$

- Now I will do the following trick. Instead of two functions $r(t)$ and $\phi(t)$ I will consider a function $r(\phi)$ — the trajectory — and use

  $$\frac{\partial}{\partial t} = \frac{\partial\phi}{\partial t}\frac{\partial}{\partial\phi} = \dot{\phi}\frac{\partial}{\partial\phi} = \frac{A}{r^2}\partial_\phi; \qquad \dot{r} = \frac{A}{r^2}\partial_\phi r = -A\partial_\phi\frac{1}{r}; \qquad \ddot{r} = -\frac{A^2}{r^2}\partial_\phi^2\frac{1}{r},$$

  then $\ddot{r} - \frac{A^2}{r^3} = 0$ becomes

  $$-\frac{A^2}{r^2}\partial_\phi^2\frac{1}{r} - \frac{A^2}{r^3} = 0, \qquad \partial_\phi^2\frac{1}{r} = -\frac{1}{r}, \qquad \frac{1}{r} = B\cos(\phi - \phi_0)$$

- This is the equation of the straight line in the polar coordinates. Why? Can you plot it?

LECTURE 11
Angular momentum. Angular velocity.

11.1. Angular momentum.

For a particle at position $\vec{r}$ which has a momentum $\vec{p}$ we can define a vector of angular momentum $\vec{J}$:

$$\vec{J} = \vec{r} \times \vec{p}.$$

Notice, that this definition requires you to chose the coordinate origin. If you shift the origin the angular momentum in the new coordinate system will be different.

This definition is given for a single particle. If we have many particles we simply sum up their angular momenta (as vectors!) to get the total angular momentum of the system of particles.

Let's see how the total angular momentum of a bunch of particles behaves under the action of internal and external forces.

- Out of all available particles we mentally separate an arbitrary collection and call this collection our object. Now we can distinguish between the inside and outside. These words refer to our object.
- Consider a bunch of particles of the object. We assume that they interact with each other by central forces. This means that the force with which the particle j acts on particle i is along the vector from the particle i to the particle j:

$$\vec{F}_{ij} \parallel \vec{r}_i - \vec{r}_j$$

.
- Also, by the Newton's third law

$$\vec{F}_{ij} = -\vec{F}_{ji}.$$

- There is also external forces $\vec{F}_i^{ex}$ acting on each particle.
- We now chose a point of space to be our coordinate origin (this MUST be done, or all the rest makes no sense!). We then can write the vector of total angular momentum

$$\vec{J} = \sum_i \vec{r}_i \times \vec{p}_i.$$

- Consider the time evolution of the vector $\vec{J}$:

$$\dot{\vec{J}} = \sum_i \dot{\vec{r}}_i \times \vec{p}_i + \sum_i \vec{r}_i \times \dot{\vec{p}}_i.$$

As $\dot{\vec{r}}_i = \vec{v}_i \parallel \vec{p}_i$, each term in the first sum is zero. So we can drop the first sum. In the second term we use $\dot{\vec{p}}_i = \sum_{j\neq i} \vec{F}_{ij} + \vec{F}_i^{ex}$ — the sum of all forces acting on the particle i.

$$\dot{\vec{J}} = \sum_i \vec{r}_i \times \dot{\vec{p}}_i = \sum_i \vec{r}_i \times \left[\sum_{j\neq i} \vec{F}_{ij} + \vec{F}_i^{ex} \right] = \sum_{i\neq j} \vec{r}_i \times \vec{F}_{ij} + \sum_i \vec{r}_i \times \vec{F}_i^{ex}$$

- If a force $\vec{F}$ is applied to a point with the position $\vec{r}$ with respect to the origin (the origin MUST be the same as for the definition of $\vec{J}$), then the torque of this force with respect to the same origin is given by
$$\vec{\tau} \equiv \vec{r} \times \vec{F}.$$
- The last term in the equation for $\dot{\vec{J}}$ is
$$\vec{\tau}_{ex} = \sum_i \vec{r}_i \times \vec{F}_i^{ex}$$
and is the torque of external forces.
- Torque is a vector. You compute the vector of external torque for each particle of our abject and then sum up all these vectors.
- Consider now the first sum in the RHS. Remember that $\vec{F}_{ij} = -\vec{F}_{ji}$
$$\sum_{i\neq j} \vec{r}_i \times \vec{F}_{ij} = \frac{1}{2} \sum_{i\neq j} \vec{r}_i \times \vec{F}_{ij} + \frac{1}{2} \sum_{i\neq j} \vec{r}_j \times \vec{F}_{ji} = \frac{1}{2} \sum_{i\neq j} (\vec{r}_i - \vec{r}_j) \times \vec{F}_{ij} = 0$$
It is zero because $\vec{F}_{ij}$ is parallel to $\vec{r}_i - \vec{r}_j$.
- So we have
$$\dot{\vec{J}} = \vec{\tau}_{ex}$$
Remember: in this equation both $\vec{J}$ and $\vec{\tau}_{ex}$ MUST be defined with respect to the same origin, which also MUST be defined before the start of any calculations.
- If the torque of external forces is zero, then the angular momentum is conserved.

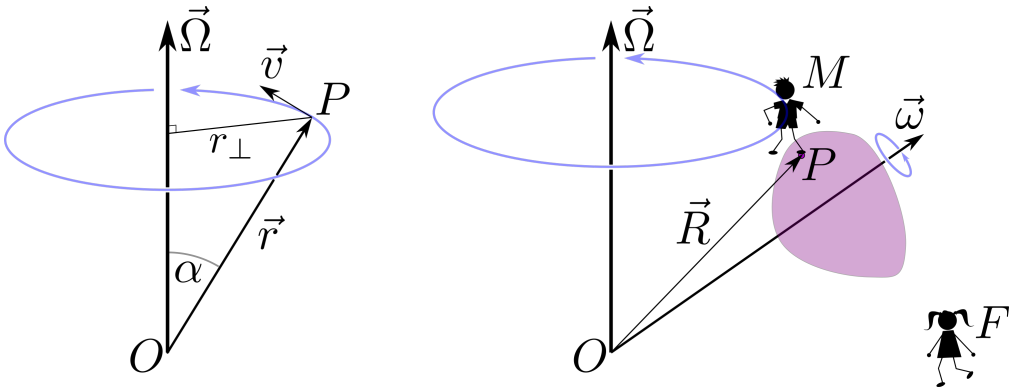
11.2. Rigid body.

Consider a rigid body which can rotate around a FIXED axis which goes through its center of mass. We apply a force $\vec{F}$ to some point of the body.

- Depending on the direction of the force the body may or may not rotate with increasing frequency.
- In any case the body as a whole will not move.
- It means that the axis must apply a force $-\vec{F}$ to the body.
- So the sum of all forces applied to the body is zero.
- What then causes the angular velocity to change?
- Consider a small piece of the body.
- Its velocity is changing! So there must be a net force acting on it.
- This is the force of interaction of our small piece with the rest of the body.
- Such forces are very difficult to compute, but
- If the body is rigid, then me know that the relative position of the points of the body does not change.

- It turns out that this observation is enough to construct the theory of the motion of a rigid body without any reference to the internal forces.
- However, if the body is not absolutely rigid, one has to go back, split the body into infinitesimally small pieces and consider the motion of each piece.

11.3. Angular velocity. Rotation of a rigid body.



Here we DEFINE the angular velocity vector. The length of this vector is the magnitude of the angular velocity and its direction is along the axis of rotation at the direction determined by the right hand rule.

- Consider a point P rotating around an axis with the angular velocity Ω (see the left panel of the figure)
- This point has velocity. The magnitude of this velocity is $v = \Omega r_\perp$, where $r_\perp$ is the distance from P to the axis of rotation.
- Let's take an arbitrary point O on the axis of rotation as our coordinate origin.
- The point P at some moment of time has a position vector $\vec{r}$.
- The distance from P to the axis is $r_\perp = r \sin\alpha$.
- So the magnitude of the point P velocity is $v = \Omega r \sin\alpha$.
- The direction of this velocity is perpendicular to both the axis of rotation and the vector $\vec{r}$.
- So if we *define* the vector of angular velocity $\vec{\Omega}$ as a vector with magnitude Ω directed along the axis (as shown), then we can write:

$$\vec{v} = \vec{\Omega} \times \vec{r}.$$

- Using the definition of the velocity $\vec{v} = \dot{\vec{r}}$ we can write the above as

$$\dot{\vec{r}} = \vec{\Omega} \times \vec{r}.$$

- Consider two arbitrary points of a body $\vec{r}$ and $\vec{r'}$. Their velocities are

$$\dot{\vec{r}} = \vec{\Omega} \times \vec{r}, \qquad \dot{\vec{r}}' = \vec{\Omega} \times \vec{r}'$$

Let's see how the distance d between the points changes with time. In order to do that we differentiate $d^2 = (\vec{r} - \vec{r}')^2$ over time

$$\dot{d}^2 = 2(\vec{r} - \vec{r}') \cdot (\dot{\vec{r}} - \dot{\vec{r}}') = 2(\vec{r} - \vec{r}') \cdot \left[\vec{\Omega} \times (\vec{r} - \vec{r}')\right] = 0.$$

So the distance between any two points of the rotating body is not changing with time!

- Notice the importance of the formula $\dot{\vec{r}} = \vec{\Omega} \times \vec{r}$:
  - It gives not only the magnitude, but also the direction of the velocity vector $\vec{v}$!
  - For a rigid body it tells us, that the velocity of every infenitesimal part of the body depends on the very same vector $\vec{\Omega}$. So instead of specifying the velocity of each part separately, we can completely describe the motion by specifying the vector $\vec{\Omega}$ and how it changes with time (both magnitude and direction!).

Now lets consider an arbitrary vector $\vec{l}$ which is constant in the rotating frame. We are interested in how it will change with time as seen from the rest (outside) frame (observer).

- We have a frame rotating with angular velocity $\vec{\omega}$ with respect to the rest frame. A vector $\vec{l}$ constant in the rotating frame will change with time in the rest frame and
$$\dot{\vec{l}} = \vec{\omega} \times \vec{l}.$$
(this is the same as for the position vector $\vec{r}$.)
- $\omega = \frac{d\phi}{dt}$, if ω is a vector $\vec{\omega}$, then $d\phi$ must be a vector $d\vec{\phi}$. Notice, that ϕ is not a vector! while $d\vec{\phi}$ is! This is the same as in the previous lecture, where we could not define the position vector, but could define the vector of infenitesimal displacement.
- The direction of the vector $d\vec{\phi}$ is along the axis of rotation according to the right hand rule.
- If we rotate one frame with respect to another by a small angle $d\vec{\phi}$, then a vector $\vec{l}$ will change by
$$d\vec{l} = d\vec{\phi} \times \vec{l}.$$

11.4. Angular velocity IS a VECTOR.

Our definition of the vector $\vec{\Omega}$ is, however, incomplete, unless we show that those vectors can be summed up and the result makes sense. Let's do just that:

- Look at the right panel of the Figure.
- Suppose we have a pink body rotating with angular velocity $\vec{\omega}$. Notice, by specifying the vector $\vec{\omega}$ I specify both the magnitude, and the axis of the rotation.
- Suppose the axis of the rotation $\vec{\omega}$ rotates with the angular velocity $\vec{\Omega}$ (see figure).
- Consider two observers F — the outside observer, and M — the observer who rotates with angular velocity $\vec{\Omega}$ (but does not rotate with $\vec{\omega}$).
- Consider the point P of the pink body.
- Let's take a moment of time t when the positions of the observer M and the position of the point P of the pink body coincide and is given by the position vector $\vec{R}$.
- From the point of view of the observer M the axis $\vec{\omega}$ is stationary. For this observer the point P has velocity
$$\vec{v}' = \vec{\omega} \times \vec{R}.$$
- As seen by the observer F, the observer M is rotating with $\vec{\Omega}$ and has a velocity
$$\vec{v}_M = \vec{\Omega} \times \vec{R}.$$
- According to Galileo the velocity $\vec{v}$ of the point P of the pink body as measured by the observer F is
$$\vec{v} = \vec{v}_M + \vec{v}'.$$

- Taken $\vec{v}_M$ and $\vec{v}'$ from above we get
$$\vec{v} = \vec{\Omega} \times \vec{R} + \vec{\omega} \times \vec{R} = \left(\vec{\Omega} + \vec{\omega}\right) \times \vec{R}.$$
- As it was done for arbitrary point P we conclude, that the outside observer F sees the pink body rotating with the angular velocity
$$\vec{\Omega}_F = \vec{\Omega} + \vec{\omega}.$$

Which is exactly what we wanted — angular velocity is a vector!

LECTURE 12

~~Moment~~ Tensor of inertia. Kinetic energy.

In the previous lecture we defined the vector of the total angular momentum of any set of particles

$$\vec{J} = \sum_i \vec{J}_i = \sum_i \vec{r}_i \times \vec{p}_i.$$

We also showed that for any a set of particles, if they interact through the central forces, the rate of change of angular momentum equals to the total torque of external forces only. In proving this statement the condition of rigidity was not used at all. The statement $\dot{\vec{J}} = \vec{\tau}^{ex}$ is very general.

In this lecture we show how to compute the angular momentum and the kinetic energy for a rigid body. Remember, that the condition of rigidity is very strong. The equation

$$\vec{v} = \vec{\omega} \times \vec{r}.$$

allows us to compute the velocity of *every* point of the body by knowing only one vector $\vec{\omega}$.

- So both the angular momentum and the kinetic energy will depend only on the vector $\vec{\omega}$ and some property of the body itself.

The total angular momentum of a body is linear in the velocities of all the particles. The velocities of all the particles are linear in the vector of angular velocity $\vec{\omega}$ of the body. So the vector of angular momentum of the body $\vec{J}$ is linear with the vector of angular velocity $\vec{\omega}$

$$\vec{J} \sim \vec{\omega}.$$

The most general linear relation between to vectors can be written through a tensor

$$J^\alpha = I^{\alpha\beta}\omega^\beta,$$

(Einstein notations are assumed) where $I^{\alpha\beta}$ is some tensor (3×3 matrix) which does not depend on $\vec{\omega}$ and depends only on the body — it is the property of the body.

- This tensor $I^{\alpha\beta}$ is called tensor of inertia.

Let's now consider the total kinetic energy of a rigid body rotating with the angular velocity $\vec{\omega}$. The total kinetic energy of the body is just a sum of the kinetic energies of all the particles of the body.

$$K = \frac{1}{2}\sum_i m_i \vec{v}_i^2 = \frac{1}{2}\sum_i m_i \vec{v}_i \cdot \vec{v}_i = \frac{1}{2}\sum_i \vec{p}_i \cdot \vec{v}_i.$$

Using $\vec{v}_i = \vec{\omega} \times \vec{r}_i$ we get

$$K = \frac{1}{2}\sum_i \vec{p}_i \cdot \vec{\omega} \times \vec{r}_i = \frac{1}{2}\sum_i \vec{\omega} \cdot \vec{r}_i \times \vec{p}_i = \frac{1}{2}\vec{\omega} \cdot \sum_i \vec{r}_i \times \vec{p}_i = \frac{1}{2}\vec{\omega} \cdot \vec{J} = \frac{1}{2}\omega^\alpha I^{\alpha\beta}\omega^\beta$$

with the SAME tensor of inertia $I^{\alpha\beta}$.

- Tensor of inertia is the property of the body which we must know in order to write equations for the rotation of the body (the same as mass m of a body which we must know in order to write equations for the translation motion of the body.)
- The expression for the kinetic energy tells us that the tensor $I^{\alpha\beta}$ must be positive definite (kinetic energy cannot be negative, as it is a sum of positive numbers.)
- Inverting the relation $J^\alpha = I^{\alpha\beta}\omega^\beta$ we get $\omega^\alpha = (\hat{I}^{-1})^{\alpha\beta}J^\beta$. Using this int he kinetic energy we get

$$K = \frac{1}{2}\left(\hat{I}^{-1}\right)^{\alpha\beta} J^\alpha J^\beta.$$

12.1. Angular momentum. ~~Moment~~ Tensor of inertia.

- Consider a rigid set of particles of masses m_i — the distances between the particles are fixed and do not change. The whole system rotates with the angular velocity $\vec{\omega}$. Each particle has a radius vector $\vec{r}_i$ with respect to the coordinate origin, which is on the axis of rotation. Let's calculate the angular momentum of the whole system.

$$\vec{J} = \sum_i m_i \vec{r}_i \times \vec{v}_i = \sum_i m_i \vec{r}_i \times [\vec{\omega} \times \vec{r}_i] = \sum_i m_i \left(\vec{\omega}\vec{r}_i^2 - \vec{r}_i(\vec{\omega} \cdot \vec{r}_i)\right)$$

  or in components (Einstein notations are assumed over Greek indexes)

$$J^\alpha = \sum_i m_i \left(\omega^\alpha \vec{r}_i^2 - r_i^\alpha \omega^\beta r_i^\beta\right) = \sum_i m_i \left(\delta^{\alpha\beta}\vec{r}_i^2 - r_i^\alpha r_i^\beta\right)\omega^\beta = I^{\alpha\beta}\omega^\beta,$$
$$I^{\alpha\beta} = \sum_i m_i \left(\delta^{\alpha\beta}\vec{r}_i^2 - r_i^\alpha r_i^\beta\right)$$

- Greek indexes α and β label the coordinate components, say x, y, and z, or 1, 2, and 3. For example, if $\alpha = x$, then $r^{\alpha=x} = x$, or if $\alpha = 1$, then $r^{\alpha=1} = x$ if x is the first coordinate.
- The Latin index i labels the particles: particle number 1, particle number 2, etc.
- The moment of inertia is a positive definite symmetric 3×3 tensor!

$$\hat{I} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix}, \qquad I^{\alpha\beta} = I^{\beta\alpha}.$$

  It transforms one vector into another:

$$\vec{J} = \hat{I}\vec{\omega}.$$

- Important: As $\hat{I}$ is a tensor and not a number the directions of the vectors $\vec{\omega}$ and $\vec{J}$ do not coincide. These two vectors can have very different directions.
- As for any symmetric tensor:

  - There are special coordinate axes in which the tensor has a diagonal form – only diagonal elements are nonzero, while all the off diagonal elements are zero. In this very specific for the body coordinate axes the tensor of inertia has the form:
$$\hat{I} = \begin{pmatrix} I_x & 0 & 0 \\ 0 & I_y & 0 \\ 0 & 0 & I_z \end{pmatrix}$$
  - These diagonal elements I_x, I_y, and I_z are called PRINCIPLE MOMENTS OF INERTIA. The corresponding axes are called PRINCIPAL AXES OF INERTIA.
  - If all the principal moments are different, then the principle axes are orthogonal to each other.
  - In a degenerate case these axes can be chosen to be orthogonal.
  - These principle axes are "attached" to the body, so if the body is rotating, then these axes are also rotating with the body.
- The direction of the angular momentum $\vec{J}$ and direction of the angular velocity $\vec{\omega}$ do not in general coincide!
- It is $\vec{J}$ which is constant when there are no external torques, not $\vec{\omega}$! Let me repeat it: If there are no external torques the vector $\vec{\omega}$ may change with time — both its direction and magnitude. But the angular momentum vector $\vec{J}$ will remain constant.

  Contrast this to the usual momentum-velocity relation
$$\vec{p} = m\vec{v}$$
  where the conservation of momentum means that the velocity is also constant. This is because the mass m is a scalar, not tensor.
- This last statement makes even the kinematics (motion with no external forces) of a rigid body very complicated and highly non-trivial.

12.2. Kinetic energy.

- Consider the kinetic energy of the rigid body.

$$K = \frac{1}{2}\sum_i m_i \vec{v}_i^2 = \frac{1}{2}\sum_i m_i [\vec{\omega} \times \vec{r}_i]^2 = \frac{1}{2}\sum_i m_i [\vec{\omega}^2 \vec{r}^2 - (\vec{\omega}\cdot\vec{r})^2] = \frac{1}{2}\sum_i m_i [\delta^{\alpha\beta}\vec{r}^2 - r^\alpha r^\beta]\omega^\alpha\omega^\beta.$$

  so we get
$$K = \frac{I^{\alpha\beta}\omega^\alpha\omega^\beta}{2}$$
  (Einstein notations are assumed over Greek indexes)
- This also shows that $\hat{I}$ is positive definite, as the kinetic energy MUST be positive (look at the first equality, it is a sum of non-negative numbers!) for ANY vector $\vec{\omega}$.
- In terms of angular momentum:
$$K = \frac{1}{2}\left(\hat{I}^{-1}\right)^{\alpha\beta} J^\alpha J^\beta.$$

12.3. Tensor of inertia for a continuous body.

Tensor of inertia is a property of the body, it DOES NOT depend on the body's motion. So we simply take the body to be stationary.

- First, we chose a system of coordinates.

- We split the body into infinitesimally small pieces. Each piece has its coordinate vector $\vec{r}$ and its mass $dm(\vec{r})$.
- Tensor of inertia of a continuous body.

$$I^{\alpha\beta} = \int \left(\delta^{\alpha\beta}\vec{r}^2 - r^\alpha r^\beta\right) dm = \int \left(\delta^{\alpha\beta}\vec{r}^2 - r^\alpha r^\beta\right) \frac{dm}{dV} dV = \int \left(\delta^{\alpha\beta}\vec{r}^2 - r^\alpha r^\beta\right) \rho(\vec{r}) dV,$$

where $\rho(\vec{r})$ is the mass density of the material at point $\vec{r}$ – it must be know as this is a characteristic of the body.

- How to compute the moment of inertia of an arbitrary body.
  - First you chose a system of coordinates registered with the body.
  - You chose which component of the tensor of inertia you want to compute. You have to compute all of them, but you need to start with something. Let's say it is I^{xy}.
  - Then in the expression $\int \left(\delta^{\alpha\beta}\vec{r}^2 - r^\alpha r^\beta\right) \rho(\vec{r}) dV$ we have $\alpha = x$ and $\beta = y$, or $I^{xy} = \int \left(\delta^{xy}\vec{r}^2 - r^x r^y\right) \rho(\vec{r}) dV$
  - The first term under the integral is then zero, as $\delta^{xy} = 0$.
  - In the second term $r^{\alpha=x} = x$, and $r^{\beta=y} = y$, so we have

$$I^{xy} = -\iiint_V xy\rho(x,y,z)dxdydz.$$

  - Let's say we want to compute I^{xx}. Then $\alpha = x$, and $\beta = x$, so the first term $\delta^{\alpha\beta}\vec{r}^2 = \delta^{xx}(x^2+y^2+z^2) = x^2+y^2+z^2$, as $\delta^{xx} = 1$, and $\vec{r}^2 = x^2+y^2+z^2$. The second term is just $r^x r^x = x^2$. So we need to compute

$$I^{xx} = \iiint \left(y^2+z^2\right) \rho(x,y,z)dxdydz.$$

12.3.1. Examples.

- Tensor of inertia for a $2D$ object. Let $\sigma(x,y)$ be the area mass density. The object is in xy plane, its z coordinate is zero $z = 0$ for all small pieces! Then

$$I^{zz} = \iint_S (x^2+y^2)\sigma(x,y)dxdy, \qquad I^{xx} = \iint_S y^2\sigma(x,y)dxdy, \qquad I^{yy} = \iint_S x^2\sigma(x,y)dxdy.$$

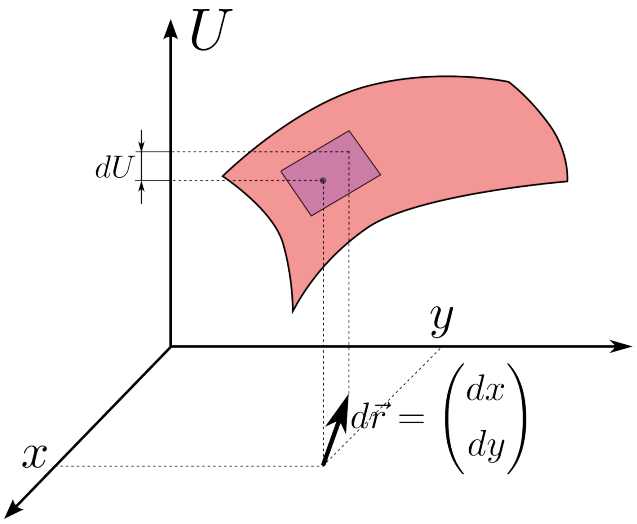
  So

$$I^{zz} = I^{xx} + I^{yy}.$$

- A thin ring: $I_{zz} = mR^2$, $I_{xx} = I_{yy} = \frac{1}{2}mR^2$, all off diagonal elements vanish.
- A disc: $I_{zz} = \frac{1}{2}mR^2$, $I_{xx} = I_{yy} = \frac{1}{4}mR^2$, all off diagonal elements vanish.
- A sphere: $I_{xx} = I_{yy} = I_{zz} = \frac{2}{5}mR^2$, all off diagonal elements vanish.
- A stick at the end: $I_{xx} = I_{yy} = \frac{1}{3}mL^2$, $I_{zz} = 0$.
- A stick at the center: $I_{xx} = I_{yy} = \frac{1}{12}mL^2$, $I_{zz} = 0$.
- Role of symmetry.

LECTURE 13
Work. Potential energy.

13.1. Mathematical preliminaries.



- Functions of many variables, say $U(x, y)$.
  - The most important part is to understand what is a function and what is not a function: A function of, say two variables x and y is a map from the (x, y) plane to a number line U. This means that a point on the (x, y) plane has ONLY ONE image in U under this map.
    We all now examples of functions. An example of a "not function" would be a multi-story building. A person in such building with coordinates x and y on the ground may have different height, as he/she can be on any floor. Thus we cannot map a position on the (x, y) plane to the position on the vertical line z. In case of the multi-story building one may try to have separate functions for each floor. However, there are stairs, so a person can move from floor to floor.
- Differential of a function of many variables.

$$dU = \frac{\partial U}{\partial x} dx + \frac{\partial U}{\partial y} dy.$$

  - The meaning of this expression is that at any ($2D$ in this case) point the surface looks like a plane.
  - We know that this is correct for any (smooth enough) function.
  - We can ask the opposite question: if we have an expression of the type as above, will there be a function which has this expression as a differential.
- Consider an expression

$$\delta G = A(x,y)dx + B(x,y)dy.$$

  where A and B are some arbitrary functions. The question is: is this a differential of some function? The answer is: not necessarily. The proof:
  - Let's assume that δG is a differential of some function U, then we must have

$$A = \frac{\partial U}{\partial x}, \qquad B = \frac{\partial U}{\partial y}.$$

  - But then

$$\frac{\partial A}{\partial y} = \frac{\partial^2 U}{\partial x \partial y} = \frac{\partial B}{\partial x}.$$

  - So δG is a differential of some function if (and only if)

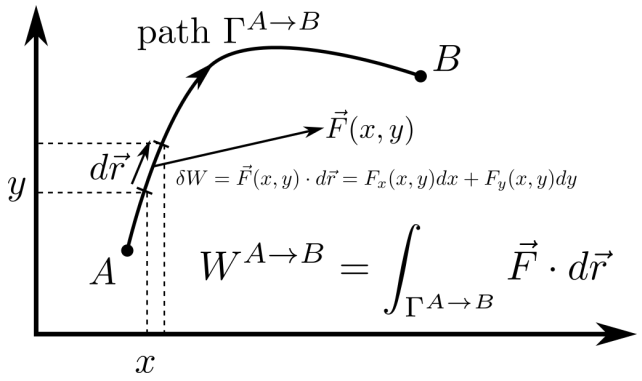
$$\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x}$$

  - In other words, if the condition above is satisfied, then there exists a function $U(x,y)$ such that

$$A(x,y) = \frac{\partial U(x,y)}{\partial x}, \qquad B(x,y) = \frac{\partial U(x,y)}{\partial y}.$$

  - Then the statement that the form δG is a differential is a very strong statement, as it tells you that in order to know two functions $A(x,y)$ and $B(x,y)$ you need to know only one function $U(x,y)$.
- Examples.
  - $\delta G = xdy + ydx$ is a differential $U = xy$.
  - $\delta G = xdy - ydx$ is not a differential. The function U does not exist.

13.2. Work.



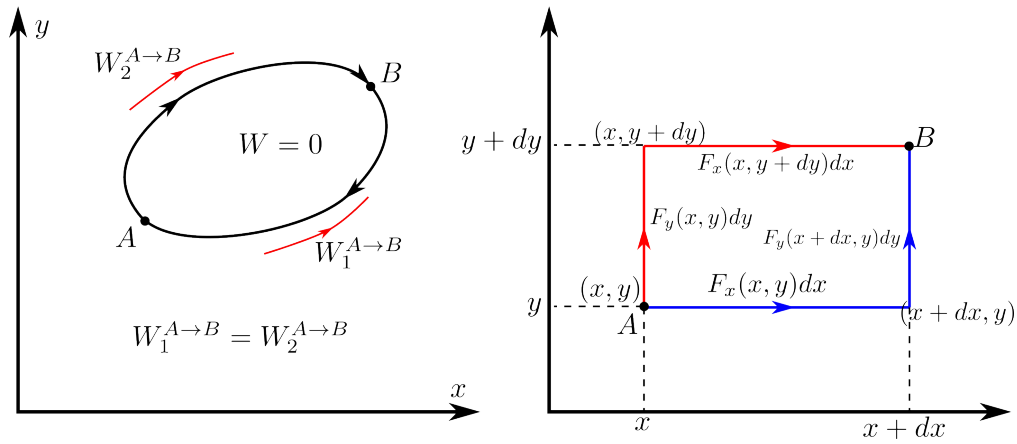
Suppose we have a force field: $\vec{F}(x,y)$ — it means that if we place a particle in a point (x,y) this particle will experience force $\vec{F}(x,y)$.

- A work done by a force: $\delta W = \vec{F} \cdot d\vec{r}$.
- Notice, that although $\delta W = F_x dx + F_y dy + F_z dz$ this is not necessarily a full differential.
- Superposition. If there are many forces, the total work is the sum of the works done by each.
- Finite displacement. Line integral.

$$W^{A\to B} = \int_{\Gamma^{A\to B}} \vec{F} \cdot d\vec{r}$$

- This formula tells us that we need to know the path $\Gamma^{A\to B}$, we split the path into infinitesimal displacements $d\vec{r}$, on each displacement we compute work $\delta W = \vec{F} \cdot d\vec{r}$, then we sum up all δWs for all displacements.
- IMPORTANT: In general case work $W^{A\to B}$ depends on path $\Gamma^{A\to B}$!!!!!

13.3. Conservative forces. Potential energy.



- Fundamental forces. Depend on coordinates, do not depend on time.
- Work done by the forces over ANY closed loop is zero!!!!! One can not get work from nothing.
- It means that work is independent of the path! see left panel of the figure.
- Look at the right panel. Consider two paths from point (x, y) to point $(x+dx, y+dy)$: first dx, then dy (blue path); first dy then dx (red path)

$$\delta W_1^{A\to B} = F_x(x, y)dx + F_y(x + dx, y)dy = F_x(x, y)dx + F_y(x, y)dy + \frac{\partial F_y}{\partial x} dydx$$
$$\delta W_2^{A\to B} = F_y(x, y)dy + F_x(x, y + dy)dx = F_y(x, y)dy + F_x(x, y)dx + \frac{\partial F_x}{\partial y} dydx.$$

where we used $F_y(x + dx, y) \approx F_y(x, y) + \frac{\partial F_y}{\partial x} dx$, and $F_x(x, y + dy) \approx F_x(x, y)dx + \frac{\partial F_x}{\partial y} dy$.

- The works must be equal to each other, $\delta W_1^{A\to B} = \delta W_2^{A\to B}$, so we must have

$$\left.\frac{\partial F_y}{\partial x}\right|_{x,y} = \left.\frac{\partial F_x}{\partial y}\right|_{x,y}$$

- So small work done by a conservative force:

$$\delta W = F_x dx + F_y dy, \qquad \frac{\partial F_y}{\partial x} = \frac{\partial F_x}{\partial y}$$

  is a full differential!
- In other words, FOR A CONSERVATIVE FORCE there exist a function U such that

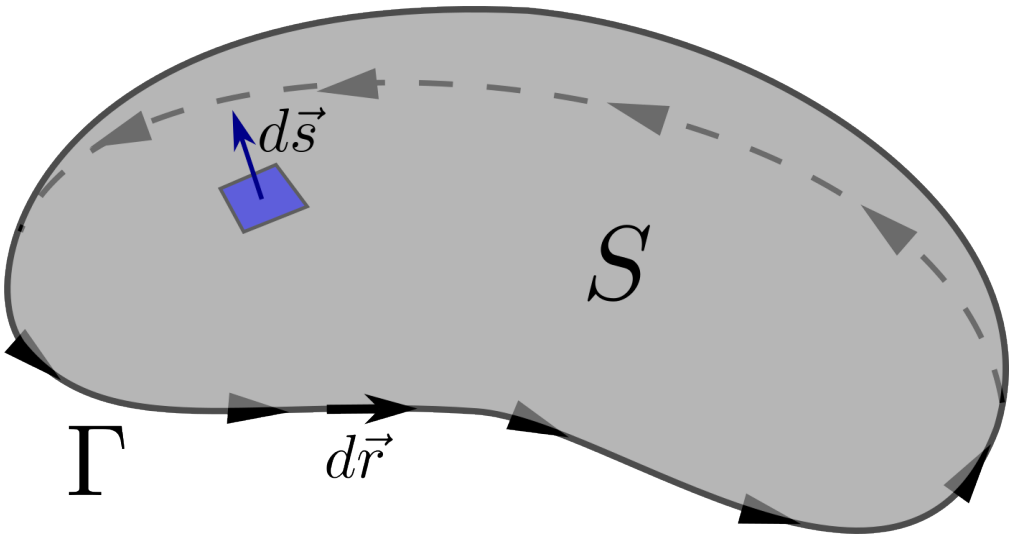
$$\delta W = -dU$$

  (the minus sign is for further convenience)
- It means that there is such a function of the coordinates $U(x, y)$, that

$$F_x = -\frac{\partial U}{\partial x}, \quad F_y = -\frac{\partial U}{\partial y}, \quad \text{or} \quad \vec{F} = -\text{grad}U \equiv -\vec{\nabla}U.$$

- This function U is called "potential energy function", or just "potential energy".

13.3.1. Stokes theorem.



One may notice, that what is done above is a part of a proof of Stokes theorem. According to this theorem the circulation of a vector field $\vec{F}$ over a closed path Γ, $\oint_\Gamma \vec{F} \cdot d\vec{r}$ equals to the flux of the field $\nabla \times \vec{F}$ through any surface S bounded by Γ, $\int_S \nabla \times \vec{F} \cdot d\vec{s}$

$$\oint_\Gamma \vec{F} \cdot d\vec{r} = \int_S \nabla \times \vec{F} \cdot d\vec{s}.$$

(the orientations of $d\vec{r}$ and $d\vec{s}$ must correspond to each other according to the right hand rule.)

Notice, that if $\vec{F}$ is a force field then what is on the left hand side of the above equation is work done by the force on the closed path Γ. If we demand that this work is zero for ANY path, then we must have $\nabla \times \vec{F} = 0$ everywhere. But this is exactly the necessary and sufficient condition for the existence of a function U such that

$$\vec{F} = -\nabla U.$$

And that work of the force $\vec{F}$ on an infinitesimal displacement $d\vec{r}$ is a full differential

$$\delta W = \vec{F} \cdot d\vec{r} = -\nabla U \cdot d\vec{r} = -\frac{\partial U}{\partial x} dx - \frac{\partial U}{\partial y} dy - \frac{\partial U}{\partial z} dz = -dU.$$

LECTURE 14
Energy Conservation. One-dimensional motion.

- Last lecture we found, that there exists a special class of forces (which depend only on coordinates) which are called "conservative forces".
  - Not all forces are conservative! Friction!
  - All fundamental forces are conservative.
- A conservative force is such a force that its work around any closed loop is zero.
- Last lecture we found that for a conservative (zero work on a closed loop) force there exists a function U — called "potential energy" such that

$$F_x = -\frac{\partial U}{\partial x}, \quad F_y = -\frac{\partial U}{\partial y}, \quad \text{or} \quad \vec{F} = -\text{grad}U \equiv -\vec{\nabla}U.$$

  Such function is not unique as one can always add an arbitrary constant to the potential energy.
- Under a small displacement $d\vec{r}$ a work done by such a force is

$$\delta W = \vec{F} \cdot d\vec{r} = F_x dx + F_y dy + F_z dz = -dU.$$

- If the force $\vec{F}(\vec{r})$ is known, then there is a test for if the force is conservative.

$$\nabla \times \vec{F} = 0.$$

14.1. Change of kinetic energy.

- If a body of mass m moves under the force $\vec{F}$, then.

$$m\frac{d\vec{v}}{dt} = \vec{F}, \qquad md\vec{v} = \vec{F}dt, \qquad m\vec{v} \cdot d\vec{v} = \vec{F} \cdot \vec{v}dt = \vec{F} \cdot d\vec{r} = \delta W.$$

  So we have

$$d\frac{mv^2}{2} = \delta W$$

- The change of kinetic energy $K = \frac{mv^2}{2}$ equals the total work done by all forces.
- In general case this is not very useful, as we need to know the path $\Gamma^{A\to B}$ from the initial point A to the final point B in order to compute work.

$$W = \int_{\Gamma^{A\to B}} \vec{F} \cdot d\vec{r}.$$

  In order to know the path $\Gamma^{A\to B}$ we need to solve the equations of motion.

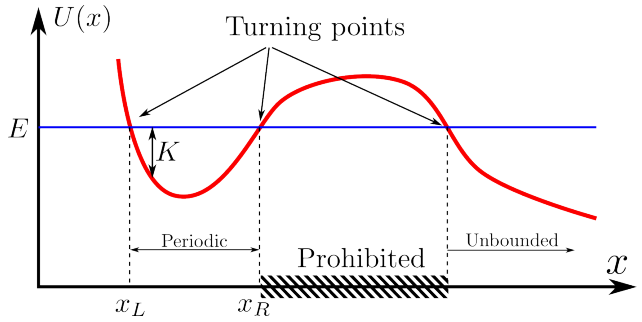
14.1.1. Conservative forces.

For the conservative forces the situation simplifies considerably.

- On a trajectory we have $dK = \delta W = -dU$, or
$$d\left(\frac{mv^2}{2} + U\right) = 0, \qquad K + U = \text{const}.$$
- What this means is that the energy $K + U$ is constant on the trajectory. While the particles moves its velocity, and hence the kinetic energy, changes, its position, and hence the potential energy, changes, but the sum of kinetic energy and the potential energy stays constant!
- As the total energy $E = K + U$ is constant it is the same at any point of time. In particular at any point of time it is the same as at the very beginning. As we know the initial conditions $\vec{r}_0$ and the initial velocity $\vec{v}_0$, we can compute this total energy $E = \frac{m\vec{v}_0^2}{2} + U(\vec{r}_0)$.
- Potential energy is defined up to a constant. In particular, one can always choose a point of space and set the potential energy at this point to be zero.
- Examples.

14.2. $1D$ motion.

For the motion in $1D$ things become even simpler, as we always know the trajectory — it is $1D$!



- In $1D$ the force that depends only on the coordinate is always conservative.
- In $1D$ in the case when the force depends only on coordinates the equation of motion can be solved in quadratures.
- The number of conservation laws is enough to solve the equations.
- If the force depends on the coordinate only $F(x)$, then there exists a function — potential energy — with the following property
$$F(x) = -\frac{\partial U}{\partial x}, \qquad U(x) = -\int_{\text{arb. point.}}^{x} F(x')dx'.$$
Such a function is not unique as one can always add an arbitrary constant to the potential energy, which means to choose the lower limit of integration to be an arbitrary point.

We want to solve the following problem: A particle of mass m can move in $1D$. There is force $F(x)$ which acts on the particle when the particle is at point x. The particle starts moving at time $t = t_0$ from the initial position x_0 with initial velocity v_0. We need to find the function $x(t)$.

- The total energy is then conserved

$$K + U = \text{const.}, \qquad \frac{m\dot{x}^2}{2} + U(x) = E, \qquad \frac{m\dot{x}^2}{2} = E - U(x).$$

- Energy E can be calculated from the initial conditions: $E = \frac{mv_0^2}{2} + U(x_0)$.
- Let's plot the function $U(x)$ and draw a horizontal line E on the same plot, see figure.
- As $\frac{mv^2}{2} > 0$ the allowed areas where the particle can be are given by $E - U(x) > 0$.
- Picture. Turning points — the solutions of the equation $E = U(x)$. Prohibited regions.
- Notice, that the equation of motion depends only on the difference $E - U(x) = \frac{mv_0^2}{2} + U(x_0) - U(x)$ of the potential energies in different points, so the zero of the potential energy (the arbitrary constant that was added to the function) does not play a role.
- From $\frac{m}{2}\left(\frac{dx}{dt}\right)^2 = E - U(x)$ we find

$$\frac{dx}{dt} = \pm\sqrt{\frac{2}{m}}\sqrt{E - U(x)}$$

- Energy conservation law cannot tell the direction of the velocity, as the kinetic energy depends only on absolute value of the velocity. In $1D$ it cannot tell which sign to use "+" or "−". You must not forget to figure it out by other means.
- We then can solve the equation

$$\pm\sqrt{\frac{m}{2}}\frac{dx}{\sqrt{E - U(x)}} = dt, \qquad t - t_0 = \pm\sqrt{\frac{m}{2}}\int_{x_0}^{x(t)}\frac{dx'}{\sqrt{E - U(x')}}$$

  Notice that
  - $x(t_0) = x_0$ initial condition is automatically satisfied.
  - the initial conditions enter the equation in two places: in the lower limit of integration and into the value of E.

  As we know E from the initial conditions and we know $U(x)$ we can take the integral and obtain the function $x(t)$ implicitly (we obtain $t(x)$ explicitly). To get the explicit function $x(t)$ we then need to solve an algebraic equation, or simply plot it on a graph.
- Examples:
  - Motion under a constant force: $U(x) = -mgx$.
  - Oscillator: $U(x) = \frac{kx^2}{2}$.
  - Pendulum: $U(x) = lgm(1 - \cos\phi)$.
- Only two types of motion are possible in $1D$ (in potential field): periodic and unbounded.
- Periodic motion. Period between two turning points x_L and x_R (period is time which it take the particle to go from the point x_L to the point x_R and back to the point x_L.)

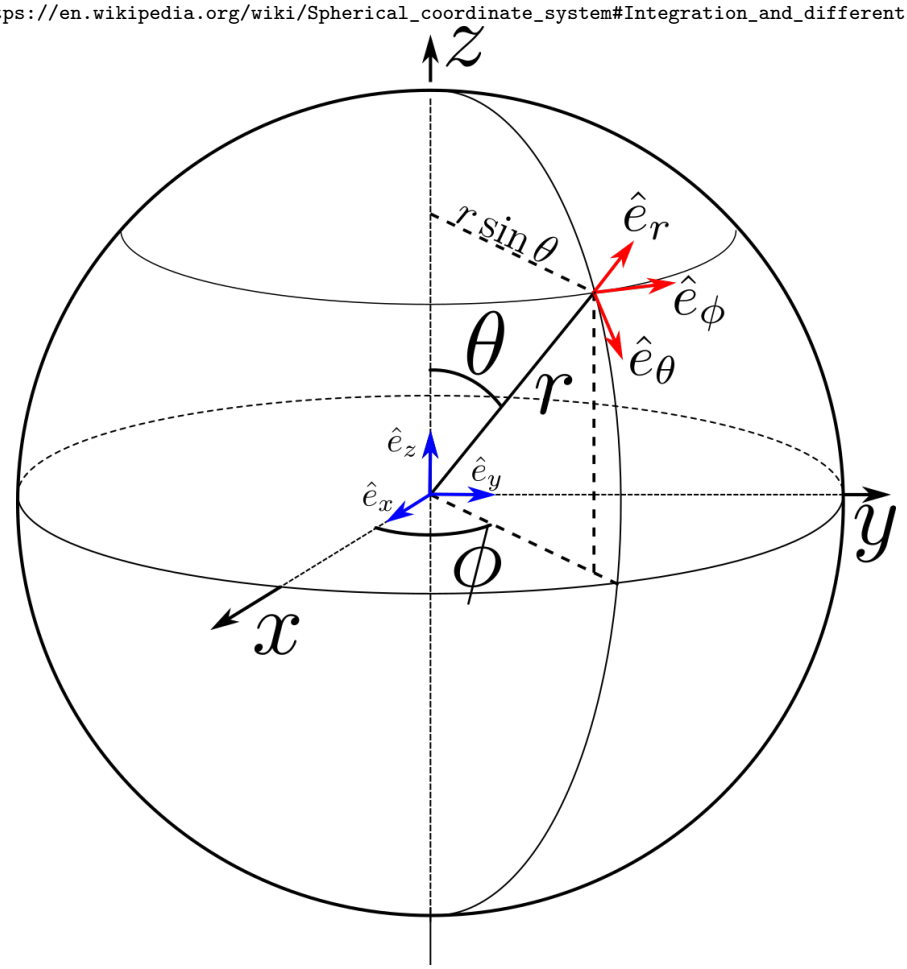
$$T = 2\sqrt{\frac{m}{2}}\int_{x_L}^{x_R}\frac{dx'}{\sqrt{E - U(x')}}, \qquad U(x_{L,R}) = E.$$

  Notice, that the dependence of the period on the energy $T(E)$ comes from two places: There is E in the integral explicitly, there is also E in the upper and lower limits, as they are the solutions of the equation $U(x_{L,R}) = E$.

LECTURE 15
Spherical coordinates.

15.1. Spherical coordinates.

https://en.wikipedia.org/wiki/Spherical_coordinate_system#Integration_and_differentiation_in_spherical_coordinates



The spherical and Cartesian coordinates are related by

$$\begin{aligned} &x = r\sin\theta\cos\phi \\ &y = r\sin\theta\sin\phi \\ &z = r\cos\theta \\ &r\in[0,\infty), \quad \theta\in[0,\pi], \quad \phi\in[0,2\pi) \end{aligned} \quad .$$

These are just the coordinates. The three numbers (r,θ,ϕ) give you any point in space. But this is not enough, as we need to describe vectors and for that we need unit coordinate vectors!

15.1.1. Coordinate vectors of spherical coordinates.

The coordinates r, θ, and ϕ can be used to specify any point.
There are corresponding unit vectors $\hat{e}_r$, $\hat{e}_\theta$, and $\hat{e}_\phi$ at EACH point (r,θ,ϕ).

The unit vectors $\hat{e}_r$, $\hat{e}_\theta$, $\hat{e}_\phi$ show the DIRECTION of where the point of space (r,θ,ϕ) shifts when we infinitesimally change each coordinate:

- The vector $\hat{e}_r$ is the unit vector along the direction where our point shifts if we infinitesimally change the coordinate r, while keeping θ and ϕ constant.
- The vector $\hat{e}_\theta$ is the unit vector along the direction where our point shifts if we infinitesimally change the coordinate θ, while keeping r and ϕ constant.
- The vector $\hat{e}_\phi$ is the unit vector along the direction where our point shifts if we infinitesimally change the coordinate ϕ, while keeping θ and r constant.

With this definitions we can now express ANY vector $\vec{a}$ at a point P of space with spherical coordinates (r,θ,ϕ) through its components a_r, a_θ, and a_ϕ

$$\vec{a} = a_r\hat{e}_r + a_\theta\hat{e}_\theta + a_\phi\hat{e}_\phi$$

where $\hat{e}_r$, $\hat{e}_\theta$, and $\hat{e}_\phi$ are coordinate vectors AT THE SAME POINT P.

15.1.2. Vector $d\vec{r}$ in spherical coordinates.

In Cartesian coordinates we have coordinate vectors $\hat{e}_x$, $\hat{e}_y$, $\hat{e}_z$ defined the same at every point. If we shift an point-like object from a point of space (x,y,z) to a point $(x+dx, y+dy, z+dz)$, then this shift is represented by a displacement vector $d\vec{r}$ and we can write

$$d\vec{r} = \hat{e}_x dx + \hat{e}_y dy + \hat{e}_z dz.$$

We want to be able to do the same in spherical coordinates.

Namely, if a point-like particle shifts from a point of space (r,θ,ϕ) to a point $(r+dr, \theta+d\theta, \phi+d\phi)$. How do we represent the corresponding displacement vector $d\vec{r}$ through coordinate vectors $\hat{e}_r$, $\hat{e}_\theta$, $\hat{e}_\phi$?

With the definitions of $\hat{e}_r$, $\hat{e}_\theta$, and $\hat{e}_\phi$ we know the DIRECTIONS of the displacement when we change the coordinates. In order to express the vector $d\vec{r}$ we also need to know the magnitudes of each displacement.

- If we change only coordinate r to $r+dr$, then the position vector $\vec{r}$ changes by $d\vec{r} = \hat{e}_r dr$, or $(d\vec{r})_r = dr$.
- If we change only coordinate θ to $\theta+d\theta$, then the position vector $\vec{r}$ changes by $d\vec{r} = \hat{e}_\theta r d\theta$, or $(d\vec{r})_\theta = r d\theta$.
- If we change only coordinate ϕ to $\phi+d\phi$, then the position vector $\vec{r}$ changes by $d\vec{r} = \hat{e}_\phi r\sin\theta d\phi$, or $(d\vec{r})_\phi = r\sin\theta d\phi$.

The vector $d\vec{r}$ then is expressed through the dr, $d\theta$ and $d\phi$ as

$$d\vec{r} = \hat{e}_r dr + \hat{e}_\theta r d\theta + \hat{e}_\phi r \sin\theta d\phi.$$

Notice, that in this formulation we do not need to have the global position vector $\vec{r}$. We can do everything with the coordinate vectors defined locally.

15.1.3. Connecting Spherical and Cartesian coordinate vectors.

Here I show how to connect $\hat{e}_x$, $\hat{e}_y$, $\hat{e}_z$, to $\hat{e}_r$, $\hat{e}_\theta$, $\hat{e}_\phi$ using only local relations.

- Using the definition of the spherical coordinates we have locally

$$\begin{aligned} dx &= dr\sin\theta\cos\phi + d\theta r\cos\theta\cos\phi - d\phi r\sin\theta\sin\phi \\ dy &= dr\sin\theta\sin\phi + d\theta r\cos\theta\sin\phi + d\phi r\sin\theta\cos\phi \\ dz &= dr\cos\theta - d\theta r\sin\theta \end{aligned}$$

- Using these expressions in $d\vec{r}$ is Cartesian coordinates $d\vec{r} = \hat{e}_x dx + \hat{e}_y dy + \hat{e}_z dz$ and collecting all the terms with dr, $d\phi$ and $d\theta$ we find

$$d\vec{r} = \left(\hat{e}_x\sin\theta\cos\phi + \hat{e}_y\sin\theta\sin\phi + \hat{e}_z\cos\theta\right)dr + \left(\hat{e}_x\cos\theta\cos\phi + \hat{e}_y\cos\theta\sin\phi - \hat{e}_z r\sin\theta\right)rd\theta$$
$$+ \left(-\hat{e}_x\sin\phi + \hat{e}_y\cos\phi\right)r\sin\theta d\phi$$

- Comparing this to the $d\vec{r}$ in spherical coordinates $d\vec{r} = \hat{e}_r dr + \hat{e}_\theta r d\theta + \hat{e}_\phi r\sin\theta d\phi$ we get

$$\begin{aligned} \hat{e}_r &= \hat{e}_x\sin\theta\cos\phi + \hat{e}_y\sin\theta\sin\phi + \hat{e}_z\cos\theta \\ \hat{e}_\theta &= \hat{e}_x\cos\theta\cos\phi + \hat{e}_y\cos\theta\sin\phi - \hat{e}_z\sin\theta \\ \hat{e}_\phi &= -\hat{e}_x\sin\phi + \hat{e}_y\cos\phi \end{aligned}$$

15.1.4. Coordinate independent definition of the gradient.

We will need to deal with the potential energy in $3D$ space. We will also need to deal with the force $\vec{F} = -\vec{\nabla}U$ in $3D$. Let's discuss these objects in some detail.

- Let's denote U a scalar function in our space. This means that in every point P of space we have a number U which is different in different points, but it changes smoothly.
- So we think of U as a map from all points of space to the space of numbers.
- Notice, that only this map matters, as a particle moving in the potential energy U knows nothing about the coordinates, it "knows" only about the point of space where it is at.
- If we use Cartesian coordinates, then a point P will have coordinates (x, y, z), and the function can be represented as $U(x, y, z)$.
- If we use spherical coordinates, then a point P will have coordinates (r, θ, ϕ), and the function can be represented as $U(r, \theta, \phi)$. But it is the same map.
- The force is $-\vec{\nabla}U$. This force is a vector field. At each point of space there is a vector $-\vec{\nabla}U$. Again the particles moving in the space knows nothing about the coordinates, but "knows" about the vector $-\nabla U$ at the point of space where it is at.
- Imagine now that we work in the spherical coordinates, and we want to find the components of a vector $\vec{\nabla}U$ in the spherical coordinates.

- In order to do that we need to define the gradient vector $\vec{\nabla}U$ in a coordinate independent way. (Remember, it is a vector, vectors are independent of coordinates!)
- Consider a function U as a function of Cartesian coordinates: $U(x,y,z)$. Then

$$dU = \frac{\partial U}{\partial x}dx + \frac{\partial U}{\partial y}dy + \frac{\partial U}{\partial z}dz = \vec{\nabla}U \cdot d\vec{r}.$$

  Notice, that we have a *coordinate independent* definition of the vector gradient. The vector of gradient $\vec{\nabla}U$ is such a vector that for *ANY* vector $d\vec{r}$ we have:

$$dU = \vec{\nabla}U \cdot d\vec{r} \quad \text{— definition of } \vec{\nabla}U.$$

  It is coordinate independent as it is a scalar/dot product which does not depend on coordinates.
- I want to make a few points about this definition.
  - This definition is constructive – it allows one to find the vector of gradient in any system of coordinates. For this it is important that $d\vec{r}$ is an *arbitrary* infinitesimal vector.
  - It connects calculus dU with geometry — the scalar product of two vectors.
  - It thus gives the geometrical meaning/picture to calculus. In particular one can see that if one chooses a vector $d\vec{r}_\perp$ which is perpendicular to the vector of the gradient at some particular point, then the function U will not change along the direction of $d\vec{r}_\perp$ (in the infinitesimal neighborhood of that point).
- Let's see how this definition works in Cartesian coordinates.
- In particular, we have our (smooth) function U, which gives a number at each point of space P, and we use Cartesian coordinates to denote the position of all points P in space.
- At each point P of space we have a vector of gradient $\vec{\nabla}U|_P$.
- This vector $\vec{\nabla}U$ at some point P of space has Cartesian components $(\vec{\nabla}U)_x$, $(\vec{\nabla}U)_y$, and $(\vec{\nabla}U)_z$.
- Using our definition of the gradient

$$dU = \vec{\nabla}U \cdot d\vec{r}$$

  we want to find those components $(\vec{\nabla}U)_x$, $(\vec{\nabla}U)_y$, and $(\vec{\nabla}U)_z$.
  - First we write the vector of gradient as

$$\vec{\nabla}U = (\vec{\nabla}U)_x\hat{e}_x + (\vec{\nabla}U)_y\hat{e}_y + (\vec{\nabla}U)_z\hat{e}_z,$$

    the components $(\vec{\nabla}U)_x$, $(\vec{\nabla}U)_y$, and $(\vec{\nabla}U)_z$ are the components which we want to find.
  - Using the vector $d\vec{r}$ in Cartesian coordinates $d\vec{r} = \hat{e}_x dx + \hat{e}_y dy + \hat{e}_z dz$ we find

$$dU = \vec{\nabla}U \cdot d\vec{r} = (\vec{\nabla}U)_x dx + (\vec{\nabla}U)_y dy + (\vec{\nabla}U)_z dz$$

  - Consider a function U as the function of Cartesian coordinates, so at the same point P which has Cartesian coordinates (x,y,z) the function is given by $U(x,y,z)$, we know from the standard calculus

$$dU = \frac{\partial U}{\partial x}dx + \frac{\partial U}{\partial y}dy + \frac{\partial U}{\partial z}dz$$

– Comparing these to results for dU (both are valid for arbitrary infinitesimal dx, dy, and dz) we find

$$(\vec{\nabla}U)_x = \frac{\partial U}{\partial x}, \quad (\vec{\nabla}U)_y = \frac{\partial U}{\partial y}, \quad (\vec{\nabla}U)_z = \frac{\partial U}{\partial z}.$$

– This is our standard formulas for the gradient in Cartesian coordinates.

- Now we can use this procedure for any other system of coordinates, as long as we know how to express $d\vec{r}$ in the corresponding coordinate vectors.

15.1.5. Gradient in spherical coordinates.

We want to find how to write the gradient vector $\vec{\nabla}U$ in spherical coordinates.

- We still have a (smooth) function U which gives a number at each point P of space.
- The gradient is a vector $\vec{\nabla}U$ at each point P of space which is defined as $dU = \vec{\nabla}U \cdot d\vec{r}$.
- We now use the spherical coordinates to denote each point of space. Our point P has a coordinates (r, θ, ϕ). At each point P we also have spherical coordinate vectors $\hat{e}_r$, $\hat{e}_\theta$, $\hat{e}_\phi$. We want to find the components of the vector $\vec{\nabla}U$ at point P of space in the coordinate vectors $\hat{e}_r$, $\hat{e}_\theta$, $\hat{e}_\phi$ at THE SAME POINT P.
- As any other vector, the vector $\vec{\nabla}U$ at a point P can be written through its components $(\vec{\nabla}U)_r$, $(\vec{\nabla}U)_\theta$, and $(\vec{\nabla}U)_\phi$ at the same point P in the spherical coordinate vectors $\hat{e}_r$, $\hat{e}_\theta$, $\hat{e}_\phi$ at THE SAME POINT P.

$$\vec{\nabla}U = (\vec{\nabla}U)_r \hat{e}_r + (\vec{\nabla}U)_\theta \hat{e}_\theta + (\vec{\nabla}U)_\phi \hat{e}_\phi,$$

Again, $(\vec{\nabla}U)_r$, $(\vec{\nabla}U)_\theta$, and $(\vec{\nabla}U)_\phi$ are the components of the vector $\vec{\nabla}U$ in the spherical coordinates at a point P. It is those components that we want to find.

- By the definition of the gradient vector, and using $d\vec{r}$ in spherical coordinates

$$d\vec{r} = \hat{e}_r dr + \hat{e}_\theta r d\theta + \hat{e}_\phi r \sin\theta d\phi.$$

we get

$$dU = \vec{\nabla}U \cdot d\vec{r} = (\vec{\nabla}U)_r dr + (\vec{\nabla}U)_\theta r d\theta + (\vec{\nabla}U)_\phi r \sin\theta d\phi$$

- On the other hand if we now consider U as a function of the spherical coordinates $U(r, \theta, \phi)$, then calculus tells us

$$dU = \frac{\partial U}{\partial r} dr + \frac{\partial U}{\partial \theta} d\theta + \frac{\partial U}{\partial \phi} d\phi$$

- Comparing the two expressions for dU we find

$$(\vec{\nabla}U)_r = \frac{\partial U}{\partial r}$$
$$(\vec{\nabla}U)_\theta = \frac{1}{r}\frac{\partial U}{\partial \theta}$$
$$(\vec{\nabla}U)_\phi = \frac{1}{r\sin\theta}\frac{\partial U}{\partial \phi}$$

- The vector of gradient in spherical coordinates is then written as

$$\vec{\nabla}U = \frac{\partial U}{\partial r}\hat{e}_r + \frac{1}{r}\frac{\partial U}{\partial \theta}\hat{e}_\theta + \frac{1}{r\sin\theta}\frac{\partial U}{\partial \phi}\hat{e}_\phi$$

LECTURE 16
Central force. Effective potential. Kepler orbits.

16.1. Results of the last lecture

- The coordinate independent definition of gradient of function U is

$$dU = \vec{\nabla}U \cdot d\vec{r}, \qquad \text{for ANY } d\vec{r}.$$

- In Cartesian coordinates this definition gives

$$\vec{\nabla}U = \frac{\partial U}{\partial x}\hat{e}_x + \frac{\partial U}{\partial y}\hat{e}_y + \frac{\partial U}{\partial z}\hat{e}_z.$$

- In Spherical coordinates the same definition gives

$$\vec{\nabla}U = \frac{\partial U}{\partial r}\hat{e}_r + \frac{1}{r}\frac{\partial U}{\partial \theta}\hat{e}_\theta + \frac{1}{r\sin\theta}\frac{\partial U}{\partial \phi}\hat{e}_\phi$$

16.2. Central force. General.

- If U is the potential energy, then in spherical coordinates

$$\vec{F} = -\vec{\nabla}U = -\frac{\partial U}{\partial r}\hat{e}_r - \frac{1}{r}\frac{\partial U}{\partial \theta}\hat{e}_\theta - \frac{1}{r\sin\theta}\frac{\partial U}{\partial \phi}\hat{e}_\phi.$$

- Consider a motion of a body under central force. This can be a Coulomb force from a point like charge, or Newtonian gravity, or any other.
- Take the coordinate origin in the center of force.
- A central force then is given by (this is the definition of the central force)

$$\vec{F} = F(r)\hat{e}_r.$$

- Such a force is always conservative: $\vec{\nabla} \times \vec{F} = 0$, so there is a potential energy U such that:

$$\vec{\nabla}U = -F(r)\hat{e}_r$$

comparing this with the $\vec{\nabla}U$ in spherical coordinate, we find

$$\frac{\partial U}{\partial r} = -F(r), \qquad \frac{\partial U}{\partial \theta} = 0, \qquad \frac{\partial U}{\partial \phi} = 0,$$

so that potential energy depends only on the distance r, $U(r)$. and

$$F(r) = -\frac{\partial U}{\partial r}$$

- The torque of the central force $\tau = \vec{r} \times \vec{F} = 0$, so the angular momentum is conserved: $\vec{J} = \text{const}$.

16.3. Motion in under central force.

Consider now a particle of mass m which is moving in the central force field. The field is completely described by the potential energy function $U(r)$. We set this function such, that $U(r \to \infty) \to 0$.

In order to set up the problem we must also specify the initial conditions. So we know that at some time $t = 0$ the velocity of the particle is $\vec{v}_0$ and the position is $\vec{r}_0$.

We have two independent conservation laws: conservation of angular momentum and conservation of energy.

$$\vec{J} = \vec{r} \times m\vec{v}, \qquad E = \frac{m\vec{v}^2}{2} + U(r).$$

Both angular momentum and energy can be computed from the initial conditions:

$$\vec{J} = \vec{r}_0 \times m\vec{v}_0, \qquad E = \frac{m\vec{v}_0^2}{2} + U(r_0).$$

So from now we treat $\vec{J}$ and E as known and constant vector and number.

Let's see how we can use these conservation laws.

16.3.1. Angular momentum conservation.

- The direction of $\vec{J}$ is perpendicular to the initial momentum and initial coordinate.
- During the motion the direction of $\vec{J}$ will not change — it is conserved.
- So during the motion at any moment the momentum and position vectors will be in the same plane perpendicular to $\vec{J}$.
- The motion is all in one plane! The plane which contains the vector of the initial velocity and the initial radius vector. As it contains the initial position vector, this plane contains the center of the force.
- We take the direction of $\vec{J}$ as our z axis. The plane of motion is then $x - y$ plane.
- The angular momentum is $\vec{J} = J\vec{e}_z$, where $J = |\vec{J}| = \text{const}$.. This constant is given by initial conditions $J = m|\vec{r}_0 \times \vec{v}_0|$.
- In the $x - y$ plane the spherical coordinate $\theta = \pi/2$ we can use only r and ϕ coordinates — the polar coordinates.
- Writing the value of the angular momentum at any moment of time, and using $\vec{v} = \dot{r}\hat{e}_r + r\dot{\phi}\hat{e}_\phi$, $\vec{r} = r\hat{e}_r$ in the polar coordinates we get

$$\vec{J} = \vec{r} \times m\vec{v} = mr\hat{e}_r \times \left(\dot{r}\hat{e}_r + r\dot{\phi}\hat{e}_\phi\right) = mr^2\dot{\phi}\hat{e}_r \times \hat{e}_\phi = mr^2\dot{\phi}\hat{e}_z$$

or

$$mr^2\dot{\phi} = J, \qquad \dot{\phi} = \frac{J}{mr^2}$$

Notice: The last equation means that if we know $r(t)$ we will be able to compute $\phi(t)$!

$$\phi(t) - \phi_0 = \frac{J}{m}\int_0^t \frac{dt'}{r^2(t')}.$$

So we only need to find $r(t)$!

- The use of the conservation of the angular momentum $\vec{J}$ allowed us to simplify the problem of the motion in three dimension to a problem of finding just one function!

16.3.2. Energy conservation.

- The velocity in these polar coordinates is

$$\vec{v} = \dot{r}\vec{e}_r + r\dot{\phi}\vec{e}_\phi = \dot{r}\vec{e}_r + \frac{J}{mr}\vec{e}_\phi$$

- The kinetic energy then is

$$K = \frac{m\vec{v}^2}{2} = \frac{m\dot{r}^2}{2} + \frac{J^2}{2mr^2}$$

- The total energy then is

$$E = K + U = \frac{m\dot{r}^2}{2} + \frac{J^2}{2mr^2} + U(r).$$

- If we introduce the effective potential energy

$$U_{eff}(r) = \frac{J^2}{2mr^2} + U(r),$$

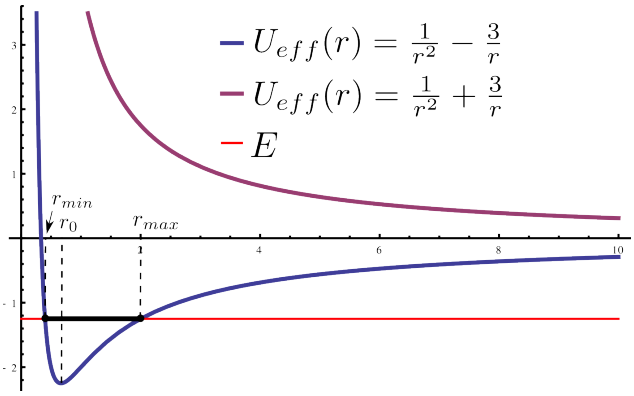
then we have

$$\frac{m\dot{r}^2}{2} + U_{eff}(r) = E, \qquad m\ddot{r} = -\frac{\partial U_{eff}}{\partial r}$$

- This is a one dimensional motion in the potential U_{eff} which was solved before! This means that we can find the exact solution of a problem of motion if ANY central force by simply using the answers we obtained previously.

Let's see how it works for a particular example of the central force — the gravity.

16.4. Kepler orbits.



Historically, the Kepler problem — the problem of motion of the bodies in the Newtonian gravitational field — is one of the most important problems in physics. It is the solution of the problems and experimental verification of the results that convinced the physics community in the power of Newton's new math and in the correctness of his mechanics. For the first time people could understand the observed motion of the celestial bodies and make accurate predictions. The whole theory turned out to be much simpler than what existed before.

- In the Kepler problem we want to consider the motion of a body of mass m in the gravitational central force due to much larger mass M.
- As $M \gg m$ we ignore the motion of the larger mass M and consider its position fixed in space (we will discuss what happens when this limit is not applicable later) For now we take the position of M as our center of the force and the coordinate origin.
- The force that acts on the mass m is given by the Newton's law of gravity:
$$\vec{F} = -\frac{GmM}{r^3}\vec{r} = -\frac{GmM}{r^2}\vec{e}_r$$
where $\vec{e}_r$ is the direction from M to m.
- The potential energy is then given by
$$U(r) = -\frac{GMm}{r}, \qquad -\frac{\partial U}{\partial r} = -\frac{GmM}{r^2}, \qquad U(r \to \infty) \to 0$$
- The effective potential is
$$U_{eff}(r) = \frac{J^2}{2mr^2} - \frac{GMm}{r},$$
where J is the angular momentum.
- For the Coulomb potential we will have the same r dependence, but for the like charges the sign in front of the last term is different — repulsion.
- In case of attraction for $J \neq 0$ the function $U_{eff}(r)$ always has a minimum for some distance r_0. It has no minimum for the repulsive interaction.
- Looking at the graph of $U_{eff}(r)$ we see, that
  - for the repulsive interaction there can be no bounded orbits. The total energy E of the body is always positive. The minimal distance the body may have with the center is given by the solution of the equation $U_{eff}(r_{min}) = E$.
  - for the attractive interaction there is a minimum of the effective potential energy at $r = r_0$ which is given by the equation
$$\left.\frac{\partial U_{eff}}{\partial r}\right|_{r=r_0} = 0, \qquad r_0 = \frac{J^2}{Gm^2M}.$$
and $U(r_0) < 0$, where $U(r \to \infty) \to 0$. Then, from the graph $U(r)$ we see
    * if $E > 0$, then the motion is not bounded. The minimal distance the body may have with the center is given by the solution of the equation $U_{eff}(r_{min}) = E$.
    * if $U_{eff}(r_0) < E < 0$, then the motion is bounded between the two real solutions of the equation $U_{eff}(r) = E$ gives both r_{min} and r_{max}. One of the solution is larger than r_0, the other is smaller.
    * if $U_{eff}(r_0) = E$, then the only solution is $r = r_0$. So the motion is around the circle with fixed radius r_0. For such motion we must have
$$ma = \frac{mv^2}{r_0} = \frac{GmM}{r_0^2}, \qquad \frac{J^2}{mr_0^3} = \frac{GmM}{r_0^2}, \qquad r_0 = \frac{J^2}{Gm^2M}.$$
Notice, that this is exactly r_0 that we found before. Also
$$U_{eff}(r_0) = E = \frac{mv^2}{2} - \frac{GmM}{r_0} = -\frac{1}{2}\frac{GmM}{r_0}.$$

LECTURE 17
Kepler orbits continued.

Kepler orbits are orbits of the planets and other bodies in the sun's gravitation field. The mass of the sun is much larger, than the mass of any other body in the solar system. So we ignore the motion of the sun (the gravitation force which acts on the body will also act on the sun, by Newton's third law).

If M is the mass of the sun, and m is the mass of the body, then the Newton gravitation force is

$$\vec{F} = -\frac{GMm}{r^2} e_r.$$

It is a central force. The corresponding potential energy $U(r)$ with the condition $U(r \to \infty) \to 0$ is

$$U(r) = -\frac{GMm}{r}.$$

It is an attractive force. The effective potential energy is

$$U_{eff}(r) = \frac{J^2}{2mr^2} + U(r) = \frac{J^2}{2mr^2} - \frac{GMm}{r}.$$

- In the motion the angular momentum and the energy are conserved

$$J = mr^2\dot{\phi}, \qquad E = \frac{m\dot{r}^2}{2} + U_{eff}(r)$$

- Accounting: 3D, total order is 6, four conserved quantities, final order is $6 - 4 = 2$.
- All motion happens in one plane.
- In that plane we describe the motion by two time dependent polar coordinates $r(t)$ and $\phi(t)$. The dynamics is given by the angular momentum conservation and the effective equation of motion for the r coordinate.
- The effective equation for the r coordinate is obtained by differentiating energy with respect to time.

$$\dot{\phi} = \frac{J}{mr^2}, \qquad m\ddot{r} = -\frac{\partial U_{eff}(r)}{\partial r},$$

  These equations must be supplied with the initial conditions — initial position $r(t = 0)$, $\phi(t = 0)$, and initial velocities $\dot{r}(t = 0)$ and $\dot{\phi}(t = 0)$. The value of the angular momentum J must be found from the initial conditions.
- This system of equations is complete. The solution will give us the functions $r(t)$ and $\phi(t)$ — the position of the body as a function of time.

- One can think of these solution as a parametric form (with t as a parameter) of a trajectory/path of the object.
- For now I am not interested in the time evolution and only want to find the trajectory (the path) of the body. This trajectory is given by the function $r(\phi)$.
- However, if we know $r(\phi)$, we can solve

$$\dot{\phi} = \frac{J}{mr^2(\phi)}, \qquad \frac{m}{J} r^2(\phi) d\phi = dt$$

and find $\phi(t)$. Then we will also have $r(t) = r(\phi(t))$. Thus one can consider finding of $r(\phi)$ as the first step in full solution.
- In order to find $r(\phi)$ I will use the trick we used before. As $r(t) = r(\phi(t))$

$$\dot{r} = \frac{dr}{dt} = \frac{d\phi}{dt}\frac{dr}{d\phi} = \frac{J}{mr^2}\frac{dr}{d\phi} = -\frac{J}{m}\frac{d(1/r)}{d\phi}, \qquad \frac{d^2r}{dt^2} = \frac{d\phi}{dt}\frac{d\dot{r}}{d\phi} = -\frac{J^2}{m^2r^2}\frac{d^2(1/r)}{d\phi^2}$$

- On the other hand

$$\frac{\partial U_{eff}}{\partial r} = -\frac{J^2}{m}(1/r)^3 + GMm(1/r)^2.$$

- Now I denote $u(\phi) = 1/r(\phi)$ and get

$$-\frac{J^2}{m}u^2\frac{d^2u}{d\phi^2} = \frac{J^2}{m}u^3 - GMmu^2$$

or, denoting $\frac{d^2u}{d\phi^2} \equiv u''$

$$u'' = -u + \frac{GMm^2}{J^2}.$$

- The general solution of this equation is

$$u = \frac{GMm^2}{J^2} + A\cos(\phi - \phi_0),$$

where A and ϕ_0 are arbitrary constants.
- We can put $\phi_0 = 0$ by redefinition of ϕ — what direction we measure ϕ from.
- Before I do that, I want to point out that this is cheating. The constants A and ϕ_0 should be obtained from the initial conditions. So unless we know how to get ϕ_0 from the initial conditions we cannot redefine our system of coordinates to measure the angle from the direction of ϕ_0. However, we know that such redefinition exists. We will discuss the issue of finding ϕ_0 from the initial conditions later and now we just go ahead and redefine ϕ.
- So by setting $\phi_0 = 0$ we have

$$\frac{1}{r} = \gamma + A\cos\phi, \qquad \gamma = \frac{GMm^2}{J^2}$$

This is the equation of the trajectory/path $r(\phi)$. Both A and γ should be found from the initial conditions.
- If $\gamma = 0$ this is the equation of a straight line in the polar coordinates. Indeed $\gamma = 0$ means that one of the masses is zero, or the angular momentum is huge, then the gravitation force has a negligible effect on the motion of the body.

- A more conventional way to write the trajectory is

$$\frac{1}{r} = \frac{1}{c}\left(1 + \epsilon \cos\phi\right), \qquad c = \frac{J^2}{GMm^2} = \frac{1}{\gamma}$$

  where $\epsilon > 0$ is dimentionless number. It defines the "shape" of the trajectory.
- c has units of length and defines the overall size of the trajectory.
- c is easily found from the initial conditions

$$c = \frac{J^2}{GMm^2}.$$

  We also need a way to find ϵ from the initial conditions (and ϕ_0, but we postpone this question.)
- In order to find ϵ we compute the minimal distance of the planet to the Sun r_{min} in two different ways.
  - From the equation of the trajectory, we see, the r is at minimum when $\cos\phi$ is at maximum. So

$$r_{min} = \frac{c}{1+\epsilon}.$$

  - On the other hand when r is at minimum $\dot{r} = 0$ and from the energy conservation $E = \frac{m\dot{r}^2}{2} + U_{eff}(r)$ we find that

$$E = U_{eff}(r_{min}) = U_{eff}\left(\frac{c}{1+\epsilon}\right).$$

    Using our function U_{eff} and $J^2 = cGMm^2$ we find

$$E = -\frac{1-\epsilon^2}{2}\frac{GMm}{c}$$

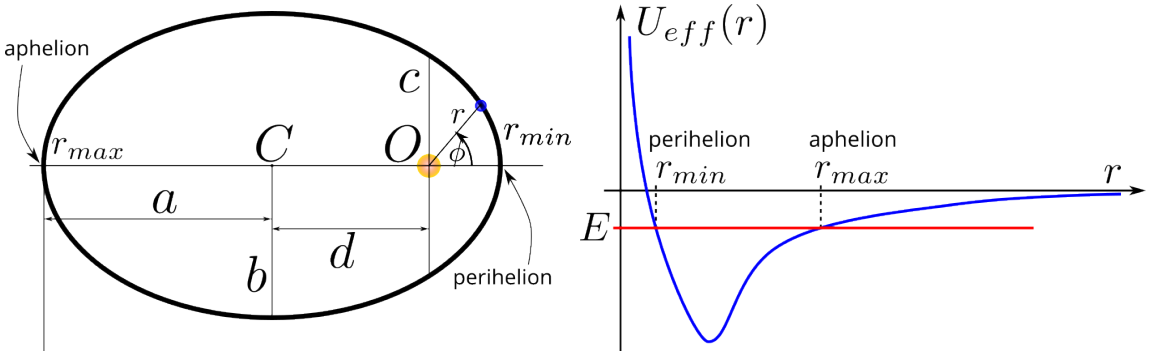
    which allows us to find ϵ from the initial conditions.
- We see that, that there are three different cases, which need to be considered separately.
  - $0 < \epsilon < 1$.
  - $\epsilon > 1$.
  - $\epsilon = 1$.
  - There is another case of $\epsilon \to \infty$. The only way to make this limit meaningful is to also take $c \to \infty$ in such a way as to have ϵ/c is finite. In this case $\frac{1}{r} = \frac{\epsilon}{c}\cos\phi$ — the trajectory is the straight line. $c \to \infty$ means $J \to \infty$. So the planet is either moving too far, or moving too fast.
  - In case $\epsilon = 0$, the orbit is just a circle $r = c$.

17.0.1. The case of $0 < \epsilon < 1$.

In this case, $\epsilon < 1$, the equation $\frac{1}{r} = \frac{1}{c}(1 + \epsilon\cos\phi)$ describes an ellipse in polar coordinates. ϵ is called the eccentricity of the ellipse, it controls the "shape" of the ellipse, while c has a dimension of length and it controls the "size" of the ellipse.

- If the minimal and maximal distance to the center — the perihelion and aphelion are at $\phi = 0$ and $\phi = \pi$ respectively.

$$r_{min} = \frac{c}{1+\epsilon}, \qquad r_{max} = \frac{c}{1-\epsilon}$$



(This is the same minimum which we used before). Both r_{min} and r_{max} are the solutions of the equation $E = U_{eff}(r)$.

- From the expression for the energy through ϵ we see that $E < 0$. We know from the previous lecture that in this case the trajectory is indeed bounded. The corresponding picture of the orbit and the potential energy are shown above.
- The ellipse can also be written as

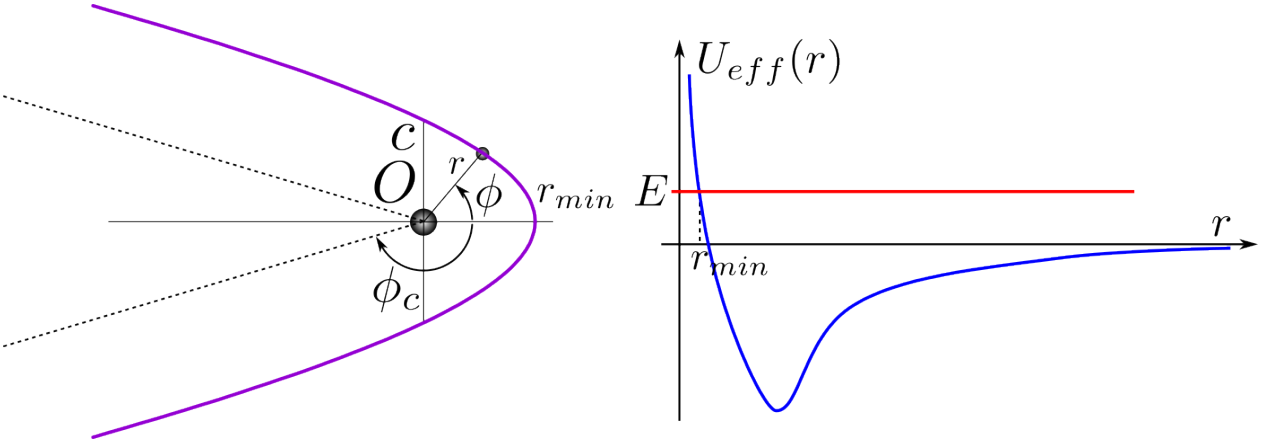
$$\frac{(x+d)^2}{a^2} + \frac{y^2}{b^2} = 1,$$

with

$$a = \frac{c}{1-\epsilon^2}, \quad b = \frac{c}{\sqrt{1-\epsilon^2}}, \quad d = a\epsilon, \quad b^2 = ac.$$

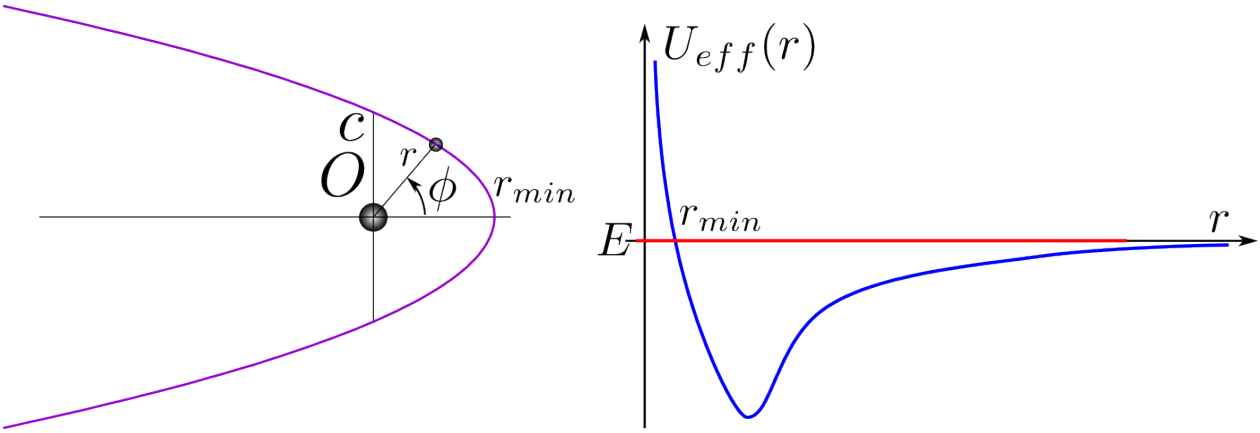
- One can check, that the position of the large mass M is one of the focuses of the ellipse — NOT ITS CENTER!
- This is the **first Kepler's law**: all planets go around the ellipses with the sun at one of the foci.

17.0.2. The case of $\epsilon > 1$.



- In this case $1 + \epsilon\cos\phi$ is zero at $\phi = \pm\phi_c$, where $\cos\phi_c = -1/\epsilon$, so $0 > \cos\phi_c > -1$.
- It means that $r(\phi \to \pm\phi_c) \to \infty$.
- So if $\epsilon > 1$, then the trajectory is unbounded.
- This also can be seen from the fact, that $E = -\frac{1-\epsilon^2}{2}\frac{GmM}{c} > 0$. As we know, for $E > 0$ the trajectory is unbounded.
- The equation $\frac{1}{r} = \frac{1}{c}(1 + \epsilon\cos\phi)$ describes a hyperbola with the sun at the focal point.

17.0.3. The case of $\epsilon = 1$.



- In this case $1 + \cos\phi$ is zero at $\phi = \pm\pi$, which is the direction straight "back".
- It means that $r(\phi \to \pm\pi) \to \infty$.
- So if $\epsilon = 1$, then the trajectory is unbounded.
- This also can be seen from the fact, that $E = -\frac{1-\epsilon^2}{2}\frac{GmM}{c} = 0$. As we know, for $E = 0$ the trajectory is unbounded.
- The equation $\frac{1}{r} = \frac{1}{c}(1 + \cos\phi)$ describes a parabola with the sun at the focal point.

LECTURE 18
Another derivation. A hidden symmetry.

18.1. Kepler's first law

In the previous lecture we found that:

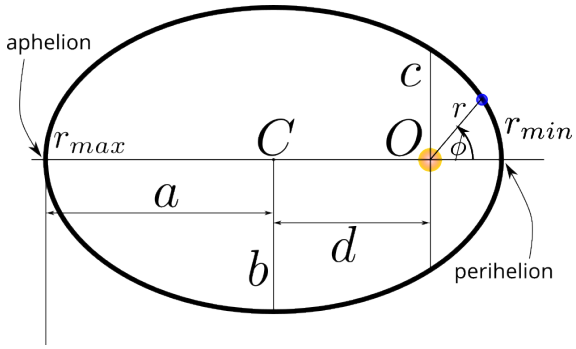
- All bodies' trajectories in the attractive potential field $U(r) \sim -1/r$ are
  - ellipses with the sun in one of the focal points if $E < 0$.
  - hyperbolas with the sun in the focal point if $E > 0$.
  - parabolas with the sun in the focal point if $E = 0$.

All these trajectories are described by

$$\frac{1}{r} = \frac{1}{c}(1 + \epsilon \cos\phi),$$

where the constants c and ϵ are computed from the initial conditions

$$c = \frac{J^2}{GMm^2}, \qquad E = -\frac{1-\epsilon^2}{2}\frac{GMm}{c}$$



For the elliptic orbit we have

$$r_{min} = \frac{c}{1+\epsilon}, \qquad r_{max} = \frac{c}{1-\epsilon}$$
$$a = \frac{c}{1-\epsilon^2}, \qquad b = \frac{c}{\sqrt{1-\epsilon^2}}$$
$$d = a\epsilon, \qquad b^2 = ac$$

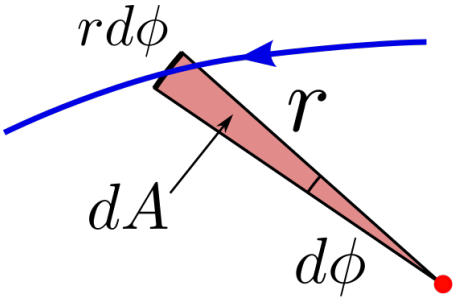
18.2. Kepler's second law

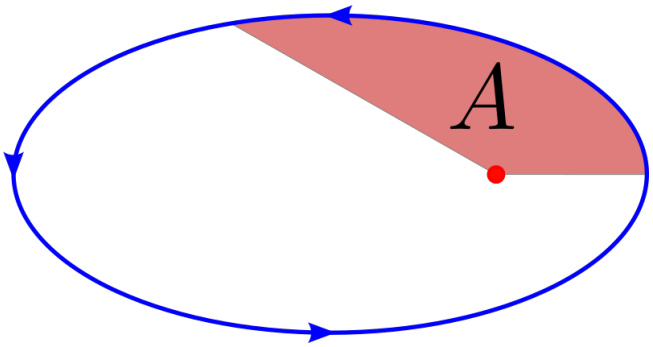
The conservation of the angular momentum reads

$$\frac{1}{2}r^2\dot{\phi} = \frac{J}{2m}.$$

We see, that in the LHS the rate at which a line from the sun to a comet or planet sweeps out area:

$$\frac{dA}{dt} = \frac{J}{2m}.$$





This rate is constant! — independent of time. So

- **Second Kepler's law:** A line joining a planet and the Sun sweeps out equal areas during equal intervals of time.

Notice, that we have used only angular momentum conservation. So this law is correct for motion in ANY central force.

18.3. Kepler's third law

Consider now the closed/elliptic orbits only. There is a period T of the rotation of a planet around the sun. We want to find this period.

The total area of an ellipse is $A = \pi ab$, so as the rate dA/dt is constant the period is

$$T = \frac{A}{dA/dt} = \frac{2\pi abm}{J},$$

Now we square the relation and use $b^2 = ac$ and $c = \frac{J^2}{GMm^2}$ to find

$$T^2 = 4\pi^2 \frac{m^2}{J^2} a^3 c = \frac{4\pi^2}{GM} a^3$$

Notice, that the mass of the planet and its angular momentum canceled out! so

- **Third Kepler's law:** For all bodies orbiting the sun the ratio of the square of the period to the cube of the semimajor axis is the same.

This is one way to measure the mass of the sun. For all planets one plots the cube of the semimajor axes as y and the square of the period as x. One then draws a straight line through all points. The slope of that line is $GM/4\pi^2$.

18.4. Another way.

Here we consider a motion in an ARBITRARY central potential $U(r)$ (central force).

- We start from the angular momentum and energy conservation laws:

$$\dot{\phi} = \frac{J}{mr^2(t)}, \qquad \frac{m\dot{r}^2}{2} + U_{eff}(r) = E, \qquad U_{eff}(r) = \frac{J^2}{2mr^2} + U(r).$$

  We know E and J from the initial conditions.
- We want to find $r(t)$ and $\phi(t)$.

- We express $\dot{r}$ from the second equation

$$\dot{r} = \sqrt{\frac{2}{m}}\sqrt{E - U_{eff}(r)}.$$

- From this equation we can find $r(t)$ the same way we solved every $1D$ problem in Lecture 14

$$\sqrt{\frac{m}{2}}\int_{r_{\text{init}}}^{r(t)} \frac{dr}{\sqrt{E - U_{eff}(r)}} = t.$$

  This will give us the function $r(t)$.
- Then we use $\dot{\phi} = \frac{J}{mr^2(t)}$ to find $\phi(t)$

$$\phi(t) - \phi_{\text{init}} = \frac{J}{m}\int_0^t \frac{dt}{r^2(t)}.$$

  This gives us the function $\phi(t)$.
- To do that we need E, $\vec{J}$, ϕ_{init}, and r_{init} (total 6) which are given by the initial conditions.
- We can also find the trajectory $r(\phi)$ directly. We divide $\dot{r} = \sqrt{\frac{2}{m}}\sqrt{E - U_{eff}(r)}$ by $\dot{\phi}$ from. We then find

$$\frac{dr}{d\phi} = \frac{dr/dt}{d\phi/dt} = \frac{\dot{r}}{\dot{\phi}} = r^2\sqrt{\frac{2m}{J^2}}\sqrt{E - U_{eff}(r)},$$

  or

$$\frac{J}{\sqrt{2m}}\frac{dr}{r^2\sqrt{E - U_{eff}(r)}} = d\phi, \qquad \frac{J}{\sqrt{2m}}\int_{r_{\text{init}}}^{r(\phi)} \frac{dr'}{r'^2\sqrt{E - U_{eff}(r')}} = \phi - \phi_{\text{init}},$$

  where E, $\vec{J}$, ϕ_{init}, and r_{init} (total 6) are given by initial conditions.
- These formulas give the trajectory for any central potential $U(r)$.
- For the potential $U(r) \sim 1/r$ the integral becomes a standard one after the substitution $x = 1/r$.

18.5. A hidden symmetry.

Let's assume, that we have some central attractive potential $U(r)$, which decays to zero at infinity.

- The problem is mapped to a one dimensional problem for the coordinate r and effective potential energy $U_{eff}(r) = \frac{J^2}{2mr^2} + U(r)$.
- For total energy $E < 0$ we have bounded motion for r between r_{min} and r_{max}.
- We can compute the time T_r for a particle to go from r_{min} to r_{max} and back

$$T_r = \sqrt{2m}\int_{r_{min}}^{r_{max}} \frac{dr}{\sqrt{E - U_{eff}(r)}}, \quad \text{where } r_{min} \text{ and } r_{max} \text{ are the solutions of } E = U_{eff}(r).$$

  Notice, that this period is a function of energy and angular momentum: $T_r(E, J)$.
- We can also compute $r(t)$, as we have done in the lecture 14.

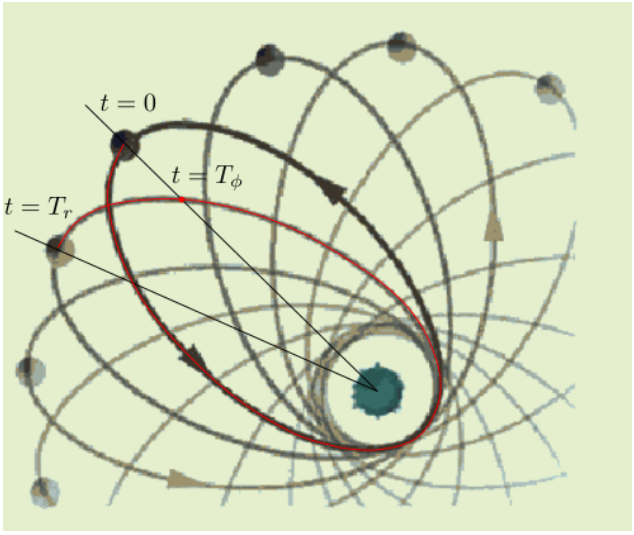
- We then can compute the time T_ϕ it takes for the angle ϕ to change by 2π

$$d\phi = \frac{J}{mr^2(t)}dt, \qquad 2\pi = \frac{J}{m}\int_0^{T_\phi}\frac{dt}{r^2(t)}.$$

  Notice, that this time will also be some function of energy and angular momentum: $T_\phi(E, J)$
- The two times $T_r(E, J)$ and $T_\phi(E, J)$ do not necessarily coincide.
- It is only for a very special potential energy function $U(r)$ that $T_r(E, J)$ and $T_\phi(E, J)$ coincide for ANY E and J!
- There are only two such functions $U(r) \sim 1/r$ and $U(r) \sim r^2$.

If $T_r \neq T_\phi$ the orbit is bounded, but not closed — this is the general situation.

It is a very special property of the gravitational (or Coulomb) potential that $T_r = T_\phi$ for ANY E and J. This symmetry requires an explanation.



If $U(r)$ is the gravitation potential energy with a small correction this discrepancy between T_r and T_ϕ is small. The orbit is almost closed, or one can say that it precesses.

LECTURE 19
Conserved Laplace-Runge-Lenz vector.

19.1. What we have learned.

- A central force field can be presented by a potential energy function $U(r)$. We choose the arbitrary constant such that $U(r \to \infty) \to 0$.
- The initial conditions $\vec{r}(t=0) = \vec{r}_{\rm init}$, $\dot{\vec{r}}(t=0) = \vec{v}_{\rm init}$.
- The vector of angular momentum is conserved and can be computed in from the initial conditions $\vec{J} = m\vec{r}_{\rm init} \times \vec{v}_{\rm init}$.
- The energy is also conserved and can be computed from the initial conditions $E = \frac{mv_{\rm init}^2}{2} + U(r_{\rm init})$.
- The motion happens in one plane perpendicular to $\vec{J}$.
- In that plane the motion is easier to describe in the polar coordinates r and ϕ. The initial conditions can be translated to:

  $$r(t=0) = r_{\rm init}, \quad \dot{r}(t=0) = \dot{r}_{\rm init}, \quad \phi(t=0) = \phi_{\rm init}, \quad \dot{\phi}(t=0) = \dot{\phi}_{\rm init}$$

  Through these initial conditions we also can compute the value/magnitude of angular momentum $J = mr_{\rm init}^2\dot{\phi}_{\rm init}$ and the total energy $E = \frac{m\dot{r}_{\rm init}^2}{2} + \frac{J^2}{2mr_{\rm init}^2} + U(r_{\rm init})$.
- If the total energy is negative, then the orbit is bounded — the object always stays at finite distance from the center.
- We can construct the effective potential energy

  $$U_{eff}(r) = U(r) + \frac{J^2}{2mr^2}$$

- We the can compute the functions $r(t)$ and $\phi(t)$.

  $$t = \pm\sqrt{\frac{m}{2}} \int_{r_{\rm init}}^{r(t)} \frac{dr'}{\sqrt{E - U_{eff}(r')}}, \qquad \phi(t) = \phi_{\rm init} + \frac{J}{2m}\int_0^t \frac{dt'}{r^2(t')}$$

  The first equation gives $r(t)$ implicitly.
- We can compute two different periods

  $$T_r(E,J) = \sqrt{2m} \int_{r_{min}}^{r_{max}} \frac{dr'}{\sqrt{E - U_{eff}(r')}}, \qquad 2\pi = \frac{J}{2m}\int_0^{T_\phi(E,J)} \frac{dt'}{r^2(t')}$$

where r_{min} and r_{max} are the shortest and the largest distances to the center of the force. The are found by solving the equation

$$U_{eff}(r_{min,max}) = E.$$

- $T_r(E, J)$ is the time it takes the object of energy E and angular momentum (magnitude) J to return to the same *distance* to the center of the force.
- $T_\phi(E, J)$ is the time it takes the object of energy E and angular momentum (magnitude) J to rotate by 2π around the center of the force.
- For general $U(r)$ these two times do not coincide:

$$T_r(E, J) \neq T_r(E, J).$$

It means that after rotating by 2π the object does not return back to the same distance! So after rotating full 2π around the sun the object is NOT at the same position where it started from. The orbit is bounded, but unless $T_r(E, J)$ and $T_\phi(E, J)$ are commensurate the orbit is NOT closed.

- In the Newtonian gravity we can find the orbit in the closed form

$$\frac{1}{r} = \frac{1}{c}(1 + \epsilon\cos(\phi - \phi_0)), \qquad c = \frac{J^2}{GMm^2}, \qquad E = -\frac{1-\epsilon^2}{2}\frac{GmM}{c}.$$

ϕ_0 must be found such, that the point r_{init} and ϕ_{init} are on the orbit. So ϕ_0 can be found from the equation $\frac{1}{r_{\text{init}}} = \frac{1}{c}(1+\epsilon\cos(\phi_{\text{init}} - \phi_0))$. So the value of ϕ_0 is conserved during the motion!

- What is also surprising is that the orbit is closed!!!!!! For any $E < 0$ and J!!!!!!
- It means, that for the $1/r$ potential we have $T_r(E, J) = T_\phi(E, J)$ for ANY $E < 0$ and J!!!!!

The last point surely requires an explanation!

19.2. Conserved Laplace-Runge-Lenz vector $\vec{A}$.

The Kepler problem has an interesting additional symmetry. This symmetry ensures that $T_r(E, J) = T_\phi(E, J)$ (for any E and J). As usual this symmetry also leads to a new conservation law. In this case the Laplace-Runge-Lenz vector $\vec{A}$ is conserved. This vector is want we want to study.

If the gravitational force is

$$\vec{F} = -\frac{k}{r^2}\vec{e}_r, \qquad k = GMm, \qquad U(r) = -\frac{k}{r}$$

(in this form we can treat k as a parameter, then it also covers the Coulomb force $k = -\frac{qQ}{4\pi\epsilon_0}$) then we define the Laplace-Runge-Lenz vector $\vec{A}$:

$$\vec{A} \equiv \vec{p} \times \vec{J} - mk\vec{e}_r,$$

where $\vec{p}$ is the momentum of the object, $\vec{J} = \vec{r} \times \vec{p}$ is its angular momentum, m is the mass of the object, and $\hat{e}_r$ is the unit vector pointing from the center of the force towards the object. The vector $\vec{A}$ can be defined for both gravitational and Coulomb forces: $k > 0$ for attraction and $k < 0$ for repulsion.

An important feature of the "inverse square force" is that this vector is conserved during the motion. Let's check it. First we notice, that $\dot{\vec{J}} = 0$, so we need to compute:

$$\dot{\vec{A}} = \dot{\vec{p}} \times \vec{J} - mk\dot{\vec{e}}_r.$$

We notice, that for ANY central force $\dot{\phi} = \frac{J}{mr^2}$ and hence $\vec{\omega} = \frac{\vec{J}}{mr^2}$. Now using

$$\dot{\vec{p}} = \vec{F}, \qquad \dot{\vec{e}}_r = \vec{\omega} \times \vec{e}_r = \frac{1}{mr^2}\vec{J} \times \vec{e}_r$$

Notice, that $1/r^2$ in the last equation comes from angular momentum conservation.

We now see

$$\dot{\vec{A}} = \vec{F} \times \vec{J} - \frac{k}{r^2}\vec{J} \times \vec{e}_r = \left(\vec{F} + \frac{k}{r^2}\vec{e}_r\right) \times \vec{J} = 0$$

So this vector is indeed conserved.

- Notice, that it is conserved ONLY for $\vec{F} = -\frac{k}{r^2}\vec{e}_r$.

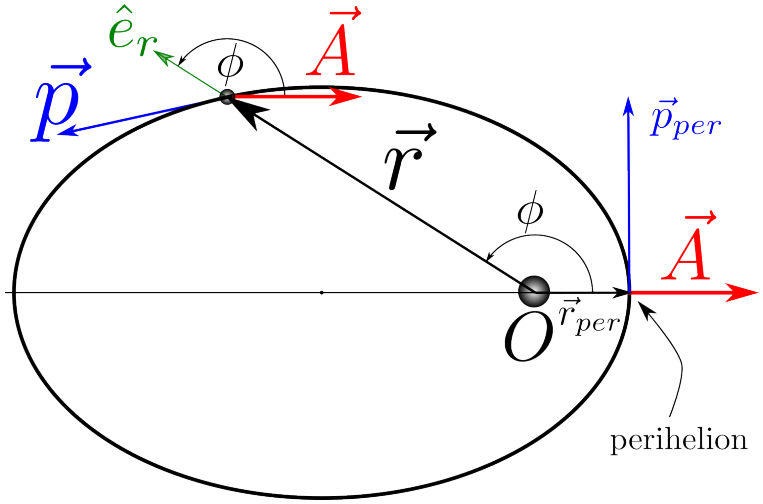
The question is: Is this conservation of vector $\vec{A}$ an independent conservation law? There are three components of the vector $\vec{A}$ are there three new conservation laws?
The answer is that not all of them are independent.

- As $\vec{J} = \vec{r} \times \vec{p}$ is orthogonal to $\vec{e}_r$, we see, that $\vec{J} \cdot \vec{A} = 0$. So the component of $\vec{A}$ perpendicular to the plane of the planet rotation is always zero $A_z = 0$. Vector $\vec{A}$ is in the plane of motion.
- Now let's calculate the magnitude of this vector

$$\vec{A} \cdot \vec{A} = \vec{p}^2\vec{J}^2 - (\vec{p} \cdot \vec{J})^2 + m^2k^2 - 2mk\vec{e}_r \cdot [\vec{p} \times \vec{J}] = \vec{p}^2\vec{J}^2 + m^2k^2 - \frac{2mk}{r}\vec{J} \cdot [\vec{r} \times \vec{p}]$$

$$= 2m\left(\frac{\vec{p}^2}{2m} - \frac{k}{r}\right)\vec{J}^2 + m^2k^2 = 2mE\vec{J}^2 + m^2k^2.$$

So we see, that the magnitude of $\vec{A}$ is not an independent conservation law.



- Using the relation between the eccentricity ϵ with $\vec{J}^2$ and E from the last lecture we find, that

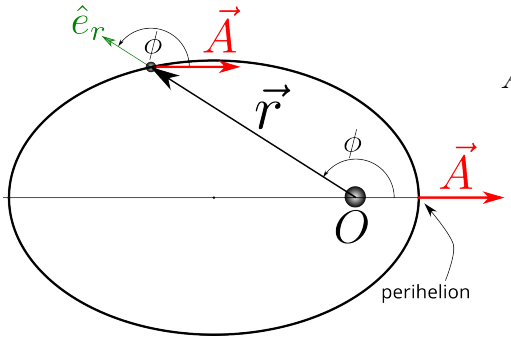
$$|\vec{A}| = \sqrt{\vec{A} \cdot \vec{A}} = \epsilon km$$

- We are left with only the direction of $\vec{A}$ within the orbit plane. Let's check this direction. As the vector is conserved it is the same at any point of the orbit, so we can chose a point where we compute it.
- So we consider the perihelion. At perihelion $\vec{p}_{per} \perp \vec{r}_{per} \perp \vec{J}$, where the subscript *per* means the value at perihelion.
- Simple examination shows that $\vec{p}_{per} \times \vec{J} = p_{per} J \vec{e}_{per}$. Then at the perihelion $\vec{A} = (p_{per} J - mk)\vec{e}_{per}$.
- However, vector $\vec{A}$ is a constant of motion, so if it has this magnitude and direction in one point it will have the same magnitude and direction at all points!
- We computed its magnitude before $|\vec{A}| = \epsilon k m$, so
$$\vec{A} = mk\epsilon \vec{e}_{per}.$$
We see, that for Kepler orbits $\vec{A}$ points to the point of the trajectory where the planet or comet is the closest to the sun.
- So we see, that $\vec{A}$ provides us with only one new independent conserved quantity.
- It also means, that if we know the velocity and the position of a planet or a comet at any time, we can compute the vector $\vec{A}$ at this moment of time and immediately know the position of the perihelion. And this position is constant — no precession. (In particular it tells us what ϕ_0 is in the equation for the orbit $\frac{1}{r} = \frac{1}{c}(1 + \epsilon \cos(\phi - \phi_0)))$.

We can also compute r_{min}, so we will know how close, say, a comet will come to the sun and where the point of the closest approach will be. We can compute this from just the initial conditions and without solving any differential equations.
But we can do more!

LECTURE 20
Kepler orbits from $\vec{A}$. Virial theorem. Panegyric to Newton.

20.1. Kepler orbits from $\vec{A}$.



Last section we showed, that for the central force $\vec{F} = -\frac{k}{r^2}\hat{e}_r$ the vector

$$\vec{A} = \vec{p} \times \vec{J} - mk\hat{e}_r,$$

is conserved.

The existence of an extra conservation law simplifies many calculations. For example we can derive equation for the trajectories without solving any differential equations. Let's do just that.

Let's derive the equation for Kepler orbits (trajectories) from our new knowledge of the conservation of the vector $\vec{A}$. For this we consider $\vec{r} \cdot \vec{A}$.

$$\vec{r} \cdot \vec{A} = \vec{r} \cdot [\vec{p} \times \vec{J}] - mkr = J^2 - mkr$$

On the other hand

$$\vec{r} \cdot \vec{A} = rA\cos\phi, \quad \text{so} \quad rA\cos\phi = J^2 - mkr$$

Or

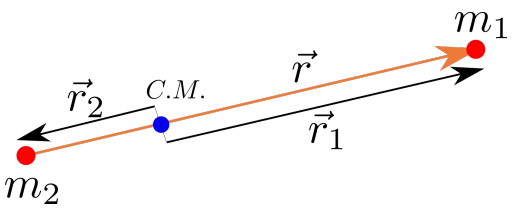
$$\frac{1}{r} = \frac{mk}{J^2}\left(1 + \frac{A}{mk}\cos\phi\right), \qquad c = \frac{J^2}{mk}, \qquad \epsilon = \frac{A}{mk}.$$

20.2. Change of orbits.

Consider a problem to change from an circular orbit Γ_1 of a radius R_1 to an orbit Γ_2 with a radius $R_2 > R_1$.

- For the transition we will use an elliptical orbit γ with $r_{min} = R_1$ and $r_{max} = R_2$.
- We need two boosts. One to go from Γ_1 to γ, and the second one to go from γ to Γ_2.
- The final speed on Γ_2 will be less than that on Γ_1.

20.3. Kepler orbits for comparable masses.



If the bodies interact only with one another and no external force acts on them, then the center of mass has a constant velocity. We then can attach our frame of reference to the center of mass and work there. This way we will only be studying the relative motion of the bodies.

Let's now consider two bodies with masses m_1 and m_2 interacting by a gravitational force. We will use center of mass system of reference and place our coordinate origin at the center of mass. If the position of m_1 is given by $\vec{r}_1$ and the position of m_2 is given by $\vec{r}_2$, then as the center of mass is in the origin we have

$$m_1\vec{r}_1 + m_2\vec{r}_2 = 0.$$

So if we know $\vec{r}_1(t)$ we immediately know $\vec{r}_2(t) = -\frac{m_1}{m_2}\vec{r}_1(t)$. So we only need to find $\vec{r}_1(t)$.

Let's write the equation of motion for $\vec{r}_1(t)$

$$m_1\ddot{\vec{r}}_1 = -\frac{k}{r^2}\vec{e}_r,$$

where $\vec{r}$ is a vector m_2 to m_1. For this vector we have

$$\vec{r} = \vec{r}_1 - \vec{r}_2 = \frac{m_1+m_2}{m_2}\vec{r}_1, \quad \text{or} \quad \vec{r}_1 = \frac{m_2}{m_1+m_2}\vec{r}.$$

Using this in the equation of motion for the mass m_1 we get

$$\frac{m_1m_2}{m_1+m_2}\ddot{\vec{r}} = -\frac{k}{r^2}\vec{e}_r, \qquad \mu\ddot{\vec{r}} = -\frac{k}{r^2}\vec{e}_r,$$

where μ is called "reduced mass"

$$\mu = \frac{m_1m_2}{m_1+m_2}$$

We then see, that the problem has reduced to a motion of a single body of a "reduced mass" μ under the same force. This is our standard problem, that we have solved before.

In the case of gravitation we can go further and use $k = Gm_1m_2 = G\frac{m_1m_2}{m_1+m_2}(m_1+m_2) = G\mu M$, where $M = m_1 + m_2$ — the total mass. So the equation of motion is

$$\mu\ddot{\vec{r}} = -\frac{G\mu M}{r^2}\vec{e}_r,$$

Or just a motion of a particle of mass μ in the gravitational field of a fixed (immovable) mass M. This is the Kepler problem which we solved before.

What one must not forget, though, is that after $\vec{r}(t)$ is found one still need to find $\vec{r}_1(t) = \frac{\mu}{m_1}\vec{r}(t)$ and $\vec{r}_2(t) = -\frac{\mu}{m_2}\vec{r}(t)$ to know the positions and motions of the real bodies.

20.4. Virial theorem

Let's consider a collection of N particles interacting with each other. Let's assume that they undergo some motion with a period T — it also means that we are in the center of mass frame of reference. Then we can define an averaged quantities as follows: Let's imagine that we have a quantity $P(\vec{r}_i, \dot{\vec{r}}_i)$ which depends on the coordinates and the velocities of all particles. We assume that we solved all equations of motion and we know all $\vec{r}_i(t)$ and hence $\dot{\vec{r}}_i(t)$. Then we define an average

$$\langle P\rangle = \frac{1}{T}\int_0^T P(\vec{r}_i(t), \dot{\vec{r}}_i(t))dt$$

- If the motion is not periodic, then we define the average a bit differently:

$$\langle P\rangle = \lim_{T\to\infty}\frac{1}{T}\int_0^T P(\vec{r}_i,\dot{\vec{r}}_i)dt$$

Now let's calculate average total kinetic energy $K=\sum_i \frac{m_i\dot{\vec{r}}_i^2}{2}$

$$\langle K\rangle = \frac{1}{T}\int_0^T\sum_i\frac{m_i\dot{\vec{r}}_i^2}{2}dt = \sum_i\frac{m_i}{2}\frac{1}{T}\int_0^T\dot{\vec{r}}_i^2dt = \sum_i\frac{m_i}{2}\frac{1}{T}\int_0^T\dot{\vec{r}}_i\cdot\dot{\vec{r}}_i dt$$

Taking the last integral by parts and using the periodicity to cancel the boundary terms we get

$$\langle K\rangle = -\frac{1}{2}\sum_i\frac{1}{T}\int_0^T\vec{r}_i\cdot m_i\ddot{\vec{r}}_i dt = -\frac{1}{2}\sum_i\frac{1}{T}\int_0^T\vec{r}_i\cdot\vec{F}_i dt = -\frac{1}{2}\frac{1}{T}\int_0^T\sum_i\vec{r}_i\cdot\vec{F}_i dt,$$

where $\vec{F}_i$ is the total force which acts on the particle i.

So we find

$$2\langle K\rangle = -\left\langle\sum_i\vec{r}_i\cdot\vec{F}_i\right\rangle.$$

So far it was all very general. Now lets assume that all the forces are the forces of Coulomb/Gravitation interaction between the particles.

$$\vec{F}_i=\sum_{j,j\neq i}\vec{F}_{ij},\qquad \vec{F}_{ij}=-\frac{k_{ij}}{r_{ij}^2}\hat{e}_{ij},\qquad k_{ij}=Gm_im_j$$

where $r_{ij}=|\vec{r}_i-\vec{r}_j|$ and $\hat{e}_{ij}$ is a unit vector pointing from j to i, $\hat{e}_{ij}=\frac{\vec{r}_i-\vec{r}_j}{r_{ij}}$. We then have for any moment of time

$$\sum_i\vec{r}_i\cdot\vec{F}_i=\sum_{i\neq j}\vec{r}_i\cdot\vec{F}_{ij}=\sum_{i>j}\vec{r}_i\cdot\vec{F}_{ij}+\sum_{i<j}\vec{r}_i\cdot\vec{F}_{ij}=\sum_{i>j}(\vec{r}_i-\vec{r}_j)\cdot\vec{F}_{ij}=-\sum_{i>j}r_{ij}\frac{k_{ij}}{r_{ij}^2}=U,$$

where U is the total potential energy of the system of the particles at the given moment of time. So we have

$$2\langle K\rangle = -\langle U\rangle$$

This is called the virial theorem.

As the total energy E is conserved — independent of time — we can write $E=\langle K\rangle+\langle U\rangle$. Using the virial theorem we find that $E=-\langle K\rangle$, and $E=\frac{1}{2}\langle U\rangle$.

- It is important, that the above relation is stated for the AVERAGES only. For example, in the perihelion of a Kepler orbit we know that $2K_{per}(1+\epsilon)=-U_{per}$.
- On the other hand for the <u>circular</u> orbit kinetic and potential energies are constant in time, so their averages are just their values.

20.5. How to see $F=\frac{GMm}{r^2}$ from Kepler's laws.

Here I will show how the Newton's gravity could be derived from the Kepler's laws. Kepler found Kepler's laws from the observations of the planet's motion. It is clear that there should be some attraction between the planets and the sun. How do we find the force of this attraction if we only know the Kepler's laws/observations and the Newton's laws of mechanics. In other words how could Newton figure out that the force of gravity is $F=\frac{GMm}{r^2}$?

The crucial observations made by Kepler were

First Law: All planets move along ellipses with the sun in the focus. Different planet's ellipses have different eccentricity and different size.

Second Law: The area swept by the position vector increases linearly with time.

Third Law: The ratio of the square of the period of orbit T to the cube of the large semi-axis a of the ellipses is the same for all planets — this ratio does not depend on the mass of the planet or the eccentricity or the size of the planet's orbit.

The argument, then is the following:

- As the ratio T^2/a^3 does not depend on eccentricity, it must be the same if a planet had a perfectly circular orbit, as a circle is just a special case of ellipse. The radius of this orbit r will play the role of the large semi-axis.
- From the second law it follows, that the speed of the planet on such an orbit must be constant.
- Let's consider this orbit of radius r. There is a force that acts on the planet $F(r)$ directed to the sun, and we must have
$$m\frac{v^2}{r} = F(r), \qquad v^2 = \frac{rF(r)}{m},$$
where m is the mass of the planet and v is its velocity.
- The period of rotation is
$$T = \frac{2\pi r}{v}.$$
- So
$$\frac{T^2}{r^3} = (2\pi)^2 \frac{r^2}{v^2}\frac{1}{r^3} = (2\pi)^2 \frac{m}{r^2 F(r)}.$$
- As this ratio must not depend neither on mass m nor on the radius r, we then must have
$$F(r) \sim \frac{m}{r^2}$$
- If the sun attracts the planet with such a force, then the planet must attract the sun with the same force. But then, according to the above formula the force must be proportional to the mass of the sun. So we have
$$F(r) = G\frac{Mm}{r^2},$$
where G is just some constant.

This is not the complete proof. We need to take the force we found, compute the arbitrary orbits, and show, that they are ellipses — just as Kepler observed.

LECTURE 21
Functions and Functionals.

21.1. Difference between functions and functionals.

- A *function* establishes a correspondence/map between elements of one set with elements of another. Usually for a number x it gives back a (single) number y according to some rule: $y = f(x)$, where f denotes this rule. So a function is a rule according to which if I give it a number it returns back a number. For example the function $f(x) = x^2$ — it is a rule, according to which if I have a number x, I need to square it and return the result back. Two different x' may return back the same number. For the previous example the numbers x and $-x$ will return the same value of $f(x)$.

  $$f: \text{ number} \longrightarrow \text{ number}.$$

  A function of many variables is a rule by which it takes a few numbers and returns one number.
- A *functional* establishes a correspondence/map between functions and numbers. Normally one has to restrict the space of functions. So a functional is a rule which one applies to a function from established space to receive back a number. Or if you give a function to a functional it returns back a number. In order to define a functional we must define the space of functions it can act on and a rule by which it returns/computes a number if we give it a function from that space.

  $$F: \text{ function} \longrightarrow \text{ number}.$$

  A functional can take more than one function as an argument.
- An *operator* takes a function from a defined space and returns back a function (from the same space).

  $$\hat{O}: \text{ function} \longrightarrow \text{ function}.$$

  We will not be dealing with operators.

21.2. Examples of functionals.

- The rule is the following. For a function $f(x)$ from the space of functions smooth on the interval $[a, b]$ the functional F returns a value of the function $f(x)$ at a point $x_0 \in (a, b)$.

  $$F[f(x)] = f(x_0).$$

This functional plays a very important role in physics. It is typically written as

$$F[f(x)] = \int_a^b \delta(x - x_0) f(x) dx = f(x_0), \qquad x_0 \in (a, b).$$

The function $\delta(x)$ is called Dirac δ-function. The above expression is the *definition* of the δ-function.

- Area under the graph: for a (integrable) functions on interval $[a, b]$ we can define a functional
$$A[f(x)] = \int_a^b f(x) dx.$$
That means, that if you have a function $f(x)$ which belongs to our space (it is integrable on the interval $[a, b]$) we can construct the number — the area under the graph. This is the rule which defines our functional.
- Length of a path in $2D$.
  - Our space is the space of smooth functions on the interval $[a, b]$.
  - For any graph $y(x)$ we can compute its length
$$\mathcal{L}[y(x)] = \int_{a,y(a)}^{b,y(b)} \sqrt{(dx)^2 + (dy)^2} = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx.$$
  - Let's now take a path $x(t)$, $y(t)$, where $t \in [a, b]$ is a parameter. Both $x(t)$, $y(t)$ are smooth. Then the length of this path is
$$\mathcal{L}[x(t), y(t)] = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} dt.$$
This is a functional of 2 functions: $x(t)$ and $y(t)$ smooth on the interval $t \in [a, b]$.
- Length of a path in $3D$.
  - Now the path is given by three functions $x(t)$, $y(t)$, and $z(t)$, where $t \in [a, b]$ is a parameter. All three functions $x(t)$, $y(t)$, and $z(t)$ are smooth. Then the length of this path is
$$\mathcal{L}[x(t), y(t), z(t)] = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} dt.$$
This is a functional of 3 functions: $x(t)$, $y(t)$, $z(t)$ smooth on the interval $t \in [a, b]$.

It is important to specify the space of functions.

LECTURE 22
More on functionals.

22.1. Examples of functionals. Continued.

- Length of a path. Invariance under reparametrization.
  - In the last lecture we considered a path $x(t)$, $y(t)$, where $t \in [a, b]$ is a parameter. Both $x(t)$, $y(t)$ are smooth. Then the length of this path is

$$\mathcal{L}[x(t), y(t)] = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt.$$

  - Let's now change this parameter. Namely we take t to be a function of another parameter τ: $t(\tau)$. The very same graph is given by $x(\tau) = x(t(\tau))$ and $y(\tau) = y(t(\tau))$. Then the length is

$$\mathcal{L}[x(\tau), y(\tau)] = \int_{a_\tau}^{b_\tau} \sqrt{\left(\frac{dx}{d\tau}\right)^2 + \left(\frac{dy}{d\tau}\right)^2}\, d\tau,$$

    where $t(a_\tau) = a$, $t(b_\tau) = b$. Using the chain rule we get $\frac{dx}{d\tau} = \frac{dx}{dt}\frac{dt}{d\tau}$ and the same for $\frac{dy}{d\tau}$, as well as $d\tau = \frac{d\tau}{dt}dt$ we will get exactly the same expression as before. So the length – the functional – is invariant under reparametrization.
  - In N dimensional space a curve is given by smooth functions $x_i(t)$, $i = 1 \dots N$. The (Euclidean) length of this curve is given by

$$\mathcal{L}[x_i(t)] = \int_a^b \sqrt{\frac{dx_i}{dt}\frac{dx_i}{dt}}\, dt.$$

    It is a functional on N functions.
    A side note: One can also define a functional

$$\mathcal{L}[x_i(t)] = \int_a^b \sqrt{\frac{dx_i}{dt} g_{ij}(\{x_i\}) \frac{dx_j}{dt}}\, dt,$$

    where $g_{ij}(\{x_i\})$ is $N \times N$ matrix which depends on the position given by all $\{x_i\}$. This matrix must be positive definite at each point of space. This matrix is called the metric tensor of the space. In Euclidean space the metric tensor is just $N \times N$ unit matrix δ_{ij}.
- Energy of a horizontal string in the gravitational field.

- Consider a rope linear density ρ. We attach it to two nails distance l apart which are on the same height. We take the rope to have a shape given by a smooth function $y(x)$ (The origin is at the left nail, x is the horizontal coordinate pointing to the right, y is the vertical coordinate pointing up.) with
$$y(0) = y(l) = 0$$
and ask the following question: What is the potential energy of the rope of this shape?
- Consider a small piece of the rope. It has a mass $\rho\sqrt{(dx)^2 + (dy)^2}$. The potential energy of this piece is $\rho g y\sqrt{(dx)^2 + (dy)^2}$. So the total potential energy is
$$U[y(x)] = \rho g \int_0^l y(x)\sqrt{1 + (y')^2}dx.$$
- It is a functional on a space of smooth functions $y(x)$ in the interval $[0, l]$ with the condition $y(0) = y(l) = 0$.
- If the total length $L > l$ of the rope is given, then the we further restrict our space of functions $y(x)$ by demanding, that they satisfy the following constraint
$$L = \mathcal{L}[y(x)].$$

- Value at a point as functional. The functional which for any function returns the value of the function at a given point.
- Functions of many variables. Area of a surface. Invariance under reparametrization.

It is important to specify the space of functions.

22.2. General form of the functionals.

- We need to establish a rule which will allow to compute a number for an arbitrary function $y(x)$ from the given space.
- General form of a functional
$$\int_{x_1}^{x_2} L(x, y, y', y'', \dots)dx.$$
- The space of functions, the integration boundaries x_1 and x_2 as well as the function L are parts of the DEFINITION of the functional. They must be given in order for the functional to be defined.
- Important: In the function L the y, y', y'' and so on are just independent variables.
- It means that we consider a function $L(x, z_1, z_2, z_3, \dots)$ of normal variables x, z_1, z_2, z_3, $\dots$ and for any function $y(x)$ at some point x we calculate $y(x)$, $y'(x)$, $y''(x)$, $\dots$ and plug x and these values instead of z_1, z_2, z_3, $\dots$ in $L(x, z_1, z_2, z_3, \dots)$.
- We do that for all values of x in the interval $[x_1, x_2]$, so that we obtain (FOR GIVEN $y(x)$!!!) a function of x: $L(x, y(x), y'(x), y''(x), \dots)$, and then we integrate this function over x in the interval $[x_1, x_2]$.

As $L(x, z_1, z_2, z_3, \dots)$ is just a KNOWN function of its arguments we can compute the partial derivatives $\frac{\partial L}{\partial x}$, $\frac{\partial L}{\partial z_1}$, $\frac{\partial L}{\partial z_2}$, etc. AFTER we computed these derivatives we can substitute the GIVEN function $y(x)$ and its derivatives instead of the corresponding variables z_1 ,z_2, z_3,

....

$$\left.\frac{\partial L(x,z_1,z_2,z_3,\dots)}{\partial z_1}\right|_{z_1=y(x),z_2=y'(x),\dots}, \qquad \left.\frac{\partial L(x,z_1,z_2,z_3,\dots)}{\partial z_2}\right|_{z_1=y(x),z_2=y'(x),\dots}, \qquad \dots$$

This procedure is DENOTED as $\frac{\partial L}{\partial y}$, $\frac{\partial L}{\partial y'}$, $\frac{\partial L}{\partial y''}$, etc.

$$\begin{aligned}
\frac{\partial L}{\partial y} &\equiv \left.\frac{\partial L(x,z_1,z_2,z_3,\dots)}{\partial z_1}\right|_{z_1=y(x),z_2=y'(x),\dots}, \\
\frac{\partial L}{\partial y'} &\equiv \left.\frac{\partial L(x,z_1,z_2,z_3,\dots)}{\partial z_2}\right|_{z_1=y(x),z_2=y'(x),\dots}, \\
\frac{\partial L}{\partial y''} &\equiv \left.\frac{\partial L(x,z_1,z_2,z_3,\dots)}{\partial z_3}\right|_{z_1=y(x),z_2=y'(x),\dots}, \\
&\dots
\end{aligned}$$

Another way of thinking about this notation is the following: You are typically given a functional in the form $\int_{x_1}^{x_2} L(x,y,y',y'',\dots)dx$. When you are taken the derivatives $\frac{\partial L}{\partial y}$, $\frac{\partial L}{\partial y'}$ etc., you treat y, y', etc. as DIFFERENT letters — different INDEPENDENT variables, and only after the derivatives are computed you recall that these variables are the function $y(x)$ and its derivatives.

22.3. Discretization. Fanctionals as functions.

Let's consider a functional $A[f(x)]$ acting on the functions from some well defined space, let's say on smooth functions on the interval $[a,b]$. We can do the following trick.

- Consider the variable x to be discretized: instead of thinking of x as a continuous variable we will select N points x_i in the interval $[a,b]$. Lat's also take $x_1 = a$, $x_N = b$.
- Eventually we will need to take a limit $N \to \infty$. This limit should be taken in such a way, that $\max(\Delta x_i) \to 0$.
- A function $f(x)$ is then represented by its values f_i at x_i: $f_i = f(x_i)$.
- Then the functional $A[f(x)]$ can be thought as a function of the values f_i: $A[f(x)] = A(f_1,\dots,f_N)$.
- We then can deal with the functional as a with the function of many variables.
- At the end we must take the limit $N \to \infty$ as described above, and make sure, that such limit does exist.

In many non-trivial cases this procedure allows one to make sense out of the calculations.

If you are to compute the value of a functional numerically, then this procedure is exactly what you have to do.

LECTURE 23
Euler-Lagrange equation

23.1. A word on notations.

We will consider the functionals of the form

$$\mathcal{A}[y(x)] = \int_{x_A}^{x_B} L(x, y(x), y'(x))dx.$$

- One MUST always specify the space of functions the functional is defined on.
- The function L under the integral is whatever is multiplied by dx in the integral.
- Function L is a FUNCTION.
- When we look at it as a function, x, $y(x)$, and $y'(x)$ are simply its variables.
- The procedure to find the value of a functional on a function is the following
  - You have a functional $\mathcal{A}[y(x)]$. What it means is that you have a defined space of functions, you have numbers x_A and x_B, and you have a function $L(x, z_1, z_2)$ — now your functional is specified.
  - You are given a function $y(x)$ from the specified space and you are asked to compute the value of the functional $\mathcal{A}$ on this function.
  - at the first step you compute $y'(x)$. Then

$$\begin{array}{ccccc} & & y(x) & y'(x) & \\ & & \downarrow & \downarrow & \\ L(x, & & z_1, & z_2) & \longrightarrow L(x, y(x), y'(x)) \longrightarrow \mathcal{A}[y(x)] = \int_{x_A}^{x_B} L(x, y(x), y'(x))dx \end{array}$$

- As for any (smooth enough) function we can differentiate the function L over its variables.
- We can take the partial derivatives of the function L with respect to its arguments. For example

$$\frac{\partial L}{\partial y(x)}.$$

- This expression simply means that you need to take the function $L(x, z_1, z_2)$, differentiate it with respect to the second argument z_1 and AFTER the differentiation substitute $z_1 = y(x)$, $z_2 = y'(x)$

$$\frac{\partial L}{\partial y(x)} \equiv \left.\frac{\partial L(x, z_1, z_2)}{\partial z_1}\right|_{z_1=y(x), z_2=y'(x)}.$$

- Analogously
$$\frac{\partial L}{\partial y'(x)} \equiv \left.\frac{\partial L(x,z_1,z_2)}{\partial z_2}\right|_{z_1=y(x),z_2=y'(x)}.$$
- IMPORTANT: The substitution $z_1 = y(x)$, $z_2 = y'(x)$ is done AFTER the partial derivative is taken.
- To shorten the notations these derivatives are written as
$$\frac{\partial L}{\partial y}, \qquad \frac{\partial L}{\partial y'}$$

23.2. Minimization problem

What kind of problems can we state with the functionals?

One of the most important problem (but not the only one) is stated as following: given a functional $\mathcal{A}[f(x)]$ (remember, that the space the functional works on is a part of its definition) which function (from the defined space) will give the smallest (or the largest) value of the functional? How do we find this function?

For an arbitrary functional such function may not exist. Moreover, generally if you change the space you will find a different answer. In many cases, if you change the space the question will not have an answer.

Notice, that this is exactly the same situation as with functions. A function may or may not have minimum or maximum on a given interval. This statement depends on the interval. For example a function $1/x$ has no maximum or minimum in the interval $[-1, 1]$, but it has a minimum and a maximum in the interval $[1, 2]$. The position of the maximum and minimum depends on the interval boundaries.

In the following examples notice the importance of defining the space of functions.

- Minimal distance between two points.
- Minimal time of travel. Ferma Principe.
- Minimal potential energy of a string.
- etc.

23.3. Minimum of a function.

Before we derive the equation for the function which minimizes a functional. Let's remember how it is done for functions.

The questions is: if we have a function $f(x)$ how do we find the position x_0 of its minimum?

There are different ways to think about it. I want to emphasize the following line of arguments:

- Let's assume, that we know the position of the minimum x_0.
- Let's consider x which is very close to x_0.
- We know that if x is close enough to x_0 the value of the function at x can be represented as a series ($\delta x \equiv x - x_0$)
$$f(x) = f(x_0) + a_1\delta x + a_2(\delta x)^2 + \dots$$
where the coefficients a_1, a_2, etc. are the coefficients of the Taylor expansion. They are some fixed numbers! We know how to compute these numbers: $a_1 = \left.\frac{df}{dx}\right|_{x=x_0}$, and $a_2 = \frac{1}{2}\left.\frac{d^2f}{dx^2}\right|_{x=x_0}$.

- In this series for δx small enough the term $a_1\delta x$ is dominant. And it's dominance is the larger the smaller δx is.
- So for very small δx we can write
$$\delta f = f(x) - f(x_0) \approx a_1\delta x.$$
- As $f(x_0)$ is the minimum, for small enough δx we must have $\delta f > 0$. This must be true for both positive and negative δx!
- The only way to have ensure this inequality is to have
$$a_1 = 0.$$
- Then the Taylor expansion starts with the term $a_2(\delta x)^2$ which is positive if $a_2 > 0$ for any δx.
- According to Taylor expansion $a_1 = \frac{\partial f}{\partial x}\Big|_{x=x_0}$. So to find the minimum we need to solve the equation
$$\frac{\partial f}{\partial x}\Big|_{x=x_0} = 0$$

Notice, that the condition which leads to the equation above is that the change of the function in the first order in δx is zero!

Notice the whole strategy to find the position of the minimum/extremum for a function $f(x)$:

- We first assume, that we know the position of the minimum x_0.
- We find how the function $f(x)$ changes, when we shift x by small δx from x_0.
- We compute the change $\delta f = f(x_0 + \delta x) - f(x)$ in LINEAR order of δx.
- We DEMAND, that x_0 be such, that this change δf vanishes (at LINEAR order in δx).
- As the result we obtain an algebraic equation for x_0.

23.4. The Euler-Lagrange equations

- The functional $\mathcal{A}[y(x)] = \int_{x_1}^{x_2} L(y(x), y'(x), x)dx$ with the boundary conditions $y(x_1) = y_1$ and $y(x_2) = y_2$.
- The problem is to find a function $y(x)$ which is the stationary "point" of the functional $\mathcal{A}[y(x)]$.
- The stationary "point" (it is a function, it is a "point" in the space of functions) of a functional $\mathcal{A}[y(x)] = \int_{x_1}^{x_2} L(x, y(x), y'(x))dx$ for the functions satisfying $y(x_1) = y_1$, $y(x_2) = y_2$ is given by the solution of Euler-Lagrange equation.
- Euler-Lagrange equation is the second order differential equation with boundary conditions $y(x_1) = y_1$, $y(x_2) = y_2$.
- Derivation of the Euler-Lagrange equation.
  - Let's assume, that we found the function $y_0(x)$ which gives us a minimum of the functional $\mathcal{A}[y(x)] = \int_{x_1}^{x_2} L(x, y(x), y'(x))dx$ for the functions satisfying $y(x_1) = y_1$, $y(x_2) = y_2$.
  - Lets shift this function a little and consider the function $y(x) = y_0(x) + \delta y(x)$, where $\delta y(x)$ is small/infinitesimal.
  - The new function $y(x)$ must be from the same space, so me must have

$$\delta y(x_1) = 0, \quad \delta y(x_2) = 0. \tag{23.1}$$

– The value of our functional on the new function is

$$\mathcal{A}[y_0(x)+\delta y(x)] = \int_{x_1}^{x_2} L(x, y_0(x)+\delta y(x), y_0'(x)+\delta y'(x))dx$$

– Let's compute $\mathcal{A}[y_0(x)+\delta y(x)]$ up to the linear order in $\delta y(x)$:

$$\mathcal{A}[y_0(x)+\delta y(x)] \approx \int_{x_1}^{x_2} \left[L(x, y_0(x), y_0'(x)) + \left.\frac{\partial L}{\partial y}\right|_{y=y_0(x)} \delta y + \left.\frac{\partial L}{\partial y'}\right|_{y'=y_0'(x)} \delta y' \right] dx$$

Here I treated $L(x, y, y')$ as just a function of its INDEPENDENT variables x, y, and y', differentiated it with respect to these variables and then plugged y_0 instead of y and y_0' instead of y'.

– To shorten notations I will use $\frac{\partial L}{\partial y_0}$ to mean $\left.\frac{\partial L}{\partial y}\right|_{y=y_0(x)}$, and the same for the primed term.

– Notice, that after this substitution $y = y_0(x)$ the functions $\frac{\partial L}{\partial y_0}$ and $\frac{\partial L}{\partial y_0'}$ are the functions of x only!

– The first term under the integral is what is in $\mathcal{A}[y_0(x)]$ — the value of the functional at the minimum.

$$\mathcal{A}[y_0(x)+\delta y(x)] \approx A[y_0(x)] + \int_{x_1}^{x_2} \left[\frac{\partial L}{\partial y_0}\delta y + \frac{\partial L}{\partial y_0'}\delta y' \right] dx$$

– Let's call $\delta\mathcal{A} = \mathcal{A}[y_0(x)+\delta y(x)] - \mathcal{A}[y_0(x)]$. It is called variation of the functional.

$$\delta\mathcal{A} \approx \int_{x_1}^{x_2} \frac{\partial L}{\partial y_0}\delta y(x)dx + \int_{x_1}^{x_2} \frac{\partial L}{\partial y_0'}\frac{d\delta y(x)}{dx}dx$$

– Notice, that in the last term in $\frac{d\delta y(x)}{dx}$ it is a full derivative over x. The function $\frac{\partial L}{\partial y_0'}$ is a function of x only, as we already plugged $y_0(x)$ instead of y and $y_0'(x)$ instead of y'.

– I will use the partial integration on that term

$$\delta\mathcal{A} \approx \int_{x_1}^{x_2} \frac{\partial L}{\partial y_0}\delta y(x)dx + \delta y(x)\left.\frac{\partial L}{\partial y_0'}\right|_{x_1}^{x_2} - \int_{x_1}^{x_2} \delta y(x)\frac{d}{dx}\frac{\partial L}{\partial y_0'}dx.$$

Notice that in this step $\frac{d}{dx}\frac{\partial L}{\partial y'}$ assumes full differentiation over x.

– Now we use the boundary conditions (23.1) and see that

$$\delta y(x)\left.\frac{\partial L}{\partial y_0'}\right|_{x_1}^{x_2} = 0.$$

– So we have

$$\delta\mathcal{A} \approx \int_{x_1}^{x_2} \delta y(x) \left[\frac{\partial L}{\partial y_0} - \frac{d}{dx}\frac{\partial L}{\partial y_0'} \right] dx.$$

– This equation tells us how the value of the functional $A[y(x)]$ changes, when we change the function from the minimum $y_0(x)$ by an ARBITRARY infinitesimal function δy (subject, of course to (23.1)).

– As the function $\delta y(x)$ is arbitrary, the value of the integral $\int_{x_1}^{x_2} \delta y(x) \left[\frac{\partial L}{\partial y_0} - \frac{d}{dx}\frac{\partial L}{\partial y_0'} \right] dx$ can be either positive or negative.

- But the function $y_0(x)$ is the minimum! If we shift from the minimum we can only go up, so the value $\delta\mathcal{A}$ must always be positive! (or non-negative in the linear order – it will become positive in the quadratic order)
- The only way to ensure that $\delta\mathcal{A}$ is non-negative for ARBITRARY $\delta y(x)$ is to demand, that

$$\frac{\partial L}{\partial y_0} - \frac{d}{dx}\frac{\partial L}{\partial y_0'} = 0.$$

The statement then is that the function $y_0(x)$ must be such as to satisfy this equation.

- The Euler-Lagrange equation reads

$$\frac{d}{dx}\frac{\partial L}{\partial y'} = \frac{\partial L}{\partial y}, \qquad y(x_1) = y_1, \quad y(x_2) = y_2$$

This is the second order differential equation with boundary conditions $y(x_1) = y_1$, $y(x_2) = y_2$.

LECTURE 24
Euler-Lagrange equation continued.

24.1. What to do.

If you have a functional which acts on functions $f(\xi)$ given by

$$\mathcal{A}[f(\xi)] = \int_{\xi_A}^{\xi_B} L(\xi, f(\xi), f'(\xi))d\xi.$$

(this means that you know ξ_A, ξ_B, and the function L of three INDEPENDENT variables), with the space of functions defined as smooth functions with

$$f(\xi_A) = f_A, \qquad f(\xi_B) = f_B.$$

(or any other way.) You can find such a function $f(\xi)$ that minimizes (or maximizes) the functional $\mathcal{A}$.

Here is what you do.

- You take the function $L(\xi, f, f')$ and differentiate it with respect to the f and f' treating f and f' as simply the names of INDEPENDENT variables. You get
$$\frac{\partial L}{\partial f}, \qquad \frac{\partial L}{\partial f'}.$$
- AFTER that you treat both f and f' as (unknown) functions of ξ: $f(\xi)$ and $f'(\xi) \equiv \frac{df(\xi)}{d\xi}$, and take the derivative
$$\frac{d}{d\xi}\frac{\partial L}{\partial f'},$$
using the chain rule (you will have to use the chain rule, as you do not know the function $f(\xi)$).
- You write the second order differential equation
$$\frac{d}{d\xi}\frac{\partial L}{\partial f'} = \frac{\partial L}{\partial f}.$$
- You solve it and find the solution which satisfies the boundary (or other) conditions
$$f(\xi_A) = f_A, \qquad f(\xi_B) = f_B.$$

This is a mathematical procedure. There is no "meaning" of the functional $\mathcal{A}$ or function f, or variable ξ. The "meaning" of these objects comes from the physical problem which you are trying to solve with this machinery.

24.2. Examples

24.2.1. Shortest path between two points in $2D$.

There are two points: point A and point B in $2D$ Euclidean space. We want to find a path between the two points which has the smallest length.

24.2.1.1. *Cartesian coordinates.*
We introduce Cartesian coordinates x and y. The points A and B have coordinates $A : (x_A, y_A)$ and $B : (x_B, y_B)$. A path is given by a function $y(x)$.

The displacement vector is $d\vec{r} = dx\hat{e}_x + dy\hat{e}_y$. The length of this vector is $dl^2 = d\vec{r}\cdot d\vec{r} = (dx)^2 + (dy)^2$. The element of length of the path $y(x)$ is $dl = \sqrt{(dx)^2 + (dy)^2} = \sqrt{1 + (y')^2}dx$. The length of the path between the points A and B given by $y(x)$ is

$$\mathcal{L}[y(x)] = \int_{x_A}^{x_B} \sqrt{1 + (y')^2}dx, \qquad y(x_A) = y_A, \quad y(x_B) = y_B.$$

We want to find which path $y(x)$ gives the shortest length. In other words, what function $y(x)$ gives the minimum of the functional $\mathcal{L}[y(x)]$.

After the question is posed this way we know what to do!

- The function $L(y, y', x)$ is given by
$$L(y, y', x) = \sqrt{1 + (y')^2}$$
It has no dependence on x and y, only on y'. This will simplify the calculations, but changes nothing about the procedure.
- We compute the partial derivatives
$$\frac{\partial L}{\partial y} = 0, \qquad \frac{\partial L}{\partial y'} = \frac{y'}{\sqrt{1 + (y')^2}}.$$
- We write the Euler-Lagrange equation
$$\frac{d}{dx}\frac{y'}{\sqrt{1 + (y')^2}} = 0$$
- At this point we treat y and y' as the functions of x (in this particular problem we do not have y).
$$\frac{y''}{\sqrt{1 + (y')^2}} - \frac{(y')^2 y''}{(1 + (y')^2)^{3/2}} = 0$$
- This is the differential equation we need to solve. We need to find a solution which satisfies the boundary conditions
$$y(x_A) = y_A, \qquad y(x_B) = y_B.$$
The rest is simply a solution of this equation.
- The easier way to solve this Euler-Lagrange equation:

$$\frac{d}{dx}\frac{y'}{\sqrt{1 + (y')^2}} = 0, \qquad \frac{y'}{\sqrt{1 + (y')^2}} = \text{const.}, \qquad y'(x) = \text{const.}, \qquad y = ax + b.$$

The constants a and b should be computed from the boundary conditions $y(x_A) = y_A$ and $y(x_B) = y_B$.

24.2.1.2. *Polar coordinates.*

We introduce polar coordinates r and ϕ. The points A and B have coordinates $A : (r_A, \phi_A)$ and $B : (r_B, \phi_B)$. A path is given by a function $r(\phi)$.

The displacement vector is $d\vec{r} = dr\hat{e}_r + rd\phi\hat{e}_\phi$. The length of this vector is $dl^2 = d\vec{r} \cdot d\vec{r} = (dr)^2 + r^2(d\phi)^2$. The element of length of the path $r(\phi)$ is $dl = \sqrt{(dr)^2 + r^2(d\phi)^2} = \sqrt{(r')^2 + r^2}d\phi$.

The length of the path between the points A and B given by $r(\phi)$ is

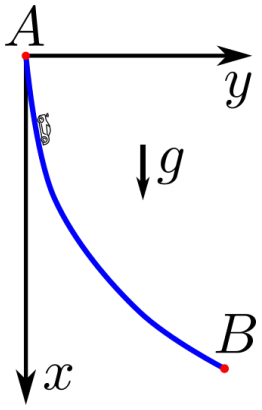
$$\mathcal{L}[r(\phi)] = \int_{\phi_A}^{\phi_B} \sqrt{(r')^2 + r^2}d\phi, \qquad r(\phi_A) = r_A, \quad r(\phi_B) = r_B.$$

We want to find which path $r(\phi)$ gives the shortest length. In other words, what function $r(\phi)$ gives the minimum of the functional $\mathcal{L}[r(\phi)]$.

After the question is posed this way we know what to do!

- The function $L(\phi, r, r')$ is
$$L = \sqrt{(r')^2 + r^2}.$$
- Treating r and r' as INDEPENDENT variables we compute
$$\frac{\partial L}{\partial r} = \frac{r}{\sqrt{(r')^2 + r^2}}, \qquad \frac{\partial L}{\partial r'} = \frac{r'}{\sqrt{(r')^2 + r^2}}.$$
- We write the Euler-Lagrange equation, treating r and r' as the functions of ϕ!!!
$$\frac{d}{d\phi}\frac{\partial L}{\partial r'} = \frac{\partial L}{\partial r}.$$
- We get
$$\frac{r''}{\sqrt{(r')^2 + r^2}} - \frac{r'(r'r'' + rr')}{((r')^2 + r^2)^{3/2}} = \frac{r}{\sqrt{(r')^2 + r^2}}.$$
This is the differential equation we need to solve. We need to find a solution which satisfies the boundary conditions
$$r(\phi_A) = r_A, \qquad r(\phi_B) = \phi_B.$$
The rest is simply a solution of this equation.
- After simple algebra we get
$$r''r^2 - 2rr'^2 - r^3 = 0.$$
- Making the substitution $r(\phi) = \frac{1}{u(\phi)}$ and hence $r' = -\frac{u'}{u^2}$, $r'' = -\frac{u''}{u^2} + 2\frac{u'^2}{u^3}$ we convert the equation into
$$u'' = -u.$$
- This is an oscillator equation with the solution $u = C\cos(\phi - \phi_0)$.
- So the solution $r(\phi)$ is
$$r(\phi) = \frac{A}{\cos(\phi - \phi_0)}.$$
The arbitrary constants A and ϕ_0 must be found from the boundary conditions.
- This is the equation of the straight line in the polar coordinates.

24.2.2. Shortest time to fall — Brachistochrone.



- What path the rail should be in order for the car to take the least amount of time to go from point A to point B under gravity if it starts with zero velocity.
- Lets take the coordinate x to go straight down and y to be horizontal, with the origin in point A. (I switched the conventional names for the axes because I know that this problem is easier to solve this way.)
- The boundary conditions: for point A: $y(0) = 0$; for point B: $y(x_B) = y_B$.
- We need to write the functional which gives the total time of travel for a rail described by a function $y(x)$. From the energy conservation (there is no friction!) $mv^2/2 = mgx$, so the velocity at "height" x is $v(x) = \sqrt{2gx}$. The length dl of the rail between "heights" x and $x + dx$ is $dl = \sqrt{(dx)^2 + (dy)^2} = \sqrt{1 + (y')^2}dx$, so the time it takes to travel from the "height" x to the "height" $x + dx$ is $dt = \frac{dl}{v(x)}$.
- The total time of travel between the points A and B along the rail given by $y(x)$ is

$$T = \int \frac{dl}{v} = \int_0^{x_B} \frac{\sqrt{1 + (y')^2}}{\sqrt{2gx}} dx.$$

- We have

$$L(y, y', x) = \frac{\sqrt{1 + (y')^2}}{\sqrt{2gx}}, \qquad \frac{\partial L}{\partial y} = 0, \qquad \frac{\partial L}{\partial y'} = \frac{1}{\sqrt{2gx}} \frac{y'}{\sqrt{1 + (y')^2}}.$$

- The Euler-Lagrange equation is

$$\frac{d}{dx}\left(\frac{1}{\sqrt{2gx}} \frac{y'}{\sqrt{1 + (y')^2}}\right) = 0.$$

- At this point we treat y and y' as the functions of x (in this particular problem we do not have y).

$$-\frac{1}{2}\frac{1}{x^{3/2}\sqrt{2g}} \frac{y'}{\sqrt{1 + (y')^2}} + \frac{1}{\sqrt{2gx}} \frac{y''}{\sqrt{1 + (y')^2}} - \frac{1}{\sqrt{2gx}} \frac{(y')^2 y''}{(1 + (y')^2)^{3/2}} = 0$$

We need to solve this second order differential equation and find the solution which satisfies the boundary conditions.

- The easier way to solve this Euler-Lagrange equation.

$$\frac{d}{dx}\left(\frac{1}{\sqrt{2gx}}\frac{y'}{\sqrt{1+(y')^2}}\right)=0, \qquad \frac{1}{x}\frac{(y')^2}{1+(y')^2}=\frac{1}{2a}, \qquad y'(x)=\sqrt{\frac{x}{2a-x}}$$

- So the path is given by

$$y(x)=\int_0^x\sqrt{\frac{\tilde{x}}{2a-\tilde{x}}}d\tilde{x}$$

- The integral is taken by substitution $\tilde{x}=a(1-\cos\tilde{\theta})$. It then becomes

$$y(x(\theta))=a\int_0^\theta(1-\cos\tilde{\theta})d\tilde{\theta}=a(\theta-\sin\theta), \qquad x=a(1-\cos\theta).$$

So the path is given by the parametric equations

$$x=a(1-\cos\theta), \qquad y=a(\theta-\sin\theta).$$

The point $x=0$, $y=0$ is already on the path for $\theta=0$. The constant a must be chosen such, that the point x_B, y_B is on the path. It means, that there is such θ_B that

$$x_A=a(1-\cos\theta_B)$$
$$y_A=a(\theta_B-\sin\theta_B)$$

Notice, that these are two equations with two unknowns θ_B and a. (These equations are transcendental. The solution does exists but cannot be expressed through any normal way. It can be found numerically.)

24.3. Reparametrization. For self-study.

The *form* of the Euler-Lagrange equation does not change under the reparametrization.

Consider a functional and corresponding E-L equation

$$A=\int_{x_1}^{x_2}L(y(x),y'_x(x),x)dx, \qquad \frac{d}{dx}\frac{\partial L}{\partial y'_x}=\frac{\partial L}{\partial y(x)}$$

Let's consider a new parameter ξ and the function $x(\xi)$ converts one old parameter x to another ξ.

Using the function $x(\xi)$ we can change the variable in the function $y(x)$

$$y(\xi)\equiv y(x(\xi)), \qquad y'_x=\frac{dy}{dx}=\frac{dy}{d\xi}\frac{d\xi}{dx}\equiv y'_\xi\frac{d\xi}{dx}, \qquad dx=\frac{dx}{d\xi}d\xi$$

Our functional then becomes

$$A=\int_{x_1}^{x_2}L(y(x),y'_x(x),x)dx=\int_{\xi_1}^{\xi_2}L\left(y(\xi),y'_\xi\frac{d\xi}{dx},x\right)\frac{dx}{d\xi}d\xi,$$

So that we have the new function (the one between the integral and $d\xi$)

$$L_\xi(y,y'_\xi,\xi)=L\left(y(\xi),y'_\xi\frac{d\xi}{dx},x\right)\frac{dx}{d\xi}$$

This is now a function of y, y'_ξ, and ξ, and

$$A = \int_{\xi_1}^{\xi_2} L_\xi(y(\xi), y'_\xi(\xi), \xi) d\xi$$

($x(\xi_1) = x_1$, $x_{\xi_2} = x_2$)

The E-L equation then is

$$\frac{d}{d\xi}\frac{\partial L_\xi}{\partial y'_\xi} = \frac{\partial L_\xi}{\partial y(\xi)}$$

Using

$$\frac{\partial L_\xi}{\partial y'_\xi} = \frac{dx}{d\xi}\frac{\partial L}{\partial y'_x}\frac{d\xi}{dx} = \frac{\partial L}{\partial y'_x}, \qquad \frac{\partial L_\xi}{\partial y(\xi)} = \frac{dx}{d\xi}\frac{\partial L}{\partial y(x)}$$

we see that E-L equation reads

$$\frac{d}{d\xi}\frac{\partial L}{\partial y'_x} = \frac{dx}{d\xi}\frac{\partial L}{\partial y(x)}, \qquad \frac{d}{dx}\frac{\partial L}{\partial y'_x} = \frac{\partial L}{\partial y(x)}.$$

So we return back to the original form of the E-L equation.

What we found is that E-L equations are invariant under the parameter change.

LECTURE 25
Lagrangian mechanics.

25.1. The Euler-Lagrange equations, for many variables.

If we have a functional of two functions $y(x)$ and $z(x)$

$$\mathcal{A}[y(x), z(x)] = \int_{x_1}^{x_2} L(x, y(x), z(x), y'(x), z'(x))dx$$

then, as we derived the Euler-Lagrange equation working with the functional variations only in the linear order, we have simply the E-L equation for each of the function

$$\frac{d}{dx}\frac{\partial L}{\partial y'} = \frac{\partial L}{\partial y}$$
$$\frac{d}{dx}\frac{\partial L}{\partial z'} = \frac{\partial L}{\partial z}$$

And so on.

- It is VERY important that the functions $y(x)$ and $z(x)$ etc. are independent from each other. We have to be able to take the variations over them INDEPENDENTLY in order for the Euler-Lagrange equations to be valid.

25.2. Problems of Newton laws.

- Not invariant when we change the coordinate system:

$$\text{Cartesian: } \begin{cases} m\ddot{x} = F_x \\ m\ddot{y} = F_y \end{cases}, \qquad \text{Cylindrical: } \begin{cases} m\left(\ddot{r} - r\dot{\phi}^2\right) = F_r \\ m\left(r\ddot{\phi} + 2\dot{r}\dot{\phi}\right) = F_\phi \end{cases}.$$

- Too complicated, too tedious. Consider two pendulums.
- Difficult to find conservation laws.
- Symmetries are not obvious.
- Cannot be used in non-classical world.

25.3. Newton second law as Euler-Lagrange equations

Second order differential equation.

25.4. Hamilton's Principle. Action.

Hamilton's Principle: For each conservative mechanical system there exists a functional, called action, which is minimal on the solution of the equations of motion

This functional — Action — has the following form:

$$\mathcal{A}[\{q_i(t)\}] = \int_{t_i}^{t_f} L(t, \{q_i(t)\}, \{\dot{q}_i(t)\})dt.$$

Let's see what it means.

- $\{q_i\}$ — a set of numbers which describes the configuration/position of our system. These numbers are called generalized coordinates.
  - A set of numbers which unambiguously describes the configuration of the system.
  - These numbers must provide the complete description. Example when it is not.
  - These numbers must be independent. Example when it is not. Bead on a rail.
- During the motion these generalized coordinates change as functions of time t. I collectively denoted the full set of these functions as $\{q_i(t)\}$.
- Correspondingly, there are generalized velocities: $\dot{q}_i = \frac{dq_i}{dt}$ for each of the coordinates. I collectively denote them as $\{\dot{q}_i(t)\}$.
- t_i is initial moment of time, t_f is the final moment.
- The function $L(t, \{q_i(t)\}, \{\dot{q}_i(t)\})$ of time t, generalized coordinates $\{q_i(t)\}$, and generalized velocities $\{\dot{q}_i(t)\}$ is called the Lagrangian of the system.
- The integration is done over time t.

The Hamilton's principle is not constructive. It states that such functional — Action $\mathcal{A}[\{q_i(t)\}]$ — exists. We still need to construct this functional. This means, that for any system, after we have chosen the coordinates $\{q_i\}$, we need to be able to construct the Lagrangian $L(t, \{q_i(t)\}, \{\dot{q}_i(t)\})$.

25.5. Lagrangian.

Before I show how to construct the Lagrangian, I want to emphasize two important points:

- Lagrangian is not energy. We do not minimize energy. We do not even minimize the Lagrangian. We minimize action!
- Lagrangian is a function of generalized coordinates $\{q_i\}$ and generalized velocities $\{\dot{q}_i\}$. There must be no momenta in Lagrangian.

The Lagrangian is constructed by the following procedure:

- After we have chosen the generalized coordinates $\{q_i\}$ and assuming, that we know the generalized velocities $\{\dot{q}_i(t)\}$ we compute the kinetic energy of our system: $K(t, \{q_i\}, \{\dot{q}_i\})$ — it may or may not explicitly depend on time.
- We also compute the potential energy $U(t, \{q_i\})$ — it also may or may not explicitly depend on time.
- The Lagrangian then is given by:

$$L(t, \{q_i\}, \{\dot{q}_i\}) = K(t, \{q_i\}, \{\dot{q}_i\}) - U(t, \{q_i\}).$$

After we constructed the Lagrangian, we can write the equation of motion for each of generalized coordinates:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i}.$$

25.6. Examples.

25.6.1. Free fall down a vertical line.

- First step: CHOSE THE COORDINATES, and CHECK:
  - Are the coordinates complete? Do they completely describe the system? If you know the coordinates do you know the configuration of your system?
  - Are the coordinates independent? Do the any values of the coordinates chosen independently describe possible configuration of the system?
- We chose our standard y vertical coordinate, to describe the position of the body.
- The kinetic energy is $K = \frac{m\dot{y}^2}{2}$.
- The potential energy is $U = mgy$.
- The Lagrangian is

$$L(y, \dot{y}) = K - U = \frac{m\dot{y}^2}{2} - mgy.$$

- The Lagrange equation:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{y}} = \frac{\partial L}{\partial y}.$$

As $\frac{\partial L}{\partial \dot{y}} = m\dot{y}$, $\frac{\partial L}{\partial y} = -mg$ we have

$$m\ddot{y} = -mg.$$

25.6.2. Motion of a particle in an arbitrary potential $U(\vec{r})$.

- First step: CHOSE THE COORDINATES, and CHECK:
  - Are the coordinates complete? Do they completely describe the system? If you know the coordinates do you know the configuration of your system?
  - Are the coordinates independent? Do the any values of the coordinates chosen independently describe possible configuration of the system?
- We chose our standard Cartesian coordinates: x, y, z.
- The kinetic energy is $K = \frac{m\dot{\vec{r}}^2}{2}$.
- The potential energy is $U(\vec{r})$.
- The Lagrangian is

$$L = \frac{m\dot{\vec{r}}^2}{2} - U(\vec{r}) = \frac{m\dot{x}^2}{2} + \frac{m\dot{y}^2}{2} + \frac{m\dot{z}^2}{2} - U(x, y, z)$$

- The Lagrange equation for the component x is

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = \frac{\partial L}{\partial x}$$

As $\frac{\partial L}{\partial \dot{x}} = m\dot{x}$, $\frac{\partial L}{\partial x} = -\frac{\partial U}{\partial x}$ we have

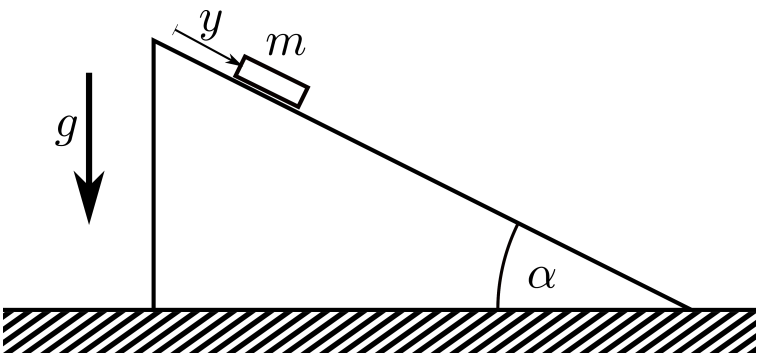
$$m\ddot{x} = -\frac{\partial U}{\partial x}.$$

- The same are for the other two components:

$$m\ddot{y} = -\frac{\partial U}{\partial y}, \qquad m\ddot{z} = -\frac{\partial U}{\partial z}.$$

- So we can write all these equations together
$$m\ddot{\vec{r}} = -\vec{\nabla}U.$$
This is Newton's equation $\vec{F} = m\vec{a}$! So we indeed reproduced the Newtonian dynamics!

25.6.3. A mass on a stationary wedge. No friction.



- First step: CHOSE THE COORDINATES, and CHECK:
  - Are the coordinates complete? Do they completely describe the system? If you know the coordinates do you know the configuration of your system?
  - Are the coordinates independent? Do the any values of the coordinates chosen independently describe possible configuration of the system?
- There is only one coordinate here y.
- The kinetic energy is $\frac{m\dot{y}^2}{2}$.
- The potential energy is $-mgy\sin\alpha$.
- The Lagrangian is $L = \frac{m\dot{y}^2}{2} + mgy\sin\alpha$.
- The Lagrange equation is
$$\frac{d}{dt}\frac{\partial L}{\partial \dot{y}} = \frac{\partial L}{\partial y}$$
As $\frac{\partial L}{\partial \dot{y}} = m\dot{y}$, $\frac{\partial L}{\partial y} = mg\sin\alpha$ we have
$$m\ddot{y} = mg\sin\alpha.$$
Notice, we did not need any forces to find this!

LECTURE 26
Lagrangian mechanics.

26.1. General strategy.

ONLY IF ALL THE FORCES ARE CONSERVATIVE!!!

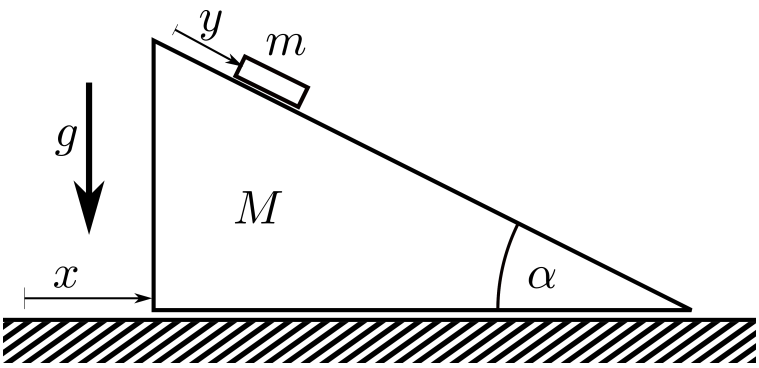
- Choose generalized coordinates $\{q_i\}$, $i = 1 \dots N$.
- Generalized coordinates:
  - A set of numbers which ambiguously describe the configuration of the system.
  - These numbers must be independent.
  - These numbers must provide the complete description.
- The number of generalized coordinates N is called the number of degrees of freedom.
- Write the total kinetic energy K of the system in terms of the generalized coordinates and their time derivatives: $\{q_i\}$ and $\{\dot{q}_i\}$.
- Write the total potential energy U in terms of the generalized coordinates $\{q_i\}$.
- Both kinetic and potential energy may or may not depend on time explicitly.
- Define the Lagrangian $L = K(t, \{\dot{q}_i\}, \{q_i\}) - U(t, \{q_i\})$.
- Write down the Lagrange equations for all/every generalized coordinates

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i}.$$

- Generally, these are coupled non-linear second order differential equations for the functions $q_i(t)$.
- The number of equations equals to the number of the coordinates q_i, the number of degrees of freedom N.
- As all equations are of the second order the general solution of the system of the equations will depend on $2N$ arbitrary constants.
- Set up the initial conditions for all generalized coordinates $\{q_i\}$ and generalized velocities $\{\dot{q}_i\}$. So each degree of freedom requires 2 initial conditions. The total number of initial conditions is $2N$.
- Solve the equations. Use THE INITIAL CONDITIONS to fix the arbitrary constants in the general solution!
- As the number of initial conditions $2N$ equals to the number of unknown constants $2N$, the equations for the constants will have a solution (but the solution is not necessarily unique.)

26.2. Examples.

26.2.1. A mass on a moving wedge. No friction.



- First step: CHOSE THE COORDINATES, and CHECK:
  - Are the coordinates complete? Do they completely describe the system? If you know the coordinates do you know the configuration of your system?
  - Are the coordinates independent? Do the any values of the coordinates chosen independently describe possible configuration of the system?
- The coordinates are x and y – see figure.
- The kinetic energy of the wedge is $\frac{M\dot{x}^2}{2}$.
- Let's compute the kinetic energy of the mass m.
  - In order to do that, we set up an auxiliary Cartesian coordinate system: X – horizontal axis and Y vertical axis pointing up.
  - The X coordinate of the mass m is $X = x + y\cos\alpha$, and the Y coordinate is $Y = -y\sin\alpha$.
  - For the mass m its horizontal velocity component $v_{\text{hor}} = \dot{X} = \dot{x} + \dot{y}\cos\alpha$.
  - Its vertical velocity component is $v_{\text{vert}} = \dot{Y} = -\dot{y}\sin\alpha$.
  - So its velocity squared is given $v^2 = v_{\text{hor}}^2 + v_{\text{vert}}^2 = \dot{x}^2 + \dot{y}^2 + 2\dot{x}\dot{y}\cos\alpha$.
  - And the kinetic energy of the mass m is $\frac{m}{2}\left(v_{\text{hor}}^2 + v_{\text{vert}}^2\right) = \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2 + 2\dot{x}\dot{y}\cos\alpha\right)$.
- So the total kinetic energy of the system is

$$K = \frac{M}{2}\dot{x}^2 + \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2 + 2\dot{x}\dot{y}\cos\alpha\right).$$

- The total potential energy is just the gravitational potential energy of mass m and it is given by $-mgy\sin\alpha$.
- The Lagrangian is

$$L = \frac{M+m}{2}\dot{x}^2 + \frac{m}{2}\dot{y}^2 + m\dot{x}\dot{y}\cos\alpha + mgy\sin\alpha.$$

- There are two Lagrange equations, for x and y. In order to derive them we compute
  - For x coordinate

$$\frac{\partial L}{\partial x} = 0, \qquad \frac{\partial L}{\partial \dot{x}} = M\dot{x} + m\dot{y}\cos\alpha$$

and

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = M\ddot{x} + m\ddot{y}\cos\alpha.$$

So the Lagrangian equation for y coordinate is

$$\frac{\partial L}{\partial \dot{x}} = M\ddot{x} + m\ddot{y}\cos\alpha = 0.$$

– For y coordinate

$$\frac{\partial L}{\partial y} = mg\sin\alpha, \qquad \frac{\partial L}{\partial \dot{y}} = m\dot{y} + m\dot{x}\cos\alpha$$

and

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{y}} = m\ddot{y} + m\ddot{x}\cos\alpha.$$

So the Lagrangian equation for y coordinate is

$$m\ddot{y} + m\ddot{x}\cos\alpha = mg\sin\alpha.$$

- The two second order differential equation together read

$$(M+m)\ddot{x} + m\ddot{y}\cos\alpha = 0$$
$$m\ddot{y} + m\ddot{x}\cos\alpha = mg\sin\alpha$$

They must be solved together and the initial conditions must be used to fix the arbitrary constants in the general solution.

26.2.2. A pendulum.

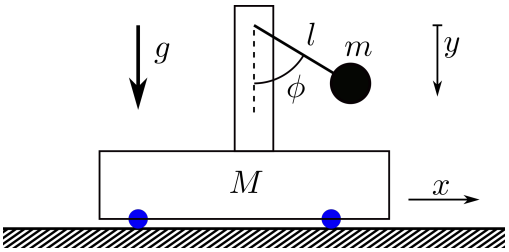
- The coordinate is ϕ — the angle the pendulum makes with the vertical line.
- The potential energy: $U(\phi) = mgl(1-\cos\phi)$.
- The kinetic energy is the rotational kinetic energy $K = \frac{I\omega^2}{2}$, where $I = ml^2$ — the moment of inertia, and $\omega = \dot{\phi}$ — angular velocity.
- The Lagrangian is

$$L = K - U = \frac{ml^2\dot{\phi}^2}{2} - mgl(1-\cos\phi).$$

- The Lagrange equation is

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} = \frac{\partial L}{\partial \phi}, \qquad \frac{\partial L}{\partial \dot{\phi}} = I\dot{\phi}, \quad \frac{\partial L}{\partial \phi} = -mgl\sin\phi$$
$$\ddot{\phi} = -\frac{g}{l}\sin\phi.$$

26.2.3. A pendulum on a cart.



- The coordinate x — the position of the Cart and ϕ — the angle of the pendulum are good generalized coordinates. A good test is the following:
  – If I know x and ϕ will I be able to draw the picture, or there will be pieces whose position I do not know?

  - If I know x can I chose ϕ arbitrarily from withing its domain? If I know ϕ can I chose x arbitrarily from withing its domain?
  - The first test is for completeness, the second is for independence.
- Next we need to find the kinetic energy K of the whole system through x, $\dot{x}$, ϕ, and $\dot{\phi}$. And the potential energy U of the whole system as a function of x and ϕ.
- The kinetic energy of the cart is $M\dot{x}^2/2$.
- To find the kinetic energy of the pendulum we need to find the velocity of the ball m through our generalized coordinates.
  - Again, we set up an auxiliary Cartesian coordinate system: X – horizontal axis and Y vertical axis pointing up.
  - The X position of the ball is $X = x + l\sin\phi$, the Y position of the ball is $Y = l\cos\phi$.
  - Then for the ball we have $v_X = \dot{X} = \dot{x} + \dot{\phi}l\cos\phi$, and $v_Y = \dot{Y} = -\dot{\phi}l\sin\phi$.
  - So $v^2 = v_X^2 + v_Y^2 = (\dot{x} + \dot{\phi}l\cos\phi)^2 + \dot{\phi}^2 l^2 \sin^2\phi$.
- The total kinetic energy is the sum of the two:

$$K = \frac{M\dot{x}^2}{2} + \frac{m}{2}\left(\dot{x}^2 + 2\dot{x}\dot{\phi}l\cos\phi + l^2\dot{\phi}^2\right).$$

- The potential energy is $U = -mgy_m = -mgl\cos\phi$. (In this case the potential energy does not depend on x. However, if, say, we had attached the cart to a wall with a spring, then we would have had the extra term in the potential energy — the potential energy of the spring — which depends on the cart's position.)
- The Lagrangian is

$$L = K - U = \frac{M\dot{x}^2}{2} + \frac{m}{2}\left(\dot{x}^2 + 2\dot{x}\dot{\phi}l\cos\phi + l^2\dot{\phi}^2\right) + mgl\cos\phi.$$

- We need to write two equations for x and ϕ.
  - For x we have:

$$\frac{\partial L}{\partial x} = 0, \quad \frac{\partial L}{\partial \dot{x}} = M\dot{x} + m\dot{x} + m\dot{\phi}l\cos\phi, \quad \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = M\ddot{x} + m\ddot{x} + m\ddot{\phi}l\cos\phi - m\dot{\phi}^2 l\sin\phi.$$

  - The first Lagrange equation is

$$M\ddot{x} + m\ddot{x} + m\ddot{\phi}l\cos\phi - m\dot{\phi}^2 l\sin\phi = 0.$$

  - For ϕ we have

$$\frac{\partial L}{\partial \phi} = -m\dot{x}\dot{\phi}l\sin\phi - mgl\sin\phi, \quad \frac{\partial L}{\partial \dot{\phi}} = m\dot{x}l\cos\phi + ml^2\dot{\phi}, \quad \frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} = m\ddot{x}l\cos\phi - m\dot{x}\dot{\phi}l\sin\phi + l^2\ddot{\phi}.$$

  - The second Lagrange equation is

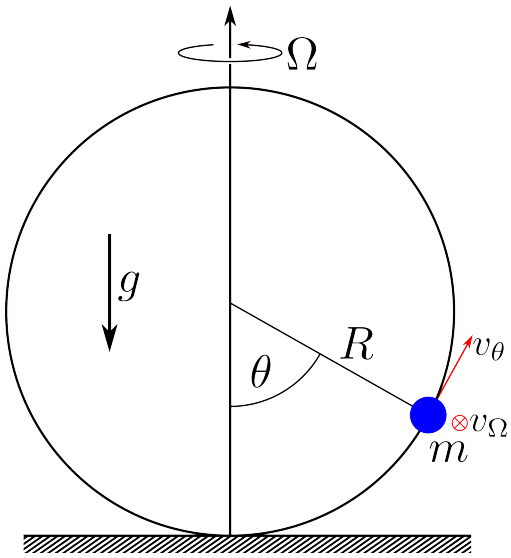
$$m\ddot{x}l\cos\phi - m\dot{x}\dot{\phi}l\sin\phi + ml^2\ddot{\phi} = -m\dot{x}\dot{\phi}l\sin\phi - mgl\sin\phi$$

- So the Lagrange equations are

$$M\ddot{x} + m\ddot{x} + m\ddot{\phi}l\cos\phi - m\dot{\phi}^2 l\sin\phi = 0$$
$$m\ddot{x}l\cos\phi + ml^2\ddot{\phi} = -mgl\sin\phi$$

LECTURE 27
Lagrangian mechanics.

27.1. Examples. A bead on a vertical rotating hoop.



We have a loop of radius R rotating with a constant and fixed(!) angular velocity Ω around a diameter in the vertical direction, see figure. There is a bead of mass m which can freely — without friction — move along the loop. There is gravity acting on the bead. We want to write the equations of motion for the system, analyze them, and see if we can learn something interesting.

"Something interesting" means that we want to learn some universal aspects. The aspects which do not depend on the details of the problem and can be used in developing intuition about more general and more complicated physical effects.

In particular, this problem illustrates a very general idea of spontaneous symmetry breaking. This idea is used very widely in physics. It is central for the Landau theory of the second order phase transitions. Such diverse phenomena as Higgs boson, magnetization in magnets, superfluidity, superconductivity, etc are all in the realm of this theory.

The phenomena mentioned above are quantum and as such requires a different machinery, but, remarkably this simple problem shows one of the most important aspects of all of them.

27.1.1. Equation of motion.

- The loop is rotating with the constant/fixed angular velocity Ω, so its motion is known and no equation required for it (Notice, that this would be different should the loop rotate freely, then its motion would be influenced by the motion of the bead and we would have to write the equations of motion for both the loop and the bead.)
- Ω is a parameter of the problem. We have full control over it.
- The position of the bead at any moment of time is then fully described by just one generalized coordinate — the angle θ.
- Lagrangian. We need potential and kinetic energies:

– The potential energy $U(\theta) = mgR(1 - \cos\theta)$.
– For the kinetic energy we notice, that the total vector velocity of the bead has two components v_θ — the velocity along the loop, and v_Ω — the velocity perpendicular to the plane of the loop, see figure. We also see that $v_\theta = R\dot\theta$, and $v_\Omega = \Omega R \sin\theta$. The two components are perpendicular to each other, the total velocity of the bead is $v^2 = R^2\dot\theta^2 + \Omega^2 R^2 \sin^2\theta$. The kinetic energy then is $K(\theta, \dot\theta) = \frac{m}{2}R^2\dot\theta^2 + \frac{m}{2}\Omega^2 R^2 \sin^2\theta$.

So the Lagrangian is

$$L = \frac{m}{2}R^2\dot\theta^2 + \frac{m}{2}\Omega^2 R^2 \sin^2\theta - mgR(1 - \cos\theta).$$

- Now we compute

$$\frac{\partial L}{\partial \dot\theta} = mR^2\dot\theta, \qquad \frac{\partial L}{\partial \theta} = m\Omega^2 R^2 \sin\theta\cos\theta - mgR\sin\theta.$$

- And the equation of motion is.

$$R\ddot\theta = (\Omega^2 R\cos\theta - g)\sin\theta.$$

27.1.2. Analysis of the motion.

The motion of the bead depends on the initial conditions. If one wants to know the full solution one has to set up initial conditions and then solve the equation of motion. This exact solution is fairly complicated and not very illuminating.

Instead we want to consider the motion around the equilibrium positions of the bead. We expect this motion to be a harmonic motion and have some universal features.

- At equilibrium the bead does not move, so $\dot\theta = 0$ and $\ddot\theta = 0$, so the right hand side of the equation of motion must be zero.
- There are four equilibrium points θ_{eq} the bead can remain stationary on the loop.

$$\sin\theta_{eq} = 0, \qquad \text{or} \qquad \cos\theta_{eq} = \frac{g}{\Omega^2 R}$$

- The first equation gives two equilibrium points $\theta_{eq} = 0$ or $\theta_{eq} = \pi$.
- Critical Ω_c. The second equation $\cos\theta_{eq} = \frac{g}{\Omega^2 R}$ gives two equilibrium points if and only if

$$\frac{g}{\Omega^2 R} < 1, \quad \text{or} \quad \Omega > \Omega_c \equiv \sqrt{g/R}, \qquad \cos\theta_{eq} = \frac{\Omega_c^2}{\Omega^2}.$$

27.1.3. Effective potential energy.

- The Lagrangian can be written as

$$L = \frac{m}{2}R^2\dot\theta^2 - U_{eff}(\theta),$$

then we can read the effective potential energy:

$$U_{eff}(\theta) = -mR^2\left(\frac{1}{2}\Omega^2\sin^2\theta + \Omega_c^2\cos\theta\right).$$

where I used $g = \Omega_c^2 R$ and ignored a constant in the potential energy.

- Notice, that the effective potential energy has a symmetry: $\theta \to -\theta$. If one changes the sign of θ the potential energy does not change!

- Using this effective potential energy the equation of motion becomes

$$mR^2\ddot{\theta} = -\frac{\partial U_{eff}}{\partial \theta}.$$

- One can check,
  - that the point $\theta_{eq} = \pi$ is always unstable — it is a maximum of the effective potential energy;
  - the point $\theta_{eq} = 0$ is stable for $\Omega < \Omega_c$, then $\theta_{eq} = 0$ is a minimum of the effective potential energy and is unstable for $\Omega > \Omega_c$, when it becomes the maximum of the effective potential energy;
  - the two points $\cos\theta_{eq} = \frac{\Omega_c^2}{\Omega^2}$ are stable (minima of the effective potential energy) when they exist.
- Notice, that the equilibriums $\theta = 0$ and $\theta = \pi$ are symmetric under $\theta \to -\theta$ — under this transformation each these points turn into themselves. However, the other two equilibriums are NOT symmetric, under this transformation they turn into each other.
- The most interesting regime is $\Omega \sim \Omega_c$. In this regime θ_{eq} is small. If we are interested in small oscillations about the equilibrium, then θ is also small.
- Assuming $\Omega \sim \Omega_c$ we are interested only in small θ. So

$$U_{eff}(\theta) \approx \frac{1}{2}mR^2(\Omega_c^2 - \Omega^2)\theta^2 + \frac{3}{4!}mR^2\Omega_c^2\theta^4$$

  Or, to make life easier and use $\Omega \approx \Omega_c$ we can write the same function as

$$U_{eff}(\theta) \approx mR^2\Omega_c(\Omega_c - \Omega)\theta^2 + \frac{3}{4!}mR^2\Omega_c^2\theta^4,$$

  One should notice, that there are two terms: one of the order of $(\Omega - \Omega_c)\theta^2$ and the other is of the order of θ^4. It seems unreasonable to keep only these terms and drop the rest. However, we will see below, that $\theta^2 \sim (\Omega - \Omega_c)$, so in fact both terms are of the same order $(\Omega - \Omega_c)^2$ and the rest of them are of the higher order.
- Spontaneous symmetry breaking. Plot the function $U_{eff}(\theta)$ for $\Omega < \Omega_c$, $\Omega = \Omega_c$, and $\Omega > \Omega_c$. Discuss universality.

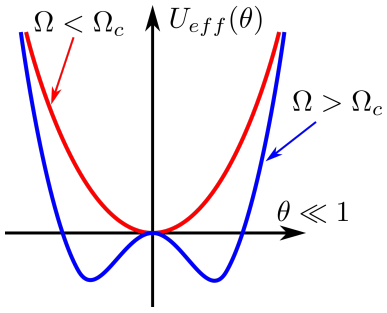


Figure 1. Effective potential for $|\theta| \ll 1$ and $\Omega \approx \Omega_c$.

27.1.4. Small oscillations about stable equilibria.

- $\Omega < \Omega_c$, the stable equilibrium position is at $\theta_0 = 0$. To study the mall oscillations around $\theta_0 = 0$, we write the equations of motion in the first order in θ

$$mR^2\ddot{\theta} = -mR^2(\Omega_c^2 - \Omega^2)\theta, \qquad \omega = \sqrt{\Omega_c^2 - \Omega^2} \approx \sqrt{2\Omega_c(\Omega_c - \Omega)}.$$

- $\Omega > \Omega_c$, the stable equilibrium position is at θ_0, $\cos\theta_0 = \Omega_c^2/\Omega^2$. To study the mall oscillations around θ_0, we write the equations of motion in the first order in $\theta - \theta_0$

$$U_{eff}(\theta) = -mR^2\left(\frac{1}{2}\Omega^2\sin^2\theta + \Omega_c^2\cos\theta\right).,$$

$$\frac{\partial U_{eff}}{\partial\theta} = -mR^2(\Omega^2\cos\theta - \Omega_c^2)\sin\theta, \qquad \frac{\partial^2 U_{eff}}{\partial\theta^2} = mR^2\Omega^2\sin^2\theta - mR^2\cos\theta(\Omega^2\cos\theta - \Omega_c^2)$$

$$\left.\frac{\partial U_{eff}}{\partial\theta}\right|_{\theta=\theta_0} = 0, \qquad \left.\frac{\partial^2 U_{eff}}{\partial\theta^2}\right|_{\theta=\theta_0} = -mR^2\left(\frac{\Omega_c^4}{\Omega^2} - \Omega^2\right)$$

So the Tylor expansion gives

$$U_{eff}(\theta \sim \theta_0) \approx \text{const} + \frac{1}{2}mR^2\left(\frac{\Omega_c^4}{\Omega^2} - \Omega^2\right)(\theta - \theta_0)^2$$

The frequency of small oscillations then is

$$\omega = \sqrt{\frac{\Omega_c^4}{\Omega^2} - \Omega^2} \approx 2\sqrt{\Omega_c(\Omega - \Omega_c)}.$$

27.1.5. Universality. Response.

- The effective potential energy for small θ and $|\Omega - \Omega_c|$

$$U_{eff}(\theta) = \frac{1}{2}a(\Omega_c - \Omega)\theta^2 + \frac{1}{4}b\theta^4.$$

- θ_0 for the stable equilibrium is given by $\partial U_{eff}/\partial\theta = 0$

$$\theta_0 = \begin{cases} 0 & \text{for} \quad \Omega < \Omega_c \\ \sqrt{\frac{a}{b}(\Omega - \Omega_c)} & \text{for} \quad \Omega > \Omega_c \end{cases}$$

Plot $\theta_0(\Omega)$. Non-analytic behavior at Ω_c.

- Response: how θ_0 responses to a small change in Ω.

$$\chi(\Omega) = \frac{\partial\theta_0}{\partial\Omega} = \begin{cases} 0 & \text{for} \quad \Omega < \Omega_c \\ \frac{1}{2}\sqrt{\frac{a}{b}}\frac{1}{\sqrt{(\Omega - \Omega_c)}} & \text{for} \quad \Omega > \Omega_c \end{cases}$$

Plot $\frac{\partial\theta_0}{\partial\Omega}$ vs Ω. The response *diverges* at Ω_c.

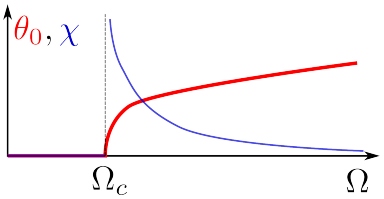
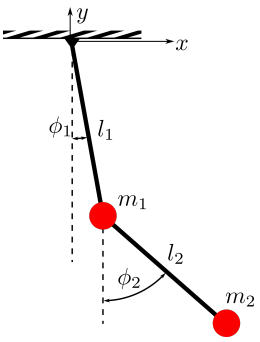


Figure 2. Stable equilibrium θ_0 as the function of Ω (red), and the response χ as the function of Ω (blue).

LECTURE 28
Lagrangian mechanics.

28.1. Example.



Here we consider one more example — a double pendulum. The strategy is same as always

- Choosing the generalized coordinates.
- Write the potential energy.
- Kinetic energy. Normally, most trouble for students.

Here the most natural choice of coordinates are the angles ϕ_1 and ϕ_2. It is also convenient to use the auxiliary x and y for the intermediate steps. So that we have

$$x_1 = l_1 \sin\phi_1, \qquad x_2 = l_1 \sin\phi_1 + l_2 \sin\phi_2$$
$$y_1 = -l_1 \cos\phi_1, \qquad y_2 = -l_1 \cos\phi_1 - l_2 \cos\phi_2$$

- Now we can write the potential energy

$$U = m_1 g y_1 + m_2 g y_2 = -(m_1 + m_2) g l_1 \cos\phi_1 - m_2 g l_2 \cos\phi_2$$

- In order to find the kinetic energy we need velocities

$$v_{1x} = \dot{x}_1 = l_1 \dot{\phi}_1 \cos\phi_1, \qquad v_{2x} = \dot{x}_2 = l_1 \dot{\phi}_1 \cos\phi_1 + l_2 \dot{\phi}_2 \cos\phi_2$$
$$v_{1y} = \dot{y}_1 = l_1 \dot{\phi}_1 \sin\phi_1, \qquad v_{2y} = \dot{y}_2 = l_1 \dot{\phi}_1 \sin\phi_1 + l_2 \dot{\phi}_2 \sin\phi_2$$

so

$$v_1^2 = v_{1x}^2 + v_{1y}^2 = l_1^2 \dot{\phi}_1^2$$
$$v_2^2 = v_{2x}^2 + v_{2y}^2 = l_1^2 \dot{\phi}_1^2 + l_2^2 \dot{\phi}_2^2 + 2 l_1 l_2 \dot{\phi}_1 \dot{\phi}_2 \cos(\phi_1 - \phi_2)$$

and the kinetic energy

$$K = \frac{(m_1 + m_2) l_1^2}{2} \dot{\phi}_1^2 + \frac{m_2 l_2^2}{2} \dot{\phi}_2^2 + m_2 l_1 l_2 \dot{\phi}_1 \dot{\phi}_2 \cos(\phi_1 - \phi_2)$$

- The Lagrangian then is

$$L = \frac{(m_1 + m_2) l_1^2}{2} \dot{\phi}_1^2 + \frac{m_2 l_2^2}{2} \dot{\phi}_2^2 + m_2 l_1 l_2 \dot{\phi}_1 \dot{\phi}_2 \cos(\phi_1 - \phi_2) + (m_1 + m_2) g l_1 \cos\phi_1 + m_2 g l_2 \cos\phi_2.$$

- Now we write the Lagrangian equations. We first compute the partial derivatives:

$$\frac{\partial L}{\partial \dot\phi_1} = (m_1+m_2)l_1^2\dot\phi_1 + m_2 l_1 l_2 \dot\phi_2 \cos(\phi_1-\phi_2), \qquad \frac{\partial L}{\partial \phi_1} = -m_2 l_1 l_2 \dot\phi_1\dot\phi_2 \sin(\phi_1-\phi_2) - (m_1+m_2) g l_1 \sin\phi_1$$

$$\frac{\partial L}{\partial \dot\phi_2} = m_2 l_2^2\dot\phi_2 + m_2 l_1 l_2 \dot\phi_1 \cos(\phi_1-\phi_2), \qquad \frac{\partial L}{\partial \phi_2} = +m_2 l_1 l_2 \dot\phi_1\dot\phi_2 \sin(\phi_1-\phi_2) - m_2 g l_2 \sin\phi_2$$

and then the full derivative for each

$$(m_1+m_2)l_1^2\ddot\phi_1 + m_2 l_1 l_2 \ddot\phi_2 \cos(\phi_1-\phi_2) + m_2 l_1 l_2 \dot\phi_2^2 \sin(\phi_1-\phi_2) = -(m_1+m_2) g l_1 \sin\phi_1$$
$$m_2 l_2^2\ddot\phi_2 + m_2 l_1 l_2 \ddot\phi_1 \cos(\phi_1-\phi_2) - m_2 l_1 l_2 \dot\phi_1^2 \sin(\phi_1-\phi_2) = -m_2 g l_2 \sin\phi_2$$

(some terms which appeared originally have canceled each other)

These are the equations of motion. They are second order coupled nonlinear differential equations. In order to complete them we need to supply also the initial conditions for both variables.

Such equations are hard to solve or analyze. Typically we are mainly interested in the small oscillations around the equilibrium position. In this case the equilibrium position is obvious: $\phi_{1,eq} = \phi_{2,eq} = 0$. So we need to linearize our equations around this point.

Linearizaton means that you only keep the linear terms in $\phi_1 - \phi_{1,eq}$ and in $\phi_2 - \phi_{2,eq}$ and their derivatives. In our case we then have

$$(m_1+m_2)l_1^2\ddot\phi_1 + m_2 l_1 l_2 \ddot\phi_2 = -(m_1+m_2) g l_1 \phi_1$$
$$m_2 l_2^2\ddot\phi_2 + m_2 l_1 l_2 \ddot\phi_1 = -m_2 g l_2 \phi_2$$

These are much simpler — they are still coupled, but at least they are linear! They can be solved by a simple Fourier transform.

28.2. Small Oscillations.

We will study the problem of small oscillation in the next semester. Here is just an overview.

A system will always have some dissipation. In many cases the dissipation can be considered to be very small. However, no matter how small it is if one waits long enough the system will find one of its equilibrium positions (there can be several.) Such equilibrium positions are the minimums of the potential energy. If $\{q_i\}$ are the set of N generalized coordinates and $U(\{q_i\})$ is the potential energy, then the equilibrium positions $\{q_{i,eq}\}$ are the solutions of N algebraic equations

$$\left.\frac{\partial U}{\partial q_i}\right|_{\{q_i = q_{i,eq}\}} = 0.$$

After one solves these equations, then for each solution one must make sure, that this is indeed the minimum, not the maximum or a saddle point.

In many cases the equilibrium position can be guessed form the problem itself, but not always!!! One has to be careful.

The Lagrangian equations of motion contain the derivative of the Lagrangian $\frac{\partial L}{\partial \dot q_i}$ and $\frac{\partial L}{\partial q_i}$. So in order for the equations of motion to be linear in the displacement of the generalized coordinates from the equilibrium positions $\{q_i - q_{i,eq}\}$ and generalized velocities one needs to write the Lagrangian in quadratic order in displacement of generalized coordinates and generalized velocities.

For example, for the problem of the double pendulum the equilibrium position is obvious $\phi_{1,eq} = \phi_{2,eq} = 0$. We can write the Lagrangian in the quadratic order in $\phi_1 - \phi_{1,eq} = \phi_1$, $\phi_2 - \phi_{2,eq} = \phi_2$ and in $\dot{\phi}_1$, $\dot{\phi}_2$.

$$L = \frac{(m_1 + m_2)l_1^2}{2}\dot{\phi}_1^2 + \frac{m_2 l_2^2}{2}\dot{\phi}_2^2 + m_2 l_1 l_2 \dot{\phi}_1\dot{\phi}_2 - \frac{1}{2}(m_1 + m_2)g l_1 \phi_1^2 - \frac{1}{2}m_2 g l_2 \phi_2^2.$$

(I dropped the constant terms from the Lagrangian.) One can see, that our linearized equations can be obtained from this Lagrangian right away, by the standard procedure.

LECTURE 29
Lagrangian mechanics.

29.1. Non uniqueness of the Lagrangian.

For any problem and any given set of generalized coordinates the Lagrangian is not uniquely defined. This is similar to the fact that the potential energy is not uniquely defined — one can always add a constant to it.

In the same way as two potential energy functions which differ only by a constant give the same equations of motion, two Lagrangians for the same problem must give the same equations of motion. So two Lagrangians are equivalent if the resulting Lagrangian equations are the same.

- Let's take a Lagrangian $L(\dot{q}, q, t)$.
- Let's take an arbitrary function $G(q, t)$.
- Let's construct a new Lagrangian $\tilde{L}(\dot{q}, q, t) = L(\dot{q}, q, t) + \dot{q}\frac{\partial G}{\partial q} + \frac{\partial G}{\partial t}$.
- The statement is that the two Lagrangians L and $\tilde{L}$ are equivalent. Equivalence means that the two Lagrangians result in exactly the same equation of motion.

29.1.1. Proof of equivalence.

- The Lagrange equation for the Lagrangian $\tilde{L}$ is

$$\frac{d}{dt}\frac{\partial \tilde{L}}{\partial \dot{q}} = \frac{\partial \tilde{L}}{\partial q}.$$

Let's use our definition of $\tilde{L}$ and see how it works

$$\frac{\partial \tilde{L}}{\partial \dot{q}} = \frac{\partial L}{\partial \dot{q}} + \frac{\partial G}{\partial q}, \qquad \frac{\partial \tilde{L}}{\partial q} = \frac{\partial L}{\partial q} + \dot{q}\frac{\partial^2 G}{\partial^2 q} + \frac{\partial^2 G}{\partial t \partial q}$$

then

$$\frac{d}{dt}\frac{\partial \tilde{L}}{\partial \dot{q}} = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} + \dot{q}\frac{\partial^2 G}{\partial^2 q} + \frac{\partial^2 G}{\partial q \partial t}$$

and we see

$$\frac{d}{dt}\frac{\partial \tilde{L}}{\partial \dot{q}} - \frac{\partial \tilde{L}}{\partial q} = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q}.$$

- So we see, that the equation we obtain using $\tilde{L}$ is exactly the same as the equation we obtain using L.

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q}.$$

29.1.2. The reason.

We want to understand why the above transformation of the Lagrangian does not change the equations of motion.

- The reason for this is the following: the expression I added to the Lagrangian $\dot{q}\frac{\partial G}{\partial q} + \frac{\partial G}{\partial t}$ is a full derivative $\dot{q}\frac{\partial G}{\partial q} + \frac{\partial G}{\partial t} = \frac{dG}{dt}$ as can be seen using the chain rule. So $\tilde{L} = L + \frac{dG}{dt}$. But then the Action changes by

$$\tilde{\mathcal{A}} = \int_{t_i}^{t_f} \tilde{L}dt = \int_{t_i}^{t_f} Ldt + \int_{t_i}^{t_f} \frac{dG}{dt}dt = \int_{t_i}^{t_f} Ldt + G(q(t_f), t_f) - G(q(t_i), t_i) = \mathcal{A} + \text{const}.$$

  So the variation of the Action does not change, and thus the condition for the extremum — the Euler-Lagrange equation — also does not change.

So one can always add a full time derivative to a Lagrangian.

The last statement is correct only in the classical mechanics. In quantum mechanics the Action itself has its own meaning (unlike the classical mechanics where we are only interested in its minimum.) and addition of a constant to the Action is not necessarily harmless.

29.2. Generalized momentum.

- Definition: For a coordinate q the generalized momentum is defined as

$$p \equiv \frac{\partial L}{\partial \dot{q}}$$

- Examples:
  - For a particle in a potential field $L = \frac{m\dot{\vec{r}}^2}{2} - U(\vec{r})$ we have

$$\vec{p} = \frac{\partial L}{\partial \dot{\vec{r}}} = m\dot{\vec{r}}$$

    The generalized momentum is just the usual momentum.
  - For a rotation around a fixed axis $L = \frac{I\dot{\phi}^2}{2} - U(\phi)$, then

$$p = \frac{\partial L}{\partial \dot{\phi}} = I\dot{\phi} = J.$$

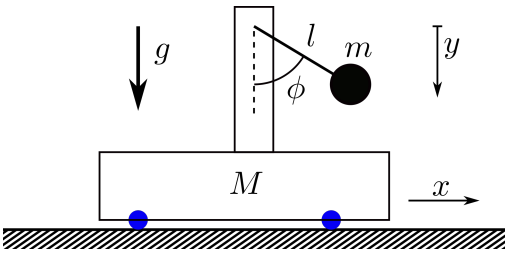
    The generalized momentum is just an angular momentum.

29.3. Ignorable coordinates. Conservation laws.

If one chooses the coordinates in such a way, that the Lagrangian does not depend on say one of the coordinates q_1 (but it still depends on $\dot{q}_1$, then the corresponding generalized momentum $p_1 = \frac{\partial L}{\partial \dot{q}_1}$ is conserved as

$$\frac{d}{dt}p_1 = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_1} = \frac{\partial L}{\partial q_1} = 0$$

29.3.1. A pendulum on a cart.



Problem of a freely horizontally moving cart of mass M with hanged pendulum of mass m and length l.

$$L = \frac{M\dot{x}^2}{2} + \frac{m}{2}\left(\dot{x}^2 + 2\dot{x}\dot{\phi}l\cos\phi + l^2\dot{\phi}^2\right) + mgl\cos\phi.$$

According to our definition of generalized momentum we can define two momenta p_ϕ and p_x:

$$p_\phi \equiv \frac{\partial L}{\partial \dot{\phi}} = ml^2\dot{\phi} + ml\dot{x}\cos\phi$$

$$p_x \equiv \frac{\partial L}{\partial \dot{x}} = (M+m)\dot{x} + m\dot{\phi}l\cos\phi.$$

We see right away, that there is no x (remember x and $\dot{x}$ are different variables for the Lagrangian) in the Lagrangian. So x is ignorable variable. It means, that the corresponding generalized momentum $p_x = \frac{\partial L}{\partial \dot{x}}$ is conserved. So we can write one of the equations of motion as

$$p_x = (M+m)\dot{x} + m\dot{\phi}l\cos\phi = \text{const}.$$

This constant should be obtained from the initial conditions.

29.4. Momentum conservation. Translation invariance

Let's consider a translationally invariant problem. For example all interactions depend only on the distance between the particles. The Lagrangian for this problem is $L(\vec{r}_1, \dots \vec{r}_i, \dot{\vec{r}}_1, \dots \dot{\vec{r}}_i, t)$. Then we add a constant vector ϵ to all coordinate vectors and define

$$L_\epsilon(\vec{r}_1, \dots \vec{r}_i, \dot{\vec{r}}_1, \dots \dot{\vec{r}}_i, t, \vec{\epsilon}) \equiv L(\vec{r}_1 + \vec{\epsilon}, \dots \vec{r}_i + \vec{\epsilon}, \dot{\vec{r}}_1, \dots \dot{\vec{r}}_i, t)$$

It is clear, that in the translationally invariant system the Lagrangian will not change under such a transformation. So we find

$$\frac{\partial L_\epsilon}{\partial \vec{\epsilon}} = 0.$$

But according to the definition

$$\left.\frac{\partial L_\epsilon}{\partial \vec{\epsilon}}\right|_{\vec{\epsilon}=0} = \sum_i \frac{\partial L}{\partial \vec{r}_i}.$$

Hence

$$\sum_i \frac{\partial L}{\partial \vec{r}_i} = 0.$$

On the other hand the Lagrange equations tell us that

$$\sum_i \frac{\partial L}{\partial \vec{r}_i} = \frac{d}{dt}\sum_i \frac{\partial L}{\partial \dot{\vec{r}}_i} = \frac{d}{dt}\sum_i \vec{p}_i,$$

so

$$\frac{d}{dt}\sum_i \vec{p}_i = 0, \qquad \sum_i \vec{p}_i = \text{const}.$$

We see, that the total momentum of the system is conserved!

29.5. Conservation laws from symmetry. For self-study.

Here I present a simplified version of Noether's theorem.

Let's assume that the Lagrangian is invariant under some continuous symmetry. It means the following: There is a parameter ϵ dependent transformation of the coordinates $\{q_i\}$ (for example a rotation around some axis by the angle ϵ, or translation of all coordinates by a constant vector, as was considered before) such that the Lagrangian has the same form in the new coordinate as in the old ones.

Let's consider the parameter ϵ to be infinitesimally small — this way we can keep only linear in ϵ terms. The transformation from the old $\{q_i\}$ to the new $\{\tilde{q}_i\}$ will have the following form

$$\tilde{q}_i = q_i + \epsilon f_i(\{q_1\}),$$

where f_i are some functions that define the transformation. (One can easily see, that these functions completely define the transformation for any finite ϵ, by simply constructing the differential equations $\frac{\partial q_i}{\partial \epsilon} = f_i(\{q_1\})$.

Our new Lagrangian is given by

$$L_\epsilon(\{q_i\},\{\dot{q}_i\}) \equiv L(\{q_i + \epsilon f_i(\{q_i\})\},\{\dot{q}_i + \epsilon \dot{q}_j \frac{\partial f_i(\{q_i\})}{\partial q_j}\}).$$

(The Einstein notations are used.) As by the condition the Lagrangian is symmetric — independent of ϵ, we can write

$$0 = \frac{\partial L_\epsilon}{\partial \epsilon} = f_i(\{q_i\})\frac{\partial L}{\partial q_i} + \dot{q}_j \frac{\partial f_i(\{q_i\})}{\partial q_j}\frac{\partial L}{\partial \dot{q}_i} = f_i(\{q_i\})\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} + \frac{\partial L}{\partial \dot{q}_i}\frac{d}{dt}f_i(\{q_i\}) = \frac{d}{dt}\left(f_i(\{q_i\})\frac{\partial L}{\partial \dot{q}_i}\right),$$

where we used the Lagrangian equations of motion. Thus we conclude, that there is a conservation law corresponding to our symmetry

$$\sum_i p_i f_i = \text{const}.$$

LECTURE 30
Lagrangian's equations for magnetic forces.

The equation of motion is

$$m\ddot{\vec{r}} = q(\vec{E} + \dot{\vec{r}} \times \vec{B})$$

The question is what Lagrangian gives such equation of motion?

30.1. Electric and magnetic fields.

In order to answer the question above we need to know a bit more about electric and magnetic fields. Classically these fields are completely described by the Maxwell equations. There are four of these equations and they are written in terms of "physical" fields: the electric field $\vec{E}(\vec{r}, t)$ and the magnetic field $\vec{B}(\vec{r}, t)$. We will need only two of the Maxwell equations: magnetic Gauss law and Faraday's law

$$\nabla \cdot \vec{B} = 0, \qquad \nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}.$$

Notice, that these are the two Maxwell equations which do not have matter — charge or current densities.

Consider first magnetic Gauss law. Which is the statement that there are no magnetic charges.

$$\nabla \cdot \vec{B} = 0$$

This equation is satisfied by the following solution

$$\vec{B} = \nabla \times \vec{A},$$

for any (smooth) vector field $\vec{A}(\vec{r}, t)$.

The Faraday's Law

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

then gives $\nabla \times \left(\vec{E} + \frac{\partial \vec{A}}{\partial t}\right) = 0$. Remembering that the curl of a gradient of any function is zero, we find

$$\vec{E} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t},$$

where ϕ is the electric potential and is again an arbitrary (smooth) function of $\vec{r}$ and t, $\phi(\vec{r}, t)$.

So instead solving four Maxwell equations for the fields $\vec{E}(\vec{r}, t)$ and $\vec{B}(\vec{r}, t)$, we can use the vector potential $\vec{A}(\vec{r}, t)$ and potential $\phi(\vec{r}, t)$ and then solve the remaining two equations

(with the boundary conditions!!) to obtain $\vec{A}(\vec{r},t)$ and $\phi(\vec{r},t)$. After that we can reconstruct the fields $\vec{E}(\vec{r},t)$ and $\vec{B}(\vec{r},t)$ by

$$\vec{E} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t},$$
$$\vec{B} = \nabla \times \vec{A}.$$

- The crucial observation: The vector potential $\vec{A}(\vec{r},t)$ and potential $\phi(\vec{r},t)$ are not uniquely defined.

One can take an arbitrary (smooth) function $F(\vec{r},t)$ and transform the potentials $\vec{A}$ and ϕ to $\vec{A}'$ and ϕ' in the following way:

$$\vec{A}' = \vec{A} + \nabla F, \qquad \phi' = \phi - \frac{\partial F}{\partial t}.$$

This transformation will not change the physical fields $\vec{E}(\vec{r},t)$ and $\vec{B}(\vec{r},t)$, as $\nabla \times A' = \nabla \times A$ and $-\nabla\phi' - \frac{\partial \vec{A}'}{\partial t} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t}$. The other two Maxwell equations contain only electric and magnetic fields (and not the potentials) so they will also not fix this freedom.

Moreover, in any experiment we can only measure electric $\vec{E}$ and magnetic $\vec{B}$ fields. This means that the potentials — vector potential $\vec{A}$ and scalar potential ϕ — cannot be measured by their own.

The transformation from one set of fields $\vec{A}$ and ϕ to another $\vec{A}'$ and ϕ'

$$\begin{matrix}\phi \\ \vec{A}\end{matrix} \longrightarrow \begin{matrix}\phi' = \phi - \frac{\partial F}{\partial t} \\ \vec{A}' = \vec{A} + \nabla F\end{matrix}$$

which leaves the "physical" fields $\vec{E}$ and $\vec{B}$ invariant is called **gauge transformation**.

The fields $\vec{A}$ and ϕ are called **gauge fields**. The freedom to chose any (smooth) function $F(\vec{r},t)$ is called **gauge freedom**. The fact, that no physical results must depend on the choice of gauge (physical quantities must be invariant under the gauge transformation) is called **gauge symmetry**.

Such gauge symmetries are extremely important in physics. A lot of constructions in modern physics involve some sort of gauge symmetry. The fields $\vec{A}$ and ϕ are called $U(1)$ gauge fields. There are many others.

As any continuous symmetry, gauge symmetry leads to conservation laws. In the case of electromagnetism it leads to the charge conservation law (we will not discuss it any further in this class).

- Important: If $\vec{B}$ and $\vec{E}$ are zero, the gauge fields do not have to be zero.

For example if $\vec{A}$ and ϕ are constants, $\vec{B} = 0$, $\vec{E} = 0$. Generally, if

$$\vec{A} = \nabla F, \qquad \phi = -\frac{\partial F}{\partial t},$$

for arbitrary (smooth) function $F(\vec{r},t)$, then $\vec{B} = 0$ and $\vec{E} = 0$.

30.2. The Lagrangian.

Now we can write the Lagrangian for a particle of mass m and charge q moving through given electric and magnetic fields.

$$L = \frac{m\dot{\vec{r}}^2}{2} - q(\phi(\vec{r},t) - \dot{\vec{r}}\cdot\vec{A}(\vec{r},t))$$

I note, that this Lagrangian has a simpler and more transparent form in the notations adopted in the special and general relativity — four dimensional space-time with Minkovskii metric.

- It is impossible to write the Lagrangian in terms of the physical fields $\vec{B}$ and $\vec{E}$!
- The expression which appears in the action $\int L dt$ is (I dropped q, as it does not matter for this consideration.)

$$\phi dt - d\vec{r}\cdot\vec{A}.$$

- Last lecture we found that if we have a full differential of a function in the integral in the action, then this function does not contribute to the equations of motion $\mathcal{A} = \cdots + \int_i^f dG = \cdots + G_f - G_i$ and the expression $G_f - G_i$ is a constant and drops out under variation.
- Now notice, that the expression $\phi dt - d\vec{r}\cdot\vec{A}$ is a full differential if and only if all "cross" derivatives equal to each other (Lecture 13)

$$-\nabla\phi - \frac{\partial\vec{A}}{\partial t} = 0, \qquad \nabla\times\vec{A} = 0.$$

  So if these conditions satisfied, then the term $q(\phi - \dot{\vec{r}}\cdot\vec{A})$ can be thrown out from the Lagrangian.
- Notice, that these conditions are exactly the conditions for the physical fields $\vec{E}$ and $\vec{B}$ be zero!

The generalized momenta are

$$\vec{P} = \frac{\partial L}{\partial\dot{\vec{r}}} = m\dot{\vec{r}} + q\vec{A}$$

(Notice, that the generalized momentum is not the same as usual momentum. Moreover, it is not gauge invariant! This is the reason I am using capital $\vec{P}$, so it will not be confused with the usual momentum $\vec{p}$)

The Lagrange equations are:

$$\frac{d}{dt}\vec{P} = \frac{\partial L}{\partial\vec{r}}$$

Let's consider the x component of the Lagrange equation

$$\frac{d}{dt}P_x = \frac{\partial L}{\partial x}, \qquad P_x = m\dot{x} + qA_x(x,y,z,t)$$

We have

$$\frac{d}{dt}P_x = m\ddot{x} + q\dot{x}\frac{\partial A_x}{\partial x} + q\dot{y}\frac{\partial A_x}{\partial y} + q\dot{z}\frac{\partial A_x}{\partial z} + q\frac{\partial A_x}{\partial t}$$
$$\frac{\partial L}{\partial x} = -q\frac{\partial\phi}{\partial x} + q\dot{x}\frac{\partial A_x}{\partial x} + q\dot{y}\frac{\partial A_y}{\partial x} + q\dot{z}\frac{\partial A_z}{\partial x}$$

and

$$m\ddot{x} + q\dot{x}\frac{\partial A_x}{\partial x} + q\dot{y}\frac{\partial A_x}{\partial y} + q\dot{z}\frac{\partial A_x}{\partial z} + q\frac{\partial A_x}{\partial t} = -q\frac{\partial \phi}{\partial x} + q\dot{x}\frac{\partial A_x}{\partial x} + q\dot{y}\frac{\partial A_y}{\partial x} + q\dot{z}\frac{\partial A_z}{\partial x}$$

$$m\ddot{x} = q\left(-\frac{\partial \phi}{\partial x} - \frac{\partial A_x}{\partial t} + \dot{y}\left[\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right] - \dot{z}\left[\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right]\right)$$

$$m\ddot{x} = q\left(E_x + \dot{y}B_z - \dot{z}B_y\right)$$

The other components can be obtained either by direct calculation or simply by cyclic permutation. The resulting three equations can be written as one single equation

$$m\ddot{\vec{r}} = q(\vec{E} + \dot{\vec{r}} \times \vec{B}).$$

LECTURE 31
Energy conservation.

31.1. Energy conservation.

We also have the time translation invariance in many systems. It means that the Lagrangian does not explicitly depend on time. So we have $L(\{q\},\{\dot{q}\})$, and not $L(\{q\},\{\dot{q}\},\mathbf{t})$. However, the generalized coordinates $\{q(t)\}$ do depend on time. So let's see how the Lagrangian on a trajectory depends on time.

Let me clarify the question. We consider the $1D$ motion. Assume that we have a generalized coordinate q and a Lagrangian $L(q,\dot{q})$. We then write Lagrangian equation of motion $\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q}$ with some initial conditions $q(t=0)=q_0$, $\dot{q}(t=0)=v_0$. For a given Lagrangian these initial conditions define a trajectory. Using these initial conditions we solve the equation of motion and obtain $q(t;q_0,v_0)$ and hence we also obtained $\dot{q}(t;q_0,v_0)=\frac{dq(t)}{dt}$. The last two arguments in functions q and $\dot{q}$ are to emphasize that these functions depend on the initial conditions. These arguments will be dropped in what follows. We then take these functions $q(t)$ and $\dot{q}(t)$ and plug them into the Lagrangian $L(q(t),\dot{q}(t))$ — this is what it means: the value of the Lagrangian on the trajectory. Now the Lagrangian becomes a function of time on the trajectory defined by q_0 and v_0. We want to see how it depends on time.

In our standard definition it means that we are interested in the full time derivative of the Lagrangian.

$$\frac{d}{dt}L(q(t),\dot{q}(t)) = \frac{\partial L}{\partial q}\dot{q} + \frac{\partial L}{\partial \dot{q}}\ddot{q} = \frac{\partial L}{\partial q}\dot{q} + \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\dot{q}\right) - \dot{q}\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\dot{q}\right) + \dot{q}\left(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right)$$

But as we are looking at the real trajectory — the function $q(t)$ is the solution of the Lagrange equation. So according to the Lagrange equation the last term is zero, so we have

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\dot{q} - L(q,\dot{q})\right) = 0$$

or

$$\frac{\partial L}{\partial \dot{q}}\dot{q} - L(q,\dot{q}) = \text{const} = E$$

Using generalized momentum we can write

$$p\dot{q} - L = E, \qquad \text{Constant on trajectory defined by } q_0 \text{ and } v_0.$$

so the value of the conserved quantity depends on the initial conditions $E(q_0,v_0)$, but it is constant during the motion.

If we have many variables q_i, then

$$E = \sum_i p_i \dot{q}_i - L$$

This is another conserved quantity. This conserved quantity is called energy.

If we perform exactly the same calculation as above, but with the Lagrangian which explicitly depends on time $L(q, \dot{q}, t)$, then the result will be

$$\frac{dE}{dt} = \frac{\partial L}{\partial t}.$$

- So if you have a Lagrangian, the only thing that you need to check is if it has explicit dependence on time or not. If there is no explicit dependence on time, then the energy is conserved!
- I want to emphasize once more, it will be very important later, the energy E is a number. This number is a constant on a trajectory, but it is different for different trajectories.

31.2. Examples:

31.2.1. A particle in a potential field.

- The Lagrangian

$$L = \frac{m\dot{\vec{r}}^2}{2} - U(\vec{r})$$

- The Lagrangian does not have explicit dependence on time t, so the energy is conserved.
- The momenta

$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x}, \qquad p_y = \frac{\partial L}{\partial \dot{y}} = m\dot{y}, \qquad p_z = \frac{\partial L}{\partial \dot{z}} = m\dot{z}.$$

- The Energy

$$E = \dot{x}p_x + \dot{y}p_y + \dot{z}p_z - L = \frac{m\dot{\vec{r}}^2}{2} + U(\vec{r})$$

31.2.2. A particle on a circle.

- The Lagrangian

$$L = \frac{mR^2}{2}\dot{\phi}^2 - U(\phi).$$

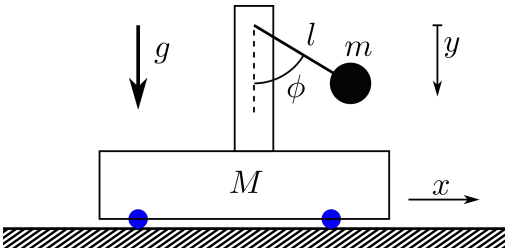
- The Lagrangian does not have explicit dependence on time t, so the energy is conserved.
- Generalized momentum

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = mR^2\dot{\phi}.$$

- The Energy

$$E = \dot{\phi}p_\phi - L = \frac{mR^2}{2}\dot{\phi}^2 + U(\phi)$$

31.2.3. A cart (mass M) with a pendulum (mass m, length l).



- The Lagrangian:

$$L = \frac{M+m}{2}\dot{x}^2 + m\dot{\phi}\dot{x}l\cos\phi + \frac{m}{2}l^2\dot{\phi}^2 - mgl(1-\cos\phi).$$

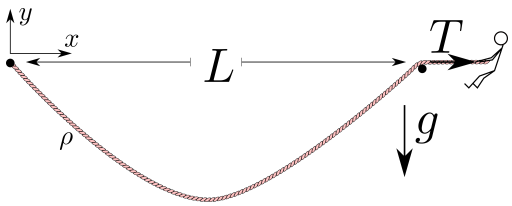
- The Lagrangian does not have explicit dependence on time t, so the energy is conserved.
- The generalized momenta:

$$p_x = \frac{\partial L}{\partial \dot{x}} = (M+m)\dot{x} + m\dot{\phi}l\cos\phi, \qquad p_\phi = \frac{\partial L}{\partial \dot{\phi}} = m\dot{x}l\cos\phi + ml^2\dot{\phi}.$$

- The Energy

$$E = \dot{x}p_x + \dot{\phi}p_\phi - L = \frac{M+m}{2}\dot{x}^2 + m\dot{\phi}\dot{x}l\cos\phi + \frac{m}{2}l^2\dot{\phi}^2 + mgl(1-\cos\phi)$$

31.2.4. A string with tension and gravity.



- The Functional

$$\int_0^L (\rho g y + T)\sqrt{1+(y')^2}\,dx.$$

- One can think of it as an Action of some mechanical system. Then for this system we identify the "Lagrangian"

$$L = (\rho g y + T)\sqrt{1+(y')^2}.$$

  We also use the letter x to denote the time in that mechanical system. The system is described by a "generalized coordinate" y and the "trajectory" is given by a function $y(x)$, which is the solution of the Euler-Lagrange equation with given boundary conditions. The "generalized velocity" is y'.
- The Lagrangian does not have explicit dependence on "time" x, so the "energy" is conserved.
- The "generalized momentum" of this system is

$$p = \frac{\partial L}{\partial y'} = \frac{\rho g y + T}{\sqrt{1+(y')^2}}y'.$$

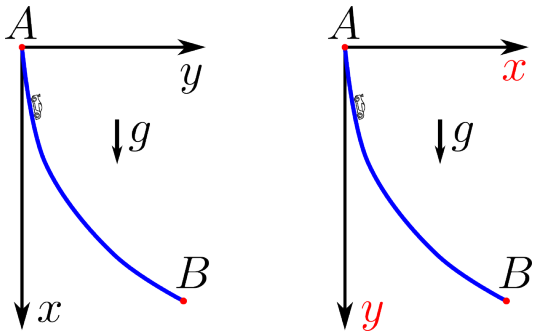
- And conserved "energy"

$$E = y'p - L = \frac{\rho g y + T}{\sqrt{1 + (y')^2}}.$$

This is now a first order differential equation which can be solved much easier, than the second order Euler-Lagrange equation.

- This conserved quantity has a physical meaning for the initial problem of the rope. It is the x component of the tension force.

31.2.5. Shortest time to fall — Brachistochrone.



In the Lecture 24 we considered the Brachistochron problem (find a path from A to B in the vertical plane which takes the least amount of time)

- The total time of travel between the points A and B along a given rail is

$$T = \int \frac{dl}{v}.$$

- In lecture 24 we used the system of coordinates with the x axis pointing down and y axis horizontal, as shown on the left panel of the figure.

  For this choice of the coordinates we have

$$T[y(x)] = \int_0^{x_B} \frac{\sqrt{1 + (y')^2}}{\sqrt{2gx}} dx.$$

We think of this functional as if it is the action of some mechanical system, then x is time, y is the coordinate, and the "Lagrangian" is

$$L(y, y', x) = \frac{\sqrt{1 + (y')^2}}{\sqrt{2gx}}.$$

We notice right away, that the "Lagrangian" does not depend on y, only on y'. So the Corresponding "momentum" is conserved:

$$p_y = \frac{\partial L}{\partial y'} = \frac{1}{\sqrt{2gx}} \frac{y'}{\sqrt{1 + (y')^2}} = \text{const}.$$

This is the first order differential equation which we then solved (see Lecture 24).

- Now lets use the coordinates with x as horizontal axis and y pointing down, as shown on the right panel of the figure.

For the rail given by $y(x)$ the time of travel now is

$$T[y(x)] = \int_0^{x_B} \frac{\sqrt{1+(y')^2}}{\sqrt{2gy}} dx.$$

We think of this functional as if it is the action of some mechanical system, then x is time, y is the coordinate, and the "Lagrangian" is

$$L(y, y', x) = \frac{\sqrt{1+(y')^2}}{\sqrt{2gy}}.$$

Now the "Lagrangian" does depend on the coordinate y, and hence the "momentum" is not conserved. However, the "Lagrangian" has no explicit dependence on "time" x. So the "Energy" is conserved. (This "Energy" has nothing to do with the true mechanical energy of the original problem. It is just the first integral.)

$$E = p_y y' - L = \frac{\partial L}{\partial y'} y' - L = \frac{1}{\sqrt{2gy}} \frac{y'^2}{\sqrt{1+y'^2}} - \frac{\sqrt{1+(y')^2}}{\sqrt{2gy}} = -\frac{1}{\sqrt{2gy}} \frac{1}{\sqrt{1+y'^2}} = \text{const}.$$

This is again the first order differential equation which can be solved to find $y(x)$. The result is the same as in Lecture 24 (with switched x and y).

Notice, that this example gives a hint that the energy and momentum conservation laws are related. It is indeed so in relativistic physics when there is a geometry (metric tensor) of the space-time.

LECTURE 32
Hamiltonian.

In this lecture we will construct a function of generalized momenta and coordinates, which is called Hamiltonian. In this lecture I will not describe how it is used — this will be done later. Here we just construct this function and consider a few examples. The most important fact:

- Hamiltonian of a given system is a FUNCTION of the generalized coordinates $\{q_i\}$ and generalized MOMENTA $\{p_i\}$!
- A Hamiltonian MUST NOT have generalized velocities $\{\dot{q}_i\}$ as arguments!

This is in contrast to Lagrangian

- Lagrangian of a given system is a FUNCTION of the generalized coordinates $\{q_i\}$ and generalized VELOCITIES $\{\dot{q}_i\}$!
- A Lagrangian MUST NOT have generalized momenta as arguments!

32.1. Hamiltonian.

In the previous lecture we studied the energy and energy conservation of a system described by a Lagrangian $L(\{q_i\}, \{\dot{q}_i\})$

- Given a Lagrangian $L(\{q_i\}, \{\dot{q}_i\})$ the energy

$$E = \sum_i p_i \dot{q}_i - L, \qquad p_i = \frac{\partial L}{\partial \dot{q}_i}$$

  is a number defined on a trajectory! One can say that it is a function of initial conditions.

Now we construct a function **a function!!** of p and q in the following way: starting from the Lagrangian $L(\{q_i\}, \{\dot{q}_i\})$ of N generalized coordinates $\{q_i\}$ and N generalized velocities $\{\dot{q}_i\}$ (it can also have explicit dependence on time $L(\{q_i\}, \{\dot{q}_i\}, t)$) we first write all N generalized momenta

$$p_i = \frac{\partial L}{\partial \dot{q}_i}.$$

We then treat these definitions as a set of N coupled generally non-linear set of N algebraic equations for all N generalized velocities $\{\dot{q}_i\}$. We solve these equations with respect to $\dot{q}_i$, we then have these functions

$$\dot{q}_i = \dot{q}_i(\{q_j\}, \{p_j\})$$

and define a function $H(\{q_i\},\{p_i\})$

$$H(\{q_i\},\{p_i\}) = \sum_i p_i\dot{q}_i(\{q_j\},\{p_j\}) - L(\{q_i\},\{\dot{q}_i(\{q_j\},\{p_j\})\}),$$

- Notice, that in this construction we have never used the equations of motion! we have treated q, $\dot{q}$ and p simply as variables, not as some functions of time.

This function is called a Hamiltonian! The Hamiltonian is a function of coordinates and momenta! THERE MUST BE NO VELOCITIES IN HAMILTONIAN!

- Hamiltonian is NOT energy. Energy is a number on a trajectory. Hamiltonian is a function of p and q — it, by itself, knows nothing about trajectories.
- Hamiltonian and energy are related to each other. The value of the Hamiltonian on a trajectory is energy.

The importance of variables:

- We have three kinds of variables:

generalized coordinates — q_i, generalized velocities — $\dot{q}_i$, generalized momenta — p_i.

- A Lagrangian is a function of generalized coordinates and **velocities**: q_i and $\dot{q}_i$. THERE MUST BE NO MOMENTA IN LAGRANGIAN!
- A Hamiltonian is a function of generalized coordinates and **momenta**: q_i and p_i. THERE MUST BE NO VELOCITIES IN HAMILTONIAN!

Here are the steps to get Hamiltonian from Lagrangian for a given system.

(a) Write down the Lagrangian $L(\{q_i\},\{\dot{q}_i\})$ – it is a function of generalized coordinates and velocities q_i, $\dot{q}_i$.
(b) Find <u>all</u> generalized momenta

$$p_i = \frac{\partial L}{\partial \dot{q}_i}.$$

(c) Treat the above definitions as equations and solve them for all $\dot{q}_i$, so for each velocity $\dot{q}_i$ you have an expression $\dot{q}_i = \dot{q}_i(\{q_j\},\{p_j\})$.
(d) Substitute these function $\dot{q}_i = \dot{q}_i(\{q_j\},\{p_j\})$ into the expression

$$\sum_i p_i\dot{q}_i - L(\{q_i\},\{\dot{q}_i\}).$$

The resulting function $H(\{q_i\},\{p_i\})$ of generalized coordinates and momenta is called Hamiltonian.

32.2. Examples.

32.2.1. A particle in a potential field.

- The Lagrangian

$$L = \frac{m\dot{\vec{r}}^2}{2} - U(\vec{r})$$

- The momenta

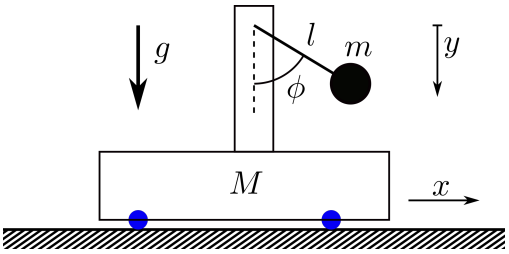
$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x}, \qquad p_y = \frac{\partial L}{\partial \dot{y}} = m\dot{y}, \qquad p_z = \frac{\partial L}{\partial \dot{z}} = m\dot{z}.$$

- The velocity
$$\dot{\vec{r}} = \frac{\vec{p}}{m}$$
- The Hamiltonian
$$H(\vec{r},\vec{p}) = \dot{x}p_x + \dot{y}p_y + \dot{z}p_z - L = \frac{\vec{p}^2}{m} - L = \frac{\vec{p}^2}{2m} + U(\vec{r})$$
- Check, that there are no velocities in the Hamiltonian.

32.2.2. A particle on a circle.

- The Lagrangian
$$L = \frac{mR^2}{2}\dot{\phi}^2 - U(\phi).$$
- Generalized momentum
$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = mR^2\dot{\phi}.$$
- The velocity
$$\dot{\phi} = \frac{p_\phi}{mR^2}$$
- The Hamiltonian
$$H(\phi, p_\phi) = \dot{\phi}p_\phi - L = \frac{p_\phi^2}{2mR^2} + U(\phi)$$
- Check, that there are no velocities in the Hamiltonian.

32.2.3. A cart (mass M) with a pendulum (mass m, length l).



- The Lagrangian:
$$L = \frac{M+m}{2}\dot{x}^2 + m\dot{\phi}\dot{x}l\cos\phi + \frac{m}{2}l^2\dot{\phi}^2 - mgl(1-\cos\phi).$$
- The generalized momenta $p_x = \frac{\partial L}{\partial \dot{x}}$ and $p_\phi = \frac{\partial L}{\partial \dot{\phi}}$:
$$p_x = (M+m)\dot{x} + m\dot{\phi}l\cos\phi,$$
$$p_\phi = m\dot{x}l\cos\phi + ml^2\dot{\phi}.$$
- We treat the above equations as a system of coupled equations for $\dot{x}$ and $\dot{\phi}$ and solve them.
- The generalized velocities $\dot{x}$ and $\dot{\phi}$ are then expressed through the generalized coordinates x and ϕ, and the generalized momenta p_x and p_ϕ:
$$\dot{x} = \frac{1}{l}\frac{lp_x - p_\phi\cos\phi}{M + m\sin^2\phi}, \qquad \dot{\phi} = \frac{1}{ml^2}\frac{(M+m)p_\phi - mlp_x}{M + m\sin^2\phi}.$$

NOTICE: There is no velocities in the right hand sides of these equations!

- The Hamiltonian

$$H = \dot{x}p_x + \dot{\phi}p_\phi - L = \frac{1}{2ml^2}\frac{ml^2p_x^2 - 2mlp_xp_\phi\cos\phi + (m+M)p_\phi^2}{M + m\sin^2\phi} + mgl(1-\cos\phi)$$

- Check, that there are no velocities in the Hamiltonian.
- The Hamiltonian is a function of the generalized coordinates x and ϕ, and the generalized momenta p_x and p_ϕ only. There is no velocities $\dot{x}$ and $\dot{\phi}$ in the Hamiltonian.

32.2.4. A particle in electro-magnetic field.

In the Lecture 30 we derived the Lagrangian for a particle of mass m and charge q moving through given electric and magnetic fields.

$$L = \frac{m\dot{\vec{r}}^2}{2} - q(\phi(\vec{r},t) - \dot{\vec{r}}\cdot\vec{A}(\vec{r},t))$$

We want to find the corresponding Hamiltonian.

- We first find the generalized momenta

$$\vec{P} = \frac{\partial L}{\partial\dot{\vec{r}}} = m\dot{\vec{r}} + q\vec{A}$$

  Important: notice the difference between the generalized momentum $\vec{P}$ and the "mechanical" momentum $\vec{p} = m\dot{\vec{r}} = \vec{P} + q\vec{A}$.
- Then we express the generalized velocity $\dot{\vec{r}}$ trough the generalized momenta

$$\dot{\vec{r}} = \frac{1}{m}\left(\vec{P} - q\vec{A}\right)$$

- Finally we compute the Hamiltonian

$$H(\vec{r},\vec{P}) = \vec{P}\cdot\dot{\vec{r}} - L\Big|_{\dot{\vec{r}}=\frac{1}{m}\left(\vec{P}+q\vec{A}\right)} = \frac{1}{2m}\left(\vec{P} - q\vec{A}\right)^2 + q\phi$$

- Check that there is no velocities in the Hamiltonian — only the generalized coordinates and generalized momenta!
- It is very tempting to use $\vec{p} = \vec{P} + q\vec{A}$ and write the Hamiltonian as $H = \frac{\vec{p}^2}{2m} + q\phi$. However, this is an incorrect form and WILL lead to many hard to find mistakes. The Hamiltonian MUST be written as a function of the generalized coordinates and generalized momenta!

32.2.5. Central symmetric potential in $3D$.

- We need to write the Lagrangian in spherical coordinates. We know

$$d\vec{r} = \vec{e}_r dr + \vec{e}_\theta r d\theta + \vec{e}_\phi r\sin\theta d\phi.$$

  Dividing this by dt we get

$$\vec{v} = \vec{e}_r\dot{r} + \vec{e}_\theta r\dot{\theta} + \vec{e}_\phi r\dot{\phi}\sin\theta,$$

  so

$$v^2 = \dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\phi}^2\sin^2\theta.$$

The Lagrangian is

$$L = \frac{m}{2}\dot{r}^2 + \frac{m}{2}r^2\dot{\theta}^2 + \frac{m}{2}r^2\dot{\phi}^2\sin^2\theta - U(r)$$

- The generalized momenta are

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r}, \qquad p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta}, \qquad p_\phi = \frac{\partial L}{\partial \dot{\phi}} = mr^2\dot{\phi}\sin^2\theta.$$

- The generalized velocities

$$\dot{r} = \frac{p_r}{m}, \qquad \dot{\theta} = \frac{p_\theta}{mr^2}, \qquad \dot{\phi} = \frac{p_\phi}{mr^2\sin^2\theta}.$$

- The Hamiltonian

$$H = \dot{r}p_r + \dot{\theta}p_\theta + \dot{\phi}p_\phi - L = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{p_\phi^2}{2mr^2\sin^2\theta} + U(r)$$

- Check, that there are no velocities in the Hamiltonian.

LECTURE 33
Hamiltonian equations.

33.1. Hamiltonian.

Here I just remind the construction we discussed last lecture.

- We start with a Lagrangian $L(\{q_i\}, \{\dot{q}_i\})$.
- We write ALL the generalized momenta

$$p_j = \frac{\partial L}{\partial \dot{q}_j}.$$

  for ALL variables.
- We treat these equations as equation for ALL $\dot{q}_i$. We solve these equations and find the functions $\dot{q}_j(\{q_i\}, \{p_i\})$ which express ALL generalized velocities through generalized coordinates and generalized momenta.
- We construct the Hamiltonian

$$H(\{q_i\}, \{p_i\}) = \sum_j p_j \dot{q}_j(\{q_i\}, \{p_i\}) - L(\{q_i\}, \{\dot{q}_j(\{q_i\}, \{p_i\})\}).$$

- The Hamiltonian is a function of generalized coordinates and generalized momenta only. There MUST be no generalized velocities in the Hamiltonian.
- If Lagrangian explicitly depends on time, then the procedure is exactly the same, but now time t will enter all the equations as a parameter. The Hamiltonian will then explicitly depend on time t.
- Mechanical system $\longrightarrow$ Lagrangian $\longrightarrow$ Hamiltonian. This procedure is well defined, so there is a Hamiltonian for any mechanical system.

33.2. New notations for partial derivatives.

Here I introduce new notation for the partial derivatives. The idea is to make it explicit what we keep fixed. Both Lagrangian and Hamiltonian have their "native" arguments, it is generalized coordinates and generalized VELOCITIES for the Lagrangian and generalized coordinates and generalized MOMENTA for the Hamiltonian. So by default when we take a partial derivative of a Lagrangian with respect to the velocity we keep all coordinates and other velocities fixed, the same for the Hamiltonian, if we take a partial derivative of a Hamiltonian with respect to a momentum, we keep fixed all coordinates and other momenta,

and so on. The need to specify and keep track of what we keep fixed arises when we start changing variables.

- The notation explicitly keeps the notion of what is kept fixed.
- The definition of momentum then is
$$p = \left(\frac{\partial L}{\partial \dot{q}}\right)_q .$$
The notation means that we differentiate the Lagrangian with respect to the variable $\dot{q}$ while keeping the variable q fixed.

  In the case of the Lagrangian this notation is overkill, as the Lagrangian only depends on q and $\dot{q}$, so the partial derivative with respect to $\dot{q}$ automatically assumes, that q is fixed. However, these notations become very useful when we start changing variables — they provide a device to keep track of what is kept fixed at every differentiation. As a side note these notations are extremely useful when one studies thermodynamics.
- The Lagrangian equation of motion using these notations is:
$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right)_q = \left(\frac{\partial L}{\partial q}\right)_{\dot{q}} .$$
- The definition of the generalized momentum is
$$p \equiv \left(\frac{\partial L}{\partial \dot{q}}\right)_q .$$
etc.

33.3. Hamiltonian equations.

Now we derive the Hamiltonian equations of motion. We will do this in $1D$ (more precise for one degree of freedom), but the extension to arbitrary number of dimensions (arbitrary number of degrees of freedom) will be obvious.

33.3.1. First Hamiltonian equation.

- Let's differentiate the Hamiltonian $H(p,q)$ with respect to momentum p, while keeping the coordinate q fixed.
- We will use $H = p\dot{q} - L(q,\dot{q})$, but we will remember, that $\dot{q}$ is the function of p and q, i.e. $\dot{q}(p,q)$.
- So we differentiate the function
$$H(p,q) = p\dot{q}(p,q) - L(q,\dot{q}(p,q))$$
with respect to p.
- Notice, that p appears in three places. When we differentiate with respect to p we need to differentiate over all of them using chain rule. So we will have three terms.
- We will also remember, that by definition $p = \left(\frac{\partial L}{\partial \dot{q}}\right)_q$.
- So we have:
$$\left(\frac{\partial H}{\partial p}\right)_q = \dot{q} + p\left(\frac{\partial \dot{q}}{\partial p}\right)_q - \left(\frac{\partial L}{\partial \dot{q}}\right)_q \left(\frac{\partial \dot{q}}{\partial p}\right)_q = \dot{q} + p\left(\frac{\partial \dot{q}}{\partial p}\right)_q - p\left(\frac{\partial \dot{q}}{\partial p}\right)_q = \dot{q}$$

This is the first Hamiltonian equation:

$$\dot{q} = \frac{\partial H}{\partial p}.$$

Here I dropped the explicit notion that in this differentiation q must be kept fixed. I dropped it, because the Hamiltonian depends only on p and q, so partial derivative of the Hamiltonian with respect to momentum p automatically assumes, that q is kept fixed.

- Notice, that deriving this equation we have not used the equation of motion, only the definition of momentum.

33.3.2. Second Hamiltonian equation.

- Now lets differentiate the Hamiltonian with respect to q, while keeping p fixed.
- Again we must remember that $\dot{q}(p,q)$ is the function of p and q.
- So using

$$H(p,q) = p\dot{q}(p,q) - L(q,\dot{q}(p,q))$$

we see, that q appears in three places. When we differentiate with respect to q we need to differentiate over all of them using chain rule. So we will have three terms. we have

$$\left(\frac{\partial H}{\partial q}\right)_p = p\left(\frac{\partial \dot{q}}{\partial q}\right)_p - \left(\frac{\partial L}{\partial q}\right)_{\dot{q}} - \left(\frac{\partial L}{\partial \dot{q}}\right)_q \left(\frac{\partial \dot{q}}{\partial q}\right)_p = -\left(\frac{\partial L}{\partial q}\right)_{\dot{q}} + \left(p - \left(\frac{\partial L}{\partial \dot{q}}\right)_q\right)\left(\frac{\partial \dot{q}}{\partial q}\right)_p.$$

- Using the definition of momentum $p = \left(\frac{\partial L}{\partial \dot{q}}\right)_q$, we see, that the last term is zero. So we have

$$\left(\frac{\partial H}{\partial q}\right)_p = -\left(\frac{\partial L}{\partial q}\right)_{\dot{q}}.$$

- According to the Lagrangian equation of motion $\left(\frac{\partial L}{\partial q}\right)_{\dot{q}} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right)_q = \dot{p}$. The last equality comes from the definition of momentum. So we have the second Hamiltonian equation:

$$\frac{\partial H}{\partial q} = -\dot{p}.$$

Again, I dropped the explicit notion that in this differentiation p must be kept fixed. I dropped it, because the Hamiltonian depends only on p and q, so partial derivative of the Hamiltonian with respect to momentum q automatically assumes, that p is kept fixed.

- Notice, that deriving this equation we DID use the equation of motion.

33.3.3. Both equations.

The two Hamiltonian equation together are

$$\dot{q} = \frac{\partial H}{\partial p},$$
$$\dot{p} = -\frac{\partial H}{\partial q}.$$

- Notice the minus sign in the second equation! This minus sign is very significant! Without this minus sign these are NOT Hamiltonian equations!
- Notice, that the equations are "self-contained" there is no notion of the generalized velocities. Everything is written in terms of the coordinates, momenta and their time dependence.
- If we have many degrees of freedom, then this pair of equations is written for each degree of freedom.
- So each degree of freedom gives us a pair of FIRST order differential equations (coupled, non-linear) So the number of initial conditions must be twice for each degree of freedom — the initial coordinate and the initial momentum. The same as for the Lagrangian formulation, where it was initial coordinate and initial velocity for each degree of freedom.

33.4. Examples.

33.4.1. A particle in a potential field.

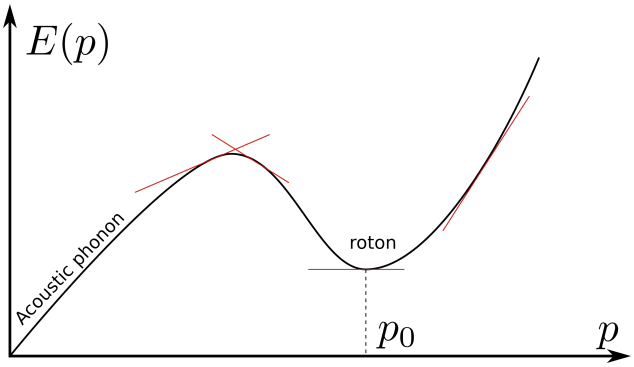
- Lagrangian→Hamiltonian.
  - The Lagrangian
$$L(\dot{\vec{r}},\vec{r}) = \frac{m\dot{\vec{r}}^2}{2} - U(\vec{r}).$$
  - The momentum
$$\vec{p} = \frac{\partial L}{\partial \dot{\vec{r}}} = m\dot{\vec{r}}.$$
  - The velocity
$$\dot{\vec{r}} = \frac{\vec{p}}{m}$$
  - The Hamiltonian
$$H(\vec{p},\vec{r}) = \vec{p}\cdot\dot{\vec{r}} - L = \frac{\vec{p}^2}{2m} + U(\vec{r})$$
- The equations of motion:
  - The Lagrangian equations of motion
$$m\ddot{\vec{r}} = -\frac{\partial U}{\partial \vec{r}}.$$
  - The Hamiltonian equations of motion
$$\dot{\vec{r}} = \frac{\partial H}{\partial \vec{p}} = \frac{\vec{p}}{m}, \qquad \dot{\vec{p}} = -\frac{\partial H}{\partial \vec{r}} = -\frac{\partial U}{\partial \vec{r}}.$$
  - Taking the time derivative of the first equation we find $\dot{\vec{p}} = m\ddot{\vec{r}}$. Using this in the second equation we find
$$m\ddot{\vec{r}} = -\frac{\partial U}{\partial \vec{r}}.$$
  - We see, that the Hamiltonian and Lagrangian equations give the same $\vec{r}(t)$!

33.4.2. Energy conservation.

- Energy is the value of the Hamiltonian on the trajectory!!!!
- What it means, is that we take a Hamiltonian, write the Hamilton equations, solve them for some initial conditions $q(t=0)=q_0$, $p(t=0)=p_0$ (and so forth if we have more degrees of freedom). We then have two functions $q(t)$ and $p(t)$.
- We now substitute these functions into the Hamiltonian $H(p,q,t)$ and obtain a function of time $H(p(t),q(t),t)$.
- Now lets differentiate this function with respect to time. This is a full derivative now
$$\frac{dH}{dt}=\frac{\partial H}{\partial p}\dot{p}+\frac{\partial H}{\partial q}\dot{q}+\frac{\partial H}{\partial t}=-\frac{\partial H}{\partial p}\frac{\partial H}{\partial q}+\frac{\partial H}{\partial q}\frac{\partial H}{\partial p}+\frac{\partial H}{\partial t}=\frac{\partial H}{\partial t}$$
where we used the Hamilton equations for the functions $p(t)$ and $q(t)$.
- So we see, that if the Hamiltonian does not explicitly depend on time (in less words $\frac{\partial H}{\partial t}=0$), then
$$\frac{dH}{dt}=0.$$
or the value of the Hamiltonian on a trajectory is constant.
- Notice the importance of the minus sign in the second of the Hamilton equations!

33.4.3. Velocity.

- In many cases the Hamiltonian is the starting point.
- The dependence of the velocity on momentum is then given by the Hamilton equation
$$\dot{q}=\frac{\partial H}{\partial p}.$$
- In particular if we have a normal "kinetic energy" $E(p)=\frac{p^2}{2m}$, then this equations gives
$$\dot{x}=\frac{\partial E}{\partial p}=p/m.$$
This is the usual $p=mv$.
- The kinetic energy as a function of momentum $E(p)$ is called dispersion relation.
- It is the dispersion relation which gives the relation between the velocity and momentum, by the Hamilton equation.
- There are cases where this is very nontrivial. For example in liquid Helium the dispersion of "exitations" is similar to the one shown in the picture. One can see, that at $p=p_0$ the momentum is not zero (it is p_0), but the velocity is zero!



LECTURE 34
Hamiltonian equations. Examples

The Hamiltonian and Lagrangian formulations of mechanics are equivalent to each other. Namely, if we know the Lagrangian we will know the Hamiltonian and if we know the Hamiltonian we will know the Lagrangian.

34.1. Lagrangian→Hamiltonian, Hamiltonian→Lagrangian.

34.1.1. $L \to H$

- We are given a Lagrangian $L(\{q_i\}, \{\dot{q}_i\})$ as a function of coordinates $\{q_i\}$ and velocities $\{\dot{q}_i\}$. There are no momenta in the Lagrangian!
- We write the definition of momenta
$$p_i = \frac{\partial L}{\partial \dot{q}_i}.$$
- We treat these equations as equations for all velocities $\{\dot{q}_i\}$ and solve them with respect to the velocities
$$\dot{q}_j = \dot{q}_j(\{p_i\}, \{q_i\}).$$
- We construct the Hamiltonian
$$H(\{p_i\}, \{q_i\}) = \sum_j p_j \dot{q}_j(\{p_{i'}\}, \{q_{i'}\}) - L(\{q_{j'}\}, \{\dot{q}_{j'}(\{p_{i'}\}, \{q_{i'}\})\}).$$

The Hamiltonian thus constructed is the function of all coordinates $\{q_i\}$ and all momenta $\{p_i\}$. There are must be no velocities in the Hamiltonian!

34.1.2. $H \to L$

- We are given a Hamiltonian $H(\{p_i\}, \{q_i\})$ as a function of all coordinates $\{q_i\}$ and all momenta $\{p_i\}$. There are no velocities in the Hamiltonian!
- We write the definition of velocity for each momentum
$$\dot{q}_i = \frac{\partial H}{\partial p_i}.$$
- We treat these equations as equations for all momenta $\{p_i\}$ and solve them with respect to the momenta
$$p_j = p_j(\{q_i\}, \{\dot{q}_i\}).$$

- We construct the Lagrangian

$$L(\{q_i\},\{\dot{q}_i\}) = \sum_j \dot{q}_j p_j(\{q_{i'}\},\{\dot{q}_{i'}\}) - H(\{p_{j'}(\{q_{i'}\},\{\dot{q}_{i'}\})\},\{q_{j'}\})$$

The Lagrangian thus constructed is the function of all coordinates $\{q_i\}$ and all velocities $\{\dot{q}_i\}$. There are must be no momenta in the Lagrangian!

34.1.3. Equations of motion.

If we have a Lagrangian and a Hamiltonian which are connected by the procedures described above, then the Lagrangian and Hamiltonian equations are equivalent — they describe the same motion!

$$\begin{array}{ccc} L(\{q_i\},\{\dot{q}_i\}) & \Longleftrightarrow & H(\{p_i\},\{q_i\}) \\ \dfrac{d}{dt}\dfrac{\partial L}{\partial \dot{q}_i} = \dfrac{\partial L}{\partial q_i} & \Longleftrightarrow & \begin{array}{l}\dot{q}_i = \frac{\partial H}{\partial p_i} \\ \dot{p}_i = -\frac{\partial H}{\partial q_i}\end{array} \\ q_i(t=0) = q_{i0}, \quad \dot{q}_i(t=0) = v_{i0} & \Longleftrightarrow & q_i(t=0) = q_{i0}, \quad p_i(t=0) = p_{i0}. \end{array}$$

Given equivalent initial conditions these equations will give exactly the same $q_i(t)$! Equivalent means that $p_{i0} = p_i(\{q_{i0}\},\{\dot{q}_{i0}\})$, or $\dot{q}_{i0} = \dot{q}(\{q_{i0}\},\{p_{i0}\})$.

34.2. Examples.

34.2.1. A particle in a potential field.

- Lagrangian→Hamiltonian.
  - The Lagrangian
$$L(\dot{\vec{r}},\vec{r}) = \frac{m\dot{\vec{r}}}{2} - U(\vec{r}).$$
  - The momentum
$$\vec{p} = \frac{\partial L}{\partial \dot{\vec{r}}} = m\dot{\vec{r}}.$$
  - The velocity
$$\dot{\vec{r}} = \frac{\vec{p}}{m}$$
  - The Hamiltonian
$$H(\vec{p},\vec{r}) = \vec{p}\cdot\dot{\vec{r}} - L = \frac{\vec{p}^2}{2m} + U(\vec{r})$$
- Hamiltonian→Lagrangian.
  - From the Hamiltonian
$$\dot{\vec{r}} = \frac{\partial H}{\partial \vec{p}} = \frac{\vec{p}}{m}.$$
  - The momentum
$$\vec{p} = m\dot{\vec{r}}.$$
  - The Lagrangian
$$L(\dot{\vec{r}},\vec{r}) = \dot{\vec{r}}\cdot\vec{p} - H = \frac{m\dot{\vec{r}}}{2} - U(\vec{r}).$$

- We found the Hamiltonian from the Lagrangian and then from Hamiltonian we found the same Lagrangian.
- The equations of motion:
  - The Lagrangian equations of motion
$$m\ddot{\vec{r}} = -\frac{\partial U}{\partial \vec{r}}.$$
  - The Hamiltonian equations of motion
$$\dot{\vec{r}} = \frac{\partial H}{\partial \vec{p}} = \frac{\vec{p}}{m}, \qquad \dot{\vec{p}} = -\frac{\partial H}{\partial \vec{r}} = -\frac{\partial U}{\partial \vec{r}}.$$
  - Taking the time derivative of the first equation we find $\dot{\vec{p}} = m\ddot{\vec{r}}$. Using this in the second equation we find
$$m\ddot{\vec{r}} = -\frac{\partial U}{\partial \vec{r}}.$$
  - We see, that the Hamiltonian and Lagrangian equations give the same $\vec{r}(t)$!

34.2.2. Rotation around a fixed axis.

- Lagrangian→Hamiltonian.
$$L(\dot{\phi}, \phi) = \frac{I\dot{\phi}^2}{2} - U(\phi).$$
  - The momentum
$$p_\phi = I\dot{\phi}.$$
  - Velocity
$$\dot{\phi} = \frac{p_\phi}{I}$$
  - The Hamiltonian
$$H(p_\phi, \phi) = p_\phi\dot{\phi} - L = \frac{p_\phi^2}{2I} + U(\phi).$$
- Hamiltonian→Lagrangian.
  - The velocity
$$\dot{\phi} = \frac{\partial H}{\partial p_\phi} = \frac{p_\phi}{I}.$$
  - The momentum
$$p_\phi = I\dot{\phi}.$$
  - The Lagrangian
$$L(\dot{\phi}, \phi) = p_\phi\dot{\phi} - H = \frac{I\dot{\phi}^2}{2} - U(\phi)$$
- The equations of motion
  - Lagrangian equation
$$I\ddot{\phi} = -\frac{\partial U}{\partial \phi}$$

– The Hamiltonian equations

$$\dot{\phi} = \frac{\partial H}{\partial p_\phi} = \frac{p_\phi}{I}, \qquad \dot{p}_\phi = -\frac{\partial H}{\partial \phi} = -\frac{\partial U}{\partial \phi}$$

Differentiating the first equation with respect to time and using the result in the second equation we get

$$I\ddot{\phi} = -\frac{\partial U}{\partial \phi}$$

– We see, that the Hamiltonian and Lagrangian equations give the same $\phi(t)$!

An example of the system considered above is a pendulum.

34.2.3. Motion in a central symmetric field.

- Lagrangian→Hamiltonian.
  – We need to write the Lagrangian in spherical coordinates. We know

$$d\vec{r} = \vec{e}_r dr + \vec{e}_\theta r d\theta + \vec{e}_\phi r \sin\theta d\phi.$$

Dividing this by dt we get

$$\vec{v} = \vec{e}_r \dot{r} + \vec{e}_\theta r\dot{\theta} + \vec{e}_\phi r\dot{\phi}\sin\theta,$$

so

$$v^2 = \dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\phi}^2\sin^2\theta.$$

  – The Lagrangian is

$$L = \frac{m}{2}\dot{r}^2 + \frac{m}{2}r^2\dot{\theta}^2 + \frac{m}{2}r^2\dot{\phi}^2\sin^2\theta - U(r)$$

  – The momenta:

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r}, \qquad p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta}, \qquad p_\phi = \frac{\partial L}{\partial \dot{\phi}} = mr^2\dot{\phi}\sin^2\theta.$$

  – The velocities

$$\dot{r} = \frac{p_r}{m}, \qquad \dot{\theta} = \frac{p_\theta}{mr^2}, \qquad \dot{\phi} = \frac{p_\phi}{mr^2\sin^2\theta}.$$

  – The Hamiltonian

$$H = \dot{r}p_r + \dot{\theta}p_\theta + \dot{\phi}p_\phi - L = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{p_\phi^2}{2mr^2\sin^2\theta} + U(r).$$

- Hamiltonian→Lagrangian.
  – The velocities

$$\dot{r} = \frac{\partial H}{\partial p_r} = \frac{p_r}{m}, \qquad \dot{\theta} = \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{mr^2}, \qquad \dot{\phi} = \frac{\partial H}{\partial p_\phi} = \frac{p_\phi}{mr^2\sin^2\theta}.$$

  – The momenta

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r}, \qquad p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta}, \qquad p_\phi = \frac{\partial L}{\partial \dot{\phi}} = mr^2\dot{\phi}\sin^2\theta.$$

  – The Lagrangian

$$L = \dot{r}p_r + \dot{\theta}p_\theta + \dot{\phi}p_\phi - H = \frac{m}{2}\dot{r}^2 + \frac{m}{2}r^2\dot{\theta}^2 + \frac{m}{2}r^2\dot{\phi}^2\sin^2\theta - U(r)$$

- The equations of motion

– The Lagrangian equations of motion

$$m\ddot{r} = mr\dot{\theta}^2 + mr\dot{\phi}^2\sin^2\theta - \frac{\partial U}{\partial r}$$
$$mr^2\ddot{\theta} + 2mr\dot{r}\dot{\theta} = mr^2\dot{\phi}^2\sin\theta\cos\theta$$
$$m\ddot{\phi}r^2\sin^2\theta + m\dot{\phi}r\dot{r}\sin^2\theta + 2r^2\dot{\phi}\dot{\theta}\sin\theta\cos\theta = 0$$

– The Hamiltonian equations of motion

$$\dot{r} = \tfrac{\partial H}{\partial p_r} = \tfrac{p_r}{m} \qquad \dot{p}_r = -\frac{\partial H}{\partial r} = \frac{p_\theta^2}{mr^3} + \frac{p_\phi^2}{mr^3\sin^2\theta} - \frac{\partial U}{\partial r}$$
$$\dot{\theta} = \tfrac{\partial H}{\partial p_\theta} = \tfrac{p_\theta}{mr^2} \qquad \dot{p}_\theta = -\frac{\partial H}{\partial \theta} = \frac{p_\phi^2\cos\theta}{mr^2\sin^3\theta}$$
$$\dot{\phi} = \tfrac{\partial H}{\partial p_\phi} = \tfrac{p_\phi}{mr^2\sin^2\theta}. \qquad \dot{p}_\phi = -\frac{\partial H}{\partial \phi} = 0$$

– You are welcome to check that these equations are equivalent to the Lagrangian equations.

34.2.4. Relativistic particle.

Consider a Hamiltonian $H(p,x) = \epsilon(p) = c\sqrt{p^2 + m_0^2c^2}$ in $1D$. We do not consider any field so the Hamiltonian does not depend on x.

- Equations of motion.

$$\dot{p} = -\frac{\partial H}{\partial x} = 0, \qquad \dot{x} = \frac{\partial H}{\partial p} = \frac{cp}{\sqrt{p^2 + m_0^2c^2}}$$

  So we see, that the momentum is conserved, but the velocity has a nontrivial dependence on momentum. In particular if $p \to \infty$ we have $\dot{x} \to c$. Moreover, the velocity can never exceed c!
- The momentum. From the last equation

$$p = \frac{m_0\dot{x}}{\sqrt{1 - \dot{x}^2/c^2}}.$$

  Notice, if we introduce a "mass" as $m = \frac{m_0}{\sqrt{1-\dot{x}^2/c^2}}$, then we have $p = m\dot{x}$ – the usual formula.
- If we use this p, substitute it into the Hamiltonian, and use our notation for m, then we get

$$E = \frac{m_0c^2}{\sqrt{1 - \dot{x}^2/c^2}} = mc^2.$$

- Lagrangian.

$$L(\dot{x}, x) = \dot{x}p - H = -m_0c\sqrt{c^2 - \dot{x}^2}$$

- Action. It is very instructive to write the Action for this problem

$$\mathcal{S} = -m_0c\int\sqrt{c^2 - \dot{x}^2}dt = -m_0c\int\sqrt{(cdt)^2 - (dx)^2}.$$

- Geometrical meaning of Action. Notice, that the action above is the length of the interval in the space-time (ct, x) with the metric $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$:

$$(ds)^2 = (cdt, dx) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} cdt \\ dx \end{pmatrix}$$

  The Action then is

$$\mathcal{A} = -m_0 c \int ds$$

- One now can easily extend this construction to the full $3 + 1$ space by using the Minkovskii metric

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

- Moreover, one is not restricted to the flat Minkovskii space and can write the Action for a particle in a curved space-time — the space-time with Einstein's gravity.

LECTURE 35
Hamiltonian equations. Examples. Phase space.

35.1. Examples.

- In the general case of a problem in Newtonian world the kinetic energy is quadratic in generalized velocities. It means that very generally the Lagrangian in generalized coordinates $\{q_i\}$ can be written in the following form.

$$L = \frac{1}{2}\dot{q}_i M_{ij}(\mathbf{q})\dot{q}_j - U(\mathbf{q}),$$

(Einstein notations are assumed) where $M_{ij}(\mathbf{q})$ is a symmetric $\mathbf{q}$-dependent positive definite matrix and I use $\mathbf{q}$ to denote the collection of all generalized coordinates $\{q_i\}$

  - The momenta

$$p_i = \frac{\partial L}{\partial \dot{q}_i} = M_{ij}(\mathbf{q})\dot{q}_j.$$

  - The velocities (matrix $\hat{M}$ is positive definite, so it has an inverse $\hat{M}^{-1}$)

$$\dot{q}_i = \left(\hat{M}^{-1}(\mathbf{q})\right)_{ij} p_j.$$

  - The Hamiltonian

$$H = \frac{1}{2} p_i \left(\hat{M}^{-1}(\mathbf{q})\right)_{ij} p_j + U(\mathbf{q})$$

  - The Hamiltonian equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i} = \left(\hat{M}^{-1}(\mathbf{q})\right)_{ij} p_j$$

$$\dot{p}_k = -\frac{\partial H}{\partial q_k} = -\frac{1}{2} p_i \frac{\partial \left(\hat{M}^{-1}(\mathbf{q})\right)_{ij}}{\partial q_k} p_j - \frac{\partial U(\mathbf{q})}{\partial q_k}$$

    Derivative of a matrix $\frac{\partial \left(\hat{M}^{-1}(\mathbf{q})\right)_{ij}}{\partial q_k}$ means simply the matrix where each matrix element is the derivative of the original matrix elements $\left(\hat{M}^{-1}(\mathbf{q})\right)_{ij}$.

- A cart (mass M) with a pendulum (mass m, length l).

$$L = \frac{M+m}{2}\dot{x}^2 + m\dot{\phi}\dot{x}l\cos\phi + \frac{m}{2}l^2\dot{\phi}^2 - mgl(1-\cos\phi).$$

This is a particular case of the example above.

– The generalized coordinates are x and ϕ. For generalized velocities we use

$$\begin{pmatrix} \dot{x} \\ \dot{\phi} \end{pmatrix}$$

– The Lagrangian then can be written as

$$L = \frac{1}{2}\left(\dot{x}, \dot{\phi}\right) \begin{pmatrix} M+m & ml\cos\phi \\ ml\cos\phi & ml^2 \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{\phi} \end{pmatrix} - mgl(1-\cos\phi)$$

– The matrix $\hat{M}$

$$\hat{M} = \begin{pmatrix} M+m & ml\cos\phi \\ ml\cos\phi & ml^2 \end{pmatrix}$$

– The inverse $\hat{M}^{-1}$ (an inverse of a 2×2 real symmetric matrix $\hat{A} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ is $\hat{A}^{-1} = \frac{1}{\det \hat{A}} \begin{pmatrix} c & -b \\ -b & a \end{pmatrix}$)

$$\hat{M}^{-1} = \frac{1}{ml^2}\frac{1}{M+m\sin^2\phi} \begin{pmatrix} ml^2 & -ml\cos\phi \\ -ml\cos\phi & m+M \end{pmatrix}$$

– The Hamiltonian

$$H = \begin{pmatrix} p_x, p_\phi \end{pmatrix} \hat{M}^{-1} \begin{pmatrix} p_x \\ p_\phi \end{pmatrix} + mgl(1-\cos\phi)$$

or

$$H = \frac{1}{2ml^2}\frac{ml^2p_x^2 - 2mlp_xp_\phi\cos\phi + (m+M)p_\phi^2}{M+m\sin^2\phi} + mgl(1-\cos\phi).$$

– etc.

35.2. Phase space. Hamiltonian vector field. Phase trajectories.

Hamiltonian equations are *the first order* differential equations! We double the number of variables and the number of equations, but each equation is now the first order differential equations. We still need two initial conditions for each degree of freedom.

- The space of all q and all p is called a *phase space* of the Hamiltonian system.
- The Hamiltonian is just a function on the *phase space*.
- Notice that we have inverted the story. The phase space comes first, only then we define a Hamiltonian as a function on the phase space. This change of perspective allows one to study the properties of the phase space itself, without referring to any Hamiltonian.

Let's consider a one dimensional problem with time independent Hamiltonian. So we have only one generalized coordinate q. The phase space is then two dimensional: (q,p). For a given Hamiltonian the equations of motion are

$$\dot{q} = \frac{\partial H}{\partial p}$$

$$\dot{p} = -\frac{\partial H}{\partial q}$$

Let's assume that a system had a phase space coordinates (q_t, p_t) at time t. The equations of motion show that at time $t + dt$ the system will be at the point

$$q_{t+dt} = q_t + \frac{\partial H}{\partial p} dt$$
$$p_{t+dt} = p_t - \frac{\partial H}{\partial q} dt$$

Let's now define the Hamiltonian vector field by

$$\vec{\mathcal{H}} = \begin{pmatrix} \frac{\partial H}{\partial p} \\ -\frac{\partial H}{\partial q} \end{pmatrix}.$$

Then we see, that a point (q_t, p_t) after time dt shifts to

$$\begin{pmatrix} q_{t+dt} \\ p_{t+dt} \end{pmatrix} = \begin{pmatrix} q_t \\ p_t \end{pmatrix} + \vec{\mathcal{H}} dt$$

So the vector $\vec{\mathcal{H}}$ is a vector of *velocity in the phase space*.

- We can compute the vector $\vec{\mathcal{H}}$ for any number of degrees of freedom.
- We can plot the vector field $\vec{\mathcal{H}}$ at every point of the phase space.
- Notice, that we do not need to solve any differential equations for that. We just need to differentiate the Hamiltonian!
- This vector field will show the *velocity in the phase space* at every point of the *phase space* of our system.

The trajectories of the system in the phase space are simply the lines which are tangential to the Hamiltonian vector field at every point of the line. Different trajectories correspond to different initial conditions.

This construction is very similar to the electric field and electric field lines.

- Motion in the phase space: we can consider the motion of a system in the phase space: we start from an initial point (q_i, p_i) and continue along the Hamiltonian vector field — along phase space trajectories.
- Trajectories do not intersect (except in isolated singular points). This is the same as for electric field lines. The phase space trajectories (electric field lines) can have one tangential vector at each point, except the points where $\vec{\mathcal{H}} = 0$ — the singular points — all the derivatives of the Hamiltonian are zero.
- On the phase trajectories the Hamiltonian is constant — the energy is conserved!

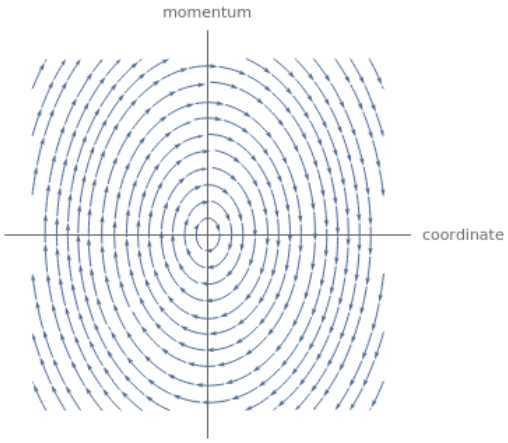
These simple rules allow one to construct the phase space trajectories for many (especially in $1D$) systems. Here are the couple of examples.

- Harmonic oscillator. The generalized coordinate is the coordinate x, the generalized momentum is p. The *phase space* is the collection of points (x, p).
  - The Hamiltonian of the Harmonic oscillator is
  $$H = \frac{p^2}{2m} + \frac{m\omega^2 x^2}{2}.$$
  - On the phase space trajectories the Hamiltonian is constant. The lines in (x, p) space are given by
  $$\frac{p^2}{2m} + \frac{m\omega^2 x^2}{2} = E$$

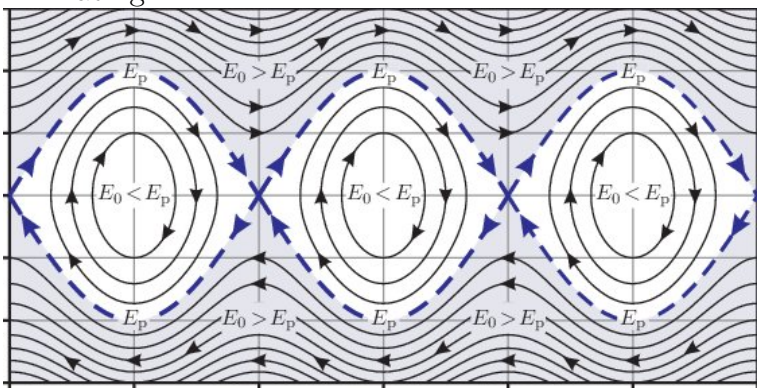
are ellipses with the semiaxes $\sqrt{2mE}$ and $\sqrt{2E/m\omega^2}$. (The area of these ellipses is $2\pi E/\omega = ET$, where T is the period.)

– The Hamiltonian vector field is

$$\vec{\mathcal{H}} = \begin{pmatrix} p/m \\ -m\omega^2 x \end{pmatrix}$$



- Pendulum. The generalized coordinate is the angle ϕ, the generalized momentum is p_ϕ. The *phase space* is the collection of points (ϕ, p_ϕ).
  – When energy is small the pendulum is a harmonic oscillator, so for small energies the trajectories are ellipses.
  – When energy grows the ellipses grow.
  – Eventually the ellipse must hit a singular point – this is when the energy of the pendulum is enough to reach the highest point.
  – If we increase the energy further the pendulum starts to rotate instead of oscillating.



LECTURE 36
Liouville's theorem. Poincaré recurrence theorem. Area law.

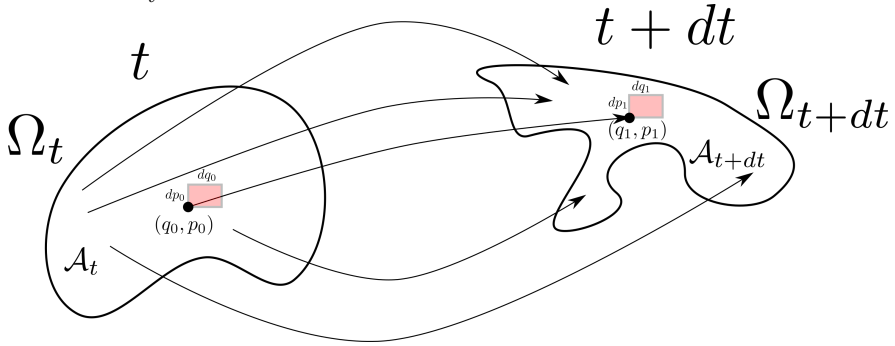
- Student's evaluation.

Before we discuss the Liouville's theorem I want to make a simple observation.

- As the Hamiltonian equations are the first order equations, any point of the phase space can be considered as the initial point of a trajectory.
- For any trajectory, any point of this trajectory can be taken as the initial point for the rest of the trajectory.
- What it means is that we do not need to know the motion in the phase space in the past in order to predict the motion in the phase space in the future.
- Compare this to the motion in the real (coordinate) space.

36.1. Liouville's theorem.

For a system of arbitrary number of degrees of freedom, consider a chunk Ω_{init} of the phase space which has a volume $\mathcal{A}_{\text{init}}$. Let's take every point of this chunk Ω_{init} as the initial condition for the Hamiltonian equations with the Hamiltonian H. After some time t every initial point of the chunk Ω_{init} will flow to some "final" point of the phase space. The collection of all these "final" points will make a chunk Ω_t of the phase space. This new chunk will have the volume $\mathcal{A}_t$.



Theorem: The phase space volume is conserved under the Hamiltonian flow. In other words $\mathcal{A}_t = \mathcal{A}_{\text{init}}$ for any t.

Proof: I show the proof for the case of one degree of freedom. But the theorem works for any number of degrees of freedom. The proof for the arbitrary number of degrees of freedom is almost identical to the one presented below.

As the trajectories do not intersect, we can consider the Hamiltonian flow as a map of the phase space on itself: any initial point $(q_{\text{init}}, p_{\text{init}})$ is mapped to a point $(q_0, p_0) = (q(t), p(t))$ after time t, where $q(t)$ and $p(t)$ are the solutions of the Hamiltonian equations with $(q_{\text{init}}, p_{\text{init}})$ as initial conditions. At this stage the initial chunk Ω_{init} became Ω_t. Let's consider what happens to the chunk Ω_t after a small additional time interval dt.

After the small time interval dt:

- The chunk Ω_t is mapped to the chunk Ω_{t+dt}.
- The boundary of the chunk Ω_t is mapped to the boundary of the chunk Ω_{t+dt}.
- A point (q_0, p_0) from the chunk Ω_t is mapped to a point (q_1, p_1) in the chunk Ω_{t+dt}.

This map $(q_0, p_0) \to (q_1, p_1)$ is given by:

$$q_1 = q_0 + \frac{\partial H}{\partial p_0} dt, \qquad p_1 = p_0 - \frac{\partial H}{\partial q_0} dt$$

- Notice, that in the linear in dt order, the derivatives on the right hand sides must be computed at the point (q_0, p_0), the derivatives above mean $\frac{\partial H}{\partial p_0} \equiv \frac{\partial H}{\partial p}\Big|_{p=p_0, q=q_0}$, etc.
- So in this picture in the linear is dt order, we fix t and dt and consider the above equations as the equations for the change of variables from (q_0, p_0) to $(q_1(q_0, p_0), p_1(q_0, p_0))$.

Consider a piece of volume at time t: $\mathcal{A}_t = \int_{\Omega_t} dq_0 dp_0$. After time dt, this volume becomes $\mathcal{A}_{t+dt} = \int_{\Omega_{t+dt}} dq_1 dp_1$.

$$\mathcal{A}_{t+dt} = \int_{\Omega_{t+dt}} dq_1 dp_1 = \int_{\Omega_t} \left(\frac{\partial q_1}{\partial q_0} \frac{\partial p_1}{\partial p_0} - \frac{\partial q_1}{\partial p_0} \frac{\partial p_1}{\partial q_0} \right) dq_0 dp_0$$

where

$$J \equiv \frac{\partial q_1}{\partial q_0} \frac{\partial p_1}{\partial p_0} - \frac{\partial q_1}{\partial p_0} \frac{\partial p_1}{\partial q_0}$$

is the Jacobian of the change of variables $(q_0, p_0) \to (q_1, p_1)$.

Using the formulas for our change of variables we find

$$\begin{aligned} \frac{\partial q_1}{\partial q_0} &= 1 + \frac{\partial^2 H}{\partial p_0 \partial q_0} dt, \\ \frac{\partial p_1}{\partial q_0} &= -\frac{\partial^2 H}{\partial q_0^2} dt, \end{aligned} \qquad \text{and} \qquad \begin{aligned} \frac{\partial q_1}{\partial p_0} &= \frac{\partial^2 H}{\partial p_0^2} dt, \\ \frac{\partial p_1}{\partial p_0} &= 1 - \frac{\partial^2 H}{\partial p_0 \partial q_0} dt. \end{aligned}$$

- Notice the minus sign in the bottom expression on the right.

We can now compute the Jacobian

$$J = \left(\frac{\partial q_1}{\partial q_0} \frac{\partial p_1}{\partial p_0} - \frac{\partial q_1}{\partial p_0} \frac{\partial p_1}{\partial q_0} \right) = 1 - \left(\frac{\partial^2 H}{\partial p_0 \partial q_0} \frac{\partial^2 H}{\partial p_0 \partial q_0} - \frac{\partial^2 H}{\partial p_0^2} \frac{\partial^2 H}{\partial q_0^2} \right) (dt)^2.$$

- Notice that the Jacobian has no term linear in dt.
- We used our equations of motion only up to a linear in dt terms.
- This means, that the quadratic in dt term is computed incorrectly, but the statement that the linear in dt term cancels out is exact.

So we can only write

$$J = 1 - D(p_0, q_0)(dt)^2,$$

where $D(p_0, q_0)$ is some unknown function.

Then we have

$$\mathcal{A}_{t+dt} = \int_{\Omega_t} dq_0 dp_0 - (dt)^2 \int_{\Omega_t} D(p_0, q_0) dq_0 dp_0,$$

The first term in this expression is the volume $\mathcal{A}_t$ of Ω_t. So the change of this volume $d\mathcal{A} = \mathcal{A}_{t+dt} - \mathcal{A}_t$ during the infinitesimally small time increment dt is

$$d\mathcal{A} = \mathcal{A}_{t+dt} - \mathcal{A}_t = -(dt)^2 \int_{\Omega_t} D(p_0, q_0) dq_0 dp_0.$$

What is important in this expression is that $d\mathcal{A} \sim (dt)^2$, or in other words, there is no linear in dt term, so that $d\mathcal{A}$ starts with the quadratic in dt term. It means, that $\frac{d\mathcal{A}}{dt} \sim dt$, so when we take the limit $dt \to 0$ we get

$$\frac{d\mathcal{A}}{dt} = 0 \quad \text{at ANY } t\text{!!!!, so} \quad \mathcal{A} = \text{const}.$$

(Calculus is great!!!)

- This is the Liouville's theorem. It states, that the volume of phase space is unchanged under the map on itself induced by the equations of motion for ANY Hamiltonian!
- It is also correct for any number of degrees of freedom.
- Notice the importance of the minus sign in the Hamiltonian equations.

36.2. Poincaré recurrence theorem.

If the available phase space for the system is finite. Let's starts the motion at some point of the phase space. Let's consider an evolution of some finite but small neighborhood of this point. The volume of the neighborhood is constant, so eventually as time progresses it will cover all of the available volume. Then the tube of the trajectories must intersect itself. But it cannot, as trajectories do not intersect. It means that it must return to the starting neighborhood (or intersect it at least partially.)

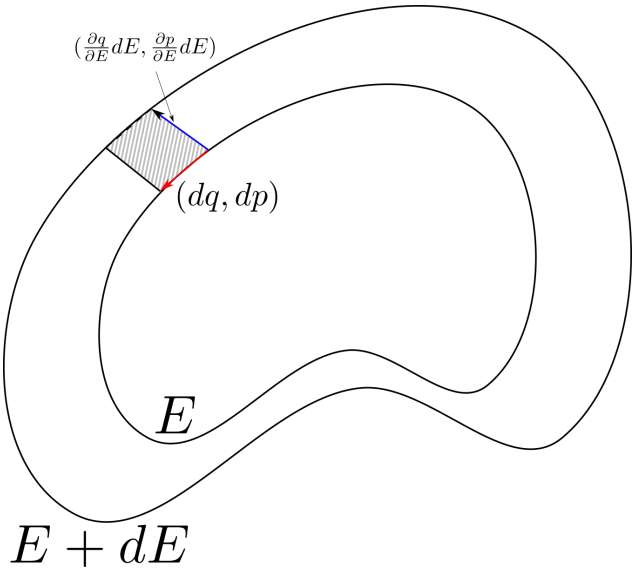
It means that under Hamiltonian dynamics the system will always return arbitrary close to the initial starting point.

The time it will take for the system to return is another matter.

36.3. Area law.

This law is valid only in $1D$. This is unlike Liouville's theorem which is correct for any number of degrees of freedom.

Let's consider a Hamiltonian motion in $1D$. We will assume, that the motion is periodic — in $1D$ the motion is either periodic, or unbounded. In the phase space picture the periodic motion means that the phase space trajectory is a closed loop (without self-crossings). We then can compute the area of the phase space $\mathcal{A} = \int dpdq$ of the loop inside the phase space trajectory of a motion with energy E. This area will depend on the energy E on the trajectory. We thus will have a function $\mathcal{A}(E)$.



If we solve the equations of motion we will get the trajectory in the phase space as functions $q(t, E)$ and $p(t, E)$ — the trajectory depends on the energy E.

If we change the energy by dE the area will change. Consider two trajectories one with the energy E and the other with the energy $E + dE$. We want to compute the difference between the areas for the two trajectories $d\mathcal{A}$.

The vector along the trajectory is $(dq, dp) = (\dot{q}, \dot{p})dt$. The vector from a point of the trajectory E to the trajectory $E + dE$ is $(\frac{\partial q}{\partial E}dE, \frac{\partial p}{\partial E}dE) = (\frac{\partial q}{\partial E}, \frac{\partial p}{\partial E})dE$.

The area of the small shaded parallelogram (see figure) is given by the vector (cross) product of the two vectors $(\dot{q}, \dot{p})dt$ and $(\frac{\partial q}{\partial E}, \frac{\partial p}{\partial E})dE$. The small shaded are is then given by

$$-dE \left(\frac{\partial q}{\partial E}\dot{p} - \frac{\partial p}{\partial E}\dot{q} \right) dt$$

.

The change of the area $d\mathcal{A}$ is then the sum of all these parallelograms along the trajectory:

$$d\mathcal{A} = -dE \oint \left(\dot{p}\frac{\partial q}{\partial E} - \dot{q}\frac{\partial p}{\partial E} \right) dt,$$

where the integral is taken along the trajectory of energy E.

Using the Hamiltonian equations of motion $\dot{q} = \frac{\partial H}{\partial p}$ and $\dot{p} = -\frac{\partial H}{\partial q}$ we get

$$d\mathcal{A} = dE \oint \left(\frac{\partial H}{\partial q}\frac{\partial q}{\partial E} + \frac{\partial H}{\partial p}\frac{\partial p}{\partial E} \right) dt.$$

where the integral is taken along the TRAJECTORY of energy E.

The Hamiltonian is the function of coordinate q and momentum p. The Hamiltonian in the last formula is the Hamiltonian on the trajectory. The trajectory depends on the energy E, so we have $H(q(t, E), p(t, E))$. This function does not depend on time t, as it is conserved. Moreover, the value of the Hamiltonian on the trajectory is energy! So

$$E = H(q(t, E), p(t, E))$$

Differentiating the above equation with respect to E using the chain rule we get

$$1 = \frac{dH(q(t, E), p(t, E))}{dE} = \frac{\partial H}{\partial q}\frac{\partial q}{\partial E} + \frac{\partial H}{\partial p}\frac{\partial p}{\partial E}$$

So we get

$$d\mathcal{A} = dE \oint dt = T dE.$$

where T is time which it take complete the closed trajectory – this is the period of motion! Thus we have our Area Law:

$$\frac{d\mathcal{A}}{dE} = T(E).$$

- In particular for oscillator we saw that $\mathcal{A} = 2\pi E/\omega = ET$.

Notice, that all the theorems/laws considered in this lecture are valid for ANY Hamiltonian! They are not a consequence of some properties of some particular Hamiltonian. They are consequence of the Hamiltonian mechanics itself.

LECTURE 37

Adiabatic invariants. For self-study.

We want to consider the following problem:

- We have a conservative $1D$ system with slowly varying parameter.
- The system is described by a Hamiltonian $H(p, q; \lambda)$, where λ is a parameter, say a spring constant, etc.
- The system undergoes a periodic motion with some period T which depends on energy E and the value of the parameter λ.
- We now start to slowly change the parameter λ as a function of time.
- What can we say about the motion?

Before we do anything we need to understand what does it mean to change the parameter "slowly". The natural definition is that the change of the parameter $\Delta\lambda$ during one period T is small in comparison to the value of the parameter itself:

$$T\frac{d\lambda}{dt} \ll \lambda.$$

Rewriting this as

$$T \ll \lambda/\dot{\lambda}$$

we see, that there are two vastly different time scales: T — typical time for the motion; $\lambda/\dot{\lambda}$ — typical time of change of the parameter λ.

What do we expect:

- If the parameter is a function of time the energy is no longer conserved.
- The rate of change of the energy averaged over the period of the motion will be very slow.
- The averaged rate of change of the energy will be proportional to $\dot{\lambda}$. If $\dot{\lambda} = 0$ — the parameter is constant — then the energy does not change, it is conserved.

So we have rapid oscillations and slow change of the parameter. Let's compute how the energy is changing. Energy is the value of the Hamiltonian on a trajectory.

$$\frac{dE}{dt} = \left(\frac{\partial H}{\partial t}\right)_{p,q} = \left(\frac{\partial H}{\partial \lambda}\right)_{p,q} \frac{d\lambda}{dt}.$$

Where in the RHS in $\left(\frac{\partial H}{\partial \lambda}\right)_{p,q}$ we must substitute the solution of the equation of motion $p(t)$ and $q(t)$. The p and q are changing rapidly with time — the typical time of change is the period T. We want to average the above expression over the period T. As $\frac{d\lambda}{dt}$ almost does

not change during the period we can take it out of the averaging

$$\frac{d\overline{E}}{dt} = \frac{d\lambda}{dt}\left(\overline{\frac{\partial H}{\partial \lambda}}\right)_{p,q}$$

While λ in $\frac{\partial H}{\partial \lambda}$ is changing just a little during the period we can do the averaging in $\overline{\frac{\partial H}{\partial \lambda}}$ assuming λ to be constant.

- So from now on we can consider the Hamiltonian system with constant λ. Which also means constant energy E.

The averaging means

$$\overline{\frac{\partial H}{\partial \lambda}} = \frac{1}{T}\int_0^T \left(\frac{\partial H(p(t),q(t),\lambda)}{\partial \lambda}\right)_{p,q,E} dt.$$

According to the Hamiltonian equation (remember λ is fixed)

$$\dot{q} = \left(\frac{\partial H}{\partial p}\right)_{q,\lambda,E}, \quad \text{or} \quad dt = \frac{dq}{(\partial H/\partial p)_{q,\lambda,E}}.$$

so we have

$$T = \int_0^T dt = \oint \frac{dq}{(\partial H/\partial p)_{q,\lambda,E}}, \qquad \int_0^T \frac{\partial H}{\partial \lambda} dt = \oint \frac{(\partial H/\partial \lambda)_{p,q,E}}{(\partial H/\partial p)_{q,\lambda,E}} dq,$$

where $\oint$ means integrating there and back. We thus have:

$$\frac{d\overline{E}}{dt} = \frac{d\lambda}{dt}\frac{\oint \frac{(\partial H/\partial \lambda)_{p,q,E}}{(\partial H/\partial p)_{q,\lambda,E}} dq}{\oint \frac{dq}{(\partial H/\partial p)_{q,\lambda}}}.$$

In the RHS the integrals must be taken on some particular trajectory. The trajectory depends on the energy E and on the parameter λ.

- This is an important point. All the integrals in the RHS are taken along a trajectory at fixed E and fixed λ!
- So we solve the Hamiltonian equations for some fixed E and λ, and find $p(t;E,\lambda)$ and $q(t;E,\lambda)$ — this is a parametric form (t is a parameter) of a phase space trajectory for given E and λ.
- On this trajectory the momentum p can be considered to be a function of the coordinate q. The phase space trajectory is given by $p(q;E,\lambda)$.

Also on a trajectory, at fixed λ the energy is conserved and $E = H(q,p(q;E,\lambda),\lambda)$ Taking the derivative of this equation with respect to λ for fixed E and q we find

$$\left(\frac{\partial H}{\partial \lambda}\right)_{q,p,E} + \left(\frac{\partial H}{\partial p}\right)_{q,\lambda,E}\left(\frac{\partial p}{\partial \lambda}\right)_{q,E} = 0,$$

or

$$\frac{(\partial H/\partial \lambda)_{p,q,E}}{(\partial H/\partial p)_{q,\lambda,E}} = -\left(\frac{\partial p}{\partial \lambda}\right)_{q,E}.$$

Also on a trajectory $\frac{\partial H}{\partial p} = \frac{\partial E}{\partial p}$. So together we have

$$\frac{d\overline{E}}{dt} = -\frac{d\lambda}{dt}\frac{\oint \left(\frac{\partial p}{\partial \lambda}\right)_{q,E} dq}{\oint \left(\frac{\partial p}{\partial E}\right)_{q,\lambda} dq}$$

or

$$\oint \left(\left(\frac{\partial p}{\partial E} \right)_{q,\lambda} \frac{d\overline{E}}{dt} + \left(\frac{\partial p}{\partial \lambda} \right)_{q,E} \frac{d\lambda}{dt} \right) dq = 0.$$

Again, considering p as $p(q; \overline{E}, \lambda)$, where the dependence of p on q comes from the solution of the Hamiltonian equations at FIXED E and λ we can write (for fixed q) $dp(q; E, \lambda) = \left(\frac{\partial p}{\partial E}\right)_{q,\lambda} d\overline{E} + \left(\frac{\partial p}{\partial \lambda}\right)_{q,E} d\lambda$. (As energy is conserved, there is no distinction between E and $\overline{E}$ in this procedure.) The above equation then is

$$\frac{d}{dt} \oint p(q; E, \lambda) dq = 0.$$

So the quantity

$$I = \frac{1}{2\pi} \oint p dq$$

is called *adiabatic invariant*. This quantity does not change during the adiabatic change of the parameters.

Let me repeat the story:

- We have a conservative $1D$ system with.
- The system is described by a Hamiltonian $H(p, q; \lambda)$, where λ is a parameter, say a spring constant, etc.
- The system undergoes a periodic motion.
- The Hamiltonian equations of motion for FIXED parameter λ conserve the energy E.
- From the equation $E = H(p, q, \lambda)$ we find $p(q; E, \lambda)$
- We compute the quantity

  $$I(E, \lambda) = \frac{1}{2\pi} \oint p dq = \frac{1}{2\pi} \oint p(q; E, \lambda) dq$$

  Notice, that all this is done at FIXED E and λ — we are solving the equations for a purely conservative system!
- If we now start to slowly change the parameter λ with time, the energy of the system will be changing in such a way, that

  $$I(E(t), \lambda(t)) = \text{const}.$$

  will remain constant.

37.1. Examples.

37.1.1. A particle in a box.

- A free $1D$ particle in a box of length L.
- We want to see how the energy of the particle depends on L if we slowly change L. Namely, we start with the particle of some particular energy E at some length L. We then slowly change the length L. How the energy of the particle will change?

We start at fixed E and L. At fixed E the momentum of the particle is $p = \sqrt{2mE}$. The adiabatic invariant then is

$$I = \frac{1}{2\pi} \oint \sqrt{2mE} dq = \frac{\sqrt{2mE}}{2\pi} \oint dq = \frac{\sqrt{2mE}}{2\pi} 2L.$$

So the combination $\sqrt{E}L$ will remain constant if we slowly change L. So will remain the combination EL^2.

In particular, lets assume, that we slowly changed L to $L + dL$. As $EL^2 = \text{const.}$, differentiating this with respect to L we find

$$dEL^2 + 2ELdL = 0,$$

or

$$dE = -2\frac{E}{L}dL.$$

Notice, that this also can be written as ($E = \frac{mv^2}{2} = \frac{1}{2}pv$)

$$dE = -2p\frac{v}{2L}dL = -\frac{2p}{T}dL.$$

but $2p/T$ is the average change of the particle's momentum during one period, so it is an average force f which the particle exerts on the wall. Then fdL is work which the system did while the wall was moving from L to $L + dL$. Accordingly, the energy of the particle has decreased by exactly the work the particle has done.

37.1.2. Oscillator.

The Hamiltonian is

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2}x^2.$$

We want to see how the energy changes if we slowly change the frequency ω.

Considering the motion at fixed E and ω we write

$$p = \pm\sqrt{2mE - m^2\omega^2x^2}$$

The adiabatic invariant is ($x_E = \sqrt{\frac{2E}{m\omega^2}}$)

$$I = \frac{1}{2\pi}2\int_{-x_E}^{x^E}\sqrt{2mE - m^2\omega^2x^2}dx = \frac{E}{\omega}.$$

So if we slowly change ω the energy will always stay proportional to the frequency

$$E \sim \omega.$$

LECTURE 38
Poisson brackets. Change of Variables. Canonical variables.

- Students' evaluations 11-14-2024, until 12-04-2024.

38.1. Poisson brackets.

To avoid the curly brackets clutter I will use the boldface letters to denote the collection of variables: $\mathbf{q} \equiv \{q_i\}$ etc., and restore the indexes when needed.

Consider a function of time, coordinates and momenta: $f(t, \mathbf{q}, \mathbf{p})$. We want to know how the value of this function changes with time on the solutions of the equations of motion. Namely, we have a Hamiltonian $H(\mathbf{q}, \mathbf{p})$ and the Hamiltonian equations of motion

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}.$$

We want to solve them with some initial conditions and find the functions $q_i(t)$ and $p_i(t)$. We then plug these functions in the function f and get $f(t, \mathbf{q}(t), \mathbf{p}(t))$, which is now a function of time – the value of the function f on the trajectory. We want to see how this value changes with time.

So we want to compute $\frac{df}{dt}$:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_i \left(\frac{\partial f}{\partial q_i}\dot{q}_i + \frac{\partial f}{\partial p_i}\dot{p}_i \right) = \frac{\partial f}{\partial t} + \sum_i \left(\frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q_i} \right) = \frac{\partial f}{\partial t} + \{H, f\}$$

where we defined the Poisson brackets for ANY two functions g and f

$$\{g, f\} = \sum_i \left(\frac{\partial g}{\partial p_i}\frac{\partial f}{\partial q_i} - \frac{\partial g}{\partial q_i}\frac{\partial f}{\partial p_i} \right).$$

- Notice, that the Poisson brackets are defined for *any* two functions f and g.

In particular we see, that

$$\{p_j, p_k\} = 0, \qquad \{q_j, q_k\} = 0, \qquad \{p_j, q_k\} = \delta_{j,k}.$$

According to the definition Poisson brackets are

- Antisymmetric.
- Bilinear.
- For a constant c, $\{f, c\} = 0$.
- $\{f_1 f_2, g\} = f_1\{f_2, g\} + f_2\{f_1, g\}$.

- Jacobi's identity. (we will talk about it later.)

38.2. Change of Variables.

We want to answer the following question. What change of variables will keep the Hamiltonian equations intact? Namely, we have our original variables $\mathbf{q}$ and $\mathbf{p}$ and the original Hamiltonian $H(\mathbf{q},\mathbf{p})$. The Hamiltonian equations are

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}.$$

We want to find the new variables $\mathbf{P}$ and $\mathbf{Q}$ (this means that the new variables are the functions of the old variables $\mathbf{P}(\mathbf{q},\mathbf{p})$, $\mathbf{Q}(\mathbf{q},\mathbf{p})$), such that the *form* of the Hamiltonian equations for the new variables have the same form

$$\dot{Q}_i = \frac{\partial H}{\partial P_i}, \qquad \dot{P}_i = -\frac{\partial H}{\partial Q_i}.$$

Let's consider an arbitrary transformation of variables: $P_i = P_i(\mathbf{q},\mathbf{p})$, and $Q_i = Q_i(\mathbf{q},\mathbf{p})$. We then have

$$\dot{P}_i = \{H, P_i\}, \qquad \dot{Q}_i = \{H, Q_i\}.$$

or

$$\dot{P}_i = \sum_k \left[\frac{\partial H}{\partial p_k}\frac{\partial P_i}{\partial q_k} - \frac{\partial H}{\partial q_k}\frac{\partial P_i}{\partial p_k} \right].$$

(the same for $\dot{Q}_i$) At this point I want to make the change of variables in the Hamiltonian. For that I invert/solve the equations for the change of variables to get $p_i = p_i(\mathbf{Q},\mathbf{P})$ and $q_i = q_i(\mathbf{Q},\mathbf{P})$ and substitute these functions into the original Hamiltonian $H(\mathbf{q},\mathbf{p})$

$$H(\mathbf{q}(\mathbf{Q},\mathbf{P}),\mathbf{p}(\mathbf{Q},\mathbf{P})) \equiv H(\mathbf{Q},\mathbf{P}),$$

and, of course, the opposite is also true

$$H(\mathbf{Q}(\mathbf{q},\mathbf{p}),\mathbf{P}(\mathbf{q},\mathbf{p})) = H(\mathbf{q},\mathbf{p}),$$

we then have by the chain rule

$$\frac{\partial H}{\partial p_k} = \sum_\alpha \left(\frac{\partial H}{\partial P_\alpha}\frac{\partial P_\alpha}{\partial p_k} + \frac{\partial H}{\partial Q_\alpha}\frac{\partial Q_\alpha}{\partial p_k} \right), \qquad \frac{\partial H}{\partial q_k} = \sum_\alpha \left(\frac{\partial H}{\partial P_\alpha}\frac{\partial P_\alpha}{\partial q_k} + \frac{\partial H}{\partial Q_\alpha}\frac{\partial Q_\alpha}{\partial q_k} \right)$$

Substituting this into our equation for $\dot{P}_i$ we get

$$\dot{P}_i = \sum_{k,\alpha} \left[\left(\frac{\partial H}{\partial P_\alpha}\frac{\partial P_\alpha}{\partial p_k} + \frac{\partial H}{\partial Q_\alpha}\frac{\partial Q_\alpha}{\partial p_k} \right)\frac{\partial P_i}{\partial q_k} - \left(\frac{\partial H}{\partial P_\alpha}\frac{\partial P_\alpha}{\partial q_k} + \frac{\partial H}{\partial Q_\alpha}\frac{\partial Q_\alpha}{\partial q_k} \right)\frac{\partial P_i}{\partial p_k} \right]$$

Collecting all terms with $\frac{\partial H}{\partial P_\alpha}$ and $\frac{\partial H}{\partial Q_\alpha}$ and using our definition of Poisson brackets we obtain

$$\dot{P}_i = \sum_\alpha \left[\frac{\partial H}{\partial P_\alpha}\{P_\alpha, P_i\} + \frac{\partial H}{\partial Q_\alpha}\{Q_\alpha, P_i\} \right].$$

Analogously,

$$\dot{Q}_i = \sum_\alpha \left[\frac{\partial H}{\partial Q_\alpha}\{Q_\alpha, Q_i\} + \frac{\partial H}{\partial P_\alpha}\{P_\alpha, Q_i\} \right]$$

We see, that the Hamiltonian equations keep their form if

$$\{P_i, Q_\alpha\} = \delta_{i,\alpha}, \qquad \{P_i, P_\alpha\} = \{Q_i, Q_\alpha\} = 0$$

- So in order for the Hamiltonian equation to have the same *form* in the new variables the Poisson brackets of the new variables must be the same as the Poisson brackets of the old variables.

38.3. Canonical variables.

The Poisson brackets

$$\{P_i, Q_\alpha\} = \delta_{i,\alpha}, \qquad \{P_i, P_\alpha\} = \{Q_i, Q_\alpha\} = 0$$

are called *canonical Poisson brackets*.

The variables that have such Poisson brackets are called the *canonical variables*, they are *canonically conjugated*. Transformations that keep the canonical Poisson brackets are called *canonical transformations*.

- Non-uniqueness of the Hamiltonian.
- Coordinates and momenta obtained from Lagrangian are always canonically conjugated.
- $L = p\dot{q} - H$ only if p and q are canonical variables.
- Canonical Poisson brackets are encoded in the $p\dot{q}$ term.

LECTURE 39
Poisson brackets structure. How to compute Poisson brackets for arbitrary functions.

- Students' evaluations 11-14-2024, until 12-04-2024.

39.1. Hamiltonian mechanics

- The Poisson brackets are property of the phase space and have nothing to do with the Hamiltonian.
- The Hamiltonian is just a function on the phase space.
- Given the phase space p_i, q_i, the Poisson brackets and the Hamiltonian. We can construct the equations of the Hamiltonian mechanics:

$$\dot{p}_i = \{H, p_i\}, \qquad \dot{q}_i = \{H, q_i\}.$$

- In this formulation there is no need to distinguish between the coordinates and momenta. So we can use $\xi_1 \dots \xi_{2N}$ instead of $q_1 \dots q_N$ and $p_1 \dots p_N$, with given Poisson brackets $\{\xi_i, \xi_j\}$.
- The equations of motion are then

$$\dot{\xi}_i = \{H, \xi_i\}.$$

- Time evolution of any function $f(\{\xi\}, t)$ is given by the equation

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \{H, f\}.$$

difference between the full and the partial derivatives!

39.2. New formulation of the Hamiltonian mechanics.

Here is the new formulation of mechanics:

- We have a phase space with coordinates ξ_i, where $i = 1 \dots 2N$ for N degrees of freedom.
- This phase space is equipped with Poisson brackets: $\{\xi_i, \xi_j\}$. What it means is that for any two coordinates ξ_i and ξ_j we know a function $\{\xi_i, \xi_j\}$ which depends on two indexes i and j and of all the coordinates.
- These functions, the Poisson brackets, must satisfy the following axioms
  - Antisymmetric.
  - Bilinear.

  - For a constant c, $\{f, c\} = 0$.
  - $\{f_1 f_2, g\} = f_1\{f_2, g\} + f_2\{f_1, g\}$.
  - Jacobi's identity. (we will talk about it later.)
- Notice, that the axioms are formulated for ANY/arbitrary functions on the phase space. So they are the property of the phase space itself.
- Any function on the phase space $H(\{\xi\}, t)$ can be a Hamiltonian (which function you take as a Hamiltonian depends on the problem you are solving.)
- Time evolution of any function $f(\{\xi\}, t)$ is given by the equation

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \{H, f\}.$$

In this formulation the phase space and the Poisson brackets play the major role. They are independent of a Hamiltonian (the are defined before the Hamiltonian even introduced) If we know the Hamiltonian we can also construct the Hamiltonian equations of time evolution of any function.

In particular the time evolution of the Hamiltonian itself is given by

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + \{H, H\} = \frac{\partial H}{\partial t}$$

as $\{H, H\} = 0$ due to antisymmetry of the Poisson brackets. So if the Hamiltonian does not explicitly depend on time, then it is conserved on the trajectories.

- In this formulation we separated the properties of the phase space (the Poisson bracket structure) from the Hamiltonian itself.
- Canonical Poisson brackets is just one example of the possible Poisson bracket structure. (In some sense, this is analogous to the statement that the Euclidean geometry is just one example of all possible geometries.)
- The Jacobi identity puts a very strong restriction on all possible Poisson brackets structures.

39.3. How to compute Poisson brackets for any two functions.

In order to use our new formulation we need a way to compute the Poisson bracket $\{f, g\}$ between any two functions f and g on the phase space if we know all $\{\xi_i, \xi_j\}$. In general the Poisson bracket $\{\xi_i, \xi_j\}$ is the function of all the phase space coordinates. We only require that all the properties listed in definition hold. We do not assume that the Poisson brackets are canonical!

The answer is:

$$\{f, g\} = \frac{\partial f}{\partial \xi_i}\frac{\partial g}{\partial \xi_j}\{\xi_i, \xi_j\}.$$

(summation over the repeated indexes is implied.) Notice the order of indexes. It is important.

Let's prove this formula.

- We start with the Poisson bracket of $\{\xi_j, g\}$, where g is an arbitrary function on the phase space (for simplicity we take that g does not depend on time explicitly).
- In order to compute it we consider ξ_j as our Hamiltonian. This Hamiltonian then gives the time evolution

$$\frac{dg}{dt} = \{\xi_j, g\}.$$

- On the other hand, by the chain rule
$$\frac{dg}{dt} = \frac{\partial g}{\partial \xi_i}\frac{d\xi_i}{dt} = \frac{\partial g}{\partial \xi_i}\{\xi_j, \xi_i\}.$$
- Comparing the two results we see, that
$$\{\xi_j, g\} = \frac{\partial g}{\partial \xi_i}\{\xi_j, \xi_i\}$$
- To compute the Poisson bracket $\{g, f\}$ we consider the function g as the Hamiltonian, then
$$\frac{df}{dt} = \{g, f\}.$$
- On the other hand, by the chain rule
$$\frac{df}{dt} = \frac{\partial f}{\partial \xi_j}\frac{d\xi_j}{dt} = \frac{\partial f}{\partial \xi_j}\{g, \xi_j\} = -\frac{\partial f}{\partial \xi_j}\frac{\partial g}{\partial \xi_i}\{\xi_j, \xi_i\}$$
so that
$$\{f, g\} = \frac{\partial f}{\partial \xi_j}\frac{\partial g}{\partial \xi_i}\{\xi_j, \xi_i\}.$$
(take notice of the order of indexes, it is important as the Poisson brackets are antisymmetric.)
- Using this rule we see, that if all the requirements for the Poisson brackets are satisfied for all $\{\xi_i, \xi_j\}$, then these requirements are satisfied for any functions f and g.

There is one more identity the Poisson brackets must satisfy – the Jacobi's identity.

LECTURE 40
The Jacobi's identity.

- Students' evaluations 11-14-2024, until 12-04-2024.
- Last HW, honors problem.

In this (and the next) lecture we discuss the last axiom: the Jacobi's identity.
Our new formulation of the Hamiltonian mechanics is:

- We have phase space with coordinates ξ_i.
- The phase space equipped with the Poisson brackets structure $\{\xi_i, \xi_j\}$. The Poisson brackets are functions on the phase space.
- The Poisson brackets must satisfy a set of axioms discussed before.
  - Antisymmetric.
  - Bilinear.
  - For a constant c, $\{f, c\} = 0$.
  - $\{f_1 f_2, g\} = f_1\{f_2, g\} + f_2\{f_1, g\}$.
  - Jacobi's identity. (we will talk about it later.)
- Any function on the phase space (and time) can be a Hamiltonian. (Which function you use for a particular problem depends on the problem.)
- In order to define mechanics (time evolution) we MUST have both: the phase space with the Poisson brackets! and the Hamiltonian.
- If the Hamiltonian $H(\boldsymbol{\xi}, t)$ is given and the Poisson brackets are defined, then for any function $g(\boldsymbol{\xi}, t)$ we can write the equation of time evolution

$$\frac{dg}{dt} = \frac{\partial g}{\partial t} + \{H, g\}.$$

- In particular the Hamiltonian equations of motion in the phase space are

$$\dot{\xi}_i = \{H, \xi_i\}.$$

40.1. Sequential evolution by two Hamiltonians.

Here we will start the discussion of the Jacobi's identity. This axiom is formulated as follows:

- The Poisson brackets must be such, that for ANY three functions f, g, and h, the following must be true

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0.$$

As with any axiomatic definitions the main requirement is that the set of axioms was minimal and self-consistent. An addition of one more axiom (Jacobi's identity in this case)

must be strongly justified by showing that without it the requirement of self-consistency will not be met. Self consistency of the set of axioms means that if we compute the same quantity in two different ways following the axioms we will always come to the same result. So in order to show that the Jacobi's identity is needed I will show that it ensures that a certain calculation done in two different ways gives the same result.

The calculation we need to consider is the change of a function g due to sequential evolution by two different Hamiltonians.

Let's consider a function $g(\boldsymbol{\xi})$. We start at some time t, then we evolve the function $g(\boldsymbol{\xi})$ with time in two steps: first we evolve it for small time dt_1 with a Hamiltonian H_1; after that, starting from time $t + dt_1$ we evolve it for small time dt_2 with the Hamiltonian H_2. We want to compute $g(t + dt_1 + dt_2)$ up to the SECOND order in dt_1, dt_2.

$$g(t) \xrightarrow{H_1} g(t + dt_1) \xrightarrow{H_2} g(t + dt_1 + dt_2)$$

Before we proceed I want to point out that in the case we considering the phase space with its Poisson brackets is defined, but there are two Hamiltonians. The time evolution, which means full time derivative, depends on the Hamiltonian. So instead of writing $\frac{d}{dt}$ I will use the notation $\frac{d}{dt}\Big|_H$ to keep the notion of which Hamiltonian gives the time evolution.

Step 1: We use the Taylor expansion up to $(dt_1)^2$

$$g(t + dt_1) = g(t) + dt_1 \left.\frac{d}{dt}\right|_{H_1} g(t) + \frac{(dt_1)^2}{2} \left.\frac{d}{dt}\right|_{H_1} \left.\frac{d}{dt}\right|_{H_1} g(t)$$

where $\frac{d}{dt}\Big|_{H_1}$ means the full time derivative induced by the Hamiltonian H_1. Using the rule for the Hamiltonian time evolution of ANY function we find

$$g(t + dt_1) = g(t) + dt_1\{H_1, g\}_t + \frac{(dt_1)^2}{2}\{H_1, \{H_1, g\}\}_t$$

where the subscript t of the Poisson bracket means that the bracket must be computed at time t — this is what the Taylor expansion demands.

Step 2: After time $t + dt_1$ we need to use the Hamiltonian H_2 for the time evolution. Using the Taylor expansion over dt_2 at time $t + dt_1$ we have

$$g(t + dt_1 + dt_2) = g(t + dt_1) + dt_2 \left.\frac{d}{dt}\right|_{H_2} g(t + dt_1) + \frac{(dt_2)^2}{2} \left.\frac{d}{dt}\right|_{H_2} \left.\frac{d}{dt}\right|_{H_2} g(t + dt_1).$$

Now I use our result of Step 1 for $g(t+dt_1)$ and keep only the terms up to the second order in time increments dt_1 and dt_2.

$$\begin{aligned} g(t + dt_1 + dt_2) &= g(t) + dt_1\{H_1, g\}_t + \frac{(dt_1)^2}{2}\{H_1, \{H_1, g\}\}_t \\ &+ dt_2\{H_2, g\}_t + \frac{(dt_2)^2}{2}\{H_2, \{H_2, g\}\}_t \\ &+ dt_1 dt_2 \left.\frac{d}{dt}\right|_{H_2} \{H_1, g\}_t \end{aligned}$$

Notice, that the blue terms in the right hand side have only H_1 and dt_1, the red terms have only dt_2 and H_2, so they do not reflect the fact that the time evolution was done sequentially by two different Hamiltonians. Only the very last term $\frac{d}{dt}\Big|_{H_2}\{H_1, g\}_t$ depends on both H_1

and H_2 — this is the term we want to concentrate on. I want to compute it in two different ways.

Way 1: Using the Hamiltonian rule for the time evolution under the Hamiltonian H_2 we find

$$\left.\frac{d}{dt}\right|_{H_2} \{H_1, g\} = \{H_2, \{H_1, g\}\}.$$

Way 2: As Poisson brackets are bi-linear, I must be able to write

$$\left.\frac{d}{dt}\right|_{H_2} \{H_1, g\} = \left\{ \left.\frac{d}{dt}\right|_{H_2} H_1, g \right\} + \left\{ H_1, \left.\frac{d}{dt}\right|_{H_2} g \right\} = \{\{H_2, H_1\}, g\} + \{H_1, \{H_2, g\}\}$$

The result for $g(t + dt_1 + dt_2)$ must not depend on how I do the computation, so we must have

$$\{\{H_2, H_1\}, g\} + \{H_1, \{H_2, g\}\} = \{H_2, \{H_1, g\}\},$$

or using antisymmetry of the Poisson brackets we rewrite the above in a more symmetric way:

$$\{H_1, \{H_2, g\}\} + \{H_2, \{g, H_1\}\} + \{g, \{H_1, H_2\}\} = 0.$$

As ANY function on the phase space can be considered as a Hamiltonian, this rule MUST be satisfied for ANY three functions.

40.2. The axiomatic definition of the Poisson brackets.

Given a $2N$ dimensional phase space $\boldsymbol{\xi}$ the Poisson brackets must satisfy:

- Antisymmetric.
- Bilinear.
- For a constant c, $\{f, c\} = 0$.
- For any functions f_1, f_2, and g the following is true $\{f_1 f_2, g\} = f_1\{f_2, g\} + f_2\{f_1, g\}$.
- Jacobi's identity: for ANY three functions f, g, and h, the following must be true $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$.

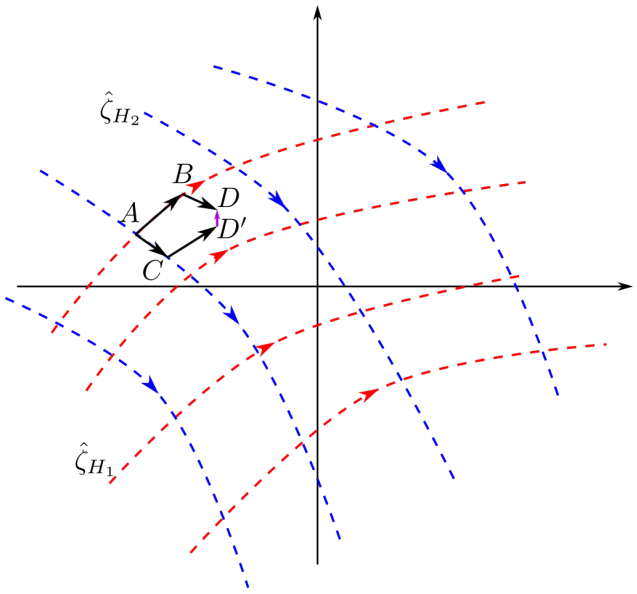
The statement is that if the axioms are satisfied by the Poisson brackets of the phase space coordinates $\{\xi_i, \xi_j\}$, then they will be satisfied for arbitrary functions.

If we know the Poisson brackets $\{\xi_i, \xi_j\}$, we can compute the Poisson brackets of any two functions f and g:

$$\{f, g\} = \frac{\partial f}{\partial \xi_i} \frac{\partial g}{\partial \xi_j} \{\xi_i, \xi_j\}.$$

(Einstein notations are used, summation over repeated indexes is implied.)

40.3. Commutation of Hamiltonian flows.



For a Hamiltonian H we can introduce the operator $\hat{\zeta}_H$ of the Hamiltonian flow by the following definition: for any function g

$$\hat{\zeta}_H g \equiv \{H, g\}$$

Using this definition for one of the coordinates (plugging ξ_i instead of g) we see, that $\hat{\zeta}_H \xi_i = \{H, \xi_i\} = \frac{d\xi_i}{dt}$. So this Hamiltonian flow induces the Hamiltonian vector field we considered earlier.

Let's now consider two Hamiltonians H_1 and H_2 and compute the commutator of their flows. Namely, for any function g we have (using Jacobi's identity)

$$\hat{\zeta}_{H_1}\hat{\zeta}_{H_2} g - \hat{\zeta}_{H_2}\hat{\zeta}_{H_1} g = \{H_1, \{H_2, g\}\} - \{H_2, \{H_1, g\}\} = \{\{H_1, H_2\}, g\} = \hat{\zeta}_{\{H_1,H_2\}} g.$$

As this is true for *any* function g we have

$$\hat{\zeta}_{H_1}\hat{\zeta}_{H_2} - \hat{\zeta}_{H_2}\hat{\zeta}_{H_1} = \hat{\zeta}_{\{H_1,H_2\}}.$$

So the commutator of the Hamiltonian flows is also a Hamiltonian flow.

On the figure

- The red dashed lines show the flow $\hat{\zeta}_{H_1}$, the blue dashed lines show the flow $\hat{\zeta}_{H_2}$.
- The operator $\hat{\zeta}_{H_2}\hat{\zeta}_{H_1}$ shifts the point A along ABD path.
- The operator $\hat{\zeta}_{H_1}\hat{\zeta}_{H_2}$ shifts the point A along ACD' path.
- So the operator $\hat{\zeta}_{H_1}\hat{\zeta}_{H_2} - \hat{\zeta}_{H_2}\hat{\zeta}_{H_1}$ shifts point D' to point D.
- This shift can be described by another Hamiltonian flow $\hat{\zeta}_{\{H_1,H_2\}}$.

LECTURE 41
The Jacobi's identity. Proof.

- Students' evaluations 11-14-2024, until 12-04-2024.

41.1. Formulation of the Hamiltonian mechanics.

- We have phase space with coordinates ξ_i.
- The phase space equipped with the Poisson brackets structure $\{\xi_i, \xi_j\}$. The Poisson brackets are functions on the phase space.
- The Poisson brackets must satisfy a set of axioms discussed before.
  - Antisymmetric.
  - Bilinear.
  - For a constant c, and any function $f(\boldsymbol{\xi}, t)$ on the phase space: $\{f, c\} = 0$.
  - For any three functions $f_1(\boldsymbol{\xi}, t)$, $f_2(\boldsymbol{\xi}, t)$, and $g(\boldsymbol{\xi}, t)$ on the phase space: $\{f_1 f_2, g\} = f_1\{f_2, g\} + f_2\{f_1, g\}$.
  - Jacobi's identity: For any three functions $f(\boldsymbol{\xi}, t)$, $g(\boldsymbol{\xi}, t)$, and $h(\boldsymbol{\xi}, t)$ on the phase space the following identity holds $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$
- If we know the Poisson brackets for the phase space coordinates $\{\xi_i, \xi_j\}$, we can compute the Poisson brackets of any two functions f and g:

$$\{f, g\} = \frac{\partial f}{\partial \xi_i} \frac{\partial g}{\partial \xi_j} \{\xi_i, \xi_j\}.$$

  (Einstein notations are used, summation over repeated indexes is implied.)
- Any function on the phase space (and time) can be a Hamiltonian. (Which function you use for a particular problem depends on the problem.)
- In order to define mechanics (time evolution) we MUST have both: the phase space with the Poisson brackets! and the Hamiltonian.
- If the Hamiltonian $H(\boldsymbol{\xi}, t)$ is given and the Poisson brackets are defined, then for any function $g(\boldsymbol{\xi}, t)$ we can write the equation of time evolution

$$\frac{dg}{dt} = \frac{\partial g}{\partial t} + \{H, g\}.$$

- In particular the Hamiltonian equations of motion are

$$\dot{\xi}_i = \{H, \xi_i\}.$$

Using the definition of the Poisson brackets in the canonical coordinates $\mathbf{p}$ and $\mathbf{q}$

$$\{f,g\} = \sum_i \left(\frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} \right)$$

it is easy, but lengthy to prove, that for any three functions f, g, and h and canonical Poisson brackets:

$$\{f,\{g,h\}\} + \{g,\{h,f\}\} + \{h,\{f,g\}\} = 0$$

As it holds for any three functions this is the property of the phase space and the Poisson brackets.

- However, we do not want to restrict ourselves to use the canonical Poisson brackets.

41.2. Proof of the Jacobi identity.

First we need to establish what we want to prove, after all the Jacobi identity is the part of the axiomatic definition of the Poisson brackets. The statement we want to prove is the following:

- If the Jacobi identity is satisfied by the Poisson brackets of the phase space coordinates, then it is satisfied for any three arbitrary functions.
- So we want to proof that if for any i, j, k

  $$\{\xi_i,\{\xi_j,\xi_k\}\} + \{\xi_j,\{\xi_k,\xi_i\}\} + \{\xi_k,\{\xi_i,\xi_j\}\} = 0$$

  then

  $$\{f,\{g,h\}\} + \{g,\{h,f\}\} + \{h,\{f,g\}\} = 0$$

  for any functions f, g, and h.

We will prove this statement in two steps. First we consider the situation when $\{\xi_i,\xi_j\} =$ const — independent of the phase space coordinates. (This case covers the canonical Poisson brackets.) Then we consider the general case.

41.2.1. The case of $\{\xi_i,\xi_j\} =$ const.

We have the phase space coordinates $\xi_1 \dots \xi_{2N}$, and $\{\xi_i,\xi_j\} =$ const. This constants will depend on i and j, but they do not depend on the coordinates $\boldsymbol{\xi}$.

- Notice, that in this case the Jacobi's identity for the coordinates $\{\xi_i,\{\xi_j,\xi_k\}\} + \{\xi_j,\{\xi_k,\xi_i\}\} + \{\xi_k,\{\xi_i,\xi_j\}\} = 0$ is satisfied trivially: all inner Poisson brackets are constants, so each term is zero.

Consider the first term in the Jacobi's identity $\{f,\{g,h\}\}$. According to our rule of computing the Poisson brackets we have

$$\{g,h\} = \frac{\partial g}{\partial \xi_j}\frac{\partial h}{\partial \xi_i}\{\xi_j,\xi_i\}.$$

- Remember, we are using the Einstein notations!

Using the same rule again we have

$$\{f,\{g,h\}\} = \frac{\partial f}{\partial \xi_p}\left(\frac{\partial}{\partial \xi_l}\{g,h\}\right)\{\xi_p,\xi_l\} = \frac{\partial f}{\partial \xi_p}\frac{\partial}{\partial \xi_l}\left(\frac{\partial g}{\partial \xi_i}\frac{\partial h}{\partial \xi_j}\{\xi_i,\xi_j\}\right)\{\xi_p,\xi_l\}$$

or

$$\{f,\{g,h\}\}=\frac{\partial f}{\partial\xi_p}\frac{\partial}{\partial\xi_l}\left(\frac{\partial g}{\partial\xi_i}\frac{\partial h}{\partial\xi_j}\right)\{\xi_i,\xi_j\}\{\xi_p,\xi_l\}+\frac{\partial f}{\partial\xi_p}\frac{\partial g}{\partial\xi_i}\frac{\partial h}{\partial\xi_j}\left(\frac{\partial}{\partial\xi_l}\{\xi_i,\xi_j\}\right)\{\xi_p,\xi_l\},$$

As $\{\xi_i,\xi_j\}=\text{const}$ the last term is zero. We will return back to this term in the second step, when we consider $\{\xi_i,\xi_j\}\neq\text{const}$ case. Taking the derivative in the first term we find

$$\{f,\{g,h\}\}=\frac{\partial f}{\partial\xi_p}\frac{\partial^2 g}{\partial\xi_i\partial\xi_l}\frac{\partial h}{\partial\xi_j}\{\xi_i,\xi_j\}\{\xi_p,\xi_l\}+\frac{\partial f}{\partial\xi_p}\frac{\partial g}{\partial\xi_i}\frac{\partial^2 h}{\partial\xi_j\partial\xi_l}\{\xi_i,\xi_j\}\{\xi_p,\xi_l\}.$$

Cycle permutations of the functions f, g, and h gives the other two terms

$$\{g,\{h,f\}\}=\frac{\partial g}{\partial\xi_p}\frac{\partial^2 h}{\partial\xi_i\partial\xi_l}\frac{\partial f}{\partial\xi_j}\{\xi_i,\xi_j\}\{\xi_p,\xi_l\}+\frac{\partial g}{\partial\xi_p}\frac{\partial h}{\partial\xi_i}\frac{\partial^2 f}{\partial\xi_j\partial\xi_l}\{\xi_i,\xi_j\}\{\xi_p,\xi_l\},$$

$$\{h,\{f,g\}\}=\frac{\partial h}{\partial\xi_p}\frac{\partial^2 f}{\partial\xi_i\partial\xi_l}\frac{\partial g}{\partial\xi_j}\{\xi_i,\xi_j\}\{\xi_p,\xi_l\}+\frac{\partial h}{\partial\xi_p}\frac{\partial f}{\partial\xi_i}\frac{\partial^2 g}{\partial\xi_j\partial\xi_l}\{\xi_i,\xi_j\}\{\xi_p,\xi_l\}$$

We want to show that the sum of all these terms is zero for *arbitrary* functions f, g, and h. In order to simplify the calculations we notice the following:

- Each term has one second derivative of one of the functions.
- The second derivatives come from the inner Poisson brackets in each term.

For example, the second derivative of the function h comes one time from the term $\{f,\{g,h\}\}$ and and the other time from the term $\{g,\{h,f\}\}$, and does not come from the term $\{h,\{f,g\}\}$, as in this term h is not in the inner Poisson bracket. So in the whole sum of Jacobi's identity the second derivative of the function h will appear twice. The same is true for the functions f and g. As the functions f, g, and h are arbitrary and independent form each other, in order for the Jacobi's identity to hold the terms that have the second derivative of the same function must cancel each other!

$$\frac{\partial g}{\partial\xi_i}\frac{\partial f}{\partial\xi_p}\frac{\partial^2 h}{\partial\xi_j\partial\xi_l}\{\xi_i,\xi_j\}\{\xi_p,\xi_l\}+\frac{\partial g}{\partial\xi_p}\frac{\partial f}{\partial\xi_j}\frac{\partial^2 h}{\partial\xi_i\partial\xi_l}\{\xi_i,\xi_j\}\{\xi_p,\xi_l\}$$

This sum MUST be zero.

- Remember, we are using the Einstein notations!
- The summations of all the indexes is assumed.
- It does not matter which latter we use to label the indexes, as we sum over all their possible values anyway.
- So we can relabel the indexes in the first term in such a way that all the derivatives in the first and the second term become the same.
- Then take the first term above an relabel the indexes according to the scheme

$$i\to p, p\to j, j\to l, l\to i.$$

We get

$$\frac{\partial g}{\partial\xi_p}\frac{\partial f}{\partial\xi_j}\frac{\partial^2 h}{\partial\xi_l\partial\xi_i}\{\xi_p,\xi_l\}\{\xi_j,\xi_i\}+\frac{\partial g}{\partial\xi_p}\frac{\partial f}{\partial\xi_j}\frac{\partial^2 h}{\partial\xi_i\partial\xi_l}\{\xi_i,\xi_j\}\{\xi_p,\xi_l\}$$

We see, that now the only difference between these two terms is that the first one has $\{\xi_j,\xi_i\}$, while the second has $\{\xi_i,\xi_j\}$. As the Poisson brackets are antisymmetric the two terms are the same, but with the opposite sign. So the sum of the two terms is zero.

The other terms are obtained by a simple cyclic permutation of the functions f, g, and h. So if the second derivative terms of h cancel each other the other terms will also cancel each other. Then the total sum of all therms is zero, as it should be by Jacobi identity.

So we proved, that in the case $\{\xi_i, \xi_j\} = \text{const}$ the Jacobi identity is satisfied for any three functions f, g, and h.

- Notice, that in this step we did not need the Jacobi identity for the coordinates.
- On the other hand if $\{\xi_i, \xi_j\} = \text{const}$ the Jacobi identity for the coordinates are satisfied automatically, as the inner Poisson brackets are constant.

41.2.2. The case of arbitrary $\{\xi_i, \xi_j\}$.

This case is different from the previous one only in one point. We cannot pull $\{\xi_i, \xi_j\}$ from under the differentiation. We then have

$$\{f,\{g,h\}\} = \frac{\partial f}{\partial \xi_p}\frac{\partial^2 g}{\partial \xi_i \partial \xi_l}\frac{\partial h}{\partial \xi_j}\{\xi_i,\xi_j\}\{\xi_p,\xi_l\} + \frac{\partial f}{\partial \xi_p}\frac{\partial g}{\partial \xi_i}\frac{\partial^2 h}{\partial \xi_j \partial \xi_l}\{\xi_i,\xi_j\}\{\xi_p,\xi_l\} + \frac{\partial f}{\partial \xi_p}\frac{\partial g}{\partial \xi_i}\frac{\partial h}{\partial \xi_j}\frac{\partial \{\xi_i,\xi_j\}}{\partial \xi_l}\{\xi_p,\xi_l\}$$

The first two terms are exactly the same as before, so as before they will cancel each other. We then need to concentrate only on the last term. Let's write all three of these terms obtained by cyclic permutation of the functions f, g, and h.

$$\frac{\partial f}{\partial \xi_p}\frac{\partial g}{\partial \xi_i}\frac{\partial h}{\partial \xi_j}\frac{\partial \{\xi_i,\xi_j\}}{\partial \xi_l}\{\xi_p,\xi_l\} + \frac{\partial g}{\partial \xi_p}\frac{\partial h}{\partial \xi_i}\frac{\partial f}{\partial \xi_j}\frac{\partial \{\xi_i,\xi_j\}}{\partial \xi_l}\{\xi_p,\xi_l\} + \frac{\partial h}{\partial \xi_p}\frac{\partial f}{\partial \xi_i}\frac{\partial g}{\partial \xi_j}\frac{\partial \{\xi_i,\xi_j\}}{\partial \xi_l}\{\xi_p,\xi_l\}.$$

- We relabel the indexes in the last two terms in such a way, that all terms have the same derivatives.
- It means, that in the second term we relabel
$$j \to p, i \to j, p \to i.$$
- In the third term we relabel
$$i \to p, j \to i, p \to j$$

We get

$$\frac{\partial f}{\partial \xi_p}\frac{\partial g}{\partial \xi_i}\frac{\partial h}{\partial \xi_j}\left(\frac{\partial \{\xi_i,\xi_j\}}{\partial \xi_l}\{\xi_p,\xi_l\} + \frac{\partial \{\xi_j,\xi_p\}}{\partial \xi_l}\{\xi_i,\xi_l\} + \frac{\partial \{\xi_p,\xi_i\}}{\partial \xi_l}\{\xi_j,\xi_l\}\right).$$

As the functions f, g, and h are arbitrary in order for the above to be zero the expression in the brackets must be zero. Let's then concentrate on the term in the brackets. In this terms we still need to sum over the index l.

To start with we take a hard look at the first term in the brackets

$$\frac{\partial \{\xi_i,\xi_j\}}{\partial \xi_l}\{\xi_p,\xi_l\}.$$

Remember, that the Poisson bracket $\{\xi_p, F\} = \frac{\partial F}{\partial \xi_l}\{\xi_p,\xi_l\}$ for ANY function F on the phase space. So treating $\{\xi_i,\xi_j\}$ as some function on the phase space we get

$$\frac{\partial \{\xi_i,\xi_j\}}{\partial \xi_l}\{\xi_p,\xi_l\} = \{\xi_p,\{\xi_i,\xi_j\}\}.$$

Applying this trick to every term in the brackets we get

$$\{\xi_p,\{\xi_i,\xi_j\}\} + \{\xi_i,\{\xi_j,\xi_p\}\} + \{\xi_j,\{\xi_p,\xi_i\}\}.$$

But this is zero by the Jacobi identity for the Poisson brackets!

So we proved, that if the Jacobi identity is satisfied by the Poisson brackets of the phase space coordinates, then it is satisfied for any three arbitrary functions.

- As it holds for any functions this is the property of the phase space and the Poisson brackets themselves.

LECTURE 42
Integrals of motion. Angular momentum.

- Students' evaluations 11-14-2024, until 12-04-2024. (7)

42.1. Time evolution of Poisson brackets.

Consider two arbitrary functions $f(\{\xi\}, t)$ and $g(\{\xi\}, t)$. We want to compute the full time derivative of their Poisson bracket

$$\frac{d}{dt}\{f, g\}.$$

It means, that

- we have a phase space with Poisson brackets.
- We also have a Hamiltonian.
- We solve the Hamiltonian equations of motion

  $$\dot{\xi}_i = \{H, \xi_i\}$$

  subject to some initial conditions, and find $\xi_i(t)$ for all is (we do not distinguish between coordinates and momenta)
- We compute the Poisson bracket for some known functions on the phase space f and g which also may depend on time explicitly: $\{f, g\}$ – it will be some function of all ξ and time.
- We substitute the solutions $\xi(t)$ in this function and then take the time derivative.

$$\frac{d}{dt}\{f, g\}.$$

Before we do that I want to compute a much simpler

$$\frac{\partial}{\partial t}\{f, g\}.$$

This is partial derivative. So we just consider the explicit dependence of $\{f, g\}$ on time. This explicit dependence comes from the explicit time dependence of f and g.

As it is only a partial derivative, we keep fixed all other variables except t, so I will leave them out to shorten the formulas. According to the definition of partial derivative we need to compute:

$$\begin{aligned}
&\{f(t+\Delta t), g(t+\Delta t)\} - \{f(t), g(t)\} \\
&= \{f(t+\Delta t), g(t+\Delta t)\} - \{f(t), g(t+\Delta t)\} + \{f(t), g(t+\Delta t)\} - \{f(t), g(t)\} \\
&= \{f(t+\Delta t) - f(t), g(t+\Delta t)\} + \{f(t), g(t+\Delta t) - g(t)\}.
\end{aligned}$$

Dividing this by Δt and taking the limit $\Delta t \to 0$ we get

$$\frac{\partial}{\partial t}\{f,g\} = \left\{\frac{\partial f}{\partial t}, g\right\} + \left\{f, \frac{\partial g}{\partial t}\right\}.$$

Notice

- The only property of the Poisson brackets which we used is its bi-linearity.

Now Let's compute the full time evolution of the Poisson bracket $\{f,g\}$ under the Hamiltonian H.

$$\frac{d}{dt}\{f,g\} = \frac{\partial}{\partial t}\{f,g\} + \{H,\{f,g\}\} = \left\{\frac{\partial f}{\partial t}, g\right\} + \left\{f, \frac{\partial g}{\partial t}\right\} + \{\{H,f\},g\} + \{f,\{H,g\}\}$$
$$= \left\{\frac{\partial f}{\partial t} + \{H,f\}, g\right\} + \left\{f, \frac{\partial g}{\partial t} + \{H,g\}\right\}$$

Notice, that in this derivation we used

- the Jacobi's identity,
- the antisymmetry,
- and the bi-linearity

of the Poisson brackets.

So we get

$$\frac{d}{dt}\{f,g\} = \left\{\frac{df}{dt}, g\right\} + \left\{f, \frac{dg}{dt}\right\}.$$

- Notice, that these are the full derivatives, not partial!!

42.2. Integrals of motion.

A conserved quantity is such a function $f(\{\xi\},t)$, that $\frac{df}{dt} = 0$ under the evolution induced by a Hamiltonian H. So if we have two conserved quantities $f(\{\xi\},t)$ and $g(\{\xi\},t)$, then

$$\frac{d}{dt}\{f,g\} = \left\{\frac{df}{dt}, g\right\} + \left\{f, \frac{dg}{dt}\right\} = 0$$

So if we have two conserved quantities we can construct a new conserved quantity! Sometimes it will turn out to be an independent conservation law!

42.3. Angular momentum.

This is an example of a case where the Poisson brackets do not have a global canonical form.

42.3.1. Poisson Brackets.

Let's calculate the Poisson brackets for the components of angular momentum: $\vec{M} = \vec{r} \times \vec{p}$.

Coordinate $\vec{r}$ and momentum $\vec{p}$ are canonically conjugated so

$$\{p^i, r^j\} = \delta_{ij}, \qquad \{p^i, p^j\} = \{r^i, r^j\} = 0.$$

As our coordinates and momenta are canonical, we can use the definition of the Poisson brackets through derivatives — the way they were introduced from the very beginning. However, I will show that we can compute the Poisson brackets between the angular momentum components algebraically — using only the properties of the Poisson brackets.

Using $M^i = \epsilon^{ijk} r^j p^k$ we write

$$\{M^i, M^j\} = \epsilon^{ilk}\epsilon^{jmn}\{r^l p^k, r^m p^n\} = \epsilon^{ilk}\epsilon^{jmn}\left(r^l\{p^k, r^m p^n\} + p^k\{r^l, r^m p^n\}\right) =$$
$$\epsilon^{ilk}\epsilon^{jmn}\left(r^l p^n\{p^k, r^m\} + r^l r^m\{p^k, p^n\} + p^k p^n\{r^l, r^m\} + p^k r^m\{r^l, p^n\}\right) =$$
$$\epsilon^{ilk}\epsilon^{jmn}\left(r^l p^n \delta_{km} - p^k r^m \delta_{ln}\right) = \left(\epsilon^{ilk}\epsilon^{jkn} - \epsilon^{ikn}\epsilon^{jlk}\right) p^n r^l = p^i r^j - r^i p^j = -\epsilon^{ijk} M^k$$

(I used $\epsilon^{ilk}\epsilon^{jnk} = \delta^{ij}\delta^{ln} - \delta^{in}\delta^{lj}$). In short the result is

$$\{M^i, M^j\} = -\epsilon^{ijk} M^k$$

Notice:

- The components of the angular momentum construct their own phase space closed under the Poisson brackets!
- Unlike the usual phase space this phase space looks odd (3) dimensional!
- This puzzle is resolved by the following observation:

$$\{M^i, \vec{M}^2\} = \{M^i, M^k M^k\} = 2\{M^i, M^k\} M^k = -2\epsilon^{ikj} M^j M^k = 0.$$

- So for any Hamiltonian H on this "angular momentum phase space" which means that H depends on the components of $\vec{M}$ only, the $\vec{M}^2$ will be conserved!

$$\frac{d\vec{M}^2}{dt} = \{H, \vec{M}^2\} = \frac{\partial H}{\partial M^i}\{M^i, \vec{M}^2\} = 0.$$

- So in $3D$ space of $\vec{M}$ all motion will happen on the spheres $\vec{M}^2 = \text{const}$..
- The sphere is $2D$ – even dimension.
- There is no way to construct <u>global</u> canonical coordinates on this space.

42.3.2. Spin in magnetic field.

We can now consider a Hamiltonian mechanics, say for the Hamiltonian

$$H = \vec{h} \cdot \vec{M} = h_j M^j.$$

In this case

$$\dot{M}^i = \{H, M^i\} = h_j\{M^j, M^i\} = -h_j \epsilon^{jik} M^k,$$

or

$$\dot{\vec{M}} = \vec{h} \times \vec{M}.$$

Notice:

- $\dot{M}^2 = 2\vec{M} \cdot \dot{\vec{M}} = 2\vec{M} \cdot \left[\vec{h} \times \vec{M}\right] = 0.$
- Energy is conserved, so $\vec{h} \cdot \vec{M} = \text{const}$. (one can check, that $\vec{h} \cdot \dot{\vec{M}} = 0$). The projection of $\vec{M}$ on the direction of $\vec{h}$ does not change with time.
- This equation (Bloch equation) describes a vector $\vec{M}$ rotating with constant angular velocity around the direction of $\vec{h}$.

42.3.3. Euler equations

Consider a free rigid body with tensor of inertia $\hat{I}$. The Hamiltonian is just the kinetic energy.

$$H = \frac{1}{2} M^i (\hat{I}^{-1})_{ij} M^j.$$

The equations of motion then are

$$\dot{M}^k = \{H, M^k\} = \frac{1}{2}\{M^i, M^k\}(\hat{I}^{-1})_{ij} M^j + \frac{1}{2} M^i (\hat{I}^{-1})_{ij} \{M^j, M^k\} = \epsilon^{kil} M^l (\hat{I}^{-1})_{ij} M^j.$$

Let's write this equation in the system of coordinates of the principal axes of the body. Then the tensor of inertia is diagonal, and for M_x component we get

$$\dot{M}^x = M^z I_{yy}^{-1} M^y - M^y I_{zz}^{-1} M^z.$$

or, using that $M^x = I_{xx}\Omega^x$, etc we get

$$I_{xx}\dot{\Omega}^x = (I_{zz} - I_{yy})\Omega^z \Omega^y,$$

and two more equations under the cyclic permutations.

- Three degrees of freedom. We must have three second order differential equations for complete description. We have only three first order equations. Three more first order equations are missing.
- The equations are written for the components of $\vec{\Omega}$ in the non-internal system of coordinates which is rotating with $\vec{\Omega}$.
- In order to find the orientation of the rigid body as a function of time we need to write and solve three more first order non-linear differential equations.
- We will do it at some point next semester.

LECTURE 43
Hamilton-Jacobi equation.

- Students' evaluations 11-14-2024, until 12-04-2024 (14).

This is the last lecture for the class. In this lecture we will tie together the concepts of Action, Lagrangian, and Hamiltonian.

43.1. Action on trajectory.

Consider an action

$$\mathcal{S} = \int_{t_0}^{t_1} L(\mathbf{q}, \dot{\mathbf{q}}, t)dt, \qquad q_i(t_0) = q_{i,0}, \quad q_i(t_1) = q_{i,1}.$$

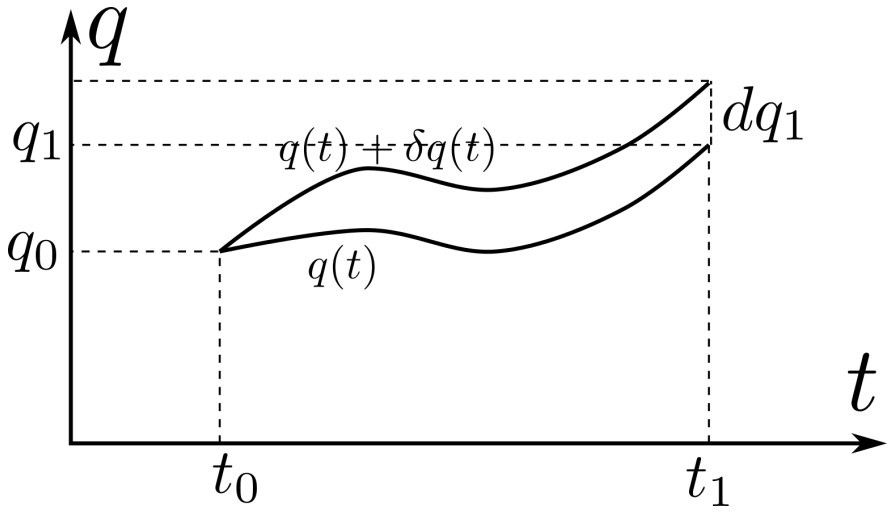
Consider the value of the action on the trajectory as a function of $q_{i,1}$ and t_1. What it means is the following:

- We have an action and thus we have a Lagrangian.
- We write the Lagrangian equations of motion with the boundary conditions: $q_i(t_0) = q_{i,0}$ and $q_i(t_1) = q_{i,1}$.
- We solve these equation of motion (with the boundary conditions)

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \qquad q_i(t_0) = q_{i,0}, \quad q_i(t_1) = q_{i,1}.$$

  and find the functions $q_i(t; t_0, \mathbf{q}_0, t_1, \mathbf{q}_1)$ — those are coordinates as function of time, the boundary conditions are the parameters the function depends on.
- We substitute those functions $q_i(t; t_0, \mathbf{q}_0, t_1, \mathbf{q}_1)$ into the action and take the integral over time t.
- The result will be a FUNCTION $\mathcal{S}(t_0, \mathbf{q}_0, t_1, \mathbf{q}_1)$ — the value of the action on the trajectory. This FUNCTION will depend on t_0, $\mathbf{q}_0$, t_1, and $\mathbf{q}_1$.
- We are interested in how this function depends on $\mathbf{q}_1$ and t_1 with all other parameters fixed.

43.2. Momentum.



- To simplify the calculation I will consider only the case of one dgree of freedom. If we have N degrees of freedom the calculation is almost identical. I will simply give the general result at the end.

We want to see how the value of the Action changes when we change the q_1, while keeping t_1 (as well as q_0 and t_0) fixed.

If we change the upper limit from q_1 to $q_1 + dq_1$ the trajectory will also change from $q(t)$ to $q(t) + \delta q(t)$, where $\delta q(t_0) = 0$, and $\delta q(t_1) = dq_1$. The change of the action then is

$$d\mathcal{S} = \int_{t_0}^{t_1} L(q+\delta q, \dot{q}+\delta\dot{q}, t)dt - \int_{t_0}^{t_1} L(q,\dot{q},t)dt = \int_{t_0}^{t_1}\left(\frac{\partial L}{\partial q}\delta q(t) + \frac{\partial L}{\partial \dot{q}}\delta\dot{q}(t)\right)dt =$$
$$\int_{t_0}^{t_1}\left(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right)\delta q(t)dt + \frac{\partial L}{\partial \dot{q}}\delta q(t)\Big|_{t_0}^{t_1} = \frac{\partial L}{\partial \dot{q}}\delta q(t)\Big|_{t_0}^{t_1} = p(t_1)dq_1.$$

So we have

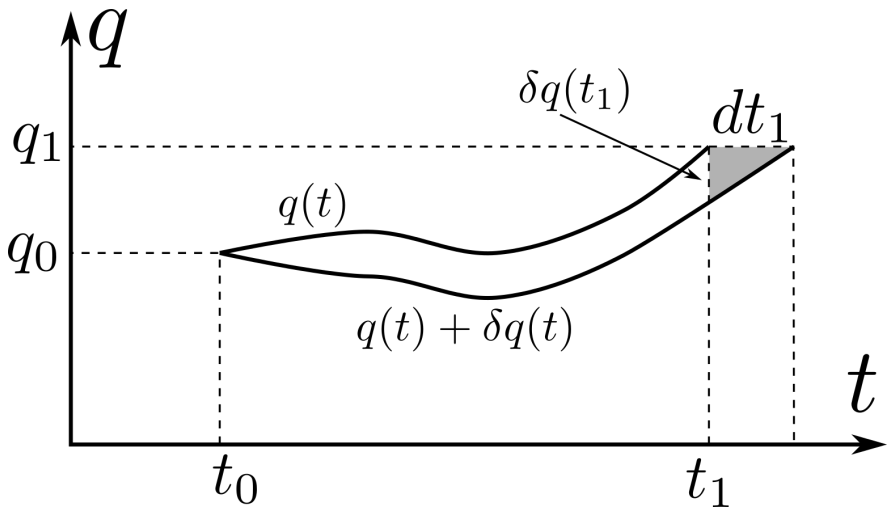
$$\frac{\partial \mathcal{S}}{\partial q} = p.$$

This is the result for the case of one degree of freedom. If we have N degrees of freedom, then

$$\frac{\partial \mathcal{S}}{\partial q_i} = p_i, \qquad i = 1 \ldots N.$$

- I want to emphasize ones more: $\mathcal{S}$ here is <u>not</u> a functional! We already substituted the solution in. It is here the function of the upper (and lower) boundary conditions.
- The momenta p_i thus obtained are the canonical momenta, as we used $p_i = \frac{\partial L}{\partial \dot{q}_i}$.

43.3. Energy.



- To simplify the calculation I will consider only the case of one dgree of freedom. If we have N degrees of freedom the calculation is almost identical. I will simply give the general result at the end.

We want to see how the value of the Action changes when we change t_1, while keeping q_1 (as well as q_0 and t_0) fixed.

Consider an action

$$\mathcal{S} = \int_{t_0}^{t_1} L(q, \dot{q}, t)dt, \qquad q(t_0) = q_0, \quad q(t_1) = q_1.$$

Consider the value of the action on the trajectory as a function of t_1.

Notice, that t_1 is there in two places: as the upper limit of integration and in the boundary condition. We do not change the value of $q = q_1$ at the upper limit! but the trajectory changes!. So we have

$$\mathcal{S}(t_1 + dt_1) = \int_{t_0}^{t_1+dt_1} L(q + \delta q, \dot{q} + \delta\dot{q}, t)dt = L(t_1)dt_1 + \int_{t_0}^{t_1} L(q + \delta q, \dot{q} + \delta\dot{q}, t)dt.$$

Using the usual trick we will get

$$d\mathcal{S} = L(t_1)dt_1 + p\delta q|_{t_0}^{t_1} = L(t_1)dt_1 + p(t_1)\delta q(t_1) = L(t_1)dt_1 - p(t_1)\dot{q}(t_1)dt_1,$$

where I used $\delta q(t_1) = -\dot{q}(t_1)dt_1$ — see picture.

So we have

$$\frac{\partial \mathcal{S}}{\partial t} = -H.$$

In the case of N degrees of freedom the result is exactly the same.

- Again, I emphasize, that $\mathcal{S}$ is not the functional here. It is the value of the functional on a trajectory. As such it is the function of the initial and final conditions.
- Notice, that this is all on a trajectory. So in the right hand side it is the value of the Hamiltonian on the trajectory, the energy at time t_1.

Putting the results of this and the previous sections together we can write

$$d\mathcal{S} = p_i dq_i - Hdt$$

where $\mathcal{S}$ is the value of the action as a function of the final coordinates q_i and time t.

43.4. Hamilton-Jacobi equation

We have on a trajectory

$$-\frac{\partial \mathcal{S}}{\partial t} = H(p, q, t),$$

but on a trajectory we also have $p = \frac{\partial \mathcal{S}}{\partial q}$, so we can write

$$-\frac{\partial \mathcal{S}}{\partial t} = H\left(\frac{\partial \mathcal{S}}{\partial q}, q, t\right).$$

This is a partial differential equation for the function $\mathcal{S}(q, t)$. This equation is called Hamilton-Jacobi equation.

We have derived it for $1D$, however, the derivation works exactly the same way for arbitrary number of degrees of freedom.

$$-\frac{\partial \mathcal{S}}{\partial t} = H\left(\frac{\partial \mathcal{S}}{\partial \mathbf{q}}, \mathbf{q}, t\right),$$

where the bold letters mean the collection of all degrees of freedom. The solution of this equation $\mathcal{S}(\mathbf{q}, t)$ depends on N coordinates $\{q_i\}$ and time t.

The function $\mathcal{S}(\mathbf{q}, t)$ at any moment of time defines a $N-1$ dimensional hypersurface $\mathcal{S}(\mathbf{q}, t) = \text{const.}$ in the N dimensional coordinate space — the space of all coordinates. With time this surface changes. One can imagine these as propagation of wave fronts — the Hamilton-Jacobi equation then is the non-linear wave equation. The momentum is the gradient of $\mathcal{S}$, so it is perpendicular to the surface of $\mathcal{S} = \text{const.}$. For a standard Newtonian world, where velocity has the same direction as momentum, the rays correspond to the trajectories.

Let's imagine, that we solved this equation and found the function $\mathcal{S}(\mathbf{q}, t, \alpha_1 \dots \alpha_N)$, where N is the number of the coordinates and α_i are N arbitrary constants. Let's see how $\frac{\partial \mathcal{S}}{\partial \alpha_i}$ depends on time.

$$\frac{d}{dt}\frac{\partial \mathcal{S}}{\partial \alpha_i} = \dot{q}_j \frac{\partial^2 \mathcal{S}}{\partial q_j \partial \alpha_i} + \frac{\partial}{\partial t}\frac{\partial \mathcal{S}}{\partial \alpha_i} = \dot{q}_j \frac{\partial^2 \mathcal{S}}{\partial q_j \partial \alpha_i} + \frac{\partial}{\partial \alpha_i}\frac{\partial \mathcal{S}}{\partial t} = \dot{q}_j \frac{\partial^2 \mathcal{S}}{\partial q_j \partial \alpha_i} - \frac{\partial}{\partial \alpha_i} H\left(\frac{\partial \mathcal{S}}{\partial \mathbf{q}}, \mathbf{q}, t\right)$$
$$= \dot{q}_j \frac{\partial^2 \mathcal{S}}{\partial q_j \partial \alpha_i} - \frac{\partial H}{\partial p_j}\frac{\partial^2 \mathcal{S}}{\partial q_j \partial \alpha_i} = \dot{q}_j \frac{\partial^2 \mathcal{S}}{\partial q_j \partial \alpha_i} - \dot{q}_j \frac{\partial^2 \mathcal{S}}{\partial q_j \partial \alpha_i} = 0.$$

Where we used the Hamilton-Jacobi equation and the canonical Hamiltonian equation $\frac{\partial H}{\partial p_j} = \dot{q}_j$. So we see, that

$$\frac{d}{dt}\frac{\partial \mathcal{S}}{\partial \alpha_i} = 0.$$

So all $\frac{\partial \mathcal{S}}{\partial \alpha_i}$ do not change with time and are constants. Then the N equations

$$\frac{\partial \mathcal{S}}{\partial \alpha_i} = \beta_i$$

are implicit definitions of the solutions of the equations of motions $q_j(t, \{\alpha_i\}, \{\beta_i\})$ that depend on $2N$ arbitrary constants, which are given by initial conditions.

43.5. Connection to quantum mechanics.

The quasiclassical approximation of quantum mechanics $\hbar \to 0$ transfers the Schrödinger equation into the Hamilton-Jacobi equation. We start with the Schrödinger equation.

$$i\hbar\frac{\partial}{\partial t}\Psi = \hat{H}\Psi$$

The Hamiltonian operator $\hat{H}$ is constructed in the following way:

- $H(p,x)$ is a polynomial of p.
- Substitute $p \to \hat{p} \equiv -i\hbar\frac{\partial}{\partial x}$ into the Hamiltonian and obtain the operator $\hat{H} = H(\hat{p},x)$. This operator is called Hamiltonian operator.

Now we want to take the classical limit $\hbar \to 0$ in the Schrödinger equation. We cannot simply put $\hbar = 0$ in the left hand side of the Schrödinger equation, as $\hbar$ is also inside the Hamiltonian and hence will be inside the function Ψ.

The limit $\hbar \to 0$ is done in the following way:

- Consider a function $\Psi = e^{\frac{i}{\hbar}\mathcal{S}(x,t)}$. So far $\mathcal{S}(x,t)$ is just some function of x and t.
- We first compute how the momentum operator $\hat{p}$ acts on this Ψ, namely $\hat{p}\Psi = -i\hbar\frac{\partial\Psi}{\partial x} = \Psi\frac{\partial\mathcal{S}}{\partial x}$.
- We also compute $\hat{p}^2\Psi = -i\hbar\frac{\partial^2\mathcal{S}}{\partial x^2}e^{\frac{i}{\hbar}\mathcal{S}} + \left(\frac{\partial\mathcal{S}}{\partial x}\right)^2 e^{\frac{i}{\hbar}\mathcal{S}} = -i\hbar\frac{\partial^2\mathcal{S}}{\partial x^2}\Psi + \left(\frac{\partial\mathcal{S}}{\partial x}\right)^2\Psi$.
- Notice, that at $\hbar \to 0$ we have $\hat{p}^2\Psi \to \Psi\left(\frac{\partial\mathcal{S}}{\partial x}\right)^2$.
- It is clear, that the same will happen for any (positive integer) power of $\hat{p}$, namely at $\hbar \to 0$ we have $\hat{p}^n\Psi \to \Psi\left(\frac{\partial\mathcal{S}}{\partial x}\right)^n$.
- As $H(p,x)$ is a polynomial of p we will have $\hat{H}\Psi = \Psi H\left(\frac{\partial\mathcal{S}}{\partial x}, x\right)$.
- Also $i\hbar\frac{\partial}{\partial t}\Psi = -\Psi\frac{\partial\mathcal{S}}{\partial t}$.
- Consider now the Schrödinger equation.

  $$i\hbar\frac{\partial}{\partial t}\Psi = \hat{H}\Psi$$

  in the limit $\hbar \to 0$ we substitute $i\hbar\frac{\partial}{\partial t}\Psi \to -\Psi\frac{\partial\mathcal{S}}{\partial t}$ and $\hat{H}\Psi \to \Psi H\left(\frac{\partial\mathcal{S}}{\partial x}, x\right)$, and obtain

  $$-\frac{\partial\mathcal{S}}{\partial t} = H\left(\frac{\partial\mathcal{S}}{\partial x}, x\right),$$

  which is the Hamilton-Jacobi equation.
- What we have shown is that in the classical limit $\hbar \to 0$ the Schrödinger equation turns into the the Hamilton-Jacobi equation. So the quantum mechanics becomes classical mechanics in the limit $\hbar \to 0$.
- We also notice, that the quantum mechanical wave function $\Psi = e^{\frac{i}{\hbar}\mathcal{S}}$. For $\hbar \to 0$ the main contribution to Ψ comes from the trajectories that minimize $\mathcal{S}$ — this is the Hamilton principle! — the classical trajectory is the trajectory which minimizes the Action! The origin of the Hamilton's principle is in quantum mechanics.

Closing remarks:

- That's it.
- Student evaluation 11-14-2024, until 12-04-2024.
- How much you have learned.