Krzysztof A. Meissner

Classical Mechanics

FACULTY OF PHYSICS
UNIVERSITY OF WARSAW

not for distribution

Warszawa, September 2024

These notes were written for the lectures of Classical Mechanics at the Undergraduate course at the Faculty of Physics at the University of Warsaw. They are based mostly on the books of L.D. Landau, J. Lifszyc it Mechanics and Wojciech Rubinowicz, Wojciech Królikowski, *Mechanika teoretyczna*. Some of the examples and derivations are taken from the DAMTP (University of Cambridge) lectures by David Tong (http://www.damtp.cam.ac.uk/user/tong/dynamics.html)

Contents

1 Newtonian mechanics

We all learned at school that the essence of classical mechanics is given by 3 laws of Newton. As I will try to argue during this set of lectures there are much better ways of formulating the classical dynamics. Before we start to describe these methods we formulate the 3 laws

- First law (given by Galileo)

  There exist reference frames (called inertial) in which a body, very distant from all other bodies, moves along a straight line with constant speed. One often encounters completely absurd definitions using the notion of force that is defined in the second law!

- Second law

  The notion of force $\mathbf{F}$ is defined as

  $$\mathbf{F} := \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} \tag{1.1}$$

  where momentum $\mathbf{p} := m\mathbf{v}$, m is 'amount of matter' and $\mathbf{v}$ is a velocity measured with respect to the inertial frame.

  1. This definition would be rather useless if not for the very fortunate fact that for two most important interactions, electromagnetic and gravitational, we can (approximately) give the expression of the LHS in terms of distances between bodies.

  2. the definition is valid also in relativistic physics but the definition of momentum changes.

  3. It is much better physically not to think in terms of forces but in terms of the flow of momenta.

- *'action is equal to reaction' – the body acting on another body with the force* $\mathbf{F}$ *is itself subject to the force* $-\mathbf{F}$ *from the other body*

  this law is a trivial application of the conservation of momentum and we will not use it in the following.

1.1 Definitions

The first law says that far away from any other bodies $\mathbf{a} = 0$ hence $\mathbf{F} \neq 0$ is a measure of interactions i.e no interactions $\Rightarrow$ force $= 0$. The arrow is only to the right since it may happen that even in the presence of interactions the flow of momenta is zero (for example when sitting on a chair).

Let us introduce some useful definitions.

For a system of bodies

$$\mathbf{P} = \sum_a \mathbf{p}_a \tag{1.2}$$

is a total momentum.

Center of mass definition

$$\mathbf{R} := \frac{\sum m_a \mathbf{r}_a}{\sum m_a} \tag{1.3}$$

Hence

$$M\dot{\mathbf{R}} = \sum m_a \mathbf{v}_a = \mathbf{P} \tag{1.4}$$

so the movement of the center of mass is uniquely given by the total momentum.

If the system is isolated i.e. $\mathbf{P}$=const then the center of mass moves with a constant velocity so its own (CM) reference frame is inertial and $\mathbf{R} = 0$ – no internal moves can change the position of the CM. It is usual to prove at this point that the RHS does not depend on the internal interactions (using $F_{ab} = -F_{ba}$ but the statement follows from the conservation of momentum and is general.

We introduce the notion of angular momentum

$$\mathbf{J}_a = \mathbf{r}_a \times \mathbf{p}_a \tag{1.5}$$

The total angular momentum is given by

$$\begin{aligned} \mathbf{J} &= \sum_a \mathbf{r}_a \times m_a \mathbf{v}_a = \sum_a (\mathbf{r}_a - \mathbf{R} + \mathbf{R}) \times m_a(\mathbf{v}_a - \mathbf{V} + \mathbf{V}) \\ &= \sum_a (\mathbf{r}_a - \mathbf{R}) \times m_a(\mathbf{v}_a - \mathbf{V}) + \mathbf{R} \times \mathbf{P} \end{aligned} \tag{1.6}$$

so it is given by the sum of CM angular momentum and the 'internal' angular momentum.

Differentiating $\mathbf{J}$ we get

$$\dot{\mathbf{J}} = \mathbf{r} \times \dot{\mathbf{p}} = \mathbf{N} \tag{1.7}$$

where we introduced moment of force

$$\mathbf{N} := \mathbf{r} \times \mathbf{F} \tag{1.8}$$

For a system of bodies we have

$$\dot{\mathbf{J}} = \sum_a \mathbf{r}_a \times \mathbf{F}_a \tag{1.9}$$

and the usual argument using the third Newton's law shows that $\mathbf{J}$ is conserved only when we have central forces i.e. $\mathbf{r}_a - \mathbf{r}_b$ is parallel to the force between a and b

$$\dot{\mathbf{J}} = \sum_a \mathbf{r}_a \times \left(\sum_{b \neq a} \mathbf{F}_{ab} \right) = \sum_{a,b,a<b} (\mathbf{r}_a \times F_{ab} + \mathbf{r}_b \times F_{ba}) = \sum_{a,b,a<b} (\mathbf{r}_a - \mathbf{r}_b) \times F_{ab} \quad (1.10)$$

But the conservation of $\mathbf{J}$ can be proved in much more general situations (by Noether's theorem to be discussed later) so we will not discuss it here.

It is however important to emphasize here the difference between the conservation of momentum and the conservation of angular momentum. The first gives as a corollary the impossibility to move center of mass position by means of internal forces only. The second does not have as a corollary that the angle with respect to some inertial frame cannot change using only internal forces and deformations, as a cat jumping and rotating clearly shows. We will discuss this issue later on.

1.2 Mechanical energy and potential

If the forces are independently given then

$$\mathbf{F}_a = m_a \frac{\mathrm{d}\mathbf{v}_a}{\mathrm{d}t} \quad (1.11)$$

and multiplying and summing we get

$$\sum \mathbf{F}_a \cdot \mathbf{v}_a = \frac{\mathrm{d}}{\mathrm{d}t} \left(\sum \frac{m_a v_a^2}{2} \right) \quad (1.12)$$

The sum on RHS is the total kinetic energy T.

$$T := \sum \frac{m_a v_a^2}{2} \quad (1.13)$$

Therefore integrating over time

$$T_f - T_i = \int \sum \mathbf{F}_a \cdot \mathbf{v}_a \mathrm{d}t = \int \sum \mathbf{F}_a \cdot \mathrm{d}\mathbf{r}_a \quad (1.14)$$

The most important class of forces are so called potential forces – when there exists a function $V(t, \mathbf{r}_, \ldots)$ such that

$$\mathbf{F}_a = -\nabla_a V(t, \mathbf{r}_1, \ldots) \quad (1.15)$$

If on top $V(t, \mathbf{r}_1, \ldots) = V(\mathbf{r}_1, \ldots)$ i.e. it does not depend explicitly on time the forces are called conservative.

Then

$$T_f - T_i = \int \sum_a \mathbf{F}_a \cdot \mathrm{d}\mathbf{r}_a = -\int \sum_a \nabla_a V(\mathbf{r}_1, \ldots) \cdot \mathrm{d}\mathbf{r}_a = -(V_f - V_i) \quad (1.16)$$

i.e.

$$E = T + V = \text{const} \tag{1.17}$$

so the total mechanical energy for conservative potentials is conserved (hence the name).

The total kinetic energy of a system is the sum of the CM kinetic energy and the internal kinetic energy

$$T = \sum_a \frac{m_a v_a^2}{2} = \sum_a \frac{m_a(\mathbf{v}_a - \mathbf{V} + \mathbf{V})^2}{2} = \sum_a \frac{m_a(\mathbf{v}_a - \mathbf{V})^2}{2} + \frac{MV^2}{2} \tag{1.18}$$

1.3 Non-potential forces

There are some forces that do not have any potential associated with them. The most common is a friction force. It is a clear example that the second law is useless if we don't know the force as a function of positions and velocities. The friction force has several approximate descriptions

- it is proportional to the normal force pressing the body to the surface with the so called coefficient of friction. It is impossible to calculate it from first principles, depends on many factors, roughness, humidity, history etc. It also depends on whether the body is at rest or moves (static and kinetic COF). Polishing the surfaces can make COF to grow and not to decrease and so on. Even the very notion of COF is an **approximate** description of the actual friction force because for larger pressures the friction force does not respond linearly!

- it is proportional to some power of velocity – usually used for friction in air or water. Here the situation is better since at least we have Navier-Stokes equations with the boundary condition that the relative velocity on the surface vanishes – the condition that does not depend sensitively on the roughness of the surface (although not totally independent, especially at larger speeds).

In any case the forces that are not of potential type are very phenomenological and much less interesting for physics with viscosity in fluids as the only exception. Therefore in what follows we will assume that the forces are of potential type and discuss fluids separately.

1.4 One dimensional motion

In one dimension with a potential independent of time we can solve the problem to the very end by quadratures.

We start with the conserved energy

$$E = T + U(x) = \frac{m\dot{x}^2}{2} + U(x) \tag{1.19}$$

Assuming that we are describing part of the trajectory with positive velocity we have

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \sqrt{\frac{2}{m}(E - U(x))} \tag{1.20}$$

hence

$$t - t_0 = \sqrt{\frac{m}{2}} \int \frac{\mathrm{d}x}{\sqrt{E - U(x)}} \tag{1.21}$$

The motion is possible only for those x for which $U(x) \leqslant E$. If there are two x i.e. $x_1(E)$ and $x_2(E)$ for which

$$U(x_i) = E \tag{1.22}$$

and in between $U(x) < E$ then particle stops there and (generically) starts to move back (it oscillates between x_1 and x_2). The period is equal to

$$T(E) = \sqrt{2m} \int\limits_{x_1(E)}^{x_2(E)} \frac{\mathrm{d}x}{\sqrt{E - U(x)}} \tag{1.23}$$

As an example let us consider a pendulum with

$$E = \frac{ml^2\dot{\theta}^2}{2} - mgl\cos\theta \tag{1.24}$$

Then writing $E = -mgl\cos\theta_0$ where θ_0 is the maximal angle we get

$$T = 4\sqrt{\frac{l}{2g}} \int\limits_0^{\theta_0} \frac{\mathrm{d}\theta}{\sqrt{\cos\theta - \cos\theta_0}} \tag{1.25}$$

Using well known formulae we get

$$T = 4\sqrt{\frac{l}{4g}} \int\limits_0^{\theta_0} \frac{\mathrm{d}\theta}{\sqrt{\sin^2\theta_0/2 - \sin^2\theta/2}} \tag{1.26}$$

Introducing $\sin(\theta/2) = \sin(\theta_0/2)\sin\xi$ we get

$$T = 4\sqrt{\frac{l}{g}} \int\limits_0^{\pi/2} \frac{\mathrm{d}\xi}{\sqrt{1 - \sin^2(\theta_0/2)\sin^2\xi}} = 4\sqrt{\frac{l}{g}} \int\limits_0^{\pi/2} \frac{\mathrm{d}\xi}{\sqrt{\cos^2\xi + \cos^2(\theta_0/2)\sin^2\xi}} \tag{1.27}$$

For small θ_0 we get

$$T = 4\sqrt{\frac{l}{g}} \int\limits_0^{\pi/2} \mathrm{d}\xi(1 + \frac{1}{8}\theta_0^2\sin^2\xi + \ldots) = 2\pi\sqrt{\frac{l}{g}}(1 + \frac{1}{16}\theta_0^2 + \ldots) \tag{1.28}$$

The full result is given by elliptic integrals.

There is a clever way of calculating this integral. We consider the integral

$$I(a,b) = \int_0^{\pi/2} \frac{\mathrm{d}\xi}{\sqrt{a^2\cos^2\xi + b^2\sin^2\xi}} \tag{1.29}$$

As we will show below we can change a and b into the arithmetical mean $(a+b)/2$ and geometrical one $\sqrt{ab}$, respectively, without changing the value of the integral

$$I(a,b) = I\left(\frac{a+b}{2}, \sqrt{ab}\right) \tag{1.30}$$

As it turns out the two means get closer to each other extremely quickly and converge to a common value a_∞. Then we get

$$I(a,b) = \int_0^{\pi/2} \frac{\mathrm{d}\xi}{\sqrt{a_\infty^2\cos^2\xi + b_\infty^2\sin^2\xi}} = \frac{\pi}{2a_\infty} \tag{1.31}$$

Proof:

$$I(a,b) = \int_0^{\pi/2} \frac{\mathrm{d}\xi}{\sqrt{a^2\cos^2\xi + b^2\sin^2\xi}} \tag{1.32}$$

We change the variable

$$x = b\tan\phi \tag{1.33}$$

Then

$$I(a,b) = \int_0^{\infty} \frac{\mathrm{d}x}{\sqrt{(a^2+x^2)(b^2+x^2)}} \tag{1.34}$$

We now introduce

$$x = \frac{1}{2}\left(t - \frac{ab}{t}\right), \qquad \mathrm{d}x = \frac{1}{2}\left(1 + \frac{ab}{t^2}\right)\mathrm{d}t \tag{1.35}$$

and calculate

$$\begin{aligned} \sqrt{x^2+ab} &= \frac{1}{2}\left(t + \frac{ab}{t}\right) \\ \sqrt{x^2 + \frac{(a+b)^2}{4}} &= \frac{1}{2t}\sqrt{(t^2+a^2)(t^2+b^2)} \end{aligned} \tag{1.36}$$

Then

$$\frac{1}{2}\int_0^{\infty} \frac{\mathrm{d}x}{\sqrt{(x^2+ab)(x^2+(a+b)^2/4)}} = \frac{1}{2}\int_0^{\infty} \frac{\mathrm{d}t}{\sqrt{(t^2+a^2)(t^2+b^2)}} \tag{1.37}$$

2 Lagrangian formalism

2.1 Hamilton's principle

It was noticed in the XIXth century that the classical trajectories can be formulated as variational problems i.e. they are extrema of some functional called action that is itself an integral over time of some function of positions and velocities called the lagrangian.

This observation led to the most fruitful formalisms in classical physics and points directly to the quantum physics as we will discuss.

We define a lagrangian as a function of positions x^A and velocities $\dot{x}^A$ where A runs over some finite set (for N particles it would be $A \in 1..3N$).

We define the lagrangian as

$$L(t, x^A, \dot{x}^A) = T(\dot{x}^A) - V(t, x^A) \tag{2.1}$$

i.e. as a difference between kinetic energy and potential energy.

We then assume that all trajectories that will be compared start at time t_i at the same point x_i^A and end at time t_f at the same point x_f^A

We define the action S as a functional

$$S = \int_{t_i}^{t_f} L(t, x^A, \dot{x}^A)\,\mathrm{d}t \tag{2.2}$$

so it depends upon the path between t_i and t_f.

Principle of Least Action says that the actual trajectory is such that it is the extremum of S.

We consider the actual path $x^A(t)$. If it is an extremum of S it means that any deviation from the trajectory does not change S up to terms linear in the deviation. We add the deviation

$$x^A(t) \to x^A(t) + \delta x^A(t) \tag{2.3}$$

and we calculate the change of the action for the perturbed trajectory (keeping the initial and final times and the end points of the trajectory unchanged)

$$\delta S = \delta \int_{t_i}^{t_f} L(t, x^A, \dot{x}^A)\,\mathrm{d}t = \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial x^A}\delta x^A + \frac{\partial L}{\partial \dot{x}^A}\delta \dot{x}^A \right)\,\mathrm{d}t \tag{2.4}$$

We integrate by parts and we get up to linear terms in δx^A

$$\delta S = \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial x^A} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}^A} \right) \right) \delta x^A \, \mathrm{d}t + \left[\frac{\partial L}{\partial \dot{x}^A} \delta x^A \right]_{t_i}^{t_f} \tag{2.5}$$

According to our assumption the endpoints of the trajectory are kept fixed so the last term vanishes. Since $\delta x^A(t)$ is arbitrary we conclude that for each A

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}^A} \right) - \frac{\partial L}{\partial x^A} = 0 \tag{2.6}$$

These equations are called Euler-Lagrange equations.

We see that adding a full time derivative to L does not change the equations of motion so we treat such lagrangians as equivalent:

$$L \equiv L + \frac{\mathrm{d}f}{\mathrm{d}t} \tag{2.7}$$

In the simplest case of one-dimensional particle

$$L = \frac{m\dot{x}^2}{2} - V(t, x) \tag{2.8}$$

the EL equations give

$$m\ddot{x} = -\frac{\mathrm{d}V}{\mathrm{d}x} \tag{2.9}$$

i.e. indeed the Newton equation.

For a free particle ($V = const$) we can see that such a lagrangian is invariant (up to total derivatives) under Galilean Transformations $\mathbf{v} \to \mathbf{v} + \mathbf{V}$ with $\mathbf{V} = \text{const}$:

$$L \to L + m\mathbf{v} \cdot \mathbf{V} + \frac{mV^2}{2} = L + \frac{\mathrm{d}}{\mathrm{d}t}\left(m\mathbf{r} \cdot \mathbf{V} + \frac{mV^2}{2} t \right) \tag{2.10}$$

However, the formulation in terms of Euler-Lagrange equations has several important advantages over the Newton formulation.

First of all it is a variational formulation what points directly to the quantum mechanical origin of these equations as we will discuss later.

Second, the equations look the same in all coordinate systems while the Newton equations are written down only in inertial frames (otherwise one has to add fictitious forces).

Let us prove the second feature

We change the coordinates x^A into y^A (with number of y equal to the number of xs) assuming this change to everywhere invertible. We have

$$\dot{x}^A = \frac{\partial x^A}{\partial y^B} \dot{y}^B + \frac{\partial x^A}{\partial t} \tag{2.11}$$

Then

$$\frac{\partial L}{\partial y^B} = \frac{\partial L}{\partial x^A}\frac{\partial x^A}{\partial y^B} + \frac{\partial L}{\partial \dot{x}^A}\left(\frac{\partial^2 x^A}{\partial y^B \partial y^C}\dot{y}^C + \frac{\partial^2 x^A}{\partial y^B \partial t}\right) \tag{2.12}$$

and

$$\frac{\partial L}{\partial \dot{y}^B} = \frac{\partial L}{\partial \dot{x}^A}\frac{\partial \dot{x}^A}{\partial \dot{y}^B} \tag{2.13}$$

Now we use the fact that $\frac{\partial \dot{x}^A}{\partial \dot{y}^B} = \frac{\partial x^A}{\partial y^B}$ what can be seen from (2.11). Therefore the EL equations read in the new coordinates

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{y}^B}\right) - \frac{\partial L}{\partial y^B} = \left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}^A}\right) - \frac{\partial L}{\partial x^A}\right)\frac{\partial x^A}{\partial y^B} \tag{2.14}$$

so they are equivalent to the original ones (assuming invertibility of the change).

2.2 Mechanical similarity and virial theorem

Assume that the potential (independent of time) has a property that

$$U(\alpha \mathbf{r}_1, \alpha \mathbf{r}_2, \ldots) = \alpha^k U(\mathbf{r}_1, \mathbf{r}_2, \ldots) \tag{2.15}$$

We substitute simultaneous change of time

$$\mathbf{r}_i \to \alpha \mathbf{r}_i, \qquad t \to \beta t \tag{2.16}$$

and require that the kinetic energy has the same factor in front as the potential i.e.

$$\frac{\alpha^2}{\beta^2} = \alpha^k \Rightarrow \beta = \alpha^{1-k/2} \tag{2.17}$$

Then the whole lagrangian is just multiplied by α^k i.e. all the EOM will be the same (with rescaled time and positions).

As an example let us quote the Coulomb potential $U = -\gamma/r$ Then

$$k = -1 \Rightarrow \beta = \alpha^{\frac{3}{2}} \tag{2.18}$$

Hence we recover Kepler's third law

$$\left(\frac{T'}{T}\right)^2 = \left(\frac{R'}{R}\right)^3 \tag{2.19}$$

Now we turn to another application – so called virial theorem.

Consider a bounded system of particles. The kinetic energy is a quadratic form of velocities so (even for non-diagonal case from the Euler theorem on homogeneous functions))

$$\sum_a \frac{\partial T}{\partial \mathbf{v}_a} = 2T \tag{2.20}$$

Introducing momenta

$$\mathbf{p}_a := \frac{\partial T}{\partial \mathbf{v}_a} \tag{2.21}$$

we can write

$$2T = \sum_a \mathbf{p}_a \cdot \mathbf{v}_a = \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_a \mathbf{p}_a \cdot \mathbf{r}_a\right) - \sum_a \mathbf{r}_a \cdot \dot{\mathbf{p}}_a \tag{2.22}$$

Let us take the average over time of this equality. The average of a full derivative tends to zero with growing time since

$$\bar{f} := \lim_{\tau\to\infty} \frac{1}{\tau}\int_0^\tau f(t)\mathrm{d}t \tag{2.23}$$

On the RHS we replace $\dot{\mathbf{p}}_a$ by derivatives of the potential and we get

$$2\overline{T} = \sum_a \overline{\mathbf{r}_a \frac{\partial U}{\partial \mathbf{r}_a}} \tag{2.24}$$

Using again our assumption on U we get the virial theorem

$$2\overline{T} = k\overline{U} \tag{2.25}$$

So that in terms of the total energy

$$\overline{T} = \frac{k}{k+2}E, \qquad \overline{U} = \frac{2}{k+2}E \tag{2.26}$$

The most famous example is the Coulomb potential where $k = -1$ and (with E negative)

$$\overline{T} = -E, \qquad \overline{U} = 2E \tag{2.27}$$

Extracting energy from the system (for example by radiation) gives more negative E so T grows – that's one of the reasons why the Sun gets hotter over time (the main one being a very sensitive dependence of the nuclear reactions in the Sun's core on the parameters of the core).

3 Examples

3.1 2-dim case in polar coordinates

We have

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2) - U(r,\phi) \tag{3.1}$$

EL eqs. read

$$\begin{aligned} m\ddot{r} &= mr\dot{\phi}^2 - \frac{\partial U}{\partial r} \\ \frac{\mathrm{d}}{\mathrm{d}t}(mr^2\dot{\phi}) &= -\frac{\partial U}{\partial \phi} \end{aligned} \tag{3.2}$$

The first equation includes the centrifugal force and the second is (in 3-dim notation)

$$\frac{\partial \mathbf{J}}{\partial t} = \mathbf{N} \tag{3.3}$$

since

$$-\nabla U = -\frac{\partial U}{\partial r}\mathbf{e}_r - \frac{1}{r}\frac{\partial U}{\partial \phi}\mathbf{e}_\phi \tag{3.4}$$

and

$$\mathbf{r} \times (-\nabla U) = -\frac{\partial U}{\partial \phi}\mathbf{e}_r \times \mathbf{e}_\phi \tag{3.5}$$

with

$$\mathbf{J} = mr^2\dot{\phi}\mathbf{e}_r \times \mathbf{e}_\phi \tag{3.6}$$

3.2 Reduced mass

If we have two bodies without any external interactions we can write

$$L = \frac{m_1\dot{\mathbf{r}}_1^2}{2} + \frac{m_2\dot{\mathbf{r}}_2^2}{2} - U(|\mathbf{r}_1 - \mathbf{r}_2|) \tag{3.7}$$

We know that without external interaction the CM moves with constant velocity. Therefore we choose CM system

$$m_1\mathbf{r}_1 + m_2\mathbf{r}_2 = 0 \tag{3.8}$$

and then we introduce

$$\mathbf{r} := \mathbf{r}_1 - \mathbf{r}_2 \tag{3.9}$$

Solving these two equations we have

$$\mathbf{r}_1 = \frac{m_2}{m_1+m_2}\mathbf{r}, \qquad \mathbf{r}_2 = -\frac{m_1}{m_1+m_2}\mathbf{r}, \tag{3.10}$$

Plugging these expressions back into the lagrangian we get

$$L = \frac{\mu \dot{\mathbf{r}}^2}{2} - U(r) \tag{3.11}$$

where the reduced mass μ

$$\mu = \frac{m_1 m_2}{m_1+m_2} \tag{3.12}$$

So the problem of two bodies boils down to the problem of one body with reduced mass.

3.3 Rotating reference frame

We have a free particle in 2 dimensions

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) \tag{3.13}$$

Introducing

$$x = x'\cos\omega t + y'\sin\omega t, \qquad y = -x'\sin\omega t + y'\cos\omega t, \tag{3.14}$$

we have

$$L = \frac{m}{2}\left((\dot{x}' - \omega y')^2 + (\dot{y}' + \omega x')^2\right) \tag{3.15}$$

The EOM read

$$\begin{aligned} \ddot{x}' - 2\omega\dot{y}' - \omega^2 x' &= 0 \\ \ddot{y}' + 2\omega\dot{x}' - \omega^2 y' &= 0 \end{aligned} \tag{3.16}$$

where we recognize the centrifugal force $\omega \times (\omega \times \mathbf{r})$ and the Coriolis force $2\omega \times \dot{\mathbf{r}}$.

3.4 Kepler orbits

We discuss a test body moving in the most important potential in 3 dimensions

$$U(r) = -\frac{GM}{r} \tag{3.17}$$

Let us recall the beautiful solution of the problem of orbits given by Laplace. We notice that the angular momentum $\mathbf{J}$ is conserved so the orbit has to lie in a plane perpendicular to $\mathbf{J}$. Introducing polar coordinates in this plane we have

$$\mathbf{r} = r\mathbf{e}_r \Rightarrow \mathbf{v} = \dot{r}\mathbf{e}_r + r\dot{\mathbf{e}}_r \tag{3.18}$$

hence

$$\mathbf{J} = m\mathbf{r} \times \mathbf{v} = mr^2 \mathbf{e}_r \times \dot{\mathbf{e}}_r \tag{3.19}$$

We now calculate the time derivative of $\mathbf{v} \times \mathbf{J}$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{v} \times \mathbf{J}) = \dot{\mathbf{v}} \times \mathbf{J} = -\frac{GM}{r^2}\mathbf{e}_r \times (mr^2 \mathbf{e}_r \times \dot{\mathbf{e}}_r) \tag{3.20}$$

We see that r^2 cancels out and using

$$\mathbf{e}_r \times (\mathbf{e}_r \times \dot{\mathbf{e}}_r) = -\dot{\mathbf{e}}_r \tag{3.21}$$

we get

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{v} \times \mathbf{J} - GMm\mathbf{e}_r) = 0 \tag{3.22}$$

so that

$$\mathbf{v} \times \mathbf{J} - GMm\mathbf{e}_r = GMm\mathbf{s} \tag{3.23}$$

where $\mathbf{s}$ is a constant vector.

Using this vector we can write

$$J^2 = \mathbf{J} \cdot (m\mathbf{r} \times \mathbf{v}) = m\mathbf{r} \cdot (\mathbf{v} \times \mathbf{J}) = GMm^2 \mathbf{r} \cdot (\mathbf{e}_r + \mathbf{s}) = GMm^2 r(1 + \epsilon \cos\phi) \tag{3.24}$$

where ϕ is an angle between $\mathbf{r}$ and $\mathbf{s}$ and ϵ is the length of $\mathbf{s}$ called the eccentricity. Therefore the orbit is given by

$$r = \frac{J^2}{GMm^2(1 + \epsilon\cos\phi)} = \frac{p}{1 + \epsilon\cos\phi}, \quad p = \frac{J^2}{GMm^2} \tag{3.25}$$

which is an ellipse with the semiaxis (obtained from $\sqrt{x^2 + y^2} + \epsilon x = p \Rightarrow (x(1-\epsilon^2) + p\epsilon)^2 + y^2(1-\epsilon^2) = p^2$)

$$a = \frac{p}{1-\epsilon^2}, \qquad b = a\sqrt{1-\epsilon^2} \tag{3.26}$$

The middle point is $a\epsilon$ from the focus.

Using

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{1}{2}\mathbf{r} \times \mathbf{v} = \frac{J}{2m} \tag{3.27}$$

we get

$$\pi ab = \frac{JT}{2m} \Rightarrow J^2T^2 = GMm^2 a(1-\epsilon^2)T^2 = 4m^2\pi^2 a^2 a^2 (1-\epsilon^2) \tag{3.28}$$

hence

$$a^3 = \frac{GM}{4\pi^2}T^2 \tag{3.29}$$

3.5 Central potential

We consider a general potential depending only on the distance $U(r)$ then

$$\mathbf{J} = \mathbf{r} \times \mathbf{p} \tag{3.30}$$

is conserved. We can choose the spherical coordinate system such that z is directed towards $\mathbf{J}$ and then the whole trajectory has to lie in the $\theta = \pi/2$ plane. We therefore neglect from now on the θ variable. We can write

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\phi}^2) - U(r) \tag{3.31}$$

The EOM for ϕ reads

$$mr^2\dot{\phi} = \text{const} \tag{3.32}$$

i.e. the conservation of $\mathbf{J}$ in these special coordinates. We shouldn't solve this equation for $\dot{\phi}$ and plug it back to L! But we can use the second conserved quantity i.e. energy - there we can do it.

$$E = \frac{m}{2}(\dot{r}^2 + r^2\dot{\phi}^2) + U(r) = \frac{m}{2}\dot{r}^2 + \frac{J^2}{2mr^2} + U(r) \tag{3.33}$$

Therefore

$$\dot{r} = \pm\sqrt{\frac{2}{m}(E - U(r)) - \frac{J^2}{m^2r^2}} \tag{3.34}$$

with the sign depending on the actual moment of motion. Therefore

$$t = \int \frac{\mathrm{d}r}{\sqrt{\frac{2}{m}(E - U(r)) - \frac{J^2}{m^2r^2}}} + \text{const} \tag{3.35}$$

or

$$\phi = \int \frac{J\mathrm{d}r}{r^2\sqrt{2m(E - U(r)) - \frac{J^2}{r^2}}} + \text{const} \tag{3.36}$$

3.6 Relativistic rocket

A rocket of (variable) mass m throws backwards Δm with velocity w ($\Delta m \neq -\mathrm{d}m$ in the relativistic case since it costs energy to throw what changes the mass of the rocket). The velocity of the rocket is v. We have to find the dependence of the rocket mass on v.

We start with conservation of momentum and energy

$$\begin{aligned} \mathrm{d}\left(\frac{mv}{\sqrt{1-\frac{v^2}{c^2}}}\right) &= \frac{u\Delta m}{\sqrt{1-\frac{u^2}{c^2}}} \\ \mathrm{d}\left(\frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}}\right) &= -\frac{\Delta mc^2}{\sqrt{1-\frac{u^2}{c^2}}} \end{aligned} \tag{3.37}$$

where

$$u = \frac{w - v}{1 - \frac{vw}{c^2}} \tag{3.38}$$

For nonrelativistic rocket we would get $\Delta m = -\mathrm{d}m$ as we would expect.

Therefore

$$\mathrm{d}\left(\frac{mv}{\sqrt{1-\frac{v^2}{c^2}}}\right) = -u\mathrm{d}\left(\frac{m}{\sqrt{1-\frac{v^2}{c^2}}}\right) \tag{3.39}$$

Expanding and multiplying by $\sqrt{1-\frac{v^2}{c^2}}$ we get

$$v\mathrm{d}m + m\mathrm{d}v + \frac{mv^2\mathrm{d}v}{(1-\frac{v^2}{c^2})c^2} = \frac{v-w}{1-\frac{vw}{c^2}}\left(\mathrm{d}m + \frac{mv\mathrm{d}v}{(1-\frac{v^2}{c^2})c^2}\right) \tag{3.40}$$

Reorganizing the terms we get a surprisingly simple equation

$$\frac{\mathrm{d}m}{m} = -\frac{\mathrm{d}v}{w(1-\frac{v^2}{c^2})} \tag{3.41}$$

with a solution

$$\frac{M}{m} = \left(\frac{1+\frac{v}{c}}{1-\frac{v}{c}}\right)^{\frac{c}{2w}} \tag{3.42}$$

For $v/c \to 0$ we get the well known result

$$\frac{M}{m} = \mathrm{e}^{\frac{v}{w}} \tag{3.43}$$

while for $w = c$ we get

$$v = c\frac{M^2 - m^2}{M^2 + m^2} \tag{3.44}$$

and for $m \to 0$ we get $v \to c$ as could be expected.

3.7 Three body problem

As is well known the three body problem is unsolved analytically - many great physicists tried to find a new integral of motion (besides energy and angular momentum) but with no success. Poincaré analyzing the system had the first idea of a chaos in deterministic systems. There are special solutions (like the 8-form solution of Christopher Moore in 1993) but generally we have to resort to numerical solutions.There are simple facts that can be drawn and we would like to point one of them.

We write the lagrangian as

$$L = T - U \tag{3.45}$$

where

$$T = \frac{m_1 v_1^2}{2} + \frac{m_2 v_2^2}{2} + \frac{m_3 v_3^2}{2} \tag{3.46}$$

and

$$U(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = -\frac{Gm_1m_2}{|\mathbf{r}_1 - \mathbf{r}_2|} - \frac{Gm_1m_3}{|\mathbf{r}_1 - \mathbf{r}_3|} - \frac{Gm_2m_3}{|\mathbf{r}_2 - \mathbf{r}_3|} \tag{3.47}$$

The total energy is conserved

$$E = T + V \tag{3.48}$$

Let us introduce (in the CM frame) the object

$$I = \frac{1}{2}\sum_i m_i r_i^2 \tag{3.49}$$

Differentiating twice wrt t we get

$$\frac{\mathrm{d}^2 I}{\mathrm{d}t^2} = \sum_i m_i v_i^2 + \sum_i \mathbf{r}_i \cdot \dot{\mathbf{v}}_i \tag{3.50}$$

Using EOM we can rewrite this as

$$\frac{\mathrm{d}^2 I}{\mathrm{d}t^2} = 2T + \sum_{ij} \mathbf{r}_i \cdot \frac{Gm_im_j}{r_{ij}^3}(\mathbf{r}_j - \mathbf{r}_i) \tag{3.51}$$

where the sum is over $i \neq j$. Expanding we get

$$\frac{\mathrm{d}^2 I}{\mathrm{d}t^2} = T + E = 2E - U \tag{3.52}$$

If $E < 0$ then

$$-U = T - E \geq -E \tag{3.53}$$

Then

$$\inf(r_{12}, r_{13}, r_{23}) \leq -\frac{G}{E}(m_1m_2 + m_1m_3 + m_2m_3) \tag{3.54}$$

If $E > 0$ then $\ddot{I}$ and hence I for large times can only grow.

$$I \geq I_1 + I_1'(t - t_1) + E(t - t_1)^2 \tag{3.55}$$

It has to go to infinity at large times so the trajectory has to be open.

3.8 Noether's theorem

Let us start with the definition of the constant of motion $G(q_a, \dot{q}_a, t)$:

$$\frac{\mathrm{d}G}{\mathrm{d}t} = \sum_i \left(\frac{\partial G}{\partial q_a}\dot{q}_a + \frac{\partial G}{\partial \dot{q}_a}\ddot{q}_a\right) + \frac{\partial G}{\partial t} = 0 \tag{3.56}$$

when the EOM can be used.

We have two straightforward examples.

Energy H

$$H := \sum_a \dot{q}_a \frac{\partial L}{\partial \dot{q}_a} - L \tag{3.57}$$

if L does not depend explicitly on t since then

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \sum_i \left(\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}_a} - \frac{\partial L}{\partial q_a} \right) \dot{q}_a = 0 \tag{3.58}$$

because of EOM.

If L does not depend on some q_b (but may depend on $\dot{q}_b$) for some b then

$$p_b := \frac{\partial L}{\partial \dot{q}_b} \tag{3.59}$$

is the constant of motion. The proof is straightforward.

The Noether's theorem is a generalization of these concepts.

If we have a one-parameter map

$$q_a(t) \to Q_a(s,t), \qquad Q_a(0,t) = q_a(t) \tag{3.60}$$

such that

$$\frac{\partial}{\partial s} L(Q_a(s,t), \dot{Q}_a(s,t), t) = 0 \tag{3.61}$$

we say that this map is a (continuous) symmetry of the theory (if time also changes under the map the argument has to be slightly generalized).

We have then

$$\begin{aligned} 0 &= \frac{\partial}{\partial s} L(Q_a(s,t), \dot{Q}_a(s,t), t) \Big|_{s=0} = \left(\frac{\partial L}{\partial q_a} \frac{\partial Q_a}{\partial s} + \frac{\partial L}{\partial \dot{q}_a} \frac{\partial \dot{Q}_a}{\partial s} \right)_{s=0} = \\ &= \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}_a} \frac{\partial Q_a}{\partial s} \right)_{s=0} \end{aligned} \tag{3.62}$$

so that

$$\left(\frac{\partial L}{\partial \dot{q}_a} \frac{\partial Q_a}{\partial s} \right)_{s=0} \tag{3.63}$$

is conserved along the trajectory.

Examples

- if the spatial translations are a symmetry

  $$\mathbf{r}_a \to \mathbf{r}_a + s\mathbf{n} \tag{3.64}$$

  then the conserved quantity is

  $$\sum_a \frac{\partial L}{\partial \dot{\mathbf{r}}_a} \cdot \mathbf{n} = \sum_a \mathbf{p}_a \cdot \mathbf{n} = \mathbf{P} \cdot \mathbf{n} \tag{3.65}$$

  i.e the total momentum in the direction $\mathbf{n}$. If $\mathbf{n}$ is arbitrary then the total momentum is conserved.

- if the spatial rotations around the axis $\mathbf{n}$ are a symmetry

$$\mathbf{r}_a \to \mathbf{r}_a + s\mathbf{n} \times \mathbf{r} \tag{3.66}$$

then the conserved quantity is

$$\sum_a \frac{\partial L}{\partial \dot{\mathbf{r}}_a} \cdot (\mathbf{n} \times \mathbf{r}_a) = \sum_a \mathbf{p}_a \cdot (\mathbf{n} \times \mathbf{r}_a) = \mathbf{n} \cdot (\mathbf{r}_a \times \mathbf{p}_a) = \mathbf{n} \cdot \mathbf{J} \tag{3.67}$$

i.e the total angular momentum in the direction $\mathbf{n}$. If $\mathbf{n}$ is arbitrary then the total angular momentum is conserved.

4 Lagrangian formalism with constraints

4.1 Types of constraints

In the previous lecture we have defined lagrangians and discussed their properties. It is very often the case that the variables are subject to constraints and such a situation requires special treatment.

There are several types of constraints

- holonomic
  - equalities (or two-sided constraints)

    $$f_\alpha(t, x^A) = 0, \quad \alpha = 1, \ldots, 3N - n \tag{4.1}$$

    or inequalities (one-side constraints)

    $$f_\alpha(t, x^A) \geqslant 0, \quad \alpha = 1, \ldots, 3N - n \tag{4.2}$$

    .
  - depending on time

    $$f_\alpha(t, x^A) = 0, \quad \alpha = 1, \ldots, 3N - n \tag{4.3}$$

    called reonomic or

    $$f_\alpha(x^A) = 0, \quad \alpha = 1, \ldots, 3N - n \tag{4.4}$$

    called scleronomic constraints
- non holonomic - all other like constraints that depend on velocities and cannot be integrated to ones depending only on positions

We will deal mostly with holonomic, two-sided, scleronomic constraints.

In the presence of constraints we define the lagrangian as

$$L_c = L(x^A, \dot{x}^A, t) + \sum_\alpha \lambda_\alpha f_\alpha(x^A, t) \tag{4.5}$$

where λ_α are auxiliary additional coordinates called Lagrange multipliers.

The EL equations wrt λ give indeed the constraint equations

$$\frac{\partial L_c}{\partial \lambda} = f_\alpha(x^A, t) = 0 \tag{4.6}$$

On the other hand the EL equations wrt to x^A have additional terms:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}^A}\right) - \frac{\partial L}{\partial x^A} = \lambda_\alpha \frac{\partial f_\alpha}{\partial x^A} \tag{4.7}$$

The RHS plays the role of additional forces coming from the presence of constraints.

If the potential U and f_α's are independent of time then the energy is conserved

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_A \dot{x}^A \frac{\partial L}{\partial \dot{x}^A} - L\right) = \sum_A \dot{q}_A \sum_\alpha \lambda_\alpha \frac{\partial f_\alpha}{\partial q_A} = -\sum_\alpha \lambda_\alpha \frac{\partial f_\alpha}{\partial t} \tag{4.8}$$

where we used $\frac{\mathrm{d}f_\alpha}{\mathrm{d}t} = 0$.

4.1.1 2-dim pendulum of length d

$$L_c = \frac{1}{2}m(\dot{z}^2 + \dot{x}^2) + mgz + \lambda(z^2 + x^2 - d^2) \tag{4.9}$$

One equation is of course the constraint equation

$$z^2 + x^2 - d^2 = 0 \tag{4.10}$$

The two other read

$$m\ddot{z} - mg = 2\lambda z, \qquad m\ddot{x} = 2\lambda x \tag{4.11}$$

Substituting

$$x = d\sin\theta, \qquad z = d\cos\theta \tag{4.12}$$

we identically satisfy the constraint equation and for the other two we get (after multiplying by sin or cos and adding/subtracting)

$$\begin{aligned} -md\ddot{\theta} - mg\sin\theta &= 0 \\ -md\dot{\theta}^2 - mg\cos\theta &= 2\lambda d \end{aligned} \tag{4.13}$$

The first one is the usual equation along the constraints hypersurface, the second one gives the reaction force perpendicular to constraints hypersurface.

4.2 Reduced lagrangians

It is very useful that if we are not interested in the reaction forces and want to solve the equations only along the constraints we can do so in an 'easy' way (this is the usual physical approach while the reaction forces usually have to be calculated in technical applications). Instead of x^A let us introduce new coordinates

$$x^A \to q_a, f_\alpha \tag{4.14}$$

If the constraints are all independent (at least in the vicinity of some point in the configuration space) one can introduce n independent variables $q_1, \dots, q_n$ and express all x^A in terms of q_i by solving these constraint equations.

$$x^A = x^A(q_1, \dots, q_n) \tag{4.15}$$

and plug these solutions to the equations of motion.

As we have proven the EL eqs. are independent of the choice of coordinates so we can immediately write

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q_a}}\right) - \frac{\partial L}{\partial q_a} = \lambda_\alpha \frac{\partial f_\alpha}{\partial q_a} \tag{4.16}$$

But in these coordinates f_α are coordinates by construction independent of q_a so the RHS vanishes. Therefore the lagrangian in terms of q_a can be considered as self-contained and the solutions will be automatically along the constraints hypersurfaces (but we cannot calculate from it the reaction from constraints).

4.3 Lagrange points

Imagine two large bodies m_1 and m_2 circulating on a circular orbit around each other. We describe the system in the rotating CM frame i.e.

$$\omega^2 = \frac{G(m_1 + m_2)}{d^3} \tag{4.17}$$

where d is the distance between the bodies. The bodies are

$$r_1 = \frac{d\mu}{m_1}, \qquad r_2 = \frac{d\mu}{m_2} \tag{4.18}$$

from the CM. The equation for ω^2 comes from the equality

$$m_1 \omega^2 r_1 = \frac{G m_2 m_1}{d^2} \tag{4.19}$$

Now we add a third very small body m_3, $m_3 \ll m_1, m_2$ and ask about the points in space where there is effectively no force from the two large bodies. We immediately see that the small body has to lie in the plane orthogonal to the rotation (otherwise it would be attracted by both large bodies). Introducing the rotating frame with x axis joining large bodies and y orthogonal to it (but in the plane of rotation) we write the lagrangian for the small body

$$L = \frac{m_3}{2}\left((\dot{x} - \omega y)^2 + (\dot{y} + \omega x)^2\right) + \frac{G m_1 m_3}{r_{13}} + \frac{G m_2 m_3}{r_{23}} \tag{4.20}$$

where

$$r_{13} = \sqrt{(x + d\mu/m_1)^2 + y^2}, \qquad r_{23} = \sqrt{(x - d\mu/m_2)^2 + y^2} \tag{4.21}$$

We write the EOM

$$\begin{aligned} m_3(\ddot{x} - \omega\dot{y}) &= m_3\omega(\dot{y} + \omega x) - \frac{Gm_1m_3(x + d\mu/m_1)}{r_{13}^3} - \frac{Gm_2m_3(x - d\mu/m_2)}{r_{23}^3} \\ m_3(\ddot{y} + \omega\dot{x}) &= -m_3\omega(\dot{x} - \omega y) - \frac{Gm_1m_3y}{r_{13}^3} - \frac{Gm_2m_3y}{r_{23}^3} \end{aligned} \tag{4.22}$$

We are looking for points (x, y) for which and $\dot{x} = \dot{y} = 0$ and $\ddot{x} = \ddot{y} = 0$ therefore

$$\begin{aligned} 0 &= \omega^2 x - \frac{Gm_1(x + d\mu/m_1)}{r_{13}^3} - \frac{Gm_2(x - d\mu/m_2)}{r_{23}^3} \\ 0 &= \omega^2 y - \frac{Gm_1y}{r_{13}^3} - \frac{Gm_2y}{r_{23}^3} \end{aligned} \tag{4.23}$$

- If $y = 0$ the second equation is trivially satisfied and we are left with

$$\omega^2 x = \frac{Gm_1(x + d\mu/m_1)}{|x + d\mu/m_1|^3} + \frac{Gm_2(x - d\mu/m_2)}{|x - d\mu/m_2|^3} \tag{4.24}$$

There are 3 solutions to this equation in the intervals $x < -d\mu/m_1$, $-d\mu/m_1 < x < d\mu/m_2$ and $x > d\mu/m_2$ (in each one solution). They are called L_2, L_1 and L_3 and one can show that they are unstable i.e. deviation from these points makes the acceleration pointing away from these points. They are used for satellites orbiting the Sun together with the Earth since being unstable they don't gather cosmic dust.

For example if $x > x_3$ (i.e. on the right of L_3) we have from the first equation $\ddot{x} > 0$ since

$$\omega^2 x\ \uparrow, \quad \frac{Gm_1(x + d\mu/m_1)}{r_{13}^3}\ \downarrow, \ \frac{Gm_2(x - d\mu/m_2)}{r_{23}^3}\ \downarrow \tag{4.25}$$

so m_3 is repelled from L_3.

- if $y \neq 0$ the second equation gives

$$\omega^2 = \frac{Gm_1}{r_{13}^3} + \frac{Gm_2}{r_{23}^3} \tag{4.26}$$

Multiplying it by x and adding to the first equation we get

$$r_{13} = r_{23} = d \tag{4.27}$$

where the last equation comes from the equation for ω^2. So we have two points forming equilateral triangle in the rotation plane. They are called L_4, and L_5 and one can show that they are stable i.e. deviation from these points makes the acceleration pointing back to these points. There are Kordylewski clouds (1961) around L5 not yet fully confirmed.

4.4 Orbits in the Schwarzschild metric

we start with the lagrangian in the Schwarzschild metric (for $\theta = \pi/2$)

$$S = -mc^2 \int \mathrm{d}t \sqrt{1 - \frac{r_g}{r} - \frac{\dot{r}^2}{c^2(1-\frac{r_g}{r})} - \frac{r^2\dot{\phi}^2}{c^2}} \tag{4.28}$$

where

$$r_g = \frac{2GM}{c^2} \tag{4.29}$$

We calculate the momenta

$$\begin{aligned} p_r &= \frac{m\dot{r}}{(1-r_r/r)\sqrt{1 - \frac{r_g}{r} - \frac{\dot{r}^2}{c^2(1-\frac{r_g}{r})} - \frac{r^2\dot{\phi}^2}{c^2}}} \\ p_\phi &= \frac{mr^2\dot{\phi}}{\sqrt{1 - \frac{r_g}{r} - \frac{\dot{r}^2}{c^2(1-\frac{r_g}{r})} - \frac{r^2\dot{\phi}^2}{c^2}}} = J \end{aligned} \tag{4.30}$$

Then the energy

$$E = \sum_i \dot{q}_i p_i - L = \frac{mc^2(1-r_g/r)}{\sqrt{1 - \frac{r_g}{r} - \frac{\dot{r}^2}{c^2(1-\frac{r_g}{r})} - \frac{r^2\dot{\phi}^2}{c^2}}} \tag{4.31}$$

Calculating $\dot{\phi}$ from J

$$\dot{\phi} = \frac{(1-r_g/r)Jc^2}{Er^2} \tag{4.32}$$

we get

$$E^2 - \left(\frac{r'^2}{1-r_g/r} + r^2\right)\frac{(1-r_g/r)J^2c^2}{r^4} = m^2c^4(1-r_g/r) \tag{4.33}$$

Writing

$$E = \tilde{E} + mc^2 \tag{4.34}$$

we get

$$\left(\tilde{E} + \frac{GMm}{r}\right)\frac{2m}{J^2} + \frac{\tilde{E}^2}{J^2c^2} + \frac{2GM}{r^3c^2} = \frac{r'^2 + r^2}{r^4} \tag{4.35}$$

Introducing $w = 1/r$ and differentiating we get

$$\frac{GMm^2}{J^2} + \frac{3GMw^2}{c^2} = w'' + w \tag{4.36}$$

This equation is exact and leads to the rotation of the perihelion of planets (for Mercury 42' per hundred years).

Without the second term on the LHS we would get the Kepler orbits $r = p/(1 + \epsilon\cos\phi)$. With the second term (very small) we substitute

$$w = A + B\cos(\alpha\phi) \tag{4.37}$$

From the classical solution we have

$$w = \frac{(1+\epsilon\cos(\alpha\phi))}{p}, \quad p = \frac{J^2}{GMm^2} = a(1-\epsilon^2) \tag{4.38}$$

where the large and small axis

$$a = \frac{p}{1-\epsilon^2}, \qquad b = a\sqrt{1-\epsilon^2} \tag{4.39}$$

Plugging this solution into (4.35) we get the coefficient in front of $\cos(\alpha\phi)$

$$1-\alpha^2 = \frac{6GM}{pc^2} = \frac{3r_S}{a(1-\epsilon^2)}, \qquad r_S = \frac{2GM}{c^2} \tag{4.40}$$

Hence

$$\alpha \sim 1 - \frac{3r_S}{2a(1-\epsilon^2)} \tag{4.41}$$

so that

$$\delta = \frac{3\pi r_S}{a(1-\epsilon^2)} \tag{4.42}$$

For the Mercury $T = 88$ days (100 years $\sim$ 415 rotations), $a = 57.9$ mln km, $\epsilon = 0.206$ so it gives $\delta \sim 43.5''/100$ years.

Using the same formula (4.35) we can derive the equation for the trajectory of light ($m = 0$).

$$\frac{\tilde{E}^2}{J^2c^2} + \frac{2GM}{r^3c^2} = \frac{r'^2 + r^2}{r^4} \tag{4.43}$$

where $\tilde{E} = h\nu$ and J are measured far away from the Sun. We therefore have an exact equation

$$w'' + w = \frac{3r_s w^2}{2} \tag{4.44}$$

Constant r is possible when

$$r_{ph} = \frac{3}{2}r_s \tag{4.45}$$

but this trajectory is unstable: $w = 2/(3r_s) + \epsilon\sinh\phi + O(\epsilon^2)$.

To derive the bending of light formula we start from $r_s = 0$ i.e.

$$w = \frac{\cos\phi}{r_0} \tag{4.46}$$

Using this on the RHS we get to the first order in r_s

$$w'' + w = \frac{3r_s\cos^2(\phi)}{2r_0^2} \tag{4.47}$$

with a solution

$$w = \frac{\cos\phi}{r_0 + r_s/2} + \frac{r_s}{2r_0^2}(1 + \sin^2\phi) \tag{4.48}$$

so that $w = 0$ for

$$\phi = \pm(\frac{\pi}{2} + \delta), \qquad \delta = \frac{r_s}{r_0} \tag{4.49}$$

so that the bending is $2r_s/r_0$. The exact solution is given in terms of the Weierstrass function that will be discussed later.

We can derive the difference in time (with respect to the far-away observer) for the Earth and for the GPS satellites. We have

$$\mathrm{d}t' = \mathrm{d}t\sqrt{1 - \frac{r_g}{r} - \frac{\dot{r}^2}{c^2(1 - \frac{r_g}{r})} - \frac{r^2\dot{\phi}^2}{c^2}} \tag{4.50}$$

where $r_g = 8.75$ mm. Assuming that we are on the Equator the time runs slower by

$$r = 6.4 \cdot 10^6 \text{ m}, \quad T = 86400 \text{ s} \Rightarrow \Delta t = 60\ \mu\text{s/day} \tag{4.51}$$

For the GPS satellites (making full circle in 12 hours) the time runs slower by

$$r = 26.6 \cdot 10^6 \text{ m}, \quad T = 43200 \text{ s} \Rightarrow \Delta t = 22\ \mu\text{s/day} \tag{4.52}$$

so the time runs slower on Earth than in satellites by 38 μs/day where 45 μs/day comes from general relativity (satellites are further away from the center of the Earth than the surface) and -7 μs/day from special relativity (satellites are faster than the surface of the Earth)

5 Oscillations

5.1 Many-body problem

Lagrangian can be written in general as

$$L = \frac{1}{2} M_{ab} \dot{q}_a \dot{q}_b - U(q) \tag{5.1}$$

where M is real, positive, symmetric, constant matrix. We can assume that M is diagonal. Let us assume that there exists an extremum of U i.e. at some point q_0 all derivatives of U vanish. So we can write in the vicinity of q_0

$$q_a(t) = q_{0a} + \eta(t) \tag{5.2}$$

and expand the EOM up to $O(\eta)$ (in matrix notation)

$$M\ddot{\eta} = -V\eta \Rightarrow \ddot{\eta} = -M^{-1}V\eta \tag{5.3}$$

where

$$V_{ab} = \left. \frac{\partial^2 U}{\partial q_a \partial q_b} \right|_{q=q_0} \tag{5.4}$$

We now look for eigenvalues of this equation. Let us first prove that they are real. We assume that for some η_k we have

$$\ddot{\eta}_k = -\lambda_k^2 \eta_k \tag{5.5}$$

where λ_k^2 and η_k can be a priori real or complex. We rewrite it as

$$-M^{-1}V\eta_k = -\lambda_k^2 \eta_k \tag{5.6}$$

so that after we multiply by $\bar{\eta}_k^T$ we get

$$\bar{\eta}_k^T V \eta_k = \lambda_k^2 \bar{\eta}_k^T M \eta_k \tag{5.7}$$

Since V and M are real symmetric matrices λ_k^2 has to be real as well. Therefore all eigenvectors can also be chosen real.

Now we distinguish two situations

- all λ_k^2 positive - the system is stable

$$\eta(t) = \sum_k A_k \eta_k \cos(\lambda_k (t - t_k)) \tag{5.8}$$

- one or more λ_k^2 is negative - the system is unstable in the direction of the eigenvector η_k.

$$\eta(t) = A_k \eta_k \left(\exp(\lambda_k (t - t_k)) + \exp(-\lambda_k (t - t_k)) \right) + \dots \tag{5.9}$$

5.2 Forced oscillations

We discuss here the problem of oscillations (with friction) under the external periodic force

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = A\cos\omega t \tag{5.10}$$

As always the general solution is given by a sum of a special solution of the inhomogeneous eq. and a general solution of the homogeneous eq. It is more convenient to write in the complex form

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = A\mathrm{e}^{\mathrm{i}\omega t} \tag{5.11}$$

Substituting

$$x_s(t) = B\mathrm{e}^{\mathrm{i}\omega t} \tag{5.12}$$

we find the special solution:

$$B = \frac{A}{-\omega^2 + \omega_0^2 + 2\mathrm{i}\gamma\omega} \tag{5.13}$$

The real part of the solution solves the original problem i.e.

$$x_s(t) = \frac{A}{\sqrt{(-\omega^2 + \omega_0)^2 + 4\gamma^2\omega^2}}\cos(\omega t + \delta) \tag{5.14}$$

where

$$\tan\delta = \frac{2\gamma\omega}{-\omega^2 + \omega_0^2} \tag{5.15}$$

5.3 Parametric resonance

Let us discuss the problem of solutions of a one-dimensional oscillator with variable parameters (for example mass or the moment of inertia for the pendulum). We can write

$$\frac{\mathrm{d}}{\mathrm{d}t}(m(t)\dot{x}) + k(t)x = 0 \tag{5.16}$$

If we introduce different time variable $\mathrm{d}\tau = \mathrm{d}t/m(t)$ we have $\frac{\mathrm{d}^2x}{\mathrm{d}\tau^2} + mkx = 0$ so that we don't lose generality if we consider

$$\ddot{x} + \omega^2(t)x = 0 \tag{5.17}$$

We assume that $\omega(t)$ is periodic with some period T i.e.

$$\omega(t+T) = \omega(t) \tag{5.18}$$

If we have two independent solutions of (5.17) then the property (5.18) requires that each $x_i(t+T)$ has to be a linear combination of these two solutions. We can always diagonalize this relation and choose these combinations in such a way that

$$x_1(t+T) = \mu_1 x_1(t), \quad x_2(t+T) = \mu_2 x_2(t) \tag{5.19}$$

where now in general μ_i and x_i can be complex as a result of diagonalization (if they are complex then necessarily $\mu_2^\star = \mu_1$ since $\omega^2(t)$ is assumed to be real). We assume that μ_i are not simultaneously equal to 1. The procedure depends on the fact that any square matrix B can be diagonalized by PBP^{-1} with P possibly complex, the only exception being when some eigenvalues have multiplicity > 1 – then it is possible that the resulting matrix is of Jordan form. Here we assume that two eigenvalues are distinct (otherwise they would have to be both equal to 1, see below)

There is a relation between μ_1 and μ_2 coming from the Wronskian of x_1 and x_2:

$$\frac{\mathrm{d}}{\mathrm{d}t}(\dot{x}_1 x_2 - \dot{x}_2 x_1) = 0 \Rightarrow \dot{x}_1 x_2 - \dot{x}_2 x_1 = \text{const} \tag{5.20}$$

But the LHS for $t \to t + T$ gets multiplied by $\mu_1\mu_2$ so we get

$$\mu_1\mu_2 = 1 \tag{5.21}$$

Therefore

$$\begin{aligned} \mu_i \text{ complex} \quad &\Rightarrow \quad |\mu_i| = 1,\ \ \mu_2 = \mu_1^\star \\ \mu_i \text{ real} \quad &\Rightarrow \quad \mu_2 = \frac{1}{\mu_1} \end{aligned} \tag{5.22}$$

If μ_i are complex their norm is one so the solutions just rotate after $t \to t+T$. However, if they are real then one of them (say, μ_1) is bigger than 1. It means that after nT it gets the factor μ_1^n i.e. it grows exponentially with time - then such a phenomenon bears the name 'parametric resonance'.

Let us discuss this phenomenon in a very well known example known from childhood - the see-saw. We very well remember that to make the amplitude bigger one has to make the leg movements with twice bigger frequency than the proper frequency of the see-saw. Let us substitute

$$\omega^2(t) = \omega_0^2(1 + h\cos(2\omega_0 + \varepsilon)t) \tag{5.23}$$

where h is small and $\varepsilon \ll \omega_0$. We substitute two independent solutions in the form

$$x = a(t)\cos(\omega_0 + \varepsilon/2)t + b(t)\sin(\omega_0 + \varepsilon/2)t \tag{5.24}$$

where $a(t$ and $b(t)$ change slowly in time, much slower than ω_0. Substituting this form and neglecting $\ddot{a}$, $\ddot{b}$ and $\cos(3\omega_0 t), \sin(3\omega_0 t)$ we get

$$-(2\dot{a} + b\varepsilon + \frac{h\omega_0}{2}b)\omega_0\sin(\omega_0 + \varepsilon/2)t + (2\dot{b} - a\varepsilon + \frac{h\omega_0}{2}a)\omega_0\cos(\omega_0 + \varepsilon/2)t = 0 \tag{5.25}$$

The functions in front of both have to be simultaneously equal to 0. We assume that

$$(a(t),\ b(t)) \sim \mathrm{e}^{st}(A, B) \tag{5.26}$$

and look for solution with $s > 0$ (and there also should be accompanying solution with $s < 0$). We get

$$s^2 = \frac{1}{4}\left[\left(\frac{h\omega_0}{2}\right)^2 - \varepsilon^2\right] \tag{5.27}$$

And indeed for

$$-\frac{h\omega_0}{2} < \varepsilon < \frac{h\omega_0}{2} \tag{5.28}$$

we have real solutions and in that interval there exists the phenomenon of parametric resonance.

If we include friction we can write

$$\ddot{x} + 2\gamma\dot{x} + \omega^2(t)x = 0 \tag{5.29}$$

We introduce

$$y(t) = \mathrm{e}^{-\gamma t}x(t) \tag{5.30}$$

and we get

$$\ddot{y} + (\omega^2(t) - \gamma^2)y = 0 \tag{5.31}$$

We can repeat the steps done before while replacing $\omega_0 \to \omega_\gamma = \sqrt{\omega_0^2 - \gamma^2}$ and then we get

$$(s-\lambda)^2 = \frac{1}{4}\left[\left(\frac{h\omega_\gamma}{2}\right)^2 - \varepsilon^2\right] \tag{5.32}$$

therefore we have parametric resonance if

$$-\sqrt{\left(\frac{h\omega_\gamma}{2}\right)^2 - 4\lambda^2} < \varepsilon < \sqrt{\left(\frac{h\omega_\gamma}{2}\right)^2 - 4\lambda^2} \tag{5.33}$$

There is also a possibility of the parametric resonance if $\omega = 2\omega_0/n$ but both the exponent s and the allowed width shrink as h^n i.e. are then much smaller. We discuss below the case $n = 2$:

$$\omega^2(t) = \omega_0^2(1 + h\cos(\omega_0 + \varepsilon)t) \tag{5.34}$$

and we substitute (note the shift in $x(t)$!)

$$x = a(t)\cos(\omega_0 + \varepsilon)t + b(t)\sin(\omega_0 + \varepsilon)t + c(t) \tag{5.35}$$

Assuming $a, b, c \sim \exp(st)$ and neglecting $\sin(\cos)(2\omega_0 + 2\epsilon)$ we get

$$c = -\frac{ha}{2}, \quad -2sa\omega_0 - 2b\omega_0\epsilon = 0, \quad -2a\omega_0\epsilon + 2sb\omega_0 - \frac{h^2\omega_0^2}{2}a = 0 \tag{5.36}$$

what gives

$$4s^2 + 4\epsilon^2 + h^2\omega_0\epsilon = 0 \tag{5.37}$$

Hence

$$s \in \mathbb{R} \quad \text{if} \quad -\frac{h^2\omega_0}{4} \leqslant \epsilon \leqslant 0 \tag{5.38}$$

so that indeed $s, \epsilon \sim h^2$.

The same parametric resonance is responsible for Faraday waves.

6 Rigid bodies

6.1 Inertia tensor

The kinetic term for a body rotating with the angular velocity ω

$$T = \frac{1}{2}\sum_a m_a \dot{\mathbf{r}}_a^2 = \frac{1}{2}\sum_a m_a(\omega \times \mathbf{r}_a)^2 = \frac{1}{2}\omega_i I^{ij}\omega_j \tag{6.1}$$

where the inertia tensor is given by

$$I^{ij} = \sum_a m_a(r_a^2\delta^{ij} - r_a^i r_a^j) \tag{6.2}$$

or for a continuous distribution

$$I^{ij} = \int \mathrm{d}^3 r \rho(\mathbf{r})(r^2\delta^{ij} - r^i r^j) \tag{6.3}$$

A symmetric real matrix can alway be diagonalized i.e. there exists an orthogonal coordinate system in which I is real and diagonal, moreover all eigenvalues are in this case non-negative (one can consider $b_i I^{ij} b_j$ what is obviously $\geqslant 0$ for arbitrary vector $\mathbf{r}$ to see this).

The sum of the eigenvalues is given by

$$\delta^{ij} I_{ij} = 2\int \mathrm{d}^3 r \rho(\mathbf{r}) r^2 \tag{6.4}$$

For example we can get the eigenvalues for the ball of radius R

$$3I_1 = 2\int \mathrm{d}^3 r \rho r^2 = \frac{8\pi}{5}\rho R^5 \Rightarrow I_1 = \frac{2}{5}MR^2 \tag{6.5}$$

For a disc and the axis perpendicular to the disk we have

$$I_3 = \rho\int r\mathrm{d}r\mathrm{d}\phi r^2 = \frac{\pi R^4\rho}{2} = \frac{MR^2}{2} \tag{6.6}$$

while for the axis in the plane of the disc

$$I_1 = I_2 = \frac{MR^2}{4} \tag{6.7}$$

since the sum has to be equal to $2\int 2\pi r\mathrm{d}r\rho r^2 = MR^2$.

If we measure I_0 wrt center of mass then wrt to any other axis there is a simple formula

$$I^{ij} = I_0^{ij} + M(c^2\delta^{ij} - c^i c^j) \tag{6.8}$$

where $\mathbf{c}$ is a vector connecting CM with the new axis.

The angular momentum is given by

$$\mathbf{J} = \sum_a m_a \mathbf{r}_a \times \dot{\mathbf{r}}_a = \sum_a m_a \mathbf{r}_a \times (\omega \times \mathbf{r}_a) = \sum_a m_a (r_a^2 \omega - (\omega \cdot \mathbf{r}_a)\mathbf{r}_a) \tag{6.9}$$

hence

$$J_i = I_{ij}\omega_j \tag{6.10}$$

Hence $\mathbf{J}$ does not have to coincide with ω and it leads sometimes to a very 'strange' motion.

6.2 Euler equations

Using a rotating coordinate frame and introducing the principal axes of the inertia tensor $\mathbf{e}_i$ in this frame with the eigenvalues I_i we can write

$$\mathbf{J} = \sum_i I_i \omega_i \mathbf{e}_i \tag{6.11}$$

Differentiating it wrt time we get

$$\sum_i I_i \dot{\omega}_i \mathbf{e}_i + \sum_i I_i \omega_i (\omega \times \mathbf{e}_i) = \mathbf{N} \tag{6.12}$$

where $\mathbf{N}$ is the moment of force.

We get in components

$$I_j \dot{\omega}_j + \sum_k I_i \omega_i \omega_k \epsilon_{ijk} = N_j \tag{6.13}$$

i.e.

$$\begin{aligned} I_1\dot{\omega_1} + (I_3 - I_2)\omega_3\omega_2 &= N_1 \\ I_2\dot{\omega_2} + (I_1 - I_3)\omega_1\omega_3 &= N_2 \\ I_3\dot{\omega_3} + (I_2 - I_1)\omega_2\omega_1 &= N_3 \end{aligned} \tag{6.14}$$

They are called Euler equations.

6.2.1 Free body

Let us analyze these equations for a free body ($\mathbf{N} = 0$).

If we multiply each of them by the respective ω_i we get the conserved energy

$$\frac{I_1\omega_1^2}{2} + \frac{I_2\omega_2^2}{2} + \frac{I_3\omega_3^2}{2} = E = \text{const} \tag{6.15}$$

If we multiply each of them by the respective $I_i\omega_i$ we get the square of the conserved angular momentum

$$I_1^2\omega_1^2 + I_2^2\omega_2^2 + I_3^2\omega_3^2 = J^2 = \text{const} \tag{6.16}$$

so we have 2 conserved quantities and only one Euler equation is independent. Analysis of such systems led to the theory of elliptic functions in the XIXth century. Let us note that

$$2E = \mathbf{J}\cdot\omega \tag{6.17}$$

Since both E and $\mathbf{J}$ are constant it means that the projection of ω on the direction of $\mathbf{J}$ is constant.

We now consider 3 cases

- When $I_1 = I_2 = I_3$ i.e. spherical body we have all $\omega_i = \text{const}$ and ω is in the direction of $\mathbf{J}$.

- when $I_1 = I_2$ then $\omega_3 = \text{const}$ and we arrive at the equation for $\omega_{1,2}$

  $$\dot\omega_1 - \omega\omega_2 = 0, \qquad \dot\omega_2 + \omega\omega_1 = 0, \tag{6.18}$$

  where

  $$\omega = \left(1 - \frac{I_3}{I_1}\right)\omega_3 \tag{6.19}$$

  Therefore

  $$\omega_1 = \omega_0 \sin\omega t, \quad \omega_2 = \omega_0\cos\omega t \tag{6.20}$$

  so that in the body frame ω precesses around $\mathbf{e}_3$ with angular frequency ω in different directions depending on whether $I_1 < I_3$ or $I_1 > I_3$

- when all of them are different $I_1 < I_2 < I_3$ we will discuss only the case when only one of the initial ω_i is large and two other very small.

  If $\omega_1 = \Omega$ is large and two other small (δ_2 and δ_3) then neglecting quadratic terms we get

  $$\delta_2 = A_2\sin\xi t, \quad \delta_3 = A_3\cos\xi t \tag{6.21}$$

  where

  $$\xi = \sqrt{\frac{(I_2-I_1)(I_3-I_1)}{I_2 I_3}}\Omega \tag{6.22}$$

  so it is stable.

  The same situation is when $\omega_3 = \Omega$ is large and two other small (δ_1 and δ_2)

  $$\delta_1 = A_1\cos\xi t, \quad \delta_2 = A_2\sin\xi t \tag{6.23}$$

  where

  $$\xi = \sqrt{\frac{(I_3-I_2)(I_3-I_1)}{I_1 I_2}}\Omega \tag{6.24}$$

and

$$A_2 = \sqrt{\frac{I_1(I_3 - I_1)}{I_2(I_3 - I_2)}} A_1 \tag{6.25}$$

so it is also stable.

We consider now the third case when $\omega_2 = \Omega$ is large and two other small (δ_1 and δ_3). Then

$$\delta_1 = A_1 \cosh \xi t, \qquad \delta_3 = A_3 \sinh \xi t \tag{6.26}$$

where

$$\xi = \sqrt{\frac{(I_3 - I_2)(I_2 - I_1)}{I_1 I_3}} \Omega \tag{6.27}$$

and

$$A_3 = -\sqrt{\frac{I_1(I_2 - I_1)}{I_3(I_3 - I_2)}} A_1 \tag{6.28}$$

so it unstable. Numerical analysis shows that later the nonlinear terms take over and finally the motion is it is periodic (with the period given approximately by $T \sim 1/\xi$ the precise value given by an elliptic integral) and 'jumps' from Ω to $-\Omega$ while two other are large at the jump.

7 Rigid bodies part II

7.1 Euler angles

We introduce now the description in the space frame X, Y, Z (and not in the body frame as before). The Euler angles are defined as subsequent rotation around z axis by ϕ then around new x' axis by θ and then again around the new axis $z''$by ψ (see the picture).

To unwind the rotation we use the matrix (note the signs of angles, reverse to the usual ones) wrt the Z axis, then X axis and again Z azis:

$$R(\psi,\theta,\phi) \;=\; \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Multiplying we get

$$R(\psi,\theta,\phi) \;=\; \begin{pmatrix} \cos\psi\cos\phi - \sin\psi\cos\theta\sin\phi & \cos\psi\sin\phi + \sin\psi\cos\theta\cos\phi & \sin\theta\sin\psi \\ -\sin\psi\cos\phi - \cos\psi\cos\theta\sin\phi & -\sin\psi\sin\phi + \cos\psi\cos\theta\cos\phi & \sin\theta\cos\psi \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{pmatrix}$$

We can use this matrix to unwind the body frame unit vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ to the space unit vectors $(\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z)$. Therefore

$$R(\psi,\theta,\phi)(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3) = \mathbf{I} \quad\Rightarrow\quad (\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3) = R(\psi,\theta,\phi)^{-1} = R(\psi,\theta,\phi)^T \tag{7.1}$$

so we have

$$\mathbf{e}_1 = \begin{pmatrix} \cos\psi\cos\phi - \sin\psi\cos\theta\sin\phi \\ \cos\psi\sin\phi + \sin\psi\cos\theta\cos\phi \\ \sin\theta\sin\psi \end{pmatrix} \tag{7.2}$$

$$\mathbf{e}_2 = \begin{pmatrix} -\sin\psi\cos\phi - \cos\psi\cos\theta\sin\phi \\ -\sin\psi\sin\phi + \cos\psi\cos\theta\cos\phi \\ \sin\theta\cos\psi \end{pmatrix} \tag{7.3}$$

$$\mathbf{e}_3 = \begin{pmatrix} \sin\theta\sin\phi \\ -\sin\theta\cos\phi \\ \cos\theta \end{pmatrix} \tag{7.4}$$

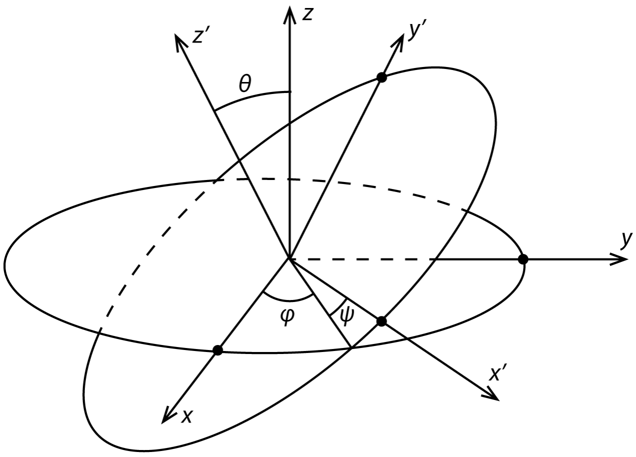

Fig. Euler angles

In the body frame we have

$$\omega = \omega_1 \mathbf{e}_1 + \omega_2 \mathbf{e}_2 + \omega_3 \mathbf{e}_3 \tag{7.5}$$

To get the expression for ω_i we use the fact that

$$\dot{\mathbf{e}}_i = \omega \times \mathbf{e}_i \tag{7.6}$$

so that for example

$$\dot{\mathbf{e}}_3 = \omega \times \mathbf{e}_3 = -\omega_1 \mathbf{e}_2 + \omega_2 \mathbf{e}_1 \tag{7.7}$$

therefore

$$\mathbf{e}_2 \cdot \dot{\mathbf{e}}_3 = -\omega_1, \mathbf{e}_1 \cdot \dot{\mathbf{e}}_3 = \omega_2 \tag{7.8}$$

Calculating the above expressions we arrive at

$$\omega = (\dot{\phi} \sin\theta \sin\psi + \dot{\theta}\cos\psi)\mathbf{e}_1 + (\dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi)\mathbf{e}_2 + (\dot{\psi} + \dot{\phi}\cos\theta)\mathbf{e}_3 \tag{7.9}$$

We can also express $\mathbf{J}$ in the space frame (for simplicity only in the case $I_1 = I_2$) using the previous expressions for $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. We get

$$\mathbf{J} = \begin{pmatrix} I_3\dot{\psi}\sin\theta\sin\phi + I_1\dot{\theta}\cos\phi + (I_3 - I_1)\dot{\phi}\cos\theta\sin\theta\sin\phi \\ -I_3\dot{\psi}\sin\theta\cos\phi + I_1\dot{\theta}\sin\phi + (I_1 - I_3)\dot{\phi}\sin\theta\cos\theta\cos\phi \\ I_3(\dot{\phi} + \dot{\psi}\cos\theta) + (I_1 - I_3)\dot{\phi}\sin^2\theta \end{pmatrix} \tag{7.10}$$

7.2 Wobbling plate

If we apply the formulae to the wobbling plate with $I_1 = I_2 = MR^2/4$ and $I_3 = MR^2/2$ we know that the frequency (in the body frame) ω_3 is constant and that ω_1 and ω_2 rotate with frequency

$$\Omega = \left(1 - \frac{I_3}{I_1}\right)\omega_3 \tag{7.11}$$

with

$$\omega_1 = \omega_0 \sin \Omega t, \qquad \omega_2 = \omega_0 \cos \Omega t \tag{7.12}$$

We have

$$\omega_1^2 + \omega_2^2 = \omega_0^2 \tag{7.13}$$

but on the other hand it is equal to

$$\omega_1^2 + \omega_2^2 = \dot\phi^2 \sin^2\theta + \dot\theta^2 \tag{7.14}$$

If we choose $\mathbf{J}$ to lie in the Z axis then $\dot\theta = 0$, $\theta = \theta_0$ (we can see it from the expression for $J_x \cos\phi + J_y \sin\phi = I_1\dot\theta = 0$). Then $\omega_0 = \dot\phi \sin\theta_0$, $\dot\phi$ is constant. From

$$\omega_1 = \dot\phi \sin\theta_0 \sin\psi = \omega_0 \sin\Omega t \tag{7.15}$$

hence

$$\Omega = \dot\psi \tag{7.16}$$

so $\dot\psi$ is also constant. Hence

$$\omega_3 = \dot\psi + \dot\phi\cos\theta_0 = \Omega + \dot\phi\cos\theta = \omega_3 - \frac{I_3}{I_1}\omega_3 + \dot\phi\cos\theta_0 \tag{7.17}$$

so that

$$\dot\phi = \frac{I_3\omega_3}{I_1 \cos\theta_0} = \frac{2\omega_3}{\cos\theta_0} \tag{7.18}$$

For small θ_0 the plate wobbles with twice the frequency of rotation.

For the Earth

$$\frac{I_3 - I_1}{I_3} \approx \frac{1}{300} \tag{7.19}$$

so we would expect the period of wobbling 300 days. It is actually around 430 days with the ω precessing around the North pole with radius about 10 m (but rather irregularly).

7.3 Heavy top

We now consider a symmetric top spinning in the gravitational field on its tip. The rotation is counted form the tip so both I_1 and I_2 are bigger by Ml^2 from the usual inertia coefficients calculated at CM. Let us write the lagrangian using ω_i but treating Euler angles as the fundamental variables

$$L = \frac{1}{2}I_1(\omega_1^2 + \omega_2^2) + \frac{1}{2}I_3\omega_3^2 + Mgl(1 - \cos\theta) \tag{7.20}$$

We use the expressions for ω_i to get

$$L = \frac{1}{2}I_1(\dot\phi^2 \sin^2\theta + \dot\theta^2) + \frac{1}{2}I_3(\dot\psi + \dot\phi\cos\theta)^2 + Mgl(1 - \cos\theta) \tag{7.21}$$

We see that there are 2 conserved momenta

$$J_3 = \frac{\partial L}{\partial \dot{\psi}} = I_3(\dot{\psi} + \dot{\phi}\cos\theta) = I_3\omega_3 = \text{const} \tag{7.22}$$

and

$$J_z = \frac{\partial L}{\partial \dot{\phi}} = I_1 \sin^2\theta \dot{\phi} + I_3\omega_3\cos\theta = \text{const} \tag{7.23}$$

We can solve for $\dot{\phi}$

$$\dot{\phi} = \frac{J_z - J_3\cos\theta}{I_1\sin^2\theta} \tag{7.24}$$

We also have the conserved energy

$$E = \frac{1}{2}I_1(\dot{\phi}^2\sin^2\theta + \dot{\theta}^2) + \frac{1}{2}I_3\omega_3^2 - Mgl(1-\cos\theta) \tag{7.25}$$

Let us rewrite this expression using the constants

$$\tilde{E} = \frac{1}{2}I_1\dot{\theta}^2 + \frac{(J_z - J_3\cos\theta)^2}{2I_1\sin^2\theta} - Mgl(1-\cos\theta) = \frac{1}{2}I_1\dot{\theta}^2 + U_{eff}(\theta) \tag{7.26}$$

where $\tilde{E} = E - \frac{1}{2}I_3\omega_3^2$.

If $J_3 \neq J_z$ then $U_{eff}(\theta) \to \infty$ for both $\theta \to 0$ and $\theta \to \pi$. Therefore there must be a minimum in between and θ oscillates between some θ_1 and θ_2 (so called nutation). The behavior of $\dot{\phi}$ (precession) depends on whether the sign of $J_z - J_3\cos\theta$ changes between θ_1 and θ_2 or not.

The question of stability in the vertical position $\theta = 0$ (then $J_z = J_3$) can be answered by expansion in θ. First we have

$$\Omega = \dot{\phi} = \frac{J_3}{2I_1} \tag{7.27}$$

and then

$$U_{eff}(\theta) \sim \frac{J_3^2}{8I_1}\theta^2 - \frac{Mgl}{2}\theta^2 \tag{7.28}$$

so that the motion is stable if

$$\omega_3^2 > \frac{4I_1Mgl}{I_3^2} \tag{7.29}$$

7.4 Balancing car wheels

We assume that we want to keep the axis fixed and ask how large momentum of force has to be exerted to arrive at this. We orient the axis of rotation along the z axis. Then the angular momentum reads

$$\mathbf{J} = \sum_a \mathbf{r}_a \times (m_a\mathbf{v}_a) = \sum_a \mathbf{r}_a \times (m_a\omega \times \mathbf{r}_a) = \sum_a m_a(\omega(r_a^2) - \mathbf{r}_a(\omega\cdot\mathbf{r}_a)) \tag{7.30}$$

what in components reads

$$(J_x, J_y, J_z) = (-\sum_a m_a z_a x_a, -\sum_a m_a z_a y_a, \sum_a m_a (x_a^2 + y_a^2)) \tag{7.31}$$

Assuming that

$$x_a = R\cos(\omega t + \phi_a), \quad y_a = R\sin(\omega t + \phi_a) \tag{7.32}$$

we get the moment of force needed to keep the axis unmoved

$$(N_x, N_y, N_z) = (\sum_a m_a z_a R\omega \sin(\omega t + \phi_a), -\sum_a m_a z_a R\omega \cos(\omega t + \phi_a), 0) \tag{7.33}$$

Measuring N_x, N_y at time $t = 0$ we get

$$N_x = \sum_a m_a z_a R\omega \sin(\phi_a), \quad N_y = -\sum_a m_a z_a R\omega \cos(\phi_a) \tag{7.34}$$

To balance the wheel i.e. cancel the moment of force we have to add some mass M at the point (z, ϕ) (ϕ is defined wrt (N_x, N_y)) such that

$$-MzR\omega \cos(\phi) + N_y = 0, \quad -MzR\omega \sin(\phi) - N_x = 0 \tag{7.35}$$

so that $\tan\phi = -N_x/N_y$ and $M = \sqrt{N_x^2 + N_y^2}/(zR\omega)$.

8 Hamiltonian formalism

8.1 Legendre transform and the Hamilton's equations

For the lagrangian we have the EL equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_a}\right) - \frac{\partial L}{\partial q_a} = 0 \tag{8.1}$$

Now we want to treat symmetrically q_a and $\dot{q}_a$. We introduce momenta

$$p_a := \frac{\partial L}{\partial \dot{q}_a} \tag{8.2}$$

It would not be correct to solve these equations and plug them back into the lagrangian. What we have to do is to make the Legendre transform.

8.2 Legendre transform

If we have a function $f(x)$ we introduce an additional variable s and we create a function

$$\tilde{f}(s,x) = sx - f(x) \tag{8.3}$$

Then

$$\mathrm{d}\tilde{f}(s,x) = x\mathrm{d}s + s\mathrm{d}x - \frac{\mathrm{d}f}{\mathrm{d}x}\mathrm{d}x \tag{8.4}$$

The differential depends on two variables unless we impose

$$s = \frac{\mathrm{d}f}{\mathrm{d}x} \tag{8.5}$$

and then we can treat $\tilde{f}$ as a function of only s

$$\mathrm{d}\tilde{f}(s) = x(s)\mathrm{d}s \tag{8.6}$$

where $x(s)$ is a solution of (8.5) and we have

$$\frac{\mathrm{d}\tilde{f}}{\mathrm{d}s} = x(s) \tag{8.7}$$

The inverse transform is

$$f(x) = xs - \tilde{f} \tag{8.8}$$

The transform

As an example take

$$f(x) = \mathrm{e}^{x/a} \tag{8.9}$$

then

$$s = \frac{1}{a}\mathrm{e}^{x/a} \quad \Rightarrow \quad x = a\ln(as) \tag{8.10}$$

Therefore

$$\tilde{f}(s) = as\ln(as) - as \tag{8.11}$$

and indeed $\tilde{f}' = x(s)$. We see that the domain of $f(x)$ i.e the whole real line $\mathbb{R}$ is different from the domain of $\tilde{f}(s)$ which is $\mathbb{R}_+$.

8.3 Hamilton's equations

We now apply the Legendre transform to the lagrangian replacing all $\dot{q}_a$'s by momenta. We write the transformed function

$$H(q_a, p_a, t) = \sum_a p_a\dot{q}_a - L(q_a, \dot{q}_a, t) \tag{8.12}$$

where all $\dot{q}_a$ are expressed as functions of p_a and q_a. Then

$$\mathrm{d}H = \dot{q}_a\mathrm{d}p_a + p_a\mathrm{d}\dot{q}_a - \left(\frac{\partial L}{\partial q_a}\mathrm{d}q_a + \frac{\partial L}{\partial \dot{q}_a}\mathrm{d}\dot{q}_a + \frac{\partial L}{\partial t}\right) \tag{8.13}$$

Using the definition of p_a and the EL equations we get

$$\mathrm{d}H = \dot{q}_a\mathrm{d}p_a - \dot{p}_a\mathrm{d}q_a - \frac{\partial L}{\partial t}\mathrm{d}t \tag{8.14}$$

and therefore we get the Hamilton's equations

$$\dot{q}_a = \frac{\partial H}{\partial p_a}, \qquad \dot{p}_a = -\frac{\partial H}{\partial q_a}, \qquad -\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t} \tag{8.15}$$

8.4 Examples

8.4.1 A particle in a potential

$$L = \frac{m\dot{\mathbf{r}}^2}{2} - V(\mathbf{r}) \tag{8.16}$$

Then

$$\mathbf{p} = m\dot{\mathbf{r}} \tag{8.17}$$

and

$$H = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}) \tag{8.18}$$

Then

$$\dot{\mathbf{r}} = \frac{\mathbf{p}}{m}, \qquad \dot{\mathbf{p}} = -\nabla V \tag{8.19}$$

8.4.2 Particle in rotating frame

We recall the lagrangian

$$L = \frac{m}{2}((\dot{x} - \omega y)^2 + (\dot{y} + \omega x)^2) \tag{8.20}$$

Therefore

$$\begin{aligned} p_x &= m(\dot{x} - \omega y) \\ p_y &= m(\dot{y} + \omega x) \end{aligned}$$

then

$$H = \dot{x}p_x + \dot{y}p_y - L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - \frac{m\omega^2(x^2 + y^2)}{2} = \frac{p_x^2 + p_y^2}{2m} + p_x\omega y - p_y\omega x \tag{8.21}$$

The HE read

$$\begin{aligned} \dot{p}_x &= \omega p_y \\ \dot{p}_y &= -\omega p_x \end{aligned}$$

therefore

$$\begin{aligned} \dot{p}_x &= m\ddot{x} - m\omega\dot{y} = p_y\omega = m\omega\dot{y} + m\omega^2 x \\ \dot{p}_y &= m\ddot{y} + m\omega\dot{x} = -p_x\omega = -m\omega\dot{x} + m\omega^2 y \end{aligned}$$

i.e. the expressions for the Coriolis and centrifugal forces.

8.4.3 Particle in an electromagnetic field

We start with the lagrangian

$$L = \frac{m\dot{\mathbf{r}}^2}{2} - q(\phi - \dot{\mathbf{r}} \cdot \mathbf{A}) \tag{8.22}$$

Then

$$\mathbf{p} = m\dot{\mathbf{r}} + q\mathbf{A} \quad \Rightarrow \quad \dot{\mathbf{r}} = \frac{\mathbf{p} - q\mathbf{A}}{m} \tag{8.23}$$

and

$$H = \mathbf{p} \cdot \dot{\mathbf{r}} - L = \frac{(\mathbf{p} - q\mathbf{A})^2}{2m} + q\phi \tag{8.24}$$

Calculating the momentum HE we get (in components)

$$\dot{p}_i = m\ddot{r}_i + q\dot{A}_i = \frac{q(p_j - qA_j)\partial_i A_j}{m} - q\partial_i\phi \tag{8.25}$$

Hence

$$m\ddot{r}_i = q\dot{r}_j(\partial_i A_j - \partial_j A_i) + q\dot{r}_j\partial_j A_i - q\partial_i\phi - q\dot{r}_j\partial_j A_i - q\dot{A}_i \tag{8.26}$$

what using $B_i = \epsilon_{ijk}\partial_j A_k$ and $E_i = -\partial_i\phi - \dot{A}_i$ is just the Lorentz force.

8.4.4 Relativistic particle in de Sitter space

We start from the lagrangian ($c = 1$)

$$L = -m\sqrt{1 - e^{2H_\Lambda t} v^2} \tag{8.27}$$

Then the momenta

$$p^i = \frac{m e^{2H_\Lambda t} v^i}{\sqrt{1 - e^{2H_\Lambda t} v^2}} = \text{const} \tag{8.28}$$

Therefore

$$H = \sum_i p^i v^i - L = \frac{m}{\sqrt{1 - e^{2H_\Lambda t} v^2}} = \sqrt{m^2 + p^2 e^{-2H_\Lambda t}} \tag{8.29}$$

The Hamilton equations

$$\dot{q}^i = \frac{\partial H}{\partial p^i} = \frac{p^i e^{-2H_\Lambda t}}{\sqrt{m^2 + p^2 e^{-2H_\Lambda t}}} \tag{8.30}$$

It can be integrated to give

$$q^i = q_0^i + \frac{p^i}{p^2 H_\Lambda}\left(\sqrt{m^2 + p^2} - \sqrt{m^2 + p^2 e^{-2H_\Lambda t}}\right) \tag{8.31}$$

Therefore the range is finite even after infinite time. The range of photons ($m = 0$) is equal to H_Λ^{-1}.

8.4.5 Relativistic particle in the Radiation Dominated Universe

We start from the lagrangian ($c = 1$)

$$L = -m\sqrt{1 - \frac{t}{t_0} v^2} \tag{8.32}$$

Then the momenta

$$p^i = \frac{m \frac{t}{t_0} v^i}{\sqrt{1 - \frac{t}{t_0} v^2}} = \text{const} \tag{8.33}$$

Therefore

$$H = \sum_i p^i v^i - L = \frac{m}{\sqrt{1 - \frac{t}{t_0} v^2}} = \sqrt{m^2 + p^2 \frac{t_0}{t}} \tag{8.34}$$

The Hamilton equations

$$\dot{q}^i = \frac{\partial H}{\partial p^i} = \frac{p^i \frac{t_0}{t}}{\sqrt{m^2 + p^2 \frac{t_0}{t}}} \tag{8.35}$$

The trajectory of photons ($m = 0$) is given by $ds^2 = 0$ – the solution is given by

$$x(t) = 2\sqrt{t t_0} \tag{8.36}$$

so the range of photons is infinite.

8.5 Conservation laws in the Hamiltonian formalism

We start with the conservation of energy. If H does not depend explicitly on time

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial q_a}\dot{q}_a + \frac{\partial H}{\partial p_a}\dot{p}_a + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t} \tag{8.37}$$

If some coordinate is cyclic (i.e. H does not depend on this coordinate) then the corresponding momentum is conserved

$$\dot{p}_a := \frac{\partial H}{\partial q_a} = 0 \tag{8.38}$$

8.6 Principle of Least Action

For the lagrangian we had the principle that the action

$$S = \int_{t_1}^{t_2} \mathrm{d}t L(q_a, \dot{q}_a, t) \tag{8.39}$$

is extremal when the variations δq_a vanish at the ends. We now have a similar principle

$$S = \int_{t_1}^{t_2} \mathrm{d}t (p_a \dot{q}_a - H(q_a, \dot{q}_a, t)) \tag{8.40}$$

where $\dot{q}_a$'s are functions of q_a and p_a. We have

$$\delta S = \int_{t_1}^{t_2} \left(\delta p_a \dot{q}_a + p_a \delta \dot{q}_a - \left(\frac{\partial H}{\partial p_a} \delta p_a - \frac{\partial H}{\partial q_a} \delta q_a \right) \right) \tag{8.41}$$

Integrating by parts we get

$$\int_{t_1}^{t_2} \left[\left(\dot{q}_a - \frac{\partial H}{\partial p_a} \delta p_a \right) + \left(-\dot{p}_a - \frac{\partial H}{\partial q_a} \delta q_a \right) \right] + p_a \delta q_a |_{t_1}^{t_2} \tag{8.42}$$

If the variations δq_a vanish at the ends we get the Hamilton's equations.

If we impose not only δq_a vanishing at the ends but also δp_a we can add to H a full derivative $\mathrm{d}F(p,q)/\mathrm{d}t$.

8.7 Adiabatic invariants

It is sometimes possible to find a set of invariants of the motion i.e. entities satisfying

$$\{I_i, H\} = 0 \tag{8.43}$$

If we introduce them as our coordinates it means that the conjugate variables θ_i are cyclic (the coordinates are then action-angle variables). If the number of such invariants is equal to the number of variables we say that the system is integrable. One dimensional systems with conserved energy are therefore always integrable , in more dimensions it is very rare.

Independently of the eqs of motion we can have objects that vary very little when we change the hamiltonian. We introduce some parameter $\lambda(t)$ that varies slowly (we will define what means slowly) and we ask what does not change to the first order in derivatives of $\lambda(t)$.

If we have a bounded motion with period T that changes slowly under the change of $\lambda(t)$ we can define the change as slow if

$$T\frac{\mathrm{d}\lambda}{\mathrm{d}t} \ll \lambda \tag{8.44}$$

over a period T. Since the parameters change with time the energy is not conserved (but very little). We can write the hamiltonian as $H(q,p;\lambda)$. Then

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{\partial H}{\partial \lambda}\frac{\mathrm{d}\lambda}{\mathrm{d}t} \tag{8.45}$$

Averaging over one period we can take the $\dot{\lambda}$ outside of the averaging and we get

$$\overline{\frac{\mathrm{d}E}{\mathrm{d}t}} = \frac{\mathrm{d}\lambda}{\mathrm{d}t}\overline{\frac{\partial H}{\partial \lambda}} \tag{8.46}$$

so that we can write

$$\overline{\frac{\mathrm{d}E}{\mathrm{d}t}} = \frac{\mathrm{d}\lambda}{\mathrm{d}t}\frac{1}{T}\int_0^T \frac{\partial H}{\partial \lambda}\mathrm{d}t \tag{8.47}$$

Using Hamilton's eqs we get

$$\mathrm{d}t = \frac{\mathrm{d}q}{\frac{\partial H}{\partial p}} \tag{8.48}$$

so that

$$\overline{\frac{\mathrm{d}E}{\mathrm{d}t}} = \frac{\mathrm{d}\lambda}{\mathrm{d}t}\frac{\oint \frac{\frac{\partial H}{\partial \lambda}}{\frac{\partial H}{\partial p}}\mathrm{d}q}{\oint \frac{\mathrm{d}q}{\frac{\partial H}{\partial p}}} \tag{8.49}$$

Now we know from the triple product formula that

$$\frac{\frac{\partial H}{\partial \lambda}}{\frac{\partial H}{\partial p}} = -\frac{\partial p}{\partial \lambda} \tag{8.50}$$

We therefore get

$$\oint\left(\frac{\partial p}{\partial E}\overline{\frac{\mathrm{d}E}{\mathrm{d}t}} + \frac{\partial p}{\partial \lambda}\frac{\mathrm{d}\lambda}{\mathrm{d}t}\right)\mathrm{d}q = 0 \tag{8.51}$$

Introducing

$$I = \frac{1}{2\pi}\oint\sum_a p_a \mathrm{d}q_a \tag{8.52}$$

we finally get

$$\overline{\frac{\mathrm{d}I}{\mathrm{d}t}} = 0 \tag{8.53}$$

We also notice that

$$\frac{\partial I}{\partial E} = \frac{T}{2\pi} \tag{8.54}$$

We can also write

$$I = \frac{1}{2\pi}\int \mathrm{d}p \wedge \mathrm{d}q \tag{8.55}$$

For example for the oscillator

$$H = \frac{p^2}{2m} + \frac{m\omega^2 q^2}{2} \tag{8.56}$$

so that for a fixed energy E we have an ellipse with semiaxis $\sqrt{2mE}$ and $\sqrt{2E/m\omega^2}$ so that the area πab divided by 2π is equal to

$$I = \frac{E}{\omega} \tag{8.57}$$

Einstein noticed it during Solvay conference 1911 that later led to the Bohr-Sommerfeld quantization rule. It is related to adiabatic invariants and the same concerns the equation

$$E = n\hbar\omega \tag{8.58}$$

We now apply this to a very slowly varying length of a pendulum. We have

$$E(t) = I\omega(t) \tag{8.59}$$

where I is an adiabatic invariant. On the other hand

$$\bar{E} = \frac{ml^2}{2}\bar{\dot{\theta}^2} + \frac{mgl}{2}\bar{\theta^2} \tag{8.60}$$

where we have taken slowly varying l out of the averaging sign. Averaging gives

$$\bar{\dot{\theta}^2} = \frac{\omega^2\theta_0^2}{2}, \qquad \bar{\theta^2} = \frac{\theta_0^2}{2} \tag{8.61}$$

so we have

$$\bar{E} = \frac{mg}{2} l\theta_0^2 \tag{8.62}$$

Dividing by $\omega = \sqrt{g/l}$ we get an adiabatic invariant i.e. constant

$$l^{3/2}\theta_0^2 = \mathrm{const} \Rightarrow l^{3/4}\theta_0 = \mathrm{const} \tag{8.63}$$

Therefore when we slowly make the pendulum longer the angle θ_0 decreases and the linear amplitude $l\theta_0$ increases.

9 Hamiltonian formalism II

9.1 Liouville theorem

Imagine the flow of (q_a, p_a) i.e. a tube of close trajectories (in the phase space). Its volume is

$$V = \mathrm{d}q_1 \dots \mathrm{d}q_n \mathrm{d}p_1 \dots \mathrm{d}p_n \tag{9.1}$$

We ask what will be this infinitesimal volume after time $\mathrm{d}t$. Then

$$q_a \to \tilde{q}_a = q_a + \frac{\partial H}{\partial p_a}\mathrm{d}t, \qquad p_a \to \tilde{p}_a = p_a - \frac{\partial H}{\partial q_a}\mathrm{d}t, \tag{9.2}$$

The jacobian from V to $\tilde{V}$ reads

$$J = \begin{pmatrix} \frac{\partial \tilde{q}_a}{\partial q_b} & \frac{\partial \tilde{q}_a}{\partial p_b} \\ \frac{\partial \tilde{p}_a}{\partial q_b} & \frac{\partial \tilde{p}_a}{\partial p_b} \end{pmatrix} \tag{9.3}$$

We now use the formula

$$\exp(\mathrm{Tr} \ln M) = \det M \tag{9.4}$$

for an arbitrary matrix M with positive eigenvalues. It can be proven using the fact that any matrix can be brought to the diagonal (or Jordan) form by some (complex) matrix A. Indeed, writing

$$M = \mathbf{1} + \delta \tag{9.5}$$

we have (M' is in the diagonal or Jordan form)

$$M' = AMA^{-1} \Rightarrow \mathrm{Tr} \ln M = \mathrm{Tr}\left(\delta + \frac{1}{2}\delta^2 + \dots\right) = \mathrm{Tr}\left(\delta' + \frac{1}{2}\delta'^2 + \dots\right) = \sum \ln \lambda_i \tag{9.6}$$

and we see that both sides of the equation (9.4) are equal to the product of the eigenvalues. In our case

$$M = \mathbf{1} + \delta \qquad \Rightarrow \qquad \det M = 1 + \mathrm{Tr}\delta + O(\delta^2) \tag{9.7}$$

but

$$\mathrm{Tr}\delta = \sum_a \left(\frac{\partial^2 H}{\partial q_a \partial p_a} - \frac{\partial^2 H}{\partial p_a \partial q_a}\right)\mathrm{d}t = 0 \tag{9.8}$$

so that

$$V = \tilde{V} \tag{9.9}$$

It says that 'squeezing' the trajectories in q requires 'expanding' them in p – it resembles quantum uncertainty relation but it is very different being purely classical.

9.2 Poincaré recurrence theorem

We now prove one of the most striking theorems in classical mechanics.

We assume that the phase space is of finite phase volume (for example of finite energy and in finite spatial volume). We consider finite time steps $0, T, 2T, \ldots$. The theorem says that for any point P_0 and for any neighborhood D_0 of P_0 in the phase space there exists such n that

$$D_n \cap D_0 \neq \emptyset \tag{9.10}$$

where D_n is D_0 transformed by H after time nT.

The proof consists in showing that since for all n regions D_n have the same volume then there must exist such n' and n'' (different from each other) for which

$$D_{n'} \cap D_{n''} \neq \emptyset \tag{9.11}$$

since otherwise the volume of the phase space would be infinite. Taking for example $n' < n''$ then acting with H backwards n times (action of the hamiltonian is reversible) we get

$$D_0 \cap D_{n''-n'} \neq \emptyset \tag{9.12}$$

what finishes the proof.

9.3 Liouville's equation

For a system of N bodies we can introduce a density (probability) on the phase space $\rho(q,p)$ such that

$$\int \rho(q,p) \mathrm{d}q_1 \ldots \mathrm{d}q_n \mathrm{d}p_1 \ldots \mathrm{d}p_n = N \tag{9.13}$$

Since the volume of the phase space is constant we get

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{\partial \rho}{\partial t} + \sum_a \left(\frac{\partial \rho}{\partial q_a} \dot{q}_a + \frac{\partial \rho}{\partial p_a} \dot{p}_a \right) = 0 \tag{9.14}$$

what gives the Liouville equation

$$\frac{\partial \rho}{\partial t} = -\sum_a \left(\frac{\partial \rho}{\partial q_a} \frac{\partial H}{\partial p_a} - \frac{\partial \rho}{\partial p_a} \frac{\partial H}{\partial q_a} \right) = -\{\rho, H\}_{PB} \tag{9.15}$$

We have introduced here the notion of a Poisson Bracket defined as

$$\{f, g\}_{PB} := \sum_a \frac{\partial f}{\partial q_a} \frac{\partial g}{\partial p_a} - \frac{\partial f}{\partial p_a} \frac{\partial g}{\partial q_a} \tag{9.16}$$

We will discuss its role in great detail later.

An important role is played by time independent (equilibrium or stationary) densities for which $\frac{\partial \rho}{\partial t} = 0$. An example of such stationary distributions is given by

$$\rho = \rho(H(q,p)) \tag{9.17}$$

where H does not depend on time. Then indeed

$$\frac{\partial \rho}{\partial t} = \frac{\partial \rho}{\partial H}\left(-\frac{\partial H}{\partial q_a}\frac{\partial H}{\partial p_a} + \frac{\partial H}{\partial p_a}\frac{\partial H}{\partial q_a}\right) = 0 \tag{9.18}$$

The most famous example of such a distribution is the Boltzmann factor in the canonical ensemble

$$\rho(H(q,p)) = \exp\left(-\frac{H(q,p)}{kT}\right) \tag{9.19}$$

which in classical statistical physics for free particles is proven to describe the equilibrium distribution for a small sytem in contact with a large reservoir of temperature T (if H. If $H(q,p)$ describes free particles, $H = \sum p^2/(2m)$, the distribution is called Maxwell-Boltzmann distribution.

It is interesting to note that in the case of a magnetic field described by the vector potential $\mathbf{A}$ the Boltzmann factor gives

$$\exp\left(-\frac{(\mathbf{p}-q\mathbf{A})^2}{2mkT}\right) = \exp\left(-\frac{m\dot{\mathbf{r}}^2}{2kT}\right) \tag{9.20}$$

and it is the same distribution in velocities with or without the magnetic field! This is the paradox that in classical physics bodies should not react to a magnetic field while obviously such a reaction exists - this is solved in quantum mechanics where there are quantized levels (Landau levels) and quantized spin degrees of freedom and the classical Boltzmann factor does not describe the real reaction of the bodies to the magnetic field.

9.4 Classical statistical physics

In classical statistical physics we are interested in the classical partition function for N particles. One distinguishes different ensembles: microcanonical, canonical and macrocanonical.

9.4.1 Microcanonical ensemble

The microcanonical ensemble is described by the number of particles N, volume of the phase space (assuming that it is finite) and finite energy U (within a small interval ΔU). The number of 'states' in classical physics is formally infinite so to make it well defined we need to appeal to quantum physics where there is a heuristic rule that a new state is possible when $\Delta q \Delta p$ differs by h (the Planck constant). Using this heuristic rule we calculate the number of states in an interval ΔU around

$$Z_N(V,U,\Delta U) = \Delta U \int \frac{\mathrm{d}^{3N}p\,\mathrm{d}^{3N}q}{N!h^{3N}}\,\delta(U - T_N - V_N) \tag{9.21}$$

$1/N!$ is the Gibbs factor, yet another factor that can be justified only in quantum physics (indistinguishability of identical particles). $Z_N(V,U)$ is then the number of states around U in the interval ΔU.

According to the famous Boltzmann formula logarithm of the number of states, i.e. logarithm of Z_N, is equal to the entropy (modulo a constant)

$$S = k \log W \tag{9.22}$$

This formula is on Boltzmann's grave in Vienna – it required an incredible ingenuity of Boltzmann to write it down in 1875, 25 years before the Planck's assumption of quantization of photon emissions and absorptions.

It can be justified by the formula (also given by Boltzmann in 1866)

$$S = -\sum P \ln P \tag{9.23}$$

and using equal (maximal) probability $P = 1/Z_N$ for all states ($\sum P = 1$).

In the following we put the Boltzmann's constant k equal to 1 (it can always be reinstated if need arises). Knowing $S(U, V, N)$ we can recover all thermodynamical functions in this ensemble by (we keep N fixed)

$$\mathrm{d}S = \frac{1}{T}\mathrm{d}U + \frac{p}{T}\mathrm{d}V \tag{9.24}$$

i.e.

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_V, \quad \frac{p}{T} = \left(\frac{\partial S}{\partial V}\right)_U \tag{9.25}$$

They are definitions of $1/T$ and p/T. If we have two subsystems with

$$\mathrm{d}S_1 = \alpha_1 \mathrm{d}U_1, \quad \mathrm{d}S_2 = \alpha_2 \mathrm{d}U_2, \tag{9.26}$$

then in equilibrium the total system should have maximal entropy under a change of U_1 and U_2:

$$\mathrm{d}(S_1 + S_2) = 0 \quad \text{when} \;\; \mathrm{d}U_1 = -\mathrm{d}U_2 \tag{9.27}$$

what gives $\alpha_1 = \alpha_2$ and we identify it with the inverse temperature.

As an example let us discuss free non-relativistic particles. Then

$$Z_N(V, U, \Delta U) = \frac{\Delta U V^N \Omega^{(3N-1)}}{N! h^{3N}} \int \mathrm{d}p\, p^{3N-1} \delta\left(U - \frac{p^2}{2m}\right) \tag{9.28}$$

where $\Omega^{(3N-1)}$ is the volume of $3N-1$-dimensional unit sphere

$$\Omega^{(3N-1)} = 2\frac{\pi^{3N/2}}{\Gamma(3N/2)} \tag{9.29}$$

and $\delta(f(x)) = \sum \delta(x - x_i)/|f'(x_i)|$ where x_i are zeroes of $f(x)$. Hence

$$Z_N(V, U, \Delta U) = \frac{V^N \Omega^{(3N-1)}}{N! h^{3N}} (2mU)^{3N/2} \tag{9.30}$$

and (using $\ln(N!) = N\ln(N) - N + \frac{1}{2}\ln(2\pi N) + O(1/N)$)

$$S = NC + N\ln(V/N) + \frac{3N}{2}\ln(U/N) + \text{const} \tag{9.31}$$

We see that without the $N!$ factor in the denominator S would not be proportional to N but there would be logarithmic corrections to S/N growing like $\ln N$. The result is the so called Sackur-Tetrode equation.

Hence

$$\frac{1}{T} = \frac{3N}{2U}, \qquad \frac{p}{T} = \frac{N}{V} \tag{9.32}$$

9.4.2 Canonical ensemble

In the canonical ensemble we do not assume that the energy is constant but that the system is in contact with a very large system of temperature T. The large system has a number of states $\exp(S(E_0))$ and if we extract energy E to the small system the number of states is equal to

$$\mathrm{e}^{S(E_0-E)} \sim \mathrm{e}^{S(E_0)-\frac{\partial S}{\partial E_0}E+\dots} \sim \mathrm{e}^{S(E_0)-E/T} \tag{9.33}$$

where we applied the definition of the temperature to the large system. Therefore we see that a probability of a given state of energy E of the small system is given by

$$P = \mathrm{e}^{\beta(F-E)} \tag{9.34}$$

the famous Gibbs-Boltzmann factor, where $\beta = 1/T$ and F is a normalizing factor. Sum of probabilities must be equal to 1 so

$$\mathrm{e}^{-\beta F(V,T)} = \int \frac{\mathrm{d}^{3N}p\,\mathrm{d}^{3N}q}{N!h^{3N}}\mathrm{e}^{-\beta H(p,q)} \tag{9.35}$$

where $1/N!$ is again the Gibbs factor. We know that the entropy S is given by

$$S = -\sum P\ln P = -\sum \beta(F-E)\mathrm{e}^{\beta(F-E)} = -\beta F + \beta U \tag{9.36}$$

Hence

$$F = U - TS \tag{9.37}$$

and it can be identified with the free energy.

Therefore

$$\begin{aligned} p &= -\left(\frac{\partial F}{\partial V}\right)_T \\ S &= -\left(\frac{\partial F}{\partial T}\right)_V \end{aligned} \tag{9.38}$$

As a first example we consider again free non-relativistic particles. Then

$$\mathrm{e}^{-\beta F(V,T)} = \int \frac{\mathrm{d}^{3N}p\mathrm{d}^{3N}q}{N!h^{3N}} \mathrm{e}^{-\beta p^2/(2m)} \tag{9.39}$$

We have

$$\mathrm{e}^{-\beta F(V,T)} = \frac{V^N \Omega^{(3N-1)}}{N!h^{3N}} \int p^{3N-1} \mathrm{e}^{-\beta p^2/(2m)} \mathrm{d}p \tag{9.40}$$

The integral is straightforward and we get

$$\mathrm{e}^{-\beta F(V,T)} = \frac{V^N \Omega^{(3N-1)}}{N!h^{3N}} (2m)^{3N/2} \frac{\Gamma(3N/2)}{\beta^{3N/2}} \tag{9.41}$$

Hence

$$-\beta F = NC' + N \ln(V/N) - \frac{3N}{2} \ln(\beta) \tag{9.42}$$

and we recover the known formulae.

As a second example we discuss a gas of photons. If the are closed in a box then the force on a wall is given by

$$F = \frac{2h\nu v_z/c^2}{2L/v_z} = \frac{h\nu}{L}\frac{v_z}{c^2} = \frac{1}{3}\frac{U}{L} \tag{9.43}$$

Hence

$$pV = \frac{1}{3}U \Rightarrow p = \frac{1}{3}\rho \tag{9.44}$$

Assuming that nothing depends on the number of photons we substitute

$$S = \alpha T^m V, \quad p = \beta T^N, \quad \Rightarrow U = 3\beta T^n V \tag{9.45}$$

Using

$$\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V \tag{9.46}$$

we get

$$3\beta n T^{n-1} V \mathrm{d}T + 3\beta T^n \mathrm{d}V = \alpha m T^m V \mathrm{d}T + \alpha T^{m+1} \mathrm{d}V - \beta T^n \mathrm{d}V \tag{9.47}$$

Comparing the expressions we get

$$\alpha = 4\beta, \quad m = 3, \quad n = 4 \tag{9.48}$$

so that

$$U = 3\beta T^4 V, \ \ p = \beta T^4, \quad S = 4\beta T^3 V \tag{9.49}$$

It turns out (from the Planck black body distribution) that

$$\beta = \frac{\pi^2 k^4}{45 c^3 \hbar^3} \Rightarrow U = \frac{\pi^2 k^4}{15 c^3 \hbar^3} T^4 V, \ \ S = \frac{4\pi^2 k^3}{45 c^3 \hbar^3} T^3 V \tag{9.50}$$

the number of photons is given by

$$N = \frac{2\zeta(3) k^2}{\pi^2 c^3 \hbar^3} T^3 V = \frac{45\zeta(3)}{2\pi^4} S \tag{9.51}$$

9.4.3 Grand canonical ensemble

In the classical setting the we assume that the system is immersed in a bath of temperature T but can also exchange particles with the reservoir. We denote the energy associated to the exchange of one particle by μ and call it the chemical potential μ. We assign a probability for N particles having energy E_N as

$$p_N(E_N) = \mathrm{e}^{\beta(\mu N - E_N)} \tag{9.52}$$

and the normalizing factor Ω defined as

$$\sum \frac{1}{N!}\mathrm{e}^{\beta(\Omega+\mu N - E_N)} = 1 \tag{9.53}$$

We define entropy as

$$S = -\sum p_i \ln p_i = -\sum \frac{1}{N!}\beta(\Omega+\mu N - E_N)\mathrm{e}^{\beta(\Omega+\mu N - E_N)} = -\beta\Omega - \beta\mu < N > +\beta < E > \tag{9.54}$$

and hence

$$\Omega = U - TS + \mu N = -pV \tag{9.55}$$

(we know from thermodynamics that $U - TS + pV + \mu N = 0$) Therefore the fundamental object in the grand canonical ensemble is

$$\mathrm{e}^{-\beta\Omega} = \sum_{N=0}^{\infty} \frac{1}{N!}\mathrm{e}^{\mu N} \int \mathrm{d}\Gamma_N \mathrm{e}^{-\beta E_N} \tag{9.56}$$

Using the grand canonical ensemble in quantum statistical physics one can derive the distribution of the number of particles at a given energy level E_j:

$$n_j = \frac{1}{\mathrm{e}^{E_j - \mu} \pm 1} \tag{9.57}$$

where $-$ is for bosons (Bose-Einstein distribution) and $+$ for fermions (Fermi-Dirac distribution).

The discussion of all three ensembles (microcanonical, canonical and grand canonical) belongs to the course on Statistical Physics and is outside of the scope of lectures on Classical Mechanics where it serves only as an illustration of the Liouville equation.

9.5 Debye theory of specific heat of solids

We will apply the derived distributions for phonons to derive the formula for the specific heat of solids (not including the electronic heat capacity).

9.5.1 classical computation

If we have a 1D string of atoms with harmonic potential and equilibrium distance a we have

$$m\ddot{x}_n = K(x_{n+1} - x_n) + K(x_{n-1} - x_n) \tag{9.58}$$

To solve this equation we substitute

$$x_n = na + \mathrm{e}^{\mathrm{i}\omega t}\sin(kna), \qquad -\frac{\pi}{a} \le k \le \frac{\pi}{a} \tag{9.59}$$

to get

$$\omega^2 = \frac{4K}{m}\sin^2(ka/2) \tag{9.60}$$

We write

$$\omega = \omega_m \sin(ka/2), \qquad \omega_m = \sqrt{\frac{4K}{m}} \tag{9.61}$$

The group velocity

$$v_g = \frac{\partial\omega}{\partial k} = \frac{\omega_m a}{2}\cos(ka/2) = \frac{Ka^2}{m}\cos(ka/2) = v_{g0}\cos(ka/2) \tag{9.62}$$

Density of states in 3D ($k = \frac{2}{a}\arcsin(\omega/\omega/m)$; there are 2 transverse and 1 longitudinal polarizations)

$$g(\omega)\mathrm{d}\omega = \frac{3V4\pi k^2\mathrm{d}k}{(2\pi)^3} = \frac{V\omega^2\mathrm{d}\omega}{2\pi^2 v_{g0}^3}\cdot\frac{1+\omega^2/(3\omega_m^2)+\ldots}{\sqrt{1-\omega^2/\omega_m^2}} \tag{9.63}$$

Einstein has used the formula for density for one specific frequency $g(\omega) = 3N\delta(\omega - \omega_E)$ – it explained the Dulong-Petit law that the heat capacity tends to $3R$ for large temperatures but was not very good in explaining low temperature behavior of heat capacity.

Debye assumed that all frequencies are present and wrote the formula (without any corrections $\sim \omega^2/\omega_m^2$) to use measured v_{g0}) and defined ω_D by

$$3N = \int_0^{\omega_D} g(\omega)\mathrm{d}\omega = \int_0^{\omega_D}\frac{3V\omega^2\mathrm{d}\omega}{2\pi^2 v_{g0}^3} = \frac{V\omega_D^3}{2\pi^2 v_{g0}^3} \tag{9.64}$$

hence

$$\omega_D = v_{g0}\left(6\rho\pi^2\right)^{\frac{1}{3}}, \qquad \rho = N/V \tag{9.65}$$

and

$$g(\omega)\mathrm{d}\omega = \frac{9N\omega^2\mathrm{d}\omega}{\omega_D^3} \tag{9.66}$$

If transverse and longitudinal speeds are different one may use the averaging

$$\frac{3}{\bar{v}^3} = \frac{2}{v_t^3} + \frac{1}{v_l^3} \tag{9.67}$$

Phonons are bosons so that the energy stored in phonons in temperature T is given by

$$E = \int_0^{\omega_D} \mathrm{d}\omega g(\omega)\hbar\omega \frac{1}{\mathrm{e}^{\frac{\hbar\omega}{kT}} - 1} \tag{9.68}$$

The heat capacity

$$c_p = \frac{\partial E}{\partial T} = \int_0^{\omega_D} \mathrm{d}\omega \frac{9N\hbar^2\omega^4}{\omega_D^3 kT^2} \frac{\mathrm{e}^{\frac{\hbar\omega}{kT}}}{\left(\mathrm{e}^{\frac{\hbar\omega}{kT}} - 1\right)^2} \tag{9.69}$$

It can be rewritten as

$$c_p = 9Nk\left(\frac{T}{\theta_D}\right)^3 \int_0^{\theta_D/T} \mathrm{d}x \frac{x^4 \mathrm{e}^x}{(\mathrm{e}^x - 1)^2} \tag{9.70}$$

where

$$\theta_D = \frac{\hbar\omega_D}{k} \tag{9.71}$$

For $T \ll \theta_D$ we have

$$c_p \to 9Nk\left(\frac{T}{\theta_D}\right)^3 \cdot \frac{4\pi^4}{15} \tag{9.72}$$

while for $T \gg \theta_D$ we recover the Dulong-Petit law $c_p \to 3R$.

This formula is in much better agreement with experimentally measured values than Einstein's but is not exact either. To have better description one has to take into account the presence of (quantum) characteristic frequencies of a given crystal or dependence of θ_D on temperature. The Debye temperatures of some of the elements (they decrease with the temperature to match the experimental values!): aluminum 433 K, beryllium 1481 K, copper 347 K, lead 105 K, gold 227 K, diamond 2200 K (in room temperature 1840 K).

10 Canonical transformations

We discuss here a very important formulation of classical mechanics that led Paul Dirac to formulate quantum mechanics in a very analogous way.

10.1 Poisson brackets

For any two functions on the phase space $f(q,p)$ and $g(q,p)$ we define a Poisson bracket as

$$\{f,g\}_P := \sum_a \frac{\partial f}{\partial q_a}\frac{\partial g}{\partial p_a} - \frac{\partial g}{\partial q_a}\frac{\partial f}{\partial p_a} \tag{10.1}$$

Poisson brackets have features that are analogous to commutators in the operator language of QM.

- antisymmetry

$$\{f,g\} = -\{g,f\} \tag{10.2}$$

- linearity

$$\{\alpha f + \beta f', g\} = \alpha\{f,g\} + \beta\{f',g\} \tag{10.3}$$

- Jacobi identity

$$\{\{f,g\},h\} + \{\{h,f\},g\} + \{\{g,h\},f\} = 0 \tag{10.4}$$

It is important to note that in analogy to QM we have

$$\{q_a,q_b\} = 0, \quad \{p_a,p_b\} = 0, \quad \{q_a,p_b\} = \delta_{ab}, \tag{10.5}$$

Its introduction is motivated by the appearance in the Hamilton's equations

$$\dot{q}_a = \{q_a,H\}_P, \qquad \dot{p}_a = \{p_a,H\}_P \tag{10.6}$$

Therefore also for any function $f(t,q,p)$ we have

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \{f,H\}_P \tag{10.7}$$

where we used the Hamilton's eqs.

If we have a function $I(q,p)$ that commutes with H

$$\{I,H\} = 0 \quad \Rightarrow \quad I = \mathrm{const} \tag{10.8}$$

i.e. it is a constant of motion. One example is a momentum p_c conjugate to a cyclic coordinate $q_c$since then

$$\{p_c, H\} = 0 \;\Rightarrow\; p_c = \text{const} \tag{10.9}$$

as expected.

The PB of two constants of motion is again a constant of motion.

$$\{\{I_1, I_2\}, H\} = -\{\{H, I_1\}, I_2\} - \{\{I_2, H\}, I_1\} = 0 \tag{10.10}$$

10.2 Canonical transformations

We noticed earlier that in the lagrangian formulation one can arbitrarily change coordinates $q \to q'$ and the EL eqs were invariant wrt this change. We now discuss what possible transformations can be applied to the pair (q_a, p_a) that lead again to the Hamilton equations (the previous $q \to q'$ constitute a small subset of these).

The canonical transformations are defined as a pair

$$q_a \to Q_a(t, q, p), \qquad p_a \to P_a(t, q, p) \tag{10.11}$$

that has canonical Poisson brackets i.e.

$$\{Q_a, Q_b\} = 0, \quad \{P_a, P_b\} = 0, \quad \{Q_a, P_b\} = \delta_{ab}, \tag{10.12}$$

If they are satisfied then the Hamilton's equations have the usual form (with possibly some new Hamilton's function).

To prove it let us introduce the symplectic structure that is present in the Hamilton's formulation. We introduce a vector

$$\mathbf{x}^T = (q_1, \ldots, q_n, p_1, \ldots, p_n) \tag{10.13}$$

that has $2n$ components. We also introduce a $2n \times 2n$ matrix J

$$J = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix} \tag{10.14}$$

It is crucial that $J^2 = -\mathbf{1}$. Then we can rewrite the Poisson brackets as

$$\{f, g\} = \frac{\partial f}{\partial x^i} J^{ij} \frac{\partial g}{\partial x^j} \tag{10.15}$$

and the Hamilton's eqs as

$$\dot{x}^i = J^{ij} \frac{\partial H}{\partial x^j} \tag{10.16}$$

Let us now transform $\mathbf{x} \to \mathbf{y}(\mathbf{x})$ where we temporarily assume that the transformation does not depend explicitly on time. Then the eqs of motion for $\mathbf{y}$ read

$$\dot{y}^i = \frac{\partial y^i}{\partial x^j} J^{jk} \frac{\partial H}{\partial y^l} \frac{\partial y^l}{\partial x^k} = (GJG^T)^{ij} \frac{\partial H}{\partial y^j} \tag{10.17}$$

Therefore we recover the Hamilton's equations in the new coordinates if

$$GJG^T = J \tag{10.18}$$

where H is the same but expressed in new coordinates.

Matrices satisfying such a condition belong to the symplectic group $Sp(n)$ - it is easy to show that they form a group of dimension $n(2n+1)$ – for example we show that the inverse element belongs to the group by the following argument (analogous to the uniqueness of the inverse matrix)

$$\begin{aligned} &G^{-1T}JG^{-1} = J \Rightarrow JG^{-1T}JG^{-1} = -\mathbf{1} \Rightarrow JG^{-1T} = -(JG^{-1})^{-1} \Rightarrow \\ &\Rightarrow JG^{-1}JG^{-1T} = -\mathbf{1} \Rightarrow G^{-1}JG^{-1T} = J \end{aligned} \tag{10.19}$$

One should specify over which field one defines the group and whether we allow for only component connected to the identity. In 2 dimensions any matrix of determinant 1 belongs to $Sp(1,\mathbb{R})$ so $Sp(1,\mathbb{R}) \equiv SL(2,\mathbb{R})$, in 4 dimensions the algebra $sp(2) \equiv so(5)$.

For a 2×2 matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{10.20}$$

gives $ad-bc = 1$ i.e. indeed the determinant should be equal to 1. We can impose further that both eigenvalues should be positive i.e. allow only for transformations connected to the identity.

The Poisson brackets for the new coordinates

$$\{f,g\} = \frac{\partial f}{\partial x^i} J^{ij} \frac{\partial g}{\partial x^j} = \frac{\partial f}{\partial y^i} (GJG^T)^{ij} \frac{\partial g}{\partial y^j} \tag{10.21}$$

so the requirement of conserving the PBs gives the same condition.

10.3 Examples of canonical transformations

- exchanging positions and momenta

$$P_a = -q_a, \quad Q_a = p_a \quad \Rightarrow G = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix} \tag{10.22}$$

  it is obviously canonical

- 'point transformations'

$$Q_a = Q_a(q) \tag{10.23}$$

  Then

$$G = \begin{pmatrix} \frac{\partial Q_a}{\partial q_b} & 0 \\ \frac{\partial P_a}{\partial q_b} & \frac{\partial P_a}{\partial p_b} \end{pmatrix} \tag{10.24}$$

The matrix

$$G = \begin{pmatrix} A & 0 \\ B & C \end{pmatrix} \tag{10.25}$$

belongs to the symplectic group if

$$C^T = A^{-1}, \quad BA^{-1} = (BA^{-1})^T \tag{10.26}$$

- linear transformation

$$q_a \to \theta^{ab} q_b, \quad p_a \to \theta^{-1}_{ba} p_a \tag{10.27}$$

with θ constant.

10.4 Harmonic oscillators on a line

We write the hamiltonian for n atoms (we assume that n is odd) bound by harmonic forces on a circle

$$H = \sum \frac{p_i^2}{2m} + \frac{m\omega_0^2}{2} \sum_i (x_{i+1} - x_i - d)^2 \tag{10.28}$$

We now introduce the canonical transformation $(x, p) \to (q, P)$

$$x_k = kd + \frac{1}{\sqrt{n}} \sum_a \mathrm{e}^{2\pi \mathrm{i} ak/n} q_a \tag{10.29}$$

$$p_k = \frac{1}{\sqrt{n}} \sum_a \mathrm{e}^{-2\pi \mathrm{i} ak/n} P_a \tag{10.30}$$

with the inverse transformation

$$q_a = \frac{1}{\sqrt{n}} \sum_a \mathrm{e}^{-2\pi \mathrm{i} ak/n} (x_k - kd) \tag{10.31}$$

$$P_a = \frac{1}{\sqrt{n}} \sum_a \mathrm{e}^{2\pi \mathrm{i} ak/n} p_k \tag{10.32}$$

SInce x_k are real we have

$$q_0^\dagger = q_0, \quad P_0^\dagger = P_0, \quad q_a^\dagger = q_{n-a}, \quad P_a^\dagger = P_{n-a} \tag{10.33}$$

We check

$$\{q_a, P_b\} = \frac{1}{n} \sum_{j,l} \mathrm{e}^{-2\pi \mathrm{i} aj/n} \mathrm{e}^{2\pi \mathrm{i} bl/n} \{x_j, p_l\} = \delta(a-b) \tag{10.34}$$

Then

$$H = \frac{1}{2m} \sum_a P_a P_a^\dagger + \frac{m\omega^2}{2n} \sum_j \sum_a (1 - \mathrm{e}^{2\pi \mathrm{i} a/n}) \mathrm{e}^{2\pi \mathrm{i} aj/n} q_a \sum_b (1 - \mathrm{e}^{-2\pi \mathrm{i} b/n}) \mathrm{e}^{-2\pi \mathrm{i} bj/n} q_b^\dagger \tag{10.35}$$

Therefore we have a diagonalized hamiltonian

$$H = \frac{1}{2m}\sum_a P_a P_a^\dagger + \frac{m\omega_0^2}{2}\sum_a 4\sin^2(\pi a/n) q_a q_a^\dagger \tag{10.36}$$

that can be written in terms of independent variables as

$$H = \frac{1}{2m}P_0^2 + \sum_{a=1}^{(n-1)/2}\left(\frac{1}{m}P_a P_a^\dagger + m\omega_0^2 4\sin^2(\pi a/n) q_a q_a^\dagger\right) \tag{10.37}$$

where P_0 corresponds to global translations and can be discarded. Dividing into real and imaginary parts

$$\tilde{q}_a = \frac{1}{\sqrt{2}}\left(q_a + q_a^\dagger\right), \quad \tilde{p}_a = \frac{1}{\sqrt{2}}\left(P_a + P_a^\dagger\right) \tag{10.38}$$

$$\tilde{r}_a = \frac{\mathrm{i}}{\sqrt{2}}\left(q_a - q_a^\dagger\right), \quad \tilde{s}_a = \frac{\mathrm{i}}{\sqrt{2}}\left(P_a - P_a^\dagger\right) \tag{10.39}$$

we get the reduced hamiltonian

$$H = \sum_{a=1}^{(n-1)/2}\left(\frac{1}{2m}(\tilde{p}_a\tilde{p}_a + \tilde{s}_a\tilde{s}_a) + \frac{m\omega_0^2}{2}4\sin^2(\pi a/n)(\tilde{q}_a\tilde{q}_a + \tilde{r}_a\tilde{r}_a\right) \tag{10.40}$$

10.5 Some identities for partial derivatives

Assume that we have 3-dim manifold with a hypersurface defined by $f(x,y,z) = 0$ and we would like to derive some identities between the partial derivatives wrt to different pairs of variables (since only 2 are independent) – they are extensively used in thermodynamics.

We start with

$$\begin{aligned} \mathrm{d}x &= \left(\frac{\partial x}{\partial y}\right)_z \mathrm{d}y + \left(\frac{\partial x}{\partial z}\right)_y \mathrm{d}z \\ \mathrm{d}y &= \left(\frac{\partial y}{\partial x}\right)_z \mathrm{d}x + \left(\frac{\partial y}{\partial z}\right)_x \mathrm{d}z \end{aligned} \tag{10.41}$$

Plugging dy from the second equation into the first we get

$$\left(\frac{\partial x}{\partial y}\right)_z = \frac{1}{\left(\frac{\partial y}{\partial x}\right)_z} \tag{10.42}$$

and the triple product formula

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1 \tag{10.43}$$

Similarly we can write

$$\begin{aligned} dx &= \left(\frac{\partial x}{\partial y}\right)_z dy + \left(\frac{\partial x}{\partial z}\right)_y dz \\ dx &= \left(\frac{\partial x}{\partial y}\right)_w dy + \left(\frac{\partial x}{\partial w}\right)_y dw \end{aligned} \tag{10.44}$$

Writing

$$dw = \left(\frac{\partial w}{\partial y}\right)_z dy + \left(\frac{\partial w}{\partial z}\right)_y dz \tag{10.45}$$

and plugging into the previous equation we get

$$\left(\frac{\partial x}{\partial y}\right)_z = \left(\frac{\partial x}{\partial y}\right)_w + \left(\frac{\partial x}{\partial w}\right)_y \left(\frac{\partial w}{\partial y}\right)_z \tag{10.46}$$

and

$$\left(\frac{\partial x}{\partial z}\right)_y = \left(\frac{\partial x}{\partial w}\right)_y \left(\frac{\partial w}{\partial z}\right)_y \tag{10.47}$$

10.6 Generating functions of canonical transformations

We will now show how to generate the canonical transformations with the use of a generating function. The argument of the function can be any pair $(q, Q), (q, P), (p, Q), (p, P)$ for definiteness we choose q, Q pair.

We choose a function $F(q, Q)$ such that the equation

$$p_a = \frac{\partial F}{\partial q_a} \tag{10.48}$$

is invertible i.e. one can get $Q = Q(q, p)$ out of this equation. Then we define

$$P_a = -\frac{\partial F}{\partial Q_a} \tag{10.49}$$

and we will now show that the pair (Q, P) satisfies the correct Poisson brackets. To avoid proliferation of indices we show it for 1 dof. Then

$$\{Q, P\} = \left(\frac{\partial Q}{\partial q}\right)_p \left(\frac{\partial P}{\partial p}\right)_q - \left(\frac{\partial Q}{\partial p}\right)_q \left(\frac{\partial P}{\partial q}\right)_p \tag{10.50}$$

Now we use the manipulations for the partial derivatives

$$\begin{aligned} \{Q, P\} &= \left(\frac{\partial Q}{\partial q}\right)_p \left(\frac{\partial P}{\partial Q}\right)_q \left(\frac{\partial Q}{\partial p}\right)_q - \left(\frac{\partial Q}{\partial p}\right)_q \left(\left(\frac{\partial P}{\partial q}\right)_Q + \left(\frac{\partial P}{\partial Q}\right)_q \left(\frac{\partial Q}{\partial q}\right)_p\right) \\ &= \left(\frac{\partial Q}{\partial p}\right)_q \left(\frac{\partial^2 F}{\partial q \partial Q}\right) = \left(\frac{\partial Q}{\partial p}\right)_q \left(\frac{\partial p}{\partial Q}\right)_q = 1 \end{aligned} \tag{10.51}$$

The definitions for the 3 remaining pairs $((q, P)$, (p, Q) and $(p, P))$ are analogous.

If the canonical transformation depends explicitly on time we have

$$\dot{P}_a = \{P_a, H\} + \frac{\partial P_a}{\partial t} = \{P_a, H\} - \frac{\partial^2 F}{\partial Q_a \partial t} = -\frac{\partial H}{\partial Q_a} - \frac{\partial^2 F}{\partial Q_a \partial t} \tag{10.52}$$

so to keep the usual Hamilton's eqs we have to modify the hamiltonian

$$H \to H' = H + \frac{\partial F}{\partial t} \tag{10.53}$$

We will use these formulae in the Hamilton-Jacobi equation in the next lecture

11 Hamilton-Jacobi equation

We recall that for the lagrangian we had the principle that the action

$$S = \int_{t_1}^{t_2} \mathrm{d}t L(q_a, \dot{q}_a, t) \tag{11.1}$$

is extremal when the variations δq_a vanish at the ends. We now have a similar principle

$$S = \int_{t_1}^{t_2} \mathrm{d}t (p_a \dot{q}_a - H(q_a, \dot{q}_a, t)) \tag{11.2}$$

where $\dot{q}_a$'s are functions of q_a and p_a. We have

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \left(\delta p_a \dot{q}_a + p_a \delta \dot{q}_a - \left(\frac{\partial H}{\partial p_a} \delta p_a - \frac{\partial H}{\partial q_a} \delta q_a \right) \right) \tag{11.3}$$

Integrating by parts we get

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \left[\left(\dot{q}_a - \frac{\partial H}{\partial p_a} \delta p_a \right) + \left(-\dot{p}_a - \frac{\partial H}{\partial q_a} \delta q_a \right) \right] + p_a \delta q_a |_{t_1}^{t_2} \tag{11.4}$$

If the variations δq_a vanish at the ends we get the Hamilton's equations.

If we impose not only δq_a vanishing at the ends but also δp_a we can add to H a full derivative $\mathrm{d}F(p, q)/\mathrm{d}t$.

11.1 Hamilton-Jacobi equation

We now treat the action S as a function of final time t and final positions $q_a(t)$ assuming initial time t_1 and initial positions $q_a(t_1)$ as fixed. We assume that it is possible to find initial velocities $\dot{q}_a(t_1)$ such that the final positions along the allowed trajectories are $q_a(t)$. Then

$$S(t, q_a(t)) = \int_{t_1}^{t} \mathrm{d}\tau L(q_a(\tau), \dot{q}_a(\tau), \tau) \tag{11.5}$$

Performing the same steps as before we have

$$\delta S = p_a \delta q_a |_{t_1}^{t} \tag{11.6}$$

Hence

$$p_a = \frac{\partial S}{\partial q_a} \tag{11.7}$$

We have

$$\frac{\mathrm{d}S}{\mathrm{d}t} = L \tag{11.8}$$

but on the other hand we have

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{\partial S}{\partial t} + \sum_a \frac{\partial S}{\partial q_a}\dot{q}_a = \frac{\partial S}{\partial t} + \sum_a p_a \dot{q}_a \tag{11.9}$$

The LHS is equal to L and therefore

$$\frac{\partial S}{\partial t} + H\left(q_a, \frac{\partial S}{\partial q_a}, t\right) = 0 \tag{11.10}$$

This is the Hamilton-Jacobi equation. It is the most efficient tool of finding conserved quantities in classical mechanics as we will see.

We can therefore write the differential of S as

$$\mathrm{d}S = -H\mathrm{d}t + \sum p_a \mathrm{d}q_a \tag{11.11}$$

what for one particle is equal to the 4-dimensional expression

$$\mathrm{d}S = p_\mu \mathrm{d}x^\mu \tag{11.12}$$

i.e. the phase differential in the quantum mechanical language. We will discuss solving the mechanical problems by the Hamilton-Jacobi equation later.

11.2 Canonical transformations and the Hamilton-Jacobi equation

We recall that for a function $F(q, Q)$ such that the equation

$$p_a = \frac{\partial F}{\partial q_a} \tag{11.13}$$

is invertible i.e. one can get $Q = Q(q, p)$ out of this equation. Then we define

$$P_a = -\frac{\partial F}{\partial Q_a} \tag{11.14}$$

and we have shown that the pair (Q, P) satisfies the correct Poisson brackets. To keep the usual Hamilton's eqs we have to modify the hamiltonian

$$H \to H' = H + \frac{\partial F}{\partial t} \tag{11.15}$$

It can also be seen from the fact that the two expressions can differ by a full differential

$$-P_a \mathrm{d}Q^a + H'\mathrm{d}t - (-p_a \mathrm{d}q^a + H\mathrm{d}t) = \mathrm{d}F \tag{11.16}$$

We now use this formula to arrive at the HJ equation again to explain the role of constants. If we choose F to be equal to the HJ function S then

$$H + \frac{\partial S}{\partial t} = 0 \Rightarrow H' = 0 \tag{11.17}$$

Therefore we know that Q_a and P_a have to be constant. Then expressing S as a function of positions and constants of integration (identified with Q_a)

$$S = S(t, q_a, \alpha_a) \tag{11.18}$$

we know that also

$$\frac{\partial S}{\partial \alpha_a} = \beta_a \tag{11.19}$$

are constant. Therefore we have the solution given by $2s + 1$ constants as it should be.

11.3 Jacobi (Maupertuis) principle

If the energy E in a given system is conserved we can write

$$S = -Et + \sum_a p_a \mathrm{d}q_a = -Et + S_0 \tag{11.20}$$

We can now formulate the principle of least action in the form

$$\delta S_0 = 0 \tag{11.21}$$

where the variations are along such trajectories that keep the energy E constant. One usually applies this proncipple to find the trajectories and not their dependence on time. Therefore we find $\mathrm{d}t$ as a function of positions q_a and differentials $\mathrm{d}q_a$ and plug it to S_0.

We illustrate the procedure by applying it to the usual lagrangian with generalized kinetic term:

$$L = \frac{1}{2}\sum_{a,b} M_{ab}\dot{q}_a\dot{q}_b - U(q) \tag{11.22}$$

The momenta are given by

$$p_a = \sum_b M_{ab}\dot{q}_b \tag{11.23}$$

and the energy

$$E = \frac{1}{2}\sum_{a,b} M_{ab}\dot{q}_a\dot{q}_b + U(q) \tag{11.24}$$

Hence

$$\mathrm{d}t = \sqrt{\frac{\sum_{a,b} M_{ab}\mathrm{d}q_a\mathrm{d}q_b}{2(E-U)}} \tag{11.25}$$

so that

$$S_0 = \int \sqrt{2(E-U)\sum_{a,b} M_{ab}\mathrm{d}q_a\mathrm{d}q_b} \tag{11.26}$$

For one particle we get

$$\delta \int \sqrt{2m(E-U)}\mathrm{d}l = 0 \tag{11.27}$$

In QM it is analogous to the Fermat principle in optics since

$$p = \frac{h}{\lambda} \qquad \Rightarrow \qquad \delta \int \frac{\mathrm{d}l}{\lambda} = 0 \tag{11.28}$$

i.e. the number of crests or troughs along a trajectory should by extremal.

11.4 Derivation of the Hamilton-Jacobi equation from quantum mechanics

It is very instructive to 'derive' classical mechanics from quantum mechanics (of course the historical path was reverse as is till today the order of teaching...).

We start with the non-relativistic Schrödinger equation for a particle in the scalar potential U

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi + U\psi \tag{11.29}$$

what comes from an operator analogy for the equality

$$E = \frac{p^2}{2m} + U \tag{11.30}$$

when we identify

$$\mathbf{p}\psi = -\mathrm{i}\hbar\nabla\psi \tag{11.31}$$

and

$$\psi = \mathrm{e}^{-\mathrm{i}Et/\hbar}\psi_E \tag{11.32}$$

We now write

$$\psi = R\mathrm{e}^{\mathrm{i}S/\hbar} \tag{11.33}$$

where both R and S are real. The real part of the SE reads

$$-\frac{\partial S}{\partial t} = \frac{1}{2m}(\nabla S)^2 + U - \frac{\hbar^2}{2m}\frac{\Delta R}{R} \tag{11.34}$$

and the momentum

$$\mathbf{p} = \nabla S + O(\hbar) \tag{11.35}$$

and neglecting $\hbar$ corrections we recover the Hamilton-Jacobi equation. The full equation is used in the pilot wave (de Broglie-Bohm) interpretation of quantum mechanics as the classical trajectory with 'quantum potential' (i.e. the last part) added.

The imaginary part reads

$$\frac{1}{R}\frac{\partial R}{\partial t} = -\frac{1}{mR}\nabla R\nabla S - \frac{1}{2m}\Delta S \Rightarrow \frac{\partial R^2}{\partial t} = -\frac{1}{m}\nabla(R^2\nabla S) \tag{11.36}$$

Hence the probability $\int R^2\mathrm{d}^3x$ is conserved in time.

12 Hamilton-Jacobi equation II

12.1 Solving HJ equation

We now discuss the equivalence of HJ equations to the lagrangian or hamiltonian formulations.

$$H + \frac{\partial S}{\partial t} = 0 \Rightarrow H' = 0 \tag{12.1}$$

We know that with a trivial hamiltonian both Q_a and P_a have to be constant. Then expressing S as a function of positions and constants of integration (identified with Q_a)

$$S = S(t, q_a, \alpha_a) \tag{12.2}$$

we know that also

$$\frac{\partial S}{\partial \alpha_a} = \beta_a, \tag{12.3}$$

being momenta P_a, are also constant. Therefore we have the solution given by $2N+1$ constants as it should be. It solves the apparent paradox since S naively depends on only $N+1$ constants being an equation with first order derivatives wrt q_a only with no $\dot{q}_a$s or p_as – the derivatives wrt these constants are also constant supplying the missing set.

Let us discuss the method of separation of variables. Let us assume that q_1 and p_1 appear in the hamiltonian only as a combination $\sigma(q_1, p_1)$ (without any dependence on time). Then we can try

$$S = S'(q_a, t) + S_1(q_1) \tag{12.4}$$

where a denotes all variables except q_1. Then the HJ equation reads

$$\frac{\partial S'}{\partial t} + H\left(t, q_a, \frac{\partial S'}{\partial q_a}, \sigma\left(q_1, \frac{\partial \sigma}{\partial q_1}\right)\right) = 0 \tag{12.5}$$

The solution can be solved only when σ is equal to a constant

$$\sigma\left(q_1, \frac{\partial \sigma}{\partial q_1}\right) = \alpha_1 \tag{12.6}$$

and then we are left with the HJ equation with one smaller number of variables

$$\frac{\partial S'}{\partial t} + H\left(t, q_a, \frac{\partial S'}{\partial q_a}, \alpha_1\right) = 0 \tag{12.7}$$

The obvious example is a cyclical coordinate – then

$$S = S'(q_a, t) + \alpha_1 q_1 \tag{12.8}$$

and the reduced HJ equation reads

$$\frac{\partial S'}{\partial t} + H\left(t, q_a, \frac{\partial S'}{\partial q_a}, \alpha_1\right) = 0 \tag{12.9}$$

If H does not depend on time then we have

$$S = -Et + S_0(q_a) \tag{12.10}$$

and the HJ equation reads

$$H\left(q_a, \frac{\partial S_0}{\partial q_a}\right) = E \tag{12.11}$$

12.2 Hamilton's evolution as a canonical transformation

We will now prove that the Hamilton's evolution is also a canonical transformation.

We consider an infinitesimal transformation (parametrized by α) which by assumption is canonical

$$\begin{aligned} q_a \to Q_a &= q_a + \alpha\sigma_a(q,p) \\ p_a \to P_a &= p_a + \alpha\tau_a(q,p) \end{aligned} \tag{12.12}$$

We require the transformation to be canonical (to first order in α)

$$GJG^T = J \tag{12.13}$$

where

$$G = \begin{pmatrix} \delta_{ab} + \alpha\frac{\partial\sigma_a}{\partial q_b} & \alpha\frac{\partial\sigma_a}{\partial p_b} \\ \alpha\frac{\partial\tau_a}{\partial q_b} & \delta_{ab} + \alpha\frac{\partial\tau_a}{\partial p_b} \end{pmatrix} \tag{12.14}$$

Therefore multiplying

$$\frac{\partial\sigma_a}{\partial q_b} = -\frac{\partial\tau_a}{\partial p_b} \tag{12.15}$$

The solution to this is

$$\sigma_a = \frac{\partial R}{\partial p_a} \qquad \tau_b = -\frac{\partial R}{\partial q_a} \tag{12.16}$$

for some $R(q,p)$ which is called the generator of the transformation.

If α is a short interval of time then we know that $R = H$ – therefore hamiltonian generates time translations

If for example $R = \sum_b \beta_b p_b$ then

$$q_a \to q_a + \alpha\beta_a, \qquad p_a \to p_a \tag{12.17}$$

i.e. momenta generate translations.

Another example is

$$R(q,p)=\sum_{a,b} q_a\theta^{ab}p_b \tag{12.18}$$

Then we recover linear point transformations

$$q_a\to\alpha\theta^{ab}q_a,\qquad p_a\to p_a-\alpha\theta^{ab}p_b \tag{12.19}$$

12.3 Example

We consider 1-dim harmonic oscillator

$$\frac{\partial S}{\partial t}+\frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2+\frac{m\omega^2}{2}x^2=0 \tag{12.20}$$

Since the energy is conserved we write

$$S=-Et+S_0\left(x,\frac{\partial S_0}{\partial x}\right) \tag{12.21}$$

so that

$$\frac{1}{2m}\left(\frac{\partial S_0}{\partial x}\right)^2+\frac{m\omega^2}{2}x^2=E \tag{12.22}$$

Then

$$S_0=\int \mathrm{d}z\sqrt{2mE-m^2\omega^2z^2}=\frac{x}{2}\sqrt{2mE-m^2\omega^2x^2}+\frac{E}{m}\arctan\left(\frac{m\omega x}{\sqrt{2mE-m^2\omega^2x^2}}\right) \tag{12.23}$$

Our constant of integration is E so we differentiate S over E and equate it to a constant

$$-t+\frac{1}{\omega}\arctan\left(\frac{m\omega x}{\sqrt{2mE-m^2\omega^2x^2}}\right)=-t_0 \tag{12.24}$$

what gives us the trajectory of the oscillator

$$x(t)=\sqrt{\frac{2E}{m\omega^2}}\cos(\omega(t-t_0)) \tag{12.25}$$

12.4 Relativistic Hamilton-Jacobi equation

The relativistic analog of the HJ equation in the presence of gravity reads

$$g^{\mu\nu}\frac{\partial S}{\partial x^\mu}\frac{\partial S}{\partial x^\nu}=-m^2c^4 \tag{12.26}$$

We will illustrate this equation by the example of the Schwarzschild metric. We recall the action for a particle in the Schwarzschild metric (for $\theta=\pi/2$)

$$S=-mc^2\int \mathrm{d}t\sqrt{1-\frac{r_s}{r}-\frac{\dot r^2}{c^2(1-\frac{r_s}{r})}-\frac{r^2\dot\phi^2}{c^2}} \tag{12.27}$$

where the Schwarzschild radius

$$r_s = \frac{2GM}{c^2} \tag{12.28}$$

We calculate the momenta

$$\begin{aligned} p_r &= \frac{m\dot{r}}{(1-r_s/r)\sqrt{1-\frac{r_s}{r}-\frac{\dot{r}^2}{c^2(1-\frac{r_s}{r})}-\frac{r^2\dot{\phi}^2}{c^2}}} \\ p_\phi &= \frac{mr^2\dot{\phi}}{\sqrt{1-\frac{r_s}{r}-\frac{\dot{r}^2}{c^2(1-\frac{r_s}{r})}-\frac{r^2\dot{\phi}^2}{c^2}}} = J \end{aligned} \tag{12.29}$$

Then the energy

$$E = \sum_i \dot{q}_i p_i - L = \frac{mc^2(1-r_s/r)}{\sqrt{1-\frac{r_g}{r}-\frac{\dot{r}^2}{c^2(1-\frac{r_s}{r})}-\frac{r^2\dot{\phi}^2}{c^2}}} \tag{12.30}$$

We make the assignments

$$\frac{\partial S}{\partial t} = -E, \quad \frac{\partial S}{\partial r} = p_r, \quad \frac{\partial S}{\partial \phi} = p_\phi \tag{12.31}$$

and indeed we have

$$g^{\mu\nu}\frac{\partial S}{\partial x^\mu}\frac{\partial S}{\partial x^\nu} = -\frac{E^2}{1-\frac{r_s}{r}} + (1-\frac{r_s}{r})c^2 p_r^2 + \frac{c^2 p_\phi^2}{r^2} = -m^2c^4 \tag{12.32}$$

For $m = 0$ we have

$$-\frac{E^2}{1-\frac{r_g}{r}} + (1-\frac{r_s}{r})c^2\left(\frac{\partial S}{\partial r}\right)^2 + \frac{c^2 p_\phi^2}{r^2} = 0 \tag{12.33}$$

and hence

$$S = \pm\int \mathrm{d}r\sqrt{\frac{E^2}{c^2(1-r_g/r)^2} - \frac{J^2}{r^2(1-r_g/r)}} - Et + J\phi \tag{12.34}$$

and we can recover the photon trajectory by

$$\frac{\partial S}{\partial J} = const = \phi_0 = \phi \mp \int \mathrm{d}r\frac{1}{r^2\sqrt{\frac{E^2}{J^2c^2}-\frac{1-r_s/r}{r^2}}} = \phi \pm \int \mathrm{d}w\frac{1}{\sqrt{\frac{E^2}{J^2c^2}-w^2(1-r_s w)}} \tag{12.35}$$

We recall the equation for the photon trajectory that we derived earlier

$$w'' + w = \frac{3r_s w^2}{2}, \quad w = \frac{1}{r} \tag{12.36}$$

and

$$w'^2 + w^2 - r_s w^3 = \frac{E^2}{J^2c^2} \tag{12.37}$$

so we recover the expression (12.35).

If we change the variable

$$w = \frac{4z + 1/3}{r_s} \tag{12.38}$$

then we have

$$\phi - \phi_0 = \int^{\infty} \frac{\mathrm{d}z}{\sqrt{4z^3 - g_2 z - g_3}}, \quad g_2 = \frac{1}{12}, \quad g_3 = \frac{2/27 - \alpha^2 r_s^2}{16} \tag{12.39}$$

and the solution is the Weierstrass (elliptic) function

$$z = \wp(\phi - \phi_0; g_2, g_3) \tag{12.40}$$

13 Fluid mechanics

13.1 Navier-Stokes equation

Fluid is described by several parameters – density, pressure, velocity, temperature etc. We divide the volume into very small domains – in each domain number of molecules is large but these parameters can be treated as constant and the whole distribution as continuous.

We have a convective time derivative (moving with the fluid)

$$\lim_{\delta t \to 0} \frac{\rho(\mathbf{r} + \mathbf{v}\delta t, t + \delta t) - \rho(\mathbf{r}, t)}{\delta t} = \frac{\partial \rho}{\partial t} + (\mathbf{v} \cdot \nabla)\rho \tag{13.1}$$

We will write this derivative as

$$\frac{D\rho}{Dt} = \frac{\partial \rho}{\partial t} + (\mathbf{v} \cdot \nabla)\rho \tag{13.2}$$

The continuity equation can be written as

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{13.3}$$

The Navier-Stokes equation reads

$$\rho \left(\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} \right) = -\nabla(\phi - \zeta \nabla \cdot \mathbf{v}) - \nabla p + \nu \left(\Delta \mathbf{v} + \frac{1}{3} \nabla (\nabla \cdot \mathbf{v}) \right) \tag{13.4}$$

where ν (shear) viscosity, sometimes written as $\mu\rho$ and ζ volume viscosity (often neglected); ϕ is the external potential. It is an unsolved problem to prove under what conditions the solutions exist...

For an incompressible fluid $\rho = \text{const}$ we have

$$\nabla \cdot \mathbf{v} = 0 \tag{13.5}$$

Then the NS equation reads

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\frac{1}{\rho}\nabla \phi - \nabla \frac{p}{\rho} + \mu \Delta \mathbf{v} \tag{13.6}$$

If on top $\nu = 0$ ('dry water') we have

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\nabla \left(\frac{\phi}{\rho} + \frac{p}{\rho} \right) \tag{13.7}$$

We can write it in a different form using the equality

$$(\mathbf{v}\cdot\nabla)\mathbf{v} = \mathbf{\Omega}\times\mathbf{v} + \frac{1}{2}\nabla v^2 \tag{13.8}$$

where

$$\mathbf{\Omega} = \nabla\times\mathbf{v} \tag{13.9}$$

Then

$$\frac{\partial\mathbf{v}}{\partial t} = \mathbf{v}\times\mathbf{\Omega} - \nabla\left(\frac{\phi}{\rho} + \frac{1}{2}v^2 + \frac{p}{\rho}\right) \tag{13.10}$$

If we have a stationary flow ($\partial\mathbf{v}/\partial t = 0$) then multiplying by $\mathbf{v}$ we get

$$\mathbf{v}\cdot\nabla\left(\frac{\phi}{\rho} + \frac{1}{2}v^2 + \frac{p}{\rho}\right) = 0 \tag{13.11}$$

which is a Bernoulli equation (the quantity inside the brackets is constant along the flow).

13.2 Propagation of sound

We now discuss the propagation of sound in the fluid (compressible, of course). We assume that the fluid is at rest and we write

$$\rho = \rho_0 + \delta\rho, \quad p = p_0 + \delta p, \quad \mathbf{v} = \delta\mathbf{v} \tag{13.12}$$

We expand the continuity equation to first order in perturbations

$$\frac{\partial\delta\rho}{\partial t} + \rho_0\nabla\cdot\delta\mathbf{v} = 0 \tag{13.13}$$

and we differentiate wrt time:

$$\frac{\partial^2\delta\rho}{\partial t^2} + \nabla\cdot\left(\rho_0\frac{\delta\mathbf{v}}{\partial t}\right) = \frac{\delta^2\rho}{\partial t^2} - \Delta\delta p = 0 \tag{13.14}$$

We can write this equation as

$$\frac{\partial^2\delta\rho}{\partial t^2} - c_0^2\Delta\delta\rho = 0 \tag{13.15}$$

where

$$c_0^2 = \left.\frac{\partial p}{\partial\rho}\right|_{p=p_0,\rho=\rho_0} \tag{13.16}$$

For adiabatic processes (compression for the sound wave is very fast) we have

$$pV^\kappa = \text{const} \Rightarrow p = \text{const}\rho^\kappa \Rightarrow \left.\frac{\partial p}{\partial\rho}\right|_{p=p_0,\rho=\rho_0} = \kappa\frac{p_0}{\rho_0} \tag{13.17}$$

what gives for air ($\kappa = 1.4$, $p = 10^5$ Pa, $\rho = 1.3$ kg/m^3)

$$c_0 = 330 \text{ m/s} \tag{13.18}$$

at $T = 0^\circ$ C.

13.3 Viscous fluid

Taking rotation of (13.7) we get

$$\frac{\partial \boldsymbol{\Omega}}{\partial t} + \nabla \times (\boldsymbol{\Omega} \times \mathbf{v}) = 0 \tag{13.19}$$

and in the case of the viscous fluid (13.6) it reads

$$\frac{\partial \boldsymbol{\Omega}}{\partial t} + \nabla \times (\boldsymbol{\Omega} \times \mathbf{v}) = \mu \Delta \boldsymbol{\Omega} \tag{13.20}$$

For dimensionless quantities (given by some characteristic length D and velocity U) we can rescale time and lengths ($x = D\tilde{x},\ v = U\tilde{v},\ t = D\tau/U$) to arrive at

$$\frac{\partial \tilde{\boldsymbol{\Omega}}}{\partial \tau} + \tilde{\nabla} \times (\tilde{\boldsymbol{\Omega}} \times \tilde{\mathbf{v}}) = \frac{1}{\mathcal{R}} \tilde{\Delta} \tilde{\boldsymbol{\Omega}} \tag{13.21}$$

where

$$\mathcal{R} = \frac{UD}{\mu} \tag{13.22}$$

is the so called Reynolds number (it is the principle of aerodynamic tunnels). For small Reynolds numbers the flow is laminar for larger turbulent.

13.4 Poiseuille flow

We have an incompressible fluid of viscosity μ in a pipe of radius R and length l with a laminar stationary flow. On the side at a radius r we have a force

$$F = -\nu 2\pi r l \frac{\mathrm{d}v}{\mathrm{d}r} \tag{13.23}$$

It has to be equal to the pressure difference inside the disc

$$F = \pi r^2 (p_1 - p_2) \tag{13.24}$$

Hence

$$\frac{\mathrm{d}v}{\mathrm{d}r} = -\frac{1}{2\nu l}(p_1 - p_2) r \tag{13.25}$$

so that

$$v(r) = \frac{p_1 - p_2}{4\nu l}(R^2 - r^2) \tag{13.26}$$

where the constant of integration was chosen to give $v(R) = 0$. The total volume per unit time that flows is given by

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \int \mathrm{d}r\, 2\pi r \frac{p_1 - p_2}{4\nu l}(R^2 - r^2) = \frac{\pi (p_1 - p_2) R^4}{8\nu l} \tag{13.27}$$

It is important to note the fourth power - if a vein has slightly smaller diameter because of for example thrombosis it can result in vastly smaller flow through the blood vessel.

13.5 Stokes' law

We now discuss the force acting on a ball with a laminar stationary flow. We start with

$$\frac{\partial \boldsymbol{\Omega}}{\partial t} + \nabla \times (\boldsymbol{\Omega} \times \mathbf{v}) = \mu \Delta \boldsymbol{\Omega} \tag{13.28}$$

We neglect the LHS (because of stationarity and the low Reynolds number) and we have to solve

$$\nabla \times (\nabla \times \Omega) = 0 \tag{13.29}$$

Far away we have

$$v_r = U \cos\theta, \quad v_\theta = -U \sin\theta \Rightarrow \Omega = 0 \tag{13.30}$$

where we used rotation in spherical coordinates

$$(\nabla \times \mathbf{A})_\phi = \frac{1}{r}\left(\frac{\partial}{\partial r}(rA_\theta) - \frac{\partial A_r}{\partial \theta}\right) \tag{13.31}$$

We assume the solution as

$$\Omega_\phi = U\frac{g(r)}{r}\sin\theta, \quad \Omega_r = \Omega_\theta = 0 \tag{13.32}$$

and we calculate the rotation

$$\nabla \times \Omega_r = 2U\frac{g}{r^2}\cos\theta, \quad \nabla \times \Omega_\theta = -U\frac{g'}{r}\sin\theta, \quad \Omega_\phi = 0 \tag{13.33}$$

and again

$$\nabla \times (\nabla \times \Omega)_\phi = -\frac{U}{r})(g'' - \frac{2g}{r})\sin\theta \tag{13.34}$$

Equating this to 0 we get the solution

$$g(r) = \frac{C}{r} \tag{13.35}$$

Now we have to find velocity. To solve the $\nabla \cdot \mathbf{v} = 0$ we assume

$$\mathbf{v} = \nabla \times \mathbf{w} \tag{13.36}$$

and we have to solve

$$\nabla \times (\nabla \times \mathbf{w}) = \frac{C}{r} \tag{13.37}$$

The solution is

$$\mathbf{w}_\phi = U(-C_1 r + \frac{C}{2} + \frac{C_3}{r^2})\sin\theta \tag{13.38}$$

Imposing the conditions at infinity and on $r = R$ we get

$$v_r = U\cos\theta(1 - \frac{3R}{2r} + \frac{R^3}{2r^3}), v_\theta = -U\sin\theta(1 - \frac{3R}{4r} - \frac{R^3}{4r^3}) \tag{13.39}$$

Then we find pressure as

$$\nabla p = \nu \nabla \times (\nabla \times \mathbf{v}) \tag{13.40}$$

Hence

$$p = p_\infty - \frac{3}{2} U R \nu \frac{\cos\theta}{r^2} \tag{13.41}$$

and the total force

$$F_l = 6\pi\nu U R \tag{13.42}$$

It should be compared with the turbulent flow

$$F_t = \frac{C_x \rho S U^2}{2} \tag{13.43}$$

Therefore for small Reynolds number $R_r = \rho 2RU/\nu$:

$$C_x = \frac{24}{R_r} \tag{13.44}$$

The assumption $\Omega_\phi = U\frac{g(r)}{r}\sin\theta$ stops to be valid at $R_r \sim 10$ - a better approximation up to $R_r < 10^4$ is

$$C_x = \frac{24}{R_r} + \frac{3.7}{\ln(2 + 4R_r)} \tag{13.45}$$

At $\sim 3 \cdot 10^5$ there is a sudden drop below 0.1 (drag crisis). If the sphere is rough the crisis appears earlier and therefore for example golf balls are made in the form of smooth polyhedrons and not round balls.

14 Deterministic chaos

14.1 Dissipative terms in the Hamilton's evolution

Up to now we discussed the equations without dissipative terms in the lagrangian or hamiltonian formulations.

We will now include in the description phenomenological method od 'damping the momentum' i.e.

$$\dot{p}_a = \frac{\partial H}{\partial q_a} - R_a(q,p) \tag{14.1}$$

with some functions R_a describing the dissipation.

We can calculate the dissipation introduced by these additional terms

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \sum_a \frac{\partial H}{\partial q_a}\dot{q}_a + \frac{\partial H}{\partial p_a}\dot{p}_a = -\sum_a R_a(q,p)\dot{q}_a \tag{14.2}$$

For example for the friction force proportional to the velocity we have

$$R_i = \gamma p_i \tag{14.3}$$

and the the 'leakage' of energy is equal to

$$\frac{\mathrm{d}H}{\mathrm{d}t} = -\frac{\gamma p^2}{m} \tag{14.4}$$

while for the 'aerodynamic' drag force proportional to v^2 we have

$$R_i = \frac{C_x \rho S}{2m^2} p p_i \tag{14.5}$$

and the the 'leakage' of energy is equal to

$$\frac{\mathrm{d}H}{\mathrm{d}t} = -\frac{C_x \rho S}{2} v^3 \tag{14.6}$$

14.2 Attractors

It may happen that the dynamics forces the trajectories in the phase space to be 'attracted' either to a point (fixed point) or to a higher dimensional hypersurface. The domain from which trajectories are 'attracted' is called the basin of attraction. We give below an example of such a behavior with the loop in the phase space

The model we consider is a so called Van der Pol's equation

$$m\ddot{y} + 2m\gamma(t)\dot{y} + m\omega^2 y = 0 \tag{14.7}$$

where

$$\gamma(t) = \gamma_0 \left(\frac{y^2(t)}{y_0^2} - 1 \right), \quad \gamma_0 > 0 \tag{14.8}$$

The 'damping term' damps the oscillations for large amplitudes but enhances them for small ones.

The equation is highly non-linear and does not have analytic solution so we will analyze it numerically. We first introduce dimensionless variables – then the equations read

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -q + (\epsilon - q^2)p \end{aligned} \tag{14.9}$$

It corresponds to

$$H = \frac{1}{2}(q^2 + p^2), \qquad R(q,p) = -(\epsilon - q^2)p \tag{14.10}$$

There is one parameter ϵ in this equation and the solution depends on the initial conditions (q_0, p_0) and ϵ.

The point $(0,0)$ is a saddle point but it is unstable. The trajectory 'around $q \sim \sqrt{\epsilon}$' is stable (a limit cycle) but it is not exactly a circle (the bigger ϵ is the more deformed it is). The question how to determine the shape of the attractor (the ultimate trajectory) is a global one and cannot be answered locally.

We can determine some properties of the attractor by some tricks for example we can use the fact that the attractor returns to its original values in the phase space after the whole turn. Therefore if we find some full derivative of any quantity then its average value shoud be zero for the attractor trajectory. We have then for example

$$< (\epsilon - q^2)p^2 >= 0 \tag{14.11}$$

For small ϵ the trajectory (as we can check numerically) is almost a circle. Plugging

$$q = R\cos t, \quad p = R\sin t \tag{14.12}$$

we get

$$\epsilon \frac{R^2}{2} - \frac{R^4}{8} = 0 \Rightarrow R = 2\sqrt{\epsilon} \tag{14.13}$$

For arbitrary ϵ we could use the fact that the trajectory is periodic and write

$$q = \sum q_n e^{in\omega t}, \quad p = \sum p_n e^{in\omega t} \tag{14.14}$$

and get a nonlinear algebraic equation for q_n and p_n.

14.3 Catastrophe theory of Thom - bifurcation of points

If we have a dynamical system given (in the matrix notation)

$$\dot{x} = F(\mu, x) \tag{14.15}$$

where $\mu = (\mu_1, \ldots, \mu_k)$ are parameters and x is an n-dimensional vector. If $k \leq 4$ then we have 7 types of possible bifurcation points i.e. the points where the character of the evolution can change. For $k = 5$ we have 11 types and for $k > 5$ there is an infinite number of possible types.

At $k \leq 4$ the bifurcation points can be described by the special points in polynomials in 1 (4 types) or 2 variables (3 types). The former are called cuspoidal the latter umbillic. The cuspoidal are given by $(F = V')$

- fold

$$V = x^3 + ax \tag{14.16}$$

- cusp

$$V = x^4 + ax^2 + bx \tag{14.17}$$

- swallowtail

$$V = x^5 + ax^3 + bx^2 + cx \tag{14.18}$$

- butterfly

$$V = x^6 + ax^4 + bx^3 + cx^2 + dx \tag{14.19}$$

The umbilic are given by

- hyperbolic

$$V = x^3 + y^3 + axy + bx + cy \tag{14.20}$$

- elliptic

$$V = x^3 - 3xy^2 + a(x^2 + y^2) + bx + cy \tag{14.21}$$

- parabolic

$$V = x^2y + y^4 + ax^2 + by^2 + cx + dy \tag{14.22}$$

For $k = 5$ we have one more cuspoidal (wigwam) and 3 more umbilic (second hyperbolic, second elliptic and symbolic).

We analyze below in some detail only a cusp.

The bifurcation can only happen at points only at points x_0 where $F(\mu, x_0) = 0$. Such a point can be stable or unstable. Then the behavior of the system depends on the matrix of second derivatives at the point x_0. If this matrix has one (or more) zero eigenvalues then such a point (or curve) is called a bifurcation point. We write the characteristic polynomial in a slightly different way

$$V' = x^3 - 3ax^2 + 2b \tag{14.23}$$

We have 3 roots of this equation and the bifurcation point is when the character of the roots changes from 1 real and 2 complex to 3 real. If we want to check when it happens we write

$$V' = (x-c)^2(x+2c) \tag{14.24}$$

and plugging into the original equation we have

$$a = c^2, \quad b = c^3 \tag{14.25}$$

or

$$a^3 - b^2 = 0 \tag{14.26}$$

and it is a condition for a bifurcation point.

14.4 Poincaré mapping

To visualize the flow it is convenient to use the notion of a Poincaré mapping. We denote a closed orbit in the phase space (the attractor) by Γ and we ask about the behavior of trajectories closed to it. We introduce the hypersurface S in some sense 'perpendicular' to Γ at some point x_0 on Γ and we choose this point to have $\tau = 0$. The neighborhood of x_0 we call S_0.Then the hypersurface will be punched in exactly the same point after T, $2T$ and so on where T is the period of the closed orbit. If we go away from the orbit Γ the other trajectories cross the hypersurface at some other time. The mapping

$$x \to \Phi(x) \tag{14.27}$$

such that $x_0 \to x_0$ after time T is for all points from S_0 called the Poincaré mapping π

$$S_0 \to \pi(S_0) = S_1 \tag{14.28}$$

Then we ask about a sequence

$$S_0 \to S_1 \to \dots S_n \tag{14.29}$$

It may happen that the sequence disperses or (as we expect for the limit cycle) it shrinks to smaller and smaller neighborhood of x_0. In order to answer the question whether the periodic orbit Γ is stable we ask about the so called characteristic multipliers of the linearized map

$$\left.\frac{\partial \pi^i}{\partial x^k}\right|_{x=x_0} \tag{14.30}$$

if all characteristic multipliers (eigenvalues of this equation) lie strictly inside the unit circle then the orbit γ is stable; if one or more lies outside then the orbit is unstable.

14.5 Bifurcation of periodic orbits

For periodic orbits there is a qualitatively new feature namely a possibility of period doubling. We now assume that the flow depends on some parameter ν. If we have a Poincaré mapping the matrix of first derivatives then if for a given ν the characteristic multipliers have all absolute value less than 1 then the orbit is stable. The interesting thing happens if for some value of ν one of the multipliers reaches -1. Then we return to the previous position in the direction of this multiplier after 2 turns and the orbit has twice bigger period (while the other directions shrink like the matrix of multipliers squared). Then we consider the Poincaré mapping after $2T$ and not T around the new 'fixed' trajectory with $\nu = 1$. It may turn out that changing ν from this new value around the new trajectory the situation repeats itself – one of the characteristic multipliers reaches -1 and we have yet another trajectory with the basic period $4T$. It may happen that the phenomenon repeats itself for smaller and smaller changes of ν and for a finite ν we reach infinite number of possible periodic orbits.

14.6 Deterministic chaos

We will describe below such a possibility for a Poincaré flow in one dimension on the most famous example of the logistic equation.

We start with some general remarks. If we measure some real value x_i, $i = 1, \ldots, n$ at the consecutive times iT then the predictive power is large if there is a strong correlation between x_1 and any later x_i even for large i. On the other hand if the correletaion is weaker and weaker then it is more and more difficult to predict the value of x_i for consecutive i's.

We can introduce a measure of this correlation by means of the following construction. For a sequence of real numbers

$$(x_1, \ldots, x_n) \tag{14.31}$$

where i denotes the time of measurement iT we assign the discrete Fourier transform numbers $\tilde{x}_\alpha$

$$\tilde{x}_\alpha := \frac{1}{\sqrt{n}} \sum_k \mathrm{e}^{-\mathrm{i}2\pi k\alpha/n}, \qquad \alpha = 1, \ldots, n \tag{14.32}$$

so that α correspond to the discrete frequency. They are complex numbers but they satisfy $\tilde{x}_\alpha^* = \tilde{x}_{n-\alpha}$. They also have the same norm

$$\sum_k x_k^2 = \sum_\alpha |\tilde{x}_\alpha|^2 \tag{14.33}$$

and there is an inverse transform

$$x_k := \frac{1}{\sqrt{n}} \sum_\alpha \mathrm{e}^{\mathrm{i}2\pi k\alpha/n} \tag{14.34}$$

We ask about the correlation in time namely about the quantity

$$C_\tau = \frac{1}{n}\sum_k x_k x_{k+\tau} \tag{14.35}$$

Plugging the expressions for x_k^* and $x_{k+\tau}$ and using the fact that $\tilde{x}_\alpha^* = \tilde{x}_{n-\alpha}$ we get

$$C_\tau = \frac{1}{n}\sum_\alpha |\tilde{x}_\alpha|^2 \mathrm{e}^{-2\pi\alpha\tau/n} = \frac{1}{n}\sum_\alpha |\tilde{x}_\alpha|^2 \cos(2\pi\alpha\tau/n) \tag{14.36}$$

The inverse transform gives

$$|\tilde{x}_\alpha|^2 = \frac{1}{n}\sum_\tau C_\tau \cos(2\pi\alpha\tau/n) \tag{14.37}$$

If C_τ goes to 0 for large τ then $|\tilde{x}_\alpha|^2$ has a continuous spectrum (and vice versa). If on the contrary C_τ does not decrease at large τ then $|\tilde{x}_\alpha|^2$ has sharp peaks around some frequencies. In the previous case we expect chaotic behavior in the latter a regular one.

14.7 Logistic equation

The logistic equation has one parameter ν

$$x_{k+1} = \nu x_k(1 - x_k) = F_1(\nu, x_k) \tag{14.38}$$

where

$$x_k \in [0,1], \quad 1 < \nu \leq 4 \tag{14.39}$$

where $\nu \leq 4$ to avoid moving out of the interval $[0,1]$. The fixed point of this transformation is for

$$\tilde{x} = \frac{\nu - 1}{\nu} \tag{14.40}$$

Hence $\nu > 1$.

The derivative

$$F_1'(\tilde{x}) = \nu(1 - 2\tilde{x}) = 2 - \nu \tag{14.41}$$

so that if $\nu < 3$ then the fixed point is stable since $|F_1'| < 1$. When $\nu_1 = 3$ we have a period doubling point so we start to analyze the new orbit $(F_1 \odot F_1)$

$$x_{k+1} = \nu^2 x_k(1 - x_k)(1 - \nu x_k(1 - x_k)) = F_2(\nu, x_k) \tag{14.42}$$

The previous fixed point $\tilde{x} = (\nu - 1)/\nu$ is unstable for $\nu > 3$ but there are new stable points

$$\tilde{x} = \frac{\nu + 1 \pm \sqrt{(\nu - 3)(\nu + 1)}}{2\nu} \tag{14.43}$$

and then the new bifurcation point is

$$F_2'(\tilde{x}) = -\nu^2 + 2\nu + 4 = -1 \quad \Rightarrow \quad \nu_2 = 1 + \sqrt{6} \approx 3.449... \tag{14.44}$$

Going further the new bifurcation points turn out to be denser and denser and at the point (discovered numerically by Feigenbaum in 1975)

$$\nu_\infty = 3.569945672... \tag{14.45}$$

at the ultimate rate

$$\delta = \lim_{k\to\infty} \frac{\nu_k - \nu_{k-1}}{\nu_{k+1} - \nu_k} = 4.669201 \tag{14.46}$$

there is a deterministic chaos - one cannot predict how the evolution will proceed.

Mathematical introduction

In the notation used in these lectures indices $i, j, k \dots$ will denote $1, 2, 3$ i.e spatial dimensions (Greek indices $\mu, \nu \dots = 0, 1, 2, 3$ will denote 4-dimensional quantities). The summation over repeated indices will always be implicitly assumed. The derivative with respect to time will be denoted by a dot and with respect to (cartesian) spatial directions by

$$\nabla_i := \frac{\partial}{\partial x^i} \equiv \partial_i \tag{14.47}$$

This operator has well defined properties under rotations and transforms tensors into tensors.

Vectors will be often denoted by boldface for example $\mathbf{r}$.

We introduce a scalar product of two vectors

$$\mathbf{A} \cdot \mathbf{B} := A_i B^i \tag{14.48}$$

with a number as a result and a vector product

$$(\mathbf{A} \times \mathbf{B})_i := \varepsilon_{ijk} A^j B^k \tag{14.49}$$

with a vector (in 3 dimensions) as a result – ε_{ijk} is a fully antisymmetric tensor with $\epsilon_{123} = 1$ (in 4 dimensions we choose the convention $\varepsilon^{0123} = 1$).

We have cyclic identity easy to prove by cyclicity of ε_{ijk}

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}) \tag{14.50}$$

We will often use the identity

$$\varepsilon_{ijk}\varepsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl} \tag{14.51}$$

Therefore, for example

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}) \tag{14.52}$$

In cylindrical coordinates

$$\mathbf{e}_\rho, \quad \mathbf{e}_\phi, \quad \mathbf{e}_z, \qquad \mathbf{r} = \rho \mathbf{e}_\rho + z \mathbf{e}_z \tag{14.53}$$

we have

$$\dot{\mathbf{e}}_\rho = \dot{\phi} \mathbf{e}_\phi, \quad \dot{\mathbf{e}}_\phi = -\dot{\phi} \mathbf{e}_\rho, \quad \dot{\mathbf{e}}_z = 0 \tag{14.54}$$

so that the velocity

$$\mathbf{v} = \dot{\rho} \mathbf{e}_\rho + \rho \dot{\mathbf{e}}_\rho + \dot{z} \mathbf{e}_z = \dot{\rho} \mathbf{e}_\rho + \rho \dot{\phi} \mathbf{e}_\phi + \dot{z} \mathbf{e}_z \tag{14.55}$$

and

$$v^2 = \dot{\rho}^2 + \rho^2 \dot{\phi}^2 + \dot{z}^2 \tag{14.56}$$

The laplacian on a scalar function f reads

$$\Delta f = \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial f}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2 f}{\partial\phi^2} + \frac{\partial^2 f}{\partial z^2} \tag{14.57}$$

The laplacian on a vector function $\mathbf{A}$ reads

$$\begin{aligned}\Delta\mathbf{A} &= \mathbf{e}_\rho\left(\Delta A_\rho - \frac{A_\rho}{\rho^2} - \frac{2}{\rho^2}\frac{\partial A_\phi}{\partial\phi}\right)\\ &+\mathbf{e}_\phi\left(\Delta A_\phi - \frac{A_\phi}{\rho^2} + \frac{2}{\rho^2}\frac{\partial A_\rho}{\partial\phi}\right)\\ &+\mathbf{e}_z\Delta A_z\end{aligned} \tag{14.58}$$

In spherical coordinates

$$\mathbf{e}_r, \quad \mathbf{e}_\theta, \quad \mathbf{e}_\phi, \quad \mathbf{r} = r\mathbf{e}_r \tag{14.59}$$

we have

$$\begin{aligned}\dot{\mathbf{e}}_r &= \dot\theta\mathbf{e}_\theta + \dot\phi\sin\theta\mathbf{e}_\phi,\\ \dot{\mathbf{e}}_\theta &= -\dot\theta\mathbf{e}_r + \dot\phi\cos\theta\mathbf{e}_\phi\\ \dot{\mathbf{e}}_\phi &= -\dot\phi\sin\theta\mathbf{e}_r - \dot\phi\cos\theta\mathbf{e}_\theta\end{aligned} \tag{14.60}$$

so that the velocity

$$\mathbf{v} = \dot r\mathbf{e}_r + r\dot{\mathbf{e}}_r = \dot r\mathbf{e}_r + r\dot\theta\mathbf{e}_\theta + r\dot\phi\sin\theta\mathbf{e}_\phi \tag{14.61}$$

and

$$v^2 = \dot r^2 + r^2\dot\theta^2 + r^2\sin^2\theta\dot\phi^2 \tag{14.62}$$

The laplacian on a scalar function f reads

$$\Delta f = \frac{1}{r}\frac{\partial^2(rf)}{\partial r^2} + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial\phi^2} \tag{14.63}$$

The laplacian on a vector function $\mathbf{A}$ reads

$$\begin{aligned}\Delta\mathbf{A} &= \mathbf{e}_r\left(\Delta A_r - \frac{2A_r}{r^2} - \frac{2}{r^2\sin\theta}\frac{\partial(A_\theta\sin\theta)}{\partial\theta} - \frac{2}{r^2\sin\theta}\frac{\partial A_\phi}{\partial\phi}\right)\\ &+\mathbf{e}_\theta\left(\Delta A_\theta - \frac{A_\theta}{r^2\sin^2\theta} + \frac{2}{r^2}\frac{\partial A_r}{\partial\theta} - \frac{2\cos\theta}{r^2\sin^2\theta}\frac{\partial A_\phi}{\partial\phi}\right)\\ &+\mathbf{e}_\phi\left(\Delta A_\phi - \frac{A_\phi}{r^2\sin^2\theta} + \frac{2}{r^2\sin\theta}\frac{\partial A_r}{\partial\phi} + \frac{2\cos\theta}{r^2\sin^2\theta}\frac{\partial A_\theta}{\partial\phi}\right)\end{aligned} \tag{14.64}$$

We define the action S as a functional

$$S = \int_{t_i}^{t_f} L(t, x^A, \dot x^A)\,\mathrm{d}t \tag{14.65}$$

so it depends upon the path between t_i and t_f.

We want to find such a trajectory that is the extremum of S.

We consider the actual path $x^A(t)$. If it is an extremum of S it means that any deviation from the trajectory does not change S up to terms linear in the deviation. We add the deviation

$$x^A(t) \to x^A(t) + \delta x^A(t) \tag{14.66}$$

and we calculate the change of the action for the perturbed trajectory (keeping the initial and final times and the end points of the trajectory unchanged)

$$\delta S = \delta \int_{t_i}^{t_f} L(t, x^A, \dot{x}^A)\,\mathrm{d}t = \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial x^A}\delta x^A + \frac{\partial L}{\partial \dot{x}^A}\delta \dot{x}^A \right)\,\mathrm{d}t \tag{14.67}$$

We integrate by parts and we get up to linear terms in δx^A

$$\delta S = \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial x^A} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}^A}\right)\right)\delta x^A\,\mathrm{d}t + \left[\frac{\partial L}{\partial \dot{x}^A}\delta x^A\right]_{t_i}^{t_f} \tag{14.68}$$

According to our assumption the endpoints of the trajectory are kept fixed so the last term vanishes. Since $\delta x^A(t)$ is arbitrary we conclude that for each A

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}^A}\right) - \frac{\partial L}{\partial x^A} = 0 \tag{14.69}$$

These equations are called Euler-Lagrange equations.

We see that adding a full time derivative to L does not change the equations of motion so we treat such lagrangians as equivalent:

$$L \equiv L + \frac{\mathrm{d}f}{\mathrm{d}t} \tag{14.70}$$