Lecture 1

Alexey Boyarsky

September 8, 2015

1 Introduction

According to the traditional view, mechanics is the branch of physics that studies the behaviour of physical bodies when subjected to forces or displacements. The first part of this course, Classical Mechanics A, has considered main principles and examples of classical mechanics in this traditional sense.

The main goal of theoretical physics however is to build the correspondence between physical phenomena in the real world and mathematical models describing some of their properties. It appears that many mathematical ideas of mechanics and its two main parts, Kinematics (that studies the possible states of physical systems and methods of their description) and Dynamics (that predicts how a given state of the system will evolve in the future) are very general and are very useful in a context that is much wider than the classical realm of mechanics.

The goal of this second part of the course is to develop this more general view on mechanics and, in particular, introduce Lagrangian and Hamiltonian formalisms, that are very important in the description of many physical systems.

2 Kinematics. Description of the states of physical system. Examples.

2.1 Point particles.

The simplest abstraction of a physical system is a point particle. Its state is characterised by its *position* at any moment of time

$$\vec{r} = \vec{r}(t)$$

In the Cartesian coordinate system $\vec{r}(t) = (x(t), y(t), z(t))$. Depending on the movement of the particles, it may be convenient to use different coordinate system, for example

spherical coordinates: (see 1.12 of [1])

$$\begin{aligned} x =& r\sin\theta\cos\varphi \\ y =& r\sin\theta\sin\varphi \\ z =& r\cos\theta \end{aligned} \tag{1}$$

and describe the state of the system by the set of functions $r(t)$, $\theta(t)$, $\varphi(t)$. Alternatively, one can use other coordinate systems,

cylindrical coordinates: (ρ, ϕ, Z) such that their relation to Cartesian coordinates

$$\begin{aligned} x &= \rho \cos \phi \\ y &= \rho \sin \phi \\ z &= Z \end{aligned} \tag{2}$$

parabolic coordinates: (σ, τ, φ)

$$\begin{aligned} x &= \sigma\tau \cos \varphi \\ y &= \sigma\tau \sin \varphi \\ z &= \frac{1}{2}\left(\tau^2 - \sigma^2\right) \end{aligned} \tag{3}$$

but in any case three independent numbers are required to describe the position of the particles of three independent functions of time to define its state at any moment.

2.2 Several particles

As the next simplest example consider two particles with the masses m_1 and m_2. We can specify position of each particle by its own radius vector $\vec{r}_1$ and $\vec{r}_2$. Or, alternatively we can introduce the *center-of-mass* position $\vec{R}_{c.m.}$ (see 7.1 of [1])

$$\vec{R}_{c.m.} = \frac{m_1\vec{r}_1 + m_2\vec{r}_2}{m_1 + m_2} \tag{4}$$

and the relative radius vector

$$\vec{r}_{12} = \vec{r}_1 - \vec{r}_2 \tag{5}$$

In either case we need *6 scalar functions to specify all positions.*

In general for the system of N particles one needs $3N$ scalar functions

2.3 Pendulum: example of a system with constraints

Consider a *pendulum* – a material point of mass m moving in the xy plane, attached to a (massless) rigid rod of length l (Fig. 8). Naively, the position of this material point is described by two coordinates $x(t)$ and $y(t)$ and two velocities: $\dot{x}(t)$ and $\dot{y}(t)$. However, it is clear that *at any moment of time* x and y are subject to a **constraint**:

$$x^2(t) + y^2(t) = l^2 \tag{6}$$

Moreover, velocity is also constraint. As the rod is rigid (i.e. it cannot change its length) the velocity vector $\vec{v} = (\dot{x}, \dot{y})$ is always directed perpendicular to the rode. This is a consequence of the constraint (73). As it is valid at all times t, differentiating this constraint with respect to time we get:

$$x(t)\dot{x}(t) + y(t)\dot{y}(t) = 0 \tag{7}$$

Thus, we see that out of two functions $x(t), y(t)$ only one is independent. And similarly, out of two velocity components $(\dot{x}, \dot{y})$ only one is independent. Of course, one can choose as such a variable $x(t)$ and $\dot{x}(t)$ and express

$$\begin{aligned} y(t) &= \pm\sqrt{l^2 - x^2(t)} \\ \dot{y}(t) &= \mp\frac{x(t)\dot{x}(t)}{\sqrt{l^2 - x^2(t)}} \end{aligned} \tag{8}$$

However, one can introduces the angle θ, such that

$$\tan\theta(t) = -\frac{x(t)}{y(t)} \tag{9}$$

This angle uniquely specifies the position:

$$\begin{aligned} x(t) &= l\cos\theta(t) \\ y(t) &= l\sin\theta(t) \end{aligned} \tag{10}$$

and velocity:

$$\begin{aligned} \dot{x}(t) &= -l\dot{\theta}\sin\theta(t) \\ \dot{y}(t) &= \quad l\dot{\theta}\cos\theta(t) \end{aligned} \tag{11}$$

We will see below (Section 6.2) that the dynamics can be expressed in terms of the same variable $\theta(t)$.

2.4 Rigid body

Next example is a rigid body (see 8.1 of [1]). Naively this is enormous collection of particles with $N \sim 10^{23}$. However, the term *rigid* means that the relative positions $\vec{r}_{ij}$ between particles i and j are fixed. As a result the state of rigid body can be given by center-of-mass co-ordinates and by 3 angles that define rigid body position (for example Euler angles, see Fig. 1 and 9.6 of [1]), so we need *6 variables.*

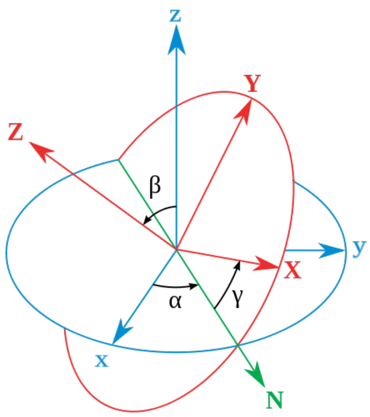

Figure 1: Proper Euler angles. The xyz (original) system is shown in blue, the XYZ (rotated) system is shown in red. The line of nodes (N) is shown in green.

2.5 Fluid

More complex example is a *fluid* – a system of interacting particles where the interaction however does not keep *shape* of any volume fixed. When describing a fluid, one is concentrated on the dynamics of *macroscopic* (albeit small) volumes of fluid, containing many molecules. The motion of these molecules defines the thermodynamics properties of a fluid (pressure, density, temperature, etc.)

In the *Lagrangian picture*, individual fluid parcels are followed through time. The fluid parcels are labelled by some (time-independent) vector field $\vec{a}$. (Often, $\vec{a}$ is chosen to be

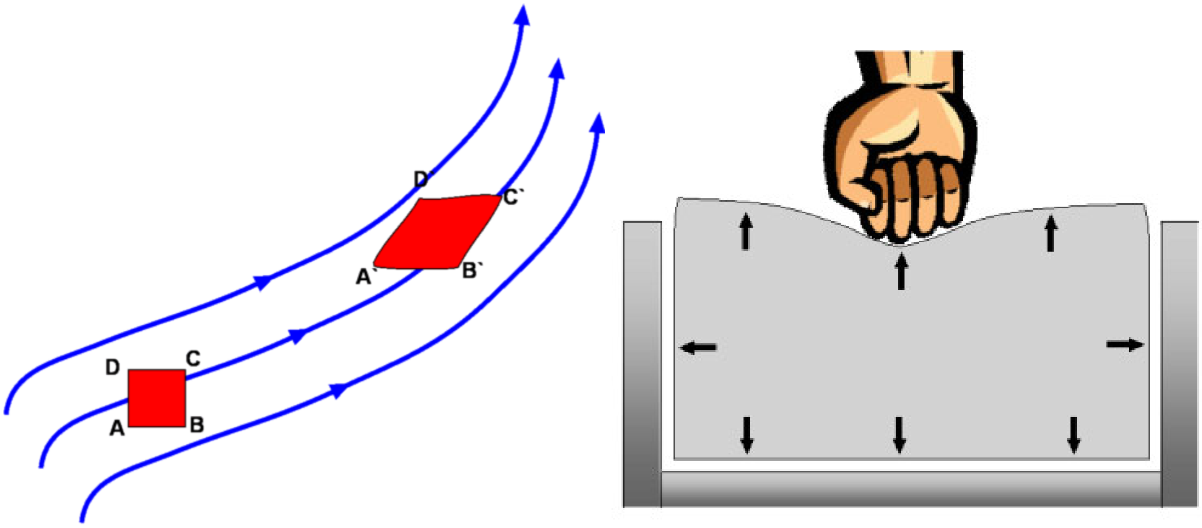

Figure 2: Lagrangian fluid description a parcel ABCD moves with the flow $\vec{v}$. If the flow is *incompressible*, the polygon ABCD deforms but its volume (area) remains constant. Mathematically incompressibility means that $\nabla \cdot \vec{v}(\vec{r}) = 0$.

the center of mass of the parcels at some initial time t_0). It is chosen in this particular manner to account for the possible changes of the shape over time. Therefore the center of mass is a good parametrization of the velocity $\vec{v}$ of the parcel.) In the Lagrangian description, the flow is described by a function $\vec{X}_{\vec{a}}(t)$ giving the position of the parcel labeled $\vec{a}$ at the time t.

In the *Eulerian picture* of the flow field, the flow quantities are depicted as a function of position $\vec{x}$ and time t. Specifically, the flow is described by a function $\vec{v}\,(\vec{x}, t)$ giving the flow velocity at position $\vec{x}$ at time t.

The two specifications are related as follows:

$$\vec{v}(\vec{x}, t) = \vec{v}\left(\vec{X}_{\vec{a}}(t)\right) = \frac{\partial \vec{X}_{\vec{a}}(t)}{\partial t} \tag{12}$$

because both sides describe the velocity of the parcel labeled $\vec{a}$ at the time t. The number of needed variables is formally infinite, one vector for each point r, but *locally* there are only there numbers:

$$\vec{v}_{\vec{x}}(t) = \vec{v}(\vec{x}, t).$$

2.6 Electromagnetic field

This is the first example of a system that can not be defined as a collection of (interacting) particles. This also also infinitely-dimensional system: a configuration of electromagnetic field is described by 2 vector fields: electric field $\vec{E}_{\vec{r}}(t) = \vec{E}(\vec{r}, t)$ and magnetic $\vec{B}_{\vec{r}}(t) = \vec{B}(\vec{r}, t)$ or 6 numbers for each point in space (6 scalar functions). *However, as we will see below, not all of these functions are independent !*

3 Equations of motion and degrees of freedom

Already in the simplest examples we see, however, that it is not enough to know only positions of particles or configuration of the the fields. For a point particle we need to know also the velocity of the particle. The same position but different velocity at the initial state will result in very different evolution of the the same particle. To define a state of a system completely, we need to specify enough information to be able to determine this state at any moment in the future. Therefore, the definition of state of a system is defined by its dynamics or equations of motion. For a wide class of physical system the equation of motion are local differential equation, second order in time. To find a specific solution of such equations (and define evolution of the physical system at hand) we need specify the initial conditions containing the both the value of the function and its time derivative at the initial moment of time, the analog of the initial position and initial velocity. For the systems that can be understood as collections of (interacting) particles, equations of motion can in principle be derived from the Newtons equations for point particles (see 2.1 of [1]):

$$m\ddot{\vec{r}} = \sum_i \vec{F}_i. \tag{13}$$

In the case of electric field:

$$m\ddot{\vec{r}} = q\vec{E} \tag{14}$$

If we want also to include magnetic field we need to write a term $q[\vec{v}\times\vec{B}]/c$ in right-hand side, where c is the speed of light:

$$\vec{F}_L = q\left(\vec{E} + \frac{1}{c}[\vec{v}\times\vec{B}]\right) \tag{15}$$

This term is proportional to v/c and vanishes in non-relativistic limit. Therefore, let us consider dynamics of relativistic particle.

3.1 Relativistic particles

For relativistic particles the equations of motion are very similar, e.g. for a particle in external electromagnetic field:

$$\begin{cases} m\dfrac{du^0}{ds} = q(\vec{u}\cdot\vec{E}) \\ m\dfrac{d\vec{u}}{ds} = q\left(u^0\vec{E} + [\vec{u}\times\vec{B}]\right) \end{cases} \tag{16}$$

where $u^\mu = (u^0, \vec{u}) = dx^\mu/ds$ is a 4-vector of velocity, s — is an interval, q — is a particle's charge. *Here the coordinates are x^μ and velocities are u^μ*:

$$u^\mu = \left(\frac{c}{\sqrt{1-v^2/c^2}}, \frac{\vec{v}}{\sqrt{1-v^2/c^2}}\right) \tag{17}$$

In terms of the the field strength $F^{\mu\nu}$ with $E^i = F^{0i}$ and $B^i = \frac{1}{2}\epsilon^{ijk}F_{jk}$ this can be rewritten in a more compact form

$$m\frac{d^2x^\mu}{ds^2} = qF^{\mu\nu}u_\nu \tag{18}$$

3.2 Particles, connected by springs

Let us consider equations of motion of a system consisting of many interacting particles, for example a system of equal particles connected by springs (see 11.4 of [1]): We will

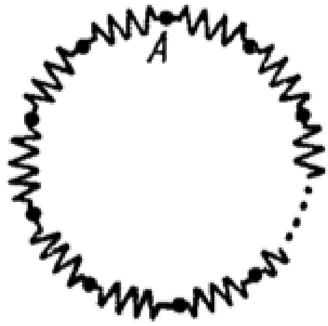

Figure 3: System of large amount of particles connected by springs

denote by x_i the displacement of the i^{th} particle from its equilibrium position (i runs from 1 to N, and $x_{N+1} = x_1$). So the equation on i^{th} particle is:

$$m\ddot{x}_i = k(x_{i-1} - x_i) - k(x_i - x_{i+1}) = k(x_{i+1} + x_{i-1} - 2x_i) \tag{19}$$

Let we denote equilibrium distance between particles as a and use continius function for deviation from equilibrium state $u(x)$, such that $x_i \to u(x)$, $x_{i-1} \to u(x-a)$ and $x_{i+1} \to u(x+a)$. In limit $a \to 0$ we have:

$$\lim_{a\to 0}\frac{k}{m}(u(x+a) + u(x-a) - 2u(x)) = \frac{d^2u}{dx^2}\lim_{a\to 0}\frac{ka^2}{m} \tag{20}$$

(to prove it use taylor series). Denoting $\lim\limits_{a\to 0}\frac{ka^2}{m}$ as κ we obtain equation of string motion (see 11.5 of [1]):

$$\frac{\partial^2 u}{\partial t^2} = \kappa\frac{\partial^2 u}{\partial x^2} \tag{21}$$

Let us introduce the number of **degrees of freedom** defined by *(one half of)* the number of functions one needs to specify at time $t = 0$ to fully determine future evolution of the system at $t > 0$.

In all the examples above, that equations of motion are second order differential equations, and for solving them we need to specify the position or *configuration function of the system* and its derivative in the initial moment of time. It is because all these systems are different generalizations of particle dynamics and Newton's second law. However, this analogy is not always the case!

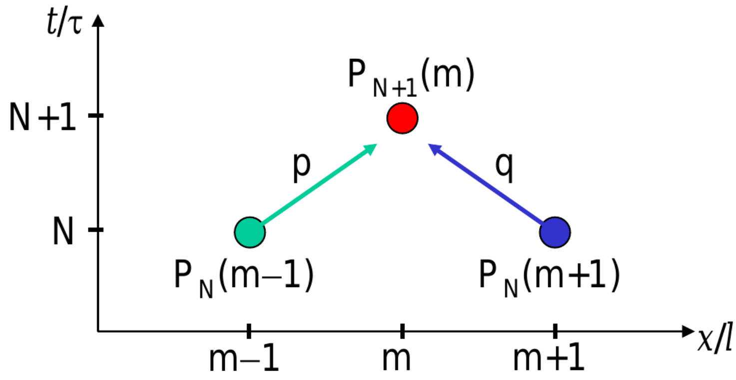

Figure 4: Random walk of the particle on $1D$ lattice.

3.3 The diffusion equation

Let us consider a very small particle in a medium (e.g. gas). The particles forming the medium hit our test particle from time to time (as they are involved in the random thermal motion) and displace it. Let us find a mathematical model describing the evolution of such a test particle in a medium. For simplicity, let us assume that the particle can move only along a line (one-dimensional motion). We can measure the position of the particle at discrete moments of time separated by equal time intervals τ : $t = 0, \tau, 2\tau, ..N\tau, \ldots$. Let us assume that during this time intervals our particle is displaced by equal distance l to the left or to the right. The direction is random (as it the displacement is caused by collisions with randomly moving thermal particles). The probability to move the particle one step right is equal to p and the probability to move one step left is $q = 1 - p$.

We are interested in the probability to find a particle in the point x at the time moment t. In our discrete model $x = ml$ and $t = \tau N$. Therefore, we are looking for the function on the lattice, $P_N(m)$. To be in at the point m at the time moment $N + 1$, our particle had either to come from the left (if it was at the moment N at the point $m - 1$ and moved to the right) or from the right (if it was at the point $m + 1$ and moved left), see Fig.4. Therefore, we have the following equation for the evolution of probability:

$$P_{N+1}(m) = pP_N(m-1) + qP_N(m+1) \tag{22}$$

The most interacting case of this theory appears when $p = q = 1/2$ (there is no directed motion in average). Subtracting $P_N(m)$ from both sides of (22):

$$P_{N+1}(m) - P_N(m) = \frac{1}{2}\left(P_N(m-1) + P_N(m+1) - 2P_N(m)\right) \tag{23}$$

Taking limit $\tau \to 0$, $l \to 0$ we can write:

$$P_{N+1}(m) - P_N(m) = \frac{\tau}{\tau}(P_{N+1}(m) - P_N(m)) \approx \tau\frac{\partial P}{\partial t} \tag{24}$$

$$\frac{l^2}{2l^2}\left(P_N(m-1) + P_N(m+1) - 2P_N(m)\right) \approx \frac{l^2}{2}\frac{\partial^2 P}{\partial x^2} \tag{25}$$

Introducing a constant $D = \lim_{l,\tau\to 0} \frac{l^2}{2\tau}$ we have diffusion equation which describes evolution of probability distribution:

$$\frac{\partial P}{\partial t} = D\frac{\partial^2 P}{\partial x^2} \tag{26}$$

If one starts with a very sharp (δ-function shaped) peak at $t = 0$, the evolution of this peak with time will be given by the following expression (see Section 3.3.1 for details):

$$P(x,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}e^{-Dk^2t}dk = \frac{1}{2\sqrt{\pi Dt}}e^{-\frac{x^2}{4D^2t^2}} \tag{27}$$

This is a Gaussian distribution with variance equal to $\sigma^2 = 2Dt$. Time dependence of this function is shown on Fig. 5.

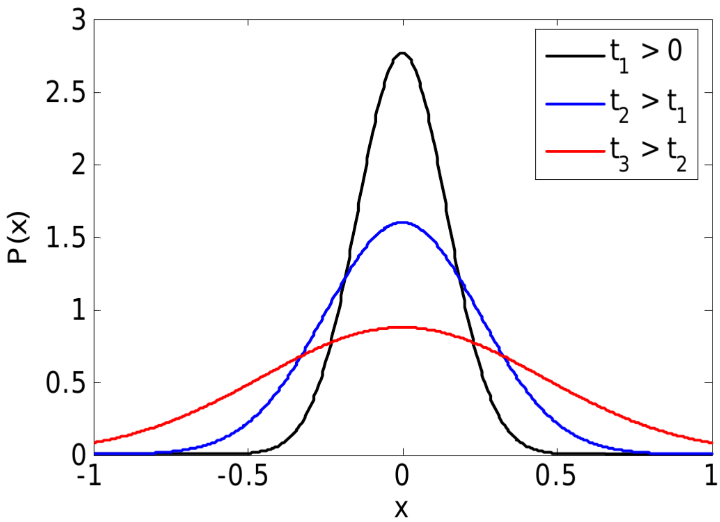

Figure 5: Diffusion evolution of initial point-like distribution.

For the mixture of gases diffusion is also described by equation of the same form:

$$\frac{\partial n_A}{\partial t} = D\frac{\partial^2 n_A}{\partial x^2} \tag{28}$$

where n_A is a number density of particles of some type A. This is the first order differential equation in time on particles concentration $n_A(x,t)$. So the only degree of freedom is a function $n_A(x,t)$.

3.3.1 A general solution of the diffusion equation

This is an additional information for those, not needed for the main course

It's not difficult to find a general solution of this equation which is not grow on infinity. Suppose $P(x,t) = X(x)T(t)$, where X and T are some unknown functions. Substituting in (26):

$$\frac{\dot{T}}{T} = D\frac{X''}{X} \tag{29}$$

Left-hand side depends only on t, whereas right-hand side depends only on x. The only possibility is that both sides are equal to some constant λ.

1. Let $\lambda = Dk^2 > 0$. Then for X we have:
$$X'' = k^2 X \tag{30}$$
and solutions are $X(x)_{1,2} = C_{1,2}e^{\pm kx}$. One of this solutions grow on $+\infty$, other on $-\infty$, so they are not interesting for us.

2. If $\lambda = 0$ we have:
$$X'' = 0 \tag{31}$$
with solution $X = Ax + B$. From condition on infinities $A = 0$, so X is some constant. For T we also have $\dot{T} = 0$, so T is a constant also.

3. In case $\lambda = -Dk^2 < 0$:
$$X'' + k^2 X = 0 \tag{32}$$
and solutions are $X(x)_{1,2} = C_{1,2}e^{\pm ikx}$, which don't grow on infinity. For T we have:
$$\dot{T} = -Dk^2 T \tag{33}$$
with solution $T(t) = Ce^{-Dk^2 t}$.

This equation is linear, so the full solution is a sum of particular ones. Taking all cases together we have:
$$P(x,t) = \sum_k C_k e^{ikx} e^{-Dk^2 t} = \int_{-\infty}^{\infty} C(k) e^{ikx} e^{-Dk^2 t} dk \tag{34}$$
because k is continuous variable. Coefficients $C(k)$ are determined by initial conditions.

Suppose initial conditions in form $P(x,0) = \delta(x)$, which means that particle has probability 1 to be in the point $x = 0$. So, for $t = 0$:
$$\delta(x) = \int_{-\infty}^{\infty} C(k) e^{ikx} dk \tag{35}$$
So it's easy to find $C(k)$:
$$C(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \delta(x) e^{-ikx} dx = \frac{1}{2\pi} \tag{36}$$

3.4 Euler equation for fluid

Let's look at equations of motion of ideal fluid.

The first one is the *continuity equation*. The mass in some (small but macroscopic) volume V is $M = \int\limits_V \rho dV$ and its change per unit time is equal to total mass flow (given by in the integral of $\rho\vec{v}$ over the surface of this volume S):
$$\frac{dM}{dt} = -\int\limits_S \rho\vec{v} \cdot \vec{n} dS, \tag{37}$$
where $\vec{n}$ is a unit vector normal to the surface and oriented out from it. Using Gauss theorem we have:
$$\frac{d}{dt} \int\limits_V \rho dV = -\int\limits_V \vec{\nabla} \cdot (\rho\vec{v}) dV \tag{38}$$

so we arrive to the continuity equation:

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \vec{v}) = 0 \tag{39}$$

If $div\ \vec{v} = \nabla \cdot \vec{v} = 0$ (incompressible fluid !) this equation can be rewritten as the

$$\frac{d\rho}{dt} \equiv \frac{\partial \rho}{\partial t} + \vec{v} \cdot \nabla \rho = 0 \tag{40}$$

It is the first equation. For second one let's use Newton's second law to our volume:

$$\int\limits_V \frac{d\mathbf{v}}{dt}\, dm = \int\limits_V \mathbf{g}\, dm - \oint\limits_S p\, d\mathbf{S} \tag{41}$$

where $\mathbf{g}$ is force per unit volume of fluid and p is pressure. Using $dm = \rho\, dV$ and Gauss theorem again we have:

$$\rho \frac{d\mathbf{v}}{dt} = \rho \mathbf{g} - \nabla p \tag{42}$$

using the identity $\frac{d\mathbf{v}(\mathbf{r},t)}{dt} \equiv \frac{\partial \mathbf{v}}{\partial t} + \frac{\partial \mathbf{v}}{\partial x_i}\frac{\partial x_i}{\partial t} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v}$ and we finally have:

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} = \mathbf{g} - \frac{1}{\rho}\nabla p \tag{43}$$

So we have 2 equations of first order on field of velocities and density, they are degrees of freedom. The pressure is defined from the equation of fluid state. Although Eq. (43) is just the second Newton's law, the equation is first order in time because it is velocity $\vec{v}_{\vec{x}}(t)$ is the coordinate in this case. The variable $\vec{x}$ is just an index, marking "fluid particle".

3.5 Systems with constraints

3.5.1 Pendulum

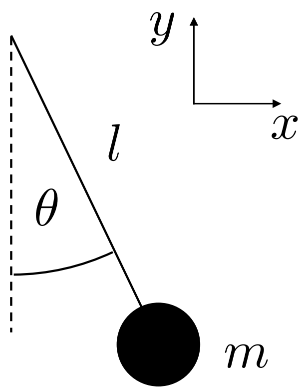

Figure 6: The pendulum.

Let us come back to the example of pendulum, whose kinematics we have already considered in Section 2.3. Let us write its equations of motion. The main problem here is that to describe the dynamics of the material point m in x, y coordinates we need to introduce some explicit forces. In this case this is the tension of the rope, $T = (T_x, T_y)$. So the laws of its motion is:

$$m\ddot{x} = -T_x \tag{44}$$

$$m\ddot{y} = T_y - mg \tag{45}$$

We have 4 unknown functions (x y, T_x, T_y) and only 2 equations. But we also have *constraints* — relations between quantities that describe state of the system. Constraints are some functions of form $F(q_1, \dots, \dot{q}_1, \dots) = 0$. For pendulum we have constraints (73)

$$x^2 + y^2 = l^2 \tag{46}$$

$$\tag{47}$$

and its time derivative, saying that velocity is orthogonal to the rod (Eq. (7)).

The problem can be simplified if we introduce coordinates that take into account constraints. It's easy to see that position of the system is fully determined by an angle θ, Eq. (9):

$$\begin{aligned} x(t) &= l\cos\theta(t) \\ y(t) &= l\sin\theta(t) \\ \dot{x}(t) &= -l\dot{\theta}\sin\theta(t) \\ \dot{y}(t) &= \quad l\dot{\theta}\cos\theta(t) \end{aligned} \tag{48}$$

and write the 2nd Newton's law in the coordinates *along* and *orthogonal to* the rod (radial coordinates). The radial movement (movement along the rod) is trivial:

$$r(t) = \sqrt{x^2(t) + y^2(t)} = l = \text{const} \tag{49}$$

while the equation, projected to the direction *orthogonal* to the rod, does not contain the unknown and redundant tension force $\vec{T}$ and can be written as

$$ml^2\ddot{\theta} = -mgl\sin\theta \tag{50}$$

One recognizes in Eq. (50) the equation for the angular momentum with the torque due to gravitational force in the right hand side.

3.5.2 Pendulum on the movable support

Consider the system of 2 particles in the gravitational field of mass M and m. The first particle moves along horizontal wire, and the second is connected to in through the rope of length l, see Fig. 7. Using Newton's laws we can write equations on system's dynamics:

$$M\ddot{x}_M = T_x \tag{51}$$

$$M\ddot{y}_M = N - Mg - T_y \tag{52}$$

$$m\ddot{x}_m = -T_x \tag{53}$$

$$m\ddot{y}_m = T_y - mg \tag{54}$$

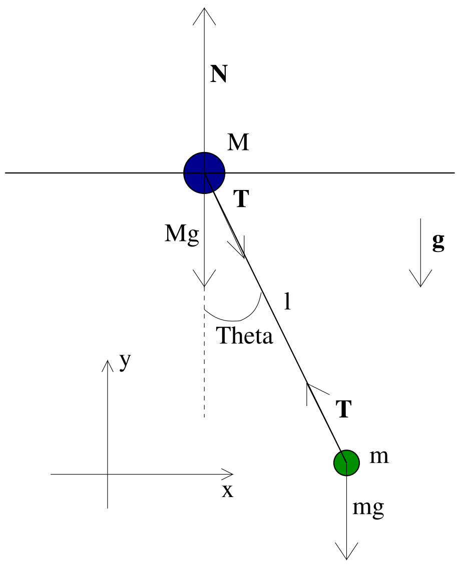

Figure 7: Pendulum on the movable support.

In this system we have 7 unknown functions of time: x_M, y_M, x_m, y_m, T_x, T_y (tension of the rope), N (reaction of the wire). Again, the situation is the same as in the previous Section 6.2: we have 3 unknown forces (T_x, T_y, N) that should be found if we are working in the redundant coordinates x_m, x_M, y_m, y_M.

We have only 4 equations (51)–(54), so where do we get other 3? Needed relations come from constraints. In our case we have 2 constraints:

$$y_M = const \tag{55}$$

$$(x_M - x_m)^2 + (y_M - y_m)^2 = l^2 \tag{56}$$

We can introduce an angle $\theta(t)$ – between the rod and the vertical line:

$$\tan\theta = -\frac{x_M - x_m}{y_M - y_m} \tag{57}$$

Differentiating constraints we can, in principle, find relations between accelerations and angle θ and plug it into equation of motions. It's easy to see, that position of the system is fully determined by 2 quantities: the x coordinate of the center of mass $x_{\text{c.m.}} = (Mx_M + mx_m)/(M + m)$ and θ.

3.6 Maxwell's equations

Electromagnetic field is even more complicated:

$$\vec{\nabla} \cdot \vec{E} = 4\pi\rho \tag{58}$$

$$\vec{\nabla} \cdot \vec{B} = 0 \tag{59}$$

$$\vec{\nabla} \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t} \tag{60}$$

$$\vec{\nabla} \times \vec{B} = \frac{1}{c}\left(4\pi\vec{j} + \frac{\partial \vec{E}}{\partial t}\right) \tag{61}$$

From the point of view of dynamics, the second pair of equations, (60)–(61) are the first order differential equations for quantities $\vec{E}$ and $\vec{B}$. Each of them requites one initial condition: $\vec{E}(x, t = 0)$ and $\vec{B}(x, t = 0)$. For any initial values of $\vec{E}$ and $\vec{B}$ we can find time derivatives such that the equations are satisfied.

But the first pair of equations (equations (58)–(59)) have no time derivative at all! They have to be valid at all moments, including the initial moment $t = 0$. If we choose an initial condition (58) and (59) may be satisfied or not. Moreover, if we define 2 components of $\vec{E}$ at $t = 0$, we can find the third one from (58). This same is true for (59) and the components of $\vec{B}$ at $t = 0$ We can say therefore that (58) and (59) constrain the space of possible initial conditions and, therefore, reduce the number of degrees of freedom. We see that we can specify only *four scalar functions* (arbitrary initial conditions for **2** components of $\vec{E}$ and for **2** components of $\vec{B}$ at each point of space). There is a way to interpret these 4 functions as **2** local degrees of freedom.[1]

It should be stressed that this is just a different example of the *system with constraints* that involves derivatives. In the examples of pendulum (Secs. 6.22.3) we solved this problem by introducing new variables that solve trivially the constraints (e.g. switching from Cartesian to polar coordinates). Let us see how this can be done in this case.

Let use the following property of vector analysis:

$$\vec{\nabla} \cdot [\vec{\nabla} \times \vec{A}] = 0 \tag{62}$$

that holds *for any vector field* $\vec{A}$.[2] We can therefore solve equation (59) by the following anzatz :

$$\vec{\nabla} \cdot \vec{B} = 0 \Longrightarrow \vec{B} = \vec{\nabla} \times \vec{A} \tag{63}$$

Then substituting this solution to the equations (60), we get

$$\vec{\nabla} \times \vec{E} = -\frac{1}{c}\frac{\partial(\vec{\nabla} \times \vec{A})}{\partial t} \tag{64}$$

[1] Recall that a mechanical degree of freedom is a pair: a coordinate and its time derivative that can be freely chosen at initial moment, *local degree of freedom* means that we have such data at every point.

[2] It is easy to demonstrate by explicit computation that $\vec{\nabla} \cdot (\vec{\nabla} \times \vec{A}) = 0$ for any vector field $\vec{A}$. Interestingly, the converse is also true: if divergence of a vector $\vec{Z}$ is equal to zero, this vector can be represented as a curl of some other vector $\vec{Z} = \vec{\nabla} \times \vec{W}$. This is known as *Poincaré lemma*.

Removing curl from both sides of (64) we find that

$$\vec{E} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} + \vec{U} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} - \vec{\nabla} A^0 \tag{65}$$

where $\vec{U}$ is a vector field such that $\vec{\nabla} \times \vec{U} = 0$ and , therefore, it can be represented as a gradient of a function A^0.[3]

As a result, we have expressed 6 functions $\vec{E}, \vec{B}$ via 4 functions $(A^0, \vec{A})$:

$$\begin{cases} \vec{E} = -\vec{\nabla} A^0 - \frac{1}{c}\frac{\partial \vec{A}}{\partial t} \\ \vec{B} = \vec{\nabla} \times \vec{A} \end{cases} \tag{66}$$

Is this relation unique? Using the property

$$\vec{\nabla} \times \vec{\nabla} f = 0 \tag{67}$$

for any function f we see that the magnetic field $\vec{B}$ does not change if we change:

$$\vec{A} \to \vec{A} + \vec{\nabla} f \tag{68}$$

However, $\vec{E}$ changes under the transformation (68). However, we along with the transformation (68) we do a transformation

$$A^0 \to A^0 - \frac{1}{c}\dot{f} \tag{69}$$

then both $\vec{E}, \vec{B}$ will not change, i.e. the pair $(A^0 - \frac{1}{c}\dot{f}, \vec{A} + \vec{\nabla} f)$ corresponds to the same $\vec{E}$, $\vec{B}$ for any scalar function $f(t, x)$. This property can be seen in the Lorentz-convariant notations. If we introduce a *4-vector of electromagnetic field* $A^\mu = (A^0, \vec{A})$ and recall that $\vec{E}, \vec{B}$ are related to the field strength tensor $F_{\mu\nu}$ via

$$\begin{aligned} E_i &= F_{0i} \\ B_i &= \frac{1}{2}\epsilon_{ijk} F^{jk} \end{aligned} \tag{70}$$

Then we can rewrite the relation (64) as

$$F_{\mu\nu} = \frac{\partial A_\mu}{\partial x^\nu} - \frac{\partial A_\nu}{\partial x^\mu} \tag{71}$$

It is then obvious that if we change

$$A_\mu \to A_\mu + \frac{\partial f}{\partial x^\mu} \tag{72}$$

(keep in mind that $x^0 = ct$) then the field-strength $F_{\mu\nu}$ (and $\vec{E}, \vec{B}$ as a consequence) does not change.

[3] Similarly to the comment above, if $\vec{\nabla} \times \vec{U} = 0$, $\vec{U}$ can be represented as a gradient of some scalar function.

The freedom to choose A^μ for given $\vec{E}$ and $\vec{B}$ is called *gauge invariance*. One needs additional conditions to uniquely relate A^μ and $\vec{E}, \vec{B}$. Such a condition is called *gauge condition* or *gauge fixing*.

In terms of the field strength $F_{\mu\nu}$ the Maxwell's equations have the form (Eqs. (58)(61)):

$$\frac{\partial F_{\mu\nu}}{\partial x^\nu} = j^\mu \tag{73}$$

where 4-vector $j^\mu = (c\rho, \vec{\jmath})$, and the pair of equations (59)–(60) gets rewritten as the so-called *Bianchi identity:*

$$\epsilon^{\mu\nu\lambda\rho}\partial_\nu F_{\lambda\rho} = 0 \tag{74}$$

Notice that the ansatz (71) solves (74) identically. If we plug the ansatz (71) into equation (73), we find

$$\left(\Delta - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) A_\mu + \partial_\mu\Big(\partial_\nu A^\nu\Big) = j^\mu \tag{75}$$

Let us choose a simple way to fix this freedom by imposing the condition

$$A^0 = 0 \tag{76}$$

Plugging this into the Maxwell's equations, we get

$$\left(\Delta - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) A_i + \partial_i\Big(\partial_j A^j\Big) = j^i \tag{77}$$

We are then with only three functions $\vec{A}$. However, this is not the final answer – our gauge freedom is not fully fixed. Indeed if we do gauge transformation (68) with a function f that does not depend on time, this will not change the condition $A^0 = 0$. Therefore, we expect another component of A_μ to be unphysical (possible to eliminate by a an additional gauge transformation. It is convenient to choose the function $f(x_i)$ such that

$$\partial_i \tilde{A}^i = \partial_i (A^i + \partial^i f) = 0 \tag{78}$$

This implicitly reduces the number of independent components of A_i (one of them can be found using the above equation if the other two are known) and simplifies the equations of motion. Let us stress however that if $A^0 = 0$, then

$$\begin{aligned} &\vec{E} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} \\ &\vec{\nabla}\cdot\vec{E} = -\frac{1}{c}\frac{\partial(\vec{\nabla}\cdot\vec{A})}{\partial t} \\ &(A^0 = 0) \end{aligned} \tag{79}$$

Therefore, we can impose both conditions $A^0 = 0$ and $\vec{\nabla}\cdot\vec{A} = 0$ *at the same time* only if $\vec{\nabla}\cdot\vec{E} = 0$ and, therefore, if $\rho = j^0 = 0$. Therefore, in the *empty space* Maxwell equations reduce to

$$\begin{cases} \Delta\vec{A} - \dfrac{1}{c^2}\dfrac{\partial^2\vec{A}}{\partial t^2} = 0 \\ \vec{\nabla}\cdot\vec{A} = 0 \end{cases} \tag{80}$$

The initial field A^μ had 4 degrees of freedom and to describe the solution of Eq. (75) one needed to specify 4 components of field and 4 time derivatives. However, A_μ are not directly observable, we need to know only they combinations that give us $\vec{E}$ and $\vec{B}$. On the other hand, the 6 components of $\vec{E}$ and $\vec{B}$ are directly observable, however are not independent. Solving explicitly the constraints on $\vec{E}$ and $\vec{B}$ with the help of A^μ we see that the system has only two degrees of freedom and we need to specify two independent functions and two independent time derivatives of these functions ("velocities") at the initial moment $t = 0$.

3.7 Quantum mechanics. Schrödinger equation

In quantum mechanics is described by the complex-valued wave function $\psi(x,t)$ and its evolution is described by the 1st order equation

$$i\hbar\frac{\partial\psi}{\partial t} = -\hbar^2\frac{\Delta\psi}{2m} + U\psi \tag{81}$$

The state of the system is $\psi(\vec{x})$ and there is no analog of velocity, as the evolution is defined by the the equation that is first order in time.

4 Lagrange mechanics

In all the examples we considered above the we have seen that the state of the system was defined as a minimal set of data required to determine the subsequent evolution of the system. For many systems this set consists of a configuration (the analog of the position of a particle) and its time derivative (the analog of velocity), however there are also deviations from this picture defined by the dynamics of the system (its equations of motion). The nature of configuration may be different: this can be a discret set of variables or a number of continuous functions (e.g. vector field). More over, the same system can be rewritten in different coordinate systems (different parametrisations of configurations), depending for example on the external forces acting on this system.

In this Section we will try to find common feature of all the above examples and find a description of the evolution of a dynamical system that will have the same form for all these systems. As the particles' dynamics is the basic example for mechanics, we will start with rewriting the dynamics of such a system of particles in an arbitrary system of coordinates and will then generalize the resulting description for a wider class of systems.

Let q_i, $i = 1\ldots n$, are new co-ordinates and f is some function of coordinates and time $f = f(q_1,\ldots,q_n,t)$.

We consider a system of particles with Cartesian co-ordinates $\vec{r}_1,\ldots,\vec{r}_N$ and we want to change them to independent generalized co-ordinates $q_1,\ldots,q_n$, $n \le 3N$. The relations between old and new co-ordinates are:

$$\vec{r}_i = \vec{r}_i(q_1,\ldots,q_n,t) \tag{82}$$

For each particle we have:

$$m_i \frac{d\vec{v}_i}{dt} = \vec{F}_i \tag{83}$$

4.1 Transformation of 2nd Newton's law under the change of variables (82)

Let we have a particle in the 3-dimensional oscillator potential. The equations of motion are:

$$m\ddot{\vec{r}} = -k\vec{r} \tag{84}$$

Let we make a general co-ordinate change:

$$x_i = f_i(q_1, q_2, q_3, t) \tag{85}$$

$$\dot{x}_i = \sum_j \frac{\partial f_i}{\partial q_j}\frac{\partial q_j}{\partial t} + \frac{\partial f_i}{\partial t} \tag{86}$$

$$\ddot{x}_i = \sum_{j,k} \frac{\partial^2 f_i}{\partial q_j \partial q_k}\frac{\partial q_j}{\partial t}\frac{\partial q_k}{\partial t} + \sum_j \frac{\partial f_i}{\partial q_j}\frac{\partial^2 q_j}{\partial t^2} + \sum_j \frac{\partial^2 f_i}{\partial t \partial q_j}\frac{\partial q_j}{\partial t} + \frac{\partial^2 f_i}{\partial t^2} \tag{87}$$

and equation of motion will be in form:

$$\begin{aligned} m\left(\frac{\partial f_i}{\partial q_1}\frac{\partial^2 q_1}{\partial t^2} + \frac{\partial f_i}{\partial q_2}\frac{\partial^2 q_2}{\partial t^2} + \frac{\partial f_i}{\partial q_3}\frac{\partial^2 q_3}{\partial t^2}\right) = \\ - kf_i - m\left(\sum_{j,k} \frac{\partial^2 f_i}{\partial q_j \partial q_k}\frac{\partial q_j}{\partial t}\frac{\partial q_k}{\partial t} + \sum_j \frac{\partial^2 f_i}{\partial t \partial q_j}\frac{\partial q_j}{\partial t} + \frac{\partial^2 f_i}{\partial t^2}\right) \end{aligned} \tag{88}$$

For example, for Galilei's transformation $x_i = q_i + v_i t$ equations don't become too difficult:

$$m\ddot{q}_i = -kq_i - kv_i t \tag{89}$$

But for rotating co-ordinate system with frequency ω (see 5.2 of [1]):

$$\begin{cases} x = x' \cos\omega t + y' \sin\omega t \\ y = -x' \sin\omega t + y' \cos\omega t \\ z = z' \end{cases} \tag{90}$$

we get:

$$\begin{cases} m\left(\ddot{x}' \cos\omega t + \ddot{y}' \sin\omega t\right) = (m\omega^2 - k)(x' \cos\omega t + y' \sin\omega t) - m\omega(-\dot{x}' \sin\omega t + \dot{y}' \cos\omega t) \\ m\left(-\ddot{x}' \sin\omega t + \ddot{y}' \cos\omega t\right) = (m\omega^2 - k)(-x' \sin\omega t + y' \cos\omega t) + m\omega(\dot{x}' \cos\omega t + \dot{y}' \sin\omega t) \\ m\ddot{z}' = -kz' \end{cases} \tag{91}$$

4.2 Covariant form

Let's multiply every of this equations by $\partial\vec{r}_i/\partial q_j$ and add results:

$$\sum_{i=1}^{N} m_i \frac{d\vec{v}_i}{dt} \cdot \frac{\partial \vec{r}_i}{\partial q_j} = \sum_{i=1}^{N} \vec{F}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} \tag{92}$$

As j changes from 1 to n we have n such equations.

$$\sum_{i=1}^{N} m_i \frac{d\vec{v}_i}{dt} \cdot \frac{\partial \vec{r}_i}{\partial q_j} = \frac{d}{dt}\sum_{i=1}^{N} m_i \vec{v}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} - \sum_{i=1}^{N} m_i \vec{v}_i \cdot \frac{d}{dt}\frac{\partial \vec{r}_i}{\partial q_j} \tag{93}$$

We can further rewrite this expression using two identities

$$\frac{\partial \vec{v}_i}{\partial \dot{q}_j} = \frac{\partial \vec{r}_i}{\partial q_j} \tag{94}$$

and

$$\frac{\partial \vec{v}_i}{\partial q_k} = \frac{d}{dt}\frac{\partial \vec{r}_i}{\partial q_k} \tag{95}$$

we have

$$\frac{d}{dt}\sum_{i=1}^{N} m_i \vec{v}_i \cdot \frac{\partial \vec{v}_i}{\partial \dot{q}_j} - \sum_{i=1}^{N} m_i \vec{v}_i \cdot \frac{\partial \vec{v}_i}{\partial q_j} = Q_j \tag{96}$$

where we have introduced notation Q_j (called *generalized forces*):

$$Q_j \equiv \sum_{i=1}^{N} \vec{F}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} \tag{97}$$

The meaning of these functions is work under infinitesimal displacement δq_j. Indeed, the work A is given by

$$A = \sum_i \vec{F}_i \delta \vec{r}_i = \sum_j \left(\sum_i \vec{F}_i \frac{\partial \vec{r}_i}{\partial q_j} \right) \delta q_j = \sum_j Q_J \delta q_j \tag{98}$$

Let's rewrite left part of (96) as

$$\frac{d}{dt}\frac{\partial}{\partial \dot{q}_j} \underbrace{\sum_{i=1}^{N} \frac{m_i v_i^2}{2}}_{\equiv T} - \frac{\partial}{\partial q_j} \sum_{i=1}^{N} \frac{m_i v_i^2}{2} = Q_j \tag{99}$$

By definition, the sum in left hand side is a kinetic energy of the system T. So, finally:

$$\boxed{\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = Q_j} \tag{100}$$

These n equations are called Lagrange equations. Q_i are called generalized forces. The main profit of this approach is that Lagrange equations are covariant — they doesn't change form after change of co-ordinate system.

Lagrange equations contains $n+1$ functions — n generalized forces Q_i and kinetic energy T. But they can be simplified for the case of potential forces $\vec{F}_i$. The potentiality of force means, that there exists scalar function of co-ordinates and time $V_i(\vec{r},t)$, that $\vec{F}_i = -\vec{\nabla} V_i$. So, if all forces are potential ones:

$$Q_j = -\sum_{i=1}^{N} \left(\frac{\partial \vec{V}_i}{\partial x_i}\frac{\partial x_i}{\partial q_j} + \frac{\partial \vec{V}_i}{\partial y_i}\frac{\partial y_i}{\partial q_j} + \frac{\partial \vec{V}_i}{\partial z_i}\frac{\partial z_i}{\partial q_j} \right) = -\frac{\partial U}{\partial q_j} \tag{101}$$

where $U = \sum_i V_i$ — potential energy of the system. Therefore:

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = -\frac{\partial U}{\partial q_j} \tag{102}$$

and, using fact that $\partial U/\partial \dot{q}_j = 0$, we finally have:

$$\boxed{\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0} \tag{103}$$

where

$$\boxed{L = T - U} \tag{104}$$

Function $L(q,\dot{q},t)$ is called *Lagrange function*, or just *Lagrangian* of the system (see 10.4 of [1]).

5 The principle of least action

It turns out that Lagrange equations of motion (103) can be obtained from conception called the *principle of least action* or Hamilton's principle (see 10.1 of [1]).

Namely, we consider a system whose state is characterized by a set of discrete or continuous variables q_j. Its dynamics is determined via the function $L(q_1, q_2, \dots, q_s, \dot{q}_1, \dot{q}_2, \dots, \dot{q}_n, t)$, or briefly $L(q,\dot{q},t)$.

Let the system occupy, at the instants t_1 and t_2, positions defined by two sets of values of the co-ordinates, $q^{(1)}$ and $q^{(2)}$. Then the condition is that the system moves between these positions in such a way that the integral

$$S = \int_{t_1}^{t_2} L(q,\dot{q},t)dt \tag{105}$$

takes the least possible value[4]. The function L is Lagrangian of the system concerned, and the integral (105) is called the action.

The fact that the Lagrangian contains only q and $\dot{q}$, but not the higher derivatives $\ddot{q}$, $\dddot{q}$, etc., expresses the result already mentioned, that the mechanical state of the system is completely defined when the co-ordinates and velocities are given.

[4]It should be mentioned that this formulation of the principle of least action is not always valid for the entire path of the system but only for any sufficiently short segment of the path. The integral (105) for entire path must be extremum, but not necessarily a minimum. This fact, however, is of no importance as regards the derivation of the equations of motion, since only the extremum condition is used.

Let us now derive the differential equations which solve the problem of minimizing the integral (105). For simplicity, we shall at first assume that the system has only one degree of freedom, so that only one function $q(t)$ has to be determined.

Let $q = q(t)$ be the function for which S is a minimum. This means that S is increased when $q(t)$ is replaced by any function of the form

$$q(t) + \delta q(t) \tag{106}$$

where $\delta q(t)$ is a function which is small everywhere in the interval of time from t_1 to t_2; it has no connection to virtual displacement from previous section and is called a variation of the function $q(t)$. Since, for $t = t_1$ and for $t = t_2$, all the functions (106) must take the values $q^{(1)}$ and $q^{(2)}$ respectively, it follows that

$$\delta q(t_1) = \delta q(t_2) = 0 \tag{107}$$

The change in S when q is replaced by $q + \delta q$ is

$$\int_{t_1}^{t_2} L(q + \delta q, \dot{q} + \delta\dot{q}, t)dt - \int_{t_1}^{t_2} L(q, \dot{q}, t)dt$$

When this difference is expanded in powers of δq and $\delta\dot{q}$ in the integrand, the leading terms are of the first order. The necessary condition for S to have a minimum is that these terms (called the first variation, or simply the variation, of the integral) should be zero. Thus the principle of least action may be written in the form

$$\delta S = \delta \int_{t_1}^{t_2} L(q, \dot{q}, t)dt = 0 \tag{108}$$

or, effecting the variation,

$$\int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta\dot{q} \right) dt = 0$$

Since $\delta q = d\delta q/dt$, we obtain, on integrating the second term by parts,

$$\delta S = \frac{\partial L}{\partial \dot{q}}\delta q\Big|_{t_1}^{t_2} + \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} \right) \delta q dt \tag{109}$$

The conditions (107) show that the integrated term in 109) is zero. There remains an integral which must vanish for all values of δq. This can be so only if the integrand is zero identically. Thus we have

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0$$

When the system has more than one degree of freedom, the n different functions $q_i(t)$ must be varied independently in the principle of least action. We then evidently obtains equations of the form

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \tag{110}$$

We see, that principle of the least action gives the right equations of motion for the system of particles. In fact, the scope of it's application is much more wider and lies from elementary particle physics up to gravity and cosmology. So we will postulate it as a fundamental principle of mechanics. Writing Lagrangian is equivalent to determining all properties of the physical system.

One further general remark should be made. Let us consider two functions $L'(q,\dot{q},t)$ and $L(q,\dot{q},t)$, differing by the total derivative with respect to time of some function $f(q,t)$ of co-ordinates and time:

$$L'(q,\dot{q},t) = L(q,\dot{q},t) + \frac{df(q,t)}{dt} \tag{111}$$

The integrals (105) calculated from these two functions are such that

$$S' = \int_{t_1}^{t_2} L'(q,\dot{q},t)dt = \int_{t_1}^{t_2} L(q,\dot{q},t)dt + \int_{t_1}^{t_2} \frac{df(q,t)}{dt}dt = S + f(q^{(2)},t_2) - f(q^{(1)},t_1)$$

i.e. they differ by a quantity which gives zero on variation, so that the conditions $\delta S' = 0$ and $\delta S = 0$ are equivalent, and the form of the equations of motion is unchanged. Thus the Lagrangian is defined only to within an additive total time derivative of any function of co-ordinates and time.

Note that we have derived the Lagrangian equations for arbitrary generalized coordinates q_i and therefore their form will remain *the same* under arbitrary change of coordinates.

From now on we will describe the dynamical system via its *Lagrangian*, introducing generalized coordinates, as dictated by symmetries of the problem.

6 Examples of Lagrangians

6.1 Lagrangian of a free particle and Gallilean invariance

Consider the Lagrangian of a free particle:

$$L = \frac{1}{2}m\vec{v}^2 \tag{112}$$

Let us perform a Gallilean transformation $\vec{v}' = \vec{v} + \vec{V}$. Naively, the form of the Lagrangian 112 changes:

$$L(v') = \frac{1}{2}m\Big(\vec{v}^2 + \vec{V}^2 + 2(\vec{v}\cdot\vec{V})\Big) \tag{113}$$

Notice, however that the additional terms form a total derivative:

$$\vec{V}^2 + 2(\vec{v}\cdot\vec{V}) = \frac{d}{dt}\Big(\vec{V}^2 t + 2(\vec{r}\cdot\vec{V})\Big) \tag{114}$$

and therefore, as discussed around Eq. (111) does not change the extremum of the action.

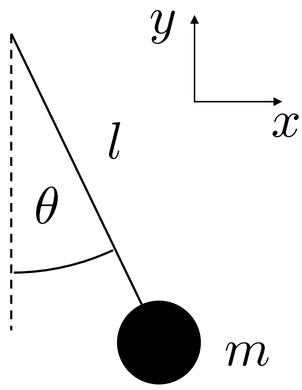

Figure 8: The pendulum.

6.2 Pendulum

So far it looks like Hamilton's principle is a very complicated way of obtaining the equation of motion, that one can write down immediately. For more complicated systems that contain certain constraints, however, such a framework is extremely useful. To give an example: Consider a pendulum, a mass m attached to a massless rod of length l that is suspended from a pivot at position $(x, y) = (0, 0)$ around which it can swing freely. The potential of the mass in the gravitational field is given by mgy. The Lagrange function of the pendulum is thus given by

$$L(x, y, \dot{x}, \dot{y}) = \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2\right) - mgy. \tag{115}$$

The Euler-Lagrange equations for the x- and y-coordinates lead to two equations of motion, $\ddot{x} = 0$ and $\ddot{y} = -g$.

Unfortunately these equations are completely wrong. What we found are the equations of motion of a free particle in 2 dimensions in a gravitational field. Solutions are e.g. trajectories of rain drops or of cannon balls but certainly not the motion of a pendulum. What went wrong? We forgot to take into account the presence of the rod that imposes a constraint, namely that $x^2 + y^2 = l^2$. A better approach would be to use a coordinate system that accounts automatically for this constraint, namely to describe the state of the pendulum by the angle $\theta(t)$ between the pendulum and the y-direction, see Fig. 8. But how does the equation of motion look in terms of this angle?

Here comes into play a great advantage of Hamilton's principle: it is independent of the coordinate system that one chooses. Suppose one goes from one coordinate system x_1, x_2,..., x_N to another coordinate system q_1, q_2,..., q_f via the transformations $\mathbf{q} = \mathbf{q}(\mathbf{x})$ and $\mathbf{x} = \mathbf{x}(\mathbf{q})$. The trajectory $\mathbf{x}(t)$ becomes then $\mathbf{q}(\mathbf{x}(t))$. The action functional can then be rewritten as

$$S[\mathbf{x}] = \int_{t_1}^{t_2} L(\mathbf{x}(t), \dot{\mathbf{x}}(t))\, dt = \int_{t_1}^{t_2} L\left(\mathbf{x}(\mathbf{q}(t)), \sum_{i=1}^{f} \frac{\partial \mathbf{x}(\mathbf{q}(t))}{\partial q_i}\dot{q}_i\right) dt. \tag{116}$$

The rhs of Eq. 116 is again of the form

$$S[\mathbf{q}] = \int_{t_1}^{t_2} \tilde{L}(\mathbf{q}(t), \dot{\mathbf{q}}(t))\, dt \tag{117}$$

with a new Lagrangian $\tilde{L}$. Also here Hamilton's principle must hold, i.e., the dynamic evolution of the system follows from the Euler-Lagrange equations

$$\frac{\partial \tilde{L}}{\partial q_i} - \frac{d}{dt}\frac{\partial \tilde{L}}{\partial \dot{q}_i} = 0 \tag{118}$$

for $i = 1, ..., f$.

If we have a system with constraints we can sometimes introduce coordinates that automatically fulfill those constraints. The equations of motion are then simply given by the Euler-Lagrange equations in these coordinates. Let us go back to the pendulum. We describe now the configuration of the pendulum by the angle $\theta(t)$, see Fig. 8. In terms of this angle the kinetic energy of the pendulum is given by $ml^2\dot{\theta}^2/2$ and the potential energy by $-mlg\cos\theta$. This leads to the following Lagrange function:

$$L\left(\theta, \dot{\theta}\right) = \frac{ml^2}{2}\dot{\theta}^2 + mgl\cos\theta. \tag{119}$$

The corresponding Euler-Lagrange equation is given by

$$\ddot{\theta}(t) = -\frac{g}{l}\sin\theta(t), \tag{120}$$

which is indeed the equation of motion of the pendulum.

6.3 Pendulum on a movable support

Consider a mass M that can move freely along a horizonal line without friction. Attached to the mass M is a pendulum of mass m via a massless connection of length l (Fig. 9). We calculate now the Lagrange equations for this system.

We first need to find a suitable coordinate system. The system has 2 degrees of freedom (you can find this number by subtracting the two constraints from the 4 degrees of freedom of the unconstrained masses). Practical coordinates are the position X of the mass M along the line and the angle θ between the pendulum and the direction of gravity. The position of the pendulum body is then given by

$$x = X + l\sin\theta \quad \text{and} \quad z = -l\cos\theta.$$

The kinetic energy is then given by:

$$T = \frac{1}{2}M\dot{X}^2 + \frac{1}{2}m\left(\dot{x}^2 + \dot{z}^2\right) = \frac{1}{2}M\dot{X}^2 + \frac{1}{2}m\left[\left(\dot{X} + l\dot{\theta}\cos\theta\right)^2 + \left(l\dot{\theta}\sin\theta\right)^2\right].$$

This simplifies to

$$T = \frac{1}{2}(m+M)\dot{X}^2 + \frac{1}{2}m\left[2l\dot{X}\dot{\theta}\cos\theta + \left(l\dot{\theta}\right)^2\right].$$

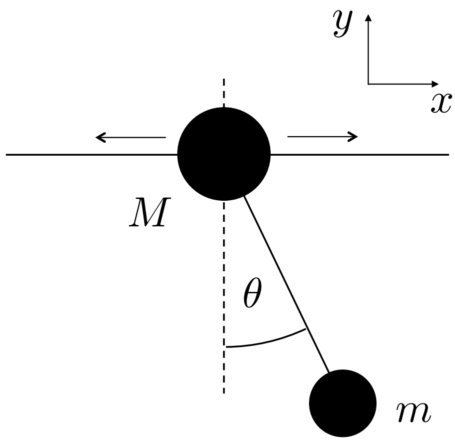

Figure 9: Example: a pendulum on a movable support.

The potential energy is given by

$$V = -mgl\cos\theta.$$

We can now obtain the equations of motions by taking derivatives of the Lagrangian with respect to the coordinates and to their time derivatives. This is done separately for the two coordinates. The Lagrange equation 118 for the coordinate X is given by:

$$\frac{d}{dt}\left[\frac{\partial\left(T-V\right)}{\partial\dot{X}}\right]-\frac{\partial\left(T-V\right)}{\partial X}=0,$$

leading to

$$\left(m+M\right)\ddot{X}+ml\frac{d}{dt}\left(\dot{\theta}\cos\theta\right)=0$$

or

$$\left(m+M\right)\ddot{X}=ml\left(\dot{\theta}^2\sin\theta-\ddot{\theta}\cos\theta\right).$$

Note that the partial derivative with respect to $\dot{X}$ is only taken on those places where this variable occurs but that the derivative with respect to the time t acts on all variables including θ en $\dot{\theta}$. Another point to note here is that quantity $\partial\left(T-V\right)/\partial\dot{X}$ is conserved (i.e. does not change with time). This follows always immediately if the Lagrangian does not depend on one of the coordinates (here X). You can check easily that this quantity is here the total momentum in the X-direction.

For the other coordinate, θ, we obtain:

$$\frac{d}{dt}\left[\frac{\partial\left(T-V\right)}{\partial\dot{\theta}}\right]-\frac{\partial\left(T-V\right)}{\partial\theta}=0,$$

leading to

$$ml\left(l\ddot{\theta}+\ddot{X}\cos\theta-\dot{X}\dot{\theta}\sin\theta\right)+ml\dot{X}\dot{\theta}\sin\theta+mgl\sin\theta=0$$

or

$$\ddot{\theta}+\frac{\ddot{X}}{l}\cos\theta+\frac{g}{l}\sin\theta=0.$$

This example shows how straightforward the equations of motion can be derived with the Lagrange formalism as compared to deriving them from Newton's formalism which involves force vectors.

6.4 Particle in a central force field

For a particle in a central field the motion takes place in a plane (see 6.1 of [1]). We choose polar coordinates. The velocity has a radial and a tangential component. The kinetic and the potential energies are given by

$$T = \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2\right) \quad \text{and} \quad V = V\left(r\right).$$

The Lagrange equation for r is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) = m\ddot{r} = \frac{\partial L}{\partial r} = mr\dot{\theta}^2 - \frac{dV}{dr}$$

and for θ:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) = m\frac{d}{dt}\left(r^2\dot{\theta}\right) = \frac{\partial L}{\partial \theta} = 0.$$

Here we find that $mr^2\dot{\theta}$, the *angular momentum*, is conserved as the Lagrangian does not depend on θ.

References

[1] Analytical Mechanics, G.R. Fowles and G.L. Cassiday 6th or 7th edition (Thomson Learning, inc., 1999), ISBN 9780534408138.