Classical Theoretical Physics I

Jörg Schmalian, Karlsruhe Institute of Technology

Winter Semester 2013/14

Contents

Preface

These lecture notes summarize the main content of the course Physics Classical Theoretical Physics I, taught at the Karlsruhe Institute of Technology during the Winter Semester 2013/14.They are partly based on the course Advanced Mechanics given at Iowa State University during the Fall of 2010.

Chapter 1

The equation of motion

1.1 Dimensional analysis

Length, mass, and time are three fundamentally different quantities which are measured in three completely independent units. It, therefore, makes no sense for a prospective law of physics to express an equality between (say) a length and a mass. In other words, the example law

$$m = l \tag{1.1}$$

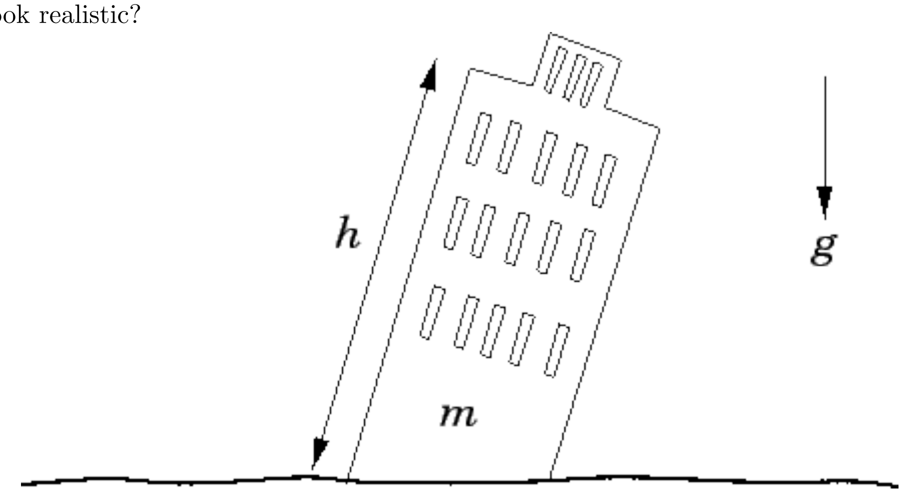
where m is a mass and l is a length, cannot possibly be correct as a law of physics. One easy way of seeing that Eq.1.1 is invalid, is to note that this equation is dependent on the adopted system of units. We use a nomenclature where $[L]$ indicates that a given quantity has dimension length. Equally $[T]$ for time and $[M]$ for mass etc. This implies for example that velocity has dimension $[L] / [T]$. A given system of units is characterized by an elementary set of quantities $[A_i]$ that suffice to determine all dimensions.

Physicists hold very strongly to the assumption that the laws of physics possess objective reality: in other words, the laws of physics are the same for all observers. One immediate rather trivial consequence of this assumption is that a law of physics must take the same form in all possible systems of units that a prospective observer might choose to employ. The only way in which this can be the case is if all laws of physics are dimensionally consistent: i.e., the quantities on the left- and right-hand sides of the equality sign in any given law of physics must have the same dimensions (i.e., the same combinations of length, mass, and time). A dimensionally consistent equation naturally takes the same form in all possible systems of units, since the same conversion factors are applied to both sides of the equation when transforming from one system to another. As an example, let us consider what is probably the most famous equation in physics:

$$E = mc^2. \tag{1.2}$$

Here, E is the energy of a body, m is its mass, and c is the velocity of light in vacuum. The dimensions of energy are $[M]\left[L^2\right]/\left[T^2\right]$, and the dimensions of velocity are $[L]/[T]$. Hence, the dimensions of the left-hand side are equal to the dimensions of the right-hand side. It follows that Eq.1.2 is indeed dimensionally consistent. It holds in any sensible set of units. Had Einstein proposed $E = m^2c$, or $E = mc^5$, then his error would have been immediately apparent to other physicists, since these prospective laws are not dimensionally consistent. Of course, this does not mean that it is impossible that E is proportional to mass and c^5. There could, at least in principle, exist a relationship $E = Gmc^5$, where the new dimensionful "coupling constant" G occurred. From our considerations we even know that the dimension of G is $[T]^3/[L]^3$. What we do know is that Eq.1.2 represents the only simple, dimensionally consistent way of combining an energy, a mass, and the velocity of light in a law of physics if we ignore overall pre-factors ($E = \sqrt{2}mc^2$ would also be dimensionally consistent as $\sqrt{2}$ is dimensionless). The last comment leads naturally to the subject of dimensional analysis: i.e., the use of the idea of dimensional consistency to guess the forms of simple laws of physics. It should be noted that dimensional analysis is of fairly limited applicability, and is a poor substitute for analysis employing the actual laws of physics; nevertheless, it is occasionally useful.

Suppose that a special effects studio wants to film a scene in which the Leaning Tower of Pisa topples to the ground. In order to achieve this, the studio might make a scale model of the tower, which is (say) 1m tall, and then film the model falling over. The only problem is that the resulting footage would look completely unrealistic, because the model tower would fall over too quickly. The studio could easily fix this problem by slowing the film down. The question is by what factor should the film be slowed down in order to make it look realistic?



Let us pretend for the moment that we do not know how to apply the laws of physics to the problem of a tower falling over. However, we can, at least, make some educated guesses as to what factors the time t_f required for this process depends on. In fact, it seems reasonable to suppose that t_f depends principally on the mass of the tower, m, the height of the tower, h, and the acceleration due to gravity, g, see figure. In other words,

$$t_f = C m^x h^y g^z \tag{1.3}$$

where C is a dimensionless constant, and x, y, and z are unknown exponents. The exponents can be determined by the requirement that the above equation be dimensionally consistent. The dimensions of an acceleration are $[L] / [T^2]$. Hence, equating the dimensions of both sides of Eq.1.3 we obtain

$$[T] = [M]^x [L]^y \frac{[L]^z}{[T^2]^z} \tag{1.4}$$

We can now compare the exponents of $[T]$, $[M]$, and $[L]$ on either side of the above expression: these exponents must all match in order for Eq.1.3 to be dimensionally consistent. Thus,

$$\begin{aligned} 0 &= x \\ 0 &= y + z \\ 1 &= -2z \end{aligned} \tag{1.5}$$

It immediately follows that $z = -\frac{1}{2}$, $y = \frac{1}{2}$, and $x = 0$. Hence,

$$t_f = C\sqrt{\frac{h}{g}} \tag{1.6}$$

Now, the actual tower of Pisa is approximately 100m tall and g is the same for both the real and the model tower. It follows that $t_f \propto \sqrt{h}$ then the 1m high model tower falls over a factor of $\sqrt{100/1} = 10$ times faster than the real tower. Thus, the film must be slowed down by a factor 10 in order to make it look more realistic.

1.2 Newton's laws

All around us we observe that all moving objects will come eventually to rest, unless we apply a force to them. We need to keep pedaling if we want to keep a bicycle going with constant speed, we need to have our engine running if we want to keep driving with a speed of 55 miles/hour. In all these cases, friction will ultimately stop any moving object, unless the friction force is canceled by the force supplied by our legs, our engine, etc. If we reduce friction, the moving object will take longer to slow down, and the force needed to overcome the friction force will be less. In the limit of no friction, our object will keep moving with a constant velocity, and no force need to be applied. This conclusion is summarized in Newton's first law:

- Consider a body on which no net force acts. If the body is at rest, it will remain at rest. If the body is moving with constant velocity, it will continue to do so.

Newton's first law is really a statement about reference frames in that it defines the kinds of reference frames in which the laws of Newtonian mechanics hold. Reference frames in which Newton's first law applies are called inertial reference frames.

One way to test whether a reference frame is an inertial reference frame, is to put a test body at rest and arrange things so that no net force acts in it. If the reference frame is an inertial frame, the body will remain at rest; if the body does not remain at rest, the reference frame is not an inertial frame. If you put a bowling ball at rest on a merry-go-round, no identifiable forces act on the ball, but it does not remain at rest. Rotating reference frames are not inertial reference frames. Strictly speaking, the earth is therefore also not an inertial frame, however, only if we consider large scale motion such as wind and ocean current do we need to take into account the non inertial character of the rotating earth.

To proceed, we use the concept of a mass point, i.e. ignore the internal structure of an object, such that its state is determined solely by its position $\mathbf{r} = (x_1, x_2, \ldots, x_d)$ in d-dimensional space. In case of N particles we have $\mathbf{r}_i$ distinct positions, corresponding to $d \times N$ degrees of freedom. Experience tells us that the knowledge of the position $\mathbf{r}$ of a mass point does not allow us to predict its future evolution. To achieve this we also need to know the velocity

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \dot{\mathbf{r}}. \tag{1.7}$$

It is conceivable that further derivatives are necessary to properly predict the temporal evolution of a particle. Experiment is however consistent with the fact that $(\mathbf{r}_i, \mathbf{v}_i)$ are sufficient to predict the future of a mechanical system. The statement of Newton's first law can now be reformulated that for a mass point on which no net force acts obeys the differential equation:

$$\frac{d\mathbf{v}}{dt} = 0 \tag{1.8}$$

To be precise, those are three differential equations for the three components of the velocity vector (d equations in d dimensions):

$$\begin{aligned} \frac{dv_x}{dt} &= 0, \\ \frac{dv_y}{dt} &= 0, \\ \frac{dv_z}{dt} &= 0. \end{aligned} \tag{1.9}$$

The solution of this differential equation is straightforward ($\alpha = x, y, z$):

$$v_\alpha(t) = v_\alpha^0 \tag{1.10}$$

i.e. the velocity components are constant and maintain their initial values. For the vector of the velocity follows accordingly

$$\mathbf{v}(t) = \mathbf{v}^0, \tag{1.11}$$

where $\mathbf{v}^0 = \left(v_x^0, v_y^0, v_z^0\right)$.

Applying a force on an object will change its velocity. The easiest way to express this is to assume that the change in the velocity (i.e. the acceleration) is proportional to the force $\mathbf{F}$.

$$\frac{d\mathbf{v}}{dt} \propto \mathbf{F}. \tag{1.12}$$

We could alternatively try to assume $\frac{d\mathbf{v}}{dt} \propto \mathbf{F}\left(\mathbf{F}\cdot\mathbf{F}\right)$ or similar. In each case we need to solve those equations and compare with experiment to decide which one is correct. It turns out that the simplest form, Eq.1.12, gives an excellent account of observations.

If we now exert the same force on several objects with different mass, we will observe different accelerations. For example, one can throw a baseball significantly further (and faster) than a ball of similar size made of lead. The unit of force is the Newton (N), and a force of 1 N is defined as the force that when applied to an object with a mass of 1 kg, produces an acceleration of 1 m/s^2. If we apply a force equal to 2 N, the corresponding acceleration is 2 m/s^2. This observation helps determine the proportionality factor in Eq.1.12 as

$$\frac{d\mathbf{v}}{dt} = \frac{1}{m}\mathbf{F} \tag{1.13}$$

Using Eq.1.7 for the velocity finally yields the well known form of Newton's second law (or simply the equation of motion of classical mechanics):

$$m\frac{d^2\mathbf{r}}{dt^2} = \mathbf{F}\left(\mathbf{r}\right). \tag{1.14}$$

We conclude that a force acting on an object produces an acceleration. The direction of the acceleration is the same as the direction of the force applied. Newton's second law includes as special case Newton's first law if $\mathbf{F} = \mathbf{0}$.

Let us solve Newton's law in the special case of a constant (i.e. position independent) force $\mathbf{F_0}$. It is now easiest to chose a coordinate system such that $\mathbf{F_0} = \left(F_0, 0, 0\right)$. Our three differential equations are obviously

$$\begin{aligned} m\frac{d^2x}{dt^2} &= F_0, \\ m\frac{d^2y}{dt^2} &= 0, \\ m\frac{d^2z}{dt^2} &= 0. \end{aligned} \tag{1.15}$$

The solution for the y- and z-components is straightforward. Lets do this for the y-component. First we consider $v_y = \frac{dy}{dt}$, i.e. the velocity in the y-direction. It follows

$$\frac{dv_y}{dt} = 0. \tag{1.16}$$

We integrate this equation on both sides. For the left hand side follows

$$\int_0^t \frac{dv_y}{dt'}dt' = v_y\left(t\right) - v_y\left(0\right). \tag{1.17}$$

There is of course no need to start integrating at time $t = 0$. However, without restriction we simply call the initial time $t = 0$. On the right hand side follows

$$\int_0^t 0dt' = 0 \int_0^t dt' = 0. \tag{1.18}$$

Equating both sides yields

$$v_y(t) = v_y(0), \tag{1.19}$$

which is our earlier result of Eq.1.11. In terms of the coordinate this result is

$$\frac{dy}{dt} = v_y(0). \tag{1.20}$$

Once again, we integrate both sides. The left hand side now yields $\int_0^t \frac{dy}{dt'} dt' = y(t) - y(0)$ while for the right hand side follows $\int_0^t v_y(0)\, dt' = v_y(0)\, t$. We finally obtain

$$y(t) = y(0) + v_y(0)\, t. \tag{1.21}$$

This is the motion of a particle without force. The y-component behaves as a particle in free space. The same is obviously true for the z-component, i.e.

$$z(t) = z(0) + v_z(0)\, t. \tag{1.22}$$

Let us now look at the x-coordinate. We proceed along the same lines and integrate the differential equation for the velocity

$$\frac{dv_x}{dt} = \frac{F_0}{m} \tag{1.23}$$

on both sides, which gives

$$v_x(t) - v_x(0) = \frac{F_0}{m} t. \tag{1.24}$$

This is a first order differential equation for the coordinate

$$\frac{dx}{dt} = v_x(0) + \frac{F_0}{m} t. \tag{1.25}$$

The last step is now to integrate one more time and we obtain

$$x(t) = x(0) + v_x(0)\, t + \frac{F_0}{2m} t^2. \tag{1.26}$$

The three dimensional trajectory of our particle can now we written as

$$\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = \begin{pmatrix} x(0) + v_x(0)\, t + \frac{F_0}{2m} t^2 \\ y(0) + v_y(0)\, t \\ z(0) + v_z(0)\, t \end{pmatrix} \tag{1.27}$$

a result that can be written more compactly in vectorial notation

$$\mathbf{r}(t) = \mathbf{r}(0) + \mathbf{v}(0)\,t + \mathbf{F}_0 \frac{t^2}{2m}. \tag{1.28}$$

The key result is that the acceleration takes place only in the direction of the applied force while the motion in the other directions is unaffected by the force. Of course, not all problems in mechanics are as simple as this one and the solution of the differential equations is in general a lot more complicated.

1.2.1 Mathematical tools 1: vectors, derivatives, and polar coordinates

Let us briefly summarize some of the main mathematical tools that we used in the previous section. First, we used the concept of a vector. In order to represent a vector, we introduce a coordinate system. The most straightforward one is the cartesian coordinate system. In three dimensions we have three mutually orthogonal unit vectors $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$. Unit vectors means that they have a length one, i.e., it holds $\mathbf{e}_i \cdot \mathbf{e}_i = 1$ for $i = x, y, z$, while for $i \neq j$ holds $\mathbf{e}_i \cdot \mathbf{e}_j = 0$. The dot between two vectors refers to the scalar product, i.e. we have with an arbitrary vector $\mathbf{a} = (a_x, a_y, a_z)$ and similar for a vector $\mathbf{b}$ that

$$\mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y + a_z b_z = \sum_{i=1}^{3} a_i b_i. \tag{1.29}$$

It also holds that

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}|\,|\mathbf{b}| \cos\varphi \tag{1.30}$$

where φ is the angle between the two vectors. Thus, using elementary trigonometry, it follows that the length of a vector is given by

$$a = |\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \sqrt{\mathbf{a}^2}. \tag{1.31}$$

The properties of cartesian unit vectors can now be summarized as

$$\mathbf{e}_{\boldsymbol{i}} \cdot \mathbf{e}_{\boldsymbol{j}} = \boldsymbol{\delta}_{\boldsymbol{ij}}, \tag{1.32}$$

where we introduced the Kronecker symbol:

$$\delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}. \tag{1.33}$$

The position of a mass point can then be expressed as

$$\begin{aligned} \mathbf{r}(t) &= x(t)\,\mathbf{e}_x + y(t)\,\mathbf{e}_y + z(t)\,\mathbf{e}_z \\ &= \sum_{i=1}^{3} x_i(t)\,\mathbf{e}_i. \end{aligned} \tag{1.34}$$

We also used the concept of the velocity as first derivative, i.e.

$$\begin{aligned} \mathbf{v}(t) &= \frac{d\mathbf{r}(t)}{dt} = \lim_{\delta\to 0}\frac{\mathbf{r}(t+\delta)-\mathbf{r}(t)}{\delta} \\ &= \sum_{i=1}^{3}\frac{dx_i(t)}{dt}\mathbf{e}_i. \end{aligned} \tag{1.35}$$

The direction of the velocity is parallel to the tangential vector of the orbit. Its magnitude

$$v = |\mathbf{v}| = \sqrt{\left(\frac{dx(t)}{dt}\right)^2+\left(\frac{dy(t)}{dt}\right)^2+\left(\frac{dz(t)}{dt}\right)^2} \tag{1.36}$$

is the change of the arc-length $ds = \sqrt{dx^2+dy^2+dz^2}$ per infinitesimal time change dt, i.e.

$$v = \frac{ds}{dt}. \tag{1.37}$$

The acceleration is then the time derivative of the velocity, i.e.

$$\mathbf{a}(t) = \frac{d\mathbf{v}(t)}{dt} = \frac{d^2\mathbf{r}(t)}{dt^2}. \tag{1.38}$$

The cartesian coordinate system is clearly not the only way to describe a vector. As an example, we consider the motion of a particle on a circular orbit in two dimensions. This can be written as

$$\mathbf{r}(t) = R\cos(\omega t)\,\mathbf{e}_x + R\sin(\omega t)\,\mathbf{e}_y. \tag{1.39}$$

Here, ω is the angular frequency. It determines the time $T = \frac{2\pi}{\omega}$ it takes the particle to orbit once around the circle of radius R. The velocity is given as

$$\mathbf{v}(t) = \frac{d\mathbf{r}(t)}{dt} = -R\omega\sin(\omega t)\,\mathbf{e}_x + R\omega\cos(\omega t)\,\mathbf{e}_y. \tag{1.40}$$

A more efficient description can be achieved by introducing the two unit vectors

$$\begin{aligned} \mathbf{e}_r &= \cos(\omega t)\,\mathbf{e}_x + \sin(\omega t)\,\mathbf{e}_y, \\ \mathbf{e}_\varphi &= -\sin(\omega t)\,\mathbf{e}_x + \cos(\omega t)\,\mathbf{e}_y. \end{aligned} \tag{1.41}$$

First we need to show that those vectors are indeed unit vectors, i.e. that $\mathbf{e}_r\cdot\mathbf{e}_r = \mathbf{e}_\varphi\cdot\mathbf{e}_\varphi = 1$ and $\mathbf{e}_r\cdot\mathbf{e}_\varphi = 0$. This follows indeed, if one uses the orthogonality of the initial unit vectors and the relation $\cos^2(x)+\sin^2(x) = 1$. The orbit and velocity are now given as

$$\begin{aligned} \mathbf{r}(t) &= R\mathbf{e}_r, \\ \mathbf{v}(t) &= R\omega\mathbf{e}_\varphi. \end{aligned} \tag{1.42}$$

How would we now calculate the acceleration? It obviously holds

$$\mathbf{a}\left(t\right)=\frac{d\mathbf{v}\left(t\right)}{dt}=R\omega\frac{d\mathbf{e}_{\varphi}}{dt}. \tag{1.43}$$

The time dependence is now obviously a consequence of the time dependent coordinate system. It follows from the definition of $\mathbf{e}_{\varphi}$ that

$$\frac{d\mathbf{e}_{\varphi}}{dt}=-\omega\cos\left(\omega t\right)\mathbf{e}_{x}-\omega\sin\left(\omega t\right)\mathbf{e}_{y}=-\omega\mathbf{e}_{r}. \tag{1.44}$$

It follows for the acceleration

$$\mathbf{a}\left(t\right)=-\omega^{2}R\mathbf{e}_{r}. \tag{1.45}$$

A particle that moves on a circular orbit is therefore constantly accelerated. In other words, a particle can only move on a circular orbit if a force acts upon it such that

$$\mathbf{F}=-f_{r}\mathbf{e}_{r}, \tag{1.46}$$

with radial component of the force fulfilling $f_r = m\omega^2 R$. This is a force pointing towards the center of the circle. One could for example realize such a force by pulling on a rope. For given pulling force magnitude f_r, length of the rope R and mass of the orbiting object m follows the frequency of the circular motion as

$$\omega=\sqrt{\frac{f_{r}}{mR}}. \tag{1.47}$$

1.3 Energy conservation

One of the central concepts of physics is energy. The reason why we pay so much attention to this quantity is that it does not change under very general conditions. Energy is conserved. This is not a concept that has to be postulated in addition to the equation of motion, it is a consequence of it. In what follows we will first analyze a one dimensional motion. In the next step we will generalize this to higher dimensions.

1.3.1 Energy conservation in one dimension

We start from our equation of motion in one dimension, i.e.

$$m\frac{d^{2}x}{dt^{2}}=F\left(x,t\right). \tag{1.48}$$

Here F is the force acting in this direction. Next we observe the following result that is obtained by simply applying the rules of differentiation

$$\frac{d}{dt}\left(\frac{dx}{dt}\right)^{2}=2\frac{dx}{dt}\frac{d}{dt}\left(\frac{dx}{dt}\right)=2\frac{dx}{dt}\frac{d^{2}x}{dt^{2}}. \tag{1.49}$$

This suggests to multiply the equation of motion by $\frac{dx}{dt}$ yielding

$$\frac{m}{2}\frac{d}{dt}\left(\frac{dx}{dt}\right)^2 = F(x,t)\frac{dx}{dt}. \tag{1.50}$$

The next step is to integrate this equation on both sides writing as usual $v(t) = \frac{dx}{dt}$. We use that

$$\int_{t_0}^{t} \frac{df(t')}{dt'}dt' = f(t) - f(t_0)$$

i.e. that the integration is the inversion of differentiation. This gives

$$\frac{m}{2}v(t)^2 - \frac{m}{2}v(t_0)^2 = \int_{t_0}^{t} F(x(t'),t')\frac{dx}{dt'}dt' \tag{1.51}$$

In the next step we make the crucial assumption that the force is not explicitly dependent on time, i.e.

$$F(x,t) = F(x). \tag{1.52}$$

Of course, the force can still change with time because the position $x(t)$ changes and then modifies the acting force. This is however a so called implicit time dependence. Now we note that one can change the integration variable of an integral according to the following recipe: Consider

$$I = \int_{u_0}^{u_1} f(u)\,du$$

We consider the relationship $u(s)$ between the original variable and a new variable s. Let s_0 be the value where $u(s_0) = u_0$ and similarly s_1 obey $u(s_1) = u_1$. Then we can change the integration variable from u to s according to

$$I = \int_{s_0}^{s_1} f(u(s))\frac{du}{ds}ds. \tag{1.53}$$

We can now use this relationship and express the above integration over time as an integration over space:

$$\int_{t_0}^{t} F(x(t'))\frac{dx}{dt'}dt' = \int_{x_0}^{x} F(x')\,dx'. \tag{1.54}$$

We finally introduce the potential $U(x)$ such that

$$F(x) = -\frac{dU(x)}{dx}. \tag{1.55}$$

The potential is therefore nothing but minus the integral of $F(x)$. Note, $U(x)$ is not uniquely defined. We get the same force if we shift the potential by a constant value. It follows

$$\int_{x_0}^{x} F(x')\,dx' = -\int_{x_0}^{x} \frac{dU(x')}{dx'}dx' = -\left(U(x) - U(x_0)\right). \tag{1.56}$$

Now we are in a position to insert the integral into our above expression, Eq. 1.51and obtain

$$\frac{m}{2}v(t)^2 - \frac{m}{2}v(t_0)^2 = -U(x(t)) + U(x(t_0)). \tag{1.57}$$

Introducing

$$E(t) = \frac{m}{2}v(t)^2 + U(x(t)) \tag{1.58}$$

our results corresponds to

$$E(t) = E(t_0). \tag{1.59}$$

Since the two times t and t_0 are completely arbitrary, we conclude that the quantity E is conserved, i.e. does not change as function of time:

$$\frac{dE}{dt} = 0. \tag{1.60}$$

E is called the energy of the system and we just demonstrated energy conservation. The first term, $\frac{m}{2}v^2$, is the kinetic energy and $U(x)$ is the potential energy. The fact that U is determined only up to an overall constant obviously doesn't affect energy conservation. The key assumption needed to conclude that the energy is conserved was that the force F and equally the potential U don't explicitly depend on time: $U(x,t) = U(x)$. Indeed, more generally, energy is conserved whenever there is no preferred time point (homogeneity of time).

1.3.2 Mathematical tools 2: vector product

In addition to the scalar product of two vectors, one can also define the vector product

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} \tag{1.61}$$

which we define in terms of the components of $\mathbf{c}$:

$$c_i = \sum_{jk} \epsilon_{ijk} a_j b_k \tag{1.62}$$

where the component of the totally antisymmetric unit tensor of rank three are

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if} \quad \{i,j,k\} = \{1,2,3\} \text{ or cyclic permutation} \\ -1 & \text{if} \quad \{i,j,k\} = \{2,1,3\} \text{ or cyclic permutation} \\ 0 & \text{otherwise} \end{cases} \tag{1.63}$$

Cyclic permutations are defined as $\{i,j,k\} \to \{k,i,j\} \to \{j,k,i\}$. Explicitly this yields

$$\begin{aligned} c_1 &= a_2b_3 - a_3b_2 \\ c_2 &= a_3b_1 - a_1b_3 \\ c_3 &= a_1b_2 - a_2b_1 \end{aligned}$$

Let us show that **c** is orthogonal to both, **a** and **b**. We write

$$\mathbf{c}\cdot\mathbf{a}=\sum_{ijk}\epsilon_{ijk}a_jb_ka_i=\sum_{jik}\epsilon_{jik}a_ib_ka_j=\sum_{jik}\epsilon_{jik}a_jb_ka_i$$

Next we use $\epsilon_{ijk}=-\epsilon_{jik}$ and obtain

$$\mathbf{c}\cdot\mathbf{a}=-\sum_{ijk}\epsilon_{ijk}a_jb_ka_i=-\mathbf{c}\cdot\mathbf{a} \tag{1.64}$$

and it follows immediately that $\mathbf{c}\cdot\mathbf{a}=0$ the proof for $\mathbf{c}\cdot\mathbf{b}=0$ can be done in full analogy. Finally we show that the magnitude of the vector product obeys

$$|\boldsymbol{c}|=|\mathbf{a}|\,|\mathbf{b}|\,|\sin\theta| \tag{1.65}$$

where θ is the angle between **a** and **b**. To proof this we use (see homework assignment):

$$\sum_i\epsilon_{ijk}\epsilon_{ilm}=\delta_{jl}\delta_{km}-\delta_{jm}\delta_{kl}. \tag{1.66}$$

The magnitude of **c** follows from

$$\begin{aligned}\mathbf{c}^2 &= \sum_i c_i^2=\sum_{ijklm}\epsilon_{ijk}\epsilon_{ilm}a_jb_ka_lb_m\\ &= \sum_{jklm}\left(\delta_{jl}\delta_{km}-\delta_{jm}\delta_{kl}\right)a_jb_ka_lb_m=\sum_{lm}\left(a_l^2b_m^2-a_lb_la_mb_m\right)\\ &= \mathbf{a}^2\mathbf{b}^2-(\mathbf{a}\cdot\mathbf{b})^2=\left|\mathbf{a}^2\right|\left|\mathbf{b}\right|^2\left(1-\cos^2\theta\right). \end{aligned} \tag{1.67}$$

Using $\cos^2\theta+\sin^2\theta=1$, we proof Eq.1.65. Thus, the vector product $\mathbf{c}=\mathbf{a}\times\mathbf{b}$ (also refereed to as cross product) creates a vector that is orthogonal to the two vectors **a** and **b**. This procedure is ill defined in case bot vectors are parallel, but then follows $\mathbf{c}=\mathbf{0}$.

1.3.3 Mathematical tools 3: partial derivatives

Another mathematical concept that occurs for the first time is that of a partial derivative. This is pretty straightforward: consider a function that depends on several variables $u(x,y,z,t)$. We denote the derivative with respect to the coordinate x etc. by

$$\frac{\partial u(x,y,z,t)}{\partial x}. \tag{1.68}$$

To use a new symbol makes sense for the following reason: if we are interested in the explicit time dependence we use $\frac{\partial u}{\partial t}$. On the other hand, it could be that the variables x, y, and z also depend on time, i.e. they are functions, e.g. $x(t)$. Suppose we are interested in the total change as function of time. The usual rules of differentiation lead to

$$\begin{aligned}\frac{du\left(x,y,z,t\right)}{dt} &= \frac{\partial u}{\partial x}\frac{dx}{dt}+\frac{\partial u}{\partial y}\frac{dy}{dt}+\frac{\partial u}{\partial z}\frac{dz}{dt}+\frac{\partial u}{\partial t} \\ &= \sum_{i=1}^{3}\frac{\partial u}{\partial x_i}\frac{dx_i}{dt}+\frac{\partial u}{\partial t} \\ &= \nabla u\cdot\frac{d\mathbf{r}}{dt}+\frac{\partial u}{\partial t}. \qquad (1.69)\end{aligned}$$

Here we used the notation

$$\nabla u=\left(\frac{\partial u}{\partial x},\frac{\partial u}{\partial y},\frac{\partial u}{\partial z}\right) \qquad (1.70)$$

for the vector of the derivatives of a function $u\left(\mathbf{r}\right)$ with respect to the components of $\mathbf{r}$. Other notations are

$$\nabla u\left(\mathbf{r}\right)=\frac{\partial}{\partial\mathbf{r}}u\left(\mathbf{r}\right)=\mathrm{grad}u\left(\mathbf{r}\right). \qquad (1.71)$$

The vector ∇u is called the gradient of u. If for example the biggest change of the function $u\left(\mathbf{r}\right)$ occurs if one changes the x-coordinate, then the x-component of ∇u is the largest. Thus, ∇u points in the direction of the biggest change of the function u.

1.3.4 Energy conservation in three dimensions

After our analysis of energy conservation in one dimension, we generalize the concept to the case of higher space dimensions. We start again from our equation of motion

$$m\frac{d^2\mathbf{r}}{dt^2}=\mathbf{F}\left(x,t\right). \qquad (1.72)$$

The individual components of the acceleration and force vectors obey of course:

$$m\frac{d^2x_i}{dt^2}=F_i\left(x,t\right), \qquad (1.73)$$

with $i=x,y,z$. Thus, F_i is the force acting in the i-th direction. We multiply these equations by $\frac{dx_i}{dt}$ yielding

$$\frac{m}{2}\frac{d}{dt}\left(\frac{dx_i}{dt}\right)^2=F_i\left(\mathbf{r},t\right)\frac{dx_i}{dt}. \qquad (1.74)$$

In analogy to the one-dimensional case we integrate this equation on both sides gives

$$\frac{m}{2}v_i\left(t\right)^2-\frac{m}{2}v_i\left(t_0\right)^2=\int_{t_0}^{t}F_i\left(\mathbf{r}\left(t'\right),t'\right)\frac{dx_i}{dt'}dt'. \qquad (1.75)$$

At this point we want to sum both sides of the equation over the vector components. It follows:

$$\frac{m}{2}\mathbf{v}(t)^2 - \frac{m}{2}\mathbf{v}(t_0)^2 = \int_{t_0}^{t} \mathbf{F}(\mathbf{r}(t')) \cdot \frac{d\mathbf{r}}{dt'} dt'. \tag{1.76}$$

Next we consider a vector the i-th component of the force is assumed to be the derivative of a function (also called scalar function since it is not a vector) We say that the vector $\mathbf{F}(\mathbf{r})$ is the gradient of the potential $U(\mathbf{r})$:

$$\mathbf{F}(\mathbf{r}) = -\nabla U(\mathbf{r}) = -\left(\frac{\partial U(\mathbf{r})}{\partial x}, \frac{\partial U(\mathbf{r})}{\partial y}, \frac{\partial U(\mathbf{r})}{\partial z}\right). \tag{1.77}$$

If this is the case, we call the force a conservative force. Below we discuss what happens in case of non-conservative forces. We make once again the assumption that the force does not depend explicitly on time. We write this explicitly as

$$\begin{aligned} \int_{t_0}^{t} \mathbf{F}(\mathbf{r}(t')) \cdot \frac{d\mathbf{r}}{dt'} dt' &= \sum_{i=1}^{3} \int_{t_0}^{t} F_x(\mathbf{r}(t')) \frac{dx_i}{dt'} dt' \\ &= -\sum_{i=1}^{3} \int_{t_0}^{t} \frac{\partial U(\mathbf{r}(t'))}{\partial x_i} \frac{dx_i}{dt'} dt' \end{aligned} \tag{1.78}$$

Now we recall that for a function that doesn't depend explicitly on time, it holds

$$\frac{du(x,y,z)}{dt} = \nabla u \cdot \frac{d\mathbf{r}}{dt} \tag{1.79}$$

which allows to write

$$\begin{aligned} \int_{t_0}^{t} \mathbf{F}(\mathbf{r}(t')) \cdot \frac{d\mathbf{r}}{dt'} dt' &= -\int_{t_0}^{t} \frac{dU(\mathbf{r}(t'))}{dt'} dt' \\ &= -U(\mathbf{r}(t)) + U(\mathbf{r}(t_0)). \end{aligned} \tag{1.80}$$

If we now insert this result into Eq.1.76, we finally obtain

$$\frac{m}{2}\mathbf{v}(t)^2 + U(\mathbf{r}(t)) = \frac{m}{2}\mathbf{v}(t_0)^2 + U(\mathbf{r}(t_0)). \tag{1.81}$$

Since the two time points are completely arbitrary, we find again that the energy

$$E = \frac{m}{2}\mathbf{v}^2 + U(\mathbf{r}) \tag{1.82}$$

is a conserved quantity. Thus, for force fields $\mathbf{F}$ that are conservative, with potential $U(\mathbf{r})$ that does not explicitly depend on time follows that the energy is conserved. The energy is once again the sum of kinetic and potential energy.

1.3.5 Mathematical tools 4: vector analysis I

In addition to the gradient, one can introduce other differential operators on vectorial functions. If we consider a vector $\mathbf{f}(\mathbf{r})$ we can build the rotation

$$\text{rot}\,\mathbf{f} = \frac{\partial}{\partial \mathbf{r}} \times \mathbf{f} = \nabla \times \mathbf{f} = \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}, \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}, \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right)$$

It can also be written using the ϵ-tensor if we express its components

$$(\nabla \times \mathbf{f})_i = \sum_{jk} \epsilon_{ijk} \frac{\partial f_k}{x_j}.$$

An important statement is Stokes -theorem that says

$$\oint_C d\mathbf{r} \cdot \mathbf{f}(\mathbf{r}) = \int_A d\mathbf{s} \cdot (\nabla \times \mathbf{f}).$$

Here C refers to a closed curve in space (i.e. $\oint_C$ refers to a closed line integral) and A is the area enclosed by C. $d\mathbf{s}$ is a surface element and its direction is along the normal vector of the surface (i.e. it is orthogonal to the surface).

It is easiest to proof this statement in the special case of a square of length a, where $d\mathbf{s} = dxdy\mathbf{e_z}$. For the contour integral follows

$$\oint_C d\mathbf{r} \cdot \mathbf{f}(\mathbf{r}) = \int_0^a dx\,(f_1(x,0) - f_1(x,a)) + \int_0^a dy\,(f_2(a,y) - f_2(0,y))$$

while the surface integral is

$$\begin{aligned} \int_A d\mathbf{s} \cdot (\nabla \times \mathbf{f}) &= \int_0^a dx \int_0^a dy\,(\nabla \times \mathbf{f})_z = \int_0^a dx \int_0^a dy \left(\frac{\partial f_2(x,y)}{\partial x} - \frac{\partial f_1(x,y)}{\partial y} \right) \\ &= \int_0^a dx\,(f_1(x,0) - f_1(x,a)) + \int_0^a dy\,(f_2(a,y) - f_2(0,y)) \end{aligned}$$

In the last step we used

$$\begin{aligned} \int_0^a dx \frac{\partial f_2(x,y)}{\partial x} &= f_2(a,y) - f_2(0,y), \\ \int_0^a dy \frac{\partial f_1(x,y)}{\partial y} &= f_1(x,a) - f_1(x,0). \end{aligned}$$

This proofs Stokes theorem for the square. A general proof can be devised by subdividing each area into a grid of infinitesimal squares. The contributions at the interfaces of the squares all cancel to zero and only the outer curve survives.

1.3.6 Conservative forces

When we discussed energy conservation in three dimensions we had to confine ourselves to forces that can be expressed as gradient of a scalar function

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

We call such forces *conservative*. We also introduce the work done by force, moving an object from $\mathbf{r}_1$ to $\mathbf{r}_2$ as

$$W = \int_{\mathbf{r_1}}^{\mathbf{r_2}} d\mathbf{r} \cdot \mathbf{F} = -\int_{\mathbf{r_1}}^{\mathbf{r_2}} \mathbf{dr} \frac{\partial \mathbf{U}}{\partial \mathbf{r}} = -\int_{\mathbf{t_1}}^{\mathbf{t_2}} \frac{\partial \mathbf{U}}{\partial \mathbf{r}} \cdot \frac{d\mathbf{r}}{dt} dt,$$

such that our earlier analysis of Newton's law (see Eq.1.76) leads to:

$$\frac{m}{2} \mathbf{v}(t)^2 - \frac{m}{2} \mathbf{v}(t_0)^2 = W.$$

Let us discuss in some more detail the implications are conditions of conservative forces. We consider

$$\begin{aligned}
\nabla \times \mathbf{F} &= \sum_{ijk} \epsilon_{ijk} \mathbf{e}_i \frac{\partial f_k}{\partial x_j} = -\sum_{ijk} \epsilon_{ijk} \mathbf{e}_i \frac{\partial^2 U}{\partial x_j \partial x_k} \\
&= -\sum_{ijk} \epsilon_{ikj} \mathbf{e}_i \frac{\partial^2 U}{\partial x_k \partial x_j} = \sum_{ijk} \epsilon_{ijk} \mathbf{e}_i \frac{\partial^2 U}{\partial x_j \partial x_k} \\
&= -\nabla \times \mathbf{F}
\end{aligned}$$

It follows therefore that $\nabla \times \mathbf{F} = 0$ for a conservative force. If we now use Stokes theorem we find

$$\oint_C d\mathbf{r} \cdot \mathbf{F}(\mathbf{r}) = 0.$$

It follows that the work done by a conservative force around a closed loop vanishes. In particular, it implies that

$$W = \int_{\mathbf{r_1}}^{\mathbf{r_2}} d\mathbf{r} \cdot \mathbf{F}$$

is independent on the path that connected the two points $\mathbf{r}_1$ and $\mathbf{r}_2$ as we can deform each contour in to an arbitrary other contour by adding closed loops. This implies that the above work is uniquely defined by properties at $\mathbf{r}_1$ and $\mathbf{r}_2$ (it is $W = U(\mathbf{r_2}) - U(\mathbf{r_1})$). This is an crucial ingredient of the concept of energy, namely a quantity that can be defined at every instant of the path. The criterion for a conservative force is then:

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

1.4 The variational principle

Next we wish to ask the question whether Newton's law, in terms of forces, is the only appropriate way to describe mechanical motion. We make the assumption that the physical path $x(t)$ of the time evolution of a mechanical system is determined by extremal properties of a scalar (Lagrange) function

$$L(x, \dot{x}, t). \tag{1.83}$$

Our above statement that only the first time derivatives are needed to characterize the time evolution is reflected in the fact that L does not depend on second or higher derivatives. The time average of L, called the action, is given by (the generalization to more than one degree of freedom is straightforward):

$$S = \int_{t_a}^{t_b} L(x, \dot{x}, t)\, dt. \tag{1.84}$$

Suppose we know the final and end position x_a and x_b, respectively. We now claim that the physical path $x(t)$ minimizes $S[x]$. Suppose we know this path, then the path

$$x(t) + \delta x(t) \tag{1.85}$$

with some small $\delta x(t)$, obeying $\delta x(t_a) = \delta x(t_b) = 0$, will increase S. The change in action is obviously

$$\begin{aligned} \Delta S &= \int_{t_a}^{t_n} L\left(x + \delta x, \dot{x} + \delta \dot{x}, t\right) dt - \int_{t_a}^{t_n} L\left(x, \dot{x}, t\right) dt. \\ &\simeq \int_{t_a}^{t_n} \left(\frac{\partial L}{\partial x} \delta x + \frac{\partial L}{\partial \dot{x}} \delta \dot{x} \right) dt \\ &= \int_{t_a}^{t_n} \left(\frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \right) \delta x dt \end{aligned} \tag{1.86}$$

In the last step we used that $\delta \dot{x} = d\delta x/dt$ and that $\delta x(t)$ vanishes at the boundaries. If indeed $x(t)$ is a minimum of S, the first derivative w.r.t. $\delta x(t)$ should vanish, which is obeyed for arbitrary x if

$$\frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} = 0. \tag{1.87}$$

This is the Euler-Lagrange equation.

An interesting observation is that the equation of motion, as determined by L is unchanged if we multiply L by an arbitrary coefficient. There is another modification of the Lagrange function that leaves the equation of motion unchanged. Suppose we add to L a total time derivative

$$L'\left(x, \dot{x}, t\right) = L\left(x, \dot{x}, t\right) + \frac{d}{dt} f(x, t). \tag{1.88}$$

It leads to an additive correction to the action that only depends on the fixed initial and final coordinates.

1.4.1 The Galilei transformation

We will postulate the transformation properties of classical mechanics and use this postulate to identify the functional form of the Lagrange function. We consider an isotropic and homogeneous system, such that the Lagrange function will not depend explicitly on the coordinate $\mathbf{r}$ on the time t and on the direction of the velocity $\mathbf{v}/\left|\mathbf{v}\right|$. Thus, we obtain

$$L\left(\mathbf{r},\mathbf{v},t\right)=L\left(v^{2}\right) \tag{1.89}$$

where $v=\left|\mathbf{v}\right|$ is the magnitude of the velocity. From the Euler-Lagrange equations follows immediately that

$$\frac{d}{dt}\frac{\partial L}{\partial\mathbf{v}}=0 \tag{1.90}$$

which implies that $\partial L/\partial\mathbf{v}=$*const.*. For a function that only depends on the magnitude of $\mathbf{v}$ follows

$$\frac{\partial L}{\partial\mathbf{v}}=\frac{\partial L}{\partial v}\frac{\partial v}{\partial\mathbf{v}}=\frac{\partial L}{\partial v}\frac{\mathbf{v}}{v} \tag{1.91}$$

Since this vector is a constant, it must hold that $\mathbf{v}$ itself is a constant. Thus, all systems that obey Eq.1.89 only allow for motions with constant velocity.

Next we want to analyze the motion with constant velocity in two distinct coordinate systems that more relative to each other with velocity $\mathbf{V}$. We postulate the Galilean Principle that relates the coordinates and times in both coordinate systems as

$$\begin{aligned}\mathbf{r} &= \mathbf{r}'+\mathbf{V}t\\ t &= t'\end{aligned} \tag{1.92}$$

The second statement expresses the assumption of an absolute time, while the first corresponds to the view that velocities in different coordinate systems are additive:

$$\dot{\mathbf{r}}=\dot{\mathbf{r}}'+\mathbf{V}. \tag{1.93}$$

Under such a transformation the Lagrange function transforms as

$$\begin{aligned}L' &= L\left(v'^{2}\right)=L\left(\left(v^{2}+2\mathbf{v}\cdot\mathbf{V}+V^{2}\right)\right)\\ &\simeq L\left(v'^{2}\right)+2\mathbf{v}\cdot\mathbf{V}\frac{\partial L}{\partial v^{2}}\end{aligned} \tag{1.94}$$

where we considered infinitesimal $\mathbf{V}$ in the last step. The Galilei transformation should be true for infinitesimal velocities as well. If the two descriptions, in terms of L or L' are physically identical, we require that the additional term is a total derivative of time of a function that only depends on $\mathbf{r}$ and t, i.e.

$$\mathbf{v}\frac{\partial L}{\partial v^{2}}=\frac{d}{dt}f\left(\mathbf{r},t\right). \tag{1.95}$$

This is only the case if $\frac{\partial L}{\partial v^2}$ is independent on v. Thus, it follows for a free particle in homogeneous space that

$$L = \frac{m}{2} v^2 \tag{1.96}$$

where the coefficient $m/2$ was introduced for convenience. m is of course the mass of the object. For the case of a free particle in homogeneous space m will not enter the equations of motion, as it is just an overall coefficient of the Lagrange function.

1.4.2 Lagrange function with potential

In case of a system with potential U we recover the Newton's equations if we write

$$L = T(\mathbf{v}) - U(\mathbf{r}) \tag{1.97}$$

where $\mathbf{r}_i$ is the position vector. The kinetic energy part is again

$$T = \frac{m\mathbf{v}^2}{2} \tag{1.98}$$

The second part is the potential energy. From the Euler Lagrange equation follows

$$m\frac{d\mathbf{v}}{dt} = \mathbf{F} \tag{1.99}$$

where

$$\mathbf{F} = -\frac{\partial U}{\partial \mathbf{r}} \tag{1.100}$$

is the force acting on the particle. Thus, we find that the variational principle offers an alternative approach to mechanics, fully equivalent to Newton's laws.

1.4.3 An application of the Lagrange formalism: motion on a cylinder

During this lecture we will not dwell too much on applications of the Lagrange formalism. This will make up a significant part of the lecture Theory B. Still, for completeness, we will solve a simple problem. We consider a particle that is forced to move on the surface of a cylinder. This reminds us of the polar coordinates we used earlier. We write for a three dimensional vector

$$\mathbf{r}(t) = x(t)\,\mathbf{e}_x + y(t)\,\mathbf{e}_y + z(t)\,\mathbf{e}_z$$

The motion on the surface of a cylinder of radius R implies that we allow only for those orbits with

$$x^2(t) + y^2(t) = R^2.$$

This suggests to introduce the following representation:

$$\begin{aligned} x(t) &= R\cos\varphi(t), \\ y(t) &= R\sin\varphi(t). \end{aligned}$$

Using $\cos^2 x + \sin^2 x = 1$, we see immediately that this parametrization fulfills our constraint. It follows for the velocity

$$\begin{aligned}\frac{d\mathbf{r}(t)}{dt} &= \frac{dx(t)}{dt}\mathbf{e}_x + \frac{dy(t)}{dt}\mathbf{e}_y + \frac{dz(t)}{dt}\mathbf{e}_z \\ &= \left(-\sin\varphi(t)\,\mathbf{e}_x + \cos\varphi(t)\,\mathbf{e}_y\right) R\frac{d\varphi(t)}{dt} + \frac{dz(t)}{dt}\mathbf{e}_z.\end{aligned}$$

For the Lagrange function we need the kinetic and potential energy.

$$\begin{aligned}T &= \frac{m}{2}\left(\frac{d\mathbf{r}(t)}{dt}\right)^2 \\ &= \frac{m}{2}R^2\left(\frac{d\varphi}{dt}\right)^2 + \frac{m}{2}\left(\frac{dz}{dt}\right)^2.\end{aligned}$$

We can always write our potential as $U(z,\varphi)$. Let us assume that in our problem the potential only depends on the height z. This gives rise to the Lagrange function

$$\begin{aligned}L &= \frac{m}{2}R^2\left(\frac{d\varphi}{dt}\right)^2 + \frac{m}{2}\left(\frac{dz}{dt}\right)^2 - U(z). \\ &= \frac{m}{2}R^2\dot{\varphi}^2 + \frac{m}{2}\dot{z}^2 - U(z).\end{aligned}$$

We can now obtain the equations of motion for the angle and height from

$$\begin{aligned}\frac{d}{dt}\frac{\partial L}{\partial \dot{z}} - \frac{\partial L}{\partial z} &= 0, \\ \frac{d}{dt}\frac{\partial L}{\partial \dot{\varphi}} - \frac{\partial L}{\partial \varphi} &= 0.\end{aligned}$$

And we find

$$\begin{aligned}m\frac{d^2 z}{dt^2} &= -\frac{\partial U}{\partial z} = F_z \\ mR^2\frac{d^2\varphi}{dt^2} &= 0\end{aligned}$$

We find the general solution

$$\varphi(t) = \varphi_0 + \omega t$$

with initial angle and angular frequency φ_0 and ω , respectively. The appeal of this formulation is that it can be efficiently used to describe motions with constraints. Of course, being solely a reformulation of Newtonian approach that is based on Forces, the Lagrange formalism does not add features that could not have been explained without it. Still, the development of modern quantum mechanics and quantum field theory is deeply based on this formulation.

We also identify a new conservation law. It follows obviously for a potential that is independent on φ that

$$\frac{d}{dt}L_z = 0,$$

where

$$L_z = mR^2\frac{d\varphi}{dt}.$$

We will see later that L_z is the z-component of the angular momentum vector.

Chapter 2

Mechanical motion in one dimension

2.1 Harmonic oscillator

One of the most important problems in physics is the harmonic oscillator. We will study it now in its most elementary form. The potential of the harmonic oscillator is

$$U(x) = \frac{k}{2}x^2$$

with corresponding force

$$F = -kx.$$

Thus, whenever a restoring force acts that gets linearly larger the bigger an object is displaced from some reference point (here $x = 0$) then one has a so called harmonic problem. Applications are restoring forces of weakly distorted sponges or a pendulum with small amplitude (see below). The coefficient k is the harmonic force constant.

Let us first solve the equation of motion

$$m\frac{d^2x(t)}{dt^2} + kx(t) = 0. \tag{2.1}$$

To solve this differential equation we notice that we are looking for functions that are proportional to their second derivatives. We notice that $\frac{d^2\cos(x)}{dx^2} = -\cos(x)$ and similarly $\frac{d^2\sin(x)}{dx^2} = -\sin(x)$. This suggests the ansatz

$$x(t) = a\cos(\omega t) + b\sin(\omega t) \tag{2.2}$$

such that

$$\begin{aligned} \frac{dx}{dt} &= -a\omega\sin(\omega t) + b\omega\cos(\omega t), \\ \frac{d^2x}{dt^2} &= -a\omega^2\cos(\omega t) - b\omega^2\sin(\omega t). \end{aligned}$$

It follows that our ansatz, Eq.2.2 obeys

$$\frac{d^2x}{dt^2} + \omega^2 x = 0.$$

This allows us to identify

$$\omega = \sqrt{\frac{k}{m}}$$

for the angular frequency of our oscillator. Since $x_0 = x\,(t = 0) = a$ and $v_0 = v\,(t = 0) = b\omega$ we write our solution as

$$x\,(t) = x_0 \cos(\omega t) + \frac{v_0}{\omega} \sin(\omega t)\,. \tag{2.3}$$

As expected, we obtain an oscillating solution. The frequency grows with $\sqrt{k}$, i.e. a stiffer spring oscillates at a higher frequency. On the other hand a heavier weight on the spring leads to smaller frequency.

2.1.1 Mathematical tools 5: complex numbers and the exponential function with complex argument

The imaginary unit i is defined via $i = \sqrt{-1}$, i.e. $i^2 = -1$. A generic complex number is then written as $z = x + iy$, where x and y are both real. Here $x = \mathrm{Re}\,(z)$ and $y = \mathrm{Im}\,(z)$ is the real and imaginary part of z, respectively. In case of $y = 0$, z is real, while it is purely imaginary for $x = 0$. The complex conjugate z^* of z is $z^* = a - ib$ such that

$$z^* z = (x - iy)(x + iy) = x^2 + y^2 \tag{2.4}$$

is real. If we depict a complex number as a point in a two dimensional coordinate system where the x-axis shows the real part and the y-axis the imaginary part, then $|z| = \sqrt{z^* z}$ is the distance of the point from the origin. It is the absolute magnitude of the complex number. Since $\frac{1}{i}$ should be the number that gives 1 if multiplied with i we obtain

$$\frac{1}{i} = -i.$$

It is an interesting exercise to analyze or appropriately generalize known functions with real argument to the case where they have complex arguments. We do this for the exponential function

$$e^z = e^{x+iy} \tag{2.5}$$

We use $e^{a+b} = e^a e^b$ and obtain $e^z = e^x e^{iy}$, i.e. we really only need to understand what the exponential function with purely imaginary argument amounts to. We use

$$\frac{d}{d\lambda} e^{\lambda z} = \lambda e^{\lambda z} \tag{2.6}$$

This implies

$$\begin{aligned}\frac{d}{dy}e^{iy} &= ie^{iy} \\ \frac{d^2}{dy^2}e^{iy} &= -e^{iy}\end{aligned} \tag{2.7}$$

Well, we find that e^{iy} is a function that changes sign after performing two derivatives. We know such functions already: $\frac{d^2}{dy^2}\cos y = -\cos y$ and $\frac{d^2}{dy^2}\sin y = -\sin y$. This suggests to try

$$e^{iy} = a\cos y + b\sin y \tag{2.8}$$

where a and b are complex coefficients that still need be determined. It holds

$$\begin{aligned}\frac{d}{dy}\left(a\cos y + b\sin y\right) &= -a\sin y + b\cos y \\ &= \frac{b}{a}\left(a\cos y - \frac{a^2}{b}\sin y\right)\end{aligned} \tag{2.9}$$

Comparing this with our above expression for the derivative of e^{iy} yields $b = ia$ and $a^2 = -b^2$. The second equation follows immediately from the first. Thus we have

$$e^{iy} = a\left(\cos y + i\sin y\right) \tag{2.10}$$

To determine a we use that $e^0 = a = 1$. Thus, we obtain

$$e^z = e^x\left(\cos y + i\sin y\right). \tag{2.11}$$

A a finite imaginary part of the argument of an exponential function yields therefore oscillatory behavior.

An interesting implication is that we have an alternative way to express an arbitrary complex number via

$$z = x + iy = re^{i\varphi} \tag{2.12}$$

where $r = |z| = \sqrt{x^2 + y^2}$ is the absolute magnitude of the complex number and $\varphi = \arctan\frac{y}{x}$ is the complex phase. In other words, one can always write $x = r\cos\varphi$ and $y = r\sin\varphi$, a result that is most obvious if one sketches z in the complex plane.

2.1.2 Mathematical tools 6: linear differential equations with constant coefficients

A class of differential equations that frequently emerges is of the form:

$$y^{(n)}(x) + a_{n-1}y^{(n-1)}(x) + \cdots + a_0 y(x) = 0. \tag{2.13}$$

This is a linear, homogeneous differential equation of n-th order with constant coefficients. It is linear as the sum of two solutions is a solution itself. It is homogeneous, because of the zero on the right hand side. The coefficients a_m in front of the derivatives $y^{(m)}(x) = \frac{d^m y(x)}{dx^m}$ are independent on x, i.e. they are constants. It is a differential equation of n-th order because the highest derivative that enters is $y^{(n)}(x)$.

Solutions of this differential equation are generically of the form

$$y(x) = y_{0,i} e^{\lambda x}, \tag{2.14}$$

where $y_{0,i}$ is a constant. In order to determine the coefficient λ we insert this solution into the differential equation, noticing that $y^{(1)}(x) = \frac{dy(x)}{dx} = \lambda e^{\lambda x}$, $y^{(2)}(x) = \frac{dy^{(1)}}{dx} = \lambda^2 e^{\lambda x}$, and generally, $y^{(m)}(x) = \lambda^m e^{\lambda x}$. This yields

$$\left(\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0\right) e^{\lambda x} = 0. \tag{2.15}$$

Since $e^{\lambda x}$ is for finite x never zero, we conclude that the coefficient λ is determined by

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0. \tag{2.16}$$

Thus, we need to determine the zeroth of a polynomial of n-th degree. This polynomial is called the characteristic polynomial of the differential equation. A general closed solution exists only for $n \leq 4$. However, numerical solutions of such polynomials can be obtained rather easily. There are in total n possible zeroth λ_i with $i = 1, \cdots, n$ of this polynomial, i.e. we found in fact n solutions

$$y_i(x) = y_{0,i} e^{\lambda_i x} \text{ with } i = 1, \cdots, n. \tag{2.17}$$

Since we have a linear equation we can write the full solution as weighted sum of these solutions:

$$y(x) = \sum_{i=1}^{n} y_i(x) = \sum_{i=1}^{n} y_{0,i} e^{\lambda_i x}, \tag{2.18}$$

where the coefficients $y_{0,i}$ are determined by the n initial or boundary conditions.

There is only one caveat in this solution that requires special attention. If two zeroth λ_i and λ_j are identical ($\lambda_i = \lambda_j$) the two solutions $y_i(x)$ and $y_j(x)$ are identical. If this is the case one choses instead $\tilde{y}_i(x) = e^{\lambda_i x}$ and $\tilde{y}_j(x) = x e^{\lambda_j x}$. An example is

$$y^{(2)}(x) - 6y^{(1)}(x) + 9y(x) = 0 \tag{2.19}$$

which leads to

$$\lambda^2 - 6\lambda + 9 = (\lambda - 3)^2 = 0 \tag{2.20}$$

with double solution $\lambda = 3$. One easily finds that $y_1(x) = e^{3x}$ and $y_2(x) = xe^{3x}$ both solve the differential equation. Clearly a mathematically more rigorous justification of this behavior would be desirable, but it requires some elementary concepts of linear algebra that simply go beyond what this physics course is able to offer.

2.1.3 Systematic solution of the undamped harmonic oscillator

We want to rederive our earlier results for the harmonic oscillator using by solving the Newtons equation as a homogeneous differential equation with constant coefficient. The equation of motion is

$$\frac{d^2x(t)}{dt^2}+\frac{k}{m}x(t)=0. \tag{2.21}$$

This is a second order equation with constant coefficients and we employ the ansatz

$$x(t)=e^{\lambda t} \tag{2.22}$$

yielding the characteristic polynomial

$$\lambda^2+\frac{k}{m}=0 \tag{2.23}$$

which yields the two solutions $\lambda=\pm i\omega$ with $\omega=\sqrt{\frac{k}{m}}$ as before. The two solutions are

$$x_\pm(t)=x_{0\pm}e^{\pm i\omega t} \tag{2.24}$$

The full solution is obviously a linear combination of these two solutions:

$$x(t)=x_{0+}e^{i\omega t}+x_{0-}e^{-i\omega t} \tag{2.25}$$

Using our earlier result for the exponential function with complex argument, it holds

$$\begin{aligned} x(t) &= (\mathrm{Re}x_{0+}+i\mathrm{Im}x_{0+})(\cos\omega t+i\sin\omega t) \\ &+ (\mathrm{Re}x_{0-}+i\mathrm{Im}x_{0-})(\cos\omega t-i\sin\omega t) \end{aligned}$$

Since the variable $x(t)$ should be real it follows that all t:

$$0=(\mathrm{Re}x_{0+}-\mathrm{Re}x_{0-})\sin\omega t+(\mathrm{Im}x_{0+}+\mathrm{Im}x_{0-})\cos\omega t.$$

This implies that the two coefficients have same real part and opposite imaginary part, i.e. they are complex conjugate of each other:

$$x_{0+}=x_{0-}^*.$$

The solution of the differential equation is then

$$x(t) = 2\mathrm{Re}x_{0+}\cos\omega t-2\mathrm{Im}x_{0+}\sin\omega t, \tag{2.26}$$

in full agreement with our earlier result. The coefficients $2\mathrm{Re}x_{0+}$ and $-2\mathrm{Im}x_{0+}$ can as usual be related to the initial position and velocity of the oscillator

$$\begin{aligned} x(t=0) &= 2\mathrm{Re}x_{0+} \\ v(t=0) &= -2\omega\mathrm{Im}x_{0+}. \end{aligned} \tag{2.27}$$

In what follows we will use the method of solving a differential equation with constant coefficients for more general settings.

2.1.4 Damped harmonic oscillator

Friction is a phenomenon where a given systems is coupled to other degrees of freedom (a street in case of tires, air in case of a pendulum or a flying object, internal excitations in case of an oscillating spring etc.). Our mechanical system will then exchange energy with those degrees of freedom and generically it will lose energy, a phenomenon that can be better described using the methods of thermodynamic and statistical mechanics. Here we will essentially ignore those subtleties and simply assume that there we a phenomenological frictional force F_{damp} that acts opposite to the velocity of the system:

$$F_{damp} = -\Gamma \frac{dx}{dt}. \tag{2.28}$$

This force cannot be expressed as derivative of a potential and one finds, as expected that the energy of the system is not conserved. After all our mechanical particle exchanges energy with the environment that is responsible for the friction. Only the total system: mechanical degree of freedom + environment are expected to respect the law of energy conservation. Newton's law is now given as

$$m\frac{d^2x}{dt^2} + \Gamma\frac{dx}{dt} + kx = 0. \tag{2.29}$$

We divide by m and with $\omega_0 = \sqrt{k/m}$ and $\gamma = \Gamma/m$ follows

$$\frac{d^2x}{dt^2} + \gamma\frac{dx}{dt} + \omega_0^2 x = 0. \tag{2.30}$$

ω_0 is of course the frequency of the system without friction. γ is a quantity of dimension 1/time (i.e. it is a rate) that characterizes the friction. This differential equation is once again linear, homogeneous and has constant coefficients i.e. we can use the ansatz

$$x\left(t\right) \propto e^{\lambda t} \tag{2.31}$$

which leads to the characteristic polynomial

$$\lambda^2 + \gamma\lambda + \omega_0^2 = 0 \tag{2.32}$$

which has the two solutions

$$\lambda_{1,2} = -\frac{\gamma}{2} \pm \sqrt{\frac{\gamma^2}{4} - \omega_0^2} \tag{2.33}$$

Depending on the nature of these zeros of the characteristic polynomials, we can distinguish three distinct regimes:

***i)* Weak damping limit:** $\omega_0 > \gamma/2$

If we introduce the real quantity $\Omega = \sqrt{\omega_0^2 - \gamma^2/4} < \omega_0$, we obtain

$$\lambda_{1,2} = -\frac{\gamma}{2} \pm i\Omega \tag{2.34}$$

and our solutions become

$$x_{1,2}(t) \propto e^{-\frac{\gamma}{2}t} e^{\pm i\Omega t}. \tag{2.35}$$

In analogy to our previous calculation we can construct the real solution

$$x(t) = e^{-\frac{\gamma}{2}t} (a \cos \Omega t + b \sin \Omega t) \tag{2.36}$$

If we consider initial values $x_0 = x(t = 0)$ and $v_0 = v(t = 0)$ we can identify

$$\begin{aligned} a &= x_0, \\ b &= \frac{1}{\Omega_0}\left(v_0 + \frac{\gamma}{2}x_0\right). \end{aligned} \tag{2.37}$$

In short, we find

- the frequency of the oscillation is reduced because of the damping, it approaches zero in the limit $\omega_0 \to \gamma/2$.
- the amplitude decays and becomes very small once t is large compared to the time scale $2/\gamma$.

ii) Strong damping limit: $\omega_0 < \gamma/2$

If the damping increases we enter the regime where $\gamma/2$ becomes larger than ω_0 and the two zeros of the characteristic polynomial are real. The solution is now a sum of two exponentials

$$x_\pm(t) = x_{0,\pm} e^{-\gamma_{1,2} t} \tag{2.38}$$

with

$$\gamma_{1,2} = -\lambda_{12} = \frac{\gamma}{2} \mp \sqrt{\frac{\gamma^2}{4} - \omega_0^2} > 0.$$

The full solution is now superposition of these two fundamental solutions and we obtain readily

$$x(t) = \frac{1}{\gamma_2 - \gamma_1}\left[(\gamma_2 x_0 + v_0) e^{-\gamma_1 t} - (\gamma_1 x_0 + v_0) e^{-\gamma_2 t}\right] \tag{2.39}$$

The oscillator is now so strongly damped that it has maximally one passage through zero and decays exponentially .

iii) The aperiodic limit: $\omega_0 = \frac{\gamma}{2}$

It is interesting to individually analyze the case $\omega_0 = \frac{\gamma}{2}$ that separates the weak and strong damping regime. In this case the two zeros of the characteristic polynomial are identical

$$\lambda_{1,2} = -\frac{\gamma}{2}. \tag{2.40}$$

In our discussion of differential equations with constant coefficients we mentioned that the two fundamental solutions are now

$$x_1(t) \propto e^{-\frac{\gamma}{2}t} \tag{2.41}$$

and

$$x_2(t) \propto te^{-\frac{\gamma}{2}t}. \tag{2.42}$$

We can now proceed to use those two solutions and determine the coefficients from the initial values, i.e. consider

$$x(t) = e^{-\frac{\gamma}{2}t}(a + bt) \tag{2.43}$$

It follows with $x_0 = x(t = 0)$ and $v_0 = v(t = 0)$ that $a = x_0$ and $b = v_0 + x_0\frac{\gamma}{2}$. An alternative way to obtain the same result is to start from the solution of the weak or strong coupling limit and perform the limit $\omega_0 \to \gamma/2$. If we start, for example, from Eq.2.39 we can write $\gamma_1 = \frac{\gamma}{2} - \epsilon$ and $\gamma_2 = \frac{\gamma}{2} + \epsilon$ we find for the coefficient

$$\frac{1}{\gamma_{2-}\gamma_1} = \frac{1}{2\epsilon}. \tag{2.44}$$

The limit $\epsilon \to 0$ seems to be singular. However, we can also consider the rest of the solution Eq.2.39 in the limit of small ϵ. To this extend we perform a Taylor expansion with respect to ϵ. It follows

$$(\gamma_2 x_0 + v_0)e^{-\gamma_1 t} - (\gamma_1 x_0 + v_0)e^{-\gamma_2 t} \approx 2\epsilon e^{-\frac{\gamma}{2}t}\left(v_0 t + \frac{\gamma}{2}x_0 t + x_0\right)$$

Since this term vanishes linearly in ϵ plus small corrections we can safely perform the limit $\epsilon \to 0$:

$$x(t) = e^{-\frac{\gamma}{2}t}\left(v_0 t + \frac{\gamma}{2}x_0 t + x_0\right), \tag{2.45}$$

which is precisely our earlier solution. Just like the strong damping limit there are no oscillations (the frequency Ω of the weak damping limit has just vanished), yet there could be one passage through zero at time $t_0 = -\frac{3x_0}{v_0 + x_0\gamma/2}$, depending on whether the initial values are such that $t_0 > 0$.

2.1.5 Driven harmonic oscillator

If an oscillator is excited by an external time dependent force $F(t)$, we have a driven system. The equation of motion is now given as

$$\frac{d^2x(t)}{dt^2} + \gamma\frac{dx(t)}{dt} + \omega_0^2 x(t) = f(t). \tag{2.46}$$

where $f(t) = F(t)/m$. This is now an inhomogeneous second order differential equation with constant coefficients. The inhomogeneity is the function $f(t)$. To solve this problem we proceed as follows: first, let $x_1(t)$ and $x_2(t)$ be the two solutions of the homogeneous problem, i.e. for $f = 0$. We determined those solutions in the previous paragraph. In addition, suppose we know some solution

$\bar{x}(t)$, not necessarily the most general one, of the inhomogeneous equation above. Then the theory of differential equations states that the most general solution can be written as

$$x(t) = \bar{x}(t) + c_1 x_1(t) + c_2 x_2(t) \tag{2.47}$$

where the two constants c_1 and c_2 are, as usual, fixed by the initial conditions

$$\begin{aligned} x(0) &= x_0 = \bar{x}(0) + c_1 x_1(0) + c_2 x_2(0) \\ v(0) &= v_0 = \left.\frac{d\bar{x}(t)}{dt}\right|_{t=0} + c_1 \left.\frac{dx_1(t)}{dt}\right|_{t=0} + c_2 \left.\frac{dx_1(t)}{dt}\right|_{t=0}. \end{aligned} \tag{2.48}$$

One finds, by inserting this solution into the differential equation, that it is indeed a solution. It is also plausible that there should be no more than two constants c_i in the final expression, as we only have two initial conditions, say $x(t=0)$ and $v(t=0)$. Clearly, these comments cannot substitute for a lecture on ordinary differential equations. At best it can serve as a motivation to attend such a lecture to overcome the unsatisfying aftertaste that such a superficial discussion may cause.

Another interesting aspect of inhomogeneous differential equations is the superposition principle. Consider two differential equations with inhomogeneities $f_1(t)$ and $f_2(t)$, yet with same homogeneous part. Thus we have two distinct external forces acting on the same system. Let the particular solution of the two problems be $\bar{x}_1(t)$ and $\bar{x}_2(t)$, respectively. One finds immediately that the particular solution of the problem with inhomogeneity

$$f(t) = f_1(t) + f_2(t) \tag{2.49}$$

is given by

$$\bar{x}(t) = \bar{x}_2(t) + \bar{x}_2(t) \tag{2.50}$$

Suppose, for example we wish to study an external harmonic force

$$f(t) = f_0 \cos(\omega t) \tag{2.51}$$

It will be very convenient to split this force as

$$f(t) = \frac{f_0}{2} e^{i\omega t} + \frac{f_0}{2} e^{-i\omega t} \tag{2.52}$$

and solve for the two parts $f_{1,2}(t) = \frac{f_0}{2} e^{\pm i\omega t}$ separately. The final solution follows with the help of Eq.2.50. If the solution for $f_1(t) = \frac{f_0}{2} e^{i\omega t}$ is known, the solution that belongs to $f_2(t) = \frac{f_0}{2} e^{-i\omega t}$ follows immediately by switching $\omega \to -\omega$ in the final result. This superposition principle is the justification for the usage of complex-valued forces, leading to complex-valued oscillator displacements. In the end, the force and displacement are of course real. The detour into the complex plane is merely a matter of convenience.

Using those remarks we have to solve

$$\frac{d^2 x(t)}{dt^2} + \gamma \frac{dx(t)}{dt} + \omega_0^2 x(t) = f_0 e^{i\omega t}. \tag{2.53}$$

Since we already know the general solutions of the homogeneous equation from the previous paragraph, we only need to find one particular solution $\bar{x}(t)$ of the inhomogeneous equation. We make the plausible ansatz

$$\bar{x}(t) = Af_0 e^{i\omega t} \tag{2.54}$$

that corresponds to a solution where the external frequency is forced upon our oscillator. We insert this solution and find

$$Af_0 \left(-\omega^2 + i\gamma\omega + \omega_0^2\right) e^{i\omega t} = f_0 e^{i\omega t} \tag{2.55}$$

Thus, we see that this is indeed a solution if we chose the amplitude as

$$A(\omega) = \frac{1}{\omega_0^2 + i\gamma\omega - \omega^2} \tag{2.56}$$

$A(\omega)$ is the response function of the system, caused by the external perturbation, represented by the force f. For a better interpretation we write the denominator in $A(\omega)$ as

$$\omega_0^2 + i\gamma\omega - \omega^2 = -(\omega - \omega_1)(\omega - \omega_2) \tag{2.57}$$

where

$$\omega_{1,2} = \pm\Omega + i\frac{\gamma}{2} \tag{2.58}$$

where $\Omega = \sqrt{\omega_0^2 - \gamma^2/4}$. Let us recall, that the solution of the homogeneous problem are in fact $x_{1,2}(t) \propto e^{i\omega_{1,2}t}$, i.e. the amplitude $A(\omega)$ has a pole at the positions of the complex frequencies of the weakly damped regime. The implication becomes most evident in the limit of small damping. If $\gamma \to 0$ the amplitude of the particular solution diverges when the external frequency approaches the oscillator frequency ω_0. This phenomenon is called resonance, where a frequency-matching perturbation induces a growth of the oscillator beyond bounds. Of course, once the amplitude A becomes very large, the assumption of a linear restoring force $F = -kx$ of the oscillator is likely not justified any longer and one needs to include higher order terms to analyze the fate of the resonantly excited oscillator.

For a more detailed analysis, we use the superposition principle and consider the real amplitude of the particular solution due to the force $f(t) = f_0 \cos\omega t$. Before we do this, we return to the full solution

$$x(t) = \bar{x}(t) + c_1 x_1(t) + c_2 x_2(t) \tag{2.59}$$

As soon as there is some damping γ, the homogeneous solutions $x_1(t)$ and $x_2(t)$ will decay exponentially and be small for $t \gg 2/\gamma$. Thus in the long time limit, we are only interested in the particular solution $\bar{x}(t)$:

$$x\left(t \gg \frac{2}{\gamma}\right) \approx \bar{x}(t)\,. \tag{2.60}$$

Using the superposition principle we find

$$\begin{aligned} \bar{x}(t) &= A(\omega)\frac{f_0}{2}e^{i\omega t} + A(-\omega)\frac{f_0}{2}e^{-i\omega t} \\ &= f_0 \mathrm{Re}\left(A(\omega)e^{i\omega t}\right), \end{aligned} \tag{2.61}$$

where we used that $A(-\omega)$ is the complex conjugate of $A(\omega)$. To proceed, we write the complex amplitude as

$$A(\omega) = |A(\omega)| e^{i\alpha(\omega)} \tag{2.62}$$

which leads to

$$\begin{aligned} \bar{x}(t) &= f_0 |A(\omega)| \mathrm{Re} e^{i(\omega t + \alpha(\omega))} \\ &= f_0 |A(\omega)| \cos(\omega t + \alpha(\omega)) \end{aligned} \tag{2.63}$$

It holds

$$\begin{aligned} |A(\omega)| &= \sqrt{A^*(\omega) A(\omega)} \\ &= \sqrt{\frac{1}{\omega_0^2 - i\gamma\omega - \omega^2}\frac{1}{\omega_0^2 + i\gamma\omega - \omega^2}} \\ &= \frac{1}{\sqrt{\left(\omega_0^2 - \omega^2\right)^2 + \gamma^2\omega^2}} \end{aligned} \tag{2.64}$$

and

$$\tan\alpha(\omega) = \frac{\mathrm{Im}A(\omega)}{\mathrm{Re}A(\omega)} = \frac{\gamma\omega}{\omega_0^2 - \omega^2}. \tag{2.65}$$

As long as $\gamma/2 < \omega_0$ the amplitude displays a maximum at

$$\omega_{max} = \Omega = \sqrt{\omega_0^2 - \frac{\gamma^2}{4}} \tag{2.66}$$

which is the frequency of the weakly damped oscillator without external force. This is the mentioned resonance effect. The long time oscillations are largest when the frequency ω of the external force equals the natural frequency of the oscillator. The damping determines the width of this resonance is, for small γ given by the damping γ, i.e. for $\gamma \ll \omega_0$ it holds that $A(\omega = \Omega \pm \gamma)$ is only a fraction of $A(\omega = \Omega)$.

The phase of the amplitude $\alpha(\omega)$ is shifted relative to the external force. We find $\alpha(\omega \to 0) \to 0$, i.e. the system responds without delay to the external force. On the other hand, $\alpha(\omega \to \infty) \to \pi$ and the system oscillates opposite to the external force. To get to this conclusion we used that $\tan\alpha(\omega \to \infty) \approx -\frac{\gamma}{\omega}$, i.e. it vanishes from below implying that the phase must be π. When the driving frequency ω equals the oscillator frequency without damping, i.e. ω_0 it holds $\tan\alpha(\omega_0) \to \infty$, i.e. the phase α is $\frac{\pi}{2}$.

2.1.6 Mathematical tools 7: the $\delta-$Distribution

Suppose our external force, acting on the oscillator is a sudden pulse of length ϵ, i.e. $f(t) = f_0 \delta_\epsilon(t)$ where

$$\delta_\epsilon(t) = \begin{cases} \frac{1}{\epsilon} & \text{if } |t| < \frac{\epsilon}{2} \\ 0 & \text{otherwise} \end{cases} \tag{2.67}$$

It turns out that one can cleverly construct the response to an arbitrary external force $f(t)$ in terms of properties that result of an appropriate limit of δ_ϵ for $\epsilon \to 0$.

To this extend we introduce the Delta-function (better Delta-distribution)

$$\delta(t) = \lim_{\epsilon \to 0} \delta_\epsilon(t). \tag{2.68}$$

The above limit is in fact not really well defined. A more careful analysis shows that $\delta(t)$ is not a function but requires the concept of distributions (mappings from the space of functions to complex numbers). A more pedestrian understanding is that one implicitly works in the limit of small but finite ϵ, where δ_ϵ is well defined and performs the limit $\epsilon \to 0$ at the end. Let us analyze a few properties of $\delta_\epsilon(t)$. It holds for $\epsilon < a$:

$$\int_{-a}^{a} dt \delta_\epsilon(t) = 1 \tag{2.69}$$

we can then perform the limit $\epsilon \to 0$ which we write symbolically as

$$\int_{-a}^{a} dt \delta(t) = 1 \text{ for all finite } a > 0. \tag{2.70}$$

This implies for an arbitrary continuous function $h(t)$ in $[-a, a]$ that

$$\begin{aligned} \int_{-a}^{a} dt h(t) \delta(t) &= \lim_{\epsilon \to 0} \int_{-a}^{a} dt h(t) \delta_\epsilon(t) \\ &= \lim_{\epsilon \to 0} \left(\int_{-a}^{-\epsilon/2} dt h(t) \delta_\epsilon(t) + \int_{\epsilon/2}^{a} dt h(t) \delta_\epsilon(t) + \int_{-\epsilon/2}^{\epsilon/2} dt h(t) \delta_\epsilon(t) \right) \\ &= 0 + 0 + \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_{-\epsilon/2}^{\epsilon/2} h(t) dt \\ &= \lim_{\epsilon \to 0} \frac{1}{\epsilon} h(0) \int_{-\epsilon/2}^{\epsilon/2} dt = h(0). \end{aligned} \tag{2.71}$$

It therefore follows

$$\int_{a}^{b} f(t') \delta(t - t') dt' = \begin{cases} f(t) & \text{if } a < t < b \\ 0 & \text{otherwise} \end{cases} \tag{2.72}$$

Another function of interest is the step function (also called Heaviside step function)

$$\theta(t) = \begin{cases} 1 & \text{if } t > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.73}$$

This implies

$$\int_{-\infty}^{t} \delta(t')\, dt' = \theta(t). \tag{2.74}$$

In this sense we can symbolically write that the δ-function is the derivative of the step function

$$\delta(t) = \frac{d}{dt}\theta(t). \tag{2.75}$$

We can also analyze the δ-function with argument $g(t)$. Then $\delta(g(t))$ is nonzero for those $t = t_i$ where $g(t_i) = 0$. This implies

$$\int_a^b dt\delta(g(t)) = \lim_{\epsilon \to 0} \sum_i \int_{g\left(t_i - \frac{\epsilon}{2}\right)}^{g\left(t_i + \frac{\epsilon}{2}\right)} dg \frac{1}{\frac{dg}{dt}} \delta(g) = \sum_i \frac{1}{\left|\frac{dg}{dt}\right|_{t=t_i}}, \tag{2.76}$$

where we sum over those t_i that are in the interval $[a, b]$.

2.1.7 Driven oscillators and Green's function

We will reconsider the problem of driven oscillators. As mentioned, we want to construct the response to an arbitrary external force $f(t)$ in terms of the response to $\delta(t - t')$. Thus we are analyzing the solution

$$\frac{d^2 x(t)}{dt^2} + \gamma \frac{dx(t)}{dt} + \omega_0^2 x(t) = \delta(t - t') \tag{2.77}$$

that corresponds to a force-pulse of unit strength at time t'. The solution to this equation with boundary condition

$$\begin{aligned} \lim_{\epsilon \to 0} x(t' - \epsilon) &= 0 \\ \lim_{\epsilon \to 0} \left.\frac{dx(t)}{dt}\right|_{t=t'-\epsilon} &= 0 \end{aligned} \tag{2.78}$$

is called the Green's function:

$$G(t, t') = x(t). \tag{2.79}$$

It is therefore at least one particular solution of the inhomogeneous equation, with δ-function inhomogeneity.

In addition holds that

$$G(t, t') = G(t - t'). \tag{2.80}$$

Suppose we shift in our differential equation $t' \to t' + \Delta t$. This shift can be absorbed into a shift in t, which in turn will not change the derivatives, i.e. we obtain

$$G(t, t' + \Delta t) = G(t - \Delta t, t') \tag{2.81}$$

The above dependence on the relative time follows immediately. Thus we analyze

$$\left(\frac{d^2}{dt^2} + \gamma \frac{d}{dt} + \omega_0^2\right) G(t - t') = \delta(t - t') \tag{2.82}$$

This equation easily allows us to analyze the behavior of $G(t - t')$ near $t = t'$. We integrate both sides over t from $t' - \epsilon$ to $t' + \epsilon$ and take the limit $\epsilon \to 0$:

$$\int_{t'-\epsilon}^{t'+\epsilon} dt \left(\frac{d^2}{dt^2} + \gamma \frac{d}{dt} + \omega_0^2\right) G(t - t') = \int_{t'-\epsilon}^{t'+\epsilon} dt \delta(t - t') \tag{2.83}$$

For the right hand side holds from the property of the δ-function

$$\int_{t'-\epsilon}^{t'+\epsilon} dt \delta(t - t') = 1$$

Now we assume that G is continuous and bounded at $t = t'$. It holds

$$\begin{aligned} \lim_{\epsilon \to 0} \int_{t'-\epsilon}^{t'+\epsilon} dt G(t - t') &= G(0) \lim_{\epsilon \to 0} \int_{t'-\epsilon}^{t'+\epsilon} dt = 0 \\ \lim_{\epsilon \to 0} \int_{t'-\epsilon}^{t'+\epsilon} dt \frac{d}{dt} G(t - t') &= \lim_{\epsilon \to 0} \left(G(\epsilon) - G(-\epsilon)\right) = 0 \end{aligned}$$

We can not fulfill the equation if the derivative of G is continuous as well, as

$$\lim_{\epsilon \to 0} \int_{t'-\epsilon}^{t'+\epsilon} dt \frac{d^2}{dt^2} G(t - t') = \lim_{\epsilon \to 0} \left(\left.\frac{dG(t)}{dt}\right|_{t=\epsilon} - \left.\frac{dG(t)}{dt}\right|_{t=-\epsilon} \right) \tag{2.84}$$

It then follows

$$\lim_{\epsilon \to 0} \left(\left.\frac{dG(t)}{dt}\right|_{t=\epsilon} - \left.\frac{dG(t)}{dt}\right|_{t=-\epsilon} \right) = 1 \tag{2.85}$$

The derivative of the Green's function must therefore have a jump at zero argument

Before we determine this Green's function we show that it allows for a determination of an arbitrary inhomogeneity. We start now from Eq.2.82 and multiply both sides with $f(t')$:

$$\left(\frac{d^2}{dt^2} + \gamma \frac{d}{dt} + \omega_0^2\right) G(t - t') f(t') = \delta(t - t') f(t') \tag{2.86}$$

Next we integrate both sides of the equation over t'.

$$\begin{aligned} \left(\frac{d^2}{dt^2} + \gamma \frac{d}{dt} + \omega_0^2\right) \int_{-\infty}^{\infty} dt' G(t - t') f(t') &= \int_{-\infty}^{\infty} dt' \delta(t - t') f(t') \\ &= f(t). \end{aligned} \tag{2.87}$$

The differential equation we have to solve is:

$$\left(\frac{d^2}{dt^2}+\gamma\frac{d}{dt}+\omega_0^2\right)x\left(t\right)=f\left(t\right). \tag{2.88}$$

Comparing both results yields that we have a particular solution

$$\bar{x}\left(t\right)=\int_{-\infty}^{\infty}dt'G\left(t-t'\right)f\left(t'\right). \tag{2.89}$$

Thus knowing the function $G\left(t-t'\right)$ allows immediately for a determination of the solution for an arbitrary inhomogeneity $f\left(t\right)$.

Let us now determine the Green's function. We first consider $t<t'$. In this limit holds $G\left(t-t'\right)=0$, as there is no force acting and the only solution of the homogeneous equation with simultaneously vanishing position and velocity is $x\left(t\right)=0$. The regime $t>t'$ can be obtained if we use our earlier general solution of the homogeneous problem

$$x_{hom}\left(t\right)=e^{-\frac{\gamma}{2}t}\left(a\cos\Omega t+b\sin\Omega t\right) \tag{2.90}$$

with $\Omega=\sqrt{\omega_0^2-\gamma^2/4}$. Since for $t>t'$ the force is again zero the Green's function must be of this form, only with $t\to t-t'$, i.e.

$$G\left(t-t'\right)=e^{-\frac{\gamma}{2}\left(t-t'\right)}\left(a\cos\left(\Omega\left(t-t'\right)\right)+b\sin\left(\Omega\left(t-t'\right)\right)\right) \tag{2.91}$$

In order to determine the boundary condition we use our earlier result for the jump of the derivative. From $G\left(t-t'\right)=0$ for $t<t'$ follows that the derivative for $t<t'$ vanishes as well, i.e. we have next to

$$\lim_{\epsilon\to 0}G\left(\epsilon\right)=0 \tag{2.92}$$

which follows from the continuity of the Green's function, the boundary condition:

$$\lim_{\epsilon\to 0}\left.\frac{dG\left(t\right)}{dt}\right|_{t=\epsilon}=1 \tag{2.93}$$

The first condition yields $a=0$ and with

$$\frac{dG\left(t\right)}{dt}=be^{-\frac{\gamma}{2}t}\left(\Omega\cos\Omega t-\frac{\gamma}{2}\sin\omega t\right) \tag{2.94}$$

follows $b=\frac{1}{\Omega}$. Thus we finally obtain

$$G\left(t-t'\right)=\theta\left(t-t'\right)\frac{e^{-\frac{\gamma}{2}\left(t-t'\right)}\sin\left(\Omega\left(t-t'\right)\right)}{\Omega} \tag{2.95}$$

We finally apply this result to our harmonic external force $f\left(t\right)=e^{-\omega t}$. It follows

$$\begin{aligned}\bar{x}(t) &= \int_{-\infty}^{\infty} dt' G(t-t') f(t') . \\ &= \int_{-\infty}^{\infty} dt' G(t-t') e^{i\omega t'} \\ &= \frac{1}{2i\Omega} \int_{-\infty}^{t} dt' e^{-\frac{\gamma}{2}(t-t')} \left(e^{i\Omega(t-t')} - e^{-i\Omega(t-t')} \right) e^{i\omega t'} \\ &= A(\omega) e^{i\omega t} \end{aligned} \tag{2.96}$$

The power of this new method is that one only has to determine the Green's function for a given homogeneous problem once. Then it's knowledge allows one to solve for an arbitrary inhomogeneity.

2.2 Integration of an arbitrary one dimensional potential, anharmonic oscillator

In case of a one dimensional motion of a single particle, the classical equations of motion can be integrated for an arbitrary potential. This approach is particularly interesting if one wants to analyze anharmonic oscillators. Since energy conservation we start from:

$$\frac{m}{2}\dot{x}^2 + U(x) = E \tag{2.97}$$

where the energy E plays the role of an integration constant. It follows

$$\frac{dx}{dt} = \pm\sqrt{\frac{2}{m}(E - U(x))}, \tag{2.98}$$

which already implies that only orbits with $U(x) \leq E$ are allowed. This makes of course sense as the total energy is the sum of the potential energy and the positive definite kinetic energy. In case of an equal sign we have a particle with zero velocity, i.e. vanishing kinetic energy. We write

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

and integrate this equation on both sides. The left hand side is trivial as

$$\int_{t_0}^{t} dt = t - t_0.$$

and it follows

$$t = t_0 \pm \sqrt{\frac{m}{2}} \int_{x(t_0)}^{x(t)} \frac{dx'}{\sqrt{E - U(x')}}, \tag{2.99}$$

where the sign is determined such that $t > t_0$, i.e. we take the upper sign if $x(t) > x(t_0)$ and the lower sign for $x(t) < x(t_0)$, respectively. The condition $U(x) = E$ determines turning points of the classical motion, as the kinetic energy at those positions vanishes (i.e. the velocity is zero). On the other hand, the particle cannot permanently rest at those positions (unless $\frac{\partial^2 U}{\partial x^2} = 0$ and no force acts), implying that the kinetic energy will only vanish for one time instant. This immediately allows for the determination of periods of closed orbits

$$T(E) = \sqrt{2m} \int_{x_-}^{x_+} \frac{dx'}{\sqrt{E - U(x')}}. \tag{2.100}$$

where $U(x_\pm) = E$ are the two turning points. Here, the period was obtained as twice the time from x_- to x_+. As an application we consider the anharmonic oscillator

$$U(x) = rx^n. \tag{2.101}$$

The turning points are given as

$$x_\pm = \pm \left(\frac{E}{r} \right)^{1/n} \tag{2.102}$$

and we find

$$\begin{aligned} T(E) &= \sqrt{2m} \int_{x_-}^{x_+} \frac{dx}{\sqrt{E - rx^n}}. \\ &= \frac{E^{1/n}}{r^{1/n}} \sqrt{\frac{2m}{E}} \int_{-1}^{1} \frac{dt}{\sqrt{1 - t^n}} \end{aligned}$$

The integral over t can be performed and yields the n-dependent constant $C_n = \sqrt{\pi}\Gamma\left(\frac{n+1}{n}\right)/\Gamma\left(\frac{2+n}{2n}\right)$ with Γ -function (one of the many special function relevant in many physics problems)

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$$

For example $C_2 = \pi$, $C_4 \approx 2.622$, $C_6 \approx 2.42865$ etc. The Γ-function with integer argument yields the factorial $\Gamma(m+1) = m! = 1 \cdot 2 \cdot 3 \cdot \cdots \cdot m$. Thus we obtain for the frequency

$$\omega = \frac{2\pi}{T} = \frac{2\pi}{C_n} \frac{r^{1/n}}{\sqrt{2m}} E^{\frac{1}{2} - \frac{1}{n}}. \tag{2.103}$$

In case of the harmonic oscillator ($n = 2$)we recover the known result $\omega = \sqrt{\frac{k}{m}}$ with $r = \frac{k}{2}$ and the above value for C_2. In case of an-harmonic oscillators, such as the case $n = 4$, we find that the frequency depends on the energy of the oscillator, where for $n > 2$ follows that ω increases with increasing energy.

Chapter 3

Mechanical motion in three dimensions: planetary orbits

As crucial example for a classical motion that takes place in three dimensions, we consider the planetary orbits around the sun. We will initially proceed in a phenomenological fashion and discuss Kepler's empirical laws. We will then determine the gravitational force that naturally follows from those laws. In a second step we will solve Newton's equation of motion with a gravitational force. Among others we will then derive the planetary orbits that are fully consistent with Kepler's laws.

3.1 Kepler's three laws

Johannes Keppler (1571-1630) deduced three statements from the observational data of the planetary motion in our solar system obtained by the danish astronomer Tycho Brahe (1546-1601).

1. All planets move on ellipses. The sun is located in one of its foci.

2. A line joining a planet and the Sun sweeps out equal areas during equal intervals of time.

3. The square of the orbital period divided by the cube of the semi-major axis of its orbit is a the same constant for all planets in the solar system.

In what follows we will, after a brief summary of the property of ellipses, analyze the implications of these observations.

3.1.1 Mathematical tools 8: ellipses

We summarize the basic facts of an ellipse. We have two foci F and F' at a distance $2c$. The ellipse is generated by all points with $r + r' = 2a$ (see figure), where $a^2 = b^2 + c^2$. Here a is the semi-major axis and b the semi-minor axis of the ellipse. We introduce the eccentricity of the ellipse

$$\epsilon = \frac{c}{a} = \frac{\sqrt{a^2 - b^2}}{a} \tag{3.1}$$

In case of a circle we have $\epsilon = 0$ and both foci merge at the center of the circle. We can express r' as

$$r' = \sqrt{(2c)^2 + r^2 + 2\,(2c)\cos\theta} \tag{3.2}$$

From $r + r' = 2a$ follows

$$(2a - r)^2 = 4\epsilon^2 a^2 + r^2 + 4\epsilon a r \cos\theta \tag{3.3}$$

The terms with r^2 cancel on both sides

$$4a^2 - 4ar = 4\epsilon^2 a^2 + 4\epsilon a r \cos\theta \tag{3.4}$$

and we obtain the following expression of an ellipse in polar coordinates

$$r = \frac{k}{1 + \epsilon\cos\theta}, \tag{3.5}$$

where $k = a\left(1 - \epsilon^2\right)$. The limit of a circle follows for $\epsilon = 0$ with $r = a$.

The area of the ellipse is

$$\begin{aligned} A &= 4\int_0^{\pi/2} d\theta \int_0^{r(\theta)} r dr \\ &= 2k^2 \int_0^{\pi/2} d\theta \frac{1}{(1 + \epsilon\cos\theta)^2} \\ &= \pi ab. \end{aligned}$$

3.1.2 Kepler's first law

One implication of Kepler's first law is that the motion of planets is in a given plane. In other words, there exists a vector that is perpendicular to this plane and that does not change in time. This vector is obviously orthogonal to the position vector $\mathbf{r}$ and to the change of the position vector, i.e. the velocity or momentum $\mathbf{p} = m\mathbf{v}$. Thus we define the angular momentum vector

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}. \tag{3.6}$$

Kepler's first law implies that

$$\frac{d\mathbf{L}}{dt} = 0. \tag{3.7}$$

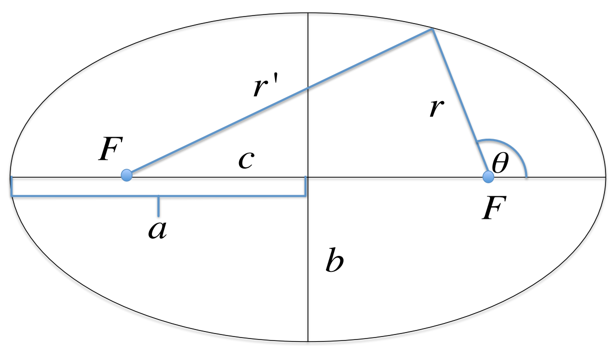


Figure 3.1: Ellipse with semi-major axis a and and semi-minor axis b. F and F' are the foci of the ellipse, c the distance between the center and the foci.

From the above definition follows

$$\begin{aligned} \frac{d\mathbf{L}}{dt} &= m\frac{d\mathbf{r}}{dt} \times \mathbf{v} + m\mathbf{r} \times \frac{d\mathbf{v}}{dt} \\ &= \mathbf{r} \times \mathbf{F}. \end{aligned} \quad (3.8)$$

The vector $\mathbf{D} = \frac{d\mathbf{L}}{dt}$ is called the torque. In case of nontrivial situations (i.e. $\mathbf{F}$ finite) implies the condition of constant angular momentum that the cross product of the force and the position vanishes everywhere. Thus the force itself is proportional to the position vector

$$\mathbf{F}(\mathbf{r}) = \Phi(\mathbf{r})\,\mathbf{e_r} \quad (3.9)$$

Since

$$\begin{aligned} \mathbf{F} &= -\nabla V \\ &= -\frac{\partial V}{\partial r}\mathbf{e}_r - \frac{1}{r}\frac{\partial V}{\partial \theta}\mathbf{e}_\theta - \frac{1}{r\sin\theta}\frac{\partial V}{\partial \varphi}\mathbf{e}_\varphi, \end{aligned} \quad (3.10)$$

it follows that a force field that only points towards or away from the origin $\mathbf{r} = \mathbf{0}$, follows from a potential that depends only on the magnitude $r = |\mathbf{r}|$ of the vector $\mathbf{r}$, not its direction. Thus, any potential of the form

$$V(\mathbf{r}) = V(r) \quad (3.11)$$

leads to a conservation of angular momentum. This conservation of angular momentum is a consequence of the isotropy of space, i.e. there is no preferred direction in this potential. The gravitational potential should therefore only depend on the distance to the origin. This will guarantee a planar motion.

3.1.3 Kepler's second law

The change in area of the line joining a planet and the Sun that follows from the motion from $\mathbf{r}$ to $\mathbf{r} + d\mathbf{r}$ is

$$dA = \frac{1}{2} |\mathbf{r} \times d\mathbf{r}| \tag{3.12}$$

This allows to determine the change of this area as function of time

$$\begin{aligned} \frac{dA}{dt} &= \frac{1}{2} |\mathbf{r} \times \mathbf{v}| = \frac{1}{2m} |\mathbf{r} \times \mathbf{p}| \\ &= \frac{1}{2m} |\mathbf{L}| \end{aligned} \tag{3.13}$$

Since the angular momentum is conserved (does not change as function of time), the change in area of the line joining the planet and the Sun is a constant of motion as well. Thus, we conclude that key aspects of the first and second Kepler's law is angular momentum conservation.

We put our plane without restriction into the $x - y$ plane and write

$$\begin{aligned} \mathbf{e}_r &= \cos\theta \mathbf{e}_x + \sin\theta \mathbf{e}_y \\ \mathbf{e}_\theta &= -\sin\theta \mathbf{e}_x + \cos\theta \mathbf{e}_y \end{aligned} \tag{3.14}$$

It follows

$$\begin{aligned} \frac{d\mathbf{e}_r}{dt} &= \frac{d\theta}{dt}\mathbf{e}_\theta \\ \frac{de_\theta}{dt} &= -\frac{d\theta}{dt}\mathbf{e}_r \end{aligned} \tag{3.15}$$

We use these relations to determine the position, velocity and acceleration

$$\begin{aligned} \mathbf{r} &= r\mathbf{e_r} \\ \mathbf{v} &= \frac{dr}{dt}\mathbf{e}_r + r\frac{d\mathbf{e}_r}{dt} = \frac{dr}{dt}\mathbf{e}_r + r\frac{d\theta}{dt}\mathbf{e}_\theta \\ \mathbf{a} &= \left(\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2\right)\mathbf{e}_r + \left(2\frac{dr}{dt}\frac{d\theta}{dt} + r\frac{d^2\theta}{dt^2}\right)\mathbf{e}_\theta \end{aligned} \tag{3.16}$$

The area-velocity now follows as

$$\frac{dA}{dt} = \frac{1}{2}r^2\frac{d\theta}{dt} \equiv \frac{1}{2}h \tag{3.17}$$

with constant h.

Now we need to take into account the actual fact that the motion is on ellipses. Since we know already that the force in only along the radial direction we can write

$$\mathbf{F} = m\left(\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2\right)\mathbf{e}_r \tag{3.18}$$

The variation of $r(t)$ for an ellipse is

$$\begin{aligned} \frac{dr}{dt} &= \frac{dr}{d\theta}\frac{d\theta}{dt} = \frac{\epsilon}{k}\sin\theta r^2\frac{d\theta}{dt} = \frac{\epsilon}{k}h\sin\theta \\ \frac{d^2r}{dt^2} &= \frac{\epsilon}{k}h\cos\theta\frac{d\theta}{dt} = \frac{\epsilon h^2}{kr^2}\cos\theta \end{aligned} \tag{3.19}$$

such that

$$\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 = \frac{h^2}{r^2}\left(\frac{\epsilon}{k}\cos\theta - \frac{1}{r}\right) = -\frac{h^2}{kr} \tag{3.20}$$

Thus, we conclude that the form of the gravitational force is

$$\mathbf{F}(\mathbf{r}) = -\frac{h^2}{kr^2}m\frac{\mathbf{r}}{r} \tag{3.21}$$

At this point we we should check whether indeed the force component along $\mathbf{e}_\theta$ vanish. It follows from

$$\frac{d\theta}{dt} = \frac{h}{r^2} \tag{3.22}$$

that

$$\frac{d^2\theta}{dt^2} = -2\frac{h}{r^3}\frac{dr}{dt} = -\frac{2}{r}\frac{d\theta}{dt}\frac{dr}{dt} \tag{3.23}$$

such that

$$2\frac{dr}{dt}\frac{d\theta}{dt} + r\frac{d^2\theta}{dt^2} = 0 \tag{3.24}$$

which demonstrated that the component of the force along $\mathbf{e}_\theta$ vanishes, as expected.

3.1.4 Kepler's third law

Based on the first and second law, the coefficient could be different for each planet (i.e. depend on its mass). However we can use the fact that the ratio T^2/a^3 is the same for all planets.

Since $\frac{1}{2}h$ is the area-velocity it follows for the period

$$\frac{h}{2}T = \pi ab = \pi\sqrt{a^3k} \tag{3.25}$$

(we used $b^2 = a^2 - c^2 = a^2\left(1 - \epsilon^2\right) = ak$), such that

$$\frac{T^2}{a^3} = \frac{4\pi^2}{h^2}k. \tag{3.26}$$

Thus, it follows that the coefficient h^2/k in the gravitational force is independent on mass. This is the force of the sun onto a planet. In reverse, there should also be a force of the planet onto the sun (actio=reactio), i.e. since the force coefficient is linear in m (the mass of the planet) it should also be linear in M (the mass of the Sun) and we write

$$\frac{h^2}{k} = GM \tag{3.27}$$

with some gravitational constant G. It finally follows

$$\mathbf{F}(\mathbf{r}) = -G\frac{Mm}{r^2}\frac{\mathbf{r}}{r}. \tag{3.28}$$

The gravitational constant is

$$G \backsimeq 6.67 \times 10^{-11}\mathrm{N}(\mathrm{m}/\mathrm{kg})^2. \tag{3.29}$$

3.2 Solution of the equations of motion for planetary motion

In what follows we solve for the equation of motion of two bodies (for example a planet and the sun, the earth and the moon) in their mutual gravitational field. In our derivation of the gravitational force

$$\mathbf{F}(\mathbf{r}) = -G\frac{Mm}{r^2}\frac{\mathbf{r}}{r}. \tag{3.30}$$

we implicitly assumed that the larger body (the sun) was fixed somewhere in space and that the lighter body moves around it. If this is the entire story it implies that there is no gravitational back-action of the earth on the sun. Such effects must however be included in a proper theory of gravity. The potential that leads to the above force is

$$V(\mathbf{r}) = -G\frac{Mm}{|\mathbf{r}|}. \tag{3.31}$$

Clearly it is more appropriate to write this as

$$V(\mathbf{r}_1 - \mathbf{r}_2) = -G\frac{Mm}{|\mathbf{r}_1 - \mathbf{r}_2|}, \tag{3.32}$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ are the coordinates of the two celestial bodies. Before we solve this problem we discuss the general concept of relative and center of gravity coordinates that allows us to simplify the motion of two bodies significantly.

3.2.1 relative and center of gravity coordinates

Let us consider two particles with coordinates $\mathbf{r}_1$ and mass m_1 as well as $\mathbf{r}_2$ and m_2, respectively, that interact via a potential $V(\mathbf{r}_1 - \mathbf{r}_2)$. The force acting on particle 1 (caused by particle 2) is

$$\mathbf{F}_1 = -\nabla_{\mathbf{r}_1} V(\mathbf{r}_1 - \mathbf{r}_2) \tag{3.33}$$

On the other hand, the force on particle 1 (caused by particle 2) is

$$\begin{aligned}\mathbf{F}_2 &= -\nabla_{\mathbf{r}_2} V(\mathbf{r}_1 - \mathbf{r}_2) \\ &= -\mathbf{F}_1.\end{aligned} \tag{3.34}$$

The Newtonian equations of motion for the two particles are

$$\begin{aligned}m_1 \frac{d^2\mathbf{r}_1}{dt^2} &= \mathbf{F}_1 \\ m_2 \frac{d^2\mathbf{r}_2}{dt^2} &= \mathbf{F}_2.\end{aligned} \tag{3.35}$$

Since the potential only depends on the relative coordinate

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \tag{3.36}$$

we first analyze the equation of motion for $\mathbf{r}$:

$$\begin{aligned}\frac{d^2\mathbf{r}}{dt^2} &= \frac{\mathbf{F}_1}{m_1} - \frac{\mathbf{F}_2}{m_2} = -\left(\frac{1}{m_1} + \frac{1}{m_2}\right) \nabla_{\mathbf{r}_1} V(\mathbf{r}_1 - \mathbf{r}_2) \\ &= -\frac{1}{\mu} \nabla_{\mathbf{r}} V(\mathbf{r})\end{aligned} \tag{3.37}$$

where we introduced the reduced mass μ via $\mu^{-1} = m_1^{-1} + m_2^{-1}$ such that

$$\mu = \frac{m_1 m_2}{m_1 + m_2}. \tag{3.38}$$

In the case where one mass is a lot larger than the other, the reduced mass is dominated by the smaller one. Thus, the equation of motion for the relative coordinate of two planets (or other mechanical objects) corresponds to the dynamics of a single particle around a fixed center. This justifies a posteriori our earlier treatment of the gravitational problem. It also holds that the relative coordinate decouples completely from the equation of the center of gravity

$$\mathbf{R} = \frac{m_1}{m_1 + m_2} \mathbf{r}_1 + \frac{m_2}{m_1 + m_2} \mathbf{r}_2 \tag{3.39}$$

It holds

$$\frac{d^2\mathbf{R}}{dt^2} = \frac{1}{m_1 + m_2} \mathbf{F}_1 + \frac{1}{m_1 + m_2} \mathbf{F}_2 = 0, \tag{3.40}$$

which corresponds to the free motion.

We can generalize these consideration to a problem where two particles interact via a potential $V(\mathbf{r}_1 - \mathbf{r}_2)$ that depends on the relative coordinate and both are in an external potential that acts on the center of gravity coordinate $U(\mathbf{R})$. This could be realized by the potential caused by the center of our galaxy. The equation for the relative coordinate is unchanged, while the center of gravity coordinate obeys

$$(m_1 + m_2)\frac{d^2\mathbf{R}}{dt^2} = -\nabla_{\mathbf{R}} U(\mathbf{R}). \tag{3.41}$$

The center of gravity dynamics is now governed by the potential U and it appears as a single particle with mass $M = m_1 + m_2$. This total mass is in case of very different masses dominated by the heavier (in distinction to the reduced mass).

3.2.2 Solution of the equation of motion for the relative coordinate

In this section we solve (we use for convenience m to refer to the reduced mass):

$$m\frac{d^2\mathbf{r}}{dt^2} = -\nabla_{\mathbf{r}} V(\mathbf{r}) \tag{3.42}$$

with

$$V(\mathbf{r}) = -G\frac{mM}{r}. \tag{3.43}$$

Since the potential is independent on the direction of $\mathbf{r}$ we know that the force $\mathbf{F} \propto \mathbf{r}$ is proportional to the position vector. If we consider the angular momentum

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \tag{3.44}$$

we find that

$$\frac{d\mathbf{L}}{dt} = \mathbf{r} \times \mathbf{F} = \mathbf{0}, \tag{3.45}$$

i.e. the angular momentum is conserved. It immediately follows that the motion is confined to a plane, where $\mathbf{L}$ points orthogonal to this plane. Without restriction we use $\mathbf{r} = (r\cos\theta, r\sin\theta, 0)$ and use our earlier results:

$$\begin{aligned}
\mathbf{r} &= r\mathbf{e}_{\mathbf{r}} \\
\mathbf{v} &= \frac{dr}{dt}\mathbf{e}_r + r\frac{d\mathbf{e}_r}{dt} = \frac{dr}{dt}\mathbf{e}_r + r\frac{d\theta}{dt}\mathbf{e}_\theta \\
\mathbf{a} &= \left(\frac{d^2 r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2\right)\mathbf{e}_r + \left(2\frac{dr}{dt}\frac{d\theta}{dt} + r\frac{d^2\theta}{dt^2}\right)\mathbf{e}_\theta
\end{aligned} \tag{3.46}$$

For the angular momentum follows immediately

$$\mathbf{L} = mr^2\frac{d\theta}{dt}\mathbf{e}_{\mathbf{r}} \times \mathbf{e}_\theta = mr^2\frac{d\theta}{dt}\mathbf{e}_z \tag{3.47}$$

The magnitude $L = mr^2 \frac{d\theta}{dt}$ of the angular momentum is therefore conserved as well. It holds

$$\frac{d\theta}{dt} = \frac{L}{mr^2} \tag{3.48}$$

where L is a constant, fixed by the initial velocity and position. If we consider the components of the force along $\mathbf{e_r}$, Newton's law reduces to

$$m\left(\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2\right) = -G\frac{Mm}{r^2} \tag{3.49}$$

Instead of writing this as equation $r(t)$ we rather want to find an equation $r(\theta)$ that determines the radius at given angle. It holds

$$\frac{dr}{dt} = \frac{dr}{d\theta}\frac{d\theta}{dt} = \frac{L}{mr^2}\frac{dr}{d\theta} \tag{3.50}$$

In addition we consider a differential equation for

$$s = \frac{1}{r} \tag{3.51}$$

with

$$\frac{dr}{d\theta} = -\frac{1}{s^2}\frac{ds}{d\theta} \tag{3.52}$$

such that

$$\frac{dr}{dt} = -\frac{L}{m}\frac{ds}{d\theta} \tag{3.53}$$

which allows us to write

$$\begin{aligned} \frac{d^2r}{dt^2} &= -\frac{L}{m}\frac{d^2s}{d\theta^2}\frac{d\theta}{dt} \\ &= -\left(\frac{L}{m}\right)^2 s^2\frac{d^2s}{d\theta^2} \end{aligned} \tag{3.54}$$

We insert this in the equation of motion and obtain

$$m\left(-\left(\frac{L}{m}\right)^2 s^2\frac{d^2s}{d\theta^2} - s^3\left(\frac{L}{m}\right)^2\right) = -GMms^2 \tag{3.55}$$

which can further be simplified as

$$\frac{d^2s}{d\theta^2} + s = s_0 \tag{3.56}$$

where $s_0 = GM\frac{m^2}{L^2}$. This is the well known differential equation for the harmonic oscillator. We need to know the solution of the homogeneous equation

$$s_{hom} = C\cos(\theta + \theta_0) \tag{3.57}$$

and a special solution of the inhomogeneous equation. For the latter we can simply chose $s_{inh} = s_0$ independent on θ. Thus, we find (assuming without restriction that $\theta_0 = 0$

$$s = s_0 + C\cos(\theta) \tag{3.58}$$

Returning to our original variable we obtain

$$r = \frac{k}{1 + \epsilon\cos(\theta)} \tag{3.59}$$

with

$$k = \frac{L^2}{GMm^2} \text{ and } \epsilon = \frac{C}{s_0} = \frac{CL^2}{GMm^2}. \tag{3.60}$$

This is the expected equation for an ellipse. The solution also allows for an explicit expression of the typical size of the ellipse k in terms of the angular momentum, masses and the gravitational constant. The eccentricity ϵ of the ellipse is determined by the integration constant (let us recall that this is really only an ellipse if $0 \le \epsilon < 1$). To determine the integration constant we use the total energy

$$\begin{aligned} E &= \frac{m}{2}\left(\frac{d\mathbf{r}}{dt}\right)^2 + V(r) \\ &= \frac{m}{2}\left(\frac{dr}{dt}\right)^2 + \frac{mr^2}{2}\left(\frac{d\theta}{dt}\right)^2 - G\frac{mM}{r} \\ &= \frac{m}{2}\left(\frac{dr}{dt}\right)^2 + \frac{L^2}{2mr^2} - G\frac{mM}{r} \end{aligned} \tag{3.61}$$

We express this result in terms of our solution

$$\begin{aligned} E &= \frac{L^2}{2m}\left(\left(\frac{ds}{d\theta}\right)^2 + s^2\right) - GmMs \\ &= \frac{L^2}{2m}\left(\left(\frac{ds}{d\theta}\right)^2 + (s - s_0)^2\right) - \frac{L^2}{2m}s_0^2 \\ &= \frac{L^2}{2m}C^2 - \frac{G^2M^2m^3}{2L^2}. \end{aligned} \tag{3.62}$$

such that

$$\epsilon = \frac{\sqrt{2m\left(E + \frac{G^2M^2m^3}{2L^2}\right)}L}{GMm^2} \tag{3.63}$$

To interpret this result we return to Eq.3.61 which corresponds to a one dimensional motion in the effective potential

$$V_{\text{eff}}(r) = \frac{L^2}{2mr^2} - G\frac{mM}{r} \tag{3.64}$$

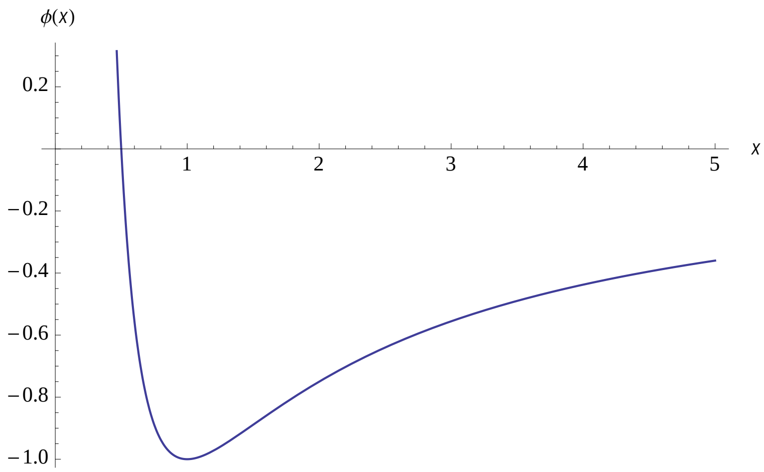


Figure 3.2: Effective potential $V_{\text{eff}}(r) = |E_0|\,\phi(r/r_0)$

The potential has a minimum at $r = r_0 = \frac{L^2}{Gm^2 M}$ (which corresponds to s_0^{-1}). The potential at this minimum is

$$E_0 = V_{\text{eff}}(r_0) = -\frac{G^2 m^3 M^2}{2L^2} \tag{3.65}$$

Thus, we can write

$$V_{\text{eff}}(r) = |E_0|\,\phi\left(\frac{r}{r_0}\right) \tag{3.66}$$

with $\phi(x) = \frac{1}{x^2} - \frac{2}{x}$ shown in Fig.3.2. We can now write for the eccentricity ∂

$$\epsilon = \sqrt{\frac{E - E_0}{|E_0|}} \tag{3.67}$$

It is obviously not possible to have an energy below E_0. For $E = E_0$, it must hold that $dr/dt = 0$. A constant radius corresponds of course to the motion on a circle. In this case follows $\epsilon(E = E_0) = 0$ as expected for a circular orbit. As long as $E_0 \leq E < 0$ we have $0 \leq \epsilon < 1$ and our orbit corresponds to an ellipse. In our effective potential the ficticious particle cannot escape to infinity as its energy is not sufficient. However, for $E > 0$ the eccentricity is larger than unity and the orbit becomes a hyperbole instead of an ellipse. Now the object can escape to infinity. This finally yields a full characterization of all potential orbits in terms of the angular momentum and energy. Both are fully determined by the initial velocity and position.

3.2.3 Perihel-rotation and Lenz-vector

Our analysis of planetary motion was based on the potential $V(r) \propto \frac{1}{r}$. The form of the potential was a consequence of Kepler's laws, i.e. based on experi-

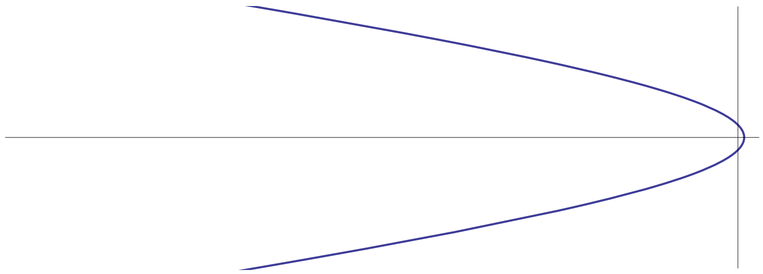


Figure 3.3: Hyperbolic orbit for $\epsilon > 1$. Here we chose $\epsilon = 1.1$.

mental observation. Experimental data always come with some uncertainty. It is therefore legitimate to ask what would happen if the gravitational potential would be not precisely governed by the $1/r$-law. Such an analysis is also of fundamental theoretical interest. It potentially reveals unique aspects of the pure $1/r$-law that one might overlook otherwise. We will therefore analyze the potential

$$V(r) = -\frac{GmM}{r}\left(1 + \frac{\eta}{r}\right) \tag{3.68}$$

i.e. we include a small correction $\propto 1/r^2$. Given the success of the conventional gravitational theory, we suspect already that η is a small quantity. Since it has dimension length, it seems reasonable to expect $\eta \ll r_0 = \frac{L^2}{Gm^2M}$ where r_0 is the typical length scale of the gravitational orbit. This corrected gravitation potential is not what happens in nature. However, the formulation of the general theory of relativity revealed that there are indeed corrections to the theory of gravity. Our little toy model and general relativity even have one aspect in common that we want to explore next, the rotation of the perihel.[1] The force associated with this potential can in case of a central force be obtained via

$$\begin{aligned} \mathbf{F}(\mathbf{r}) &= -\frac{dV(r)}{dr}\mathbf{e_r} \\ &= -G\frac{Mm}{r^2}\left(1 + \frac{2\eta}{r}\right) \end{aligned} \tag{3.69}$$

In order to solve this problem we modify our earlier equation of motion

$$m\left(\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2\right) = -G\frac{Mm}{r^2}\left(1 + \frac{2\eta}{r}\right) \tag{3.70}$$

We transform from $r(t)$ to $s(\theta)$, where $s = 1/r$ and obtain

[1] The perihel is the point on the orbit of a planet that is closest to the sun. The earth passes its perihel each year between the 3rd and 5th of January.

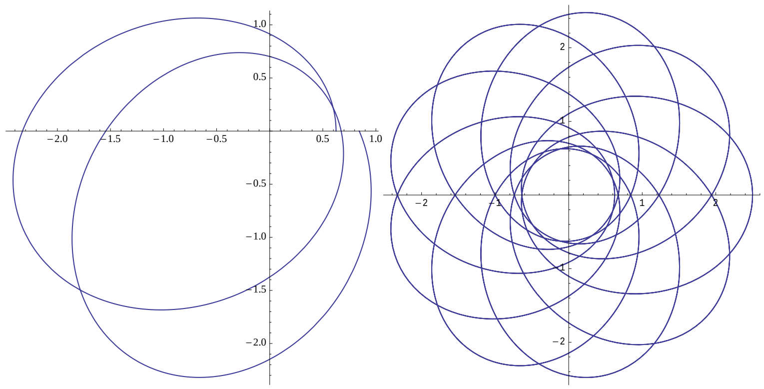


Figure 3.4: Orbit of a fictitious gravitational potential with additional $\frac{1}{r^2}$ contribution to the potential. Note, orbits are no longer closed. The left panel shows the path after two orbits, the right panel after 20 orbits. We used $\epsilon = 0.6$ and $\alpha = 0.9$.

$$m\left(-\left(\frac{L}{m}\right)^2 s^2\frac{d^2s}{d\theta^2} - s^3\left(\frac{L}{m}\right)^2\right) = -GMms^2\left(1+2\eta s\right) \tag{3.71}$$

This simplifies to

$$\frac{d^2s}{d\theta^2} + s\left(1-\frac{2\eta}{r_0}\right) = \frac{1}{r_0} \tag{3.72}$$

The solution of this differential equation is once again the sum of the homogeneous solution and a special inhomogeneous solution. The former is (with the choice that we can always chose the initial angle equal zero):

$$s_{hom} = C\cos\left(\sqrt{1-\frac{2\eta}{r_0}}\theta\right) \tag{3.73}$$

For the inhomogeneous solution we use again a constant: $s_{inh} = \frac{1}{r_0\left(1-\frac{2\eta}{r_0}\right)}$. We then find for the orbit

$$r\left(\theta\right) = \frac{r_0\left(1-\frac{2\eta}{r_0}\right)}{1+\epsilon\cos\left(\alpha\theta\right)} \tag{3.74}$$

where

$$\alpha = \sqrt{1-\frac{2\eta}{r_0}}. \tag{3.75}$$

This is no longer an ellipse. Plotting the orbit in polar coordinates also demonstrates that we don't have a closed orbit any longer. This is shown in Fig.3.4.

The perihel can be obtained from

$$\frac{dr}{d\theta} = \frac{r_0\left(1-\frac{2\eta}{r_0}\right)\alpha\epsilon\sin(\alpha\theta)}{\left(1+\epsilon\cos(\alpha\theta)\right)^2} = 0 \tag{3.76}$$

i.e. for those angles where $\alpha\theta = 2\pi n$ (note solutions with $\alpha\theta = (2n+1)\pi$ correspond to the aphel, the farthest point from the sun). For $\alpha = 1$ the perihel always occurs for the same angle. In the case of $\alpha < 1$ the shift in the perihel direction is

$$\delta_p = \frac{2\pi}{\alpha} - \frac{2\pi}{1} = \frac{1-\alpha}{\alpha}2\pi. \tag{3.77}$$

A shift in the perihel was indeed observed in case of mercury in the middle of the 19th century. The observed value was approximately 530 arc seconds[2] (modern value 571.91 arc second). This corresponds to $\alpha = 0.999559$, i.e. a very tiny effect (vastly exaggerated in our plots). The current explanation of the perihel rotation is subtle. 280 arc second are due to the gravitational force of the nearby Venus. 150 arc second are due to Jupiter and about 100 arc seconds due to the other planets. All this still doesn't account for the observed number. The modern value for the discrepancy is 43.11 arc seconds. This discrepancy can be explained as being due to effects caused by general relativity. In our theory such a description amounts to $a = 0.999967$. With a distance earth sun of $597,870,700.00$ km follows nevertheless for the length scale $\eta = 19,886.5$km.

While our corrected potential is not a correct description for the relativistic correction to gravity it revealed that the closed orbits that we found in case of the $1/r$ law are rather special. We already learned that whenever something "special" happens in physics, chances are that a conservation law is behind it. Indeed. The closed orbits of the Newtonian gravity are "protected" by another conserved quantity, the so called Lenz vector

$$\mathbf{A} = \mathbf{p}\times\mathbf{L} - Gm^2M\frac{\mathbf{r}}{r} \tag{3.78}$$

Let us analyze the time dependence of $\mathbf{A}$:

$$\frac{d\mathbf{A}}{dt} = \frac{d\mathbf{p}}{dt}\times\mathbf{L} + \mathbf{p}\times\frac{d\mathbf{L}}{dt} - Gm^2M\left(\frac{1}{r}\frac{d\mathbf{r}}{dt} - \frac{\mathbf{r}}{r^2}\frac{dr}{dt}\right) \tag{3.79}$$

It holds

$$\frac{d\mathbf{r}\cdot\mathbf{r}}{dt} = 2\mathbf{r}\cdot\frac{d\mathbf{r}}{dt} = 2r\frac{dr}{dt} \tag{3.80}$$

such that

$$\frac{dr}{dt} = \frac{\mathbf{r}}{r}\cdot\frac{d\mathbf{r}}{dt} \tag{3.81}$$

We use next that

$$\frac{d\mathbf{p}}{dt} = -GmM\frac{\mathbf{r}}{r^3} \tag{3.82}$$

[2] An arc second is $\frac{1}{3600}$ of a degree.

and $\frac{d\mathbf{L}}{dt} = 0$ and obtain

$$\frac{d\mathbf{A}}{dt} = -\frac{Gm^2M}{r^3}\left[\mathbf{r}\times\left(\mathbf{r}\times\frac{d\mathbf{r}}{dt}\right)+\left(\mathbf{r}\cdot\mathbf{r}\frac{d\mathbf{r}}{dt}-\mathbf{r}\mathbf{r}\cdot\frac{d\mathbf{r}}{dt}\right)\right] \tag{3.83}$$

It holds for two vectors $\mathbf{a}$ and $\mathbf{b}$ that (we use a notation where we sum over repeated indices):

$$\mathbf{a}\times(\mathbf{a}\times\mathbf{b}) = \mathbf{e}_\mu\epsilon_{\mu\nu\gamma}a_\nu\epsilon_{\alpha\beta\gamma}a_\alpha b_\beta \tag{3.84}$$

It holds

$$\begin{aligned}\epsilon_{\mu\nu\gamma}\epsilon_{\alpha\beta\gamma} &= \epsilon_{\mu\nu\gamma}\epsilon_{\gamma\alpha\beta}\\ &= \delta_{\mu\alpha}\delta_{\nu\beta}-\delta_{\mu\beta}\delta_{\nu\alpha}\end{aligned} \tag{3.85}$$

It follows

$$\begin{aligned}\mathbf{a}\times(\mathbf{a}\times\mathbf{b}) &= \mathbf{e}_\mu\left(\delta_{\mu\alpha}\delta_{\nu\beta}-\delta_{\mu\beta}\delta_{\nu\alpha}\right)a_\nu a_\alpha b_\beta\\ &= \mathbf{e}_\mu a_\beta a_\mu b_\beta - \mathbf{e}_\mu a_\alpha a_\alpha b_\mu\\ &= \mathbf{a}\mathbf{a}\cdot\mathbf{b}-\mathbf{b}\mathbf{a}\cdot\mathbf{a}\end{aligned} \tag{3.86}$$

This lead to

$$\mathbf{r}\times\left(\mathbf{r}\times\frac{d\mathbf{r}}{dt}\right) = \mathbf{r}\mathbf{r}\cdot\frac{d\mathbf{r}}{dt}-\mathbf{r}\cdot\mathbf{r}\frac{d\mathbf{r}}{dt}, \tag{3.87}$$

which finally implies that

$$\frac{d\mathbf{A}}{dt} = 0. \tag{3.88}$$

The Lenz vector $\mathbf{A}$ is therefore another conserved quantity for the gravity problem with $1/r$ potential. Let us determine this vector for the elliptical orbit. We use

$$\begin{aligned}\mathbf{L} &= L\mathbf{e}_z\\ \mathbf{p} &= m\mathbf{v} = m\left(\frac{dr}{dt}\mathbf{e}_r + r\frac{d\theta}{dt}\mathbf{e}_\theta\right)\\ &= m\left(\frac{dr}{dt}\mathbf{e}_r+\frac{L}{mr}\mathbf{e}_\theta\right)\end{aligned} \tag{3.89}$$

and obtain

$$\begin{aligned}\mathbf{A} &= m\left(\frac{dr}{dt}\mathbf{e}_r+\frac{L}{mr}\mathbf{e}_\theta\right)\times L\mathbf{e}_z - Gm^2M\frac{\mathbf{r}}{r}\\ &= -m\frac{dr}{dt}\mathbf{e}_\theta+\left(\frac{L^2}{r}-Gm^2M\right)\mathbf{e}_r,\end{aligned} \tag{3.90}$$

where we used that $\mathbf{e}_r\times\mathbf{e}_z = -\mathbf{e}_\theta$ and $\mathbf{e}_\theta\times\mathbf{e}_z = \mathbf{e}_r$. Since this vector is conserved we can evaluate it at any given time point and it must stay the same. We chose the point at the perihel where the planet is is closest to the sun. This is the point where r has a minimum, i.e. $\frac{dr}{dt} = 0$ which is $r = r_p$. For our choice of the coordinate system $\mathbf{e}_r = \mathbf{e}_x$, i.e. the Lenz vector points directly towards the perihel. Since $\mathbf{A}$ is conserved, the perihel cannot rotate and the orbit is closed. As in other cases discussed in this lecture, the conservation of the Lenz vector is related to a symmetry. In this case it is the transformation

$$t \to \lambda^3 t \;, \mathbf{r} \to \lambda^2 \mathbf{r}, \text{ and } \mathbf{p} \to \lambda^{-1} \mathbf{p}. \tag{3.91}$$

While this scale transformation rescales the energy $E \to \lambda^{-2} E$ and the angular momentum $L \to \lambda L$, the combination $L^2 E$ that determines the eccentricity and the Lenz vector remain unchanged.