An Introduction to Analytical Mechanics

MARTIN CEDERWALL

Institute for Theoretical Physics
Göteborg University and Chalmers University of Technology
Göteborg, Sweden

3rd corrected edition, Göteborg 1997

Preface

This compendium is intended to be a complement to the textbook *"An Introduction to Mechanics"* by D. Kleppner and R.J Kolenkow (KK) for the course "Mekanik F del B" given in the first year of the Master of Science program for Physical Engineering (Teknisk Fysik) at Chalmers University of Technology, Gothenburg.

Apart from what is contained in KK, this course also encompasses an elementary understanding of analytical mechanics, especially the lagrangian formulation. In order not to be too narrow, this text contains not only what is specified as compulsory for the Master of Science program, but tries to give a somewhat more general overview of the subject of analytical mechanics. The intention is that an interested student should be able to read additional material that may be useful in more advanced courses or simply interesting by itself.

The compulsory part of the text is the part on the lagrangian formulation of newtonian mechanics and its applications (Chapters 1-5), together with the part on variational (action) principles (Chapter 6). The chapter on the hamiltonian formulation is *not* compulsory, but it is recommended for the student who wants a deeper theoretical understanding of the subject and is very relevant for the connection between classical mechanics ("classical" here denoting both Newton's and Einstein's theories) and quantum mechanics.

The mathematical rigour is kept at a minimum, hopefully for the benefit of physical understanding and clarity. Notation is consistent with KK, unless explicitely stated.

The text is organized as follows: In Chapter 1 a background is given. Chapters 2, 3 and 4 contain the general setup needed for the lagrangian formalism. In Chapter 5 Lagrange's equation are derived and Chapter 6 gives their interpretation in terms of an action. Chapters 7 and 8 contain further developments of analytical mechanics, namely the hamiltonian formulation and a lagrangian treatment of constrained systems. Exercises are given at the end of each chapter. Finally, a translation table from English to Swedish of some terms used is found.

Table of Contents

1. Introduction

In Newtonian Mechanics, we have encountered some different equations for the motions of objects of different kinds. The simplest case possible, a pointlike particle moving under the influence of some force, is governed by the vector equation

$$\dot{\mathbb{p}} = \mathbb{F} \ . \tag{1.1}$$

This equation of motion can not be *derived* from some other equation. It is *postulated*, *i.e.* it is taken as an "axiom", or a fundamental truth of Newtonian Mechanics (one can also take the point of view that it *defines* one of the three quantities $\mathbb{F}$, m (the inertial mass) and $\mathbb{a}$ in terms of the other two).

Equation (1.1) is the fundamental equation in Newtonian Mechanics. If we consider other situations, *e.g.* the motion of a rigid body, the equations of motion

$$\dot{\mathbb{L}} = \vec{\tau} \tag{1.2}$$

can be obtained from it by imagining the body to be put together of a great number of small, approximately pointlike, particles whose relative positions are fixed (rigidity condition). If you don't agree here, you should go back and check that the only *dynamical* input in eq. (1.2) is eq. (1.1). What more is needed for eq. (1.2) is the *kinematical* rigidity constraint and a suitable definition for the torque $\vec{\tau}$. We have also seen eq. (1.1) expressed in a variety of forms obtained by expressing its components in non-rectilinear bases (*e.g.* polar coordinates). Although not immediately recognizable as eq. (1.1), these obviously contain no additional information, but just represent a choice of coordinates convenient to some problem. Furthermore, we have encountered the principles of energy, momentum and angular momentum, which tell that under certain conditions some of these quantities (defined in terms of masses and velocities, *i.e. kinematical*) do not change with time, or in other cases predict the rate at which they change. These are also consequences of eq. (1.1) or its derivates, *e.g.* eq. (1.2). Go back and check how the equations of motion are integrated to get those principles! It is very relevant for what will follow.

Taken all together, we see that although a great variety of different equations have been derived and used, they all have a common root, the equation of motion of a single pointlike particle. The issue for the subject of analytical mechanics is to put all the different forms of the equations of motion applying in all the different contexts on an equal footing. In fact, they will all be expressed as the same, identical, (set of) equation(s), Lagrange's equation(s), and, later, Hamilton's equation(s). In addition, these equations will be *derived* from a fundamental principle, the *action principle*, which then can be seen as the fundament of newtonian mechanics.

We will also see one of the most useful and important properties of Lagrange’s and Hamilton’s equations, namely that they take the same form independently of the choice of coordinates. This will make them extremely powerful when dealing with systems whose degrees of freedom most suitably are described in terms of variables in which Newton’s equations of motion are difficult to write down immediately, and they often dispense with the need of introducing forces whose only task is to make kinematical conditions fulfilled, such as for example the force in a rope of constant length (”constrained systems”). We will give several examples of these types of situations.

2. Generalized Coordinates

A most fundamental property of a physical system is its number of *degrees of freedom*. This is the minimal number of variables needed to completely specify the positions of all particles and bodies that are part of the system, *i.e.* its *configuration*. If the number of degrees of freedom is N, any set of variables $q_1, \ldots, q_N$ specifying the configuration is called a set of *generalized coordinates*.

> <u>Example:</u> *A point particle moving on a line has one degree of freedom. A generalized coordinate can be taken as x, the coordinate along the line. A particle moving in three dimensions has three degrees of freedom. Examples of generalized coordinates are the usual rectilinear ones, $\mathbb{r} = (x, y, z)$, and the spherical ones, $\mathbb{r} = (r, \theta, \phi)$, where $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\theta$.*

> <u>Example:</u> *A rigid body in two dimensions has three degrees of freedom – two "translational" which give the position of some specified point on the body and one "rotational" which gives the orientation of the body. An example, the most common one, of generalized coordinates is (x_c, y_c, ϕ), where x_c and y_c are rectilinear components of the position of the center of mass of the body, and ϕ is the angle from the x axis to a line from the center of mass to another point (x_1, y_1) on the body.*

> <u>Example:</u> *A rigid body in three dimensions has six degrees of freedom. Three of these are translational and correspond to the degrees of freedom of the center of mass. The other three are rotational and give the orientation of the rigid body. We will not discuss how to assign generalized coordinates to the rotational degrees of freedom (one way is the so called Euler angles), but the number should be clear from the fact that one needs a vector ω with three components to specify the rate of change of the orientation.*

The number of degrees of freedom is equal to the number of equations of motion one needs to find the motion of the system. Sometimes it is suitable to use a larger number of coordinates than the number of degrees of freedom for a system. Then the coordinates must be related via some kind of equations, called constraints. The number of degrees of freedom in such a case is equal to the number of generalized coordinates minus the number of constraints. We will briefly treat constrained systems in Chapter 8.

> <u>Example:</u> *The configuration of a mathematical pendulum can be specified using the rectilinear coordinates (x, y) of the mass with the fixed end of the string as origin. A natural generalized coordinate, however, would be the angle from the vertical. The number of degrees of freedom is only one, and (x, y) are subject to the constraint $x^2 + y^2 = l^2$, where l is the length of the string.*

In general, the generalized coordinates are chosen according to the actual problem one is interested in. If a body rotates around a fixed axis, the most natural choice for generalized coordinate is the rotational angle. If something moves rectilinearly, one choses a linear coordinate, *etc.* For composite systems, the natural choices for generalized coordinates are often mixtures of different types of variables, of which linear and angular ones are most common. The strength of the lagrangian formulation of Newton's mechanics, as we will soon see, is that the nature of the generalized coordinates is not reflected in the corresponding equation of motion. The way one gets to the equations of motion is identical for all generalized coordinates.

Generalized velocities are defined from the generalized coordinates exactly as ordinary velocity from ordinary coordinates:

$$v_i = \dot{q}_i \ , \quad i = 1, \ldots, N \ . \tag{2.1}$$

Note that the dimension of a generalized velocity depends on the dimension of the corresponding generalized coordinate, so that *e.g.* the dimension of a generalized velocity for an angular coordinate is $(\text{time})^{-1}$ – it is an angular velocity. In general, $(v_1, \ldots, v_N)$ is not the velocity vector.

> Example: *With polar coordinates (r, ϕ) as generalized coordinates, the generalized velocities are $(\dot{r}, \dot{\phi})$, while the velocity vector is $\dot{r}\hat{r} + r\dot{\phi}\hat{\phi}$.*

Exercises

1. Two masses m_1 and m_2 connected by a spring are sliding on a frictionless plane. How many degrees of freedom does this system have? Introduce a set of generalized coordinates!
2. Try to invent a set of generalized coordinates for a rigid body!

3. Generalized Forces

Suppose we have a system consisting of a number of point particles with coordinates $x_1, \ldots, x_N$, and that the configuration of the system also is described by the set of generalized coordinates $q_1, \ldots, q_N$. Since both sets of coordinates specify the configuration, there must be a relation between them:

$$\begin{aligned} x_1 &= x_1(q_1, q_2, \ldots, q_N) = x_1(q) \ , \\ x_2 &= x_2(q_1, q_2, \ldots, q_N) = x_2(q) \ , \\ &\vdots \\ x_N &= x_N(q_1, q_2, \ldots, q_N) = x_N(q) \ , \end{aligned} \tag{3.1}$$

compactly written as $x_i = x_i(q)$. To make the relation between the two sets of variable specifying the configuration completely general, the functions x_i could also involve an explicit time dependence. We choose not to include it here. The equations derived in Chapter 5 are valid also in that case. If we make a small (infinitesimal) displacement dq_i in the variables q_i, the chain rule implies that the corresponding displacement in x_i is

$$dx_i = \sum_{j=1}^{N} \frac{\partial x_i}{\partial q_j} dq_j \ . \tag{3.2}$$

The infinitesimal work performed by a force during such a displacement is the sum of terms of the type $\mathbb{F} \cdot \mathbb{r}$, *i.e.*

$$dW = \sum_{i=1}^{N} F_i dx_i = \sum_{j=1}^{N} \mathscr{F}_j dq_j \ , \tag{3.3}$$

where $\mathscr{F}$ is obtained from (3.2) as

$$\mathscr{F}_j = \sum_{i=1}^{N} F_i \frac{\partial x_i}{\partial q_j} \ . \tag{3.4}$$

$\mathscr{F}_j$ is the *generalized force* associated to the generalized coordinate q_j. As was the case with the generalized velocities, the dimensions of the $\mathscr{F}_j$'s need not be those of ordinary forces.

Example: *Consider a mathematical pendulum with length l, the generalized coordinate being ϕ, the angle from the vertical. Suppose that the mass moves* *an angle $d\phi$ under the influence of a force $\mathbb{F}$. The displacement of the mass is $d\mathbb{r} = l d\phi\hat{\phi}$ and the infinitesimal work becomes $dW = \mathbb{F}\cdot d\mathbb{r} = F_\phi l d\phi$. The generalized force associated with the angular coordinate ϕ obviously is $\mathscr{F}_\phi = F_\phi l$, which is exactly the torque of the force.*

The conclusion drawn in the example is completely general – the generalized force associated with an angular variable is a torque.

If the force is conservative, we may get it from a potential V as

$$F_i = -\frac{\partial V}{\partial x_i}\ . \tag{3.5}$$

If we then insert this into the expression (3.4) for the generalized force, we get

$$\mathscr{F}_j = -\sum_{i=1}^{N}\frac{\partial V}{\partial x_i}\frac{\partial x_i}{\partial q_j} = -\frac{\partial V}{\partial q_j}\ . \tag{3.6}$$

The relation between the potential and the generalized force looks the same whatever generalized coordinates one uses.

Exercises

3. A particle is moving without friction at the curve $y = f(x)$, where y is vertical and x horizontal, under the influence of gravity. What is the generalized force when x is chosen as generalized coordinate?

4. Kinetic Energy and Generalized Momenta

We will examine how the kinetic energy depends on the generalized coordinates and their derivatives, the generalized velocities. Consider a single particle, so that $N = 3$ in the above description. The kinetic energy is

$$K = \tfrac{1}{2}m \sum_{i=1}^{3} \dot{x}_i^2 \ . \tag{4.1}$$

Eq. (3.2) in the form

$$\dot{x}_i = \sum_{j=1}^{3} \frac{\partial x_i}{\partial q_j} \dot{q}_j \tag{4.2}$$

tells us that $\dot{x}_i$ is a function of q_j, $\dot{q}_j$ and time (time enters only if the transition functions (3.1) involve time explicitely). We may write the kinetic energy in terms of the generalized coordinates and velocities as

$$K = \tfrac{1}{2}m \sum_{j,k=1}^{3} A_{jk}(q)\dot{q}_j\dot{q}_k \tag{4.3}$$

(or in matrix notation $K = \frac{1}{2}m\dot{q}^t A\dot{q}$), where the symmetric matrix A is given by

$$A_{jk} = \sum_{i=1}^{3} \frac{\partial x_i}{\partial q_j} \frac{\partial x_i}{\partial q_k} \ . \tag{4.4}$$

It is important to note that although the relations between the rectilinear coordinates x_i and the generalized coordinates q_j may be non-linear, the kinetic energy is always a bilinear form in the generalized velocities with coefficients (A_{jk}) that depend only on the generalized coordinates.

> Example: *We look again at plane motion in polar coordinates. The relations to rectilinear ones are*
>
> $$\begin{aligned} x &= r\cos\phi \ , \\ y &= r\sin\phi \ , \end{aligned} \tag{4.5}$$
>
> *so the matrix A becomes (after a little calculation)*
>
> $$A = \begin{bmatrix} A_{rr} & A_{r\phi} \\ A_{r\phi} & A_{\phi\phi} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix} \ , \tag{4.6}$$

and the obtained kinetic energy is in agreement with the well known

$$K = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2)\ . \tag{4.7}$$

If one differentiates the kinetic energy with respect to one of the (ordinary) velocities $v_i = \dot{x}_i$, one obtains

$$\frac{\partial K}{\partial \dot{x}_i} = m\dot{x}_i\ , \tag{4.8}$$

i.e. a momentum. The *generalized momenta* are defined in the analogous way as

$$p_j = \frac{\partial K}{\partial \dot{q}_j}\ . \tag{4.9}$$

Example: The polar coordinates again. Differentiating K *of eq.* (4.7) *with respect to* r *and* ϕ *yields*

$$p_r = m\dot{r}\ , \quad p_\phi = mr^2\dot{\phi}\ . \tag{4.10}$$

The generalized momentum to r *is the radial component of the ordinary momentum, while the one associated with* ϕ *is the angular momentum, something which by now should be no surprise.*

The fact that the generalized momentum associated to an angular variable is an angular momentum is a completely general feature.

We now want to connect back to the equations of motion, and formulate them in terms of the generalized coordinates. This will be done in the following chapter.

Exercises

4. Find the expression for the kinetic energy in terms of spherical coordinates (r, θ, ϕ), $(x, y, z) = (r\sin\theta\cos\phi\,,\ r\sin\theta\sin\phi\,,\ r\cos\theta)$.
5. Find the kinetic energy for a particle moving at the curve $y = f(x)$.

5. Lagrange's Equations

5.1. A Single Particle

The equation of motion, as we know it so far, is given by (1.1). We would like to recast it in a form that is possible to generalize to generalized coordinates. Remembering how the momentum was obtained from the kinetic energy, we rewrite (1.1) in the form

$$\frac{d}{dt}\frac{\partial K}{\partial \dot{x}_i} = F_i \ . \tag{5.1}$$

A first guess would be that this, or something very similar, holds if the coordinates are replaced by the generalized coordinates and the force by the generalized force. We therefore calculate the left hand side of (5.1) with q instead of x and see what we get:

$$\begin{aligned}\frac{d}{dt}\frac{\partial K}{\partial \dot{q}_j} &= \sum_{i=1}^{3}\frac{d}{dt}\left(\frac{\partial K}{\partial \dot{x}_i}\frac{\partial \dot{x}_i}{\partial \dot{q}_j}\right) = \sum_{i=1}^{3}\frac{d}{dt}\left(\frac{\partial K}{\partial \dot{x}_i}\frac{\partial x_i}{\partial q_j}\right) = \\ &= \sum_{i=1}^{3}\left(\frac{\partial x_i}{\partial q_j}\frac{d}{dt}\frac{\partial K}{\partial \dot{x}_i} + \frac{d}{dt}\frac{\partial x_i}{\partial q_j}\frac{\partial K}{\partial \dot{x}_i}\right) = \sum_{i=1}^{3}\left(\frac{\partial x_i}{\partial q_j}\frac{d}{dt}\frac{\partial K}{\partial \dot{x}_i} + \frac{\partial \dot{x}_i}{\partial q_j}\frac{\partial K}{\partial \dot{x}_i}\right) = \\ &= \sum_{i=1}^{3}\frac{\partial x_i}{\partial q_j}\frac{d}{dt}\frac{\partial K}{\partial \dot{x}_i} + \frac{\partial K}{\partial q_j} \ .\end{aligned} \tag{5.2}$$

Here, we have used the chain rule and the fact that K depends on $\dot{x}_i$ and not on x_i in the first step. Then, in the second step, we use the fact that x_i are functions of the q's and not the $\dot{q}$'s to get $\frac{\partial \dot{x}_i}{\partial \dot{q}_j} = \frac{\partial x_i}{\partial q_j}$. The fourth step uses this again to derive $\frac{d}{dt}\frac{\partial x_i}{\partial q_j} = \frac{\partial \dot{x}_i}{\partial q_j}$, and the last step again makes use of the chain rule on K. Now we can insert the form (5.1) for the equations of motion of the particle:

$$\frac{d}{dt}\frac{\partial K}{\partial \dot{q}_j} = \sum_{i=1}^{3}\frac{\partial x_i}{\partial q_j}F_i + \frac{\partial K}{\partial q_j} \ , \tag{5.3}$$

and arrive at Lagrange's equations of motion for the particle:

$$\frac{d}{dt}\frac{\partial K}{\partial \dot{q}_j} - \frac{\partial K}{\partial q_j} = \mathscr{F}_j \ . \tag{5.4}$$

> Example: *A particle moving under the force $\mathbb{F}$ using rectilinear coordinates. Here one must recover the known equation $m\mathbb{a} = \mathbb{F}$. Convince yourself that this is true.*

Example: To complete the series of examples on polar coordinates, we finally derive the equations of motion. From (4.7)*, we get*

$$\begin{array}{ll} \frac{\partial K}{\partial \dot{r}} = m\dot{r}\ , & \frac{\partial K}{\partial r} = mr\dot{\phi}^2\ , \\ \frac{\partial K}{\partial \dot{\phi}} = mr^2\dot{\phi}\ , & \frac{\partial K}{\partial \phi} = 0\ . \end{array} \tag{5.5}$$

Lagrange's equations now give

$$\begin{array}{l} m(\ddot{r} - r\dot{\phi}^2) = F_r\ , \\ m(r^2\ddot{\phi} + 2r\dot{r}\dot{\phi}) = \tau\ (= rF_\phi)\ . \end{array} \tag{5.6}$$

The lagrangian formalism is most useful in cases when there is a potential energy, *i.e.* the forces are conservative and mechanical energy is conserved. Then the generalized forces can be written as $\mathscr{F}_j = -\frac{\partial V}{\partial q_j}$ and Lagrange's equations read

$$\frac{d}{dt}\frac{\partial K}{\partial \dot{q}_j} - \frac{\partial K}{\partial q_j} + \frac{\partial V}{\partial q_j} = 0\ . \tag{5.7}$$

The potential V can not depend on the generalized velocities, so if we form

$$L = K - V\ , \tag{5.8}$$

the equations are completely expressible in terms of L:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0 \tag{5.9}$$

The function L is called the *Lagrange function* or the *lagrangian*. This form of the equations of motion is the one most often used for solving problems in mechanics. There will be examples in a little while.

Example: Suppose that one, for some strange reason, wants to solve for the motion of a particle with mass m moving in a harmonic potential with spring constant k using the generalized coordinate $q = x^{1/3}$ instead of the (inertial) coordinate x. In order to derive Lagrange's equation for $q(t)$, one first has to express the kinetic and potential energies in terms of q and $\dot{q}$. One gets $\dot{x} = \frac{\partial x}{\partial q}\dot{q} = 3q^2\dot{q}$ and thus $K = \frac{9m}{2}q^4\dot{q}^2$. The potential is $V = \frac{1}{2}kx^2 = \frac{1}{2}kq^6$, so that $L = \frac{9m}{2}q^4\dot{q}^2 - \frac{1}{2}kq^6$. Before writing down Lagrange's equations we *need* $\frac{\partial L}{\partial q} = 18mq^3\dot{q}^2 - 3kq^5$ *and* $\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = \frac{d}{dt}\left(9mq^4\dot{q}\right) = 9mq^4\ddot{q} + 36mq^3\dot{q}^2$. *Finally,*

$$\begin{aligned} 0 &= \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = \\ &= 9mq^4\ddot{q} + 36mq^3\dot{q}^2 - 18mq^3\dot{q}^2 + 3kq^5 = \\ &= 9mq^4\ddot{q} + 18mq^3\dot{q}^2 + 3kq^5 = \\ &= 3q^2\left(3mq^2\ddot{q} + 6mq\dot{q}^2 + kq^3\right) \end{aligned} \quad . \tag{5.10}$$

If one was given this differential equation as an exercise in mathematics, one would hopefully end up by making the change of variables $x = q^3$*, which turns it into*

$$m\ddot{x} + kx = 0 \ , \tag{5.11}$$

which one recognizes as the correct equation of motion for the harmonic oscillator.

The example just illustrates the fact that Lagrange's equations give the correct result for any choice of generalized coordinates. This is certainly not the case for Newton's equations. If x fulfills eq. (5.11), it certainly doesn't imply that any $q(x)$ fulfills the same equation!

Exercises

6. Write down Lagrange's equations for a freely moving particle in spherical coordinates!
7. A particle is constrained to move on the sphere $r = a$. Find the equations of motion in the presence of gravitation.
8. A bead is sliding without friction along a massless string. The endpoint of the string are fixed at $(x, y) = (0, 0)$ and $(a, 0)$ and the length of the string is $a\sqrt{2}$. Gravity acts along the negative z-axis. Find the stable equilibrium position and the frequency for small oscillations around it!
9. Write down Lagrange's equations for a particle moving at the curve $y = f(x)$. The y axis is vertical. Are there functions that produce harmonic oscillations? What is the angular frequency of small oscillations around a local minimum $x = x_0$?

5.2. Lagrange’s equations with any number of degrees of freedom

In a more general case, the system under consideration can be any mechanical system: any number of particles, any number of rigid bodies etc. The first thing to do is to determine the number of degrees of freedom of the system. In three dimensions, we already know that a particle has three translational degrees of freedom and that a rigid body has three translational and three rotational ones. This is true as long as there are no kinematical constraints that reduce these numbers. Examples of such constraints can be that a mass is attached to the end of an unstretchable string, that a body slides on a plane, that a particle is forced to move on the surface of a sphere, that a rigid body only may rotate about a fixed axis,...

Once the number n of degrees of freedom has been determined, one tries to find the same number of variables that specify the configuration of the system, the ”position”. Then these variables are *generalized coordinates* for the system. Let us call them $q_1, q_2, \ldots, q_n$. The next step is to find an expression for the kinetic and potential energies in terms of the q_k’s and the $\dot{q}_k$’s (we confine to the case where the forces are conservative – for dissipative forces the approach is not as powerful). Then the lagrangian is formed as the difference $L = K - V$. The objects $p_k = \frac{\partial L}{\partial \dot{q}}$ are called *generalized momenta* and $v_k = \dot{q}_k$ *generalized velocities* (if q_k is a rectilinear coordinate, p_k and v_k coincide with the ordinary momentum and velocity components). Lagrange’s equations for the systems are

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = 0 \ , \qquad k = 1, \ldots, n \ , \tag{5.12}$$

or, equivalently,

$$\dot{p}_k - \frac{\partial L}{\partial q_k} = 0 \ . \tag{5.13}$$

We state these equations without proof. The proof is completely along the lines of the one-particle case, only that some indices have to be carried around. Do it, if you feel tempted!

In general, the equations (5.12) lead to a system of n coupled second order differential equations. We shall take a closer look at some examples.

Example: *A simple example of a constrained system is the "mathematical pendulum", consisting of a point mass moving under the influence of gravitation and attached to the end of a massless unstretchable string whose other end is fixed at a point. If we consider this system in two dimensions, the particle moves in a plane parameterized by two rectilinear coordinates that we may label x and y. The number of degrees of freedom here is not two, however. The constant length l of the rope puts a constraint on the position of the particle, which we can write as $x^2 + y^2 = l^2$ if the fixed end of the string is taken as origin. The number of degrees of freedom is* one, *the original two minus one constraint. It is possible, but not recommendable, to write the equations of motion using these rectilinear coordinates. Then one has to introduce a string force that has exactly the right value to keep the string unstretched, and then eliminate it. A better way to proceed is to identify the single degree of freedom of the system as the angle ϕ from the vertical (or from some other fixed line through the origin). ϕ is now the generalized coordinate of the system. In this and similar cases, Lagrange's equations provide a handy way of deriving the equations of motion. The velocity of the pointmass is $v = l\dot{\phi}$, so its kinetic energy is $K = \frac{1}{2}ml^2\dot{\phi}^2$. The potential energy is $V = -mgl\cos\phi$. We form the lagrangian as*

$$L = K - V = \tfrac{1}{2}ml^2\dot{\phi}^2 + mgl\cos\phi \ . \tag{5.14}$$

We form

$$\begin{aligned} \frac{\partial L}{\partial \phi} &= -mgl\sin\phi \ , \\ p_\phi = \frac{\partial L}{\partial \dot{\phi}} &= ml^2\dot{\phi} \ , \\ \frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} &= ml^2\ddot{\phi} \ . \end{aligned} \tag{5.15}$$

Lagrange's equation for ϕ now gives $ml^2\ddot{\phi} + mgl\sin\phi = 0$, i.e.

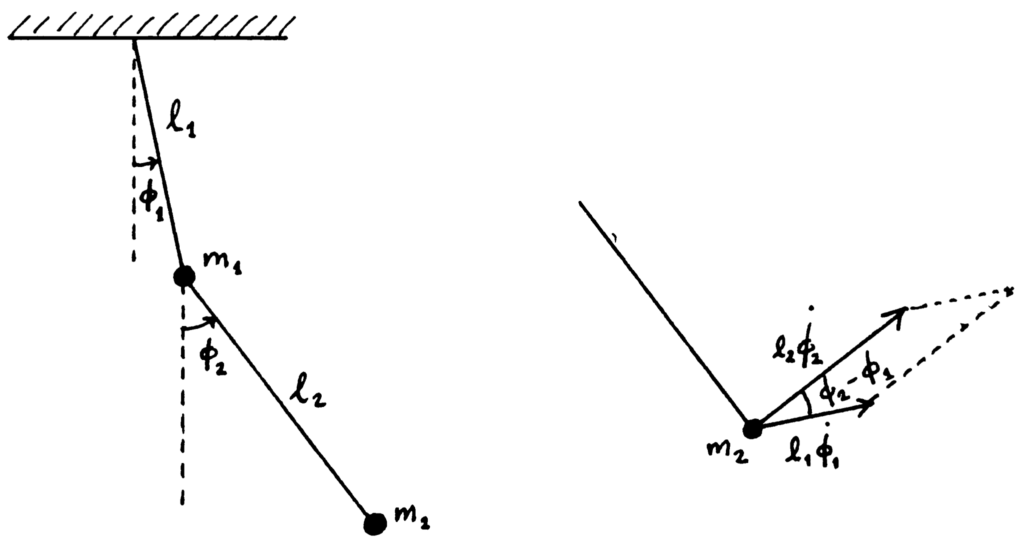
$$\ddot{\phi} + \frac{g}{l}\sin\phi = 0 \ . \tag{5.16}$$

This equation should be recognized as the correct equation of motion for the mathematical pendulum. In the case of small oscillations, one approximates $\sin\phi \approx \phi$ and get harmonic oscillations with angular frequency $\sqrt{g/l}$.

Some comments can be made about this example that clarifies the lagrangian approach. First a dimensional argument: the lagrangian always has the dimension of energy. The generalized velocity here is $v_\phi = \dot{\phi}$, the angular velocity, with dimension (time)$^{-1}$. The generalized momentum p_ϕ, being the derivative of L with respect to $\dot\phi$, obviously hasn't the dimension of an ordinary momentum, but (energy) $\times$ (time) = (mass) $\times$ (length)^2 $\times$ (time)$^{-1}$ = (mass) $\times$ (length) $\times$ (velocity). This is the same dimension as an angular momentum component (recall "$\mathbb{L} = m\mathbb{r} \times \mathbb{v}$"). If we look back at eq. (5.15), we see that p_ϕ is indeed the angular momentum with respect to the origin. This fenomenon is quite general: *if the generalized coordinate is an angle, the associated generalized momentum is an angular momentum.* It is not difficult to guess that the generalized force should be the torque, and this is exactly what we find by inspecting $-\frac{\partial V}{\partial\phi} = -mgl\sin\phi$.

The example of the mathematical pendulum is still quite simple. It is easy to solve without the formalism of Lagrange, best by writing the equation for the angular momentum (which is what Lagrange's equation above achieves) or, alternatively, by writing the force equations in polar coordinates. By using Lagrange's equations one doesn't have to worry about *e.g.* expressions for the acceleration in non-rectilinear coordinates. That comes about automatically.

There are more complicated classes of situations, where the variables are not simply an angle or rectilinear coordinates or a combination of these. Then Lagrange's equations makes the solution much more easier. We shall look at another example, whose equations of motion are cumbersome to derive using forces or torques, a coupled double pendulum.



<u>Example:</u> *Consider two mathematical pendulums one at the end of another, with masses and lengths as indicated in the figure. The number of degrees of freedom of this system is two (as long as the strings are stretched), and we need to find two variables that completely specify the configuration of it, i.e.*

the positions of the two masses. The two angles ϕ_1 and ϕ_2 provide one natural choice, which we will use, although there are other possibilities, e.g. to use instead of ϕ_2 the angle ${\phi_2}' = \phi_2 - \phi_1$ which is zero when the two strings are aligned. The only intelligent thing we have to perform now is to write down expressions for the kinetic and potential energies, then Lagrange does the rest of the work. We start with the kinetic energy, which requires knowledge of the velocities. The upper particle is straightforward, it has the velocity $v_1 = l_1\dot{\phi}_1$. The lower one is trickier. The velocity gets two contributions, one from ϕ_1 changing and one from ϕ_2 changing. Try to convince yourselves that those have absolute values $l_1\dot{\phi}_1$ and $l_2\dot{\phi}_2$ respectively, and that the angle between them is $\phi_2 - \phi_1$ as in the figure. The first of these contributions depend on ϕ_2 being defined from the vertical, so that when only ϕ_1 changes, the lower string gets parallel transported but not turned. Now the cosine theorem gives the square of the total velocity for the lower particle:

$$ {v_2}^2 = {l_1}^2{\dot{\phi}_1}^2 + {l_2}^2{\dot{\phi}_2}^2 + 2l_1 l_2 \dot{\phi}_1\dot{\phi}_2\cos(\phi_2 - \phi_1)\ , \tag{5.17} $$

so that the kinetic energy becomes

$$ K = \tfrac{1}{2}m_1{l_1}^2{\dot{\phi}_1}^2 + \tfrac{1}{2}m_2\left[{l_1}^2{\dot{\phi}_1}^2 + {l_2}^2{\dot{\phi}_2}^2 + 2l_1 l_2\dot{\phi}_1\dot{\phi}_2\cos(\phi_2 - \phi_1)\right]\ . \tag{5.18} $$

The potential energy is simpler, we just need the distances from the "roof" to obtain

$$ V = -m_1 g l_1\cos\phi_1 - m_2 g(l_1\cos\phi_1 + l_2\cos\phi_2)\ . \tag{5.19} $$

Now the intelligence is turned off, the lagrangian is formed as $L = K - V$, and Lagrange's equations are written down. We leave the derivation as an exercise (a good one!) and state the result:

$$ \begin{aligned} &\ddot{\phi}_1 + \frac{m_2}{m_1+m_2}\frac{l_2}{l_1}\left[\ddot{\phi}_2\cos(\phi_2-\phi_1) - {\dot{\phi}_2}^2\sin(\phi_2-\phi_1)\right] + \frac{g}{l_1}\sin\phi_1 = 0\ ,\\ &\ddot{\phi}_2 + \frac{l_1}{l_2}\left[\ddot{\phi}_1\cos(\phi_2-\phi_1) + {\dot{\phi}_1}^2\sin(\phi_2-\phi_1)\right] + \frac{g}{l_2}\sin\phi_2 = 0\ . \end{aligned} \tag{5.20} $$

One word is at place about the way these equations are written. When one gets complicated expressions with lots of parameters hanging around in different places, it is good to try to arrange things as clearly as possible. Here, we have combined masses and lengths to get dimensionless factors as far as possible, which makes a dimensional analysis simple. This provides a check for errors – most calculational errors lead to dimensional errors! The *equations (5.20) are of course not analytically solvable. For that we need a computer simulation. What we can obtain analytically is a solution for small angles* ϕ_1 *and* ϕ_2*. We will do this calculation for two reasons. Firstly, it learns us something about how to linearize equations, and secondly, it tells us about interesting properties of coupled oscillatory systems. In order to linearize the equations, we throw away terms that are not linear in the angles or their time derivatives. To identify these, we use the Maclaurin expansions for the trigonometric functions. The lowest order terms are enough, so that* $\cos x \approx 1$ *and* $sinx \approx x$*. The terms containing* $\dot{\phi}_1^{\,2}$ *or* $\dot{\phi}_2^{\,2}$ *go away (these can be seen to represent centrifugal forces, that do not contribute when the strings are approximately aligned or the angular velocities small). The linearized equations of motion are thus*

$$\begin{aligned} &\ddot{\phi}_1 + \frac{m_2}{m_1+m_2}\frac{l_2}{l_1}\ddot{\phi}_2 + \frac{g}{l_1}\phi_1 = 0 \ , \\ &\frac{l_1}{l_2}\ddot{\phi}_1 + \ddot{\phi}_2 + \frac{g}{l_2}\phi_2 = 0 \ . \end{aligned} \tag{5.21}$$

We now have a system of two coupled linear second order differential equations. They may be solved by standard methods. It is important to look back and make sure that you know how that is done. The equations can be written on matrix form

$$M\ddot{\Phi} + K\Phi = 0 \ , \tag{5.22}$$

where

$$\Phi = \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} \ , \qquad M = \begin{bmatrix} 1 & \frac{m_2}{m_1+m_2}\frac{l_2}{l_1} \\ \frac{l_1}{l_2} & 1 \end{bmatrix} \ , \qquad K = \begin{bmatrix} \frac{g}{l_1} & 0 \\ 0 & \frac{g}{l_2} \end{bmatrix} \ . \tag{5.23}$$

The ansatz one makes is $\Phi = Ae^{\pm i\omega t}$ *with* A *a column vector containing "amplitudes", which gives*

$$(-M\omega^2 + K)A = 0 \ . \tag{5.24}$$

Now one knows that this homogeneous equation has non-zero solutions for A *only when the determinant of the "coefficient matrix"* $(-M\omega^2 + K)$ *is zero, i.e. the rows are linearly dependent giving two copies of the same equation. The vanishing of the determinant gives a second order equation for* ω^2 *whose* *solutions, after some work (do it!), are*

$$\omega^2 = \frac{g}{2m_1l_1l_2}\Big\{(m_1+m_2)(l_1+l_2)\pm \\ \pm\sqrt{(m_1+m_2)\big[m_1(l_1-l_2)^2+m_2(l_1+l_2)^2\big]}\Big\}\ . \tag{5.25}$$

These are the eigenfrequencies *of the system. It is generic for coupled system with two degrees of freedom that there are two eigenfrequencies. If one wants, one can check what A is in the two cases. It will turn out, and this is also generic, that the lower frequency corresponds to the two masses moving in the same direction and the higher one to opposite directions. It is not difficult to imagine that the second case gives a higher frequency – it takes more potential energy, and thus gives a higher "spring constant". What is one to do with such a complicated answer? The first thing is the dimension control. Then one should check for cases where one knows the answer. One such instance is the* single *pendulum, i.e. $m_2 = 0$. Then the expression* (5.25) *should boil down to the frequency $\sqrt{g/l_1}$ (see exercise 14). One can also use ones physical imagination to deduce what happens e.g. when m_1 is much smaller than m_2. After a couple of checks like this one can be almost sure that the expression obtained is correct. This is possible for virtually every problem.*

The above example is very long and about as complicated a calculation we will encounter. It may seem confusing, but give it some time, go through it systematically, and you will see that it contains many ingredients and methods that are useful to master. If you really understand it, you know most of the things you need to solve many-variable problems in Lagrange's formalism.

The Lagrange function is the difference between kinetic and potential energy. This makes energy conservation a bit obscure in Lagrange's formalism. We will explain how it comes about, but this will become clearer when we move to Hamilton's formulation. Normally, in one dimension, one has the equation of motion $m\ddot{x} = F$. In the case where F only depends on x, there is a potential, and the equation of motion may be integrated using the trick $\ddot{x} = a = v\frac{dv}{dx}$ which gives $mvdv = Fdx$, $\frac{1}{2}mv^2 - \int Fdx = C$, conservation of energy. It must be possible to do this in Lagrange's formalism too. If the lagrangian does not depend on t, we observe that

$$
\begin{aligned}
&\frac{d}{dt}\left[\dot{x}\frac{\partial L}{\partial \dot{x}} - L(x,\dot{x})\right] = \\
&= \ddot{x}\frac{\partial L}{\partial \dot{x}} + \dot{x}\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \dot{x}\frac{\partial L}{\partial x} - \ddot{x}\frac{\partial L}{\partial \dot{x}} = \\
&= \dot{x}\left[\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x}\right]
\end{aligned}
\quad . \tag{5.26}
$$

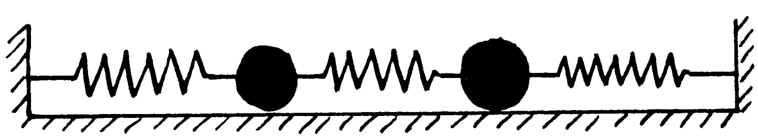
Therefore, Lagrange's equation implies that the first quantity in square brackets is conserved. As we will see in Chapter 7, it is actually the energy. In a case with more generalized coordinates, the energy takes the form

$$
E = \sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L = \sum_i \dot{q}_i p_i - L \ . \tag{5.27}
$$

It is only in one dimension that energy conservation can replace the equation of motion — for a greater number of variables it contains less information.

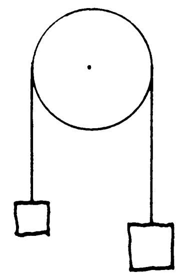
Exercises

10. Find and solve the equations of motion for a homogeneous sphere rolling down a slope, assuming enough friction to prevent sliding.
11. A particle is connected to a spring whose other end is fixed, and free to move in a horizontal plane. Write down Lagrange's equations for the system, and describe the motion qualitatively.
12. Find Lagrange's equations for the system in exercise 1.
13. Two masses are connected with a spring, and each is connected with a spring to a fixed point. Find the equations of motion, and describe the motion qualitatively. Solve for the possible angular frequencies in the case when the masses are equal and the spring constants are equal. There is no friction.

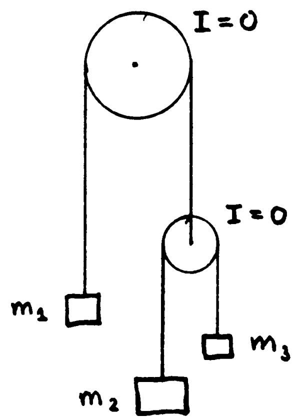


14. Consider the double pendulum in the limit when either of the masses is small compared to the other one, and interpret the result.
15. Find the equations of motion for a particle moving on an elliptic curve $(\frac{x}{a})^2 + (\frac{y}{b})^2 = 1$ using a suitable generalized coordinate. Check the case when $a = b$.

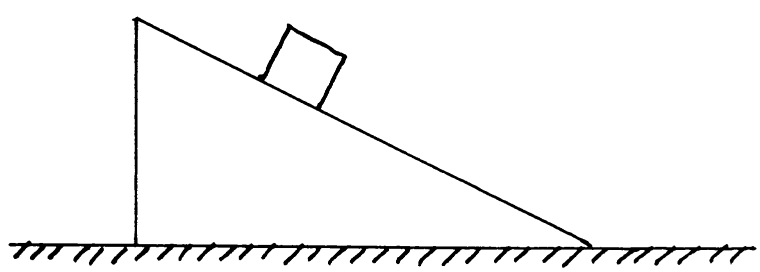
16. Consider Atwood's machine. The two masses are m_1 and m_2 and the moment of inertia of the pulley is I. Find the equation of motion using Lagrange's formulation. Note the simplification that one never has to consider the internal forces.



17. Calculate the accelerations of the masses in the double Atwood machine.



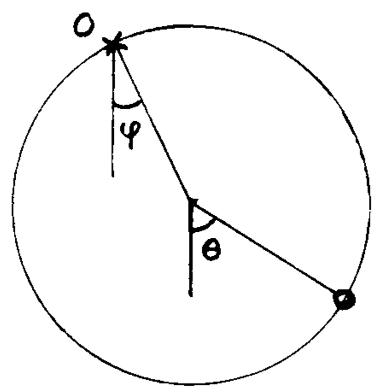
18. A particle of mass m is sliding on a wedge, which in turn is sliding on a horizontal plane. No friction. Determine the relative acceleration of the particle with respect to the wedge.



19. A pendulum is suspended in a point that moves horizontally according to $x = a \sin \omega t$. Find the equation of motion for the pendulum, and specialize to small angles.

20. A particle slides down a stationary sphere without friction beginning at rest at the top of the sphere. What is the reaction from the sphere on the particle as a function of the angle θ from the vertical? At what value of θ does the particle leave the surface?

21. A small bead of mass m is sliding on a smooth circle of radius a and mass m which in turn is freely moving in a vertical plane around a fixed point O on its periphery. Give the equations of motion for the system, and solve them for small oscillations around the stable equilibrium. How should the initial conditions be chosen for the system to move as a rigid system? For the center of mass not to leave the vertical through O?



6. The Action Principle

In this chapter we will formulate a fundamental principle leading to the equations of motion for any mechanical system. It is the *action principle*. In order to understand it, we need some mathematics that goes beyond ordinary analysis, so called functional analysis. This is nothing to be afraid of, and the mathematical strictness of what we are doing will be minimal.

Suppose we have a mechanical system – for simplicity we can think of a particle moving in a potential – and we do not yet know what the path $\mathbb{r}(t)$ of it will be, once the initial conditions are given (it is released at a certain time t_0 with given position $\mathbb{r}(t_0) = \mathbb{r}_0$ and velocity $\mathbb{v}(t_0) = \mathbb{v}_0$). For *any* path $\mathbb{r}(t)$ fulfilling the initial conditions we define a number S by

$$S = \int_{t_0}^{\infty} dt L \ , \tag{6.1}$$

where L is the lagrangian $K - V$. This is the *action*. In the case of a single particle in a potential, the action is

$$S = \int_{t_0}^{\infty} dt \Big[\tfrac{1}{2} m \dot{x}(t)^2 - V\big(x(t)\big) \Big] \ . \tag{6.2}$$

The action is a function whose argument is a function and whose value is a number (carrying dimension (energy)×(time)). Such a function is called a *functional*. When the argument is written out, we we enclose it in square brackets, *e.g.* "$S\,[x(t)]$", to mark the difference from ordinary functions.

The action principle now states that *the path actually taken by the particle must be a stationary point of the action*. What does this mean? Recall how one determines when an ordinary function $f(x)$ has a local extremum. If we make an infinitesimal change δx in the argument of the function, the function itself does not change, so that $f(x + \delta x) = f(x)$. This is the same as saying that the derivative is zero, since $f'(x) = \lim_{\varepsilon \to 0} \frac{f(x+\varepsilon) - f(x)}{\varepsilon}$. When now the function we want to "extremize" is a functional instead of an ordinary function, we must in the same way demand that a small change in the argument $\mathbb{r}(t)$ of the functional does not change the functional. Therefore we chose a new path for the particle $\mathbb{r}(t) + \vec{\varepsilon}(t)$ (we have to take $\vec{\varepsilon}(t_0) = \dot{\vec{\varepsilon}}(t_0) = 0$ not to change the given initial conditions) which differs infinitesimally from $\mathbb{r}(t)$ at every time, and see how the action S is changed. For simplicity we consider rectilinear motion, so that there is only one coordinate $x(t)$. We get

$$S[x(t)+\varepsilon(t)]-S[x(t)]=\int_{t_0}^{\infty}dt\,\big[L\big(x(t)+\varepsilon(t),\dot{x}(t)+\dot{\varepsilon}(t)\big)-L\big(x(t),\dot{x}(t)\big)\big]\ . \quad (6.3)$$

Taking ε to be infinitesimally small, one can save parts linear in ε only, to obtain $L\big(x(t)+\varepsilon(t),\dot{x}(t)+\dot{\varepsilon}(t)\big)=L\big(x(t),\dot{x}(t)\big)+\varepsilon(t)\frac{\partial L}{\partial x}(t)+\dot{\varepsilon}(t)\frac{\partial L}{\partial \dot{x}}(t)$. Inserting this into eq. (6.3) gives

$$\begin{aligned}S[x(t)+\varepsilon(t)]-S[x(t)]&=\int_{t_0}^{\infty}dt\left[\varepsilon(t)\frac{\partial L}{\partial x}(t)+\dot{\varepsilon}(t)\frac{\partial L}{\partial \dot{x}}(t)\right]=[\text{partial integration}]=\\&=\int_{t_0}^{\infty}dt\varepsilon(t)\left[\frac{\partial L}{\partial x}(t)-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}(t)\right]\end{aligned} \quad (6.4)$$

(some boundary term at infinity has been thrown away, but never mind). If the path $x(t)$ is to be a stationary point, this has to vanish for all possible infinitesimal changes $\varepsilon(t)$, which means that the entity inside the square brackets in the last expression in eq. (6.4) has to vanish for all times. We have rederived Lagranges equations as a consequence of the action principle. The derivation goes the same way if there are more degrees of freedom (do it!).

The above derivation actually shows that Lagrange's equations are true also for non-rectilinear coordinates. We will not present a rigorous proof of that, but think of the simpler analog where a function of a number of variables has a local extremum in some point. If we chose different coordinates, the function *itself* is of course not affected – the solution to the minimization problem is still that all derivatives of the function vanish at the local minimum. The only difference for our action *functional* is that the space in which we look for stationary points is infinite-dimensional.

One may say a word about the nature of the stationary points. Are they local minima or maxima? In general, they need not be either. The normal situation is that they are "terrace points", comparable to the behaviour of the function x^3 at $x=0$. Paths that are "close" to the actual solution may have either higher or lower value of the action. The only general statement one can make about the solution is that it is a stationary point of the action, *i.e.* that an infinitesimal change in the path gives no change in the action, analogously to the statement that a function has zero derivative in some point.

Analogously to the way one defines derivatives of functions, one can define *functional derivatives* of functionals. A functional derivative $\frac{\delta}{\delta x(t)}$ is defined so that a change in the argument $x(t)$ by an infinitesimal function $\varepsilon(t)$ gives a change in the functional $F[x(t)]$:

$$F[x+\varepsilon]-F[x]=\int dt\varepsilon(t)\frac{\delta F}{\delta x(t)}\ . \tag{6.5}$$

The functional derivative of F is a functional with an explicit t-dependence. Compare this definition with what we did in eqs. (6.3) and (6.4). We then see that the action principle can be formulated as

$$\frac{\delta S}{\delta x(t)}=0 \tag{6.6}$$

in much the same way as an ordinary local extremum is given by $\frac{df}{dx}=0$. Eqn. (6.6) is Lagrange's equation.

Variational principles are useful in many areas, not only in Newtonian Mechanics. The action formulation of the dynamics of a system is the dominating one when one formulates *field theories*. Elementary particles are described by relativistic quantum fields, and their motion and interaction are almost always described in terms of an action.

Exercises

22. Using a variational method, find the shortest path between two given points.
23. Find the shortest path between two points on a sphere.

7. Hamilton's Equations

When we derived Lagrange's equations, the variables we used were the generalized coordinates $q_1, \ldots, q_N$ and the generalized velocities $\dot{q}_1, \ldots, \dot{q}_N$. The lagrangian L was seen as a function of these, $L(q_i, \dot{q}_i)$. This set of variables is not unique, and there is one other important choice, that is connected to the *hamiltonian* formulation of mechanics. Hamilton's equations, as compared to Lagrange's equations, do not present much, if any, advantage when it comes to problem solving in Newtonian Mechanics. Some things become clearer, though, *e.g.* the nature of conserved quantities. Hamilton's formulation is often used in quantum mechanics, where it leads to a complementary and equivalent picture to one that uses Lagrange's variables.

We depart from the lagrangian, as defined in chapter 5, and define the generalized momenta corresponding to the coordinates q_i according to

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \ . \tag{7.1}$$

In a rectilinear coordinate system, p_i are the usual momenta, $p_i = m\dot{q}_i$, but, as we have seen, this is not true for other types of generalized coordinates.

Our situation now is that we want to change the fundamental variables from (generalized) coordinates q_i and velocities v_i to coordinates and momenta p_i. We will soon see that it is natural to consider an other function than the Lagrangian when this change of variables is performed. To illustrate this, consider a situation with only one coordinate q. The differential of the lagrangian $L(q, v)$ is

$$dL = \frac{\partial L}{\partial q} dq + \frac{\partial L}{\partial v} dv = \frac{\partial L}{\partial q} dq + p dv \ . \tag{7.2}$$

In a framework where the fundamental variables are q and p we want the differential of a function to come out naturally as (something)dq+(something)dp. Consider the new function H defined by

$$H = vp - L = \dot{q}p - L \ . \tag{7.3}$$

H is the *hamiltonian*. Its differential is

$$dH = dvp + vdp - dL = dvp + vdp - \frac{\partial L}{\partial q} dq - pdv = -\frac{\partial L}{\partial q} dq + vdp \ , \tag{7.4}$$

so we have

$$\frac{\partial H}{\partial q} = -\frac{\partial L}{\partial q} \ , \quad \frac{\partial H}{\partial p} = v = \dot{q} \ . \tag{7.5}$$

The change of function of the type (7.3) associated with this kind of change of variables as is called a *Legendre transform*. It is important here to remember that when we change variables to q and p, we have to express the function H in terms of the new variables in eq. (7.3) so that every occurrence of v is eliminated. By using Lagrange's equation $\frac{\partial L}{\partial q} = \dot{p}$ we find *Hamilton's equations* from eq. (7.5):

$$\dot{q} = \frac{\partial H}{\partial p} \ , \quad \dot{p} = -\frac{\partial H}{\partial q} \ . \tag{7.6}$$

The many-variable case is completely analogous, one just hangs an index i on (almost) everything; the proof is almost identical. The general form is

$$H = \sum_i \dot{q}_i p_i - L \ , \tag{7.7}$$

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \ , \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \ . \tag{7.8}$$

Example: *In order to understand these equations, we examine what they say for a rectilinear coordinate, where we have* $K = \frac{1}{2}m\dot{x}^2$ *and* $L = \frac{1}{2}m\dot{x}^2 - V(x)$. *Then* $p = \frac{\partial L}{\partial \dot{x}} = m\dot{x}$ *and*

$$H = \dot{x}p - L = \frac{p^2}{m} - \tfrac{1}{2}m(\frac{p}{m})^2 + V(x) = \frac{p^2}{2m} + V(x) \ , \tag{7.9}$$

which is the sum of kinetic and potential energy. This statement is actually general (as long as there is no explicit time-dependence in L*). Hamilton's equations take the form*

$$\dot{x} = \frac{\partial H}{\partial p} = \frac{p}{m} \ , \quad \dot{p} = -\frac{dV}{dx} \ . \tag{7.10}$$

We notice that instead of one second order differential equation we get two first order ones. The first one can be seen as defining p*, and when it is inserted in the second one one obtains*

$$m\ddot{x} = -\frac{dV}{dx} \ , \tag{7.11}$$

which of course is the "usual" equation of motion.

Some things become very clear in the hamiltonian framework. In particular, conserved quantities (quantities that do not change with time) emerge naturally. Consider, for example, the case where the hamiltonian does not depend on a certain coordinate q_k. Then Hamilton's equations immediately tells us that the corresponding momentum is conserved, since $\dot{p}_k = -\frac{\partial H}{\partial q_k} = 0$.

Example: *For a rectilinear coordinate q, when the potential does not depend on it, we obtain the conserved quantity p associated with q. This is not surprising — we know that momentum is conserved in the absence of force.*

Example: *In polar coordinates, with a central potential, we have*

$$H = \frac{1}{2m}\big(p_r^2 + \frac{p_\phi^2}{r^2}\big) + V(r) \ , \tag{7.12}$$

which is independent of ϕ, so that the associated momentum p_ϕ, the angular momentum, is conserved. This is exactly what we are used to in the absence of torque. Note that $\frac{p_\phi^2}{2mr^2} + V(r)$ is the "effective potential energy" used for motion under a central force.

It should be mentioned that this also can be seen in Lagrange's framework — if the lagrangian does not depend on q_k, Lagrange's equation associated with that variable becomes $\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} = 0$, telling that $p_k = \frac{\partial L}{\partial \dot{q}_k}$ is conserved.

We can also draw the conclusion that the hamiltonian H itself is conserved:

$$\dot{H} = \sum_i \frac{\partial H}{\partial q_i}\dot{q}_i + \frac{\partial H}{\partial p_i}\dot{p}_i = \sum_i \frac{\partial H}{\partial q_i}\frac{\partial H}{\partial p_i} + \frac{\partial H}{\partial p_i}(-\frac{\partial H}{\partial q_i}) = 0 \ . \tag{7.13}$$

This states the conservation of energy, which is less direct in Lagrange's formulation.

Finally, we will do some formal development in the hamiltonian formalism. Exactly as we calculated the time derivative of the hamiltonian in equation (7.13), the time derivative of any function on phase space $A(q_i, p_i)$ is calculated as

$$\dot{A} = \sum_i \frac{\partial A}{\partial q_i}\dot{q}_i + \frac{\partial A}{\partial p_i}\dot{p}_i = \sum_i \frac{\partial A}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial H}{\partial q_i} \ . \tag{7.14}$$

If one defines the *Poisson bracket* between two functions A and B on phase space as

$$\{A, B\} = \sum_i \frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial B}{\partial q_i} \ , \tag{7.15}$$

the equation of motion for any function A is stated as

$$\dot{A} = \{A, H\} \ . \tag{7.16}$$

The equations of motion for q_i and p_i, Hamilton's equations (7.8), are special cases of this (show it!). The Poisson brackets for the phase space variables q_i and p_i are

$$\begin{aligned} \{q_i, q_j\} &= 0 \ , \\ \{q_i, p_j\} &= \delta_{ij} \ , \\ \{p_i, p_j\} &= 0 \ . \end{aligned} \tag{7.17}$$

This type of formal manipulations do not have much relevance to actual problem-solving in classical mechanics. It is very valuable, though, when analyzing the behavior of some given system in field or particle theory. It is also a powerful tool for handling systems with constraints. Here, it is mainly mentioned because it opens the door towards quantum mechanics. *One* way of going from a classical to a quantum system is to replace the Poisson bracket by $(-i)$ times the *commutator* $[A, B] = AB - BA$. This means that one has $xp - px = -i$, position and momentum no longer commute. The momentum can actually be represented as a space derivative, $p = i\frac{\partial}{\partial x}$. The two variables become "operators", and their values can not be simultaneously given specific values, because ordinary numbers commute. This leads to Heisenberg's "uncertainty principle", stating the impossibility of performing measurements on both x and p simultaneously beyond a maximal precision.

Exercises

24. Any exercise from Chapter 5, with special emphasis on finding the conserved quantities.

8. Systems with Constraints

In quite many applications it happens that one does not simply want to minimize a certain functional, but to do it under certain conditions. We will investigate how this is done.

> Example: *A mathematical pendulum is confined to move on constant radius from the attachment point of the string. It is then easy to choose just the angle from the vertical as a generalized coordinate and write down the lagrangian. Another formulation of the problem would be to extremize the integral of the lagrangian* $L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2) + mgr\cos\phi$ *under the extra constraint that* $r = a$.

In the above example, the formulation with a constraint was unnecessary, since it was easy to find a generalized coordinate for the only degree of freedom. Sometimes it is not so. Consider an other example.

> Example: *A particle (a small bead) moves in two dimensions* (x, y), *where* x *is horizontal and* y *vertical, and it slides on a track whose form is described by a function* $y = f(x)$. *A natural generalized coordinate would be the distance from one specific point on the curve measured along the track, the* x*-coordinate, or something else. This can all be done, but it is easier to formulate the problem the other way: extremize the action* $S = \int dtL$ *with* $L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - mgy$ *under the constraint that* $y = f(x)$.

Let us turn to how these constraints are treated. We would like to have a new action that *automatically* takes care of the constraint, so that it comes out as one of the equations of motion. This can be done as follows: we introduce a new coordinate λ that enters in tha lagrangian multiplying the constraint. The time derivative of λ does not enter the lagrangian at all. If we call the unconstrained lagrangian L_0 and the constraint $\Phi = 0$, this means that

$$L = L_0 + \lambda\Phi \ . \tag{8.1}$$

The equation of motion for λ then just gives the constraint:

$$0 = \frac{\delta S}{\delta\lambda} = -\frac{d}{dt}\frac{\partial L}{\partial\dot{\lambda}} + \frac{\partial L}{\partial\lambda} = \frac{\partial L}{\partial\lambda} = \Phi \ . \tag{8.2}$$

The extra variable λ is called a *Lagrange multiplier*. We can examine how this works in the two examples.

<u>Example:</u> *For the pendulum, we get according to this scheme,*

$$L = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2) + mgr\cos\phi + \lambda(r-a) \ , \tag{8.3}$$

from which the equations of motion follow:

$$\begin{aligned} r:&\quad m(\ddot{r} - r\dot{\phi}^2) - mg\cos\phi - \lambda = 0 \ , \\ \phi:&\quad \frac{d}{dt}(mr\dot{\phi}) + mgr\sin\phi = 0 \ , \\ \lambda:&\quad r - a = 0 \ . \end{aligned} \tag{8.4}$$

This is not yet exactly the equations we want. By inserting the last equation (the constraint) in the other two we arrive at

$$\begin{aligned} &\lambda = m(a\dot{\phi}^2 + gcos\phi) \ , \\ &\ddot{\phi} + \frac{g}{a}\sin\phi = 0 \ . \end{aligned} \tag{8.5}$$

The second of these equations is the equation of motion for ϕ, the only degree of freedom of the pendulum, and the first one gives no information about the motion, it only states what λ is expressed in ϕ.

The pattern in the example is completely general, the equation of motion for the Lagrange multiplier gives the constraint, and the equation of motion for the constrained variable gives an expression for the Lagrange multiplier in terms of the real degrees of freedom (in this case ϕ). Let us also examine the other example.

<u>Example:</u> *In the same way as above, the new lagrangian becomes*

$$L = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) - mgy + \lambda\big(y - f(x)\big) \ , \tag{8.6}$$

with the resulting equations of motion

$$\begin{aligned} x:&\quad m\ddot{x} + \lambda f'(x) = 0 \ , \\ y:&\quad m\ddot{y} + mg - \lambda = 0 \ , \\ \lambda:&\quad y - f(x) = 0 \ . \end{aligned} \tag{8.7}$$

As before, we insert the constraint in the other equations, and get, using $\frac{d^2}{dt^2}f(x) = \frac{d}{dt}(\dot{x}f'(x)) = \ddot{x}f'(x) + \dot{x}^2 f''(x)$:

$$\begin{aligned} &\ddot{x} + \frac{\lambda}{m}f'(x) = 0 \ , \\ &\ddot{x}f'(x) + \dot{x}^2 f''(x) + g - \frac{\lambda}{m} = 0 \ . \end{aligned} \tag{8.8}$$

The second of these equations can be seen as solving λ. Inserting back in *the first one yields*

$$\ddot{x} + \big(\ddot{x}f'(x) + \dot{x}^2 f''(x) + g\big) f'(x) = 0 \ , \tag{8.9}$$

or, equivalently,

$$\ddot{x} + \frac{f'(x)}{1 + f'(x)^2} \big(g + \dot{x}^2 f''(x)\big) \ , \tag{8.10}$$

which is the simplest form of the equation of motion for this system, and as far as we can get without specifying the function $f(x)$.

There are often tricks for identifying degrees of freedom and writing the lagrangian in terms of them. The Lagrange multiplier method makes that unnecessary — one just has to use the same variational principle as usual on a modified lagrangian, and everything comes out automatically.

It turns out to be theoretically very fruitful to treat constrained systems in a hamiltonian formalism. We will not touch upon that formulation here.

As a last example, we will solve a (classical) mechanical problem that is not a dynamical one.

<u>Example:</u> *Consider a string whose ends are fixed in two given points. What shape will the string form? We suppose that the string is unstretchable and infinitely flexible (i.e. it takes an infinite amount of energy to stretch it and no energy to bend it). It is clearly a matter of minimizing the potential energy of the string, under the condition that the length is some fixed number. The input for calculating the potential energy is the shape $y(x)$ of the string, so it is a functional. The principle for finding the solution must then be*

$$\frac{\delta V}{\delta y(x)} = 0 \ , \tag{8.11}$$

where $V[y(x)]$ is the potential energy. We need an explicit expression for V, and also for the length, that will be constrained to a certain value L using the Lagrange multiplier method. The length of an infinitesimal part of the string is

$$ds = \sqrt{(dx)^2 + (dy)^2} = dx\sqrt{1 + y'(x)^2} \ , \tag{8.12}$$

so that the total length and potential energy are

$$\begin{aligned} L &= \int_a^b dx\sqrt{1+y'(x)^2}\ , \\ V &= -\rho g \int_a^b dx y(x)\sqrt{1+y'(x)^2}\ . \end{aligned} \tag{8.13}$$

The technique is again to add a Lagrange multiplier term to V:

$$\begin{aligned} U = &\int_a^b dx u(y,y') = \\ &-\rho g \int_a^b dx y(x)\sqrt{1+y'(x)^2} + \lambda\left(L - \int_a^b dx\sqrt{1+y'(x)^2}\right)\ . \end{aligned} \tag{8.14}$$

Exactly as we recovered Lagrange's equations from $\frac{\delta S}{\delta x(t)} = 0$*, the variation of* U *gives the equation for* y

$$\frac{d}{dx}\frac{\partial u}{\partial y'} - \frac{\partial u}{\partial y} = 0\ . \tag{8.15}$$

Since there is no explicit x*-dependence one may use the integrated form of the equations as described in Chapter 5.2,*

$$0 = \frac{d}{dx}\left[u - y'\frac{\partial u}{\partial y'}\right] = \frac{d}{dx}\left[-(\lambda+\rho g y)\sqrt{1+y'^2} + \frac{(\lambda+\rho g y)y'}{\sqrt{1+y'^2}}\right]\ . \tag{8.16}$$

We shift y *by a constant value to get the new vertical variable* $z = y + \rho g/\lambda$*, and equation (8.16) gives*

$$\frac{d}{dx}\frac{z}{\sqrt{1+z'^2}} = 0 \quad \Longrightarrow \quad \sqrt{1+z'^2} = kz\ , \tag{8.17}$$

so that $z' = \pm\sqrt{k^2z^2-1}$*, and*

$$\frac{dx}{dz} = \pm\frac{1}{\sqrt{k^2z^2-1}}\ , \tag{8.18}$$

with the solutions

$$x(z)+c=\pm\text{arcosh}\, kz\ ,\quad z=\frac{1}{k}\cosh k(x+c)\ . \tag{8.19}$$

One may of course go back to the variable y and solve for the Lagrange multiplier λ, but here we are only interested in the types of curves formed by the hanging string — they are hyperbolic cosine curves.

Exercises

25. Show that the action $S=\int d\tau\lambda\dot{x}_\mu\dot{x}^\mu$ describes a massless relativistic particle.

Answers to Exercises

1. Four. For example the center of mass coordinates, the distance between the masses and an angle.
3. $\mathscr{F} = -mgf'(x)$
4. $E_k = \frac{1}{2}m\big(\dot{r}^2 + (r\dot{\theta})^2 + (r\dot{\phi}\sin\theta)^2\big)$
5. $E_k = \frac{1}{2}m\dot{x}^2\big(1 + f'(x)^2\big)$
6. $\ddot{r} - r(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta) = 0$
   $\frac{d}{dt}(r^2\dot{\theta}) - r^2\dot{\phi}^2\sin\theta\cos\theta = 0$
   $\frac{d}{dt}(r^2\dot{\phi}\sin^2\theta) = 0$
7. $\ddot{\theta} - \dot{\phi}^2\sin\theta\cos\theta + \frac{g}{a}\sin\theta = 0$
   $\frac{d}{dt}(\dot{\phi}\sin^2\theta) = 0$
8. $\omega = \sqrt{\frac{g}{a}}$
9. $\ddot{x}(1 + f'^2) + \dot{x}^2 f'f'' + gf' = 0$
   No.
   $\omega = \sqrt{gf''(x_0)}$
10. $a = \frac{5}{7}g\sin\alpha$
11. $m\ddot{r} - mr\dot{\phi}^2 + k(r - l) = 0$
    $\frac{d}{dt}mr^2\dot{\phi} = 0$
12. Translation + motion as in exercise 11. Note the appearance of the "reduced mass" $\mu = \frac{m_1 m_2}{m_1 + m_2}$.
13. $\omega = \omega_0$ or $\omega_0\sqrt{3}$ where $\omega_0 = \sqrt{\frac{k}{m}}$
15. $\ddot{\phi} + \frac{(a^2 - b^2)\sin\phi\cos\phi}{a^2\sin^2\phi + b^2\cos^2\phi}\dot{\phi}^2 = 0$
16. $a = \frac{m_1 - m_2}{m_1 + m_2 + \frac{I}{R^2}}g$
17. $a_1 = \frac{1-\gamma}{1+\gamma}g$ (downwards), where $\gamma = \frac{4m_2m_3}{m_1(m_2+m_3)}$
    $a_2 = -a_1 + \frac{2}{1+\gamma}\frac{m_2 - m_3}{m_2 + m_3}$
18. $a_{rel} = \frac{(M+m)\sin\alpha}{M + m\sin^2\alpha}g$
19. $\phi \approx 0\ :\ \ddot{\phi} + \frac{g}{l}\phi = \frac{a\omega^2}{l}\sin\omega t$ (forced oscillations)

20. $\arccos\frac{2}{3} \approx 49°$
21. $\phi = A\sin(\omega_1 t + \alpha_1) + B\sin(\omega_2 t + \alpha_2)$
$\theta = -2A\sin(\omega_1 t + \alpha_1) + B\sin(\omega_2 t + \alpha_2)$
$A = 0$
$B = 0$
25. Consider the meaning of the obtained constraint!

References

There is a great number of good books relevant for the further study of analytical mechanics. I just list a few here.

A.P. Arya, *"Introduction to Classical Mechanics"* (Allyn and Bacon, Boston, 1990).

E.T. Whittaker, *"A Treatise on the Analytical Dynamics of Particles and Rigid Bodies"* (Cambridge Univ. Press, 1988).

F. Eriksson, *"Variationskalkyl"* (kompendium, CTH).

G.M. Ewing, *"Calculus of Variations with Applications"* (Norton, New York, 1969).

Translation Table

English term	*Swedish term*	*Symbol*
acceleration	acceleration	$\mathbb{a}$
angular frequency	vinkelfrekvens	ω
angular momentum	rörelsemängdsmoment	$\mathbb{L}$
angular velocity	vinkelhastighet	ω , $\vec{\omega}$
center of mass	masscentrum	$\mathbb{R}$
coefficient of friction	friktionskoefficient	μ , f
collision	stöt	
elastic	elastisk	
inelastic	oelastisk	
conservation	bevarande	
constant of motion	rörelsekonstant	
constraint	tvång	
cross section	tvärsnitt	
curl	rotation	$\nabla\times$, curl
damping	dämpning	
density	densitet	ρ
derivative	derivata	
partial	partiell	
displacement	förskjutning	
energy	energi	E
kinetic	kinetisk, rörelse-	K
potential	potentiell, läges-	U , V
equilibrium	jämvikt	
event	händelse	
force	kraft	$\mathbb{F}$
central	central-	
conservative	konservativ	
fictitious	fiktiv	
inertial	tröghets-	
nonconservative	ickekonservativ	
four-vector	fyrvektor	
gradient	gradient	∇ , grad
gravity	tyngdkraft, gravitation	
harmonic	harmonisk	
impulse	impuls	I , $\int_{t_1}^{t_2} \mathbb{F} dt$
inertia	tröghet	
inertial system	inertialsystem	

initial conditions	begynnelsevillkor	
interaction	växelverkan	
mass	massa	m
gravitational	tung	
inertial	trög	
rest	vilo-	
magnitude	belopp	$\| \ \|$
moment of inertia	tröghetsmoment	I
(linear) momentum	rörelsemängd	$\mathbb{p}$
orbit	omlopp, bana	
oscillation	svängning	
path	väg, bana	
perturbation	störning	
power	effekt	P
pulley	talja	
resonance	resonans	
rigid body	stel kropp	
scattering angle	spridningsvinkel	
simultaneity	samtidighet	
speed	fart	v
spring	fjäder	
tension	spänning	
tensor of inertia	tröghetstensor/-matris	$\tilde{\boldsymbol{I}}$
torque	vridande moment	$\vec{\tau}$
trajectory	bana	
velocity	hastighet	$\mathbb{v}$
wedge	kil	
weight	tyngd	
work	arbete	W