Hanno Essén

The Theory of Lagrange's Method

KTH MECHANICS
STOCKHOLM 2013

1 Introduction

This compendium is the result of a merger in 2006 of two shorter texts by me. Input form Arne Nordmark 2013 has resulted in a minor revision. The first part derives and discusses Lagrange's method for a single particle (Sections 2 – 11). After that the general theory for systems of particles (in particular rigid bodies) is presented. These two parts have simply been concatenated so there is some redundancy.

Joseph Louis Lagrange (1763-1813) discovered a method that makes it possible to simply derive the equations of motion for systems with constraints, *i.e.* systems of particles and rigid bodies that are connected in various ways. One obtains the equations of motion using Lagrange's method by differentiating energy expressions. One of its main advantages is that one can use arbitrary coordinates as long as they fully describe the configuration of the system. In this way one easily handles constraints that can be expressed by demanding that certain coordinates do not vary, *i.e.* have fixed values. Such constraints are called *holonomic*. Lagrange's method then gives equations of motion for the remaining unconstrained coordinates.

Normally Lagrange's method is not needed for problems with a single particle. For such problems its strength lies mainly on a fundamental level. Practically the method is best suited for problems involving *mechanisms*, linked systems of rigid bodies. However, when deriving Lagrange's method for an arbitrary system of particles, it is easy to loose orientation among all the, seemingly, uncountable summations and differentiations, that are done. Therefore we will derive the method here for first *a single particle*. Later we show that an N-particle system can be viewed mathematically as a single particle in a space with $3N$ dimensions. Most of the single particle definitions and derivations can then be reused with small adjustments.

2 Generalized coordinates and constraints

The position of a particle in three dimensional space can be specified by means of the values of three coordinates,

$$q_1, q_2, q_3, \tag{1}$$

that are suitably defined. Examples are Cartesian, or rectangular, coordinates (x, y, z), cylindrical coordinates (ρ, φ, z) and spherical coordinates (r, φ, θ). The position vector, $\boldsymbol{r}$, of the particle can then be viewed as a function of these coordinates,

$$\boldsymbol{r} = \boldsymbol{r}(q_1, q_2, q_3). \tag{2}$$

Such arbitrary coordinates are called *generalized coordinates*.

Three coordinates are not always needed to determine the position of the particle. For a simple pendulum, for example, all you need is a single angle (φ), for a particle in a horizontal plane you need two coordinates (x, y), etc. One then says that there are *holonomic constraints* reducing the mobility of the particle. The position vector can the be written as $\boldsymbol{r}(q_1)$, or as $\boldsymbol{r}(q_1, q_2)$, depending on whether one or two coordinates are needed. The number, n, of coordinates required is called the number of of *degrees-of-freedom*. For a single particle then this number can be zero, one, two, or three ($n = 0, 1, 2,$ or, 3). We will thus write

$$\boldsymbol{r} = \boldsymbol{r}(q_1, \ldots, q_n) = \boldsymbol{r}(q) \tag{3}$$

for the position vector; by just q we mean collectively all n generalized coordinates of the problem.

If s is the *number of constraints* we have that,

$$n = 3 - s. \tag{4}$$

I.e. the number of degrees-of-freedom n is three minus the number of constraints s. Here three is the dimension of space and a constraint is given by an equation that limits the mobility of the particle. If *e.g.* the particle must stay at the distance R from the origin one has,

$$|\boldsymbol{r}| = R. \tag{5}$$

This equation is a constraint and if it is obeyed there remain two degrees-of-freedom – the particle can move on the surface of a sphere of radius R. In general the *constraints* are given by s expressions, or equations,

$$f_k(\boldsymbol{r}) = 0, \qquad k = 1, \ldots, s, \tag{6}$$

that the position vector $\boldsymbol{r}$ of the particle must obey.

3 Curves, surfaces and tangent vectors

It is intuitively important to understand that the expression $\boldsymbol{r}(q_1)$ can be understood as the parametric equation of a *curve* (with q_1 as parameter). This curve is a one-dimensional "manifold" on which the particle can move while obeying the constraints, when $s = 2$. Note that a *tangent vector* of this curve can be obtained through differentiation,

$$\boldsymbol{\tau}_1 = \frac{\partial \boldsymbol{r}}{\partial q_1}. \tag{7}$$

Similarly the expression $\boldsymbol{r}(q_1, q_2)$ gives parametric equations for a *surface*. The surface is the two dimensional manifold that the particle can move on, in accordance with the constraint ($s = 1$). A surface has two (linearly independent) tangent vectors,

$$\boldsymbol{\tau}_1 = \frac{\partial \boldsymbol{r}}{\partial q_1}, \quad \boldsymbol{\tau}_2 = \frac{\partial \boldsymbol{r}}{\partial q_2}, \tag{8}$$

see Figure 1. If there are no constraints one can view $\boldsymbol{r}(q) = \boldsymbol{r}(q_1, q_2, q_3)$ as a parametrization of three dimensional space., and the corresponding three tangent vectors,

$$\boldsymbol{\tau}_i(q) = \frac{\partial \boldsymbol{r}}{\partial q_i}, \quad i = 1, 2, 3, \tag{9}$$

represent the three different directions in which one moves when one increases the coordinate q_i, while keeping the two others constant.

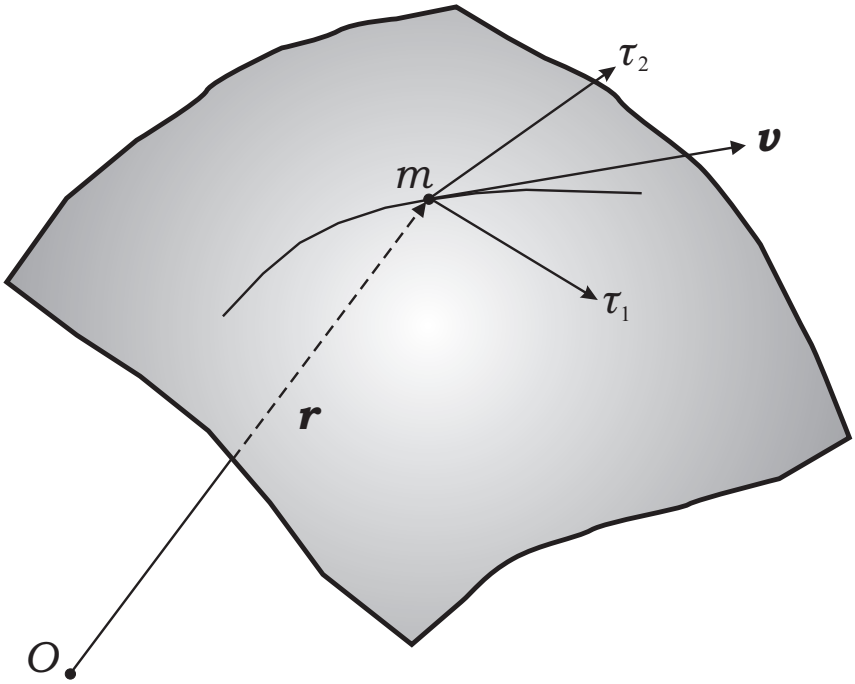

Figure 1: A particle, m, moves on a fixed surface $\boldsymbol{r} = \boldsymbol{r}(q_1, q_2)$. The two vectors $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2$ are tangent vectors to the surface at the point $\boldsymbol{r}(q_1, q_2)$. The trajectory of the particle, the curve $\boldsymbol{r}(t) = \boldsymbol{r}(q_1(t), q_2(t))$, is is also indicated. The velocity vector $\boldsymbol{v}$ is tangent to this curve, and for time independent constraints (*i.e.* fixed surface, $\boldsymbol{\tau}_t = \boldsymbol{0}$) the velocity vector lies in the plane spanned by $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2$. According to (12) we here have $\boldsymbol{v} = \dot{q}_1\boldsymbol{\tau}_1 + \dot{q}_2\boldsymbol{\tau}_2$.

4 Generalized velocities

One gets the velocity vector of a particle by time differentiating the position vector $\boldsymbol{r} = \boldsymbol{r}(q,t)$. The position vector will change in time because the generalized coordinates $q_i = q_i(t)$ change their values as the particle alters its position. The position vector may, however, also depend explicitly on time t. It will do so if the constraints are time dependent. The curves or surfaces defined by the constraint equations will then move in space independently of the particle motion. We can thus write the velocity vector in the form,

$$\boldsymbol{v} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \sum_{j=1}^{n} \frac{\partial \boldsymbol{r}}{\partial q_j}\frac{\mathrm{d}q_j}{\mathrm{d}t} + \frac{\partial \boldsymbol{r}}{\partial t}. \tag{10}$$

Here the last term is the contribution due to time dependent constraints. The time derivatives of the generalized coordinates, $\dot{q}_i = \mathrm{d}q_i/\mathrm{d}t$, are called *generalized velocities*. If we introduce the notation,

$$\boldsymbol{\tau}_t(q,t) = \frac{\partial \boldsymbol{r}}{\partial t}, \tag{11}$$

and use the definition of tangent vectors in the equations (7) – (9), we can write,

$$\boldsymbol{v} = \sum_{j=1}^{n} \dot{q}_j \boldsymbol{\tau}_j + \boldsymbol{\tau}_t. \tag{12}$$

Note that tangent vectors in general depend on q, and, in the time dependent constraint case, also on time t, so that,

$$\boldsymbol{\tau}_i = \boldsymbol{\tau}_i(q,t). \tag{13}$$

We can thus write our velocity vector,

$$\boldsymbol{v}(q,\dot{q},t) = \sum_{j=1}^{n} \dot{q}_j \boldsymbol{\tau}_j(q,t) + \boldsymbol{\tau}_t(q,t). \tag{14}$$

It is important to notice that its dependence on the *generalized velocities* is linear. This linearity implies that the derivative,

$$\frac{\partial \boldsymbol{v}}{\partial \dot{q}_i} = \boldsymbol{\tau}_i(q,t), \tag{15}$$

is independent of the generalized velocities $\dot{q}$. Note that the two results (9) and (15), with $\boldsymbol{v} = \dot{\boldsymbol{r}}$, together give,

$$\frac{\partial \boldsymbol{r}}{\partial q_i} = \frac{\partial \dot{\boldsymbol{r}}}{\partial \dot{q}_i} = \boldsymbol{\tau}_i, \tag{16}$$

i.e. two ways of expressing the tangent vectors. Formally the expression says that it is allowed to ”cancel the dots”.

5 The equations of motion

The equations of motion for one particle are given by

$$\dot{\boldsymbol{p}} = \boldsymbol{F}, \tag{17}$$

where, $\boldsymbol{p} = m\boldsymbol{v}$. This vector equation has normally three components along suitable basis vectors. But if we have constraints we only really need one equation of motion per generalized coordinate q_i. The number of equations of motion needed is thus equal to the number of degrees-of-freedom n.

Once the choice of suitable generalized coordinates q_i has been made the best basis vectors are in general the tangent vectors $\boldsymbol{\tau}_i$. If one then projects the vector equation (17) on the n tangent vectors one gets n equations of motion that are adapted to the n generalized coordinates. These projected equations of motion can be written,

$$\dot{\boldsymbol{p}} \cdot \boldsymbol{\tau}_i = \boldsymbol{F} \cdot \boldsymbol{\tau}_i, \qquad i = 1, \ldots, n, \tag{18}$$

by means of the scalar product.

Notice that this is analogous to using the components of acceleration along, so called, "moving" basis vectors. In *e.g.* cylindrical coordinates these are $\boldsymbol{e}_\rho(\varphi), \boldsymbol{e}_\varphi(\varphi)$ and the procedure gives the components of the equations of motion along the radial and the transverse (or azimuthal) directions. Here we assume the constraint $z = 0$. Note however, that the tangent vectors $\boldsymbol{\tau}_i$, as defined above, not necessarily are unit vectors; this is clearly irrelevant for our purposes.

6 Lagrange's method

To get explicit equations of motion using Eq. (18) can be difficult. Lagrange discovered an elegant short cut to these projected equations of motion. One can the left hand sides by differentiating the kinetic energy in a specific way. We will now show that,

$$\dot{\boldsymbol{p}} \cdot \boldsymbol{\tau}_i = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_i}\right) - \left(\frac{\partial T}{\partial q_i}\right), \tag{19}$$

where T is the kinetic energy. We do this by studying the two terms of the right hand side one at the time.

The kinetic energy can be written in the form,

$$T = \frac{1}{2} m\boldsymbol{v} \cdot \boldsymbol{v} \tag{20}$$

and therefore we get,

$$\frac{\partial T}{\partial \dot{q}_i} = \frac{1}{2} m \frac{\partial}{\partial \dot{q}_i}\boldsymbol{v} \cdot \boldsymbol{v} = \frac{1}{2} m \left(\frac{\partial \boldsymbol{v}}{\partial \dot{q}_i} \cdot \boldsymbol{v} + \boldsymbol{v} \cdot \frac{\partial \boldsymbol{v}}{\partial \dot{q}_i}\right) = m\boldsymbol{v} \cdot \frac{\partial \boldsymbol{v}}{\partial \dot{q}_i} = \boldsymbol{p} \cdot \boldsymbol{\tau}_i, \tag{21}$$

where we have used, $\boldsymbol{p} = m\boldsymbol{v}$ and Eq. (15) in the last step. If we now take the time derivative of this we get,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_i}\right) = \frac{\mathrm{d}}{\mathrm{d}t} m\boldsymbol{v}\cdot\boldsymbol{\tau}_i = m\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}\cdot\boldsymbol{\tau}_i + m\boldsymbol{v}\cdot\frac{\mathrm{d}\boldsymbol{\tau}_i}{\mathrm{d}t} = \dot{\boldsymbol{p}}\cdot\boldsymbol{\tau}_i + \boldsymbol{p}\cdot\frac{\mathrm{d}\boldsymbol{\tau}_i}{\mathrm{d}t}, \quad (22)$$

but,

$$\frac{\mathrm{d}\boldsymbol{\tau}_i}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \boldsymbol{r}}{\partial q_i} = \frac{\partial}{\partial q_i}\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \frac{\partial \boldsymbol{v}}{\partial q_i} \quad (23)$$

since one can interchange the order of the two differentiations. Summarizing, we have obtained,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_i}\right) = \dot{\boldsymbol{p}}\cdot\boldsymbol{\tau}_i + \boldsymbol{p}\cdot\frac{\partial \boldsymbol{v}}{\partial q_i}. \quad (24)$$

If we can now get rid of the last term we have something that is the projection of $\dot{\boldsymbol{p}}$ on $\boldsymbol{\tau}_i$.

In order to arrive at Eq. (19) we now calculate the second term on the right hand side. We get,

$$\frac{\partial T}{\partial q_i} = \frac{1}{2}m\frac{\partial}{\partial q_i}\boldsymbol{v}\cdot\boldsymbol{v} = \frac{1}{2}m\left(\frac{\partial \boldsymbol{v}}{\partial q_i}\cdot\boldsymbol{v} + \boldsymbol{v}\cdot\frac{\partial \boldsymbol{v}}{\partial q_i}\right) = m\boldsymbol{v}\cdot\frac{\partial \boldsymbol{v}}{\partial q_i} = \boldsymbol{p}\cdot\frac{\partial \boldsymbol{v}}{\partial q_i}, \quad (25)$$

and we immediately see that if this is subtracted from (24) we get $\dot{\boldsymbol{p}}\cdot\boldsymbol{\tau}_i$. This is what we wanted to show.

7 Generalized forces

We have now shown that the equations of motion, according to (18) and (19), can be written,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_i}\right) - \left(\frac{\partial T}{\partial q_i}\right) = \boldsymbol{F}\cdot\boldsymbol{\tau}_i. \quad (26)$$

If we now define the *generalized forces*,

$$Q_i = \boldsymbol{F}\cdot\boldsymbol{\tau}_i, \quad (27)$$

we get the following expression for the n equations of motion:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_i}\right) - \left(\frac{\partial T}{\partial q_i}\right) = Q_i \quad i = 1, \ldots, n. \quad (28)$$

This is the most *general form* of Lagrange's equations.

How are we to understand the generalized forces? Recall that we are assuming that the position of the particle can be written $\boldsymbol{r} = \boldsymbol{r}(q,t)$. A small displacement of the position is then given by:

$$\mathrm{d}\boldsymbol{r} = \sum_{j=1}^{n} \frac{\partial \boldsymbol{r}}{\partial q_j}\mathrm{d}q_j + \frac{\partial \boldsymbol{r}}{\partial t}\mathrm{d}t = \sum_{j=1}^{n} \boldsymbol{\tau}_j \mathrm{d}q_j + \boldsymbol{\tau}_t \mathrm{d}t. \tag{29}$$

The work done by the force on the particle is thus,

$$dW = \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r} = \boldsymbol{F} \cdot \left(\sum_{j=1}^{n} \boldsymbol{\tau}_j \mathrm{d}q_j + \boldsymbol{\tau}_t \mathrm{d}t \right). \tag{30}$$

By means of (27), and the notation,

$$Q_t = \boldsymbol{F} \cdot \boldsymbol{\tau}_t, \tag{31}$$

we now find that,

$$dW = \sum_{j=1}^{n} Q_j \mathrm{d}q_j + Q_t \mathrm{d}t. \tag{32}$$

We thus find that the *generalized force* Q_i has the property that,

$$dW_i = Q_i \mathrm{d}q_i \tag{33}$$

is the *work* that is done on the particle if q_i is increased by $\mathrm{d}q_i$ while the remaining q:s, and time, are held constant.

8 Conservative forces

If the work W, at least locally, is a function of q and t, and this function is,

$$W = W(q,t), \tag{34}$$

then, by definition, the *differential* of W is,

$$\mathrm{d}W = \sum_{j=1}^{n} \frac{\partial W}{\partial q_j}\mathrm{d}q_j + \frac{\partial W}{\partial t}\mathrm{d}t. \tag{35}$$

If we compare this expression with (32) and put $dW = \mathrm{d}W$, then we can immediately identify the coefficients in front of coordinate differentials and get that,

$$Q_i = \frac{\partial W}{\partial q_i}. \tag{36}$$

Note carefully that this is *not true* in the general case. The infinitesimal work dW is normally not a differential of a function of q and t. Instead it may also depend, for example, on the velocities.

If the force on the particle is *conservative* the work on the particle has the property that it is the negative of the potential energy function,

$$V(q,t) = -\int_{\boldsymbol{r}_0}^{\boldsymbol{r}(q,t)} \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r}. \tag{37}$$

Note that the (line or curve)integral here must be independent of the path of integration between an arbitrary start point $\boldsymbol{r}_0$ and the end point $\boldsymbol{r}(q,t)$ in order to define a function of the end point coordinates only. When this is the case one finds that the work is given by

$$W(q,t) = -V(q,t). \tag{38}$$

If we now compare with Eq. (36) we find that the *generalized force*,

$$Q_i = -\frac{\partial V}{\partial q_i}. \tag{39}$$

For *conservative* forces therefore we find that the generalized force Q_i is minus the partial derivative of the potential energy with respect to q_i.

9 The Lagrange function

In the conservative case we now have that the equations of motion (28) can be written,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_i}\right) - \left(\frac{\partial T}{\partial q_i}\right) = -\frac{\partial V}{\partial q_i} \quad i = 1, \ldots, n. \tag{40}$$

If we form the new function,

$$L = L(q,\dot{q}) = T(q,\dot{q}) - V(q), \tag{41}$$

the, so called, *Lagrange function*, or the Lagrangian, these equations take the form,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \left(\frac{\partial L}{\partial q_i}\right) = 0 \quad i = 1, \ldots, n. \tag{42}$$

This elegant form of the equations of motion, often referred to as the *Euler-Lagrange equations*, can thus be used when we are dealing with conservative forces.

The Euler-Lagrange equations can also be obtained in a fundamentally different way. They turn out to be the equations that determine the solution $q(t)$ that minimizes the *action*,

$$S[q(t)] = \int L(q, \dot{q})\, \mathrm{d}t, \tag{43}$$

a functional of the path $q(t)$ (*i.e.* a mapping from a path to a real number). Among all possible paths $q(t)$, the one that the system actually follows is the one that minimizes the action functional. This is called the *principle of least action*. The variation of the action at the correct path must be zero, $\delta S = 0$. This is presented in more detail in Sec. 18.1.

10 Velocity dependent forces

In the previous section we showed that it is possible to introduce a Lagrange function $L(q, \dot{q})$ when the forces are conservative with potential energy $V(q)$. Should the forces be velocity dependent this simple recipe will not work. Known cases of velocity dependent forces are the *Lorentz force* on a charged particle in a magnetic field, $\boldsymbol{B}$,

$$\boldsymbol{F}_L = \frac{e}{c}\boldsymbol{v} \times \boldsymbol{B}, \tag{44}$$

(Gaussian units, c is the speed of light) and the fictitious *Coriolis force*,

$$\boldsymbol{F}_C = 2m\boldsymbol{v} \times \boldsymbol{\omega}, \tag{45}$$

that acts on a particle according to an observer using a rotating reference frame, rotating with angular velocity vector $\boldsymbol{\omega}$. We now show how such forces can be included in the Lagrangian formalism.

When the generalized force Q_i is the negative partial derivative of a scalar function (a potential energy) then we can introduce a Lagrange function (41) and get the equations of motion on the form (42). But the assumption $Q_i = -\partial V/\partial q_i$ is clearly not the most general possible. If we start from the Lagrange equations in the general form (28) we see that, generalized forces on the form,

$$Q_i = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial U}{\partial \dot{q}_i}\right) - \left(\frac{\partial U}{\partial q_i}\right), \tag{46}$$

allow equations of motion on the form (42) using a Lagrange function defined by

$$L(q, \dot{q}) = T(q, \dot{q}) - U(q, \dot{q}). \tag{47}$$

The function $U(q, \dot{q})$ is then called the *work function*. One notes that the conservative case, with a potential energy function V, is obtained for the special case that U does not depend on the generalized velocities. Let us now consider the two main cases of interest for this formalism: a charged particle in an electromagnetic field, and a particle moving in an accelerated reference frame.

10.1 Particle in electromagnetic field

A particle with electric charge e that moves under the influence of external electric and magnetic fields, can be characterized by the Lagrange function,

$$L(\boldsymbol{r}, \dot{\boldsymbol{r}}, t) = \frac{1}{2} m \dot{\boldsymbol{r}}^2 - e \left(\Phi(\boldsymbol{r}, t) - \frac{1}{c} \dot{\boldsymbol{r}} \cdot \boldsymbol{A}(\boldsymbol{r}, t) \right). \tag{48}$$

Here, Φ, is the electric potential and, $\boldsymbol{A}$, the so called vector potential. These are defined so that the electric, $\boldsymbol{E}$, and magnetic, $\boldsymbol{B}$, fields are obtained by differentiations,

$$\boldsymbol{E} = -\nabla\Phi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}, \tag{49}$$

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}. \tag{50}$$

Here, $\nabla = \frac{\partial}{\partial x}\boldsymbol{e}_x + \frac{\partial}{\partial y}\boldsymbol{e}_y + \frac{\partial}{\partial z}\boldsymbol{e}_z$, is the "nabla" or del-operator. The work function of the Lagrangian, $L = T - U$, is in this case is given by,

$$U(\boldsymbol{r}, \dot{\boldsymbol{r}}, t) = e \left(\Phi(\boldsymbol{r}, t) - \frac{1}{c} \dot{\boldsymbol{r}} \cdot \boldsymbol{A}(\boldsymbol{r}, t) \right). \tag{51}$$

After some calculation one gets for the x-component of the equations of motion,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{x}}\right) - \left(\frac{\partial T}{\partial x}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial U}{\partial \dot{x}}\right) - \left(\frac{\partial U}{\partial x}\right), \tag{52}$$

the result,

$$m\ddot{x} = eE_x + \frac{e}{c}(\dot{y}B_z - \dot{z}B_y). \tag{53}$$

Adding this, and the corresponding analogous y and z-components, vectorially, gives the vector equation of motion,

$$m\ddot{\boldsymbol{r}} = e\boldsymbol{E} + \frac{e}{c}\dot{\boldsymbol{r}} \times \boldsymbol{B}. \tag{54}$$

This is the equation of motion for a charged particle moving in an electromagnetic field.

10.2 Particle in accelerated reference system

Assume that we chose to study the motion of a particle in an accelerated reference frame. We select a coordinate system at rest in this frame. This system may then have a translational acceleration $\boldsymbol{a}_0$ relative to an inertial frame. It may also rotate relative to the inertial frame with an angular velocity (vector) $\boldsymbol{\omega}$. We select our origin in the accelerated coordinate system so that the origin is on the rotation

axis. In elementary mechanics one then derives the equations of motion, valid in this accelerated system, by means of Coriolis' theorem.

One then finds that the translational acceleration gives rise to a fictitious force,

$$\boldsymbol{F}_a = -m\boldsymbol{a}_0, \tag{55}$$

which is of the same character as the acceleration due to gravity. This force thus has potential energy,

$$V_a(\boldsymbol{r}) = m\boldsymbol{a}_0 \cdot \boldsymbol{r}. \tag{56}$$

The rotation requires three fictitious forces

$$\begin{aligned} \boldsymbol{F}_r &= -m\,\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}) = m\,\omega^2 \boldsymbol{r}_\perp, & (57)\\ \boldsymbol{F}_t &= m\,\boldsymbol{r} \times \dot{\boldsymbol{\omega}}, & (58)\\ \boldsymbol{F}_C &= 2m\,\boldsymbol{v} \times \boldsymbol{\omega}. & (59) \end{aligned}$$

The first of these, the *centrifugal force,* is directed radially outwards from the rotation axis. The centrifugal force has a potential energy,

$$V_r(\boldsymbol{r}) = -\frac{1}{2}m(\boldsymbol{\omega} \times \boldsymbol{r})^2 = -\frac{1}{2}m\omega^2 r_\perp^2. \tag{60}$$

The second of these forces only appear if the angular velocity is not constant. The third is the, non-working, *Coriolis force,* a fictitious force that is perpendicular to the relative velocity $\boldsymbol{v}$.

The equations of motion for a particle in such an accelerated system can be obtained from the Lagrange function,

$$L(\boldsymbol{r}, \dot{\boldsymbol{r}}, t) = \frac{1}{2}m\dot{\boldsymbol{r}}^2 - U(\boldsymbol{r}, \dot{\boldsymbol{r}}, t), \tag{61}$$

where the work function U is given by,

$$U(\boldsymbol{r}, \dot{\boldsymbol{r}}, t) = V(\boldsymbol{r}, t) + V_a(\boldsymbol{r}, t) + V_r(\boldsymbol{r}, t) - m\dot{\boldsymbol{r}} \cdot (\boldsymbol{\omega} \times \boldsymbol{r}), \tag{62}$$

or, more explicitly,

$$U(\boldsymbol{r}, \dot{\boldsymbol{r}}, t) = V(\boldsymbol{r}, t) + m\boldsymbol{a}_0(t) \cdot \boldsymbol{r} - \frac{1}{2}m[\boldsymbol{\omega}(t) \times \boldsymbol{r}]^2 - m\dot{\boldsymbol{r}} \cdot [\boldsymbol{\omega}(t) \times \boldsymbol{r}]. \tag{63}$$

Here $V(\boldsymbol{r}, t)$ is the potential energy of the real (non-fictitious) forces that act on the particle even when viewed from an inertial reference frame.

The equations of motion corresponding to this Lagrangian are,

$$m\ddot{\boldsymbol{r}} = -\nabla V + \boldsymbol{F}_a + \boldsymbol{F}_r + \boldsymbol{F}_t + \boldsymbol{F}_C. \tag{64}$$

Here it is assumed that the position $\boldsymbol{r}$, the velocity $\boldsymbol{v} = \dot{\boldsymbol{r}}$, and the acceleration $\ddot{\boldsymbol{r}}$, are all measured relative (by an observer fixed in) the accelerated system. It is also assumed that the origin of the relative position vector $\boldsymbol{r}$ is on the rotation axis.

11 Generalized quantities

The following generalized quantities are introduced in the Lagrange formalism:

- Generalized coordinates: q_i
- Generalized velocities: $\dot{q}_i = \frac{\mathrm{d}q_i}{\mathrm{d}t}$
- Generalized forces: $Q_i = -\frac{\partial V}{\partial q_i}$, or $Q_i = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial U}{\partial \dot{q}_i}\right) - \left(\frac{\partial U}{\partial q_i}\right)$
- Generalized momenta: $p_i = \frac{\partial L}{\partial \dot{q}_i}$

The expressions for the generalized forces assume that there are no dissipative forces acting. For more general cases one must use $Q_i = \boldsymbol{F} \cdot \boldsymbol{\tau}_i$.

Only the *generalized momenta*, in the list above, have not been introduced before. Using them Lagrange's equations, in the non-dissipative case (42), can be written,

$$\dot{p}_i = \frac{\partial L}{\partial q_i}, \quad i = 1, \ldots, n. \tag{65}$$

It may happen that a certain generalized coordinate q_i does not appear in the Lagrange function, although the corresponding generalized velocity $\dot{q}_i$ does appear. In that case the partial derivative, $\partial L/\partial q_i = 0$, and from Eq. (65) there then follows that the corresponding generalized momentum is constant: $p_i = \text{const}$. The coordinate q_i is then said to be *cyclic*, and p_i is a *constant of the motion*.

12 Introduction – Lagrange's method for systems

Except for of certain advanced variational principles there are no deeper principles in mechanics than those given by Newton's laws. Lagrange's equations, that we are going, to present here, are therefore best seen as mathematical consequences of these laws. Some textbooks emphasize that Lagrange's equations arise when Newton's laws are transformed to arbitrary curvilinear coordinates and this is at least part of the truth.

The intellectually most economic point of view is that Lagrange's equations arise by projection of Newton's equations ($\dot{\boldsymbol{p}}_k = \boldsymbol{F}_k,\ k = 1, \ldots, K$) from some original unconstrained configuration space ($3K$-dimensional for K particles) to a lower dimensional configuration space consistent with the (holonomic) constraints. This space can be seen as a surface embedded in the original space.

Here we start with a discussion about constraints and degrees of freedom. We then introduce a notation for treating systems of particles in a compact way. Normally texts on analytical mechanics are full of sums over x,y, and z as well as sums over the particles of a system. With suitable vector notation most of these can be avoided.

13 Degrees of Freedom, Constraints and Rigidity

The number of *degrees of freedom* of a

$$\text{system of particles } = \{\{m_k, \boldsymbol{r}_k\};\ k = 1, 2, \ldots, K\}. \tag{66}$$

is the number of coordinates needed to completely specify its position in space. An arrangement which forces particles to behave in some special way, and thus reduces the degrees of freedom is called a *constraint*. These concepts are best illustrated by some examples.

A particle which can move freely in space has *three* degrees of freedom since its position is completely specified by, for example, its Cartesian coordinates $\boldsymbol{r} = (x, y, z)$. A particle that is restricted to move on a two dimensional surface has *two* degrees of freedom. There is then *one* constraint, namely the equation for the surface $f(\boldsymbol{r}) = 0$. Possible coordinates, q_1, q_2, are then the two parameters in a parameter expression

$$\boldsymbol{r} = \boldsymbol{r}(q_1, q_2) = (x(q_1, q_2), y(q_1, q_2), z(q_1, q_2)) \tag{67}$$

for the surface. A particle restricted to move on a *curve*,

$$\boldsymbol{r} = \boldsymbol{r}(q_1) = (x(q_1), y(q_1), z(q_1))\,, \tag{68}$$

has *one* degree of freedom. There are then *two* constraints since any curve can be seen as the line of intersection between two surfaces $f_1(\boldsymbol{r}) = 0, f_2(\boldsymbol{r}) = 0$.

A system of two particles that can move freely has *six* degrees of freedom. Should they both be restricted to move on some surface the system will have only *four* degrees of freedom, two for each particle. If the two particles are connected by a stiff rod of negligible mass one has the constraint

$$|\boldsymbol{r}_1 - \boldsymbol{r}_2| = r_{12} = const. \tag{69}$$

which fixes the distance between the particles. One can also see this as restricting the second particle to move on a spherical surface of radius r_{12} around particle one. This system has *five* degrees of freedom (six for two particles minus one for the constraint). As coordinates one might choose the three Cartesian coordinates of particle one, plus two angles giving the direction to particle two from particle one.

Consider a stiff, light straight rod with N particles fixed along it so that they all are restricted to lie along a straight line with fixed distances. This system also has *five* degrees of freedom since it is still enough to know the position of one particle (3 coordinates) plus the direction of the rod (2 angles) in order to completely fix the positions of all particles. This is clearly independent of the number of particles so we conclude that any matter distribution where the matter can be thought of as being fixed along a *straight line* has *five* degrees of freedom.

Consider again the two particles connected by a stiff rod. Let us now add a third particle by connecting it to the two first with two more light stiff rods so that the system of three particles is restricted to make up a triangle of fixed shape. How many degrees of freedom does this system have? The first two particles need the five coordinates found above, so what we need to know is how many further are required to fix the position of the third particle. The geometric situation is such that the third particle is restricted to move on a circle around an axis defined by the first two. All we need to know is *one* angle to fix its position along this circle. This gives us a total of *six* degrees of freedom. We can arrive at this number as follows: Three particles would have $3 \cdot 3 = 9$ degrees of freedom if they where free to move. The three rods that connect them, however, give rise to three constraints

$$r_{12} = c_1,\ r_{13} = c_2,\ r_{23} = c_3, \tag{70}$$

fixing the inter-particle distances to be constants. The remaining degrees of freedom are now $9 - 3 = 6$.

Imagine now that we add a fourth particle to the rigid triangle we built above in such a way that this fourth particle has fixed distances to the three particles in the triangle. In this way we get a four particle system with particles at the corners of a rigid tetrahedron. Since this means that we get three more constraints

$$r_{14} = c_4,\ r_{24} = c_5,\ r_{34} = c_6, \tag{71}$$

all the three new degrees of freedom of the new particle are gone; the system still has *six* degrees of freedom. If the position of the first three are fixed the position of the fourth will also be so.

If we continue to add particles to the system above in such a way that each new particle has three distances, to already present particles, fixed we thus do not add to the degrees of freedom. One realizes that the result is a system in which all inter-particle distances are constant

$$|\boldsymbol{r}_i - \boldsymbol{r}_j| = r_{ij} = \text{constants}\ , \quad (i, j = 1, \ldots, N). \tag{72}$$

Such a system is said to be *rigid*. In general a rigid system (or body) of more than two particles has *six* degrees of freedom, the exception is one of linear shape which has five as we saw above.

The conclusion that a rigid system has six degrees of freedom is independent of the number of particles as is seen by the following count:

particle number:	1	coordinates:	3	constraints:	0	
particle number:	2	coordinates:	3	constraints:	1	
particle number:	3	coordinates:	3	constraints:	2	
particle number:	4	coordinates:	3	constraints:	3	
particle number:	5	coordinates:	3	constraints:	3	(73)
...	...	...	...	...	...	
particle number:	N	coordinates:	3	constraints:	3	
In total:		coordinates:	3N	constraints:	3N-6	

Independently of the value of N we thus get

$$[3N \text{ coordinates}\,] \ - \ [(3N-6) \text{ constraints}\,] \ = \ 6 \text{ degrees of freedom.} \tag{74}$$

This is still valid in the limit $N \to \infty$ so a continuum $\varrho(\boldsymbol{r})$ can also be thought of as rigid.

Should the motion of a rigid body be constrained to a plane the coordinates needed to specify its position are two coordinates fixing some point of the body in the plane plus one angle fixing the direction of a line in the body with respect to some reference direction in the plane. The planar (or two-dimensional) motion of the rigid body thus has *three* degrees of freedom. The results of the findings about degrees of freedom in this section are summarized in table 1.

In general systems treated in analytical mechanics can be seen as systems of particles and rigid bodies. This means that they can be viewed as K-particle systems with some number, N_c, of constraints $f_1(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_K) = 0$, $f_2(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_K) = 0, \ldots, f_{N_c}(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_K) = 0$. These constraints together define a surface in the

System	*Dimension*	*Translational +*	*Rotational =*	*Degrees of freedom*
Particle	3	3	0	3
N Particles	3	$3N$	0	$3N$
N Particles	2	$2N$	0	$2N$
N Particles	1	N	0	N
Rigid body	3	3	3	6
Rigid body	2	2	1	3
Rigid body	1	1	0	1
Rigid line	3	3	2	5
Rigid line	2	2	1	3

Table 1: Summary of results about the number of degrees of freedom of various systems. Translational degrees of freedom are those that can be described by Cartesian coordinates while rotational are described by angles and relate to directions rather than position.

$3K$-dimensional configuration space of the system on which the system is allowed to move. This surface has dimension $n = 3K - N_c$, the number of degrees of freedom of the system. We will assume that we can parameterize the surface by introducing n *generalized coordinates*, $q_1, q_2, \ldots, q_n$ so that it is given by

$$(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_K) = (\boldsymbol{r}_1(q_1, \ldots, q_n), \boldsymbol{r}_2(q_1, \ldots, q_n), \ldots, \boldsymbol{r}_K(q_1, \ldots, q_n)) \,. \quad (75)$$

This is thus an expression for an n-dimensional surface embedded in a $3K$-dimensional space.

14 Notation and Background

We shall assume that the mechanical system of interest can be obtained by adding constraints to an underlying unconstrained K-particle system. Let the masses of these particles be m_k, their position vectors in some coordinate system $\boldsymbol{r}_k$, and the (total) forces on them $\boldsymbol{F}_k$, where $k = 1, \ldots, K$. To simplify the expressions in the rest of this article we introduce the following notation:

$$\vec{R} \equiv \begin{pmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \vdots \\ \boldsymbol{r}_K \end{pmatrix}, \quad \vec{V} \equiv \begin{pmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \\ \vdots \\ \boldsymbol{v}_K \end{pmatrix}, \quad \vec{P} \equiv \begin{pmatrix} m_1\boldsymbol{v}_1 \\ m_2\boldsymbol{v}_2 \\ \vdots \\ m_K\boldsymbol{v}_K \end{pmatrix} = \begin{pmatrix} \boldsymbol{p}_1 \\ \boldsymbol{p}_2 \\ \vdots \\ \boldsymbol{p}_K \end{pmatrix}, \quad \vec{F} \equiv \begin{pmatrix} \boldsymbol{F}_1 \\ \boldsymbol{F}_2 \\ \vdots \\ \boldsymbol{F}_K \end{pmatrix}. \quad (76)$$

Then, clearly, $\vec{V} = \dot{\vec{R}}$, and with

$$\hat{\mathbf{M}} = \begin{pmatrix} m_1\mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & m_2\mathbf{1} & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & m_K\mathbf{1} \end{pmatrix}, \tag{77}$$

where **1** denotes the 3×3 unit matrix and **0** the 3×3 zero matrix, we can write

$$\vec{P} = \hat{\mathbf{M}}\vec{V}. \tag{78}$$

Using this notation Newton's equations of motion for the system can be written

$$\dot{\vec{P}} = \vec{F}. \tag{79}$$

Following Lesser (1992) we introduce a scalar product on the $3K$-dimensional unconstrained configuration space as the sum of the K ordinary scalar products. We will denote this scalar product by $\bullet$ so that

$$\vec{A} \bullet \vec{B} \equiv \sum_{k=1}^{K} \boldsymbol{a}_k \cdot \boldsymbol{b}_k. \tag{80}$$

The kinetic energy can, for example, be written in the compact form

$$T = \frac{1}{2}\hat{\mathbf{M}}\vec{V} \bullet \vec{V} = \frac{1}{2}\vec{P} \bullet \vec{V} \tag{81}$$

when this scalar product is used.

We now assume that the holonomic constraints of the system are taken into account by giving a parametrization $q = q_1, \ldots, q_n$ of an n-dimensional surface embedded in the unconstrained configuration space as was discussed above equation (75). With the present notation that equation can be written

$$\vec{R} = \vec{R}(q,t) \equiv \begin{pmatrix} \boldsymbol{r}_1(q,t) \\ \vdots \\ \boldsymbol{r}_K(q,t) \end{pmatrix}. \tag{82}$$

For time-dependent constraints one gets one such surface for each value of the time t. If the constraints are time-independent (scleronomic system) the time t will not appear in the parametrization. This surface is the (instantaneous) configuration space of the system. Since a single parametrization cannot, in general, cover the entire configuration space, the theory given here is local.

15 Projection onto Coordinate Tangent Vectors

A constraint represents knowledge about the motion of the system that one has prior to solving the equations of motion. Since one already knows that the motion will take place on the embedded surface (82) implied by the holonomic constraints, there is no need to solve all $3K$ equations of motion (79). It is enough to solve the equations that are the components of the vector equation (79) along the tangent vectors of the surface (82).

One obtains n linearly independent tangent vectors of the (instantaneous) configuration space (surface) by taking the partial derivatives of the position vector (82) with respect to the generalized coordinates

$$\vec{\tau}_a(q,t) \equiv \frac{\partial \vec{R}}{\partial q_a}, \quad a = 1, \ldots, n. \tag{83}$$

The projected equations of motion that will be of interest are thus

$$\dot{\vec{P}}(\dot{q},q,t) \bullet \vec{\tau}_a(q,t) = \vec{F}(\dot{q},q,t) \bullet \vec{\tau}_a(q,t) \qquad a = 1, \ldots, n \tag{84}$$

since these describe the motion along directions allowed by the constraints. The ideas involved here are not really more advanced than those used when Newton's equations are projected onto so called moving basis vectors associated with polar coordinates. Yet these equations are, in fact, the (Euler-) Lagrange equations as we will show below. One notes that, even if there are no constraints, the procedure is useful just for changing coordinates.

16 Lagrange's Equations

We denote the partial time derivative of the position vector by

$$\vec{\tau}_t(q,t) \equiv \frac{\partial \vec{R}}{\partial t}. \tag{85}$$

The total time derivative $\mathrm{d}/\mathrm{d}t$, sometimes denoted simply by an over-dot, should be thought of as the operator

$$\frac{\mathrm{d}}{\mathrm{d}t} \equiv \sum_{a=1}^{n} \left(\ddot{q}_a \frac{\partial}{\partial \dot{q}_a} + \dot{q}_a \frac{\partial}{\partial q_a} \right) + \frac{\partial}{\partial t}, \tag{86}$$

when it acts on a function of the independent variables $\dot{q}, q$, and t. Here $\dot{q}_a \equiv \mathrm{d}q_a/\mathrm{d}t$ are the so called *generalized velocities*. The $3K$-dimensional velocity vector, $\vec{V} = \mathrm{d}\vec{R}/\mathrm{d}t$, is now given by

$$\vec{V}(q,\dot{q},t) = \sum_{a=1}^{n} \dot{q}_a \, \vec{\tau}_a(q,t) + \vec{\tau}_t(q,t). \tag{87}$$

The generalized velocities, which are regarded as independent variables in the Lagrange formalism, can thus be seen as coordinates in velocity space; when they vary the vector $\vec{V}$ sweeps over an n-dimensional hyperplane. For the case of a scleronomic system the vector $\vec{\tau}_t$ is a null vector ($\vec{\tau}_t = \vec{0}$) and this hyperplane can be identified with the tangent plane of the configuration surface (82).

In terms of the independent variables $q, \dot{q}$, and t we can now write the kinetic energy, $T = \frac{1}{2}\hat{\mathbf{M}}\vec{V} \bullet \vec{V}$, as follows

$$T(q, \dot{q}, t) = \sum_{a=1}^{n} \sum_{b=1}^{n} \frac{1}{2} g_{ab}(q, t)\, \dot{q}_a \dot{q}_b + \sum_{a=1}^{n} a_a(q, t)\, \dot{q}_a + b(q, t) \tag{88}$$

where we have introduced

$$g_{ab}(q, t) \equiv \hat{\mathbf{M}}\vec{\tau}_a \bullet \vec{\tau}_b, \tag{89}$$

$$a_a(q, t) \equiv \hat{\mathbf{M}}\vec{\tau}_a \bullet \vec{\tau}_t, \tag{90}$$

$$b(q, t) \equiv \frac{1}{2}\hat{\mathbf{M}}\vec{\tau}_t \bullet \vec{\tau}_t. \tag{91}$$

In the scleronomic case only the first term appears and the kinetic energy is purely quadratic in the generalized velocities. One notes that the definition (89) implies symmetry: $g_{ab} = g_{ba}$.

We now proceed to differentiate the kinetic energy to find Lagrange's equations. We first note that

$$\frac{\partial \vec{V}}{\partial \dot{q}_a} = \vec{\tau}_a = \frac{\partial \vec{R}}{\partial q_a} \tag{92}$$

according to (87) and (83). This is sometimes referred to as the 'cancelation of the dots'. Using this one finds that

$$\frac{\partial T}{\partial \dot{q}_a} = \frac{\partial}{\partial \dot{q}_a} \frac{1}{2} \hat{\mathbf{M}}\vec{V} \bullet \vec{V} = \vec{P} \bullet \frac{\partial \vec{V}}{\partial \dot{q}_a} = \vec{P} \bullet \frac{\partial \vec{R}}{\partial q_a} = \vec{P} \bullet \vec{\tau}_a \equiv p_a. \tag{93}$$

The quantities p_a are the *generalized momenta* and the above equation shows that these are projections of the $3K$-dimensional momentum vector $\vec{P}$ onto the tangent vector $\vec{\tau}_a$. If we now take the total time derivative of this we get

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial T}{\partial \dot{q}_a} = \frac{\mathrm{d}}{\mathrm{d}t} \left(\vec{P} \bullet \vec{\tau}_a \right) = \frac{\mathrm{d}\vec{P}}{\mathrm{d}t} \bullet \vec{\tau}_a + \vec{P} \bullet \frac{\mathrm{d}\vec{\tau}_a}{\mathrm{d}t}. \tag{94}$$

If we differentiate T with respect to q_a we find

$$\frac{\partial T}{\partial q_a} = \frac{\partial}{\partial q_a} \frac{1}{2} \hat{\mathbf{M}}\vec{V} \bullet \vec{V} = \vec{P} \bullet \frac{\partial \vec{V}}{\partial q_a}. \tag{95}$$

We now form the difference

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{q}_a} - \frac{\partial T}{\partial q_a} \;=\; \frac{\mathrm{d}\vec{P}}{\mathrm{d}t} \bullet \vec{\tau}_a + \vec{P} \bullet \left(\frac{\mathrm{d}\vec{\tau}_a}{\mathrm{d}t} - \frac{\partial \vec{V}}{\partial q_a} \right) \tag{96}$$

and note that the vector

$$\left(\frac{\mathrm{d}\vec{\tau}_a}{\mathrm{d}t} - \frac{\partial \vec{V}}{\partial q_a} \right) = \left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial q_a} - \frac{\partial}{\partial q_a}\frac{\mathrm{d}}{\mathrm{d}t} \right) \vec{R} = \vec{0} \tag{97}$$

is the null vector since the differential operators commute (see equation (86)). If one now makes the natural definition that the projection of the $3K$-dimensional force vector $\vec{F}$ onto the tangent vector $\vec{\tau}_a$,

$$\frac{\mathrm{d}\vec{P}}{\mathrm{d}t} \bullet \vec{\tau}_a = \vec{F} \bullet \vec{\tau}_a \;\equiv\; Q_a, \tag{98}$$

is the *generalized force* Q_a one finally ends up with

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{q}_a} - \frac{\partial T}{\partial q_a} = Q_a, \quad a = 1, \ldots, n \tag{99}$$

i.e. the *(Euler-) Lagrange equations* in one of their traditional forms. These are thus simply the projected Newton equations (84).

17 Generalized Forces and the Conservative Case

It should be clear that the definition of generalized force Q_a in Eq. (98),

$$Q_a = \vec{F} \bullet \vec{\tau}_a = \sum_{k=1}^{K} \boldsymbol{F}_k \cdot \frac{\partial \boldsymbol{r}_k}{\partial q_a}, \tag{100}$$

in practice means that a number of forces $\boldsymbol{F}_k$ vanish from the problem. They vanish either because $\vec{F}$ is perpendicular to the tangent vectors, or because the forces in the sum implied by the scalar products occur pairwise withe opposite sign but equal $\partial \boldsymbol{r}_k / \partial q_a$. The former case is related to normal forces, the latter to Newton's third law of action and reaction. We now discuss some properties of generalized forces that do not vanish.

The infinitesimal work, $\mathrm{d}W$, done when a K-particle system is displaced by $\mathrm{d}\vec{R} \equiv (\mathrm{d}\boldsymbol{r}_1, \mathrm{d}\boldsymbol{r}_2, \ldots, \mathrm{d}\boldsymbol{r}_K)^T$ (the superscript T here means matrix transposition so that $\mathrm{d}\vec{R}$ denotes a column matrix as usual) can be written

$$\mathrm{d}W = \vec{F} \bullet \mathrm{d}\vec{R} \equiv \sum_{k=1}^{K} \mathbf{F}_k \cdot \mathrm{d}\boldsymbol{r}_k. \tag{101}$$

There is then a *potential energy*, V, for the system if the (negative) work done when going from some point $\vec{R}_0 = (\boldsymbol{r}_1(0), \boldsymbol{r}_2(0), \ldots, \boldsymbol{r}_K(0))^T$ to some final point $\vec{R}$, i.e.

$$V[\vec{R}] = -\int_{\vec{R}_0}^{\vec{R}} \vec{F} \bullet \mathrm{d}\vec{R} \tag{102}$$

is independent of the path from $\vec{R}_0$ to $\vec{R}$. One can show that this will, in general, be true if the individual forces on the K particles are conservative. One then has that the force on particle k is given by

$$\boldsymbol{F}_k = -\nabla_k V \equiv -\left(\frac{\partial V}{\partial x_k}, \frac{\partial V}{\partial y_k}, \frac{\partial V}{\partial z_k}\right). \tag{103}$$

The $3K$ dimensional force is thus given by

$$\vec{F} = -\vec{\nabla} V \equiv (-\nabla_1 V, -\nabla_2 V, \ldots, -\nabla_K V)^T \tag{104}$$

and the infinitesimal work can be written

$$\mathrm{d}W = -\vec{\nabla} V \bullet \mathrm{d}\vec{R} = -\mathrm{d}V, \tag{105}$$

i.e. as (minus) the total differential of the scalar function V on the $3K$-dimensional space.

Now assume that there are constraints so that the system only can move on the n-dimensional surface $\vec{R} = \vec{R}(q_1, \ldots, q_n)$ in the $3K$-dimensional space. Also assume that these constraints are such that the corresponding constraint forces do not perform work for time independent constraints. If the constraints are time dependent, $\vec{R} = \vec{R}(q_1, \ldots, q_n, t)$, then they should not do work for displacements $\mathrm{d}q_a$ at a fixed time $t =$ const. (so called, virtual displacements). One then speaks about *smooth* constraints. This is only true if the sliding friction forces in the joints and bearings of the system can be neglected. Under these circumstances the work done on the system for some displacement is given by the same expression (102) as above. The only difference is that now the start and end points, as well as the path between them, must lie on the n-dimensional surface, $\vec{R} = \vec{R}(q_1, \ldots, q_n)$, so that they are consistent with the constraints. There is thus still a potential energy which we can write

$$V[\vec{R}] = V[\vec{R}(q_1, \ldots, q_n)] \equiv V(q_1, \ldots, q_n). \tag{106}$$

According to the chain rule we now have

$$-\frac{\partial V}{\partial q_a} = -\sum_{k=1}^{K} \nabla_k V \cdot \frac{\partial \boldsymbol{r}_k}{\partial q_a} = \sum_{k=1}^{K} \mathbf{F}_k \cdot \frac{\partial \boldsymbol{r}_k}{\partial q_a} = \vec{F} \bullet \vec{\tau}_a. \tag{107}$$

Comparing with equation (98) we thus see that

$$-\frac{\partial V}{\partial q_a} = Q_a \tag{108}$$

simply is the generalized force Q_a, in this, conservative, case.

In general the work done during a small displacement is not the total differential of some function. One sometimes stresses this fact by not using the ordinary differential sign (d). Instead one might write

$$\delta W = \vec{F} \bullet \mathrm{d}\vec{R} = \vec{F} \bullet \sum_{a=1}^{n} \frac{\partial \vec{R}}{\partial q_a} \mathrm{d}q_a = \sum_{a=1}^{n} (\vec{F} \bullet \vec{\tau}_a) \mathrm{d}q_a = \sum_{a=1}^{n} Q_a \mathrm{d}q_a. \tag{109}$$

This expression tells us that, whether the forces are conservative or not, the generalized forces can be obtained by writing down the infinitesimal work, δW, done for small changes, $\mathrm{d}q_a$, of the generalized coordinates. The generalized force, Q_a, is then simply the coefficient in front of $\mathrm{d}q_a$ in the expression for δW. Should the working forces all be conservative $\delta W = \mathrm{d}W = -\mathrm{d}V$.

18 The Lagrange Function and Conservative Systems

If we assume that the system is conservative we can now write the Lagrange's equations (99) in the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{q}_a} - \frac{\partial T}{\partial q_a} = -\frac{\partial V}{\partial q_a}, \quad a = 1, \ldots, n \tag{110}$$

We now move $-\frac{\partial V}{\partial q_a}$ to the left hand side of the equation. If we then define the Lagrange function

$$L = L(q, \dot{q}) \equiv T(q, \dot{q}) - V(q), \tag{111}$$

and note that since V does not depend on $\dot{q}_a$, we can rewrite the equations (110) as

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_a} - \frac{\partial L}{\partial q_a} = 0, \quad a = 1, \ldots, n. \tag{112}$$

This is the form that Lagrange's equations take when the system is conservative. The case of velocity dependent forces that was presented for the one particle case in Sec. 10 may also be of interest for systems of particles. Equations of motion on the form (112) are then also valid with $V(q)$ replaced by $U(q, \dot{q})$.

Example 1: Consider a physical pendulum of mass m. Let J be the moment of inertia about the horizontal axis and denote by h the distance from the axis to the

center of mass. If we chose the angle, φ, between the vertical downward direction and the line from the axis to the center of mass, as generalized coordinate, the kinetic energy is $T = \frac{1}{2}J\dot{\varphi}^2$ and the potential energy is $V = mgh(1 - \cos\varphi)$.

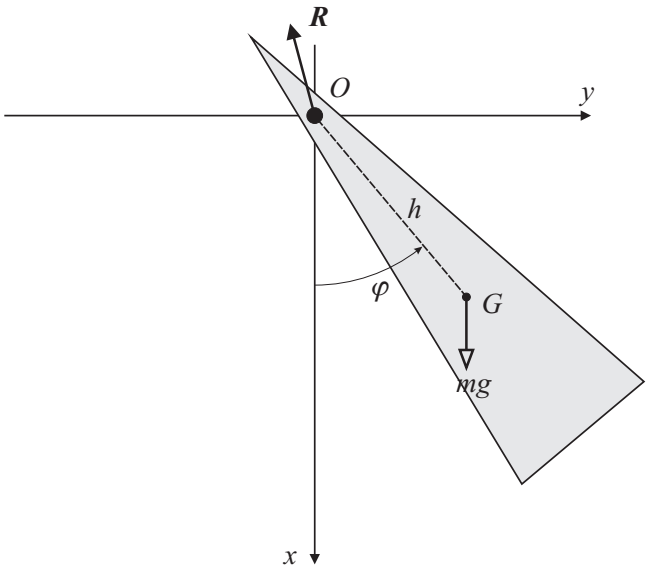

Figure 2: A physical pendulum constructed from a wedge shaped body. It is hinged so that it can rotate about a smooth horizontal axis (the z-axis). The x-axis is in the vertical downward direction. The center of mass is G and its distance from the origin O on the rotation axis is h. The external forces acting are the weight mg and the reaction (constraint) force $\boldsymbol{R}$ at the rotation axis.

From this one finds that $L = T - V = \frac{1}{2}J\dot{\varphi}^2 - mgh(1 - \cos\varphi)$ and thus that

$$\frac{\partial L}{\partial \dot{\varphi}} = J\dot{\varphi}, \quad \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\varphi}} = J\ddot{\varphi}, \quad \text{and} \quad \frac{\partial L}{\partial \varphi} = -mgh\sin\varphi$$

Thus equation (112) in this case becomes

$$J\ddot{\varphi} + mgh\sin\varphi = 0,$$

which is the well known equation of motion for the physical pendulum. **End of example 1**

18.1 The Variational Principle and Euler-Lagrange's equations

This form of the equations can be obtained directly from a variational principle: the principle of least action. The *action* is defined as the time integral of L:

$$S[q(t)] = \int_{t_1}^{t_2} L(q(t), \dot{q}(t))\mathrm{d}t \tag{113}$$

for some path $q(t)$ between two fixed end-points $q_1 = q(t_1)$ and $q_2 = q(t_2)$. Demanding that the change of $\delta S = S[q(t) + \delta q(t)] - S[q(t)] = 0$, for a small change of the real path, $q(t)$, to some nearby path $q(t) + \delta q(t)$, with $\delta q(t_1) = \delta q(t_2) = 0$, see Figure 3, leads directly to the equations (112).

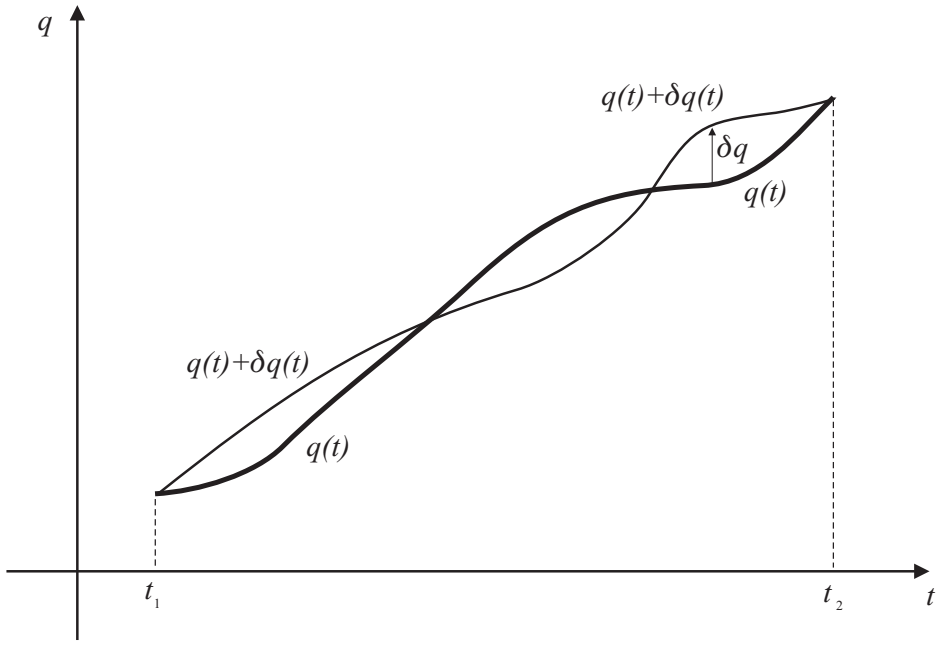

Figure 3: This figure shows a (one-dimensional) trajectory $q(t)$ and an example of a variation of it $q(t) + \delta q(t)$ that obeys $\delta q(t_1) = \delta q(t_2) = 0$.

To prove this we consider

$$\delta S[q(t)] = \int_{t_1}^{t_2} L(q + \delta q, \dot{q} + \delta\dot{q})\mathrm{d}t - \int_{t_1}^{t_2} L(q, \dot{q})\mathrm{d}t = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta\dot{q} \right) \mathrm{d}t. \tag{114}$$

Since $\delta\dot{q} = \dfrac{\mathrm{d}\delta q}{\mathrm{d}t}$ we get, by integrating by parts,

$$\delta S[q(t)] = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} \right) \delta q \mathrm{d}t + \left[\frac{\partial L}{\partial \dot{q}}\delta q \right]_{t_1}^{t_2}. \tag{115}$$

If we now use that $\delta q(t_1) = \delta q(t_2) = 0$ the integrated part vanishes. Since $\delta q(t)$ otherwise is arbitrary the remaining integral is identically zero, and thus $\delta S = 0$, only if the quantity within the parentheses is zero. This gives (112). Note that the calculation was done as if there was only one q but it is easily generalized to the case of several degrees of freedom.

19 Energy and the Hamiltonian formalism

Here we investigate conservation laws and in particular conservation of energy as these concepts appear in the Lagrange formalism. We then briefly present the Hamiltonian formalism.

19.1 Energy in the Lagrange formalism

If we have a conservative system we know that the energy, $E = T + V$, is a conserved quantity. Assume that we have a Lagrangian $L(q, \dot{q}) = T(q, \dot{q}) - V(q)$ for a system that moves according to Lagrange's equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_a} - \frac{\partial L}{\partial q_a} = 0, \quad a = 1, \ldots, n. \tag{116}$$

How can we find a constant of the motion that corresponds to the energy, if we pretend that we do not know it already?

Consider the time derivative of L,

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \sum_{a=1}^{n}\left(\frac{\partial L}{\partial q_a}\frac{\mathrm{d}q_a}{\mathrm{d}t} + \frac{\partial L}{\partial \dot{q}_a}\frac{\mathrm{d}\dot{q}_a}{\mathrm{d}t}\right). \tag{117}$$

According to (116) we have that

$$\frac{\partial L}{\partial q_a} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_a}, \quad a = 1, \ldots, n, \tag{118}$$

when the functions $q_a(t)$ actually obey the equations of motion. If we put this into (117) we get

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \sum_{a=1}^{n}\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_a}\dot{q}_a + \frac{\partial L}{\partial \dot{q}_a}\frac{\mathrm{d}\dot{q}_a}{\mathrm{d}t}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\sum_{a=1}^{n}\left(\frac{\partial L}{\partial \dot{q}_a}\dot{q}_a\right), \tag{119}$$

so the terms of the sum are time derivatives of the products $(\partial L/\partial \dot{q}_a)\dot{q}_a$. If we move $\mathrm{d}L/\mathrm{d}t$ to the right hand side, we now find that

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{a=1}^{n}\frac{\partial L}{\partial \dot{q}_a}\dot{q}_a - L\right), \tag{120}$$

as long as the functions $q_a(t)$ actually obey the equations of motion. Since this time derivative then is zero the quantity inside must be a constant of the motion, i.e. independent of time. When we integrate (120) with respect to time the result is

$$\sum_{a=1}^{n}\frac{\partial L}{\partial \dot{q}_a}\dot{q}_a - L = \mathcal{E}, \tag{121}$$

where $\mathcal{E}$ is the constant of integration. We have thus found that the function

$$\mathcal{E}(q(t), \dot{q}(t)) \equiv \sum_{a=1}^{n} \frac{\partial L(q, \dot{q})}{\partial \dot{q}_a} \dot{q}_a - L(q, \dot{q}), \tag{122}$$

is a constant of the motion. We now show that this is, in fact, the *energy*.

Assuming there are no time dependent constraints the kinetic energy is given by

$$T(q, \dot{q}) = \frac{1}{2} \sum_{bc} g_{bc}(q)\, \dot{q}_b \dot{q}_c. \tag{123}$$

Here $\sum_{bc}$ is shorthand notation for the double sum $\sum_{b=1}^{n} \sum_{c=1}^{n}$. In this case we can thus write the conservative Lagrangian L as follows

$$L(q, \dot{q}) = T(q, \dot{q}) - V(q) = \frac{1}{2} \sum_{bc} g_{bc}(q)\, \dot{q}_b \dot{q}_c - V(q). \tag{124}$$

Now calculate

$$\frac{\partial L}{\partial \dot{q}_a} = \frac{\partial T}{\partial \dot{q}_a} = \frac{\partial}{\partial \dot{q}_a} \left(\frac{1}{2} \sum_{bc} g_{bc}(q)\, \dot{q}_b \dot{q}_c \right). \tag{125}$$

Since

$$\frac{\partial}{\partial \dot{q}_a} \dot{q}_b = \delta_{ab}, \tag{126}$$

i.e., it is zero if $a \neq b$ and one if $a = b$, we find

$$\frac{\partial L}{\partial \dot{q}_a} = \frac{1}{2} \sum_{bc} g_{bc}(q)\ (\delta_{ab} \dot{q}_c + \dot{q}_b \delta_{ac}) = \frac{1}{2} \left(\sum_{bc} g_{bc}(q)\, \delta_{ab} \dot{q}_c + \sum_{bc} g_{bc}(q)\, \dot{q}_b \delta_{ac} \right). \tag{127}$$

In the first sum we perform the b-summation. Only one term survives; the term with $b = a$. In the second we perform the c-summation and here only the term with $c = a$ will be non-zero. We thus have

$$\frac{\partial L}{\partial \dot{q}_a} = \frac{1}{2} \left(\sum_{c} g_{ac}(q)\, \dot{q}_c + \sum_{b} g_{ba}(q)\, \dot{q}_b \right). \tag{128}$$

The mass matrix, defined in Eq. (89), is symmetric, $g_{ba}(q) = g_{ab}(q)$, so we can change the dummy index in the first sum to b instead of c, and get

$$\frac{\partial L}{\partial \dot{q}_a} = \frac{1}{2} \left(\sum_{b} g_{ab}(q)\, \dot{q}_b + \sum_{b} g_{ab}(q)\, \dot{q}_b \right) = \sum_{b} g_{ab}(q)\, \dot{q}_b. \tag{129}$$

If we now calculate

$$\sum_{a=1}^{n} \frac{\partial L}{\partial \dot{q}_a} \dot{q}_a = \sum_{a=1}^{n} \left(\sum_{b=1}^{n} g_{ab}(q)\, \dot{q}_b \right) \dot{q}_a = \sum_{ab} g_{ab}(q)\, \dot{q}_b \dot{q}_a, \tag{130}$$

we find that, according to equation (123),

$$\sum_{a=1}^{n} \frac{\partial L}{\partial \dot{q}_a} \dot{q}_a = 2T. \tag{131}$$

Finally then,

$$\mathcal{E} = \sum_{a=1}^{n} \frac{\partial L}{\partial \dot{q}_a} \dot{q}_a - L = 2T - L = 2T - (T - V) = T + V, \tag{132}$$

so that the constant of motion, $\mathcal{E}$, we found above, is in fact the energy. This is what we wanted to show.

19.2 The Hamiltonian formalism

A conservative system with a Lagrangian function $L(q, \dot{q})$ that does *not* depend on some generalized coordinate q_b is said to possess a *cyclic coordinate*. The equation of motion (116) for this coordinate then becomes

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}_b} = 0. \tag{133}$$

If we note that

$$\frac{\partial L}{\partial \dot{q}_b} \equiv p_b \tag{134}$$

is the *generalized momentum*, see formula (93), corresponding to q_b we find that the solution of equation (133) is simply p_b =constant. Thus cyclic coordinates correspond to generalized momenta that are *constants of the motion*.

We know that a conservative system has another constant of the motion, the energy, $E = T + V$. How does one find this in the Lagrange formalism? One way of doing this is to change the independent coordinates from $q, \dot{q}$ to q, p. This means that one changes from generalized velocities $\dot{q}$ as independent coordinates to generalized momenta p as independent coordinates. One then forms the function

$$H(q, p) = \sum_a p_a \dot{q}_a - L \tag{135}$$

where we assume that equation (134) has been used to express $\dot{q} = \dot{q}(q, p)$ so that everything on the left hand side are functions of q and p. If we now form the total differential of H we find, first by definition,

$$\mathrm{d}H(q, p) = \sum_a \left(\frac{\partial H}{\partial q_a} \mathrm{d}q_a + \frac{\partial H}{\partial p_a} \mathrm{d}p_a \right). \tag{136}$$

On the other hand using (135) we get

$$\mathrm{d}H(q,p) = -\sum_a \left(\frac{\partial L}{\partial q_a}\mathrm{d}q_a + \frac{\partial L}{\partial \dot{q}_a}\mathrm{d}\dot{q}_a - p_a \mathrm{d}\dot{q}_a - \dot{q}_a \mathrm{d}p_a \right). \tag{137}$$

But according to Lagrange's equations (116) we have $\dot{p}_a = \partial L/\partial q_a$ (this is equation (133) for the general case when the coordinate q_a is not cyclic). Using this, and (134) turns our last differential into

$$\mathrm{d}H(q,p) = -\sum_a \left(\dot{p}_a \mathrm{d}q_a + p_a \mathrm{d}\dot{q}_a - p_a \mathrm{d}\dot{q}_a - \dot{q}_a \mathrm{d}p_a \right) = \sum_a \left(\dot{q}_a \mathrm{d}p_a - \dot{p}_a \mathrm{d}q_a \right). \tag{138}$$

Comparing with (136) we can now identify the coefficients in front of the differentials and get

$$\dot{p}_a = -\frac{\partial H}{\partial q_a}, \qquad \dot{q}_a = \frac{\partial H}{\partial p_a}. \tag{139}$$

These are *Hamilton's equations* for the system. Dividing equation (138) by $\mathrm{d}t$ we find

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \sum_a \left(\dot{q}_a \dot{p}_a - \dot{p}_a \dot{q}_a \right) = 0. \tag{140}$$

Thus $H(q,p)$ is a *constant of the motion*. This constant can be identified with *the energy*.

20 Small Oscillations about an Equilibrium Position

Many systems are to a first approximation in a stable static equilibrium configuration. These include man made structures of all kinds as well as molecules. When slightly perturbed these systems execute small oscillations (or vibrations) about the equilibrium position. The nature of these vibrations will be studied here.

We assume that we have a Lagrangian of the form

$$L(q,\dot{q}) = T - V = \frac{1}{2}\sum_{a,b=1}^{n} g_{ab}(q)\dot{q}_a\dot{q}_b - V(q). \tag{141}$$

In order for there to be an equilibrium position we must further assume that there is at least one solution $q_a = q_a^0$ to the equations

$$\frac{\partial V}{\partial q_a} = 0, \qquad a = 1, \ldots, n; \tag{142}$$

this means that the forces all are zero at that position. This solution will be a stable position provided that it corresponds to a a local minimum of potential energy V.

It is only then that $V(q^0) \leq V(q)$, for all q sufficiently close to q^0 so that the forces tend to carry the system back to q^0 when the deviation from this position is sufficiently small.

If we now make a Taylor expansion of $V(q)$ around q^0 we get

$$V(q) = V(q^0) + \frac{1}{2}\sum_{a,b=1}^{n}\left(\frac{\partial^2 V}{\partial q_a \partial q_b}\right)_{q=q^0}(q_a - q_a^0)(q_b - q_b^0) + \dots . \tag{143}$$

The linear terms are zero because q^0 is a solution of (142). Since we will be considering small oscillations about q^0 we assume that cubic and higher order terms are negligible. If we now define

$$\left(\frac{\partial^2 V}{\partial q_a \partial q_b}\right)_{q=q^0} \equiv K_{ab}, \qquad u_a \equiv q_a - q_a^0 \tag{144}$$

and discard the irrelevant constant $V(q^0)$ we see that we can assume the potential energy to be of the form

$$V(u) = \frac{1}{2}\sum_{a,b=1}^{n} K_{ab} u_a u_b, \tag{145}$$

with constant coefficients K_{ab}.

Example 2: Consider the potential energy

$$V(x,y) = 4x^2 + 2xy + y^2 + 12x + 9 \tag{146}$$

We first calculate the minimum values of x and y using equations (142). We find

$$\frac{\partial V}{\partial x} = 8x + 2y + 12 = 0 \tag{147}$$

$$\frac{\partial V}{\partial y} = 2x + 2y = 0 \tag{148}$$

Solving this system we find the solution $x_0 = -2,\ y_0 = 2$. At this point the forces are zero on a particle with this potential energy. Since the potential is quadratic with positive coefficients the point should be a minimum. We now calculate the expansion (143). This gives

$$V(x,y) = V(x_0,y_0) + \frac{1}{2}\left[\left(\frac{\partial^2 V}{\partial x^2}\right)_{x_0,y_0}(x-x_0)^2 + 2\left(\frac{\partial^2 V}{\partial x \partial y}\right)_{x_0,y_0}(x-x_0)(y-y_0) + \left(\frac{\partial^2 V}{\partial y^2}\right)_{x_0,y_0}(y-y_0)^2\right] \tag{149}$$

$$= -3 + \frac{1}{2}[8(x+2)^2 + 2\cdot 2(x+2)(y-2) + 2(y-2)^2].$$

This expression is the same as (146) since the Taylor expansion of a second order polynomial to second order is exact; there are no neglected terms. The constants are such that $V(x_0, y_0) = V(-2, 2) = -3$, so this is the irrelevant constant $V(q^0)$ in (143).

We now skip the constant term -3 and introduce $u_1 = x + 2,\ u_2 = y - 2$ according to equation (144). This finally gives us

$$V(u) = \frac{1}{2}\sum_{a,b=1}^{n} K_{ab}u_a u_b = \frac{1}{2}(8u_1^2+4u_1u_2+2u_2^2) = \frac{1}{2}(8u_1^2+2u_1u_2+2u_2u_1+2u_2^2), \tag{150}$$

for the expression in equation (145). The matrix K_{ab} is then seen to have the elements $K_{11} = 8,\ K_{22} = 2,\ K_{12} = K_{21} = 2$. **End of example 2**

Example 3: Consider a coplanar double pendulum, see Figure 4, consisting of a particle of mass m_1 suspended in a string of length l_1. From m_1 a second string of length l_2 is suspended with a particle of mass m_2 at the other end. If we denote the angle between the first string and the vertical by φ_1 and that between the second and the vertical by φ_2 we find that the potential energy can be written

$$V(\varphi_1, \varphi_2) = m_1 g l_1(1 - \cos\varphi_1) + m_2 g[l_1(1 - \cos\varphi_1) + l_2(1 - \cos\varphi_2)]. \tag{151}$$

Clearly the values $\varphi_{10} = 0,\ \varphi_{20} = 0$ correspond to the minimum. Here the constants are chosen so that $V(0, 0) = 0$ so that the irrelevant constant $V(q^0)$ in (143) is zero from the beginning. We now make Taylor expansions of the cosines ($\cos x = 1 - x^2/2 + \ldots$) and get

$$V = \frac{1}{2}m_1 g l_1\varphi_1^2 + \frac{1}{2}m_2 g[l_1\varphi_1^2 + l_2\varphi_2^2] + \ldots = \frac{1}{2}g[(m_1+m_2)l_1\varphi_1^2 + m_2 l_2\varphi_2^2] + \ldots. \tag{152}$$

We also see that the minimum already corresponds to the values $\varphi_1 = \varphi_2 = 0$ so these can be taken as the coordinates $u_a,\ a = 1, 2$ in equation (144). The matrix K_{ab} has the elements $K_{11} = g(m_1 + m_2)l_1,\ K_{22} = gm_2 l_2,\ K_{12} = K_{21} = 0$. **End of example 3**

Let us return to the Lagrangian (141). We now assume that (145) is a good approximation for the potential energy. This is true when deviations from the equilibrium positions are small. But small deviations will normally mean small forces and thus small velocities. In what follows we also assume that the velocities can be assumed small. Note that $\dot{q}_a = \dot{u}_a$. If we now make a Taylor expansion of $g_{ab}(q)$ near q^0,

$$g_{ab}(q) = g_{ab}(q^0) + \sum_{c=1}^{n}\left(\frac{\partial g_{ab}}{\partial q_c}\right)_{q=q^0}(q_c - q_c^0) + \ldots, \tag{153}$$

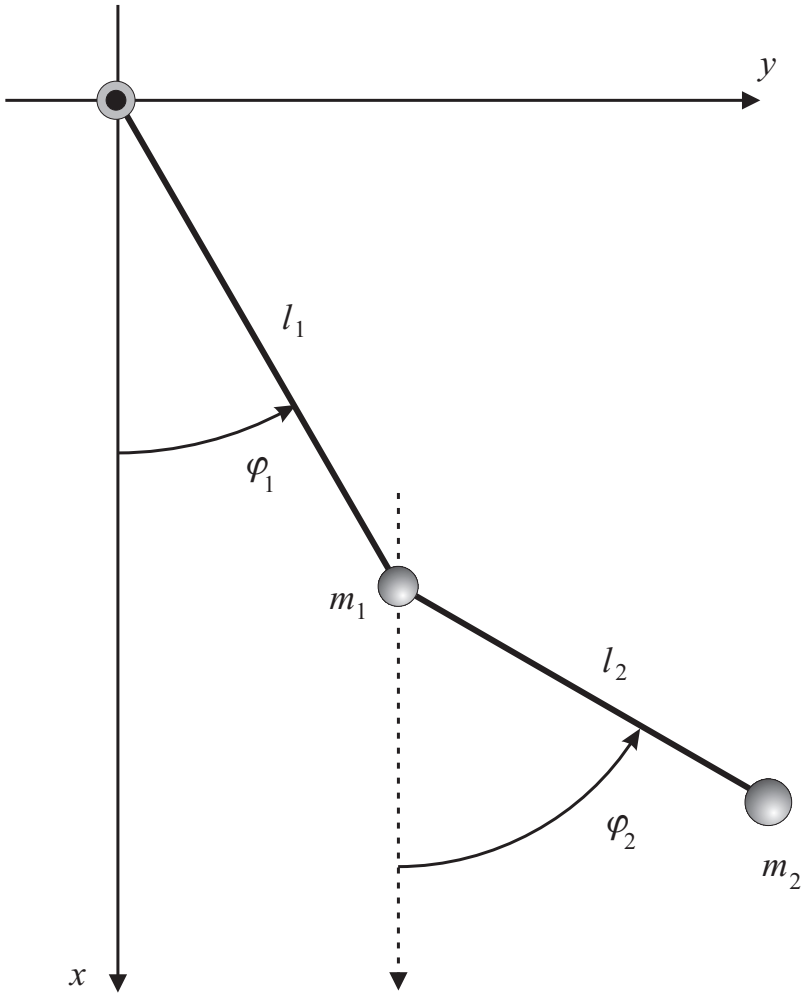

Figure 4: This figure shows the double pendulum treated in Examples 3, 4, and 5.

we see that the linear terms in the expansion of g_{ab} will contribute terms of type $u_c\dot{u}_a\dot{u}_b$ i.e. cubic terms in positions and velocities, to the kinetic energy T. Higher order terms in the expansion contribute terms of even higher order. The assumption of small displacements and velocities thus makes it reasonable to keep only the constant terms in the expansion of g_{ab}. If we now introduce the notation

$$g_{ab}(q^0) \equiv M_{ab}, \tag{154}$$

we can write the Lagrangian for a system with small oscillations about an equilibrium on the form

$$L(u,\dot{u}) = T(\dot{u}) - V(u) = \frac{1}{2}\sum_{a,b=1}^{n} M_{ab}\dot{u}_a\dot{u}_b - \frac{1}{2}\sum_{a,b=1}^{n} K_{ab}u_a u_b. \tag{155}$$

Here the coefficients M_{ab} and K_{ab} are constants.

Example 4: Consider again the coplanar double pendulum of Examples 2 and 3 above, see Figure 4. A particle of mass m_1 suspended in a string of length l_1. From m_1 a second string of length l_2 is suspended with a particle of mass m_2 at the other end. Denote the angle between the first string and the vertical by φ_1 and that between the second and the vertical by φ_2 and find the kinetic energy of the system!

Introduce Cartesian axes with the x-axis vertically down, the y-axis horizontal in the plane of the pendulum and the origin at the point of suspension. The Cartesian coordinates of the two particles are given by

$$x_1 = l_1 \cos\varphi_1, \tag{156}$$
$$y_1 = l_1 \sin\varphi_1, \tag{157}$$
$$x_2 = l_1 \cos\varphi_1 + l_2 \cos\varphi_2, \tag{158}$$
$$y_2 = l_1 \sin\varphi_1 + l_2 \sin\varphi_2. \tag{159}$$

The velocities are then,

$$\dot{x}_1 = -l_1\dot{\varphi}_1 \sin\varphi_1, \tag{160}$$
$$\dot{y}_1 = l_1\dot{\varphi}_1 \cos\varphi_1, \tag{161}$$
$$\dot{x}_2 = -l_1\dot{\varphi}_1 \sin\varphi_1 - l_2\dot{\varphi}_2 \sin\varphi_2, \tag{162}$$
$$\dot{y}_2 = l_1\dot{\varphi}_1 \cos\varphi_1 + l_2\dot{\varphi}_2 \cos\varphi_2. \tag{163}$$

The kinetic energy is, by definition, given by

$$T = T_1 + T_2 = \frac{1}{2}m_1(\dot{x}_1^2 + \dot{y}_1^2) + \frac{1}{2}m_2(\dot{x}_2^2 + \dot{y}_2^2). \tag{164}$$

Inserting the above expressions into the definition gives after patient calculations:

$$T = T_1 + T_2 = \frac{1}{2}m_1 l_1^2\dot{\varphi}_1^2 + \frac{1}{2}m_2[l_1^2\dot{\varphi}_1^2 + l_2^2\dot{\varphi}_2^2 + 2l_1 l_2 \cos(\varphi_1 - \varphi_2)\dot{\varphi}_1\dot{\varphi}_2] \tag{165}$$
$$\Rightarrow\ T = \frac{1}{2}(m_1 + m_2)l_1^2\dot{\varphi}_1^2 + \frac{1}{2}m_2 l_2^2\dot{\varphi}_2^2 + m_2 l_1 l_2 \cos(\varphi_1 - \varphi_2)\dot{\varphi}_1\dot{\varphi}_2. \tag{166}$$

The elements of the matrix $g_{ab}(q) = g_{ab}(\varphi_1, \varphi_2)$ of equation (141) are thus $g_{11} = (m_1+m_2)l_1^2,\ g_{22} = m_2 l_2^2,\ g_{12} = g_{21} = m_2 l_1 l_2 \cos(\varphi_1 - \varphi_2)$. Since the equilibrium position is $\varphi_1 = \varphi_2 = 0$ we see that the mass matrix M_{ab} has the elements $M_{11} = (m_1 + m_2)l_1^2,\ M_{22} = m_2 l_2^2,\ M_{12} = M_{21} = m_2 l_1 l_2$. **End of example 4**

For a single degree of freedom ($n = 1$) we get $L = (1/2)(M_{11}\dot{u}_1^2 - K_{11}u_1^2)$ and the equation of motion is $M_{11}\ddot{u}_1 = -K_{11}u_1$. This is the equation of motion for

a mass M_{11} kept at $u_1 = 0$ by a spring with stiffness (spring constant) K_{11}. This explains the notation. (Should u_1 be an angle then mass is replaced by moment of inertia and stiffness by torsional stiffness.) Collectively one speaks of the M_{ab} as the mass matrix and the K_{ab} as the stiffness matrix. The definitions of these matrices imply that they are both symmetric, $M_{ab} = M_{ba}$ because of (154) and (89), and $K_{ab} = K_{ba}$ because of the definition as the mixed partial derivatives of the potential in Eq. (144).

We now introduce matrix notation. We put

$$\mathbf{u} \equiv \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, \mathbf{M} \equiv \begin{pmatrix} M_{11} & M_{12} & \dots & M_{1n} \\ M_{21} & M_{22} & \dots & M_{2n} \\ \vdots & \vdots & & \vdots \\ M_{n1} & M_{n2} & \dots & M_{nn} \end{pmatrix}, \mathbf{K} \equiv \begin{pmatrix} K_{11} & K_{12} & \dots & K_{1n} \\ K_{21} & K_{22} & \dots & K_{2n} \\ \vdots & \vdots & & \vdots \\ K_{n1} & K_{n2} & \dots & K_{nn} \end{pmatrix}. \tag{167}$$

We denote the transpose with a T-superscript so that $\mathbf{u}^T \equiv (u_1\ u_2\ \dots\ u_n)$. $\mathbf{M}$ is then symmetric if and only if $\mathbf{M} = \mathbf{M}^T$. With this notation the Lagrangian (155) can be written

$$L(u, \dot{u}) = \frac{1}{2}\left(\dot{\mathbf{u}}^T \mathbf{M} \dot{\mathbf{u}} - \mathbf{u}^T \mathbf{K} \mathbf{u}\right). \tag{168}$$

The equations of motion, $(\mathrm{d}/\mathrm{d}t)(\partial L/\partial \dot{u}_a) - (\partial L/\partial u_a) = 0$, corresponding to (155) are

$$\sum_{b=1}^{n} (M_{ab}\ddot{u}_b + K_{ab}u_b) = 0, \quad a = 1, \dots, n. \tag{169}$$

Using the matrix notation we can write them as a column matrix of equations

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{0}, \tag{170}$$

if $\mathbf{0}$ denotes a column of n zeroes.

We now discuss how to solve these equations. We do this by assuming that a solution can be written on the form

$$\mathbf{u}(t) = \mathbf{a}\cos(\omega t + \phi) \tag{171}$$

where $\mathbf{a}^T = (a_1 \dots a_n)$ is a matrix of constants. We here take for granted that the solutions must be oscillating. Without knowledge of the nature of the stiffness matrix $\mathbf{K}$ an ansatz with $\mathbf{a}\exp(\lambda t)$ could be used. Imaginary λ would then correspond to oscillations, but real to unstable motion. Putting (171) into (170) we find

$$\left(-\mathbf{M}\omega^2 + \mathbf{K}\right)\mathbf{a} = \mathbf{0}. \tag{172}$$

This homogeneous system of linear equations will have non-trivial solutions **a** only if the determinant of the coefficient matrix is zero:

$$\det\Big(-\mathbf{M}\omega^2+\mathbf{K}\Big)=\begin{vmatrix} M_{11}\omega^2-K_{11} & M_{12}\omega^2-K_{12} & \dots & M_{1n}\omega^2-K_{1n} \\ M_{21}\omega^2-K_{21} & M_{22}\omega^2-K_{22} & \dots & M_{2n}\omega^2-K_{2n} \\ \vdots & \vdots & & \vdots \\ M_{n1}\omega^2-K_{n1} & M_{n2}\omega^2-K_{n2} & \dots & M_{nn}\omega^2-K_{nn} \end{vmatrix}=0. \tag{173}$$

This equation, which sometimes is called the *secular equation*, for historical reasons, will in general have n roots $\omega_i^2,\ (i=1,\dots,n)$.

Example 5: The double pendulum of examples 3 and 4 gave us

$$\mathbf{K}=\begin{pmatrix} g(m_1+m_2)l_1 & 0 \\ 0 & gm_2l_2 \end{pmatrix},\quad \mathbf{M}=\begin{pmatrix} (m_1+m_2)l_1^2 & m_2l_1l_2 \\ m_2l_1l_2 & m_2l_2^2 \end{pmatrix}$$

so equation (173) becomes

$$\begin{vmatrix} (m_1+m_2)l_1^2\omega^2-g(m_1+m_2)l_1 & m_2l_1l_2\omega^2 \\ m_2l_1l_2\omega^2 & m_2l_2^2\omega^2-gm_2l_2 \end{vmatrix}=0,$$

If we put $l_1=l$, $l_2=\alpha l$, $m_1=m$, and $m_2=\beta m$, we find the roots

$$\omega_{1,2}^2=\frac{g}{l}\frac{(1+\alpha)(1+\beta)}{2\alpha}\left(1\pm\sqrt{1-\frac{4\alpha}{(1+\alpha)^2(1+\beta)}}\right)$$

and the square roots of these values are thus the two angular frequencies of the double pendulum. **End of example 5**

Each root ω_i^2 can then be inserted back into equation (172) which then becomes

$$\Big(-\mathbf{M}\omega_i^2+\mathbf{K}\Big)\,\mathbf{a}_i=\mathbf{0}. \tag{174}$$

The determinant is now zero so there are solutions and, if the ω_i are different, one can show that the components of $\mathbf{a}_i$ can be taken as n minors (cofactors) of any row of the determinant $\det(-\mathbf{M}\omega_i^2+\mathbf{K})$. If a solution $\mathbf{a}_i$ has been found it is easily seen that also $\lambda_i\mathbf{a}_i$ are solutions for arbitrary non-zero constants λ_i. This arbitrariness can be removed for example by arbitrarily choosing some value for one of the non-zero components of $\mathbf{a}_i$. The result of the ansatz (171) is thus n different solutions

$$\mathbf{u}_i(t)=\mathbf{a}_i\cos(\omega_i t+\phi_i),\quad i=1,\dots,n. \tag{175}$$

The *general solution* to the linear problem (170) is then an arbitrary linear combination of these:

$$\mathbf{u}(t) = \sum_{i=1}^{n} c_i \mathbf{a}_i \cos(\omega_i t + \phi_i). \tag{176}$$

The $2n$ constants c_i and ϕ_i must be determined by the initial conditions.

20.1 Normal Modes

M is a real symmetric matrix[1]. It can therefore be diagonalized by an orthonormal matrix (transformation) **U** (obeying $\mathbf{U}^{-1} = \tilde{\mathbf{U}}$, i.e. the inverse is equal to the transpose). This means that $\mathbf{M} = \mathbf{U}\mathbf{D}\tilde{\mathbf{U}}$ where **D** is a diagonal matrix. If we define $\mathbf{D}^{1/2}$ to be the diagonal matrix with the square roots of the elements of **D** along the diagonal we have that $\mathbf{D} = \mathbf{D}^{1/2}\mathbf{D}^{1/2}$ and thus we get

$$\mathbf{M} = \mathbf{U}\mathbf{D}^{1/2}\mathbf{D}^{1/2}\tilde{\mathbf{U}} = \mathbf{U}\mathbf{D}^{1/2}\mathbf{1}\mathbf{D}^{1/2}\tilde{\mathbf{U}} = \mathbf{U}\mathbf{D}^{1/2}\tilde{\mathbf{U}}\mathbf{U}\mathbf{D}^{1/2}\tilde{\mathbf{U}} \equiv \mathbf{M}^{1/2}\mathbf{M}^{1/2}. \tag{177}$$

Here we have used that $\tilde{\mathbf{U}}\mathbf{U} = \mathbf{1}$ is the unit matrix. We have thus defined the symmetric matrix $\mathbf{M}^{1/2} \equiv \mathbf{U}\mathbf{D}^{1/2}\tilde{\mathbf{U}}$. Since **M** is positive definite, so is $\mathbf{M}^{1/2}$ and the inverse $\mathbf{M}^{-1/2}$ therefore exists.

Consider the matrix equation (170). If we multiply it to the left by $\mathbf{M}^{-1/2}$ and insert the unit matrix $\mathbf{1} = \mathbf{M}^{-1/2}\mathbf{M}^{1/2}$ to the right of **M** and **K** we get

$$\mathbf{M}^{-1/2}\mathbf{M}\mathbf{M}^{-1/2}\mathbf{M}^{1/2}\ddot{\mathbf{u}} + \mathbf{M}^{-1/2}\mathbf{K}\mathbf{M}^{-1/2}\mathbf{M}^{1/2}\mathbf{u} = \mathbf{M}^{-1/2}\mathbf{0}. \tag{178}$$

If we now define $\mathbf{w} \equiv \mathbf{M}^{1/2}\mathbf{u}$ and $\mathbf{H} \equiv \mathbf{M}^{-1/2}\mathbf{K}\mathbf{M}^{-1/2}$ we get

$$\ddot{\mathbf{w}} + \mathbf{H}\mathbf{w} = \mathbf{0}. \tag{179}$$

Here **H** is a real symmetric (constant) matrix. The equation has the formal solution

$$\mathbf{w}(t) = \exp(\mathrm{i}\mathbf{H}^{1/2}t)\mathbf{w}(0), \tag{180}$$

where the exponential of the matrix is defined through the power series expansion.

Since **H** is a real symmetric matrix it can (just as **M** above) be diagonalized by an orthonormal transformation **N** such that $\mathbf{\Omega}^2 \equiv \mathbf{N}\mathbf{H}\tilde{\mathbf{N}}$ is a diagonal matrix. We now multiply (179) to the left by **N** and insert the unit matrix $\mathbf{1} = \tilde{\mathbf{N}}\mathbf{N}$ between **H** and **w**. We get

$$\mathbf{N}\ddot{\mathbf{w}} + \mathbf{N}\mathbf{H}\tilde{\mathbf{N}}\mathbf{N}\mathbf{w} = \mathbf{N}\mathbf{0}. \tag{181}$$

[1] In this Section one must assume that the mass and stiffness matrices are such that all elements have the same physical dimension. This is not always the case, but it is always possible to chose to work with dimensionless quantities.

If we now put $\mathbf{Q} \equiv \mathbf{N}\mathbf{w}$ we find the simple form

$$\ddot{\mathbf{Q}} + \mathbf{\Omega}^2\mathbf{Q} = \mathbf{0} \tag{182}$$

for the equations of motion. Note that $\mathbf{Q} = \mathbf{N}\mathbf{M}^{1/2}\mathbf{u}$.

Since $\mathbf{\Omega}^2$ is diagonal this last matrix equation of motion is simply a set of equations for the n components Q_i of $\mathbf{Q}$,

$$\ddot{Q}_i + \omega_i^2 Q_i = 0, \quad i = 1, \ldots, n. \tag{183}$$

The coordinates Q_i are called *normal coordinates*. The motion of the system when only one normal coordinate, Q_i say, has time dependence is called the ith *normal mode*. Finally it is easy to see that equations of motion of the form (183) must arise from a Lagrangian

$$L(Q, \dot{Q}) = \frac{1}{2}(\dot{\mathbf{Q}}^T\dot{\mathbf{Q}} - \mathbf{Q}^T\mathbf{\Omega}^2\mathbf{Q}) = \sum_{i=1}^{n} \frac{1}{2}(\dot{Q}_i^2 - \omega_i^2 Q_i^2). \tag{184}$$

Note that first a rescaling has made the mass matrix a unit matrix, and then the resulting stiffness matrix has been diagonalized.

21 Impact Problems

A force is said to be impulsive if it large but acts for a very short time interval τ (between t_i and $t_f = t_i + \tau$) in such a way that the change in position, of the system, during this interval is negligible, $q_a(t_i) \approx q_a(t_f) = q_a(t_i + \tau)$, but the accelerations are so large that the velocities have changed significantly $\dot{q}_a(t_i) \neq \dot{q}_a(t_f)$. The approximation that $q_a(t_i) = q_a(t_f)$, or, equivalently $\tau \to 0$, is called the *impact* approximation.

21.1 Impact Problems with Vector Methods

Assume that the impulsive forces $\boldsymbol{F}_j^{\mathrm{i}}$ act on a rigid body at points $\boldsymbol{r}_j$ relative to its center of mass, G. Let $\boldsymbol{p}$ be the momentum of the body ($\boldsymbol{p} = m\boldsymbol{v}_G$). Then

$$\dot{\boldsymbol{p}} = \boldsymbol{F}^{\mathrm{e}} + \sum_j \boldsymbol{F}_j^{\mathrm{i}}, \tag{185}$$

where $\boldsymbol{F}^{\mathrm{e}}$ is the total ordinary (non-impulsive) external force. If we integrate this with respect to time between t_i and $t_f = t_i + \tau$, we find

$$\boldsymbol{p}(t_f) - \boldsymbol{p}(t_i) = \int_{t_i}^{t_f} \boldsymbol{F}^{\mathrm{e}}(t)\,\mathrm{d}t + \sum_j \int_{t_i}^{t_f} \boldsymbol{F}_j^{\mathrm{i}}(t)\,\mathrm{d}t. \tag{186}$$

The first integral gives

$$\int_{t_i}^{t_i+\tau} \boldsymbol{F}^{\mathrm{e}}(t)\,\mathrm{d}t \approx \boldsymbol{F}^{\mathrm{e}}(t_i)\tau \tag{187}$$

and thus goes to zero when $\tau \to 0$, since one can assume that the ordinary forces vary slowly with time. The impact forces on the other hand are large during the impact interval and give finite integrals

$$\int_{t_i}^{t_i+\tau} \boldsymbol{F}^{\mathrm{i}}_j(t)\,\mathrm{d}t \equiv \boldsymbol{I}_j. \tag{188}$$

In the limit $\tau \to 0$ we must assume that $\boldsymbol{F}^{\mathrm{i}}_j(t_i + \frac{1}{2}\tau) \to \infty$ so that the integral remains non-zero. The vectors $\boldsymbol{I}_j$ are the *impulses* of the impact forces $\boldsymbol{F}^{\mathrm{i}}_j$. In summary we find that

$$\boldsymbol{p}(t_f) - \boldsymbol{p}(t_i) = m[\boldsymbol{v}_G(t_f) - \boldsymbol{v}_G(t_i)] = \sum_j \boldsymbol{I}_j, \tag{189}$$

and the total impulse, $\boldsymbol{I} = \sum_j \boldsymbol{I}_j$, divided by the total mass, m,

$$\boldsymbol{v}_G(t_f) - \boldsymbol{v}_G(t_i) = \boldsymbol{I}/m, \tag{190}$$

gives the change in center of mass velocity.

Now consider the law of angular momentum

$$\dot{\boldsymbol{L}} = \boldsymbol{M}^{\mathrm{e}} + \sum_j \boldsymbol{r}_j \times \boldsymbol{F}^{\mathrm{i}}_j \tag{191}$$

where the center of mass of the body is base point. Time integration of this gives

$$\boldsymbol{L}(t_f) - \boldsymbol{L}(t_i) = \sum_j \boldsymbol{r}_j \times \boldsymbol{I}_j, \tag{192}$$

since, again we can assume that $\boldsymbol{M}^{\mathrm{e}}(t)$ varies slowly with time, and that the positions, $\boldsymbol{r}_j$, change negligibly during impact. We call the vectors,

$$\mathbf{H}_j \equiv \boldsymbol{r}_j \times \boldsymbol{I}_j, \tag{193}$$

the *angular impulses* and $\mathbf{H} = \sum_j \mathbf{H}_j$ the total angular impulse, so that

$$\boldsymbol{L}(t_f) - \boldsymbol{L}(t_i) = \mathbf{H}, \tag{194}$$

Since $\boldsymbol{L} = \hat{\mathbf{J}}\boldsymbol{\omega}$, where $\hat{\mathbf{J}}$ is the inertia tensor, we find that

$$\boldsymbol{\omega}(t_f) - \boldsymbol{\omega}(t_i) = \hat{\mathbf{J}}^{-1}\mathbf{H}, \tag{195}$$

gives the change of angular velocity, $\boldsymbol{\omega}$, caused by the impact.

21.2 Impact Problems with Lagrange's Method

Recall the definition of generalized force, $\vec{F} \bullet \vec{\tau}_a \equiv Q_a$ of equation (98). If the forces of $\vec{F}$ can be divided into ordinary $\vec{F}^{\mathrm{e}}$ and impulsive $\vec{F}^{\mathrm{i}}$, so that $\vec{F} = \vec{F}^{\mathrm{e}} + \vec{F}^{\mathrm{i}}$, we find that so can the generalized forces

$$Q_a = Q_a^{\mathrm{e}} + Q_a^{\mathrm{i}}. \tag{196}$$

If the Lagrange equations (99) are now integrated with respect to time one gets

$$\int_{t_i}^{t_i+\tau} \left(\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial T}{\partial \dot{q}_a} \right) \mathrm{d}t - \int_{t_i}^{t_i+\tau} \left(\frac{\partial T}{\partial q_a} \right) \mathrm{d}t = \int_{t_i}^{t_i+\tau} Q_a^{\mathrm{e}} \, \mathrm{d}t + \int_{t_i}^{t_i+\tau} Q_a^{\mathrm{i}} \, \mathrm{d}t. \tag{197}$$

Using the definition (134) of generalized momentum, p_a, the first integral on the left hand side gives

$$\int_{t_i}^{t_i+\tau} \left(\frac{\mathrm{d}p_a}{\mathrm{d}t} \right) \mathrm{d}t = p_a(t_f) - p_a(t_i). \tag{198}$$

The integrands of the second and third integrals of (197) are finite during impact so in the limit $\tau \to 0$ they give zero. The generalized impact forces, on the other hand, must be assumed to become infinite during impact and thus give finite non-zero results,

$$\int_{t_i}^{t_i+\tau} Q_a^{\mathrm{i}} \, \mathrm{d}t \equiv I_a \tag{199}$$

which we call *generalized impulses*. We thus find that

$$p_a(t_f) - p_a(t_i) = I_a, \quad a = 1, \ldots, n \tag{200}$$

is the impact version of the Lagrange equations.

How does one find the generalized impulses? Assume as above that (external) impulsive forces $\boldsymbol{F}_j^{\mathrm{i}}$ act on the system at points $\boldsymbol{r}_j$. We then get that

$$Q_a^{\mathrm{i}} = \vec{F}^{\mathrm{i}} \bullet \vec{\tau}_a = \sum_j \boldsymbol{F}_j^{\mathrm{i}} \cdot \frac{\partial \boldsymbol{r}_j(q)}{\partial q_a}. \tag{201}$$

Time integration of this over the duration of the impact gives

$$I_a = \int_{t_i}^{t_f} \left(\sum_j \boldsymbol{F}_j^{\mathrm{i}} \cdot \frac{\partial \boldsymbol{r}_j(q)}{\partial q_a} \right) \mathrm{d}t. \tag{202}$$

If we now use the fact that the positions $\boldsymbol{r}_j$, and their derivatives with respect to q_a, do not change during impact, the definition (188) gives us

$$I_a = \sum_j \boldsymbol{I}_j \cdot \frac{\partial \boldsymbol{r}_j(q)}{\partial q_a}, \tag{203}$$

for the generalized impulses, I_a, in terms of the impulse vectors, $\boldsymbol{I}_j$.

22 Passing to the Continuum

Consider a collection of N particles that can move along a line and which are connected with identical springs of stiffness (force constant) k. Let u_b be the displacement of the bth particle from its equilibrium position. If the particles have mass m the Lagrangian will be

$$L(u, \dot{u}) = \frac{1}{2} \sum_{b=1}^{N} [m\dot{u}_b^2 - k(u_{b+1} - u_b)^2]. \tag{204}$$

Let a be the separation between equilibrium positions. We can then write

$$L(u, \dot{u}) = \sum_{b=1}^{N} a \frac{1}{2} \left[\frac{m}{a} \dot{u}_b^2 - ka \left(\frac{u_{b+1} - u_b}{a} \right)^2 \right] \equiv \sum_{b=1}^{N} a\mathcal{L}_b. \tag{205}$$

Here $\mathcal{L}_b$ is the linear Lagrangian density, i.e. the Lagrangian per unit length.

To pass to a continuous mechanical system with an infinity of degrees of freedom we consider the limit when a goes to zero. We thus make the following replacements

$$a \to \mathrm{d}x, \quad \frac{m}{a} \to \lambda = \text{ linear mass density}, \tag{206}$$

$$\frac{u_{b+1} - u_b}{a} \to \frac{\partial u}{\partial x}, \quad ka \to Y = \text{Young's modulus}. \tag{207}$$

This gives us

$$L = \int \mathcal{L} \, \mathrm{d}x, \tag{208}$$

where

$$\mathcal{L} = \frac{1}{2} \left[\lambda \dot{u}^2 - Y \left(\frac{\partial u}{\partial x} \right)^2 \right]. \tag{209}$$

We note that the generalized coordinates u_b have become a function (or field) of the continuous parameters x and t. In the continuum Lagrangian formalism we still treat $u(x, t)$ as a generalized 'coordinate'.

We now consider the variational principle for this Lagrangian, as we did for the discrete case in subsection 18.1. We thus consider

$$\delta \int_{t_1}^{t_2} \mathcal{L} \, \mathrm{d}t = \delta \int_{t_1}^{t_2} \mathrm{d}t \int \mathrm{d}x \, \mathcal{L} \left(u, \dot{u}, \frac{\partial u}{\partial x} \right). \tag{210}$$

The variation δu of $u(x, t)$ is assumed to vanish at t_1 and t_2 and also at the boundaries of the x-integration. We now do the variation and use the same trick as in

subsection 18.1. This gives

$$\delta \int_{t_1}^{t_2} \mathcal{L} \mathrm{d}t = \int_{t_1}^{t_2} \mathrm{d}t \int \mathrm{d}x \left\{ \frac{\partial \mathcal{L}}{\partial u} \delta u + \frac{\partial \mathcal{L}}{\partial(\partial u/\partial x)} \delta\left(\frac{\partial u}{\partial x}\right) + \frac{\partial \mathcal{L}}{\partial(\partial u/\partial t)} \delta\left(\frac{\partial u}{\partial t}\right) \right\} \quad (211)$$

$$= \int_{t_1}^{t_2} \mathrm{d}t \int \mathrm{d}x \left\{ \frac{\partial \mathcal{L}}{\partial u} \delta u - \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial(\partial u/\partial x)}\right) \delta u - \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial(\partial u/\partial t)}\right) \delta u \right\} \quad (212)$$

and the integrations by part in the second expression are justified by the fact that the variations vanish at the limits of the integration intervals. If this variation is to vanish for arbitrary δu we must have

$$\frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial(\partial u/\partial x)}\right) + \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial(\partial u/\partial t)}\right) - \frac{\partial \mathcal{L}}{\partial u} = 0. \quad (213)$$

This is the Euler-Lagrange equation for the case of a field $u(x,t)$, i.e. a continuum of degrees of freedom. For the case of our Lagrange density (209) equation (213) becomes

$$Y \frac{\partial^2 u}{\partial x^2} - \lambda \frac{\partial^2 u}{\partial t^2} = 0. \quad (214)$$

This should be recognized as the wave equation for the propagation of a disturbance with velocity $\sqrt{Y/\lambda}$.