

# PHY422/820: Classical Mechanics

Fall Semester 2019

H. Hergert  
December 2, 2019



**MICHIGAN STATE**  

---

**U N I V E R S I T Y**

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Themes . . . . .	1
1.2	Brief Recap of Newtonian Mechanics . . . . .	1
<b>2</b>	<b>Lagrangian Mechanics</b>	<b>2</b>
2.1	Constraints . . . . .	2
2.2	Geometrical Derivation of the Lagrange Equations . . . . .	4
2.2.1	The Single-Particle Case . . . . .	5
2.2.2	Lagrange Equations for Multiple Particles and Degrees of Freedom . . . . .	11
2.2.3	Examples . . . . .	12
2.3	The Principle of Least Action . . . . .	15
2.3.1	Elements of Variational Calculus . . . . .	17
2.3.2	Examples . . . . .	19
2.3.3	The Principle of Least Action for Mechanical Systems . . . . .	21
2.4	Constraints Revisited: Lagrange Equations of the First and Second Kind . . . . .	21
2.4.1	Examples . . . . .	22
2.5	Symmetries and Invariances . . . . .	28
2.5.1	Cyclic Coordinates and Noether's Theorem . . . . .	28
2.5.2	Spatial Translations . . . . .	30
2.5.3	Galilean Boosts . . . . .	31
2.5.4	Translations in Time . . . . .	32
2.5.5	Deriving Lagrangians from Symmetries . . . . .	36
2.6	Velocity-Dependent Potentials . . . . .	38
2.7	Dissipation . . . . .	40
2.7.1	The Dissipation Function . . . . .	40
2.7.2	Viscous Friction and Rayleigh's Dissipation Function . . . . .	42
2.7.3	Coulomb or Dry Friction . . . . .	44
<b>3</b>	<b>Oscillations</b>	<b>46</b>
3.1	Simple Oscillators . . . . .	46
3.1.1	Damping . . . . .	46
3.1.2	Periodic Driving Forces and Resonance . . . . .	48
3.1.3	Phase Space . . . . .	49
3.1.4	General Driving Forces . . . . .	49
3.2	Coupled Oscillators . . . . .	52

3.2.1	Example: Two Coupled Pendula . . . . .	52
3.2.2	Normal Coordinates and Normal Modes . . . . .	55
3.2.3	Example: A Linear Triatomic Molecule . . . . .	59
<b>4</b>	<b>Central Forces</b>	<b>63</b>
4.1	Introduction . . . . .	63
4.2	Reducing Two-Body Problems to Equivalent One-Body Problems . . . . .	63
4.3	Trajectories in the Central-Force Problem . . . . .	65
4.3.1	General Solution of the Equations of Motion . . . . .	65
4.3.2	Geometry of Central-Force Trajectories . . . . .	68
4.3.3	Stability of Circular Orbits . . . . .	70
4.3.4	Closed Orbits and Bertrand's Theorem . . . . .	72
4.3.5	The Kepler Problem . . . . .	75
4.3.6	Perihelion Precession from General Relativity . . . . .	75
4.4	Scattering . . . . .	76
4.5	Orbital Dynamics . . . . .	76
4.5.1	Transfer Orbits . . . . .	76
4.5.2	Gravitational Slingshot . . . . .	76
4.6	The Three-Body Problem . . . . .	76
<b>5</b>	<b>The Rigid Body</b>	<b>79</b>
<b>6</b>	<b>Hamiltonian Mechanics</b>	<b>80</b>
6.1	The Hamiltonian . . . . .	80
<b>A</b>	<b>Lagrange Multipliers</b>	<b>81</b>
A.1	General Procedure . . . . .	81
A.1.1	The Second-Derivative Test . . . . .	81
A.2	Examples . . . . .	81
A.2.1	Rectangle Inscribed in an Ellipse . . . . .	81
A.2.2	Extrema of the Mexican Hat Potential . . . . .	83



# Chapter 1

## Introduction

1.1 Themes

1.2 Brief Recap of Newtonian Mechanics

## Chapter 2

# Lagrangian Mechanics

### 2.1 Constraints

In many applications of classical mechanics, we are dealing with constrained motion. Naively, we would assign Cartesian coordinates to all masses of interest because that is easy to visualize, and then solve the equations of motion resulting from Newton's Second Law. In this process, any restrictions to the motion will be modeled with the help of **constraint forces** like the normal force, string tension, etc. The problem with that approach is that the constraint forces can only be determined once the dynamical equations have been solved. It would be much more efficient to exploit the constraints immediately, so that we could describe the motion using the actual degrees of freedom.

An extreme example is the description of any rigid body, e.g., a chair. Microscopically, a chair weighing a few kilograms consists on the order of  $10^{27}$  atoms, which can naively move in all three spatial directions. This leaves us with  $(10^{27})^3 = 10^{81}$  coordinates, so any attempt to describe the motion of the atoms using the tools of classical mechanics (neglecting quantum effects) is completely out of the question. However, the motion of these "classical atoms" is severely constrained: For each pair of particles, there is a constraint of the form

$$(\vec{r}_i - \vec{r}_j)^2 - c_{ij}^2 = 0 \quad (2.1)$$

that fixes their distance. Naively, we then have  $3N$  coordinates and  $N(N-1)/2$  pairwise constraints, but for  $N > 3$  the distance constraints become redundant, i.e., the distances between an added particle and any 3 particles in the rigid body completely determine all other distances. Thus, the actual number of degrees of freedom for  $N \geq 3$  is 6, which can be interpreted as the translational and rotational degrees of freedom of the solid. That is quite the reduction from  $10^{81}$  coordinates!

A related situation is that we have a system of  $N$  particles whose center of mass is held fixed at all times. In that case, we have a constraint that not just involves pairs of particles, as above, but all particle coordinates at once:

$$\frac{1}{M} \sum_i m_i \vec{r}_i = 0. \quad (2.2)$$

If this constraint were combined with the previous one to describe, e.g., a system of  $N$  connected masses, that system could then only rotate around the center of mass.

[More to fill in...]

### Configuration Space and Configuration Manifold

Let us consider the description of a system consisting of  $N$  particles. Each of the particles can move in three spatial directions, so we have  $3N$  coordinates in total, hence we can indicate the *state* of our system at a time  $t$  by a point in a  $3N$ -dimensional real vector space  $\mathbb{R}^{3N}$  that we will refer to as the **configuration space**. We can write

$$\vec{r}(t) = \begin{pmatrix} \vec{r}_1(t) \\ \vdots \\ \vec{r}_N(t) \end{pmatrix} = \begin{pmatrix} x_1(t) \\ y_1(t) \\ z_1(t) \\ \vdots \\ x_N(t) \\ y_N(t) \\ z_N(t) \end{pmatrix} \quad (2.3)$$

[...] Now let us assume that we impose a set of  $C$  constraints on our system. In general, constraint can be functions of the particle's trajectories, their velocities, etc., as well as time, and we can write

$$f_c(\vec{r}_1, \dots, \vec{r}_N, \dots, t) = 0, \quad c = 1, \dots, C. \quad (2.4)$$

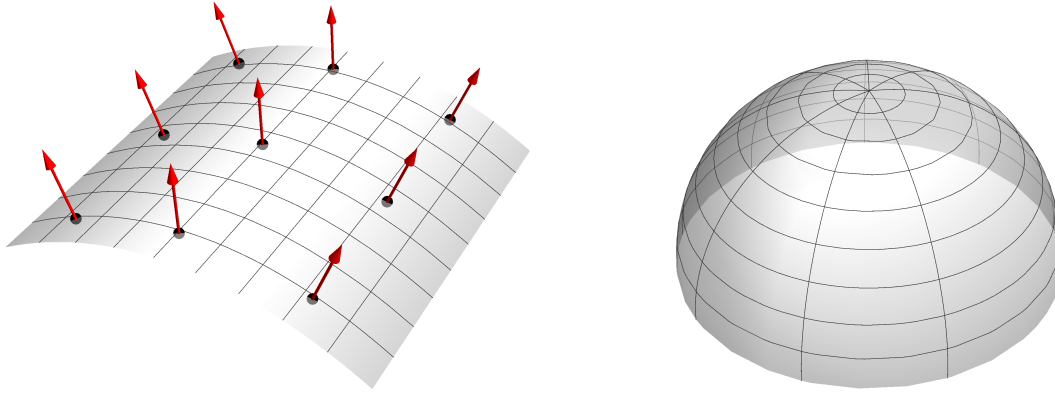
Such constraints will restrict the possible configurations of the system to a *manifold* that is defined by the set of all points that satisfy the constraints:

$$\mathcal{M} = \{ \vec{r} \in \mathbb{R}^{3N} \mid f_1(\vec{r}, t) = 0 \wedge \dots \wedge f_C(\vec{r}, t) = 0 \} \quad (2.5)$$

We shall refer to  $\mathcal{M}$  as the **configuration manifold**<sup>1</sup>

---

<sup>1</sup> Other authors and textbooks sometimes refer to this manifold as the configuration space. That is problematic because one cannot define vectors that lie within the manifold and perform the usual algebraic operations like constructing linear combinations or multiplying by a scalar, because one would usually have to leave the manifold in the process (think, e.g., of the surface of a sphere). It is, however, possible to define **tangent vector spaces** at each point of  $\mathcal{M}$ . Structurally, a vector space is always a manifold, but a manifold need not be a vector space.



**Figure 2.1:** Constraint manifolds and constraint forces.

[...]

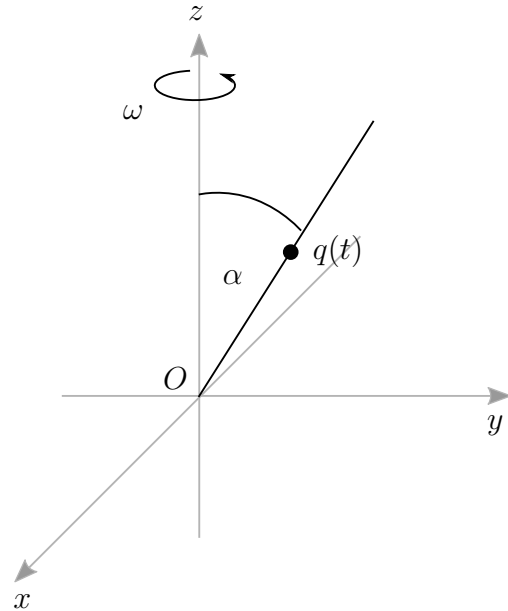
$$x(t) = q(t) \sin \alpha \cos \omega t, \quad (2.6)$$

$$y(t) = q(t) \sin \alpha \sin \omega t, \quad (2.7)$$

$$z(t) = q(t) \cos \alpha. \quad (2.8)$$

In Cartesian coordinates, the constraints are

$$\tan \alpha = \frac{\sqrt{x^2 + y^2}}{z}. \quad (2.9)$$



**Figure 2.2:** Bead sliding on rotating wire.

## 2.2 Geometrical Derivation of the Lagrange Equations

In the present section, we will derive **D'Alembert's principle** and the **Lagrange** equations using the geometric view of constraints we introduced in the previous section.



**Box 2.1: Summary: Types of Constraints**

Type	Form	Examples
<b>holonomic</b>		
scleronomic	$f(q_1, \dots, q_n) = 0$	rigid body, pendulum ...
rheonomic	$f(q_1, \dots, q_n, t) = 0$	bead on rotating wire, pendulum with moving suspension
<b>nonholonomic</b>		
	$f(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n) = 0$	disk rolling without slipping
	$f(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t) = 0$	processes with friction
	$f(q_1, \dots, q_n) \geq 0$	mass sliding off a semisphere

**2.2.1 The Single-Particle Case**

We follow the discussion of Beers [1], and start by considering a bead that can slide without friction on a spiral wire with cross section radius  $a$  (see Fig. 2.3). The coordinates of the bead are

$$x = a \cos \phi, \quad (2.10)$$

$$y = a \sin \phi, \quad (2.11)$$

$$z = b\phi, \quad (2.12)$$

where we choose  $\phi$  as our generalized coordinate. We can parameterize multiple loops of the spiral by allowing  $\phi \in ]-\infty, \infty[$  — otherwise, we would have to eliminate the angle in favor of the coordinate  $z$ , which leads to messier equations for  $x$  and  $y$ .)

The definition of  $z$  is one of the constraints on the bead; the other is given by

$$x^2 + y^2 = a^2. \quad (2.13)$$

Together, the two constraints define the constraint manifold, i.e., the spiral wire.

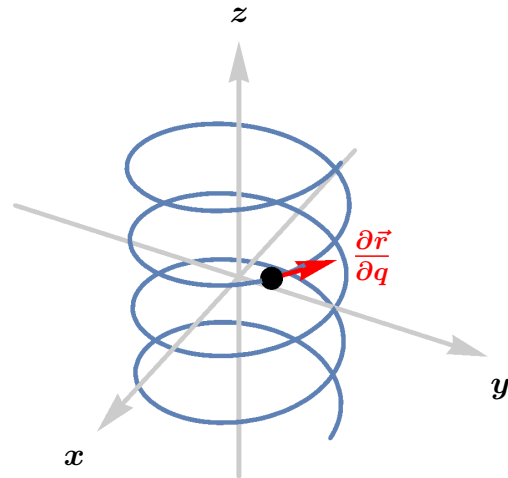
Now consider Newton's Second Law,

$$m\ddot{\vec{r}} = \vec{F}_A + \vec{F}_C, \quad (2.14)$$

where  $\vec{F}_A$  is the total “applied” force acting on the bead, and  $\vec{F}_C$  is the total constraint force. Here, the applied force is gravity, so

$$\vec{F}_A = -mg\vec{e}_z, \quad (2.15)$$

and the explicit form of  $\vec{F}_C$  can only be determined *after* we have solved the equations of motion for the bead. We do know from our discussion in Sec. 2.1 that the constraint forces



**Figure 2.3:** Bead gliding on a spiral wire without friction.

are proportional to the *gradient* of the *orthogonal* to the manifold itself in each point, so the scalar product with the *tangent* to the manifold must vanish. We obtain a tangent vector to the manifold the usual way, by taking the derivative of the coordinates with respect to  $\phi$ ,

$$\frac{\partial \vec{r}}{\partial \phi} = (-a \sin \phi, a \cos \phi, b)^T = a\vec{e}_\phi + b\vec{e}_z, \quad (2.16)$$

where we have introduced the usual unit vectors of a cylindrical coordinate system, and we have

$$\vec{F}_C \cdot \frac{\partial \vec{r}}{\partial \phi} = 0 \quad (2.17)$$

(details are worked out in Box 2.2).

We can exploit the orthogonality condition (2.17) by projecting Eq. (2.14) onto the tangent vector:

$$m\ddot{\vec{r}} \cdot \frac{\partial \vec{r}}{\partial \phi} = \vec{F}_A \cdot \frac{\partial \vec{r}}{\partial \phi} + \underbrace{\vec{F}_C \cdot \frac{\partial \vec{r}}{\partial \phi}}_{=0}. \quad (2.18)$$

In this way, we do not have to worry about the constraint forces, at least for the time being. In fact, we can actually drop the distinction between applied and constraint forces in this equation and simply write

$$m\ddot{\vec{r}} \cdot \frac{\partial \vec{r}}{\partial \phi} = \vec{F} \cdot \frac{\partial \vec{r}}{\partial \phi}, \quad (2.19)$$

because the constraint forces are projected out from  $\vec{F}$  automatically!

Let us now consider the left-hand side of Eq. (2.19). The trajectory of the bead can be written as  $\vec{r}(t) = \vec{r}(\phi(t))$ , i.e., the time dependence is entirely contained in the time dependence of the generalized coordinate  $\phi$ , but all steps would work analogously if  $\vec{r}$  were explicitly time dependent. Using the chain rule, we have

$$\dot{\vec{r}} = \frac{\partial \vec{r}}{\partial \phi} \dot{\phi}, \quad (2.20)$$

$$\ddot{\vec{r}} = \frac{\partial^2 \vec{r}}{\partial \phi^2} \dot{\phi}^2 + \frac{\partial \vec{r}}{\partial \phi} \ddot{\phi}. \quad (2.21)$$

The vector in the first term can be written as

$$\frac{\partial^2 \vec{r}}{\partial \phi^2} = (-a \cos \phi, -a \sin \phi, 0)^T = -\vec{e}_\rho, \quad (2.22)$$

so we see that

$$\frac{\partial^2 \vec{r}}{\partial \phi^2} \cdot \frac{\partial \vec{r}}{\partial \phi} = -\vec{e}_\rho \cdot (a\vec{e}_\phi + b\vec{e}_z) = 0 \quad (2.23)$$

because of the orthogonality of the unit vectors. Consequently, we obtain

$$m\ddot{\vec{r}} \cdot \frac{\partial \vec{r}}{\partial \phi} = \ddot{\phi} \left| \frac{\partial \vec{r}}{\partial \phi} \right|^2 = \ddot{\phi}(a^2 + b^2). \quad (2.24)$$

The right-hand side is obtained easily:

$$\vec{F} \cdot \frac{\partial \vec{r}}{\partial \phi} = -mg\vec{e}_z \cdot \frac{\partial \vec{r}}{\partial \phi} = -mgb. \quad (2.25)$$

**Box 2.2: Constraint Forces for a Bead on a Spiral Wire**

Let us compute the constraint forces acting on the bead. Using cylindrical coordinates for convenience, we have two constraint equations that define our manifold:

$$f_1(\rho, \phi, z) = \rho - a = 0, \quad (\text{B2.2-1})$$

$$f_2(\rho, \phi, z) = z - b\phi = 0. \quad (\text{B2.2-2})$$

The gradient in cylindrical coordinates is

$$\nabla = \vec{e}_\rho \frac{\partial}{\partial \rho} + \vec{e}_\phi \frac{1}{\rho} \frac{\partial}{\partial \phi} + \vec{e}_z \frac{\partial}{\partial z}, \quad (\text{B2.2-3})$$

so we obtain

$$\nabla f_1 = \vec{e}_\rho, \quad (\text{B2.2-4})$$

$$\nabla f_2 = \vec{e}_z - \frac{b}{\rho} \vec{e}_\phi. \quad (\text{B2.2-5})$$

Now we can compute the scalar products with the tangent vector (2.16):

$$\frac{\partial \vec{r}}{\partial \phi} \cdot \nabla f_1 = \vec{e}_\rho \cdot (a\vec{e}_\phi + b\vec{e}_z) = 0, \quad (\text{B2.2-6})$$

$$\begin{aligned} \frac{\partial \vec{r}}{\partial \phi} \cdot \nabla f_2 &= \left( \vec{e}_z - \frac{b}{\rho} \vec{e}_\phi \right) \cdot (a\vec{e}_\phi + b\vec{e}_z) \\ &= b - \frac{b}{\rho} a = b - \frac{b}{a} a = 0, \end{aligned} \quad (\text{B2.2-7})$$

where we have used the orthogonality of the unit vectors  $\{\vec{e}_\rho, \vec{e}_\phi, \vec{e}_z\}$ .

Thus, the **equation of motion** for the bead becomes

$$\begin{aligned} m(a^2 + b^2)\ddot{\phi} &= -mgb \\ \Rightarrow \ddot{\phi} &= -\frac{b}{a^2 + b^2}g, \end{aligned} \quad (2.26)$$

which has the general solution

$$\phi(t) = \phi_0 + \omega_0 t - \frac{1}{2} \frac{bg}{a^2 + b^2} t^2. \quad (2.27)$$

When we multiply  $\phi(t)$  with  $b$  (cf. Eq. (2.12)), we essentially obtain a “free” fall in  $z$  with a reduced acceleration.

**D’Alembert’s Principle**

We can rewrite our projected form of Newton’s Second Law, Eq. (2.19), as

$$\left( \vec{F} - m\ddot{\vec{r}} \right) \cdot \frac{\partial \vec{r}}{\partial q} = 0, \quad (2.28)$$

**Box 2.3: Virtual Work**

Traditionally, d'Alembert's principle has been stated in the following form:

*The constraint forces perform no virtual work.*

**Virtual work** is defined as the work done by the forces along virtual displacements that are compatible with the constraints:

$$\delta W \equiv \vec{F} \cdot \delta \vec{r} = \vec{F}_A \cdot \delta \vec{r}. \quad (\text{B2.3-1})$$

It can be a useful concept in the determination and analysis of *static equilibrium* configurations of mechanical systems, which are defined by

$$\delta W = 0. \quad (\text{B2.3-2})$$

In his original work, d'Alembert generalized this idea by including the inertial force  $-m\ddot{\vec{r}}$  (or  $-\dot{\vec{p}}$  if the mass is allowed to change) in his balancing equation for the virtual work, giving rise to Eq. (2.30). Mathematically, this moves the inertial term in Newton's Law from one side of the equation to the other, but philosophically, it removes the special role that the inertial term plays in Newtonian mechanics and instead treats it on an equal footing as all other forces.

where we switched to an arbitrary generalized coordinate  $q$ . Now we can recall the definition of the virtual displacements along the constraint manifolds from Sec. 2.1, which states

$$\delta \vec{r} = \frac{\partial \vec{r}}{\partial q} \delta q. \quad (2.29)$$

Thus, we can simply multiply the geometric condition (2.28) by  $\delta q$  on both sides, and obtain

$$\left( \vec{F} - m\ddot{\vec{r}} \right) \cdot \delta \vec{r} = 0. \quad (2.30)$$

This is **d'Alembert's principle** for a single particle in the form that is usually found in textbooks. It is statement of the fundamental laws of classical motion, and while it is equivalent to Newton's Laws for essentially all intents and purposes, *it cannot be derived from them*. It is also fundamentally related to the Principle of Least Action that we will discuss below, but in fact more general since it applies to systems with non-holonomic constraints. For further reading, I recommend the beautiful (and celebrated) book by Cornelius Lanczos [2], as well as the more recent work by Jennifer Coopersmith [3].

### The Lagrange Equations

A drawback of Newton's Second Law is that the form of the resulting equations of motion for each coordinate of a particle are in general *not invariant* under *point transformations* between two sets of independent coordinates,

$$q_i(s_1, \dots, s_n) \rightarrow s_i(q_1, \dots, q_n), \quad i = 1, \dots, n. \quad (2.31)$$

For instance, the term  $m\ddot{\vec{r}}$  already becomes much more complicated if we switch to curvilinear coordinates to describe motion in an inertial frame, and in non-inertial frames *fictitious* terms like the centrifugal or Coriolis forces appear.

As an example, consider a mass moving in a two-dimensional plane. In Cartesian coordinates, Newton's Second Law yields the following equations of motion:

$$\ddot{x} = \frac{F_x}{m}, \quad (2.32)$$

$$\ddot{y} = \frac{F_y}{m}. \quad (2.33)$$

If we use polar coordinates,

$$x = r \cos \phi, \quad y = r \sin \phi, \quad (2.34)$$

we have

$$\vec{r} = r\vec{e}_r, \quad (2.35)$$

$$\dot{\vec{r}} = \dot{r}\vec{e}_r + r\dot{\vec{e}}_r = \dot{r}\vec{e}_r + r\dot{\phi}\vec{e}_\phi, \quad (2.36)$$

$$\ddot{\vec{r}} = (\ddot{r} - r\dot{\phi}^2)\vec{e}_r + (2\dot{r}\dot{\phi} + r\ddot{\phi})\vec{e}_\phi, \quad (2.37)$$

from which we obtain the equations of motion in the form

$$\ddot{r} - r\dot{\phi}^2 = \frac{F_r}{m}, \quad (2.38)$$

$$2\dot{r}\dot{\phi} + r\ddot{\phi} = \frac{F_\phi}{m}, \quad (2.39)$$

with  $F_r = \vec{F} \cdot \vec{e}_r$  and  $F_\phi = \vec{F} \cdot \vec{e}_\phi$ . The dynamics of the mass will be the same in both descriptions, of course.

D'Alembert's principle (2.30), on the other hand, can be rewritten in a way that makes it form-invariant under point transformations between generalized coordinates. For simplicity, we will consider a holonomic system with one generalized coordinate first, and generalize our result afterwards.

We start with the force term:

$$\vec{F} \cdot \delta\vec{r} = \vec{F} \cdot \frac{\partial\vec{r}}{\partial q} \delta q \equiv Q\delta q. \quad (2.40)$$

Here, we have defined the **generalized force**  $Q$  associated with the variable  $q$ . Just like a generalized coordinate need not have the dimensions of a length,  $Q$  need not have the dimensions of a force, but  $Q\delta q$  will *always* have the dimensions of an energy.

Let us now rewrite the inertial term. Using the product rule, we have

$$m\ddot{\vec{r}} \cdot \frac{\partial\vec{r}}{\partial q} = \frac{d}{dt} \left( m\dot{\vec{r}} \cdot \frac{\partial\vec{r}}{\partial q} \right) - m\dot{\vec{r}} \cdot \frac{d}{dt} \frac{\partial\vec{r}}{\partial q}. \quad (2.41)$$

Consider the time derivative of  $\vec{r}$ , which we can evaluate using the chain rule:

$$\dot{\vec{r}} = \frac{\partial\vec{r}}{\partial q} \dot{q}. \quad (2.42)$$

Since we have holonomic constraints,  $\vec{r}$  does not depend on  $\dot{q}$ , and  $\dot{\vec{r}}$  is a function that only depends on  $\dot{q}$  *linearly*. If we take the partial derivative with respect to  $\dot{q}$  on both sides of Eq. (2.42), we obtain

$$\frac{\partial \dot{\vec{r}}}{\partial \dot{q}} = \frac{\partial \vec{r}}{\partial q}. \quad (2.43)$$

This identity is often referred to as a *cancellation of dots*, but keep in mind that it only applies under certain circumstances. Now consider the second term on the right-hand side of Eq. (2.41). The derivatives commute (even for non-holonomic constraints and explicitly time dependent  $\vec{r}$ ), so we have

$$\frac{d}{dt} \frac{\partial \vec{r}}{\partial q} = \frac{\partial \dot{\vec{r}}}{\partial q}. \quad (2.44)$$

Using identities (2.43) and (2.44), we can rewrite Eq. (2.41) as

$$m\ddot{\vec{r}} \cdot \frac{\partial \vec{r}}{\partial q} = \frac{d}{dt} \left( m\dot{\vec{r}} \cdot \frac{\partial \vec{r}}{\partial \dot{q}} \right) - m\dot{\vec{r}} \cdot \frac{\partial \dot{\vec{r}}}{\partial q}. \quad (2.45)$$

If we now introduce the kinetic energy  $T(q, \dot{q}) = \frac{1}{2}m\dot{\vec{r}}^2$ , we recognize its partial derivatives with respect to  $q$  and  $\dot{q}$  on the right-hand side:

$$\begin{aligned} m\ddot{\vec{r}} \cdot \frac{\partial \vec{r}}{\partial q} &= \frac{d}{dt} \frac{\partial}{\partial \dot{q}} \left( \frac{1}{2}m\dot{\vec{r}}^2 \right) - \frac{\partial}{\partial \dot{q}} \left( \frac{1}{2}m\dot{\vec{r}}^2 \right) \\ &= \frac{d}{dt} \frac{\partial T}{\partial \dot{q}} - \frac{\partial T}{\partial q}. \end{aligned} \quad (2.46)$$

Putting everything together, we see that d'Alembert's principle satisfies

$$\begin{aligned} (m\ddot{\vec{r}} - F) \cdot \delta \vec{r} &= (m\ddot{\vec{r}} - F) \cdot \frac{\partial \vec{r}}{\partial q} \delta q \\ &= \left( \frac{d}{dt} \frac{\partial T}{\partial \dot{q}} - \frac{\partial T}{\partial q} - Q \right) \delta q = 0. \end{aligned} \quad (2.47)$$

Since this equation has to hold for *arbitrary* virtual displacements  $\delta q$ , the expression in the parenthesis must vanish, which leads us to the **Lagrange equation** for a system with one degree of freedom:

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}} - \frac{\partial T}{\partial q} - Q = 0. \quad (2.48)$$

If the forces acting on the system are *conservative*,  $Q$  can be derived from a potential  $V(q)$

$$Q = -\frac{\partial V}{\partial q}. \quad (2.49)$$

In this case, we can define the **Lagrangian**  $L$  of the system,

$$L(q, \dot{q}) \equiv T(q, \dot{q}) - V(q), \quad (2.50)$$

and rewrite Eq. (2.47) as

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0. \quad (2.51)$$

### 2.2.2 Lagrange Equations for Multiple Particles and Degrees of Freedom

The derivations of the previous section can be readily generalized to multiple particles and degrees of freedom without going through the math in detail again. In an  $N$ -particle system, we obviously have positions and velocities for each particle, corresponding to  $3N$  coordinates:

$$\vec{r}, \dot{\vec{r}} \quad \longrightarrow \quad \{\vec{r}_i, \dot{\vec{r}}_i\}_{i=1, \dots, N}. \quad (2.52)$$

If the system is subject to  $k$  holonomic constraints, the total number of degrees of freedom is  $n = 3N - k$ , and to each of them we associate an independent generalized coordinate and velocity:

$$q, \dot{q} \quad \longrightarrow \quad \{q_j, \dot{q}_j\}_{j=1, \dots, 3N-k}. \quad (2.53)$$

In general, each  $\vec{r}_i$  is a function of *all* generalized coordinates, because the motion of each particle could be related to that of all others by the constraint — just think of the case of the rigid body. As a consequence, differentials and virtual displacements will have the generalization

$$\delta \vec{r} = \frac{\partial \vec{r}}{\partial q} \delta q \quad \longrightarrow \quad \delta \vec{r}_i = \sum_{j=1}^{3N-k} \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j, \quad (2.54)$$

$$d\vec{r} = \frac{\partial \vec{r}}{\partial q} dq + \frac{\partial \vec{r}}{\partial t} dt \quad \longrightarrow \quad d\vec{r}_i = \sum_{j=1}^{3N-k} \frac{\partial \vec{r}_i}{\partial q_j} dq_j + \frac{\partial \vec{r}}{\partial t} dt. \quad (2.55)$$

Using these rules for the coordinates, we can generalize the kinetic and potential energies as

$$T = \frac{1}{2} m \dot{\vec{r}}^2 \quad \longrightarrow \quad \frac{1}{2} \sum_i m_i \dot{\vec{r}}_i^2 \quad (2.56)$$

and

$$V(\vec{r}) \quad \longrightarrow \quad V(\vec{r}_1, \dots, \vec{r}_N), \quad (2.57)$$

respectively. Here we have assumed that the potential does not depend on the velocity: This is the case for most of our applications, although we will discuss an important counter-example in Sec. 2.6. Since the particle coordinates depend on the  $q_i$ , we can also express the potential energy in terms of the generalized coordinates instead:

$$V(q) \quad \longrightarrow \quad V(q_1, \dots, q_{3N-k}). \quad (2.58)$$

The generalized forces are extended via

$$Q = \vec{F} \cdot \frac{\partial \vec{r}}{\partial q} \quad \longrightarrow \quad Q_j = \sum_{i=1}^N \vec{F}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j}. \quad (2.59)$$

For conservative forces, we have

$$\vec{F}_i = -\nabla_i V(\vec{r}_1, \dots, \vec{r}_N), \quad (2.60)$$

where  $\nabla_i$  acts on the coordinates of particle  $i$ . Plugging this into the definition of the generalized force, we obtain

$$Q_j = - \sum_{i=1}^N (\nabla_i V) \cdot \frac{\partial \vec{r}_i}{\partial q_j} = - \sum_{i=1}^N \sum_{k=1}^3 \frac{\partial V}{\partial x_{ik}} \frac{\partial x_{ik}}{\partial q_j} = - \frac{\partial V}{\partial q_j}, \quad (2.61)$$

where we have written out the scalar product in components, and used the chain rule in the final step (noting again that  $V$  cannot depend on  $\dot{q}_j$  or  $t$ ).

Finally, we can state the many-particle version of d'Alembert's principle,

$$\sum_i \left( \vec{F}_i - \dot{\vec{p}}_i \right) \cdot \delta \vec{r}_i = 0, \quad (2.62)$$

as well as the Lagrange equations for each generalized coordinate:

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} - Q_j = 0. \quad j = 1, \dots, 3N - k, \quad (2.63)$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0. \quad j = 1, \dots, 3N - k \quad (2.64)$$

### 2.2.3 Examples

Let us now demonstrate the Lagrange formalism in action by working through some examples.

#### Bead on a Spiral Wire

In our first example, we revisit the problem of a bead on a spiral wire, which we used to derive d'Alembert's principle. We will, however, make one small alteration: Instead of starting from Cartesian coordinates, we will take the *symmetries* of the system into account and work in cylindrical coordinates  $\{r, \phi, z\}$  instead (cf. Fig. 2.3).

With this coordinate choice, the constraints of the motion can be expressed as

$$\rho - a = 0, \quad (2.65)$$

$$z - b\phi = 0, \quad (2.66)$$

and the polar angle  $\phi$  is the generalized coordinate. We recall that can let  $\phi$  perform an arbitrary amount of revolutions, so that we can cover the full height of the spiral wire, i.e., the range of  $z$  coordinates the spiral wire encompasses.

The trajectory of the bead can be written as

$$\vec{r} = \rho \vec{e}_\rho + z \vec{e}_z = a \vec{e}_\rho + b\phi \vec{e}_z, \quad (2.67)$$

which leads to the following expression for the velocity:

$$\dot{\vec{r}} = a \dot{\vec{e}}_\rho + b \dot{\phi} \vec{e}_z = a \dot{\phi} \vec{e}_\phi + b \dot{\phi} \vec{e}_z. \quad (2.68)$$



We can easily compute the square of the velocity vector, exploiting the orthonormality of the unit vectors,

$$\left(a\dot{\phi}\vec{e}_\phi + b\dot{\phi}\vec{e}_z\right)^2 = (a^2 + b^2)\dot{\phi}^2. \quad (2.69)$$

In this way, we obtain the kinetic energy

$$T = \frac{1}{2}m(a^2 + b^2)\dot{\phi}^2. \quad (2.70)$$

The potential energy is given by

$$V = mgz = mgb\phi, \quad (2.71)$$

so our Lagrangian is

$$L = T - V = \frac{1}{2}m(a^2 + b^2)\dot{\phi}^2 - mgb\phi. \quad (2.72)$$

Next, we compute the Lagrangian's partial derivatives with respect to  $\phi$  and  $\dot{\phi}$ ,

$$\frac{\partial L}{\partial \phi} = -mgb, \quad (2.73)$$

$$\frac{\partial L}{\partial \dot{\phi}} = m(a^2 + b^2)\dot{\phi}, \quad (2.74)$$

and plugging these into the Lagrange equation, we obtain

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} = m(a^2 + b^2)\ddot{\phi} + mgb = 0. \quad (2.75)$$

The equation of motion for  $\phi$  can be rearranged in the form

$$\ddot{\phi} = -\frac{gb}{a^2 + b^2}, \quad (2.76)$$

which is of course our result from Sec. 2.2.

### Block Sliding on a Gliding Wedge

As a second example, we consider a block of mass  $m$  sliding without friction on a wedge with inclination  $\alpha$  that can itself glide on a frictionless plane (Fig. 2.4). We can consider the motion of block and wedge in two dimensions if neither of them starts spinning while it moves. The coordinates for the wedge are  $X$ , its distance from the origin in the horizontal plane, and

$$Z = \text{const.}, \quad (2.77)$$

which is defined by our choice of coordinate system and acts as a constraint of the motion. The coordinates of the block are

$$x = X + s \cos \alpha, \quad z = H - s \sin \alpha, \quad (2.78)$$

where  $H$  is the height of the wedge. (We could have eliminated this constant by shifting the coordinate system in  $z$  direction — this will only give an offset to the potential energy that has no consequences for the dynamics.)

The time derivatives of the coordinates are

$$\dot{x} = \dot{X} + \dot{s} \cos \alpha, \quad (2.79)$$

$$\dot{z} = -\dot{s} \sin \alpha, \quad (2.80)$$

so the kinetic energy is

$$\begin{aligned} T &= \frac{1}{2} M \dot{X}^2 + \frac{1}{2} m (\dot{x}^2 + \dot{z}^2) \\ &= \frac{1}{2} M \dot{X}^2 + \frac{1}{2} m (\dot{X}^2 + \dot{s}^2 \cos^2 \alpha + 2\dot{X}\dot{s} \cos \alpha + \dot{s}^2 \sin^2 \alpha) \\ &= \frac{1}{2} (m + M) \dot{X}^2 + \frac{1}{2} m \dot{s}^2 + m \dot{X} \dot{s} \cos \alpha. \end{aligned} \quad (2.81)$$

The potential energy is given by

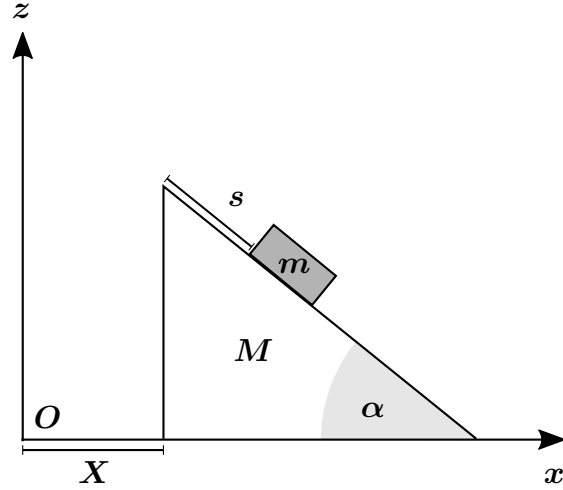
$$V = mg(h - s \sin \alpha), \quad (2.82)$$

so we obtain the Lagrangian

$$L = \frac{1}{2} (m + M) \dot{X}^2 + \frac{1}{2} m \dot{s}^2 + m \dot{X} \dot{s} \cos \alpha - mg(h - s \sin \alpha). \quad (2.83)$$

Now let us derive the Lagrange equations. We immediately notice that  $L$  does not explicitly depend on  $X$ , so we have

$$\frac{\partial L}{\partial X} = 0 = \frac{d}{dt} \frac{\partial L}{\partial \dot{X}}. \quad (2.84)$$



**Figure 2.4:** A block sliding without friction on a wedge that can itself glide on a frictionless plane.

### Box 2.4: Recipe for Solving Problems in Lagrangian Mechanics

As we have seen from our discussion of the examples in Sec. 2.2.3, the general procedure for solving problems in Lagrangian mechanics consists of the following steps:

1. Choose **convenient coordinates** for your problem, e.g., by exploiting **symmetries**.
2. Formulate the **constraints**.
3. Construct the **Lagrangian**.
4. Use the **Lagrange equations** to derive the **equations of motion** and identify **conserved quantities**.
5. Solve the equations of motion, and **analyze your solutions**.

This means that

$$\frac{\partial L}{\partial \dot{X}} = (m + M)\dot{X} + m\dot{s} \cos \alpha \quad (2.85)$$

is a conserved quantity. It is easy to see that Eq. (2.85) is the total momentum in the horizontal direction, which is conserved because there is *no external force* acting on the system in  $x$  direction. (The internal forces between the block and the wedge cancel because of Newton's Third Law.)

The second Lagrange equation is obtained from

$$\frac{\partial L}{\partial s} = mg \sin \alpha, \quad (2.86)$$

$$\frac{\partial L}{\partial \dot{s}} = m\dot{s} + m\dot{X} \cos \alpha, \quad (2.87)$$

which yields

$$\ddot{s} + \ddot{X} \cos \alpha = g \sin \alpha \quad (2.88)$$

( $s$  increases as the block slides down the slope). The conservation law (2.85) can be used to eliminate  $\ddot{X}$ :

$$(m + M)\ddot{X} = -m\ddot{s} \cos \alpha \quad \Rightarrow \quad \ddot{X} = -\frac{m}{m + M}\ddot{s} \cos \alpha, \quad (2.89)$$

so

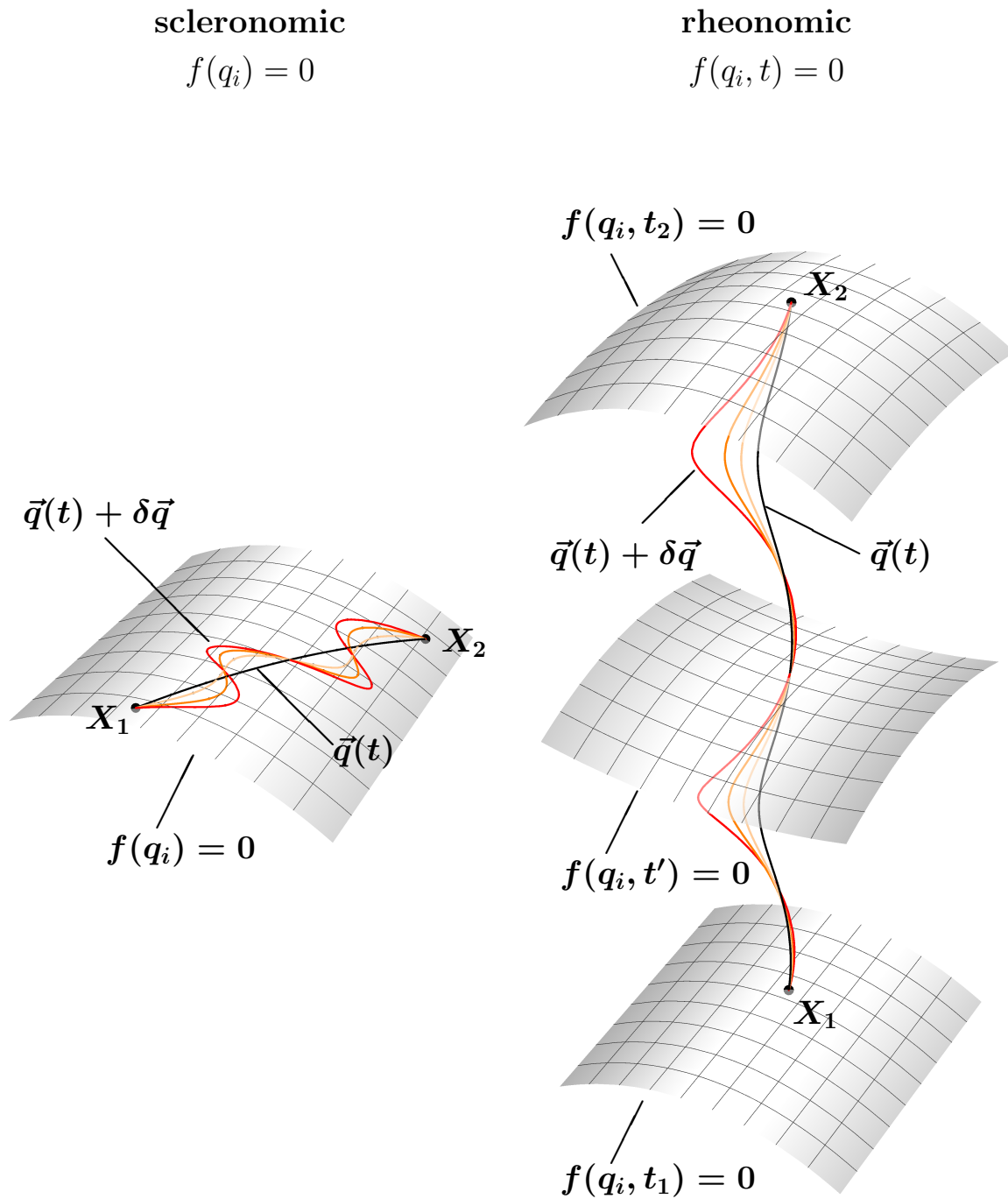
$$\ddot{s} - \frac{m}{m + M}\ddot{s} \cos^2 \alpha = g \sin \alpha. \quad (2.90)$$

Rearranging, we obtain the equation of motion

$$\ddot{s} = \frac{(m + M) \sin \alpha}{m \sin^2 \alpha + M} g. \quad (2.91)$$

## 2.3 The Principle of Least Action

[More stuff to fill in...]



**Figure 2.5:** Variation of trajectories with fixed endpoints in configuration manifolds defined by holonomic constraints.

### 2.3.1 Elements of Variational Calculus

#### Varying the Functional of a Curve

The calculus of variations aims to determine the function  $y(x)$  for which the integral

$$\mathcal{I}[y] \equiv \int_{x_1}^{x_2} dx f(x, y(x), y'(x)) \quad (2.92)$$

becomes stationary,  $\delta I = 0$ . The integral  $\mathcal{I}[y]$  is also referred to as a **functional** on the space of curves  $y(x)$  that are compatible with the boundary conditions, i.e., that have the same values at  $x_1$  and  $x_2$ .

Let us assume we already know the solution. We can define variations of this curve in the vicinity of the solution by defining

$$y(x, \epsilon) = y(x, 0) + \epsilon \eta(x), \quad \epsilon \ll 1, \quad (2.93)$$

where we choose an auxiliary function  $\eta(x)$  that is twice continuously derivable, to avoid singularities and general pathological behavior. We also demand that

$$\eta(x_1) = \eta(x_2) = 0, \quad (2.94)$$

so that the boundary conditions are automatically satisfied.

In this way,  $\mathcal{I}[y]$  becomes a function of  $\epsilon$ ,

$$\mathcal{I}(\epsilon) = \int_{x_1}^{x_2} dx f(x, y(x, \epsilon), y'(x, \epsilon)), \quad (2.95)$$

and since  $y(x, 0)$  is supposed to make the functional stationary, we can perform a Taylor expansion around  $\epsilon = 0$ :

$$\mathcal{I}(\epsilon) = \int_{x_1}^{x_2} dx \left( f(x, y(x, 0), y'(x, 0)) + \epsilon \left. \frac{\partial f}{\partial y} \right|_{\epsilon=0} \eta(x) + \epsilon \left. \frac{\partial f}{\partial y'} \right|_{\epsilon=0} \eta'(x) + O(\epsilon^2) \right). \quad (2.96)$$

The stationarity condition  $\delta \mathcal{I} = 0$  implies that

$$0 = \frac{d\mathcal{I}(\epsilon)}{d\epsilon} = \int_{x_1}^{x_2} dx \left( \left. \frac{\partial f}{\partial y} \right|_{\epsilon=0} \eta(x) + \left. \frac{\partial f}{\partial y'} \right|_{\epsilon=0} \eta'(x) \right). \quad (2.97)$$

The second term in the integrand can be rewritten using integration by parts, leading to

$$\begin{aligned} 0 &= \left. \frac{\partial f}{\partial y'} \eta(x) \right|_{x_1}^{x_2} - \int_{x_1}^{x_2} dx \left( \frac{\partial f}{\partial y} + \frac{d}{dx} \frac{\partial f}{\partial y'} \right) \eta(x) \\ &= \int_{x_1}^{x_2} dx \left( \frac{d}{dx} \frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y} \right) \eta(x), \end{aligned} \quad (2.98)$$

where we have used that the first term (sometimes referred to as the *boundary term*) vanishes at the boundaries, i.e., the starting and end points of the curve, because of the condition (2.94). Since Eq. (2.98) must hold for arbitrary  $\eta(x)$ , the expression in the parenthesis must vanish<sup>2</sup>, and we obtain the **Euler-Lagrange equation**

<sup>2</sup>This is properly proven in the fundamental lemma of the calculus of variations

$$\frac{d}{dx} \frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y} = 0, \quad (2.99)$$

This is both a *necessary and sufficient* condition that a curve  $y(x)$  must satisfy in order to make  $\mathcal{I}[y]$  stationary. The left-hand side of the Euler-Lagrange equation can also be used to define the **functional derivative**

$$\frac{\delta \mathcal{I}}{\delta y} \equiv \frac{d}{dx} \frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y}. \quad (2.100)$$

### Euler-Lagrange Equations for Multiple Degrees of Freedom and Variables

The extension of the Euler-Lagrange equations to multiple variables — i.e., multiple particles and degrees of freedom — is straightforward. A general curve will be characterized by the values of all coordinates  $y_1(x), \dots, y_n(x)$  as a function of the variable  $x$  that is used to parameterize it, and the functional generalizes to

$$\mathcal{I}[y_i] = \int_1^2 dx f(x, y_1, \dots, y_n, y_1', \dots, y_n'). \quad (2.101)$$

Variations of a curve that makes  $\mathcal{I}$  stationary are written as

$$y_i(x, \epsilon_i) = y_i(x) + \epsilon_i \eta_i(x) \equiv y_i(x, 0) + \delta y_i(x, \epsilon), \quad (2.102)$$

The  $\delta y_i$  must vanish at the start and end points of the curves, i.e.,

$$\delta y_i(x_1) = \delta y_i(x_2) = 0. \quad (2.103)$$

The stationarity condition can be expressed for independent (but infinitesimal) variations by introducing *epsilon* =  $(\epsilon_1, \dots, \epsilon_n)^T$ , and we obtain

$$0 = \delta \mathcal{I} = \nabla \mathcal{I}(\vec{\epsilon}) \cdot \vec{\epsilon} = \int_{x_1}^{x_2} \sum_{i=1}^n \left( \frac{\partial f}{\partial y_i} \delta y_i + \frac{\partial f}{\partial y_i'} \delta y_i' \right), \quad (2.104)$$

where  $\delta y_i' = \epsilon_i \eta_i'$ . Partially integrating as in the one-dimensional case, we get

$$0 = \int_{x_1}^{x_2} dx \sum_{i=1}^n \left( \frac{d}{dx} \frac{\partial f}{\partial y_i'} - \frac{\partial f}{\partial y_i} \right) \delta y_i, \quad (2.105)$$

and since the  $\delta y_i$  are independent, *all* parentheses must vanish separately, leading to the **Euler-Lagrange equations**

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx} \frac{\partial f}{\partial y_i'} = 0, \quad i = 1, \dots, n. \quad (2.106)$$

### 2.3.2 Examples

#### Brachistochrone

We parameterize the trajectory using the arc length  $s$ , which is the variable of choice if we are interested in the shape of a curve. Since  $ds = v dt$ , the time required to move from the start to the end of the curve is given by the functional

$$\mathcal{T} = \int_{t_1}^{t_2} dt = \int_1^2 ds \frac{1}{v}. \quad (2.107)$$

Energy conservation implies

$$E = \frac{mv^2}{2} + mgy = mgy_1, \quad (2.108)$$

so we can solve for  $v$  and obtain

$$v = \sqrt{2g(y_1 - y)}. \quad (2.109)$$

The differential can be rewritten as

$$ds = \sqrt{dx^2 + dy^2} = dx \sqrt{1 + y'(x)^2}, \quad y' \equiv \frac{dy}{dx}, \quad (2.110)$$

and plugging in our expression for the velocity, the functional  $\mathcal{T}$  becomes

$$\mathcal{T} = \int_1^2 ds \frac{1}{v} = \int_{x_1}^{x_2} dx \frac{\sqrt{1 + y'(x)^2}}{\sqrt{2g(y_1 - y(x))}}. \quad (2.111)$$

We need to find the trajectory that minimizes this integral, so we set

$$f(x, y, y') = \sqrt{\frac{1 + y'^2}{2g(y_1 - y)}}. \quad (2.112)$$

First, we note that  $f$  does not explicitly depend on  $x$ , which implies the so-called **Beltrami identity**:

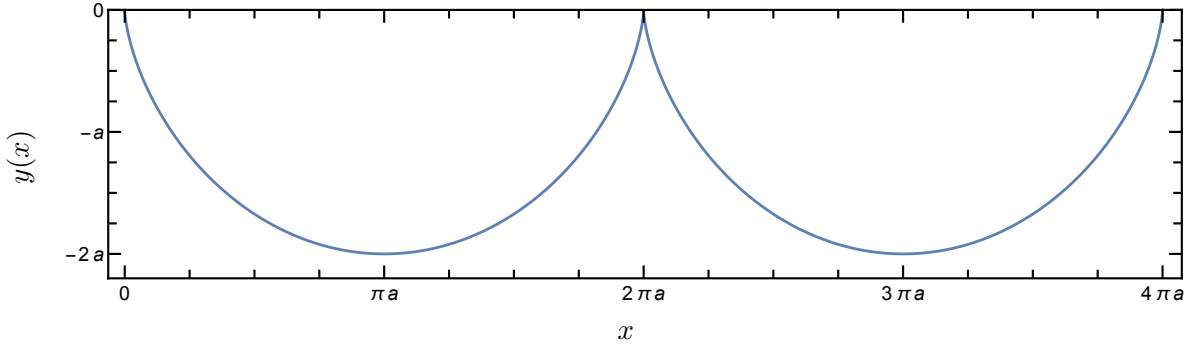
$$f - \frac{\partial f}{\partial y'} y' = \text{const.} \quad (2.113)$$

The proof is straightforward:

$$\begin{aligned} \frac{d}{dx} \left( f - \frac{\partial f}{\partial y'} y' \right) &= \frac{\partial f}{\partial y} y' + \frac{\partial f}{\partial y'} y'' - \frac{\partial f}{\partial y'} y'' - \frac{d}{dx} \left( \frac{\partial f}{\partial y'} \right) y' \\ &= \left( \frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} \right) y' = 0. \end{aligned} \quad (2.114)$$

Using the relation (2.113), we have

$$\begin{aligned} \text{const.} &= \sqrt{2g} \left( f - \frac{\partial f}{\partial y'} y' \right) \\ &= \sqrt{\frac{1 + y'^2}{y_1 - y}} - \frac{y'^2}{\sqrt{y_1 - y} \sqrt{1 + y'^2}} \end{aligned}$$



**Figure 2.6:** The cycloid defined by Eqs. (2.117), (2.118).

$$= \frac{1}{\underbrace{\sqrt{y_1 - y} \sqrt{1 + y'^2}}_{=1}} \left( 1 + y'^2 - y'^2 \right) \quad (2.115)$$

Squaring both sides, the equation can be rearranged as

$$(y_1 - y) (1 + y'^2) = \text{const.} \quad (2.116)$$

This differential equation is solved by the following *cycloid* trajectory (see Fig. 2.6):

$$x = a(t - \sin t) \quad (2.117)$$

$$y = a(\cos t - 1). \quad (2.118)$$

We can show that the cycloid satisfies Eq. (2.116) by plugging in Eqs. (2.117) and (2.118):

$$y' = \frac{dy}{dx} = \frac{dy}{dt} \frac{dt}{dx} = \frac{\dot{y}}{\dot{x}} = \frac{-a \sin t}{a(1 - \cos t)}, \quad (2.119)$$

hence

$$1 + y'^2 = 1 + \frac{a^2 \sin^2 t}{a^2(1 - \cos t)^2} = \frac{a^2(1 - 2 \cos t + \cos^2 t + \sin^2 t)}{y^2} \quad (2.120)$$

$$= \frac{2a^2}{y^2}(1 - \cos t) = -\frac{2a}{y}. \quad (2.121)$$

Noting that  $y_1 = 0$ , Eq. (2.116) now reads

$$(y_1 - y) (1 + y'^2) = (-y) \left( \frac{-2a}{y} \right) = 2a = \text{const.}, \quad (2.122)$$

as required.

### Shortest Line Connecting Two Points

We again start from the line element  $ds = \sqrt{dx^2 + dy^2} = dx \sqrt{1 + y'^2}$ , which defines

$$f = \sqrt{1 + y'^2}. \quad (2.123)$$



## 2.4. CONSTRAINTS REVISITED: LAGRANGE EQUATIONS OF THE FIRST AND SECOND KIND<sup>21</sup>

Plugging this into the Euler-Lagrange equation (2.99) yields

$$\frac{d}{dx} \frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y} = \frac{d}{dx} \frac{y'}{\sqrt{1+y'^2}} = 0 \quad \implies \quad \frac{y'}{\sqrt{1+y'^2}} = c. \quad (2.124)$$

We square both sides and rearrange the equation, obtaining

$$y'^2 = \frac{c^2}{1-c^2} \quad \implies \quad y' = \sqrt{\frac{c^2}{1-c^2}} = \text{const.} \quad (2.125)$$

This implies

$$y(x) = ax + b, \quad (2.126)$$

where  $a = c/\sqrt{1-c^2}$  and  $b$  is a constant obtained upon integration of the differential equation. Thus, the shortest trajectory connecting two points in a plane is a line. The same procedure can be used to compute the shortest connections — the so-called **geodesics** — between two points in arbitrary smooth manifolds.

### 2.3.3 The Principle of Least Action for Mechanical Systems

[...]

From the principle of least action we immediately see that the dynamics of a holonomic mechanical system remain invariant under the addition of a total time derivative to the Lagrangian: If we introduce

$$\tilde{L}(q_i, \dot{q}_i, t) = L(q_i, \dot{q}_i, t) + \frac{dF(q_i, t)}{dt}, \quad (2.127)$$

the corresponding action functional reads

$$\tilde{S} = S + \int_{t_1}^{t_2} dt \frac{dF(q_i, t)}{dt} = S + F_2 - F_1, \quad (2.128)$$

where  $F_1, F_2$  are constants. Thus, the variation of the action  $\tilde{S}$  is identical to that of the original action  $S$ :

$$\delta\tilde{S} = \delta S + \underbrace{\delta(F_2 - F_1)}_{=0}. \quad (2.129)$$

This means that the Lagrange equations of a holonomic system will remain invariant as well<sup>3</sup>.

## 2.4 Constraints Revisited: Lagrange Equations of the First and Second Kind

[...]

The holonomic constraints can be directly coupled to the Lagrangian by defining

$$\tilde{L}(\vec{q}, \dot{\vec{q}}, t, \vec{\lambda}) = L(\vec{q}, \dot{\vec{q}}, t) + \sum_{\alpha=1}^r \lambda_{\alpha} f_{\alpha}(\vec{q}, t). \quad (2.130)$$

---

<sup>3</sup>In a homework problem, this invariance was proven directly using the Lagrange equations. This is actually a stronger statement, because it allows the extension to systems with nonholonomic constraints.

Since the constraint equations are independent of the generalized velocity, the modified Lagrangian is compatible with the Principle of Least Action, and the variation of the action

$$S' = \int_{t_1}^{t_2} dt \tilde{L}(\vec{q}, \dot{\vec{q}}, t, \vec{\lambda}) = \int_{t_1}^{t_2} dt \left( L(\vec{q}, \dot{\vec{q}}, t) + \sum_a \lambda_a f_a(\vec{q}, t) \right) \quad (2.131)$$

will yield the Lagrange equations

$$\frac{d}{dt} \frac{\partial \tilde{L}}{\partial \dot{q}_j} - \frac{\partial \tilde{L}}{\partial q_j} = 0 \quad \Leftrightarrow \quad \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = \sum_{a=1}^r \lambda_a \frac{\partial f_a}{\partial q_j}, \quad (2.132)$$

and

$$\frac{d}{dt} \underbrace{\frac{\partial \tilde{L}}{\partial \dot{\lambda}_a}}_{=0} - \frac{\partial \tilde{L}}{\partial \lambda_a} = 0 \quad \Leftrightarrow \quad \frac{\partial \tilde{L}}{\partial \lambda_a} = f_a(\vec{q}, t) = 0 \quad (2.133)$$

[...]

Unfortunately, the same kind of construction cannot be used in the case of velocity-dependent nonholonomic constraints (and certainly not for constraint that are defined by inequalities). The treatment of such constraints is a long-standing problem that has led to several false starts and continuing misconceptions<sup>4</sup> — a discussion of the issues can be found, for instance, in a series of papers by M. Flannery [5, 6, 7].

[...]

The Lagrange equations of the first kind for  $r$  holonomic and  $s$  nonholonomic, velocity-dependent constraints (no inequalities!) are

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = \sum_{a=1}^r \lambda_a \frac{\partial f_a}{\partial q_j} + \sum_{b=1}^s \mu_b \frac{\partial g_b}{\partial \dot{q}_j}. \quad (2.134)$$

## 2.4.1 Examples

### Spinning a Mass on a String

Consider a mass that is being spun around on a string of length  $l$  with a constant angular velocity  $\omega$ , parallel to the ground. In polar coordinates, the trajectory of the mass is

$$\vec{r} = r \vec{e}_r, \quad (2.135)$$

and therefore

$$\dot{\vec{r}} = \dot{r} \vec{e}_r + r \dot{\phi} \vec{e}_\phi = \dot{r} \vec{e}_r + r \dot{\phi} \vec{e}_\phi. \quad (2.136)$$

The kinetic energy is

$$T = \frac{1}{2} m \left( \dot{r}^2 + r^2 \dot{\phi}^2 \right). \quad (2.137)$$

---

<sup>4</sup>Note that the discussion of nonholonomic constraints in the context of the principle of least action in [4] is erroneous as well - the specific example turns out to be correct if one foregoes the attempt to include the variational principle, but instead directly builds the constraint into the Lagrange equations of the first kind, Eqs. (2.134), relying on d'Alembert's principle.

#### 2.4. CONSTRAINTS REVISITED: LAGRANGE EQUATIONS OF THE FIRST AND SECOND KIND 23

Since the motion occurs in a plane that is parallel to the ground, the potential is constant, and we can choose the origin of our coordinate system such that  $V = 0$ . The constraints of the motion are

$$r - l = 0, \quad \phi - \omega t = 0, \quad (2.138)$$

so we can define the modified Lagrangian

$$\tilde{L} = \frac{1}{2}m \left( \dot{r}^2 + r^2 \dot{\phi}^2 \right) + \lambda_r (r - l) + \lambda_\phi (\phi - \omega t). \quad (2.139)$$

The Lagrange equations of the first kind are now given by

$$\frac{d}{dt} \frac{\partial \tilde{L}}{\partial \dot{r}} - \frac{\partial \tilde{L}}{\partial r} = m\ddot{r} - mr\dot{\phi}^2 - \lambda_r = 0, \quad (2.140)$$

$$\frac{d}{dt} \frac{\partial \tilde{L}}{\partial \dot{\phi}} - \frac{\partial \tilde{L}}{\partial \phi} = \frac{d}{dt} (mr^2\dot{\phi}) - \lambda_\phi = 0, \quad (2.141)$$

$$-\frac{\partial \tilde{L}}{\partial \lambda_r} = r - l = 0, \quad (2.142)$$

$$-\frac{\partial \tilde{L}}{\partial \lambda_\phi} = \phi - \omega t = 0. \quad (2.143)$$

From the first two equations, we obtain the Lagrange multipliers

$$\lambda_r = m\ddot{r} - mr\dot{\phi}^2 \quad (2.144)$$

and

$$\lambda_\phi = \frac{d}{dt} mr^2\dot{\phi}. \quad (2.145)$$

The right-hand side of this equation shows that  $\lambda_\phi$  is the time derivative of an angular momentum, i.e., a torque. We can now plug in the constraint equations for  $\phi$  and  $r$ , and obtain

$$\lambda_\phi = \frac{d}{dt} ml^2\omega = 0, \quad (2.146)$$

i.e., the torque vanishes. This is consistent:  $\omega = \text{const.}$  implies conservation of the angular momentum around the  $z$  axis, so there cannot be an external torque acting on the system.

From the Lagrange equation (2.144) we obtain

$$\lambda_r = \underbrace{m\ddot{r}}_{=0} - ml\dot{\phi}^2 = -ml\omega^2 = \text{const.} \quad (2.147)$$

We see that  $\lambda_r$  has the dimensions of a force, and it points in negative  $r$  direction, i.e., toward the hub of the circle — this means that  $\lambda_r$  is the **centripetal force** for the circular motion, which is provided by the string tension in the present case.

### Cylinder Rolling Down an Inclined Plane

Let us consider a cylinder of mass  $M$  and radius  $R$  that is rolling down an inclined plane (see figure). Its kinetic energy can be written as the sum of the center-of-mass translation and the cylinder's rotation around the symmetry axis through the center of mass (see Chapter 5):

$$T = \frac{1}{2}M\dot{s}^2 + \frac{1}{2}I\dot{\phi}^2, \quad (2.148)$$

where  $s$  is the distance the cylinder has rolled down the plane,  $\phi$  is the rotation angle, and  $I = \frac{1}{2}MR^2$ .

We now assume that the cylinder is **rolling without slipping**, which means that

$$ds = Rd\phi. \quad (2.149)$$

This is a nonholonomic constraint that connects the generalized velocities,

$$g(\dot{s}, \dot{\phi}) = \dot{s} - R\dot{\phi} = 0. \quad (2.150)$$

The potential energy of the cylinder is given by

$$V = V_0 - Mgs \sin \alpha, \quad (2.151)$$

where  $\alpha$  is the inclination angle. (Note that we have not explicitly considered the radius of the cylinder in the potential energy, which would just cause another constant offset that has no impact on the dynamics.) Thus, the Lagrangian is given by

$$L = \frac{1}{2}M\dot{s}^2 + \frac{1}{4}M\dot{\phi}^2 - V_0 + Mgs \sin \alpha. \quad (2.152)$$

Since the constraint is nonholonomic, we cannot couple it to the Lagrangian but add it to the Lagrange equation (see Eq. (2.134)):

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{s}} - \frac{\partial L}{\partial s} = \mu \frac{\partial g}{\partial \dot{s}} \Rightarrow M\ddot{s} - Mg \sin \alpha = \mu, \quad (2.153)$$

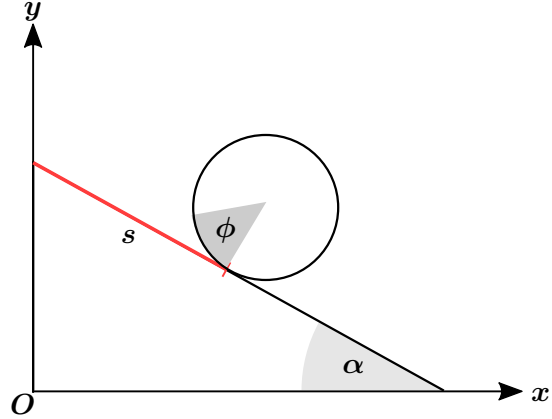
$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} = \mu \frac{\partial g}{\partial \dot{\phi}} \Rightarrow \frac{1}{2}MR^2\ddot{\phi} = -\mu R. \quad (2.154)$$

From the second equation and the time derivative of the constraint (2.150), we obtain

$$\mu = -\frac{1}{2}MR\ddot{\phi} = -\frac{1}{2}M\ddot{s}. \quad (2.155)$$

We plug this into the first equation of motion, which now becomes

$$M\ddot{s} - Mg \sin \alpha = -\frac{1}{2}M\ddot{s}, \quad (2.156)$$



**Figure 2.7:** Cylinder rolling without slipping on an inclined plane.

and rearranging, we have

$$\frac{3}{2}\ddot{s} - g \sin \alpha = 0 \quad \Rightarrow \quad \ddot{s} = \frac{2}{3}g \sin \alpha. \quad (2.157)$$

(Note that the acceleration is in growing  $s$  direction, i.e., downward along the plane, so the signs are correct.) Thus, the general solution to the equation of motion is given by

$$s(t) = s_0 + v_0 t + \frac{1}{2}\left(\frac{2}{3}g \sin \alpha\right)t^2 = s_0 + v_0 t + \frac{1}{3}(g \sin \alpha)t^2, \quad (2.158)$$

and assuming the cylinder starts rolling from rest at the top of the plane, we have

$$s(t) = \frac{1}{3}(g \sin \alpha)t^2. \quad (2.159)$$

We conclude the discussion of this example by noting that the nonholonomic constraint discussed here is actually **integrable**, i.e., a holonomic constraint in disguise. Since we only have two variables  $s$  and  $\phi$ , we can integrate the constraint equation Eq. (2.150),

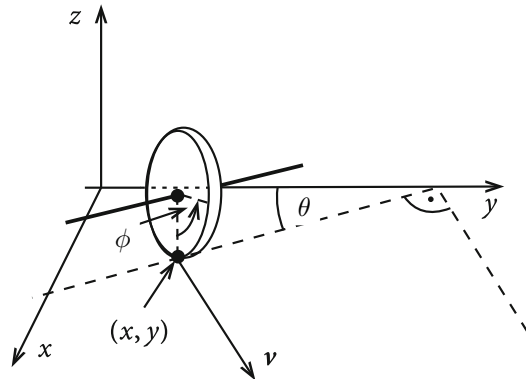
$$ds = R d\phi \quad \Rightarrow \quad s = R\phi + s_0, \quad (2.160)$$

and use it to eliminate the  $s$  or  $\phi$  in favor of the other coordinate. In the next example, we will also consider rolling without slipping, but we will *not* be able to integrate the nonholonomic constraint.

### Disk Rolling on a Plane

We consider a disk that is rolling without slipping in a horizontal plane while remaining in an upright position, so that the rotational axis remains parallel to the plane (see figure). As generalized coordinates, we can choose

- the  $x$  and  $y$  coordinates of the disk's center of mass, which also correspond to the support point in the plane,
- the angle  $\theta$  between the rotational axis and the  $x$  axis, and
- the angle  $\phi$  characterizing the disk's rotation around its axis.



**Figure 2.8:** Disk rolling without slipping on a plane.

The condition for rolling without slipping relates the change of the center of mass's position to the change in the angle  $\phi$  due to the rotation, just like in the previous example:

$$ds = R d\phi, \quad (2.161)$$

or in terms of velocities

$$\dot{s} = |\dot{\vec{v}}| = R\dot{\phi}. \quad (2.162)$$

We also constrain  $\vec{v}$  to be perpendicular to the rotational axis, which implies

$$\dot{x} = \dot{s} \cos \theta, \quad (2.163)$$

$$\dot{y} = \dot{s} \sin \theta. \quad (2.164)$$

The directional constraints can be combined with the rolling condition to yield

$$g_1(\vec{q}, \dot{\vec{q}}) \equiv \dot{x} - R\dot{\phi} \cos \theta = 0, \quad (2.165)$$

$$g_2(\vec{q}, \dot{\vec{q}}) \equiv \dot{y} - R\dot{\phi} \sin \theta = 0. \quad (2.166)$$

The Lagrangian for the disk's unconstrained motion is the sum of the translational and rotational terms (see Chapter 5),

$$L = \frac{1}{2}M(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I_\phi\dot{\phi}^2 + \frac{1}{2}I_\theta\dot{\theta}^2, \quad (2.167)$$

where  $I_\phi$  is the moment of inertia for the rotation around the disk's horizontal axis, and  $I_\theta$  the moment of inertia for rotation around the vertical axis through the disk's center of mass and support point in the plane.

Using the Lagrangian (2.167) and the constraints (2.165), (2.166), we obtain the following Lagrange equations of the first kind:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = \sum_{b=1}^2 \mu_b \frac{\partial g_b}{\partial \dot{x}} \quad \Rightarrow \quad M\ddot{x} = \mu_1, \quad (2.168)$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{y}} - \frac{\partial L}{\partial y} = \sum_{b=1}^2 \mu_b \frac{\partial g_b}{\partial \dot{y}} \quad \Rightarrow \quad M\ddot{y} = \mu_2, \quad (2.169)$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} = \sum_{b=1}^2 \mu_b \frac{\partial g_b}{\partial \dot{\phi}} \quad \Rightarrow \quad I_\phi\ddot{\phi} = -\mu_1 R \cos \theta - \mu_2 R \sin \theta, \quad (2.170)$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = \sum_{b=1}^2 \mu_b \frac{\partial g_b}{\partial \dot{\theta}} \quad \Rightarrow \quad I_\theta\ddot{\theta} = 0. \quad (2.171)$$

Together with Eqs. (2.165) and (2.166) we now have six equations for six unknowns. From the equation of motion for  $\theta$ , we obtain

$$\theta(t) = \omega t + \theta_0 = \omega t, \quad (2.172)$$

where we have assumed the initial condition  $\theta_0 = 0$  for simplicity.

Next, we differentiate the constraint equations again with respect to time, obtaining

$$\ddot{x} = R\ddot{\phi} \cos \omega t - R\omega\dot{\phi} \sin \omega t, \quad (2.173)$$

$$\ddot{y} = R\ddot{\phi} \sin \omega t + R\omega\dot{\phi} \cos \omega t. \quad (2.174)$$

These expressions can be used in the equations of motion for  $x$  and  $y$  to determine the Lagrange multipliers,

$$\mu_1 = MR \left( \ddot{\phi} \cos \omega t - \omega\dot{\phi} \sin \omega t \right), \quad (2.175)$$

2.4. CONSTRAINTS REVISITED: LAGRANGE EQUATIONS OF THE FIRST AND SECOND KIND 27

$$\mu_2 = MR \left( \ddot{\phi} \sin \omega t + \omega \dot{\phi} \cos \omega t \right), \quad (2.176)$$

and plugging these expressions into the equation of motion for  $\phi$ , we finally obtain

$$\begin{aligned} I_\phi \ddot{\phi} &= -MR^2 \cos \omega t \left( \ddot{\phi} \cos \omega t - \omega \dot{\phi} \sin \omega t \right) - MR^2 \sin \omega t \left( \ddot{\phi} \sin \omega t + \omega \dot{\phi} \cos \omega t \right) \\ &= -MR^2 \ddot{\phi} \end{aligned} \quad (2.177)$$

i.e.,

$$(I_\phi + MR^2) \ddot{\phi} = 0 \quad \Rightarrow \quad \dot{\phi}_0 = \text{const.}, \quad (2.178)$$

regardless of the moment of inertia of the disk. We still have to integrate the equations of motion for  $x, y$ , which read

$$\ddot{x} = -R\omega \dot{\phi}_0 \sin \omega t, \quad (2.179)$$

$$\ddot{y} = R\omega \dot{\phi}_0 \cos \omega t. \quad (2.180)$$

This means

$$\dot{x} = R\dot{\phi}_0 \cos \omega t + \dot{x}_0, \quad (2.181)$$

$$\dot{y} = R\dot{\phi}_0 \sin \omega t + \dot{y}_0, \quad (2.182)$$

and integrating once more, we have

$$x(t) = \frac{R}{\omega} \dot{\phi}_0 \sin \omega t + \dot{x}_0 t + x_0, \quad (2.183)$$

$$y(t) = -\frac{R}{\omega} \dot{\phi}_0 \cos \omega t + \dot{y}_0 t + y_0, \quad (2.184)$$

We can even determine the constraint forces which ensure that the disk rolls without slipping and remains upright.

$$\mu_1 = -MR\omega \dot{\phi} \sin \omega t, \quad (2.185)$$

$$\mu_2 = MR\omega \dot{\phi} \cos \omega t, \quad (2.186)$$

Finally, let us consider that the disk rolls in a straight line, so that  $\omega$ , the rotation around the disk's vertical axis, vanishes. Then

$$x(t) = \dot{x}_0 t + x_0, \quad (2.187)$$

$$y(t) = \dot{y}_0 t + y_0, \quad (2.188)$$

i.e., uniform linear motion of the center-of-mass with fixed velocity, and

$$\mu_1 = \mu_2 = 0. \quad (2.189)$$

## 2.5 Symmetries and Invariances

### 2.5.1 Cyclic Coordinates and Noether's Theorem

As we have seen in several applications of the Lagrange formalism, the structure of the Lagrange equations implies the existence of conserved quantities whenever our system's Lagrangian depends on a generalized velocity  $\dot{q}_i$  but not on the associated  $q_i$ :

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \underbrace{\frac{\partial L}{\partial q_i}}_{=0} = 0 \quad \Rightarrow \quad \frac{\partial L}{\partial \dot{q}_i} = \text{const.} \quad (2.190)$$

We refer to such  $q_i$ 's as **cyclic coordinates**.

The existence of cyclic coordinates is fundamentally related to invariances of the Lagrangian under transformations of the generalized coordinates, as first proven by Emmy Noether<sup>5</sup>. To see this, we introduce a general transformation of the coordinates that defines a set of trajectories labeled by the parameter  $\alpha$ ,

$$q_i(t) \longrightarrow q_i(t, \alpha), \quad (2.191)$$

and demand that the Lagrangian remains invariant up to a total time derivative:

$$L(q_i(t, \alpha), \dot{q}_i(t, \alpha), t) = L(q_i(t, 0), \dot{q}_i(t, 0), t) + \frac{dF(q_i(t, \alpha), t)}{dt}. \quad (2.192)$$

In this case, we speak of a symmetry of the Lagrangian under the transformation (2.191), which in turn implies a symmetry of the action for holonomic systems (see Sec. 2.3.3), or more generally a symmetry of d'Alembert's principle. The Lagrange equations are therefore invariant as well, and

$$\left( \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \frac{dq_i}{d\alpha} - \frac{dF}{d\alpha} \right) \Big|_{\alpha=0} = \text{const.} \quad (2.193)$$

is a **conserved quantity**. If no total time derivative appears under the symmetry transformation (2.191), we obtain a conserved quantity with the simpler form

$$\left( \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \frac{dq_i}{d\alpha} \right) \Big|_{\alpha=0} = \text{const.} \quad (2.194)$$

The appearance of conserved quantities associated with the symmetries of a Lagrangian is the central statement of **Noether's theorem**, which is of fundamental importance in many domains of physics.

The proof of Noether's theorem is straightforward: The Lagrangian is supposed to be invariant under changes of  $\alpha$  up to a total time derivative, so Eq. (2.192) implies that

$$\frac{d}{d\alpha} L(q(t, \alpha), \dot{q}(t, \alpha), t) \Big|_{\alpha=0} = \frac{d^2 F}{d\alpha dt} \Big|_{\alpha=0} = \frac{d\dot{F}}{d\alpha} \Big|_{\alpha=0}. \quad (2.195)$$

<sup>5</sup>A translated version is available as Ref. [8, 9], or as an updated preprint in arXiv:physics/0503066.



Considering the left-hand side, we have

$$\begin{aligned}
& \left. \frac{d}{d\alpha} L(q_i(t, \alpha), \dot{q}_i(t, \alpha), t) \right|_{\alpha=0} \\
&= \sum_{i=1}^n \left( \frac{\partial L}{\partial q_i} \frac{dq_i}{d\alpha} + \frac{\partial L}{\partial \dot{q}_i} \frac{d\dot{q}_i}{d\alpha} \right) \Big|_{\alpha=0} \\
&= \sum_{i=1}^n \left( \frac{\partial L}{\partial q_i} \frac{dq_i}{d\alpha} + \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \frac{dq_i}{d\alpha} \right) - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) \frac{dq_i}{d\alpha} \right) \Big|_{\alpha=0} \\
&= \sum_{i=1}^n \underbrace{\left( \frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right)}_{=0} \frac{dq_i}{d\alpha} \Big|_{\alpha=0} + \frac{d}{dt} \left( \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \frac{dq_i}{d\alpha} \right) \Big|_{\alpha=0} \tag{2.196}
\end{aligned}$$

For the term on the right-hand side, we recall that  $F(q_i(t, \alpha), t)$  does not depend on  $\dot{q}(t, \alpha)$ , hence its time derivative can only have a linear dependence on the generalized velocities,

$$\dot{F}(q_i(t, \alpha), \dot{q}_i(t, \alpha), t) = \sum_i \left( \frac{\partial F}{\partial q_i} \dot{q}_i + \underbrace{\frac{\partial F}{\partial \dot{q}_i}}_{=0} \ddot{q}_i \right) + \frac{\partial F}{\partial t}, \tag{2.197}$$

and the ‘‘cancellation of dots’’ holds:

$$\frac{\partial \dot{F}}{\partial \dot{q}_i} = \frac{\partial F}{\partial q_i}. \tag{2.198}$$

Applying this to Eq. (2.195), we have

$$\begin{aligned}
\left. \frac{d}{d\alpha} \dot{F}(q_i(t, \alpha), \dot{q}_i(t, \alpha), t) \right|_{\alpha=0} &= \sum_{i=1}^n \left( \frac{\partial \dot{F}}{\partial q_i} \frac{dq_i}{d\alpha} + \frac{\partial \dot{F}}{\partial \dot{q}_i} \frac{d\dot{q}_i}{d\alpha} \right) \Big|_{\alpha=0} \\
&= \sum_{i=1}^n \left( \frac{\partial \dot{F}}{\partial q_i} \frac{dq_i}{d\alpha} + \frac{d}{dt} \left( \frac{\partial \dot{F}}{\partial \dot{q}_i} \frac{dq_i}{d\alpha} \right) - \left( \frac{d}{dt} \frac{\partial \dot{F}}{\partial \dot{q}_i} \right) \frac{dq_i}{d\alpha} \right) \Big|_{\alpha=0} \\
&= \sum_{i=1}^n \left( \frac{d}{dt} \left( \frac{\partial \dot{F}}{\partial \dot{q}_i} \frac{dq_i}{d\alpha} \right) \right) \Big|_{\alpha=0} \\
&= \frac{d}{dt} \left( \sum_{i=1}^n \frac{\partial F}{\partial q_i} \frac{dq_i}{d\alpha} \right) \Big|_{\alpha=0}, \tag{2.199}
\end{aligned}$$

where we have used that  $\dot{F}$  satisfies the Lagrange equations. Rearranging Eq. (2.195), and using our results, we see that

$$\begin{aligned}
0 &= \frac{d}{d\alpha} \left( L(q_i(t, \alpha), \dot{q}_i(t, \alpha), t) - \dot{F}(q_i(t, \alpha), \dot{q}_i(t, \alpha), t) \right) \Big|_{\alpha=0} \\
&= \frac{d}{dt} \left( \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \frac{dq_i}{d\alpha} - \sum_{i=1}^n \frac{\partial F}{\partial q_i} \frac{dq_i}{d\alpha} \right) \Big|_{\alpha=0} \\
&= \frac{d}{dt} \left( \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \frac{dq_i}{d\alpha} - \frac{dF}{d\alpha} \right) \Big|_{\alpha=0}, \tag{2.200}
\end{aligned}$$

which completes the proof. Let us now use Noether’s theorem to study various invariances.

### 2.5.2 Spatial Translations

Consider a Lagrangian that is invariant under a translation of the coordinate system i.e., a change of coordinates

$$\vec{r}_i(t) \longrightarrow \vec{r}_i(t, \alpha) = \vec{r}_i(t) + \alpha \vec{e}, \quad (2.201)$$

where  $\alpha$  is time independent and  $\vec{e}$  is a constant unit vector in  $\mathbb{R}^3$ . For simplicity, we consider an  $N$ -particle system without constraints, but the conclusions are readily generalized to holonomic systems.

The kinetic energy is obviously invariant under translation because  $\dot{\vec{r}}_i(t, \alpha) = \dot{\vec{r}}_i$  (also see Sec. 2.5.5), and the potential energy is invariant if it only depends on the particles' relative coordinates  $\vec{r}_{ij} \equiv \vec{r}_i - \vec{r}_j$ . Thus, a translationally invariant Lagrangian can be written as

$$L(\vec{r}_i, \dot{\vec{r}}_i) = \frac{1}{2} \sum_{i=1}^N m_i \dot{\vec{r}}_i^2 - V(\vec{r}_1 - \vec{r}_2, \dots, \vec{r}_i - \vec{r}_j, \dots), \quad (2.202)$$

We have

$$\left. \frac{\partial \vec{r}_i}{\partial \alpha} \right|_{\alpha=0} = \vec{e} \quad (2.203)$$

for *each particle* and therefore Noether's theorem (2.194) implies that

$$\text{const.} = \sum_{i=1}^N \sum_{k=1}^3 \frac{\partial L}{\partial \dot{x}_{ik}} \left. \frac{\partial x_{ik}}{\partial \alpha} \right|_{\alpha=0} = \sum_{k=1}^3 \underbrace{\sum_{i=1}^N \frac{\partial L}{\partial \dot{x}_{ik}}}_{\equiv P_k} e_k = \vec{P} \cdot \vec{e}, \quad (2.204)$$

i.e., the component of the total momentum  $\vec{P}$  in the direction of  $\vec{e}$  is conserved (think of the block sliding down a wedge discussed in Sec. 2.2.3, where we found conservation of the total momentum in  $x$  direction).

If the Lagrangian is invariant for an  $\vec{e}$  *arbitrary* unit vector in  $\mathbb{R}^3$ , we have a **spatially homogeneous** system in which no single point is preferred. In that case, each component of  $\vec{P}$  must be conserved, and we obtain three **constants of motion** (see Box 2.6). Thus, we obtain a deep connection between the fundamental structure of space and a conservation law — we will come back to this in Sec. 2.5.5.

### Spatial Rotations

Next, we consider a Lagrangian that remains invariant under rotations by an angle  $\alpha$  around a spatially fixed axis  $\vec{e} = \vec{\omega}/\omega$ . We can express such a rotation in vectorial fashion as

$$\vec{r}_i(t) \longrightarrow \vec{r}_i(t, \alpha) = \vec{r}_i(t) \cos \alpha + \vec{e}(\vec{e} \cdot \vec{r}_i)(1 - \cos \alpha) + (\vec{e} \times \vec{r}_i) \sin \alpha. \quad (2.205)$$

To see this, we choose a spherical coordinate system such that the  $z$  axis is aligned with the rotational axis,  $\vec{e} = \vec{e}$ . The rotation by an angle  $\alpha$  around this axis can then be expressed in matrix form as

$$\vec{r}(\alpha) = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \cos \alpha - y \sin \alpha \\ x \sin \alpha + y \cos \alpha \\ z \end{pmatrix}. \quad (2.206)$$

With  $\vec{e} = \vec{e}_z$ , Eq. (2.205) becomes

$$\vec{r}(\alpha) = \begin{pmatrix} x \cos \alpha \\ y \cos \alpha \\ z \cos \alpha \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ z(1 - \cos \alpha) \end{pmatrix} + \begin{pmatrix} -y \sin \alpha \\ x \sin \alpha \\ 0 \end{pmatrix} = \begin{pmatrix} x \cos \alpha - y \sin \alpha \\ x \sin \alpha + y \cos \alpha \\ z \end{pmatrix}. \quad (2.207)$$

Noting that

$$\left. \frac{\partial \vec{r}_i(t, \alpha)}{\partial \alpha} \right|_{\alpha=0} = \vec{e} \times \vec{r}_i(t), \quad (2.208)$$

Noether's theorem (2.194) yields

$$\text{const.} = \sum_{i=1}^N m_i \dot{\vec{r}}_i \cdot (\vec{e} \times \vec{r}_i) = \vec{e} \cdot \underbrace{\sum_{i=1}^N m_i (\vec{r}_i \times \dot{\vec{r}}_i)}_{\equiv \vec{L}} = \vec{e} \cdot \vec{L}, \quad (2.209)$$

where we have introduced the total angular momentum  $\vec{L}$  and used

$$\vec{a} \cdot (\vec{b} \times \vec{c}) = \vec{b} \cdot (\vec{c} \times \vec{a}) = \vec{c} \cdot (\vec{a} \times \vec{b}). \quad (2.210)$$

We see that the total angular momentum along the rotational axis  $\vec{e}$  is conserved.

If the Lagrangian is invariant under rotations around an *arbitrary* axis  $\vec{e}$  in  $\mathbb{R}^3$ , no direction is preferred and we call a system **spatially isotropic**. An example would be a Lagrangian containing a potential that only depends on the relative *distances*  $|\vec{r}_i - \vec{r}_j|$  of the particles. In this case all three components of the total angular momentum  $\vec{L}$  are conserved, and we have three constants of motion (see Box 2.6).

### 2.5.3 Galilean Boosts

According to the special principle of relativity, the laws of physics must be the same in any inertial frame, i.e., any frame moving with a constant velocity with respect to the observer's frame. In non-relativistic mechanics, transformations between such coordinate frames are referred to as **Galilean boosts**, and they have the form

$$\vec{r}_i(t, \alpha) = \vec{r}_i(t) + \alpha \vec{u}_0 t \quad (2.211)$$

with a fixed velocity  $\vec{u}_0$ .

A potential of the form  $V(\vec{r}_1 - \vec{r}_2, \dots, \vec{r}_i - \vec{r}_j, \dots)$  will be invariant under Galilean boosts. The kinetic energy, however, is *not* invariant under such transformations, because

$$T(\alpha = 0) = \sum_{i=1}^N \frac{1}{2} m_i \dot{\vec{r}}_i^2 \quad \longrightarrow \quad T(\alpha) = \sum_{i=1}^N \frac{1}{2} m_i (\dot{\vec{r}}_i + \alpha \vec{u}_0)^2. \quad (2.212)$$

We show that the kinetic energy difference between the two frames is a total time derivative, and therefore Noether's theorem in the general form (2.193) can be applied.

$$L(\vec{r}_i(t, \alpha), \dot{\vec{r}}_i(t, \alpha), t)$$

$$\begin{aligned}
&= \frac{1}{2} \sum_{i=1}^N m_i (\dot{\vec{r}}_i + \alpha \vec{u}_0)^2 - V(\vec{r}_1 + \alpha \vec{u}_0 t - \vec{r}_2 - \alpha \vec{u}_0 t, \dots, \vec{r}_{N-1} + \alpha \vec{u}_0 t - \vec{r}_N - \alpha \vec{u}_0 t) \\
&= \frac{1}{2} \sum_{i=1}^N m_i \dot{\vec{r}}_i^2 + \sum_{i=1}^N m_i \left( \alpha \dot{\vec{r}}_i \cdot \vec{u}_0 + \frac{1}{2} \alpha^2 \vec{u}_0^2 \right) - V(\vec{r}_1 - \vec{r}_2, \dots, \vec{r}_{N-1} - \vec{r}_N) \\
&= L(\vec{r}_i, \dot{\vec{r}}_i, t) + \underbrace{\frac{d}{dt} \left( \sum_{i=1}^N m_i \left( \alpha \vec{u}_0 \cdot \vec{r}_i + \frac{1}{2} \alpha^2 \vec{u}_0^2 t \right) \right)}_{\equiv F}. \tag{2.213}
\end{aligned}$$

Thus, Eq. (2.193) implies

$$\begin{aligned}
\left( \sum_{j=1}^n \frac{\partial L}{\partial \dot{q}_j} \frac{dq_j}{d\alpha} - \frac{dF}{d\alpha} \right) \Big|_{\alpha=0} &= \sum_{i=1}^N m_i \dot{\vec{r}}_i \cdot \vec{u}_0 t - \sum_{i=1}^N m_i \vec{r}_i \cdot \vec{u}_0 \\
&= \vec{u}_0 (\vec{P}t - M\vec{R}) = \text{const.}, \tag{2.214}
\end{aligned}$$

where we have introduced the center of mass position vector

$$\vec{R} = \frac{1}{M} \sum_{i=1}^N m_i \vec{r}_i. \tag{2.215}$$

For arbitrary  $\vec{u}_0$ , we obtain three constants of motion:

$$\vec{P}t - M\vec{R} = \text{const.} \tag{2.216}$$

(see Box 2.6). If we write the constant as  $-M\vec{R}_0$ , we can solve for  $\vec{R}$  and find

$$\vec{R}(t) = \frac{1}{M} \vec{P}t + \vec{R}_0, \tag{2.217}$$

which is the trajectory of the center of mass undergoing uniform linear motion. If the direction of  $\vec{u}_0$  is fixed, we only obtain one constant of motion, and the motion of the center of mass is only uniform along the direction of  $\vec{u}_0$ .

#### 2.5.4 Translations in Time

A system whose properties are invariant under temporal translations

$$t \longrightarrow t + \alpha \tag{2.218}$$

is called **homogeneous in time**. This means that the results of any measurement are independent of the specific time at which it is conducted. Since time is the variable that parameterizes the trajectories of the generalized coordinates (and velocities) and not a generalized coordinate itself, we cannot directly apply Noether's theorem in either form, but most follow a slightly different approach.

Consider

$$\frac{d}{dt} L(q_i, \dot{q}_i, t) = \sum_{i=1}^n \left( \frac{\partial L}{\partial q_i} \dot{q}_i + \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i \right) + \frac{\partial L}{\partial t}$$

$$\begin{aligned}
&= \sum_{i=1}^n \left( \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) \dot{q}_i + \frac{\partial L}{\partial \dot{q}_i} \frac{d}{dt} \dot{q}_i \right) + \frac{\partial L}{\partial t} \\
&= \frac{d}{dt} \left( \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) + \frac{\partial L}{\partial t}, \tag{2.219}
\end{aligned}$$

where we have used the Lagrange equations in the second line. Collecting the total time derivatives on the left-hand side, we have

$$\frac{d}{dt} \left( \sum_{i=1}^n \left( \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) - L \right) = - \frac{\partial L}{\partial t}. \tag{2.220}$$

If the Lagrangian does not explicitly depend on time, i.e.,  $\frac{\partial L}{\partial t} = 0$ , we obtain another constant of motion<sup>6</sup>:

$$\bar{H}(q_i, \dot{q}_i) = \sum_{i=1}^n \left( \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) - L = \text{const.}, \tag{2.221}$$

This is the **Hamiltonian** of the mechanical system — or rather, it will be after we make a change from the generalized coordinates and velocities to the proper variables, as discussed in Chapter 6. We will disregard this subtle distinction for now.

To understand the Hamiltonian's physical meaning, we need to analyze the terms  $\frac{\partial L}{\partial \dot{q}_i} \dot{q}_i$  in Eq. (2.221). For now, we restrict the discussion to systems without velocity-dependent potentials or dissipation — we will reconsider the Hamiltonian for such systems in Sections 2.6 and 2.7, respectively.

### Scleronomic Constraints

In a holonomic system with scleronomic constraints, we have  $\vec{r} = \vec{r}(q_1, \dots, q_n)$ , which implies

$$\dot{\vec{r}}_i = \sum_{j=1}^n \frac{\partial \vec{r}_i}{\partial q_j} \dot{q}_j \tag{2.222}$$

and

$$T = \frac{1}{2} \sum_{i=1}^N m_i \dot{\vec{r}}_i \cdot \dot{\vec{r}}_i = \frac{1}{2} \sum_{j,k=1}^n \left( \sum_{i=1}^N m_i \frac{\partial \vec{r}_i}{\partial q_j} \cdot \frac{\partial \vec{r}_i}{\partial q_k} \right) \dot{q}_j \dot{q}_k \equiv \frac{1}{2} \sum_{j,k=1}^n M_{jk} \dot{q}_j \dot{q}_k, \tag{2.223}$$

where we have introduced the **mass tensor**  $\mathbf{M}$ . The kinetic energy's partial derivative with respect to  $\dot{q}_j$  is given by

$$\begin{aligned}
\frac{\partial T}{\partial \dot{q}_j} &= \frac{1}{2} \sum_{kl} M_{kl} \frac{\partial}{\partial \dot{q}_j} (\dot{q}_k \dot{q}_l) = \frac{1}{2} \sum_{kl} M_{kl} \left( \frac{\partial \dot{q}_k}{\partial \dot{q}_j} \dot{q}_l + \dot{q}_k \frac{\partial \dot{q}_l}{\partial \dot{q}_j} \right) \\
&= \frac{1}{2} \sum_{kl} M_{kl} (\delta_{jk} \dot{q}_l + \delta_{jl} \dot{q}_k) = \sum_k M_{jk} \dot{q}_k, \tag{2.224}
\end{aligned}$$

---

<sup>6</sup>Note that this is equivalent to the Beltrami identity, Eq. (2.113), which we used in our discussion of the brachistochrone problem.

### Box 2.5: Euler's Homogeneous Function Theorem

A homogeneous function  $F(x_1, \dots, x_n)$  of degree  $k$  has the property

$$F(\lambda x_1, \dots, \lambda x_n) = \lambda^k F(x_1, \dots, x_n), \quad (\text{B2.5-1})$$

which means that

$$\sum_{i=1}^n x_i \frac{\partial F}{\partial x_i} = kF, \quad (\text{B2.5-2})$$

Thus, for holonomic systems with scleronomic constraints, the kinetic energy is a homogeneous function of degree 2 in the generalized velocities.

where we have used the symmetry of the mass tensor ( $M_{jk} = M_{kj}$ ) and the freedom to rename summation variables. This means that

$$\sum_j \frac{\partial T}{\partial \dot{q}_j} \dot{q}_j = \sum_{jk} M_{jk} \dot{q}_j \dot{q}_k = 2T, \quad (2.225)$$

and since we only consider velocity-independent potentials, the Hamiltonian becomes

$$\bar{H}(q_j, \dot{q}_j) = \sum_{j=1}^n \frac{\partial L}{\partial \dot{q}_j} \dot{q}_j - L = \sum_{j=1}^n \frac{\partial T}{\partial \dot{q}_j} \dot{q}_j - L = 2T - (T - V) = T + V. \quad (2.226)$$

Thus, invariance with respect to translations in time implies the **conservation of energy** in holonomic systems with scleronomic constraints, and the Hamiltonian is the *total energy* of such a system (see Box 2.6).

### Rheonomic Systems

Let us now consider holonomic systems with rheonomic constraints. In this case, we have  $\vec{r} = \vec{r}(\mathbf{q}, t)$ , and

$$\dot{\vec{r}}_i = \sum_{j=1}^n \frac{\partial \vec{r}_i}{\partial q_j} \dot{q}_j + \frac{\partial \vec{r}_i}{\partial t}. \quad (2.227)$$

The kinetic energy now becomes

$$T = \frac{1}{2} \sum_{i=1}^N m_i \left( \sum_{k=1}^n \frac{\partial \vec{r}_i}{\partial q_k} \dot{q}_k + \frac{\partial \vec{r}_i}{\partial t} \right) \cdot \left( \sum_{l=1}^n \frac{\partial \vec{r}_i}{\partial q_l} \dot{q}_l + \frac{\partial \vec{r}_i}{\partial t} \right). \quad (2.228)$$

Computing the first term in the Hamiltonian, we find

$$\sum_{j=1}^n \frac{\partial L}{\partial \dot{q}_j} \dot{q}_j = \sum_{j=1}^n \frac{\partial T}{\partial \dot{q}_j} \dot{q}_j = \sum_{i=1}^N \sum_{j,k=1}^n m_i \underbrace{\left( \frac{\partial \vec{r}_i}{\partial q_j} \frac{\partial \dot{q}_k}{\partial \dot{q}_j} \right)}_{=\delta_{jk}} \cdot \underbrace{\left( \sum_{l=1}^n \frac{\partial \vec{r}_i}{\partial q_l} \dot{q}_l + \frac{\partial \vec{r}_i}{\partial t} \right)}_{=\dot{\vec{r}}_i} \dot{q}_j$$

$$\begin{aligned}
&= \sum_{i=1}^N m_i \left( \sum_{j=1}^n \frac{\partial \vec{r}_i}{\partial q_j} \dot{q}_j \right) \cdot \dot{\vec{r}}_i = \sum_{i=1}^N m_i \underbrace{\left( \sum_{j=1}^n \frac{\partial \vec{r}_i}{\partial q_j} \dot{q}_j + \frac{\partial \vec{r}_i}{\partial t} \right)}_{=\dot{\vec{r}}_i} \cdot \dot{\vec{r}}_i - \sum_{i=1}^N m_i \frac{\partial \vec{r}_i}{\partial t} \cdot \dot{\vec{r}}_i \\
&= 2T - \sum_{i=1}^N m_i \frac{\partial \vec{r}_i}{\partial t} \cdot \dot{\vec{r}}_i, \tag{2.229}
\end{aligned}$$

and in total

$$\bar{H}(\mathbf{q}, \dot{\mathbf{q}}) = \sum_{j=1}^n \frac{\partial L}{\partial \dot{q}_j} \dot{q}_j - L = T + V - \sum_{i=1}^N m_i \dot{\vec{r}}_i \cdot \frac{\partial \vec{r}_i}{\partial t}. \tag{2.230}$$

Thus, the Hamiltonian does *not* correspond to the total energy in this case unless the third term vanishes. This term is the projection of the momentum on the tangent vector  $\frac{\partial \vec{r}_i}{\partial t}$ , which results from the time dependence of the constraint in the rheonomic case (compare Eqs. (2.222) and (2.227)). Thus, if the change of the constraint is orthogonal to the direction of motion,  $\bar{H}$  would still be the total energy.

### Example: Bead on a Rotating Wire

As an example for a rheonomic system, we consider a bead on a rotating wire, as previously discussed in Sec. 2.1 (see Fig. 2.3). For simplicity, we choose  $\alpha = \pi/2$ , so that the coordinates are

$$x = q \cos \omega t, \quad y = q \sin \omega t, \quad z = 0, \tag{2.231}$$

and the Lagrangian becomes

$$\begin{aligned}
L &= \frac{1}{2} m (\dot{x}^2 + \dot{y}^2) \\
&= \frac{1}{2} m (\dot{q}^2 (\cos^2 \omega t + \sin^2 \omega t) + q^2 \omega^2 (\sin^2 \omega t + \cos^2 \omega t)) \\
&= \frac{1}{2} m (\dot{q}^2 + \omega^2 q^2). \tag{2.232}
\end{aligned}$$

Now

$$\frac{\partial L}{\partial \dot{q}} = m \dot{q}, \quad \frac{\partial L}{\partial q} = m q \omega^2, \tag{2.233}$$

and the Lagrange equation yields the equation of motion

$$\ddot{q} - \omega^2 q = 0. \tag{2.234}$$

The Hamiltonian becomes

$$\bar{H}(q, \dot{q}) = \frac{\partial L}{\partial \dot{q}} \dot{q} - L(q, \dot{q}) = \frac{1}{2} m (\dot{q}^2 - q^2 \omega^2), \tag{2.235}$$

which is conserved because

$$\frac{d\bar{H}}{dt} = -\frac{\partial L}{\partial t} = 0 \tag{2.236}$$

**Box 2.6: Constants of Motion**

For a closed system of  $N$  (non-relativistic) particles that only interact through conservative forces which depend on  $\vec{r}_i - \vec{r}_j$ , Noether's theorem implies the existence of **10 constants of motion**:

Symmetry Transformation		Conserved Quantity	
temporal translation	$t' = t + \tau$	$E = \text{const.}$	total energy
spatial translation	$\vec{r}' = \vec{r} + \vec{a}$	$\vec{P} = \text{const.}$	total momentum
rotation	$\vec{r}' = R\vec{r}$	$\vec{L} = \text{const.}$	total angular momentum
Galilean boost	$\vec{r}' = \vec{r} + \vec{u}t$	$M\vec{R} - \vec{P}t = \text{const.}$	

(see Eq. (2.220)). However, it is *not* the total energy, which we can see by considering the relation (2.230). Since  $V = 0$ , the total energy is identical to the kinetic energy,

$$T = \frac{1}{2}m(\dot{q}^2 + \omega^2 q^2). \quad (2.237)$$

The momentum is

$$m\dot{\vec{r}} = m(\dot{q} \cos \omega t - q\omega \sin \omega t, \dot{q} \sin \omega t + q\omega \cos \omega t, 0)^T \quad (2.238)$$

and the explicit time dependence of the coordinate due to the rheonomic constraint imposed by the rotating wire yields

$$\frac{\partial \vec{r}}{\partial t} = (-q\omega \sin \omega t, q\omega \cos \omega t, 0)^T, \quad (2.239)$$

so we obtain

$$m\dot{\vec{r}} \cdot \frac{\partial \vec{r}}{\partial t} = m\dot{q}^2 \omega^2. \quad (2.240)$$

The time derivative of the total energy is given by

$$\frac{dE}{dt} = \frac{d(T + V)}{dt} = \frac{d}{dt} \left( \frac{1}{2}m(\dot{q}^2 + \omega^2 q^2) \right) = m\dot{q}(\ddot{q} + \omega^2 q), \quad (2.241)$$

and plugging in Eq. (2.234), we obtain

$$\frac{dE}{dt} = \frac{d}{dt} m\omega^2 q^2 = 2m\omega^2 q\dot{q}. \quad (2.242)$$

This is the change in the energy the motor needs to provide to keep the wire spinning at a constant angular velocity as the pearl slides to different positions  $q$  along the wire.

### 2.5.5 Deriving Lagrangians from Symmetries

We have seen that the laws of classical mechanics can be cast in the form of d'Alembert's principle or (with certain limitations) the form of the principle of least action. However, nothing in these principles enforces a particular shape of the Lagrangian, aside perhaps from the idea that the action should be *minimal* in the latter case.



The discussion in the present section has shed light on the fundamental relationship between symmetries and invariances of the Lagrangian. Moreover, we have seen that we had to *impose* certain properties on the potentials appearing in the Lagrangian to produce a desired invariance — or, in other words, to make the potentials compatible with the fundamental symmetries of space. This is a very powerful idea that we want to explore in a bit more detail.

Consider the Lagrangian governing the dynamics of a free particle, which should be a function of the particle's position, velocity, and possibly time,

$$L_0 = L_0(\vec{r}, \dot{\vec{r}}, t). \quad (2.243)$$

First, we make use of the **homogeneity of space**, which implies that the dynamics of a particle cannot depend on the choice of coordinate system that we use to describe its motion. In fact, this means that  $L_0$  cannot explicitly depend on  $\vec{r}$  at all. Likewise, **time is homogeneous** if we do not have external forces, so the free particle's dynamics cannot depend explicitly on the time we make an observation either. Thus, we must have

$$L_0 = L_0(\dot{\vec{r}}). \quad (2.244)$$

Next, we use the **isotropy of space**, which means that there is no preferred direction and the dynamics of the particle cannot depend on the orientation of the coordinate system we use to describe it. Thus, the Lagrangian can only be a function of the magnitude of the velocity vector,

$$\dot{\vec{r}} \cdot \dot{\vec{r}} = |\dot{\vec{r}}|^2 = v^2, \quad (2.245)$$

i.e.,

$$L_0 = L_0(v^2). \quad (2.246)$$

Finally, we consider the principle of (Newtonian) relativity, which implies the invariance of the Lagrangian under Galilean boosts, so that observers in different inertial systems will derive the same equations of motion for the particle. If the particle is observed from a coordinate system that is moving with a small relative velocity  $\vec{u}$  with respect to ours, the Lagrangian will be

$$L'_0 = L_0((\vec{v} + \vec{u})^2) = L_0(v^2) + 2\vec{u} \cdot \vec{v} \frac{\partial L_0}{\partial v^2} + O(u^2) \quad (2.247)$$

Since the equations of motion for  $L'_0$  must be identical to those for  $L_0$ , the additional term must be a total time derivative. We can rewrite it as

$$2 \left( \frac{d\vec{r}}{dt} \right) \cdot \vec{u} \frac{\partial L_0}{\partial v^2} = \frac{d}{dt} \left( 2\vec{r} \cdot \vec{u} \frac{\partial L_0}{\partial v^2} \right) - 2\vec{r} \cdot \vec{u} \left( \frac{d}{dt} \frac{\partial L_0}{\partial v^2} \right). \quad (2.248)$$

The first term on the right-hand side is our total time derivative, so the second term must vanish. The direction of  $\vec{u}$  is arbitrary, so the scalar product  $\vec{r} \cdot \vec{u}$  is not vanishing in general, and we must have

$$\frac{d}{dt} \frac{\partial L_0}{\partial v^2} = 0 \quad \Rightarrow \quad \frac{\partial L_0}{\partial v^2} = \text{const.} \quad (2.249)$$

This actually means that the Lagrangian must be a scalar multiple of the square of the velocity, i.e.,

$$L_0(v^2) = cv^2, \quad (2.250)$$

and we have

$$L'_0(v^2) = c(v^2 + 2\vec{v} \cdot \vec{u} + u^2) = cv^2 + c \frac{d}{dt} (2\vec{r} \cdot \vec{u} + u^2t) . \quad (2.251)$$

[...]

## 2.6 Velocity-Dependent Potentials

When we derived the Lagrange equations from d'Alembert's principle, we first obtained them in the form

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = Q_j . \quad (2.252)$$

Assuming that the generalized forces can be written as the gradients of a potential  $V(\mathbf{q})$ ,

$$Q_j = \frac{\partial V}{\partial q_j} , \quad (2.253)$$

we moved them to the left-hand side, using that  $\frac{\partial V}{\partial \dot{q}_j} = 0$ :

$$\frac{d}{dt} \frac{\partial(T - V)}{\partial \dot{q}_j} - \frac{\partial(T - V)}{\partial q_j} = 0 . \quad (2.254)$$

It is easy to see that the Lagrange equations (i.e., d'Alembert's principle) would also be satisfied for more general forces of the form

$$Q_j = -\frac{\partial U}{\partial q_j} + \frac{d}{dt} \frac{\partial U}{\partial \dot{q}_j} , \quad (2.255)$$

where  $U(\mathbf{q}, \dot{\mathbf{q}})$  is a velocity-dependent potential. The definition of the Lagrangian simply becomes

$$L \equiv T - U , \quad (2.256)$$

since  $V(\mathbf{q})$  would be a special case of  $U$ .

### Example: Particle Moving in an Electromagnetic Field

Let us consider a particle with mass  $m$  and charge  $q$  that moves in an *external* electromagnetic field. We not impose any constraints, hence we can work in Cartesian coordinates:

$$(q_1, q_2, q_3) = (x_1, x_2, x_3) = (x, y, z) . \quad (2.257)$$

The Lagrangian of this particle is given by

$$L(\vec{r}, \dot{\vec{r}}, t) = \frac{1}{2} m \dot{\vec{r}}^2 - q\phi(\vec{r}, t) + q\vec{A}(\vec{r}, t) \cdot \dot{\vec{r}} , \quad (2.258)$$

with the explicitly velocity-dependent potential

$$U(\vec{r}, \dot{\vec{r}}) = q\phi - q\vec{A} \cdot \dot{\vec{r}} . \quad (2.259)$$

In this case, we have

$$\frac{\partial L}{\partial \dot{x}_i} = m\dot{x}_i + qA_i(\vec{r}, t) , \quad (2.260)$$

and

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} = m\ddot{x}_i + q \left( \sum_{k=1}^3 \frac{\partial A_i}{\partial x_k} \dot{x}_k + \frac{\partial A_i}{\partial t} \right) . \quad (2.261)$$

We also have

$$\frac{\partial L}{\partial x_i} = -q \frac{\partial \phi}{\partial x_i} + q \sum_{k=1}^3 \frac{\partial A_k}{\partial x_i} \dot{x}_k . \quad (2.262)$$

Combining the derivatives and switching to a vectorial form, we obtain the Lagrange equation

$$m \left( \ddot{\vec{r}} + q \left( (\dot{\vec{r}} \cdot \nabla) \vec{A} + \frac{\partial \vec{A}}{\partial t} \right) \right) = -q \nabla \phi + q \nabla (\vec{A} \cdot \dot{\vec{r}}) . \quad (2.263)$$

It is left as an exercise to show that this equation reduces to the usual equation of motion under the influence of the **Lorentz force**,

$$m\ddot{\vec{r}} = q(\vec{E} + \dot{\vec{r}} \times \vec{B}) , \quad (2.264)$$

where the electric and magnetic fields are defined as

$$\vec{E} = -\nabla \phi - \frac{\partial \vec{A}}{\partial t} \quad (2.265)$$

and

$$\vec{B} = \nabla \times \vec{A} . \quad (2.266)$$

Thus, the Lorentz force can indeed be derived from the velocity-dependent potential (2.259)

$$F_i = q(\vec{E} + \dot{\vec{r}} \times \vec{B})_i = -\frac{\partial U}{\partial x_i} + \frac{d}{dt} \frac{\partial U}{\partial \dot{x}_i} . \quad (2.267)$$

We conclude our discussion by computing the Hamiltonian  $\bar{H}(\vec{r}, \dot{\vec{r}})$ :

$$\begin{aligned} \bar{H}(\vec{r}, \dot{\vec{r}}, t) &= \left( m\dot{\vec{r}} + q\vec{A}(\vec{r}, t) \right) \cdot \dot{\vec{r}} - \frac{1}{2} m \dot{\vec{r}}^2 + q \left( \phi(\vec{r}, t) - \dot{\vec{r}} \cdot \vec{A}(\vec{r}, t) \right) \\ &= \frac{1}{2} m \dot{\vec{r}}^2 + q\phi(\vec{r}, t) . \end{aligned} \quad (2.268)$$

Since the fields can be time dependent,  $\bar{H}$  is generally *not conserved*, although it does represent the total energy. Changes of the external fields  $\phi(\vec{r}, t)$ ,  $\vec{A}(\vec{r}, t)$  require some form of work from the source that generates them, so energy is added to or removed from our system.

## 2.7 Dissipation

### 2.7.1 The Dissipation Function

In realistic mechanical systems, dissipative forces like dry or viscous friction will resist the *relative* motion of extended solids, surfaces, or fluid layers, causing a loss of mechanical energy. While the details of frictional mechanisms are usually microscopic in nature and beyond the scope of Classical Mechanics, a wide class of frictional phenomena can be modeled by forces of the form [10, 11]

$$\vec{F}^D = -\mu(v)\frac{\vec{v}}{v}, \quad (2.269)$$

pointing in the opposite direction of the relative velocity  $\vec{v}$  between the moving object and the environment. Here,  $\mu(v)$  is a positive function that could also depend on the coordinates. If the environment is static, we can identify

$$\vec{v} = \dot{\vec{r}}. \quad (2.270)$$

Since friction forces are nonconservative — i.e., the work done against these forces depends on the trajectory — they must be treated explicitly as generalized forces in the Lagrange equations:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = Q_j^D, \quad (2.271)$$

where we assumed that all *conservative* forces are included in the Lagrangian. The generalized dissipative forces are obtained from (cf. Sec. 2.2.2)

$$Q_j^D = \sum_{i=1}^N \vec{F}_i^D \cdot \frac{\partial \vec{r}_i}{\partial q_j} = \sum_{i=1}^N \vec{F}_i^D \cdot \frac{\partial \vec{v}_i}{\partial \dot{q}_j} = - \sum_{i=1}^N \mu_i(v_i) \frac{\vec{v}_i}{v_i} \cdot \frac{\partial \vec{v}_i}{\partial \dot{q}_j}, \quad (2.272)$$

where we have used the “cancellation of dots”,

$$\frac{\partial \vec{v}_i}{\partial \dot{q}_j} = \frac{\partial \vec{r}_i}{\partial q_j}. \quad (2.273)$$

Noticing that

$$\vec{v}_i \cdot \frac{\partial \vec{v}_i}{\partial \dot{q}_j} = \frac{1}{2} \frac{\partial}{\partial \dot{q}_j} (\vec{v}_i \cdot \vec{v}_i) = \frac{1}{2} \frac{\partial v_i^2}{\partial \dot{q}_j} = v_i \frac{\partial v_i}{\partial \dot{q}_j}, \quad (2.274)$$

we obtain

$$Q_j^D = - \sum_{i=1}^N \mu_i(v_i) \frac{\partial v_i}{\partial \dot{q}_j}. \quad (2.275)$$

We can rewrite this expression further: First, we note that

$$\mu_i(v_i) = \frac{\partial}{\partial v_i} \left( \int_0^{v_i} \mu_i(v') dv' \right), \quad (2.276)$$

so the chain rule implies

$$\mu_i(v_i) \frac{\partial v_i}{\partial \dot{q}_j} = \frac{\partial v_i}{\partial \dot{q}_j} \mu_i(v_i) = \frac{\partial}{\partial \dot{q}_j} \left( \int_0^{v_i} \mu_i(v') dv' \right), \quad (2.277)$$

i.e.,

$$Q_j^D = -\frac{\partial}{\partial \dot{q}_j} \sum_{i=1}^N \int_0^{v_i} \mu_i(v') dv'. \quad (2.278)$$

We see that instead of treating friction using the multi-component generalized forces  $Q_j^D$ , we can introduce a scalar **dissipation function**  $D(\mathbf{q}, \dot{\mathbf{q}})$ ,

$$D = \sum_{i=1}^N \int_0^{v_i} \mu_i(v') dv', \quad (2.279)$$

and write the  $Q_j^D$  as its derivatives with respect to  $\dot{q}_j$ . The Lagrange equations with the generalized friction forces now become

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = -\frac{\partial D}{\partial \dot{q}_j}, \quad (2.280)$$

and we can apply them to systems with friction by specifying  $L$  and  $D$ .

### Interpretation of the Dissipation Function

To understand the physical meaning of  $D$ , we consider the rate of change of the total energy ( $T + V$ ):

$$\frac{d}{dt}(T + V) = \sum_{j=1}^s \left( \frac{\partial T}{\partial q_j} \dot{q}_j + \frac{\partial T}{\partial \dot{q}_j} \ddot{q}_j \right) + \frac{dV}{dt}. \quad (2.281)$$

The second term can be rewritten with the usual trick,

$$\sum_{j=1}^n \frac{\partial T}{\partial \dot{q}_j} \ddot{q}_j = \frac{d}{dt} \left( \sum_{j=1}^n \frac{\partial T}{\partial \dot{q}_j} \dot{q}_j \right) - \sum_{j=1}^n \dot{q}_j \frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j}. \quad (2.282)$$

If we assume scleronomic constraints for simplicity, the kinetic energy  $T$  is a homogeneous function of degree 2 in the generalized velocities, and the parenthesis simply gives us  $2T$  (see Sec. 2.5.4 and Box 2.5). We can use this result along with Eqs. (2.63) and (2.280) to rewrite Eq. (2.281):

$$\begin{aligned} \frac{d}{dt}(T + V) &= \sum_{j=1}^n \left( \frac{\partial T}{\partial q_j} - \frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} \right) \dot{q}_j + \frac{d}{dt}(2T) + \frac{dV}{dt} \\ &= \sum_{j=1}^n \left( \frac{\partial T}{\partial q_j} - \frac{\partial T}{\partial q_j} + \frac{\partial V}{\partial q_j} + \frac{\partial D}{\partial \dot{q}_j} \right) \dot{q}_j + \frac{d}{dt}(2T) + \frac{dV}{dt} \\ &= \sum_{j=1}^n \frac{\partial V}{\partial q_j} \dot{q}_j + \frac{d}{dt}(2T + V) + \sum_{j=1}^n \frac{\partial D}{\partial \dot{q}_j} \dot{q}_j \\ &= 2 \frac{d}{dt}(T + V) + \sum_{j=1}^n \frac{\partial D}{\partial \dot{q}_j} \dot{q}_j, \end{aligned} \quad (2.283)$$

and rearranging, we have

$$\frac{d}{dt}(T + V) = - \sum_{j=1}^n \frac{\partial D}{\partial \dot{q}_j} \dot{q}_j. \quad (2.284)$$

The work that a dynamical system of  $N$  particles must do against the frictional forces under an infinitesimal displacement along the generalized coordinates is

$$dW^D = \sum_{i=1}^N \vec{F}_i^D \cdot d\vec{r}_i = \sum_j Q_j^D dq_j = - \sum_j \frac{\partial D}{\partial \dot{q}_j} dq_j \quad (2.285)$$

and therefore

$$\frac{dW^D}{dt} = - \sum_j \frac{\partial D}{\partial \dot{q}_j} \dot{q}_j, \quad (2.286)$$

so we can also write

$$\frac{d}{dt}(T + V) = \frac{d}{dt} W^D. \quad (2.287)$$

Thus, the change in the systems total energy is equal to the work it does against dissipative forces. Since we restricted ourselves to scleronomic constraints above, the total energy is identical to the ‘‘Hamiltonian’’  $\bar{H}(\mathbf{q}, \dot{\mathbf{q}})$ , and we also have

$$\frac{d}{dt} \bar{H}(\mathbf{q}, \dot{\mathbf{q}}) = - \sum_{j=1}^n \frac{\partial D}{\partial \dot{q}_j} \dot{q}_j. \quad (2.288)$$

### 2.7.2 Viscous Friction and Rayleigh’s Dissipation Function

Rayleigh considered the special case where the friction forces acting on a particle  $i$  are linear in the velocities, i.e.,  $\mu(v) = bv$ . Assuming that the friction force is identical for all particles in the system, but possibly dependent on the direction in which they are moving, the force on particle  $i$  can be written as

$$\vec{F}_i = -\mathbf{B} \dot{\vec{r}}_i, \quad (2.289)$$

where  $\mathbf{B}$  is a symmetric  $3 \times 3$  matrix, or in components

$$F_r^{(i)} = - \sum_{s=1}^3 B_{rs} \dot{x}_s^{(i)}, \quad r, s = 1, 2, 3, \quad (2.290)$$

with  $B_{rs} = B_{sr}$ . Switching to generalized velocities and forces, we obtain

$$Q_j^D = \sum_{i=1}^N \vec{F}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} = - \sum_{i=1}^N \sum_{k=1}^n \mathbf{B} \frac{\partial \vec{r}_i}{\partial q_k} \dot{q}_k \cdot \frac{\partial \vec{r}_i}{\partial q_j} = - \sum_{k=1}^n \beta_{jk} \dot{q}_k, \quad (2.291)$$

where we have defined the symmetric dissipation matrix in generalized coordinates,

$$\beta_{jk} \equiv \sum_{i=1}^N \mathbf{B} \frac{\partial \vec{r}_i}{\partial q_j} \cdot \frac{\partial \vec{r}_i}{\partial q_k} = \sum_{r,s=1}^3 \frac{\partial x_s^{(i)}}{\partial q_j} B_{rs} \frac{\partial x_r^{(i)}}{\partial q_k} \quad (2.292)$$

We can use  $\beta_{jk}$  to write the quadratic dissipation function — also referred to as **Rayleigh's dissipation function** — in a compact form:

$$D = \frac{1}{2} \sum_{j,k=1}^n \beta_{jk} \dot{q}_j \dot{q}_k . \quad (2.293)$$

Since  $\beta_{jk}$  does not depend on the generalized velocities, we readily obtain Eq. (2.291) when we evaluate  $Q_j^D = \frac{\partial D}{\partial \dot{q}_j}$ .

We note that Rayleigh's dissipation function is a homogeneous function of degree 2 (see Box 2.5) so the change in the total energy (and the Hamiltonian) is given by

$$\frac{dE}{dt} = \frac{d}{dt} \bar{H}(\mathbf{q}, \dot{\mathbf{q}}) = -2D \quad (2.294)$$

(see Eqs. (2.284), (2.288)).

### Example: Stokes's Law

As an example, we consider a sphere of mass  $m$  that is moving through a liquid at slow velocity, experiencing a drag force due to the fluid's *laminar flow* around its surface. The drag force is<sup>7</sup>

$$\vec{F}_D = -6\pi\eta R\vec{v} \equiv -\beta\vec{v}, \quad (2.295)$$

where  $\eta$  is the viscosity of the fluid and  $R$  the radius of the sphere. Considering the motion in one dimension, with  $z$  increasing in downward direction, the Lagrangian becomes

$$L = T - V = \frac{m}{2} \dot{z}^2 + mgz , \quad (2.296)$$

and the drag can be modeled by the dissipation function

$$D = \frac{1}{2} \beta v^2 = \frac{1}{2} \beta \dot{z}^2 . \quad (2.297)$$

Thus, we obtain the Lagrange equation

$$m\ddot{z} - mg = -\beta\dot{z} . \quad (2.298)$$

We can write

$$\frac{d}{dt} \dot{z} = g - \frac{\beta}{m} \dot{z} \quad \Rightarrow \quad dt = -\frac{d\dot{z}}{\frac{\beta}{m} \dot{z} - g} , \quad (2.299)$$

and integrate:

$$t - t_0 = -\frac{m}{\beta} \ln \frac{\beta\dot{z} + mg}{\beta v_0 - mg} . \quad (2.300)$$

With the initial conditions  $t_0 = 0, v_0 = 0$ , we can exponential to obtain

$$\exp\left(-\frac{\beta}{m}t\right) = \frac{\beta\dot{z} - mg}{-mg} \quad (2.301)$$

<sup>7</sup> The form of this drag force was first derived in 1851 by G. Stokes, who computed the friction between a viscous fluid and a solid sphere at the sphere's surface using fluid dynamics. Thus, such drag forces are referred to as Stokes's drag or Stokes's friction in the literature.

$$\Leftrightarrow \dot{z} = \frac{mg}{\beta} \left[ 1 - \exp\left(-\frac{\beta}{m}t\right) \right]. \quad (2.302)$$

Thus, we see that the velocity remains finite for large  $t$ , and approaches the *terminal velocity*

$$v_\infty = \frac{mg}{\beta}. \quad (2.303)$$

### Example: Drag Due to Turbulent Flow

If the flow of a fluid or gas around an object is turbulent instead of laminar, e.g., due to the object's high velocity, the drag force is no longer linear as in Stokes's law, but quadratic:

$$\vec{F}_D = -\frac{1}{2}C_D\rho Av^2\frac{\vec{v}}{v} \equiv -\frac{1}{2}\beta v^2\frac{\vec{v}}{v}, \quad (2.304)$$

where  $\rho$  is the density of the fluid,  $A$  the cross-section area of the object orthogonal to its direction of motion,  $v$  the relative velocity of object and fluid flow, and  $C_D$  the so-called drag coefficient. The dissipation function is now cubic,

$$D = \frac{1}{6}\beta v^3. \quad (2.305)$$

Repeating the analysis of the previous example for  $v = \dot{z}$ , we obtain the nonlinear equation of motion

$$m\ddot{z} - mg = -\frac{1}{2}\beta\dot{z}^2. \quad (2.306)$$

[Make this an exercise... Solve the EOM and find the terminal velocity?]

### 2.7.3 Coulomb or Dry Friction

The simple models for static and kinetic friction forces that are discussed in introductory mechanics classes are examples of **dry** or **Coulomb friction**. Coulomb modeled the frictional forces between dry surfaces with the ansatz

$$\vec{F}_{s,k} = -\mu_{s,k}N\frac{\vec{v}}{v}, \quad (2.307)$$

where  $\mu_s$  and  $\mu_k$  are the (constant) static and kinetic friction coefficients that are tabulated for a variety of materials and surface types, and  $N$  is the absolute value of the normal force pressing the surfaces together. We recall that static friction is only considered for masses *at rest*, to define a critical force that is required to set objects in motion. Once the static friction force is overcome, the model switches from the static to the kinetic friction coefficient.

In the framework described in this section, we can obtain the kinetic friction force by setting

$$\mu(v) = \mu_k N = \text{const.} \quad (2.308)$$

and the dissipation function becomes

$$D = \mu_k N v. \quad (2.309)$$

In principle, we can also model the transition from static to kinetic friction with a steep but still smooth change in the friction function  $\mu(v)$  at small velocities (see next example.)

Overall, Coulomb's model is a significant simplification of the underlying microscopic effects that nevertheless proves to be not only versatile, but adequate for many physical systems.



**Example: Wooden Block on a Conveyor Belt**

Consider a block of mass  $m$  that is sitting on a conveyor belt. At  $t = 0$ , a worker pushes it with the force  $\vec{F} = (F_x, F_y)^T$ . We describe the block's motion from the rest frame of the worker, assign coordinates  $x, y$  to it. The absolute value of the relative velocity between block and conveyor belt is given by

$$v_r = \sqrt{(\dot{x} - v_0)^2 + \dot{y}^2}, \quad (2.310)$$

where  $v_0$  is the speed of the belt.

We use the dry friction force

$$\vec{F}_f = -\mu(v) \frac{\vec{v}}{v}, \quad (2.311)$$

The friction function for wood on belt rubber is given by [11]

$$\mu(v) = \left( \frac{\mu_0 - \mu_\infty}{1 + av} + \mu_\infty \right) N = \left( \frac{\mu_0 - \mu_\infty}{1 + av} + \mu_\infty \right) mg, \quad (2.312)$$

where  $\mu_0, \mu_\infty, v_0$  and  $a$  are positive constants, and we have plugged in the magnitude of the normal force,  $|\vec{N}| = mg$ . Note that

$$\mu(v) \xrightarrow{v \rightarrow \infty} \mu_\infty N, \quad \mu(v) \xrightarrow{v \rightarrow 0} \mu_0 N, \quad (2.313)$$

so  $\mu_0$  essentially corresponds to the coefficient of static friction in a Coulomb model, as discussed above.

According to Eq. (2.279), the dissipation function is now given by

$$D = mg \int_0^{v_r} \mu(u) du = mg \left( \frac{\mu_0 - \mu_\infty}{a} \ln \left( 1 + a \sqrt{(\dot{x} - v_0)^2 + \dot{y}^2} \right) + \mu_\infty \sqrt{(\dot{x} - v_0)^2 + \dot{y}^2} \right), \quad (2.314)$$

and its derivatives are

$$\begin{aligned} \frac{\partial D}{\partial \dot{x}} &= mg \left( \frac{\mu_0 - \mu_\infty}{a} \frac{1}{1 + a \sqrt{(\dot{x} - v_0)^2 + \dot{y}^2}} + \mu_\infty \right) \frac{\dot{x} - v_0}{\sqrt{(\dot{x} - v_0)^2 + \dot{y}^2}}, \\ \frac{\partial D}{\partial \dot{y}} &= mg \left( \frac{\mu_0 - \mu_\infty}{a} \frac{1}{1 + a \sqrt{(\dot{x} - v_0)^2 + \dot{y}^2}} + \mu_\infty \right) \frac{\dot{y}}{\sqrt{(\dot{x} - v_0)^2 + \dot{y}^2}}. \end{aligned} \quad (2.315)$$

Thus, the equations of motion for the block read

$$m\ddot{x} = F_x - mg \left( \frac{\mu_0 - \mu_\infty}{a} \frac{1}{1 + a \sqrt{(\dot{x} - v_0)^2 + \dot{y}^2}} + \mu_\infty \right) \frac{\dot{x} - v_0}{\sqrt{(\dot{x} - v_0)^2 + \dot{y}^2}}, \quad (2.316)$$

$$m\ddot{y} = F_y - mg \left( \frac{\mu_0 - \mu_\infty}{a} \frac{1}{1 + a \sqrt{(\dot{x} - v_0)^2 + \dot{y}^2}} + \mu_\infty \right) \frac{\dot{y}}{\sqrt{(\dot{x} - v_0)^2 + \dot{y}^2}}. \quad (2.317)$$

# Chapter 3

## Oscillations

[...]

Oscillatory motion is also so prevalent in physics because it frequently occurs when a system is displaced out of a stable equilibrium. Consider the potential as a function of the generalized coordinates and perform a Taylor expansion around an equilibrium point:

$$V(\vec{q}_0 + \Delta\vec{q}) = V(\vec{q}_0) + \sum_{i=1}^n \left. \frac{\partial V}{\partial q_i} \right|_{\vec{q}_0} \Delta q_i + \frac{1}{2} \sum_{ij} \left. \frac{\partial^2 V}{\partial q_i \partial q_j} \right|_{\vec{q}_0} \Delta q_i \Delta q_j + O(\Delta q^3). \quad (3.1)$$

In an equilibrium point, the (generalized) forces on the system vanish, i.e.,

$$Q_i = -\frac{\partial V}{\partial q_i} = 0, \quad (3.2)$$

and the potential is approximately by a quadratic near an equilibrium point. If the matrix of second derivatives (i.e., the Hessian) is positive definite at  $\vec{q}_0$ , the potential will be that of an harmonic oscillator — but that is precisely the condition under which the equilibrium is stable!

[...]

### 3.1 Simple Oscillators

#### 3.1.1 Damping

Damped oscillators are modelled by including a friction term with linear velocity dependence in the equation of motion. For reasons that will become clear below, we write it as

$$\ddot{x} + \omega_0^2 x = -2\gamma\omega_0 \dot{x} \quad \Leftrightarrow \quad \ddot{x} + 2\gamma\omega_0 \dot{x} + \omega_0^2 x = 0. \quad (3.3)$$

We can determine the **fundamental solutions** by making the ansatz

$$x(t) = e^{\alpha t}, \quad \alpha \in \mathbb{C} \quad (3.4)$$

and deriving the characteristic equation of the ODE. Since we have a second-order ODE, the characteristic equation is quadratic:

$$\alpha^2 e^{\alpha t} + 2\gamma\omega_0 \alpha e^{\alpha t} + \omega_0^2 e^{\alpha t} = 0 \quad \Rightarrow \quad \alpha^2 + 2\gamma\omega_0 \alpha + \omega_0^2 = 0. \quad (3.5)$$

Its solutions are readily found to be

$$\alpha_{1/2} = -\gamma\omega_0 \pm \sqrt{\gamma^2\omega_0^2 - \omega_0^2} = \left(-\gamma \pm \sqrt{\gamma^2 - 1}\right)\omega_0. \quad (3.6)$$

Thus, the general solution of the oscillator ODE can now be written as

$$x(t) = e^{-\gamma\omega_0 t} \left( A e^{\sqrt{\gamma^2 - 1}\omega_0 t} + B e^{-\sqrt{\gamma^2 - 1}\omega_0 t} \right). \quad (3.7)$$

Let us now distinguish several cases.

**No Damping:**  $\gamma = 0$

In the trivial case without damping, we have

$$\alpha_{1/2} = \pm\omega_0\sqrt{-1} = \pm i\omega_0, \quad (3.8)$$

and the general solution of the ODE is given by

$$x(t) = A e^{i\omega_0 t} + A^* e^{-i\omega_0 t}, \quad A \in \mathbb{C}. \quad (3.9)$$

Here,  $A^*$  denotes the complex conjugate of  $A$ . The coefficients of the solution must be complex conjugates because a second-order ODE is uniquely defined by *two* coefficients that are determined via the initial or boundary conditions, which are the real and imaginary parts of  $A$  in our case. (If we had allowed unrelated complex amplitudes for the fundamental solutions, there would be *four* coefficients that cannot be fixed using two conditions.)

**Underdamping:**  $\gamma^2 - 1 < 0$

In this case, the oscillator performs oscillations that decay in time. We have

$$\alpha_{1/2} = \omega_0 \left( -\gamma \pm \sqrt{\gamma^2 - 1} \right) = (-\gamma \pm i\sqrt{1 - \gamma^2})\omega_0, \quad (3.10)$$

and the general solution can be written as

$$x(t) = e^{-\gamma\omega_0 t} \left( A e^{i\omega t} + A^* e^{-i\omega t} \right), \quad (3.11)$$

where we have introduced the **shifted frequency**

$$\omega \equiv \sqrt{1 - \gamma^2}\omega_0. \quad (3.12)$$

**Overdamping:**  $\gamma^2 - 1 > 0$

In this case, the damping is so strong that no oscillation occurs. The discriminant in the solutions of the characteristic equation is real:

$$\alpha_{1/2} = \omega_0 \left( -\gamma \pm \sqrt{\gamma^2 - 1} \right) = (-\gamma \pm \sqrt{\gamma^2 - 1})\omega_0, \quad (3.13)$$

and the general solution can be written as

$$x(t) = e^{-\gamma\omega_0 t} \left( A e^{\lambda t} + B e^{-\lambda t} \right), \quad (3.14)$$

where we have introduced the **decay constant**

$$\lambda \equiv \sqrt{\gamma^2 - 1}\omega_0. \quad (3.15)$$

Note that the positive exponential term counteracts the overall damping prefactor  $e^{-\gamma\omega_0 t}$

**Critical Damping:**  $\gamma^2 = 1$ 

The critically damped oscillator does not undergo any oscillations either, but in this case the decay to zero amplitude is shortest because the decay constant  $\lambda$  we defined for the overdamped case vanishes exactly. We now note a peculiar issue: while we obtained two solutions for the under- and overdamped cases, as required for a second-order ODE, we now only seem to have one.

To resolve the issue, we consider the limit of the overdamped solution for infinitesimally small  $\lambda = \epsilon$ :

$$\begin{aligned} x(t) &= e^{-\gamma\omega_0 t} (Ae^{\epsilon t} + Be^{-\epsilon t}) = e^{-\gamma\omega_0 t} (A(1 + \epsilon t) + B(1 - \epsilon t) + O(\epsilon^2)) \\ &= e^{-\gamma\omega_0 t} (A + B + (A - B)\epsilon t + O(\epsilon^2)) . \end{aligned} \quad (3.16)$$

We can replace the linear combinations  $A \pm B$  with new coefficients

$$x(t) = (C_1 + C_2 t) e^{-\gamma\omega_0 t}, \quad C_1, C_2 \in \mathbb{R}. \quad (3.17)$$

**3.1.2 Periodic Driving Forces and Resonance**

[...] Let us now consider an oscillator with a periodic driving force, which is given by the inhomogeneous ODE

$$\ddot{x} + 2\gamma\omega_0 \dot{x} + \omega_0^2 x = \frac{F_0}{m} \cos(\omega t + \phi). \quad (3.18)$$

For compactness, we switch to the complex coordinate  $z(t)$  and use a complex ansatz for the driving force:

$$\ddot{z} + 2\gamma\omega_0 \dot{z} + \omega_0^2 z = C e^{i\omega t}, \quad C \in \mathbb{R}. \quad (3.19)$$

We have already determined the possible solutions of the homogeneous ODE, so we only need to find a particular solution of the inhomogeneous ODE here. We make the ansatz  $z(t) = z_0 e^{i\omega t}$  and obtain

$$(-\omega^2 z_0 + 2i\gamma\omega\omega_0 + \omega_0^2 z_0) e^{i\omega t} = C e^{i\omega t}. \quad (3.20)$$

After rearranging, the complex amplitude is given by

$$z_0 = \frac{C}{\omega_0^2 - \omega^2 + 2i\gamma\omega\omega_0}. \quad (3.21)$$

We can isolate its real and imaginary parts:

$$z_0 = \frac{C(\omega_0^2 - \omega^2 - 2i\gamma\omega\omega_0)}{(\omega_0^2 - \omega^2)^2 + (2\gamma\omega\omega_0)^2}, \quad (3.22)$$

i.e.,

$$\operatorname{Re} z_0 = \frac{C(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + (2\gamma\omega\omega_0)^2}, \quad \operatorname{Im} z_0 = \frac{-2C\gamma\omega\omega_0}{(\omega_0^2 - \omega^2)^2 + (2\gamma\omega\omega_0)^2}. \quad (3.23)$$

Thus,  $z_0$  is real if  $\gamma = 0$ , i.e., if we don't have any damping. To obtain the real-valued solution, we express  $z_0$  in polar representation,

$$z_0 = |z_0| e^{i\phi}, \quad (3.24)$$

with

$$\begin{aligned} |z_0| &= \sqrt{(\operatorname{Re} z_0)^2 + (\operatorname{Im} z_0)^2} = C \sqrt{\frac{(\omega_0^2 - \omega^2)^2 + (2\gamma\omega\omega_0)^2}{\left[(\omega_0^2 - \omega^2)^2 + (2\gamma\omega\omega_0)^2\right]^2}} \\ &= \frac{C}{\sqrt{(\omega_0^2 - \omega^2)^2 + (2\gamma\omega\omega_0)^2}} \equiv A(\omega), \end{aligned} \quad (3.25)$$

and

$$\tan \phi(\omega) = \frac{\operatorname{Im} z_0}{\operatorname{Re} z_0} = \frac{-2\gamma\omega\omega_0}{\omega_0^2 - \omega^2}. \quad (3.26)$$

Thus, the particular solution to the driven oscillator ODE is given by

$$z_p(t) = A(\omega)e^{i(\omega t + \phi(\omega))}, \quad (3.27)$$

and its real part is

$$x_p(t) = A(\omega) \cos(\omega t + \phi(\omega)). \quad (3.28)$$

The general solutions for underdamping, critical damping and overdamping are obtained by adding this particular solution to the homogeneous solutions obtained in the previous section.

### 3.1.3 Phase Space

### 3.1.4 General Driving Forces

[...] The **Fourier transform** is defined as

$$\tilde{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt e^{-i\omega t} f(t). \quad (3.29)$$

with the **inverse transform**

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega e^{i\omega t} \tilde{f}(\omega), \quad (3.30)$$

[...] The Fourier transform can be used to convert ODEs into algebraic equations. Let us apply this to the damped oscillator with a general driving force,

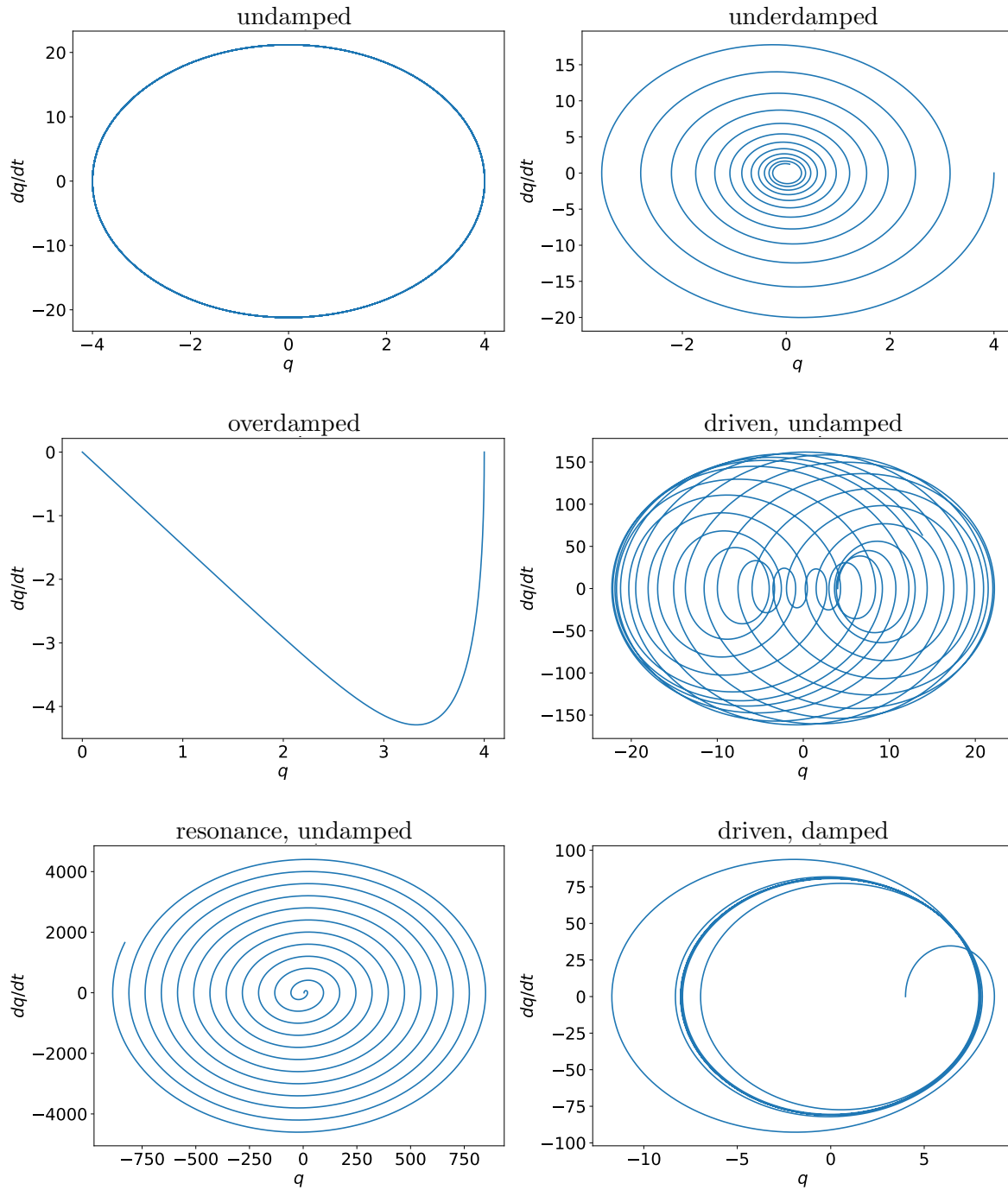
$$\ddot{z} + 2\gamma\omega_0\dot{z} + \omega_0^2 z = f(t). \quad (3.31)$$

We express  $z(t)$  and its derivatives in terms of its Fourier transforms:

$$z(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega e^{i\omega t} \tilde{z}(\omega), \quad (3.32)$$

$$\begin{aligned} \dot{z}(t) &= \frac{1}{\sqrt{2\pi}} \frac{d}{dt} \int_{-\infty}^{\infty} d\omega e^{i\omega t} \tilde{z}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega \frac{d}{dt} (e^{i\omega t} \tilde{z}(\omega)) \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega i\omega e^{i\omega t} \tilde{z}(\omega), \end{aligned} \quad (3.33)$$

$$\ddot{z}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega (-\omega^2) e^{i\omega t} \tilde{z}(\omega), \quad (3.34)$$



**Box 3.1: The Dirac Distribution**

The **Dirac  $\delta$  distribution**, also referred to as the  **$\delta$  function** [...] Dirac introduced the function as a generalization of the Kronecker symbol  $\delta_{ij}$  to continuous variables. It is formally defined by its action in an integral.

$$\int_a^b dx f(x)\delta(x-x_0) = \begin{cases} f(x_0) & \text{for } x_0 \in [a, b], \\ 0 & \text{else.} \end{cases} \quad (\text{B3.1-1})$$

- It is not a function in the proper sense. It is implied that  $\delta(x-x_0)$  appears under an integral, so that the definition (B3.1-1) can be applied.
- Schwartz developed the formal theory of  $\delta(x-x_0)$  and other distributions by showing that they define functionals on a space of differentiable functions.
- One can view it as the limit of a series of functions, e.g., Gaussians,

$$g_a(x) = \frac{1}{\sqrt{2\pi a}} e^{-\frac{x^2}{2a^2}}, \quad (\text{B3.1-2})$$

so that

$$\lim_{a \rightarrow \infty} g_a(x) = \delta(x), \quad (\text{B3.1-3})$$

but the object obtained in the limit is not contained in any space of proper functions.

- 

$$\delta(t-t_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega e^{i\omega(t-t_0)}. \quad (\text{B3.1-4})$$

where we have used that the time derivative commutes with the integration over  $\omega$  and can therefore be moved into the integral. Expressing  $f(t)$  as a Fourier transform as well and collecting all terms on one side, we have

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega \left[ (-\omega^2 + 2i\gamma\omega\omega_0 + \omega_0^2) \tilde{z}(\omega) - \tilde{f}(\omega) \right] e^{i\omega t} = 0. \quad (3.35)$$

For general functions, this relationship can only hold if the integrand vanishes, i.e.,

$$(\omega_0^2 - \omega^2 + 2i\gamma\omega\omega_0) \tilde{z}(\omega) = \tilde{f}(\omega). \quad (3.36)$$

Solving this equation for  $\tilde{z}(\omega)$ , we obtain

$$\tilde{z}(\omega) = \frac{\tilde{f}(\omega)}{\omega_0^2 - \omega^2 + 2i\gamma\omega\omega_0}. \quad (3.37)$$

Now we only need to perform the inverse Fourier transform of this result to obtain  $z(t)$ .

As an example, we apply this procedure to the periodic driving force with a single frequency, as discussed in Sec. 3.1.2.

## 3.2 Coupled Oscillators

### 3.2.1 Example: Two Coupled Pendula

We start the discussion of coupled oscillators by recalling the procedure for solving the equations of motion of such systems. As an example, we consider two equal pendula of length  $l$  and mass  $m$  that perform small oscillations while being coupled by an ideal spring with spring constant  $k$  and natural length  $d$  (see Fig. 3.1). Letting the  $y$  axis point upward, the coordinates of the masses are given by

$$x_1 = l \sin \theta_1, \quad \dot{x}_1 = l \dot{\theta}_1 \cos \theta_1, \quad (3.38)$$

$$y_1 = l(1 - \cos \theta_1), \quad \dot{y}_1 = l \dot{\theta}_1 \sin \theta_1, \quad (3.39)$$

$$x_2 = d + l \sin \theta_2, \quad \dot{x}_2 = l \dot{\theta}_2 \cos \theta_2, \quad (3.40)$$

$$y_2 = l(1 - \cos \theta_2), \quad \dot{y}_2 = l \dot{\theta}_2 \sin \theta_2, \quad (3.41)$$

so the kinetic energy reads

$$T = \frac{1}{2}m(\dot{x}_1^2 + \dot{y}_1^2 + \dot{x}_2^2 + \dot{y}_2^2) = \frac{1}{2}ml^2(\dot{\theta}_1^2 + \dot{\theta}_2^2). \quad (3.42)$$

For the potential term, we first consider the spring potential:

$$\begin{aligned} V_{\text{sp}} &= \frac{k}{2} \left( \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} - d \right)^2 \\ &= \frac{k}{2} \left( (x_2 - x_1) \sqrt{1 + \left( \frac{y_2 - y_1}{x_2 - x_1} \right)^2} - d \right)^2, \end{aligned} \quad (3.43)$$

where we have used that  $x_2 - x_1 > 0$  for small-angle oscillations. For small angles, we also have

$$x_2 - x_1 = d + l\theta_2 - l\theta_1 + O(\theta^3), \quad (3.44)$$

$$y_2 - y_1 = l\frac{\theta_2^2}{2} - l\frac{\theta_1^2}{2} + O(\theta^4) = \frac{l}{2}(\theta_2^2 - \theta_1^2) + O(\theta^4). \quad (3.45)$$

We can expand the ratio in the potential,

$$\begin{aligned} \frac{y_2 - y_1}{x_2 - x_1} &= \frac{\frac{l}{2}(\theta_2^2 - \theta_1^2) + O(\theta^4)}{d + l(\theta_2 - \theta_1) + O(\theta^3)} = \frac{l}{2d}(\theta_2^2 - \theta_1^2 + O(\theta^4)) \frac{1}{1 + \frac{l}{d}(\theta_2 - \theta_1) + O(\theta^3)} \\ &= \frac{l}{2d}(\theta_2^2 - \theta_1^2 + O(\theta^4)) \left( 1 - \frac{l}{d}(\theta_2 - \theta_1) + O(\theta^2) \right) \\ &= \frac{l}{2d}(\theta_2^2 - \theta_1^2) + O(\theta^3), \end{aligned} \quad (3.46)$$

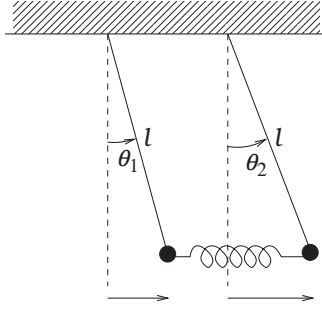
so

$$\sqrt{1 + \left( \frac{y_2 - y_1}{x_2 - x_1} \right)^2} = 1 + \frac{1}{2} \left( \frac{y_2 - y_1}{x_2 - x_1} \right)^2 + \dots = 1 + O(\theta^4) \quad (3.47)$$

Now the spring potential reads

$$V_{\text{sp}} = \frac{k}{2} (x_2 - x_1 - d + O(\theta^4))^2 = \frac{k}{2} (l\theta_2 + d - l\theta_1 - d + O(\theta^3)) \quad (3.48)$$





**Figure 3.1:** Coordinates for two pendula that are coupled by a spring.

and the total potential becomes

$$\begin{aligned} V &= \frac{k}{2} l^2 (\theta_2 - \theta_1 + O(\theta^3))^2 + mgl \left( \frac{\theta_1^2}{2} + \frac{\theta_2^2}{2} \right) + O(\theta^3) \\ &= \frac{k}{2} l^2 (\theta_2 - \theta_1)^2 + mgl \left( \frac{\theta_1^2}{2} + \frac{\theta_2^2}{2} \right) + O(\theta^3). \end{aligned} \quad (3.49)$$

Putting everything together, we obtain the Lagrangian

$$L = \frac{1}{2} ml^2 (\dot{\theta}_1^2 + \dot{\theta}_2^2) - \frac{k}{2} l^2 (\theta_2 - \theta_1)^2 - mgl \left( \frac{\theta_1^2}{2} + \frac{\theta_2^2}{2} \right). \quad (3.50)$$

Next, we derive the Lagrange equations. The partial derivatives are

$$\frac{\partial L}{\partial \dot{\theta}_j} = ml^2 \dot{\theta}_j, \quad \frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}_j} = ml^2 \ddot{\theta}_j, \quad (3.51)$$

$$\frac{\partial L}{\partial \theta_1} = kl^2 (\theta_2 - \theta_1) - mgl \theta_1, \quad (3.52)$$

$$\frac{\partial L}{\partial \theta_2} = -kl^2 (\theta_2 - \theta_1) - mgl \theta_2, \quad (3.53)$$

and the equations of motion read

$$ml^2 \ddot{\theta}_1 - kl^2 (\theta_2 - \theta_1) + mgl \theta_1 = 0, \quad (3.54)$$

$$ml^2 \ddot{\theta}_2 + kl^2 (\theta_2 - \theta_1) + mgl \theta_2 = 0. \quad (3.55)$$

The usual procedure for solving this system of equation is to make the ansatz

$$\begin{pmatrix} \theta_1(t) \\ \theta_2(t) \end{pmatrix} = \begin{pmatrix} \theta_{1,0} \\ \theta_{2,0} \end{pmatrix} e^{i\omega t}. \quad (3.56)$$

Note that the same frequency  $\omega$  is used for both components of the solution. We have

$$\ddot{\theta}_j(t) = -\omega^2 \theta_j(t), \quad (3.57)$$

and we can use this to write the system of equations of motion in matrix form:

$$\begin{pmatrix} -ml^2 \omega^2 + mgl + kl^2 & -kl^2 \\ -kl^2 & -ml^2 \omega^2 + mgl + kl^2 \end{pmatrix} \begin{pmatrix} \theta_{1,0} \\ \theta_{2,0} \end{pmatrix} = 0. \quad (3.58)$$

This is an eigenvalue problem whose solutions are the **characteristic frequencies**  $\omega^2$  and associated **characteristic vectors** that define the **normal modes** of the system of coupled oscillators.

The characteristic polynomial of the matrix is

$$\begin{aligned} \det \begin{pmatrix} -ml^2\omega^2 + mgl + kl^2 & -kl^2 \\ -kl^2 & -ml^2\omega^2 + mgl + kl^2 \end{pmatrix} \\ = (-ml^2\omega^2 + mgl + kl^2)^2 - (-kl^2)^2 = 0. \end{aligned} \quad (3.59)$$

Moving the final term to the right and taking the square root on both sides, we end up with the equation

$$-ml^2\omega^2 + mgl + kl^2 = \pm kl^2. \quad (3.60)$$

Its solutions are

$$\omega_+^2 = \frac{g}{l}, \quad \omega_-^2 = \frac{g}{l} + \frac{2k}{m}. \quad (3.61)$$

We can determine the characteristic vectors by plugging the frequencies back into Eq. (3.58). For  $\omega_+^2$ , we have

$$\begin{aligned} 0 &= \begin{pmatrix} -ml^2\frac{g}{l} + mgl + kl^2 & -kl^2 \\ -kl^2 & -ml^2\frac{g}{l} + mgl + kl^2 \end{pmatrix} \begin{pmatrix} \rho_1^{(+)} \\ \rho_2^{(+)} \end{pmatrix} \\ &= \begin{pmatrix} kl^2 & -kl^2 \\ -kl^2 & kl^2 \end{pmatrix} \begin{pmatrix} \rho_1^{(+)} \\ \rho_2^{(+)} \end{pmatrix}, \end{aligned} \quad (3.62)$$

which means that the components of the characteristic vector need to satisfy

$$\rho_1^{(+)} = \rho_2^{(+)}. \quad (3.63)$$

Thus, a normalized solution is

$$\vec{\rho}^{(+)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (3.64)$$

For  $\omega_-^2$ , we find

$$\begin{aligned} 0 &= \begin{pmatrix} -ml^2 \left( \frac{g}{l} + 2\frac{k}{m} \right) + mgl + kl^2 & -kl^2 \\ -kl^2 & -ml^2 \left( \frac{g}{l} + 2\frac{k}{m} \right) + mgl + kl^2 \end{pmatrix} \begin{pmatrix} \rho_1^{(-)} \\ \rho_2^{(-)} \end{pmatrix} \\ &= \begin{pmatrix} -kl^2 & -kl^2 \\ -kl^2 & -kl^2 \end{pmatrix} \begin{pmatrix} \rho_1^{(-)} \\ \rho_2^{(-)} \end{pmatrix}, \end{aligned} \quad (3.65)$$

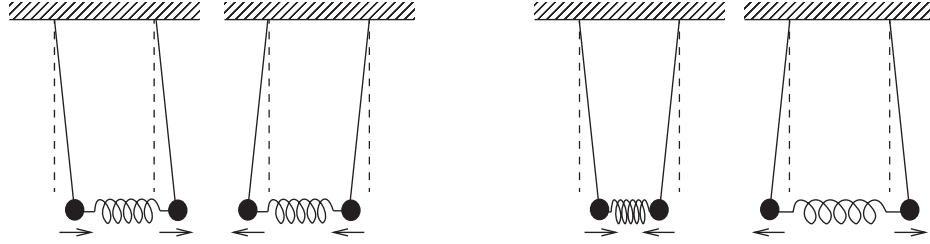
which implies

$$\rho_1^{(-)} = -\rho_2^{(-)}, \quad (3.66)$$

so a solution for the second characteristic vector is

$$\vec{\rho}^{(-)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (3.67)$$

Figure 3.2 visualizes the two normal modes we have found. For  $\vec{\rho}^{(+)}$ , the two pendula are in sync, while the masses swing in opposite directions for  $\vec{\rho}^{(-)}$ , as indicated by the signs of the characteristic vectors.



**Figure 3.2:** Normal modes for two pendula that are coupled by a spring.

### 3.2.2 Normal Coordinates and Normal Modes

In the example discussed in the previous section, we determined the normal modes of a system of coupled oscillators *after* deriving the equations of motion. The characteristic vectors associated with these modes constitute a complete basis for the description of the system — in fact, if we make a change of basis from our initial choice of coordinates and basis vectors to the normal mode basis, the equations of motion are *decoupled* and can be solved independently.

The central premise of the Lagrangian formalism, however, is to formulate the problem in the most efficient choice of generalized coordinates from the very beginning (unless we need to deal with some constraints explicitly). If we can identify the matrix appearing in the equations of motion of a system of oscillators already in the Lagrangian, we can diagonalize it and express  $L$  directly in terms of normal coordinates with respect to the characteristic vectors. The Lagrange equations for the normal coordinates are then particularly easy to derive, because they correspond to simple, uncoupled oscillators.

[...] Consider a general Lagrangian of the form

$$L = \frac{1}{2} \sum_{jk} M_{jk} \dot{q}_j \dot{q}_k - V(\vec{q}), \quad (3.68)$$

where  $\vec{q} = (q_1, \dots, q_n)$  and  $\mathbf{M}$  is the mass tensor first defined in Eq. (2.223):

$$M_{jk} = \sum_{i=1}^A m_i \frac{\partial \vec{r}_i}{\partial q_j} \cdot \frac{\partial \vec{r}_i}{\partial q_k} = M_{kj}. \quad (3.69)$$

Let us consider displacements  $\eta_j$  around an equilibrium position of the system:

$$q_j = q_{0j} + \eta_j, \quad (3.70)$$

$$\dot{q}_j = \dot{\eta}_j. \quad (3.71)$$

The potential can be expanded as

$$\begin{aligned} V(\vec{q}) &= V(\vec{q}_0) + \sum_j \underbrace{\frac{\partial V}{\partial q_j} \Big|_{\vec{q}_0}}_{=0} \eta_j + \frac{1}{2} \sum_{jk} \frac{\partial^2 V}{\partial q_j \partial q_k} \Big|_{\vec{q}_0} \eta_j \eta_k + O(\eta^3) \\ &\equiv V(\vec{q}_0) + \frac{1}{2} \sum_{jk} V_{jk} \eta_j \eta_k + O(\eta^3), \end{aligned} \quad (3.72)$$

where we have used that the gradient of the potential vanishes in equilibrium, and introduced the symmetric matrix

$$V_{jk} = \frac{\partial^2 V}{\partial q_j \partial q_k} = V_{kj}, \quad (3.73)$$

which is nothing but the Hessian of the system evaluated at  $\vec{q}_0$ . Dropping cubic and higher terms in  $\eta$ , a general potential therefore looks like a system of coupled oscillators near an equilibrium configuration. For a system of coupled harmonic oscillators, the higher-order terms vanish and the expression for the potential is exact.

Introducing

$$T_{jk} \equiv M_{jk}|_{\vec{q}_0} = T_{kj}, \quad (3.74)$$

and dropping the constant  $V(\vec{q}_0)$  we can write the Lagrangian as a **quadratic form**,

$$L = \frac{1}{2} \dot{\vec{\eta}}^T \mathbf{T} \dot{\vec{\eta}} - \frac{1}{2} \vec{\eta}^T \mathbf{V} \vec{\eta} = \frac{1}{2} \sum_{jk} \dot{\eta}_j T_{jk} \dot{\eta}_k - \frac{1}{2} \sum_{jk} \eta_j V_{jk} \eta_k. \quad (3.75)$$

### Equations of Motion and Generalized Eigenvalue Problem

The partial derivatives are

$$\begin{aligned} \frac{\partial L}{\partial \eta_j} &= \frac{1}{2} \sum_{kl} \frac{\partial}{\partial \eta_j} (\eta_k V_{kl} \eta_l) = \frac{1}{2} \sum_{kl} \left( \frac{\partial \eta_k}{\partial \eta_j} V_{kl} \eta_l + \eta_k V_{kl} \frac{\partial \eta_l}{\partial \eta_j} \right) \\ &= \frac{1}{2} \sum_{kl} (\delta_{jk} V_{kl} \eta_l + \eta_k V_{kl} \delta_{jl}) \\ &= \sum_k V_{jk} \eta_k, \end{aligned} \quad (3.76)$$

and

$$\frac{\partial L}{\partial \dot{\eta}_j} = \sum_k T_{jk} \dot{\eta}_k, \quad (3.77)$$

where we have used the properties of the Kronecker delta and combined terms by using the freedom to rename summation indices. In terms of the matrices  $\mathbf{T}$  and  $\mathbf{V}$ , the Lagrange equations therefore read

$$\sum_k (T_{jk} \ddot{\eta}_k + V_{jk} \eta_k) = 0, \quad j = 1, \dots, n. \quad (3.78)$$

To identify the normal modes, we make the usual switch to complex coordinates

$$\eta_j \longrightarrow z_j = z_{0j} e^{i\omega t}, \quad z_j \in \mathbb{C}, \quad (3.79)$$

and plug them into the equations of motion to obtain

$$\sum_k (-\omega^2 T_{jk} z_{0k} + V_{jk} z_{0k}) = 0, \quad (3.80)$$

or in matrix form

$$(\mathbf{V} - \omega^2 \mathbf{T}) \vec{z}_0 = 0. \quad (3.81)$$

This is, in fact, a **generalized eigenvalue problem** since  $\mathbf{T}$  appears in place of the identity matrix  $\mathbb{1}$ . It contains the information about the **curvature of the configuration manifold** itself in the vicinity of the equilibrium configuration we expand around. If the configuration manifold is flat, e.g., because we use a Cartesian basis to describe the system,  $\mathbf{T}$  will be simply a diagonal matrix with the masses of the oscillators on the diagonal. If  $\mathbf{T}$  is invertible, we can in principle convert Eq. (3.82) into a regular eigenvalue problem by multiplying with  $\mathbf{T}^{-1}$  from the left,

$$(\mathbf{T}^{-1} \mathbf{V} - \omega^2 \mathbb{1}) \vec{z}_0 = 0. \quad (3.82)$$

Solving Eq. (3.82), we obtain the characteristic frequencies  $\omega_s^2$  as the eigenvalues, and the characteristic vectors  $\vec{\rho}^{(s)}$ . Since  $\mathbf{T}$  and  $\mathbf{V}$  are real and symmetric, the  $\omega_s^2$  are guaranteed to be real as well. The characteristic vectors will be real as well, since their entries are just linear combinations of the original coordinates we used to describe the oscillators. A general oscillation of the system can then be expanded in the basis of normal modes instead of our initial choice of basis vectors (denoted here by  $\{\vec{e}^{(1)}, \dots, \vec{e}^{(n)}\}$ ):

$$\vec{\eta}(t) = \sum_{j=1}^n \eta_j(t) \vec{e}^{(j)} = \sum_{s=1}^n \zeta_s(t) \vec{\rho}^{(s)}, \quad (3.83)$$

where  $\zeta_s$  are the **normal coordinates** of the problem.

### Determination of the Normal Coordinates

Let us now assume that we have a Lagrangian given by Eq. (3.75), and introduce normal coordinates that are related to our initial displacement coordinates by a matrix  $\mathbf{A}$  that is to be determined:

$$\vec{\eta} = \mathbf{A} \vec{\zeta}, \quad \eta_k, \zeta_k \in \mathbb{R}. \quad (3.84)$$

Plugging this relation into the Lagrangian, we have

$$L = \frac{1}{2} \dot{\vec{\zeta}}^T \mathbf{A}^T \mathbf{T} \mathbf{A} \dot{\vec{\zeta}} - \frac{1}{2} \vec{\zeta}^T \mathbf{A}^T \mathbf{V} \mathbf{A} \vec{\zeta}. \quad (3.85)$$

If the normal coordinates are to be decoupled,  $\mathbf{A}^T \mathbf{T} \mathbf{A}$  and  $\mathbf{A}^T \mathbf{V} \mathbf{A}$  must each be diagonal because they depend on  $\dot{\vec{\zeta}}$  and  $\vec{\zeta}$ , respectively, so there can be no cancellations in off-diagonal matrix elements between the two terms.

To proceed, we first define the **inner product induced by  $\mathbf{T}$** :

$$(\vec{a}, \vec{b}) \equiv \vec{a}^T \mathbf{T} \vec{b} = \sum_{kl} a_k T_{kl} b_l \quad (3.86)$$

It generalizes the standard scalar product from a flat space to the geometry of configuration manifold or, more precisely, the tangent vector spaces to the configuration manifold at the equilibrium configuration. For  $T_{kl} = \delta_{kl}$ , we get back the regular scalar product.

The characteristic vectors  $\vec{\rho}^{(s)}$  are orthonormal with respect to the new inner product:

$$(\vec{\rho}^{(r)}, \vec{\rho}^{(s)}) = \sum_{kl} \rho_k^{(r)} T_{kl} \rho_l^{(s)} = \delta_{rs}. \quad (3.87)$$

[Add proofs - math “appendix” subsection?] Let us now define the **modal matrix**  $\mathbf{A}$  in terms of the characteristic vectors

$$\mathbf{A} \equiv \begin{pmatrix} \rho_1^{(1)} & \cdots & \rho_1^{(n)} \\ \vdots & \ddots & \vdots \\ \rho_n^{(1)} & \cdots & \rho_n^{(n)} \end{pmatrix}, \quad (3.88)$$

or componentwise

$$A_{ks} = \rho_k^{(s)}. \quad (3.89)$$

Then

$$\sum_{kl} \rho_k^{(r)} T_{kl} \rho_l^{(s)} = \sum_{kl} A_{kr} T_{kl} A_{ls} = \sum_{kl} A_{rk}^T T_{kl} A_{ls} \stackrel{!}{=} \delta_{rs}, \quad (3.90)$$

which means that  $\mathbf{T}$  will be represented by the identity matrix in the basis of characteristic vectors:

$$\mathbf{A}^T \mathbf{T} \mathbf{A} = \mathbb{1}. \quad (3.91)$$

For the potential term, the eigenvalue problem implies

$$\sum_{kl} V_{kl} \rho_l^{(s)} = \omega_s^2 \sum_l T_{kl} \rho_l^{(s)}. \quad (3.92)$$

Defining the diagonal matrix

$$\mathbf{W} = \text{diag}(\omega_1^2, \dots, \omega_n^2), \quad (3.93)$$

we can rewrite this as

$$\sum_{kl} V_{kl} A_{ls} = \sum_l T_{kl} A_{ls} = \sum_{lr} T_{kl} A_{lr} W_{rs}, \quad (3.94)$$

or in matrix form as

$$\mathbf{V} \mathbf{A} = \mathbf{T} \mathbf{A} \mathbf{W}. \quad (3.95)$$

Multiplying from the left by  $\mathbf{A}^T$  and using the fact that  $\mathbf{T}$  is the identity matrix in the basis of characteristic vectors, we have

$$\mathbf{A}^T \mathbf{V} \mathbf{A} = \underbrace{\mathbf{A}^T \mathbf{T} \mathbf{A}}_{=\mathbb{1}} \mathbf{W} = \mathbf{W}, \quad (3.96)$$

so the modal matrix also renders  $\mathbf{V}$  diagonal.

Plugging our results for the transformed matrices back into Eq. (3.85), the Lagrangian becomes

$$L = \frac{1}{2} \dot{\vec{\zeta}}^T \mathbf{A}^T \mathbf{T} \mathbf{A} \dot{\vec{\zeta}} - \frac{1}{2} \vec{\zeta}^T \mathbf{A}^T \mathbf{V} \mathbf{A} \vec{\zeta} = \frac{1}{2} \dot{\vec{\zeta}}^T \cdot \dot{\vec{\zeta}} - \frac{1}{2} \vec{\zeta}^T \mathbf{W} \vec{\zeta} = \frac{1}{2} \sum_s \left( \dot{\zeta}_s^2 - \omega_s^2 \zeta_s^2 \right), \quad (3.97)$$

and the Lagrange equations yield the decoupled equations of motion

$$\ddot{\zeta}_s + \omega_s^2 \zeta_s = 0, \quad s = 1, \dots, n. \quad (3.98)$$

as desired. The normal coordinates can now be determined from original ansatz (3.84) by mutiplying it with  $\mathbf{A}^T \mathbf{T}$  from the left-hand side and using  $\mathbf{A}^T \mathbf{T} \mathbf{A} = \mathbb{1}$ . We obtain

$$\vec{\zeta} = \mathbf{A}^T \mathbf{T} \vec{\eta} \quad (3.99)$$

using the coefficient matrix  $\mathbf{T}$  and the modal matrix  $\mathbf{A}$  obtained from the characteristic vectors of our generalized eigenvalue problem.

[Add remark about additional generalized forces / cf. homework.]

### 3.2.3 Example: A Linear Triatomic Molecule

We now demonstrate the introduction of normal coordinates in the Lagrangian for the example of a linear triatomic molecule, which we describe as a central mass  $M$  that is connected to two equal masses  $m$  via identical springs with spring constants  $k$ . The spring potential can be viewed as the harmonic approximation to the (classical) intermolecular interaction close to an equilibrium configuration, i.e.,

$$V(|x_i - x_j|) = V_0 + \frac{1}{2} \sum_{ij} \frac{\partial^2 V}{\partial x_i \partial x_j} \Big|_{|x_i,0-x_j,0|} \Delta x_i \Delta x_j + O(\Delta x^3). \quad (3.100)$$

Referring to the coordinates introduced in Fig. 3.3, we have

$$V = \frac{k}{2} (x_2 - x_1 - l)^2 + \frac{k}{2} (x_3 - x_2 - l)^2, \quad (3.101)$$

where  $l$  is the natural length of the spring. If we expand the squares in  $V$ , we get

$$V = \frac{k}{2} (x_1^2 + 2x_2^2 + x_3^2 + 2l^2 - 2x_1x_2 + 2lx_1 - 2lx_2 - 2x_2x_3 - 2lx_2 + 2lx_3), \quad (3.102)$$

which cannot be written as a quadratic form in the coordinates due to the presence of terms like  $lx_i$ . Thus, we first have to introduce coordinates that directly measure the displacement out of equilibrium:

$$\eta_1 = x_1 + l, \quad \eta_2 = x_2, \quad \eta_3 = x_3 - l \quad (3.103)$$

(other choices are possible as long as they allow us to write  $L$  as a quadratic form.) In terms of the new coordinates, the Lagrangian of the system can be written as

$$\begin{aligned} L &= \frac{1}{2} m (\dot{x}_1^2 + \dot{x}_3^2) + \frac{1}{2} M \dot{x}_2^2 - \frac{k}{2} (x_2 - x_1 - l)^2 - \frac{k}{2} (x_3 - x_2 - l)^2 \\ &= \frac{1}{2} m (\dot{\eta}_1^2 + \dot{\eta}_3^2) + \frac{1}{2} M \dot{\eta}_2^2 - \frac{k}{2} (\eta_2 - \eta_1)^2 - \frac{k}{2} (\eta_3 - \eta_2)^2, \end{aligned} \quad (3.104)$$

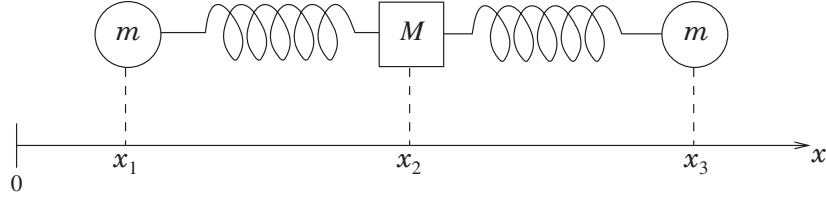
or

$$L = \frac{1}{2} \dot{\vec{\eta}}^T \mathbf{T} \dot{\vec{\eta}} - \frac{1}{2} \vec{\eta}^T \mathbf{V} \vec{\eta} \quad (3.105)$$

with the matrices

$$\mathbf{T} = \begin{pmatrix} m & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & m \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix}. \quad (3.106)$$

Note that we have distributed terms of the form  $2\eta_i\eta_j$  to the upper and lower off-diagonal parts of  $\mathbf{V}$  so that the matrix is symmetric.



**Figure 3.3:** Coordinates for the linear triatomic molecule

Now we solve the generalized eigenvalue problem for  $\mathbf{V} - \omega^2\mathbf{T}$ , starting with the determination of the characteristic polynomial:

$$\begin{aligned}
 0 &= \det \begin{pmatrix} k - m\omega^2 & -k & 0 \\ -k & 2k - M\omega^2 & -k \\ 0 & -k & k - m\omega^2 \end{pmatrix} \\
 &= (k - m\omega^2) [(2k - M\omega^2)(k - m\omega^2) - (-k)^2] - (-k) [(-k)(k - m\omega^2)] \\
 &= (k - m\omega^2) [(2k - M\omega^2)(k - m\omega^2) - 2k^2] \\
 &= (k - m\omega^2) [2k^2 - k(2m + M)\omega^2 + mM\omega^4 - 2k^2] \\
 &= \omega^2 (k - m\omega^2) [-k(2m + M) + mM\omega^2]. \tag{3.107}
 \end{aligned}$$

We can immediately read off the solutions

$$\omega_1^2 = 0, \quad \omega_2^2 = \frac{k}{m}, \quad \omega_3^2 = \frac{k}{m} \left(1 + \frac{2m}{M}\right). \tag{3.108}$$

Plugging these frequencies into the eigenvalue equation, we can determine the characteristic vectors. For  $\omega_1^2 = 0$ , we have

$$\begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{pmatrix} = 0, \tag{3.109}$$

and the first and third rows give us the equations

$$\rho_1 = \rho_2, \quad \rho_2 = \rho_3. \tag{3.110}$$

Thus, our first characteristic vector is

$$\vec{\rho}^{(1)} = C_1 (1 \ 1 \ 1) \tag{3.111}$$

with some normalization constant  $C_1$ .

Analogously,  $\omega_2^2 = k/m$  yields

$$\begin{pmatrix} k - km\frac{k}{m} & -k & 0 \\ -k & 2k - M\frac{k}{m} & -k \\ 0 & -k & k - m\frac{k}{m} \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{pmatrix} = \begin{pmatrix} 0 & -k & 0 \\ -k & k(2 - \frac{M}{m}) & -k \\ 0 & -k & 0 \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{pmatrix} = 0, \tag{3.112}$$



and we obtain the equations

$$-k\rho_2 = 0, \quad (3.113)$$

$$-k(\rho_1 + \rho_3) + k\left(2 - \frac{m}{M}\right)\rho_2 = 0. \quad (3.114)$$

Thus, our second characteristic vector is

$$\vec{\rho}^{(2)} = C_2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}. \quad (3.115)$$

Finally, we plug  $\omega_3^2 = k/m(1 + 2m/M)$  into the eigenvalue equation and obtain

$$\begin{aligned} 0 &= \begin{pmatrix} k - k\left(1 + \frac{2m}{M}\right) & -k & 0 \\ -k & 2k - k\frac{M}{m}\left(1 + \frac{2m}{M}\right) & -k \\ 0 & -k & k - k\left(1 + \frac{2m}{M}\right) \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{pmatrix} \\ &= \begin{pmatrix} -k\frac{2m}{M} & -k & 0 \\ -k & -k\frac{M}{m} & -k \\ 0 & -k & -k\frac{2m}{M} \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{pmatrix} \end{aligned} \quad (3.116)$$

Thus, the components of the characteristic vector must satisfy the equations

$$-\frac{2m}{M}\rho_1 - \rho_2 = 0, \quad (3.117)$$

$$-k(\rho_1 + \rho_3) - \frac{M}{m}\rho_2 = 0, \quad (3.118)$$

$$-\rho_2 - \frac{2m}{M}\rho_3 = 0, \quad (3.119)$$

and we have

$$\vec{\rho}^{(3)} = C_3 \begin{pmatrix} 1 \\ -\frac{2m}{M} \\ 1 \end{pmatrix}. \quad (3.120)$$

The vectors  $\rho^{(k)}$  are associated with different frequencies, so they are guaranteed to be orthonormal in the inner product induced by  $\mathbf{T}$ . The normalization condition yields

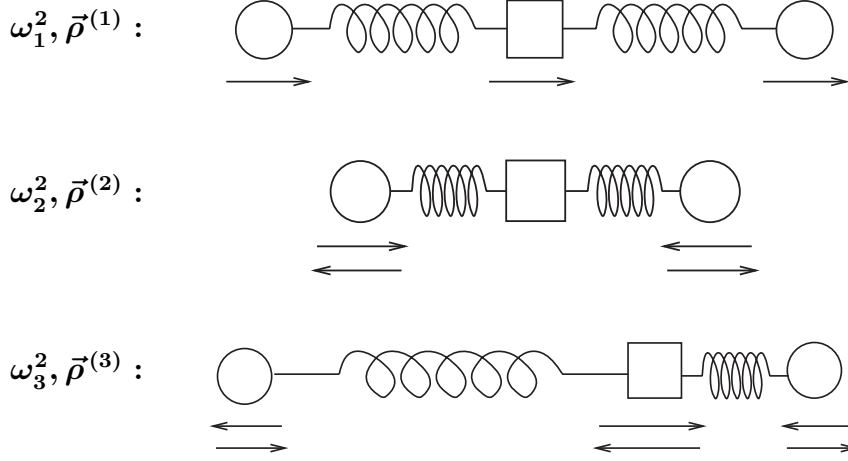
$$\left(\vec{\rho}^{(s)}, \vec{\rho}^{(s)}\right) = \vec{\rho}^{(s)\top} \mathbf{T} \vec{\rho}^{(s)} = m\left(\rho_1^{(s)2} + \rho_3^{(s)2}\right) + M\rho_2^{(s)2} = 1, \quad (3.121)$$

which allows us to determine the normalization constants  $C_s$ :

$$C_1 = \frac{1}{\sqrt{2m(1 + M/2m)}}, C_2 = \frac{1}{\sqrt{2m}}, C_3 = \frac{1}{\sqrt{2m(1 + 2m/M)}}. \quad (3.122)$$

Assembling the modal matrix  $\mathbf{A}$  from the characteristic vectors, we have

$$\mathbf{A} = \left(\vec{\rho}^{(1)} \quad \vec{\rho}^{(2)} \quad \vec{\rho}^{(3)}\right) = \frac{1}{\sqrt{2m}} \begin{pmatrix} \frac{1}{\sqrt{1+M/2m}} & 1 & \frac{1}{\sqrt{1+2m/M}} \\ \frac{1}{\sqrt{1+M/2m}} & 0 & -\frac{2m/M}{\sqrt{1+2m/M}} \\ \frac{1}{\sqrt{1+M/2m}} & -1 & \frac{1}{\sqrt{1+2m/M}} \end{pmatrix} \quad (3.123)$$



**Figure 3.4:** Normal modes of the linear triatomic molecule

and we can compute the normal coordinates

$$\vec{\zeta} = \mathbf{A}^T \mathbf{T} \vec{\eta} \Rightarrow \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{pmatrix} = \frac{1}{\sqrt{2m}} \begin{pmatrix} \frac{1}{\sqrt{1+M/2m}} & \frac{1}{\sqrt{1+M/2m}} & \frac{1}{\sqrt{1+M/2m}} \\ 1 & 0 & -1 \\ \frac{1}{\sqrt{1+2m/M}} & -\frac{2m/M}{\sqrt{1+2m/M}} & \frac{1}{\sqrt{1+2m/M}} \end{pmatrix} \begin{pmatrix} m\eta_1 \\ M\eta_2 \\ m\eta_3 \end{pmatrix} \quad (3.124)$$

The three normal modes of the molecule are illustrated in Fig. 3.4. The mode  $(\omega_1^2, \vec{\rho}^{(1)})$  is actually not vibrational, but corresponds to a rigid uniform translation of the entire molecule. This is reflected in the normal coordinate  $\zeta_1$ : carrying out the matrix-vector product in Eq. (3.124) and plugging in the definition of the displacement coordinates in terms of the initial coordinates  $x_i$ , we have

$$\zeta_1 = \frac{m\eta_1 + M\eta_2 + m\eta_3}{\sqrt{2m + M}} = \frac{mx_1 + Mx_2 + mx_3}{\sqrt{2m + M}} = \sqrt{2m + M} X, \quad (3.125)$$

where  $X$  is the center of mass coordinate of the molecule. The equation of motion for the normal coordinate  $\zeta_1$  therefore becomes

$$\ddot{\zeta}_1 + \underbrace{\omega_1^2}_{=0} \zeta_1 = 0 \Rightarrow \ddot{X} = 0, \quad (3.126)$$

which is of course solved by

$$X(t) = X_0 + V_0 t. \quad (3.127)$$

In mode  $(\omega_2^2, \vec{\rho}^{(2)})$ , the outer masses move in opposite direction from each other while the center mass stays fixed, while in mode  $(\omega_3^2, \vec{\rho}^{(3)})$  the outer masses move in sync in the opposite direction as the center mass.

# Chapter 4

## Central Forces

### 4.1 Introduction

[...] In the present chapter, we will discuss the **central-force problem**. An obvious example is the motion of a mass or particle in a spherically symmetric potential, but the importance of the following discussion extends far beyond this case, because fundamental forces of nature like gravity or electromagnetism are pairwise interactions between objects that only depend on their relative distance. For this reason, we will first discuss the reduction of such pairwise interactions to equivalent one-body problems in relative coordinates.

### 4.2 Reducing Two-Body Problems to Equivalent One-Body Problems

Let us consider a system of two masses  $m_1$  and  $m_2$  whose positions are given by  $\vec{r}_1$  and  $\vec{r}_2$ , respectively. Their center of mass is given by

$$\vec{R} = \frac{1}{M} (m_1 \vec{r}_1 + m_2 \vec{r}_2), \quad (4.1)$$

where the total mass is  $M = m_1 + m_2$ . In the center-of-mass system, the coordinates of the masses are given by

$$\vec{r}'_i = \vec{r}_i - \vec{R}, \quad (4.2)$$

and specifically

$$\vec{r}'_1 = \vec{r}_1 - \left( \frac{m_1}{M} \vec{r}_1 + \frac{m_2}{M} \vec{r}_2 \right) = \frac{m_2}{M} (\vec{r}_1 - \vec{r}_2), \quad (4.3)$$

$$\vec{r}'_2 = \vec{r}_2 - \left( \frac{m_1}{M} \vec{r}_1 + \frac{m_2}{M} \vec{r}_2 \right) = \frac{m_1}{M} (\vec{r}_2 - \vec{r}_1). \quad (4.4)$$

Introducing the relative distance vector

$$\vec{r} \equiv \vec{r}_2 - \vec{r}_1, \quad (4.5)$$

we see that

$$\vec{r}'_2 - \vec{r}'_1 = \frac{m_1}{M} \vec{r} - \frac{m_2}{M} (-\vec{r}) = \frac{m_1 + m_2}{M} \vec{r} = \vec{r}, \quad (4.6)$$

i.e., the relative distance of the two masses does not depend on our choice of coordinate system, which makes sense. We can use  $\vec{r}$  to express the coordinates in the center-of-mass system as

$$\vec{r}'_1 = -\frac{m_2}{M}\vec{r}, \quad \vec{r}'_2 = \frac{m_1}{M}\vec{r}. \quad (4.7)$$

Now we can express the kinetic energy of the two masses in center-of-mass and relative coordinates. First, we have

$$\begin{aligned} T &= \frac{1}{2}m_1\dot{\vec{r}}_1^2 + \frac{1}{2}m_2\dot{\vec{r}}_2^2 \\ &= \frac{1}{2}m_1\left(\dot{\vec{r}}_1 + \dot{\vec{R}}\right)^2 + \frac{1}{2}m_2\left(\dot{\vec{r}}_2 + \dot{\vec{R}}\right)^2 \\ &= \frac{1}{2}m_1\left(\dot{\vec{r}}_1^2 + 2\dot{\vec{r}}_1 \cdot \dot{\vec{R}} + \dot{\vec{R}}^2\right) + \frac{1}{2}m_2\left(\dot{\vec{r}}_2^2 + 2\dot{\vec{r}}_2 \cdot \dot{\vec{R}} + \dot{\vec{R}}^2\right) \\ &= \frac{1}{2}M\dot{\vec{R}}^2 + \left(m_1\dot{\vec{r}}_1 + m_2\dot{\vec{r}}_2\right) \cdot \dot{\vec{R}} + \frac{1}{2}m_1\dot{\vec{r}}_1^2 + \frac{1}{2}m_2\dot{\vec{r}}_2^2 \\ &= \frac{1}{2}M\dot{\vec{R}}^2 + \left(\frac{m_1m_2}{M}\dot{\vec{r}} - \frac{m_2m_1}{M}\dot{\vec{r}}\right) \cdot \dot{\vec{R}} + \frac{1}{2}m_1\dot{\vec{r}}_1^2 + \frac{1}{2}m_2\dot{\vec{r}}_2^2 \\ &= \frac{1}{2}M\dot{\vec{R}}^2 + \frac{1}{2}m_1\dot{\vec{r}}_1^2 + \frac{1}{2}m_2\dot{\vec{r}}_2^2, \end{aligned} \quad (4.8)$$

i.e., the kinetic energy is the sum of the center-of-mass kinetic energy and the kinetic energy of the masses as expressed in the center-of-mass system — the so-called **intrinsic kinetic energy**. This result extends to general  $N$ -particle systems. For the two-body system, the intrinsic term can be rewritten using Eq. (4.7) as

$$T_{\text{intr}} = \frac{1}{2}m_1\frac{m_2^2}{M^2}\dot{\vec{r}}^2 + \frac{1}{2}m_2\frac{m_1^2}{M^2}\dot{\vec{r}}^2 = \frac{1}{2}(m_1 + m_2)\frac{m_1m_2}{M^2}\dot{\vec{r}}^2 \equiv \frac{1}{2}\mu\dot{\vec{r}}^2, \quad (4.9)$$

with the reduced mass

$$\mu \equiv \frac{m_1m_2}{m_1 + m_2} = \frac{m_1m_2}{M}. \quad (4.10)$$

With a potential that only depends on the relative distance of the two masses, the Lagrangian now can be written as

$$L = \frac{1}{2}M\dot{\vec{R}}^2 + \frac{1}{2}\mu\dot{\vec{r}}^2 - V(\vec{r}) = L_{\text{com}} + L_{\text{intr}} \quad (4.11)$$

Since  $L$  only depends on  $\dot{\vec{R}}$  but not on  $\vec{R}$ , the center-of-mass motion decouples from the intrinsic motion, and the Lagrange equations imply

$$\frac{\partial L}{\partial \vec{R}} = 0 = \frac{d}{dt} \frac{\partial L}{\partial \dot{\vec{R}}} \Rightarrow M\dot{\vec{R}} = \text{const.} \quad (4.12)$$

Thus, the center-of-mass motion is uniform and linear. The intrinsic Lagrangian is equivalent to a one-body problem in the relative coordinate  $\vec{r}$ .

## 4.3 Trajectories in the Central-Force Problem

### 4.3.1 General Solution of the Equations of Motion

As we have seen in the previous section, the motion of a particle in a central force field can be modeled by the Lagrangian

$$L = \frac{1}{2}m\dot{\vec{r}}^2 - V(\vec{r}), \quad (4.13)$$

where  $\vec{r}$  can either indicate the coordinates of a single object of mass  $m$  in the external potential  $V(\vec{r})$ , or the relative coordinate of a two-body system with reduced mass  $m = \mu$ .

In general, a central force does not have to be spherically symmetric or conservative, but a conservative force will always be spherically symmetric. To see this, we first assume that the central force is conservative, i.e., that it can be written as

$$\vec{F}(\vec{r}) = -\nabla V(\vec{r}) = -\left(\vec{e}_r \frac{\partial}{\partial r} + \vec{e}_\theta \frac{1}{r} \frac{\partial}{\partial \theta} + \vec{e}_\phi \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi}\right) V(\vec{r}), \quad (4.14)$$

where we have used the gradient in spherical coordinates. Since  $\vec{F}(\vec{r})$  must be parallel to  $\vec{e}_r$ , the partial derivatives of  $V(\vec{r})$  with respect to the angles  $\theta$  and  $\phi$  must vanish, implying that  $V(\vec{r}) = V(|\vec{r}|) = V(r)$ . Conversely, if the force is spherically symmetric,  $\vec{F}(\vec{r}) = f(r)\vec{e}_r$ , then

$$\begin{aligned} \nabla \times \vec{F}(r) &= \left(\vec{e}_r \frac{\partial}{\partial r} + \vec{e}_\theta \frac{1}{r} \frac{\partial}{\partial \theta} + \vec{e}_\phi \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi}\right) \times \frac{f(r)}{r} \vec{r} \\ &= \vec{e}_r \times \left(\frac{f'(r)r - f(r)}{r} \vec{e}_r + \frac{f(r)}{r} \vec{e}_r\right) + \vec{e}_\theta \times \frac{f(r)}{r} \vec{e}_\theta + \vec{e}_\phi \times \frac{f(r)}{r^2 \sin \theta} r \sin \theta \vec{e}_\phi \\ &= 0, \end{aligned} \quad (4.15)$$

where we have used that

$$\frac{\partial \vec{r}}{\partial r} = \vec{e}_r, \quad \frac{\partial \vec{r}}{\partial \theta} = r\vec{e}_\theta, \quad \frac{\partial \vec{r}}{\partial \phi} = r \sin \theta \vec{e}_\phi. \quad (4.16)$$

Combining the proofs in both directions, we find that a central force is conservative if and only if it is spherically symmetric.

Assuming a conservative, spherically symmetric central force going forward, we can write the Lagrangian in spherical coordinates as

$$L = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2 \theta \dot{\phi}^2\right) - V(r). \quad (4.17)$$

The Lagrangian does not depend on  $\phi$ , hence

$$\frac{\partial L}{\partial \phi} = 0 = \frac{d}{dt} \frac{\partial L}{\partial \dot{\phi}} \Rightarrow \frac{\partial L}{\partial \dot{\phi}} = mr^2 \sin^2 \theta \dot{\phi} \equiv l_z = \text{const.}, \quad (4.18)$$

where  $l_z$  denotes the angular momentum around the  $z$  axis of our coordinate system. However, since the Lagrangian is spherically symmetric, we can orient our coordinate system at will, so *all* components of the angular momentum vector must be conserved. Since

$$\vec{l} = \vec{r} \times \vec{p} = \text{const.}, \quad (4.19)$$

the motion of the mass(es) will always be confined to a plane that is perpendicular to  $\vec{l}$ . Choosing our coordinate system such that

$$\vec{l} = l\vec{e}_z, \quad (4.20)$$

we have  $\theta = \frac{\pi}{2}$  in spherical coordinates, and our Lagrangian (4.17) simplifies to

$$L = \frac{1}{2}m \left( \dot{r}^2 + r^2 \underbrace{\dot{\theta}^2}_{=0} + r^2 \underbrace{\sin^2 \theta}_{=1} \dot{\phi}^2 \right) - V(r) = \frac{1}{2}m \left( \dot{r}^2 + r^2 \dot{\phi}^2 \right) - V(r). \quad (4.21)$$

The Lagrange equations now read

$$\begin{aligned} \frac{d}{dt} \frac{\partial L}{\partial \dot{r}} - \frac{\partial L}{\partial r} = 0 &\Rightarrow m\ddot{r} - mr\dot{\phi}^2 + V'(r) = 0, \\ \frac{d}{dt} \frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} = 0 &\Rightarrow mr^2\dot{\phi} = l = \text{const.} \end{aligned} \quad (4.22)$$

Expressing  $\dot{\phi}$  by the conserved angular momentum  $l$ , and introducing  $f(r) = V'(r)$ , we can write the radial equation of motion as

$$m\ddot{r} - \frac{l^2}{mr^3} + f(r) = 0. \quad (4.23)$$

Since the force field we consider here is conservative, the total energy of the moving mass must be conserved. Alternatively, we can note that  $\frac{\partial L}{\partial t} = 0$  since  $L$  does not explicitly depend on time, which implies conservation of the total energy as discussed in Sec. 2.5.4. Starting from the energy conservation law

$$\begin{aligned} E &= T + V \\ &= \frac{1}{2}m \left( \dot{r}^2 + r^2 \dot{\phi}^2 \right) + V(r) \\ &= \frac{1}{2}m\dot{r}^2 + \frac{l^2}{2mr^2} + V(r) = \text{const.}, \end{aligned} \quad (4.24)$$

we can solve for  $\dot{r}$ ,

$$\dot{r} = \frac{dr}{dt} = \pm \sqrt{\frac{2}{m} \left( E - \frac{l^2}{2mr^2} - V(r) \right)}, \quad (4.25)$$

where the sign depends on the initial or boundary conditions, and must be chosen so that  $r \geq 0$  at all times. Separating the variables, we obtain

$$dt = \pm \frac{dr}{\sqrt{\frac{2}{m} \left( E - \frac{l^2}{2mr^2} - V(r) \right)}} \quad (4.26)$$

which can be integrated to give

$$t - t_0 = \pm \int_{r_0}^{r(t)} \frac{dr'}{\sqrt{\frac{2}{m} \left( E - \frac{l^2}{2mr'^2} - V(r') \right)}}. \quad (4.27)$$

Analogously, we can separate the variables in the angular momentum conservation law (4.22) and integrate to obtain

$$\phi - \phi_0 = \int_{t_0}^t \frac{l}{mr(t')^2} dt'. \quad (4.28)$$

Together, Eqs. (4.27) and (4.28) constitute the general solution of any central force problem. As we might guess, these equations can be solved analytically only under very special circumstances. However, we can continue the discussion of the general solutions qualitatively if we introduce the **effective potential**

$$V_{\text{eff}}(r) = \frac{l^2}{2mr^2} + V(r) \quad (4.29)$$

and write Eq. (4.27) as

$$t - t_0 = \pm \int_{r_0}^{r(t)} \frac{dr'}{\sqrt{\frac{2}{m}(E - V_{\text{eff}}(r'))}}. \quad (4.30)$$

In this form, the integral equation is equivalent to the general solution for one-dimensional motion of an object in the effective potential  $V_{\text{eff}}$  - here, this is one-dimensional **radial** motion. The object is performing **angular** motion around the center of the potential due to the angular-momentum dependent term in  $V_{\text{eff}}$ .

Figure (4.1) shows the effective potential for the Kepler problem,

$$V(r) = -\frac{\kappa}{r}, \quad \kappa > 0, \quad (4.31)$$

as well as the main types of radial trajectories as a function of the energy  $E$ . We can distinguish the following cases:

- $E = V_{\text{eff}}(r_{\min})$ : According to the figure, the potential has a global minimum at  $r = r_{\min}$ , and for a trajectory with  $E = V_{\text{eff}}(r_{\min})$ , the entire energy resides in the potential term. Thus, the radial kinetic energy vanishes, which implies that there is no radial motion and the radius of the trajectory stays fixed at all times:

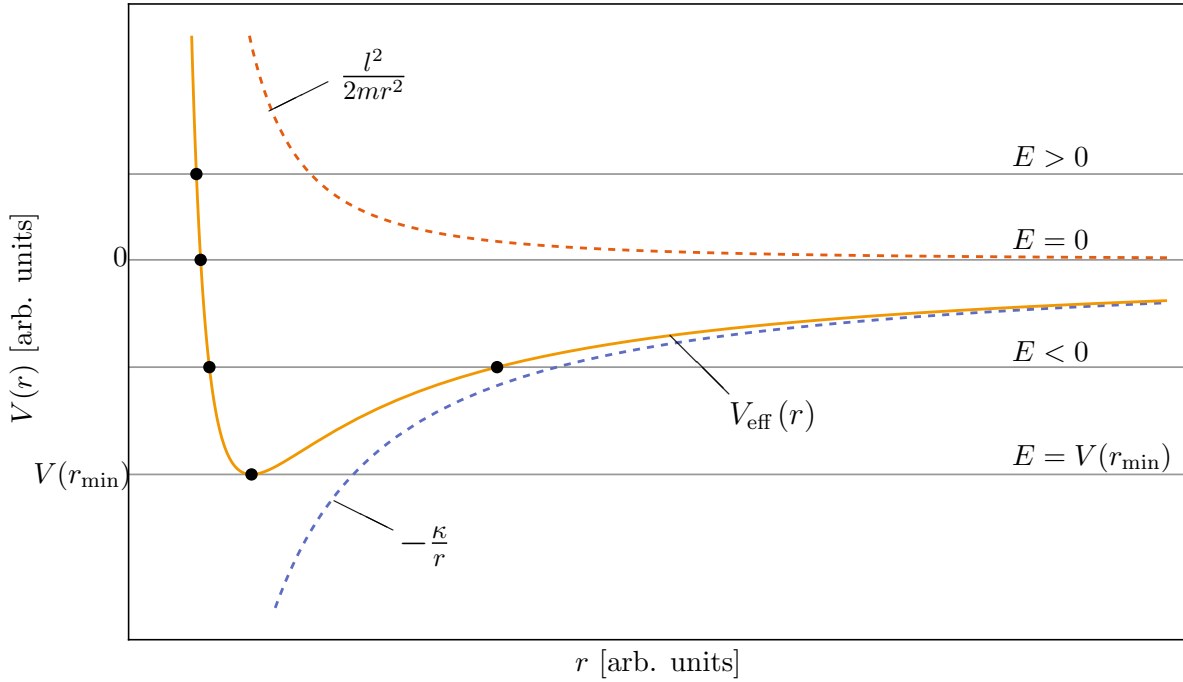
$$T_r = \frac{1}{2}m\dot{r}^2 = 0 \quad \Rightarrow r = r_{\min} = \text{const}. \quad (4.32)$$

This means that the trajectory of the object is a circle around the center of the potential.

We also note that trajectories with  $E < V(r_{\min})$  are physically forbidden because they would require a negative kinetic energy.

- $V(r_{\min}) < E < 0$ : In this case, the trajectory is *bounded* radially, and the distance of the mass from the origin of the potential varies periodically between the turning points of the motion, which are defined by the conditions  $E = V_{\text{eff}}(r_p) = V_{\text{eff}}(r_a)$ . Borrowing terminology that was originally introduced specifically for the Kepler problem, we refer to the turning point at the minimum distances as the **periapsis** (from Greek *peri-*, “near”, and *apsis*, “orbit”) and **apoapsis** (*apo-*, “away from”). At any point  $r_p < r < r_a$ , the difference between  $E$  and  $V_{\text{eff}}(r)$  corresponds to the radial kinetic energy of the motion,

$$\left. \frac{1}{2}m\dot{r}^2 \right|_r = E - V_{\text{eff}}(r). \quad (4.33)$$



**Figure 4.1:** Effective potential as a function of  $r$  and trajectories with fixed energy  $E$ .

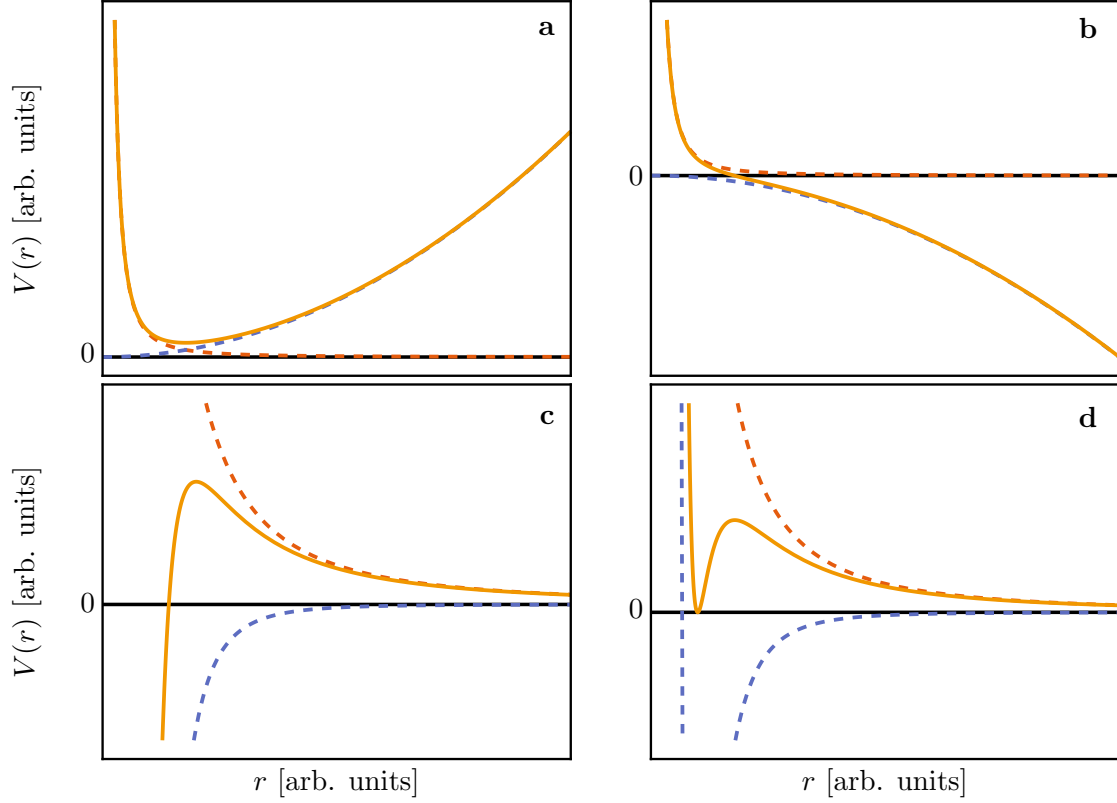
- $E \geq 0$ : The trajectories are *unbounded* in radially direction, i.e., the object approaches the potential center up to some closest distance  $r_p$  that can be determined from  $E = V_{\text{eff}}(r_p)$ , but will then move to  $r \rightarrow \infty$  and outside of the potential's influence, in general. Thus, the motion corresponds to a potential **scattering process** (see Sec. 4.4). The shape of the scattering trajectory depends on the potential, and may be distinct for  $E = 0$  and  $E > 0$ . For the Kepler potential, a trajectory with  $E = 0$  is a parabola, while trajectories with  $E > 0$  are hyperbolas (see Sec. 4.3.5).

It is understood that the nature of the radial trajectories depends on the form of the effective potential, and not all types of orbits need to be possible for each  $V_{\text{eff}}$ . Figure 4.2 shows additional examples of (more or less) physical radial potentials. Figure 4.2a is an isotropic oscillator, which only has bounded solutions, while Fig. 4.2b is an inverted parabolic potential which leads to an effective potential that only admits scattering solutions. Highly singular inverse power law potentials like the one shown in Fig. 4.2c appear in effective theories of the strong interaction, and inverse power laws with alternating signs (Fig. 4.2d) are used to model molecular interactions. The latter two examples are particularly interesting because whether a trajectory with a given energy  $E$  is a bound or scattering solution depends on the region of the potential an object is located.

### 4.3.2 Geometry of Central-Force Trajectories

The general solutions (4.30) and (4.28) allow us to determine the trajectories of objects in central force fields as functions of time by determining  $r(t)$  from (4.30) and plugging it into





**Figure 4.2:** Different types of effective potentials.

Eq. (4.28) to find  $\phi(t)$ . Very often, however, we are only interested in determining the *geometry* of trajectories for a given force field— or, conversely, in finding what type of force generates a particular observed trajectory. For central forces, a unique relationship between the force law and the shape of a trajectory exists, which we want to derive now.

First, consider the time derivatives of the radial coordinate. Using the chain rule, we can write

$$\dot{r} = \frac{dr}{d\phi} \frac{d\phi}{dt} = \frac{dr}{d\phi} \dot{\phi}, \quad (4.34)$$

$$\ddot{r} = \frac{d^2r}{dt^2} = \frac{d^2r}{d\phi^2} \dot{\phi}^2 + \frac{dr}{d\phi} \ddot{\phi}. \quad (4.35)$$

From angular momentum conservation (Eq. (4.22)), we obtain

$$\frac{d}{dt} mr^2 \dot{\phi} = 2mr\dot{r}\dot{\phi} + mr^2\ddot{\phi} = 0 \quad \Rightarrow \quad \ddot{\phi} = -\frac{2}{r}\dot{r}\dot{\phi} = -\frac{2}{r} \frac{dr}{d\phi} \dot{\phi}^2, \quad (4.36)$$

and plugging this into Eq. (4.35), we have

$$\ddot{r} = \frac{d^2r}{d\phi^2} \dot{\phi}^2 - \frac{2}{r} \left( \frac{dr}{d\phi} \right)^2 \dot{\phi}^2. \quad (4.37)$$

The angular momentum conservation law (4.22) also allows us to express  $\dot{\phi}$  in terms of  $l$ ,

$$\dot{\phi}^2 = \frac{l^2}{m^2 r^4} \quad (4.38)$$

and plugging both this relation and Eq. (4.37) into the radial equation of motion, we find

$$f(r) = m\ddot{r} - \frac{l^2}{mr^3} = m \left( \frac{d^2 r}{d\phi^2} - \frac{2}{r} \left( \frac{dr}{d\phi} \right)^2 \right) \frac{l^2}{mr^4} - \frac{l^2}{mr^3} \quad (4.39)$$

and after minor rearrangement

$$f(r) = \left( \frac{d^2 r}{d\phi^2} - \frac{2}{r} \left( \frac{dr}{d\phi} \right)^2 - r \right) \frac{l^2}{mr^4}. \quad (4.40)$$

For a given orbit  $r(\phi)$ , Eq. (4.40) immediately yields the underlying force law.

To determine the geometry of the orbit for a given force or potential, we can derive an integral equation by noting that  $\frac{dr}{d\phi}$  is related to the ratio of the energy and angular momentum conservation laws:

$$\frac{dr}{d\phi} = \frac{\dot{r}}{\dot{\phi}} = \frac{\pm \sqrt{\frac{2}{m} (E - V_{\text{eff}}(r))}}{l/mr^2} = \pm r^2 \frac{\sqrt{2m}}{l} \sqrt{E - V_{\text{eff}}(r)}. \quad (4.41)$$

Separating the variables, we have

$$d\phi = \pm \frac{l dr}{r^2 \sqrt{2m (E - V_{\text{eff}}(r))}}, \quad (4.42)$$

and integrating, we obtain

$$\phi - \phi_0 = \pm \int_{r(\phi_0)}^{r(\phi)} dr' \frac{l}{r'^2 \sqrt{2m (E - V_{\text{eff}}(r'))}} \quad (4.43)$$

### 4.3.3 Stability of Circular Orbits

As discussed in Sec. 4.3.1, a circular orbit is obtained at an extremum of the effective potential

$$V_{\text{eff}}(r) = \frac{l^2}{2mr^2} + V(r). \quad (4.44)$$

Thus, the radius  $R$  of the circular orbit is obtained by solving

$$V'_{\text{eff}}(R) = -\frac{l^2}{mR^3} + V'(R) = 0. \quad (4.45)$$

Whether the extremum is a minimum, maximum, or saddle point can be determined by considering the second derivative, which is given by

$$V''_{\text{eff}}(R) = \frac{3l^2}{mR^4} + V''(R) = \frac{3}{R} V'(R) + V''(R). \quad (4.46)$$

Comparing with the radial equation of motion (cf. Eq. (4.23))

$$m\ddot{r} = \frac{l^2}{mr^3} - V'(r), \quad (4.47)$$

we see that the extremum of the effective potential corresponds to a radial equilibrium position, since

$$m\ddot{R} = 0. \quad (4.48)$$

Naturally, we can only have a circular trajectory with fixed radius  $R$  if there is no acceleration in radial direction.

Let us now consider a perturbation of the circular orbit,

$$r(t) = R + \epsilon(t), \quad (4.49)$$

and plug it into Eq. (4.23). We obtain

$$\begin{aligned} 0 &= m\ddot{\epsilon} - \frac{l^2}{m(R+\epsilon)^3} - f(R+\epsilon) \\ &= m\ddot{\epsilon} - \frac{l^2}{mR^3} \left(1 - 3\frac{\epsilon}{R} + O\left(\frac{\epsilon^2}{R^2}\right)\right) - (f(R) + f'(R)\epsilon + O(\epsilon^2)) = 0 \\ &\approx m\ddot{\epsilon} - \underbrace{\frac{l^2}{mR^3} - f(R)}_{=0} + \left(\frac{3l^2}{mR^4} - f'(R)\right)\epsilon = 0. \end{aligned} \quad (4.50)$$

This is the equation of motion of a harmonic oscillator if

$$\omega^2 \equiv \frac{1}{m} \left( \frac{3l^2}{mR^4} - f'(R) \right) = \frac{1}{m} \left( \frac{3l^2}{mR^4} + V''(R) \right) > 0. \quad (4.51)$$

or

$$\omega^2 = \frac{1}{m} V''_{\text{eff}}(R) > 0. \quad (4.52)$$

Thus, an orbit with radius  $R$  will be stable if  $V_{\text{eff}}$  has a minimum at  $R$ . For  $V(r)$ , the condition (4.51) implies that we must have

$$\frac{3l^2}{mR^4} + V''(R) = \frac{3}{R} V'(R) + V''(R) > 0, \quad (4.53)$$

where we have used Eq. (4.48). In terms of the forces, the inequality reads

$$\frac{3}{R} f(R) + f'(R) < 0. \quad (4.54)$$

### Example: Power-Law Potentials

As an example, we consider power-law potentials of the form

$$V(r) = Ar^n, \quad n \in \mathbb{N}. \quad (4.55)$$

Thus, the radial force field and its derivative are

$$f(r) = -nAr^{n-1}, \quad f'(r) = -n(n-1)Ar^{n-2}. \quad (4.56)$$

Plugging these functions into Eq. (4.54), we find that  $n$  has to satisfy

$$\frac{3}{R}nAR^{n-1} + n(n-1)AR^{n-2} > 0, \quad (4.57)$$

i.e.,

$$3n + n(n-1) > 0 \quad \Rightarrow \quad n > -2. \quad (4.58)$$

Thus, only power-law potentials with  $n > -2$  can support stable circular orbits.

#### 4.3.4 Closed Orbits and Bertrand's Theorem

Let us now discuss the conditions under which an orbit is closed, which will lead us to **Bertrand's theorem**. We will see that only the Kepler and isotropic oscillator potentials allow closed orbits.

We start from the integral equation (4.43) for the trajectory. For a full period of the motion, the orbiting mass moves from the the periapsis  $r_p$  to the apoapsis  $r_a$  and back, advancing by the angle

$$\Delta\phi = \pm 2 \int_{r_p}^{r_a} dr \frac{l}{r^2 \sqrt{2m(E - V_{\text{eff}}(r))}}. \quad (4.59)$$

Thus, an orbit will only be closed if

$$m\Delta\phi = n \cdot 2\pi, \quad m, n \in \mathbb{N}. \quad (4.60)$$

To solve the integral, we substitute

$$u = \frac{1}{r}, \quad du = -\frac{1}{r^2} dr, \quad (4.61)$$

obtaining

$$\Delta\phi = 2 \int_{u_p}^{u_a} du \frac{l}{\sqrt{2m(E - W(u))}}, \quad (4.62)$$

with

$$W(u) \equiv \frac{l^2 u^2}{2m} + V\left(\frac{1}{u}\right). \quad (4.63)$$

We focus on the positive branch for simplicity (the sign only matters if we need to know whether the angular motion is clockwise or counter-clockwise).

Now consider a circular orbit, which occurs at an extremum of the effective potential, as discussed in the Sec. 4.3.3. To describe the near-circular orbit, we write it as a circular orbit with an added perturbation ( $U = 1/R$ ):

$$u(\phi) = U + \epsilon(\phi). \quad (4.64)$$

Note that  $\epsilon(\phi)$  is an inverse length here. Next, we expand the energy difference in the square root denominator. Denoting the energy of the circular orbit by  $E_c$ , we have

$$E - W(u) = E_c + \Delta E - \left( W(U) + W'(U)\epsilon + \frac{1}{2}W''(U)\epsilon^2 \right) = \Delta E - \frac{1}{2}W''(U)\epsilon^2, \quad (4.65)$$

where we have used that

$$E_c = W(U), \quad W'(U) = 0. \quad (4.66)$$

In terms of  $u$ , the relations for the first and second derivatives of the effective potential for circular orbits read

$$W'(U) = \frac{l^2}{m}U - \frac{1}{U^2}V' \left( \frac{1}{U} \right) = \frac{1}{R} \left( \frac{l^2}{m} - R^3V'(R) \right) = 0 \quad (4.67)$$

and

$$W''(U) = \frac{l^2}{m} + \frac{2}{U^3}V' \left( \frac{1}{U} \right) + \frac{1}{U^4}V'' \left( \frac{1}{U} \right) = 3R^3V'(R) + R^4V''(R) = R^4V''_{\text{eff}}(R), \quad (4.68)$$

(recall that derivatives of  $W(u)$  are with respect to  $u$ , those of  $V_{\text{eff}}(r)$  with respect to  $r$ ).

The integral can now be written as

$$\Delta\phi = \frac{2l}{\sqrt{mW''(U)}} \int_{u_p}^{u_a} du \frac{1}{\sqrt{\frac{2\Delta E}{W''(U)} - \epsilon^2}}. \quad (4.69)$$

The perturbation is going to cause oscillatory motion around the circular trajectory which can be parameterized as a function of the *polar angle*  $\phi$  that characterizes the *circular trajectory* instead of time. For each angle  $\phi$  on the circular trajectory, the near-circular orbit will be

$$u(\phi) = U + \epsilon(\phi) = U - \epsilon_0 \cos(\phi), \quad (4.70)$$

where we have made the assumption that

$$u_p = U - \epsilon_0, \quad u_a = U + \epsilon_0. \quad (4.71)$$

From energy conservation for the perturbed orbit at the periapsis, we have

$$\underbrace{E - W(U)}_{=0} + \Delta E - \underbrace{W'(U)(-\epsilon_0)}_{=0} - \frac{1}{2}W''(U)\epsilon_0^2 = 0, \quad (4.72)$$

which means that

$$\epsilon_0^2 = \frac{2\Delta E}{W''(U)}. \quad (4.73)$$

Plugging everything into the integral and changing the integration variable from  $u$  to the angle,

$$\frac{du}{d\phi'} = \epsilon_0 \sin \phi' \quad (4.74)$$

we have

$$\Delta\phi = \frac{2l}{\sqrt{mW''(U)}} \int_0^\pi d\phi' \frac{\epsilon_0 \sin \phi'}{\sqrt{\epsilon_0^2 (1 - \cos^2 \phi')}} = \pm \frac{2l}{\sqrt{mW''(U)}} \int_0^\pi d\phi' = \pm \frac{2\pi l}{\sqrt{mW''(U)}}. \quad (4.75)$$

Using Eq. (4.68), we obtain

$$\Delta\phi = \pm \frac{2\pi l}{R^2 \sqrt{mV''_{\text{eff}}(R)}} = \pm 2\pi \sqrt{\frac{V'(R)}{3V'(R) + RV''(R)}}. \quad (4.76)$$

We see that  $\Delta\phi$  diverges for a saddle point, or becomes complex for a maximum ( $V''_{\text{eff}}(R) < 0$ ) — as discussed in the previous section, the circular orbit would be unstable in these cases, any perturbation would grow exponentially and prevent a closed orbit.

According to the closed-orbit condition (4.60), we must have

$$\Delta\phi = \pm 2\pi \sqrt{\frac{V'(R)}{3V'(R) + RV''(R)}} = \frac{n}{m} \cdot 2\pi. \quad (4.77)$$

Squaring and inverting the equation, we have

$$\frac{V'(R)}{3V'(R) + RV''(R)} = \frac{n^2}{m^2}. \quad (4.78)$$

Switching to forces and rearranging, we obtain

$$\frac{R}{f(R)} f'(R) = \frac{m^2}{n^2} - 3 \equiv \beta^2 - 3. \quad (4.79)$$

Since we can vary  $R$  smoothly by varying the angular momentum  $l$  (cf. Eq. (4.45)) this equation should hold for general  $r$ , so we obtain the differential equation

$$\frac{r}{f(r)} f'(r) = \frac{r}{f} \frac{df}{dr} = \beta^2 - 3. \quad (4.80)$$

Separating the variables, we have

$$\frac{1}{f} df = (\beta^2 - 3) \frac{1}{r} dr, \quad (4.81)$$

and integrating, we obtain

$$\ln f(r) - \ln f(r_0) = (\beta^2 - 3) (\ln r - \ln r_0). \quad (4.82)$$

We can rewrite this as

$$\ln \frac{f(r)}{f(r_0)} = (\beta^2 - 3) \ln \frac{r}{r_0} = \ln \left( \frac{r}{r_0} \right)^{\beta^2 - 3} \quad (4.83)$$

and exponentiate, which yields

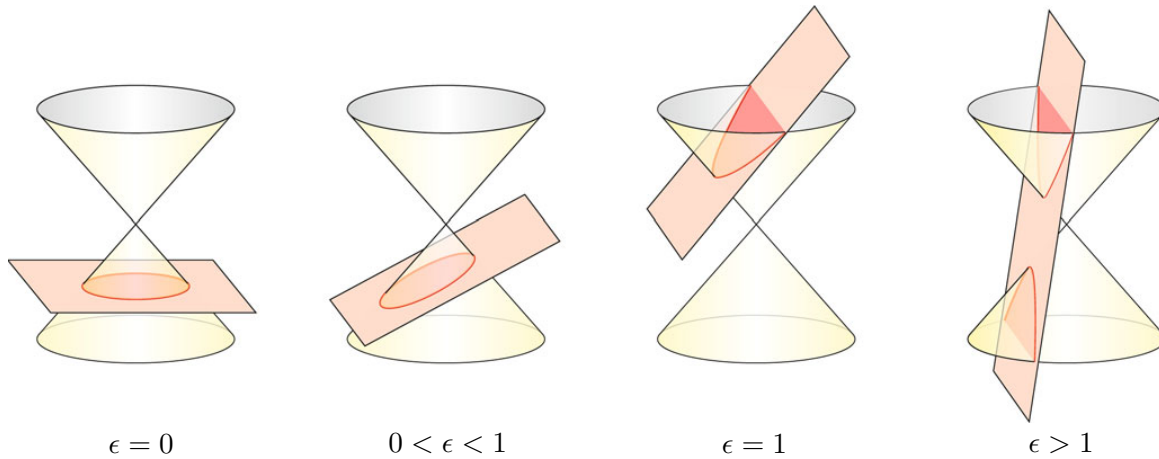
$$\frac{f(r)}{f(r_0)} = \left( \frac{r}{r_0} \right)^{\beta^2 - 3}. \quad (4.84)$$

Thus, only power-law forces of the form

$$f(r) = Ar^{\beta^2 - 3} \quad (4.85)$$

will satisfy the conditions for closed orbits. This means that the potential either has a power-law form as well,

$$V_\alpha(r) \equiv -\frac{A}{\alpha} r^\alpha \equiv -\frac{A}{\beta^2 - 2} r^{\beta^2 - 2}, \quad (4.86)$$



**Figure 4.3:** Trajectories of the Kepler problem corresponding to conic sections.

or it is logarithmic

$$V_0(r) \equiv A \ln \frac{r}{r_0} \quad (4.87)$$

where we have ignored irrelevant constants. Note that for the power-law potentials, we have  $\alpha > -2$  for a stable near-circular orbit, which agrees with our result from Sec. 4.3.3.

Plugging the logarithmic potential into Eq. (4.76), we obtain

$$\Delta\phi = \pm 2\pi \sqrt{\frac{A/R}{3A/R - RA/R^2}} = \pm 2\pi \frac{1}{\sqrt{2}}, \quad (4.88)$$

which is not rational and will therefore not yield a closed orbit. For the power-law potentials, we find

$$\Delta\phi = \pm 2\pi \sqrt{\frac{-AR^{\alpha-1}}{-3AR^{\alpha-1} - (\alpha-1)AR^{\alpha-1}}} = \pm 2\pi \sqrt{\frac{1}{2+\alpha}}. \quad (4.89)$$

[Finish argument.]

### 4.3.5 The Kepler Problem

[...] The effective potential for the Kepler problem is shown in Fig. 4.1. [...]

### 4.3.6 Perihelion Precession from General Relativity

[...]

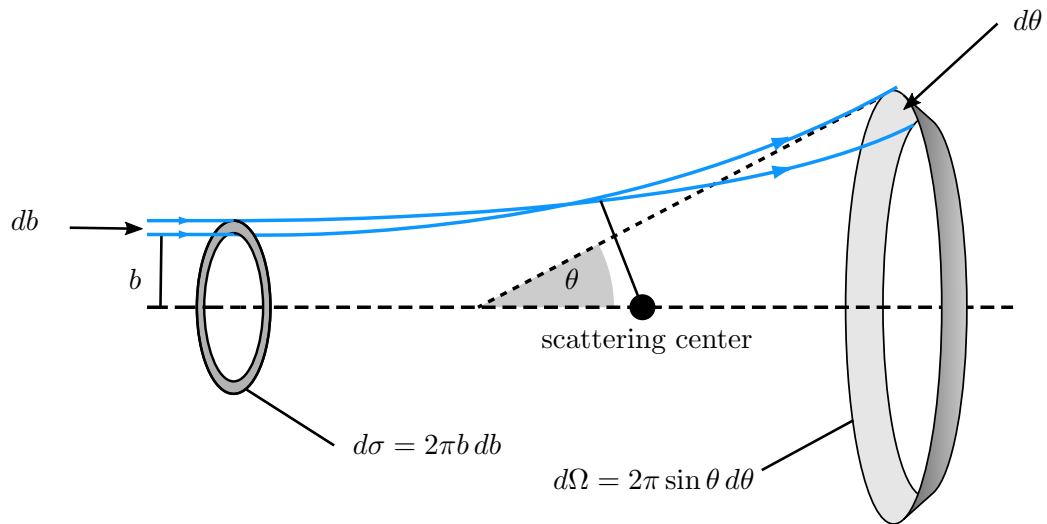


Figure 4.4: Differential cross section.

## 4.4 Scattering

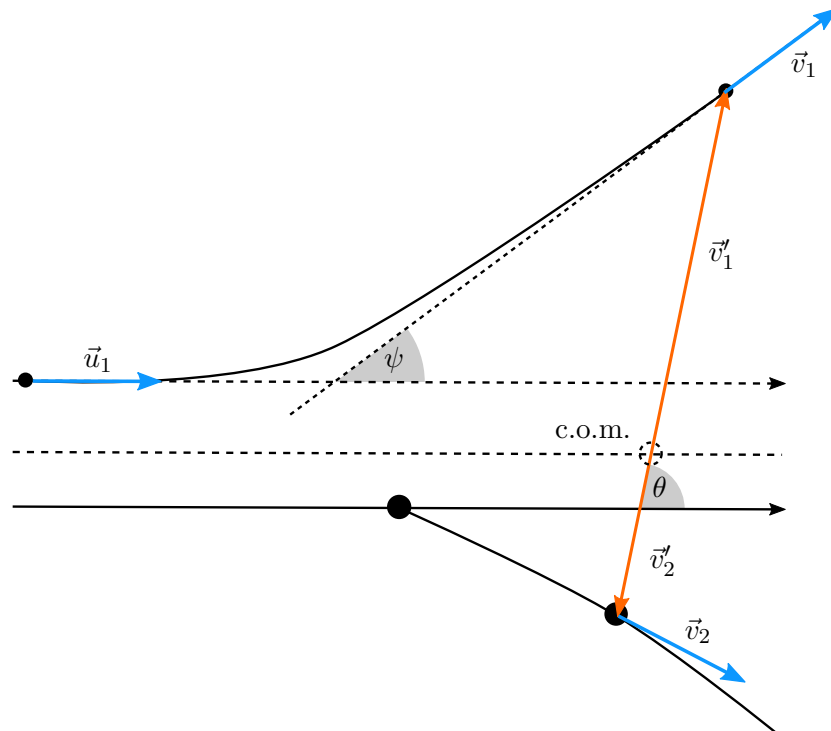
## 4.5 Orbital Dynamics

### 4.5.1 Transfer Orbits

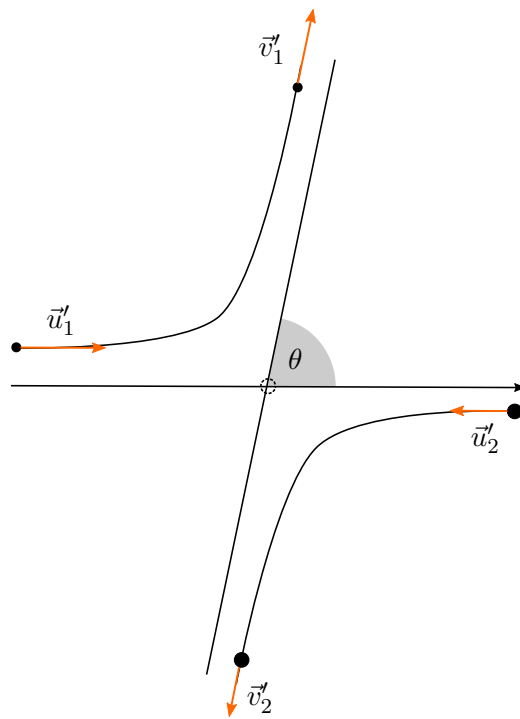
### 4.5.2 Gravitational Slingshot

## 4.6 The Three-Body Problem

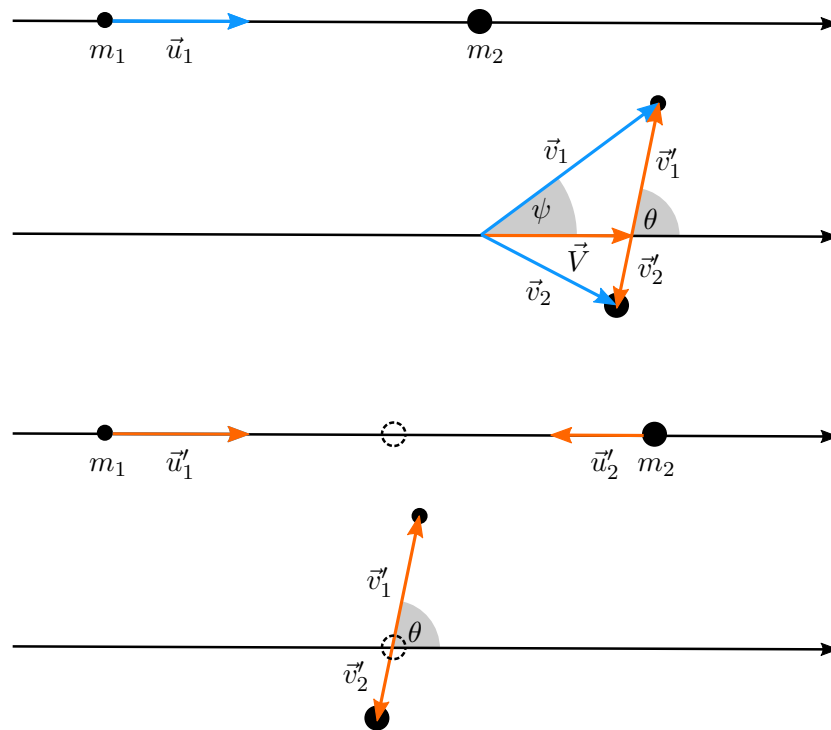




**Figure 4.5:** Scattering in the laboratory frame.



**Figure 4.6:** Scattering in the center-of-mass frame.



**Figure 4.7:** Scattering in the center-of-mass frame.

## Chapter 5

# The Rigid Body

## Chapter 6

# Hamiltonian Mechanics

### 6.1 The Hamiltonian

[cf. 2.5.4...] The **Hamiltonian** is defined in terms of the canonical coordinates and their conjugate momenta as

$$H(q_i, p_i, t) \equiv \sum_{i=1}^n (p_i \cdot \dot{q}_i(q_i, p_i, t)) - \bar{L}(q_i, p_i, t), \quad (6.1)$$

where  $\bar{L}(q_i, p_i, t)$  is the Lagrangian after a change of variables from  $\dot{q}_i$  to the  $p_i$ .

# Appendix A

## Lagrange Multipliers

Lagrange multipliers are a convenient choice for performing optimizations under constraints.  
[Fill in...]

### A.1 General Procedure

#### A.1.1 The Second-Derivative Test

### A.2 Examples

#### A.2.1 Rectangle Inscribed in an Ellipse

Consider the problem of inscribing the largest possible rectangle into an ellipse with semi-major axes  $a$  and  $b$ , which is defined by the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \quad (\text{A.1})$$

The area of the rectangle with corners  $(-x, -y)$ ,  $(-x, y)$ ,  $(x, -y)$ ,  $(x, y)$  is given by the function

$$A(x, y) = (2x)(2y) = 4xy, \quad x, y \geq 0. \quad (\text{A.2})$$

We introduce a Lagrange multiplier and couple the constraint (A.1) to this function,

$$\tilde{A}(x, y, \lambda) = A(x, y) + \lambda\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1\right), \quad (\text{A.3})$$

and proceed to maximize  $\tilde{A}$ . The partial derivatives are

$$\frac{\partial \tilde{A}}{\partial x} = 4y + 2\lambda \frac{x}{a^2}, \quad (\text{A.4})$$

$$\frac{\partial \tilde{A}}{\partial y} = 4x + 2\lambda \frac{y}{b^2}, \quad (\text{A.5})$$

$$\frac{\partial \tilde{A}}{\partial \lambda} = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1. \quad (\text{A.6})$$

Let us first consider the boundaries of our domain, which are given by  $(x = 0, y)$  and  $(x, y = 0)$ , respectively. The cases corresponds to the limit in which the rectangle turns into a line. While the are vanishes, the constraint can still be satisfied. On the first boundary, for example, we have

$$\frac{\partial \tilde{A}}{\partial \lambda} = \frac{0^2}{a^2} + \frac{y^2}{b^2} - 1 = 0 \quad \Rightarrow \quad y = \pm b. \quad (\text{A.7})$$

Our system of equations now becomes

$$\left. \frac{\partial \tilde{A}}{\partial x} \right|_{x=0} = 0 \stackrel{!}{=} 0, \quad (\text{A.8})$$

$$\left. \frac{\partial \tilde{A}}{\partial y} \right|_{x=0} = 2\lambda \frac{y}{b^2} \stackrel{!}{=} 0, \quad (\text{A.9})$$

$$\left. \frac{\partial \tilde{A}}{\partial \lambda} \right|_{x=0} = \frac{y^2}{b^2} - 1 \stackrel{!}{=} 0. \quad (\text{A.10})$$

The first equation is trivially true, while the second and third equations are solved by

$$x_1 = 0, \quad y_1 = b, \quad \lambda_1 = 0. \quad (\text{A.11})$$

Analogously, we obtain

$$x_2 = a, \quad y_2 = 0, \quad \lambda_2 = 0. \quad (\text{A.12})$$

on the other boundary. Mathematically, these are global minima of  $\tilde{A}(x, y, \lambda)$  and  $A(x, y)$ .

Now consider the interior of the domain, where  $x > 0, y > 0$ . We start with Eq. (A.5) and obtain

$$4x + 2\lambda \frac{y}{b^2} = 0 \quad \Rightarrow \quad x = -\lambda \frac{y}{2b^2}. \quad (\text{A.13})$$

Note that this implies  $\lambda < 0$ , since  $y$  must be positive. Plugging this solution into Eq. (A.4), we obtain

$$4y - \lambda^2 \frac{y}{a^2 b^2} = 0 \quad \Rightarrow \quad y \left( 4 - \frac{\lambda^2}{a^2 b^2} \right) = 0. \quad (\text{A.14})$$

The solutions to this equation are  $y = 0$ , which we considered separately above, and  $\lambda = -2ab$  ( $\lambda = 2ab$  is ruled out).

Let's proceed with  $\lambda = -2ab$ , which yields

$$x = -\lambda \frac{y}{2b^2} = y \frac{a}{b^2}. \quad (\text{A.15})$$

Using this relation in Eq. (A.6), we have

$$\frac{y^2 a^2}{b^2 a^2} + \frac{y^2}{b^2} - 1 = 0 \quad \Rightarrow \quad y = \frac{b}{\sqrt{2}}. \quad (\text{A.16})$$

Thus, the extremum in the domain's interior is

$$E_3 : \quad x_3 = \frac{a}{\sqrt{2}}, \quad y_3 = \frac{b}{\sqrt{2}}, \quad \lambda_3 = -2ab. \quad (\text{A.17})$$

The area of the resulting rectangle is

$$A(x_3, y_3) = 2ab, \quad (\text{A.18})$$

and the constrained stationary point is obviously a maximum.

Although the nature of the extremum is clear, let us still perform the **second-derivative test** for practice. The bordered Hessian reads

$$\mathcal{H}(x, y, \lambda) = \begin{pmatrix} \frac{\partial^2 \tilde{A}}{\partial \lambda^2} & \frac{\partial^2 \tilde{A}}{\partial \lambda \partial x} & \frac{\partial^2 \tilde{A}}{\partial \lambda \partial y} \\ \frac{\partial^2 \tilde{A}}{\partial x \partial \lambda} & \frac{\partial^2 \tilde{A}}{\partial x^2} & \frac{\partial^2 \tilde{A}}{\partial x \partial y} \\ \frac{\partial^2 \tilde{A}}{\partial y \partial \lambda} & \frac{\partial^2 \tilde{A}}{\partial x \partial y} & \frac{\partial^2 \tilde{A}}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 0 & \frac{2x}{a^2} & \frac{2y}{b^2} \\ \frac{2x}{a^2} & \frac{2\lambda}{a^2} & 4 \\ \frac{2y}{b^2} & 4 & \frac{2\lambda}{b^2} \end{pmatrix}. \quad (\text{A.19})$$

We have  $n = 2$  variables and  $m = 1$  constraint, so **the second-derivative test for constrained optimization** requires us to look at the principal minors of the bordered Hessian for  $k = \min(2m + 1, m + n), \dots, m + n$ . Here, this means  $k = 3$ , and the principal minor is  $\mathcal{H}$  itself. At the extremum  $E_3$ , we have

$$\det \mathcal{H} \left( \frac{a}{\sqrt{2}}, \frac{b}{\sqrt{2}}, -2ab \right) = \det \begin{pmatrix} 0 & \frac{\sqrt{2}}{a} & \frac{\sqrt{2}}{b} \\ \frac{\sqrt{2}}{a} & -\frac{4b}{a} & 4 \\ \frac{\sqrt{2}}{b} & 4 & -\frac{4a}{b} \end{pmatrix} = \frac{32}{ab} > 0, \quad (\text{A.20})$$

so the test indeed correctly identifies the extremum as a constrained maximum.

## A.2.2 Extrema of the Mexican Hat Potential

### Extrema Without Constraints

Let us consider the function

$$V(x, y) = -40(x^2 + y^2) + (x^2 + y^2)^2, \quad (\text{A.21})$$

an example of a so-called *mexican hat* potential that is frequently used to discuss symmetry breaking phenomena in physics (see Fig. A.1). Let us first compute the extrema of  $V(x, y)$ . The partial derivatives with respect to the variables are

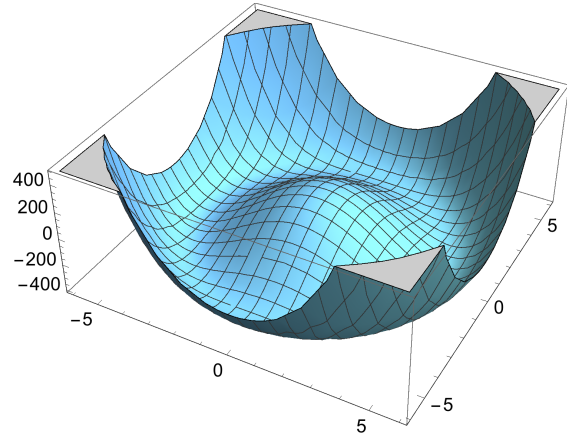
$$\frac{\partial V}{\partial x} = -80x + 4(x^2 + y^2)x, \quad (\text{A.22})$$

$$\frac{\partial V}{\partial y} = -80y + 4(x^2 + y^2)y. \quad (\text{A.23})$$

The extrema are obtained by solving

$$0 = -80x + 4(x^2 + y^2)x = x(4(x^2 + y^2) - 80), \quad (\text{A.24})$$

$$0 = -80y + 4(x^2 + y^2)y = y(4(x^2 + y^2) - 80). \quad (\text{A.25})$$



**Figure A.1:** Mexican hat potential.

The solutions of Eq. (A.24) are

$$x_1 = 0 \quad (\text{A.26})$$

and all points on a circle in the  $xy$ -plane with radius  $r = \sqrt{20}$ .

$$4(x^2 + y^2) - 80 = 0 \quad \Rightarrow \quad x^2 + y^2 = 20. \quad (\text{A.27})$$

From Eq. (A.25), we again obtain the definition of the circle, as well as

$$y_1 = 0. \quad (\text{A.28})$$

By inspecting the figure, we see that  $(x_1, y_1) = (0, 0)$  is a **local maximum**, while points on the circle are **degenerate global minima** in radial direction, and **saddle points** in any direction tangential to the circle. If we do not have a figure at hand, or a more complicated function, we can compute the Hessian and check its definiteness:

$$\mathcal{H}(x, y) = \begin{pmatrix} \frac{\partial^2 V}{\partial x^2} & \frac{\partial^2 V}{\partial x \partial y} \\ \frac{\partial^2 V}{\partial x \partial y} & \frac{\partial^2 V}{\partial y^2} \end{pmatrix} = \begin{pmatrix} -80 + 4(x^2 + y^2) + 8x^2 & 8xy \\ 8xy & -80 + 4(x^2 + y^2) + 8y^2 \end{pmatrix}. \quad (\text{A.29})$$

For  $(x_1, y_1) = (0, 0)$ ,  $\mathcal{H}(0, 0)$  is diagonal and we can read off the doubly degenerate eigenvalue  $h_{1/2} = -80$ . Since all the eigenvalues are negative, the  $\mathcal{H}(0, 0)$  is negative definite, the point is a local maximum.

For any point on the ring, the Hessian becomes

$$\mathcal{H}(x, y)|_C = \begin{pmatrix} -80 + 4 \cdot 20 + 8x^2 & 8xy \\ 8xy & -80 + 4 \cdot 20 + 8y^2 \end{pmatrix} = \begin{pmatrix} 8x^2 & 8xy \\ 8xy & 8y^2 \end{pmatrix}. \quad (\text{A.30})$$

The eigenvalues are  $h_3 = 0$  and  $h_4 = 8(x^2 + y^2) = 160$ , so the matrix is indefinite. Parameterizing the circle by  $(r, \phi) = (\sqrt{20}, \phi)$ , we find the expected result that  $\vec{e}_3 = \vec{e}_\phi = (-\sin \phi, \cos \phi)^T$  are eigenvectors associated with  $h_3$ , and  $\vec{e}_4 = \vec{e}_r = (\cos \phi, \sin \phi)^T$  are eigenvectors associated with  $h_4$ .

### Extrema Under Constraints

Let us now determine the extrema of the mexican hat potential under the constraint

$$y = x \quad \Rightarrow \quad f(x, y) \equiv y - x = 0. \quad (\text{A.31})$$

We couple the constraint to the function  $V(x, y)$  with a Lagrange multiplier  $\lambda$ :

$$\tilde{V}(x, y, \lambda) = V(x, y) + \lambda f(x, y). \quad (\text{A.32})$$

The partial derivatives of  $\tilde{V}$  and the resulting conditions for extrema are

$$\frac{\partial \tilde{V}}{\partial x} = -80x + 4(x^2 + y^2)x - \lambda = 0, \quad (\text{A.33})$$

$$\frac{\partial \tilde{V}}{\partial y} = -80y + 4(x^2 + y^2)y + \lambda = 0, \quad (\text{A.34})$$



$$\frac{\partial \tilde{V}}{\partial \lambda} = y - x = 0. \quad (\text{A.35})$$

Plugging the solution of (A.35) obviously into either of the other equations, e.g., Eq. (A.33), we can determine  $\lambda$ :

$$-80x + 8x^3 - \lambda = 0 \quad \Rightarrow \quad \lambda = 8x^3 - 80x. \quad (\text{A.36})$$

The remaining equation now becomes

$$-80x + 8x^3 + 8x^3 - 80x = 16x(x^2 - 10) = 0, \quad (\text{A.37})$$

with the solutions

$$x_1 = 0, x_{2/3} = \pm\sqrt{10}, \quad (\text{A.38})$$

so our extrema are formally given by

$$(x_1, y_1, \lambda_1) = (0, 0, 0), \quad (x_2, y_2, \lambda_2) = (\sqrt{10}, \sqrt{10}, 0), \quad (x_3, y_3, \lambda_3) = (-\sqrt{10}, -\sqrt{10}, 0). \quad (\text{A.39})$$

**(Note: in general, the Lagrange multiplier(s) need not vanish at the stationary points; see App. A.2.1.)**

The bordered Hessian of  $\tilde{V}$  now reads

$$\mathcal{H}(x, y, \lambda) = \begin{pmatrix} 0 & -1 & 1 \\ -1 & -80 + 8x^2 + 4(x^2 + y^2) & 8xy \\ 1 & 8xy & -80 + 8x^2 + 4(x^2 + y^2) \end{pmatrix} \quad (\text{A.40})$$

and it must be evaluated at the different stationary points. We have  $n = 2$  variables and  $m = 1$  constraints, so **the second-derivative test for constrained optimization discussed above** requires us to look at the principal minors  $\mathcal{H}$  for  $k = \min(2m + 1, m + n), \dots, m + n$ . In the present case, the range gives us  $k = 3$  and the principal minor is  $\mathcal{H}$  itself.

At  $(x, y, \lambda) = (0, 0, 0)$ , we have

$$\det \mathcal{H}(0, 0, 0) = \det \begin{pmatrix} 0 & -1 & 1 \\ -1 & -80 & 0 \\ 1 & 0 & -80 \end{pmatrix} = 160 > 0, \quad (\text{A.41})$$

so the point is a constrained maximum **[sign rule]**. At  $(\sqrt{10}, \sqrt{10}, 0)$  and  $(-\sqrt{10}, -\sqrt{10}, 0)$  we obtain the same bordered Hessian, and its determinant is

$$\det \mathcal{H}(\pm\sqrt{10}, \pm\sqrt{10}, 0) = \det \begin{pmatrix} 0 & -1 & 1 \\ -1 & 80 & 80 \\ 1 & 80 & 80 \end{pmatrix} = -320 < 0, \quad (\text{A.42})$$

implying that these points are constrained minima **[sign rule]**. This makes sense, of course: We precisely get the two global minima and the local maxima of  $V(x, y)$  that are compatible with the constraint  $y - x = 0$ .

# Bibliography

- [1] B. L. Beers, *American Journal of Physics* **40**, 1636 (1972).
- [2] C. Lanczos, *The Variational Principles of Mechanics*, 4th ed. (Dover Publications, 1949).
- [3] J. Coopersmith, *The Lazy Universe* (Oxford University Press, 2017).
- [4] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*, 3rd ed. (Addison-Wesley, 2002).
- [5] M. R. Flannery, *American Journal of Physics* **73**, 265 (2005).
- [6] M. R. Flannery, *American Journal of Physics* **79**, 932 (2011).
- [7] M. R. Flannery, *Journal of Mathematical Physics* **52**, 032705 (2011).
- [8] E. Noether, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* **1918**, 235 (1918).
- [9] E. Noether, *Transport Theory and Statistical Physics* **1**, 186 (1971).
- [10] A. I. Lurie, *Analytical Mechanics* (Springer, 2002).
- [11] E. Minguzzi, *European Journal of Physics* **36**, 035014 (2015).