Lecture Notes

Classical Mechanics

PHYS 571

Prepared By:

Fatma Alhazmi

Lecturer:

Prof. Hocine Bahlouli

Academic Year: 2024–2025

Department of Physics

King Fahd University of Petroleum and Minerals (KFUPM)

Contents

Chapter 1

Survey of the Elementary Principles

Elementary classical mechanics treats the motion of particles and rigid bodies in regimes where quantum and relativistic effects are negligible: macroscopic length scales, speeds $v \ll c$ with $c \approx 3 \times 10^8\,\mathrm{m/s}$, and kinetic energies $\frac{1}{2}mv^2 \ll mc^2$. In this idealized Newtonian picture, specifying initial position and velocity determines a unique trajectory (deterministic dynamics).

While linear systems admit superposition, many mechanical systems are intrinsically nonlinear, and even very simple ones can exhibit sensitive dependence on initial conditions (chaos). Such behavior is discussed in standard references (e.g., Chapter 11 of Ref. [1]); it is not part of this course's syllabus.

1.1. Mechanics of a Single Particle

A **particle** is an idealized point mass m. Relative to a chosen inertial frame, it has

- a **position vector** $\mathbf{r}(t)$,
- a **velocity** $\mathbf{v}(t) = \dot{\mathbf{r}}(t)$.

We summarize the kinematics and dynamics as

$$\mathbf{v}(t) = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}, \tag{1.1}$$

$$\mathbf{a}(t) = \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2}, \tag{1.2}$$

$$\mathbf{p}(t) = m\,\mathbf{v}(t), \tag{1.3}$$

$$\mathbf{F}_{\mathrm{ext}}(t) = \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} \overset{m\,\mathrm{const.}}{=} m\,\mathbf{a}(t), \tag{1.4}$$

where (1.4) is **Newton's second law** for constant mass.

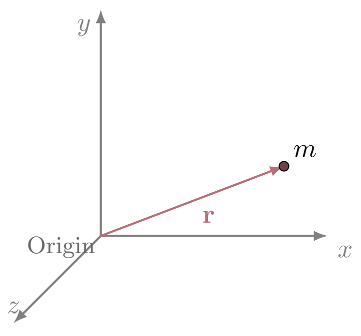


Figure 1.1: A particle of mass m at position $\mathbf{r}$ in an inertial frame.

This is a second-order ordinary differential equation in time. Strictly speaking, *inertial* frames are non-accelerating frames in which Newton's laws hold in their simplest form. If we imagine a fixed star as a reference, Newton's laws are valid in any frame moving with constant velocity relative to that star (Galilean inertial frames). Earth-fixed frames are only approximately inertial.

Remark 1.1 – Inertial frames on Earth

Frames attached to Earth's surface are rotating; over short durations and modest length scales they are often treated as inertial, but over long times or large scales (e.g., Foucault pendulum) rotation-induced fictitious forces (Coriolis, centrifugal) become relevant.

1.2. Conservation Laws

A quantity $Q(t)$ is **conserved** if

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = 0 \quad \Longrightarrow \quad Q(t) = \text{constant}. \tag{1.5}$$

1.2.1. Conservation of Linear Momentum

From (1.4),

$$\mathbf{F}_{\text{net}} = \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t}. \tag{1.6}$$

If $\mathbf{F}_{\text{net}} = \mathbf{0}$, then $\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = \mathbf{0}$ and $\mathbf{p}$ is constant (linear momentum is conserved).

1.2.2. Conservation of Angular Momentum

Define the **angular momentum**

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} = m\,\mathbf{r} \times \mathbf{v}. \tag{1.7}$$

Its rate of change (for a particle) is

$$\begin{aligned}\frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{r} \times \mathbf{p}) = \mathbf{v} \times \mathbf{p} + \mathbf{r} \times \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} \\ &= \mathbf{r} \times \mathbf{F}_{\text{ext}} \equiv \boldsymbol{\tau}_{\text{ext}},\end{aligned} \tag{1.8}$$

since $\mathbf{v} \times \mathbf{p} = \mathbf{v} \times (m\mathbf{v}) = \mathbf{0}$. If $\boldsymbol{\tau}_{\text{ext}} = \mathbf{0}$, then $\mathbf{L}$ is conserved.

1.2.3. Work, Energy, and Conservative Forces

The work done by a force along a path $\mathcal{C}$ from 1 to 2 is

$$W_{1\to 2} = \int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{s}. \tag{1.9}$$

Parametrizing by time, $\mathrm{d}\mathbf{s} = \mathbf{v}(t)\,\mathrm{d}t$ and, for a single particle, $\mathbf{F} = m\,\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t}$, hence

$$\begin{aligned}W_{1\to 2} &= \int_{t_1}^{t_2} m\,\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} \cdot \mathbf{v}\,\mathrm{d}t = \frac{m}{2}\int_{t_1}^{t_2} \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{v}\cdot\mathbf{v})\,\mathrm{d}t \\ &= \frac{m}{2}\left(v_2^2 - v_1^2\right) \equiv T_2 - T_1 = \Delta T,\end{aligned} \tag{1.10}$$

where $T = \frac{1}{2}mv^2$ is the kinetic energy (work–energy theorem).
A force is **conservative** if the work between two points is path independent, equivalently

$$\oint_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{s} = 0 \quad \text{for every closed curve } \mathcal{C}. \tag{1.11}$$

By Stokes' theorem, for any surface $\mathcal{S}$ with boundary $\partial\mathcal{S} = \mathcal{C}$,

$$\oint_{\partial\mathcal{S}} \mathbf{F} \cdot \mathrm{d}\mathbf{s} = \iint_{\mathcal{S}} (\boldsymbol{\nabla} \times \mathbf{F}) \cdot \mathrm{d}\mathbf{a}. \tag{1.12}$$

Thus $\boldsymbol{\nabla} \times \mathbf{F} = \mathbf{0}$ implies (1.11); on simply connected domains this is equivalent to the existence of a scalar potential V with

$$\mathbf{F} = -\boldsymbol{\nabla} V(\mathbf{r}). \tag{1.13}$$

For such forces and time-independent V, mechanical energy $E = T + V$ is conserved.

Remark 1.2 – Curl-free fields and topology

The condition $\nabla \times \mathbf{F} = \mathbf{0}$ guarantees a potential *locally*. Global path independence requires the domain to be simply connected. Non-conservative forces (e.g. kinetic friction) yield $\oint \mathbf{F} \cdot d\mathbf{s} < 0$, dissipating mechanical energy.

See here an example of using conservation laws in orbital dynamics.

1.3. System of Particles

Consider a system of N particles labeled by $i = 1, 2, \ldots, N$, each with a constant mass m_i. Newton's second law for the i-th particle is

$$m_i \, \ddot{\mathbf{r}}_i = \mathbf{F}_i, \tag{1.14}$$

where $\mathbf{F}_i$ is the total force acting on the i-th particle.
The force $\mathbf{F}_i$ can be decomposed into two components,

$$\mathbf{F}_i = \mathbf{F}_i^{\text{ext}} + \sum_{\substack{j=1 \\ j \neq i}}^{N} \mathbf{F}_{ij}, \tag{1.15}$$

where

- $\mathbf{F}_i^{\text{ext}}$ is the external force acting on particle i,
- $\mathbf{F}_{ij}$ is the internal force on i due to particle j.

If the internal forces satisfy the weak form of Newton's third law,

$$\mathbf{F}_{ij} = -\,\mathbf{F}_{ji}, \tag{1.16}$$

then the total contribution of internal forces vanishes:

$$\sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \mathbf{F}_{ij} = \mathbf{0}. \tag{1.17}$$

Summing Newton's second law over all particles, we find

$$\sum_{i=1}^{N} m_i \, \ddot{\mathbf{r}}_i = \sum_{i=1}^{N} \mathbf{F}_i = \sum_{i=1}^{N} \mathbf{F}_i^{\text{ext}} \equiv \mathbf{F}^{\text{ext}}, \tag{1.18}$$

where $\mathbf{F}^{\text{ext}}$ is the total external force acting on the system.

1.3.1. Center of Mass Motion

The total mass of the system is

$$M = \sum_{i=1}^{N} m_i. \tag{1.19}$$

The position of the center of mass (C.M.) is defined as

$$\mathbf{R} = \frac{1}{M} \sum_{i=1}^{N} m_i \, \mathbf{r}_i. \tag{1.20}$$

Differentiating twice with respect to time and using (1.18), we find

$$M \, \ddot{\mathbf{R}} = \sum_{i=1}^{N} m_i \, \ddot{\mathbf{r}}_i = \mathbf{F}^{\text{ext}}. \tag{1.21}$$

Thus, the motion of the center of mass is governed by

$$M\,\ddot{\mathbf{R}} = \mathbf{F}^{\text{ext}}, \tag{1.22}$$

i.e., the center of mass moves as if all the mass were concentrated at $\mathbf{R}$ and acted upon by the net external force $\mathbf{F}^{\text{ext}}$.
If $\mathbf{F}^{\text{ext}} = \mathbf{0}$, the center of mass moves with constant velocity:

$$\ddot{\mathbf{R}} = \mathbf{0} \quad \Longrightarrow \quad \mathbf{V} \equiv \dot{\mathbf{R}} = \text{constant}. \tag{1.23}$$

1.3.2. Total Linear Momentum

The total linear momentum of the system is defined as

$$\mathbf{P} = \sum_{i=1}^{N} m_i\,\dot{\mathbf{r}}_i. \tag{1.24}$$

Using the center of mass velocity $\mathbf{V} = \dot{\mathbf{R}}$ from (1.20), one also has

$$\mathbf{P} = M\,\mathbf{V}. \tag{1.25}$$

Differentiating with respect to time and using (1.21),

$$\dot{\mathbf{P}} = M\,\ddot{\mathbf{R}} = \mathbf{F}^{\text{ext}}. \tag{1.26}$$

Thus, the rate of change of the total linear momentum equals the net external force acting on the system.
If $\mathbf{F}^{\text{ext}} = \mathbf{0}$, then

$$\dot{\mathbf{P}} = \mathbf{0} \quad \Longrightarrow \quad \mathbf{P} = \text{constant}. \tag{1.27}$$

Hence, the total linear momentum of the system is conserved when no external forces act.

1.3.3. Moments of Force and Momentum

Starting with Newton's second law for each particle,

$$\dot{\mathbf{p}}_i = \mathbf{F}_i, \qquad \mathbf{p}_i \equiv m_i\,\dot{\mathbf{r}}_i, \tag{1.28}$$

take the cross product with $\mathbf{r}_i$ and sum over all particles:

$$\sum_{i=1}^{N} \mathbf{r}_i \times \dot{\mathbf{p}}_i = \sum_{i=1}^{N} \mathbf{r}_i \times \mathbf{F}_i. \tag{1.29}$$

For the left-hand side, use the product rule,

$$\begin{aligned} \sum_i \mathbf{r}_i \times \dot{\mathbf{p}}_i &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_i \mathbf{r}_i \times \mathbf{p}_i\right) - \sum_i \dot{\mathbf{r}}_i \times \mathbf{p}_i \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_i \mathbf{r}_i \times \mathbf{p}_i\right), \end{aligned} \tag{1.30}$$

since $\dot{\mathbf{r}}_i \times \mathbf{p}_i = \dot{\mathbf{r}}_i \times (m_i \dot{\mathbf{r}}_i) = \mathbf{0}$. Define the total angular momentum

$$\mathbf{L} \equiv \sum_{i=1}^{N} \mathbf{r}_i \times \mathbf{p}_i, \tag{1.31}$$

so that the left-hand side becomes

$$\frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t}. \tag{1.32}$$

For the right-hand side, substitute (1.15):

$$\sum_i \mathbf{r}_i \times \mathbf{F}_i = \sum_i \mathbf{r}_i \times \mathbf{F}_i^{\text{ext}} + \sum_i \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathbf{r}_i \times \mathbf{F}_{ij}. \tag{1.33}$$

Using the pairwise antisymmetry (1.16), the internal term can be symmetrized:

$$\sum_i \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathbf{r}_i \times \mathbf{F}_{ij} = \frac{1}{2} \sum_{\substack{i,j=1 \\ i\neq j}}^{N} (\mathbf{r}_i - \mathbf{r}_j) \times \mathbf{F}_{ij}. \tag{1.34}$$

If the forces satisfy the *strong* form of Newton's third law (internal forces are central, i.e. $\mathbf{F}_{ij} \parallel \mathbf{r}_i - \mathbf{r}_j$), then each pairwise term vanishes and

$$\sum_i \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathbf{r}_i \times \mathbf{F}_{ij} = \mathbf{0}. \tag{1.35}$$

Therefore,

$$\frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t} = \sum_{i=1}^{N} \mathbf{r}_i \times \mathbf{F}_i^{\text{ext}} \equiv \tau^{\text{ext}}, \tag{1.36}$$

where we defined the total external torque

$$\tau^{\text{ext}} = \sum_{i=1}^{N} \mathbf{r}_i \times \mathbf{F}_i^{\text{ext}}. \tag{1.37}$$

If $\tau^{\text{ext}} = \mathbf{0}$, then the total angular momentum is conserved:

$$\mathbf{L} = \text{constant}. \tag{1.38}$$

1.3.4. Energy of the System

The work–energy theorem states

$$\mathbf{F} \cdot \mathrm{d}\mathbf{r} \;=\; \mathrm{d}\big(\tfrac{1}{2} m v^2\big) \;=\; \mathrm{d}T, \tag{1.39}$$

where T is the kinetic energy. For a system of N particles, summing over all particles gives

$$\mathrm{d}T \;=\; \sum_{i=1}^{N} \mathbf{F}_i \cdot \mathrm{d}\mathbf{r}_i. \tag{1.40}$$

Substituting the decomposition (1.15),

$$\mathrm{d}T \;=\; \sum_i \mathbf{F}_i^{\text{ext}} \cdot \mathrm{d}\mathbf{r}_i + \sum_i \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathbf{F}_{ij} \cdot \mathrm{d}\mathbf{r}_i. \tag{1.41}$$

If the external forces are conservative, there exists a potential $V(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ such that

$$\mathbf{F}_i^{\text{ext}} \;=\; -\boldsymbol{\nabla}_i V, \tag{1.42}$$

where $\boldsymbol{\nabla}_i$ is the gradient with respect to $\mathbf{r}_i$.
For internal forces:

- In full generality (conservative internal forces), there exists an *internal* potential $U(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ with the *total* internal force on particle i given by

$$\mathbf{F}_i^{\text{int}} \;\equiv\; \sum_{j\neq i} \mathbf{F}_{ij} \;=\; -\boldsymbol{\nabla}_i U. \tag{1.43}$$

- If the interactions are pairwise and central, one may write $U = \frac{1}{2}\sum_{i\neq j} U_{ij}(\mathbf{r}_i - \mathbf{r}_j)$ with

$$\mathbf{F}_{ij} \;=\; -\boldsymbol{\nabla}_i U_{ij}. \tag{1.44}$$

Using (1.42) and (1.43) in (1.41) (and assuming no explicit time dependence of V, U),

$$\mathrm{d}T \;=\; -\,\mathrm{d}V \;-\; \mathrm{d}U. \tag{1.45}$$

Thus, the total mechanical energy is conserved:

$$T + V + U \;=\; \text{constant}. \tag{1.46}$$

1.3.5. Energy and Conservation Laws

For N particles, the total kinetic energy is

$$T \;=\; \frac{1}{2}\sum_{i=1}^{N} m_i v_i^2. \tag{1.47}$$

If both external and internal forces are conservative, they are gradients of scalar potentials:

$$\mathbf{F}_i^{\text{ext}} = -\,\boldsymbol{\nabla}_i V, \qquad\qquad \mathbf{F}_{ij} = -\,\boldsymbol{\nabla}_i U_{ij} \quad \text{(pairwise case)}. \tag{1.48}$$

The total work along a displacement from point 1 to 2 is

$$W_{1\to2} \;=\; \sum_i \int_1^2 \mathbf{F}_i \cdot \mathrm{d}\mathbf{r}_i. \tag{1.49}$$

With $\mathbf{F}_i = \mathbf{F}_i^{\text{ext}} + \sum_{j\neq i} \mathbf{F}_{ij}$,

$$W_{1\to2} = \sum_i \int_1^2 \mathbf{F}_i^{\text{ext}} \cdot \mathrm{d}\mathbf{r}_i + \sum_i \int_1^2 \sum_{j\neq i} \mathbf{F}_{ij} \cdot \mathrm{d}\mathbf{r}_i. \tag{1.50}$$

For conservative external forces,

$$\sum_i \int_1^2 \mathbf{F}_i^{\text{ext}} \cdot \mathrm{d}\mathbf{r}_i \;=\; -\,\Delta V. \tag{1.51}$$

For pairwise internal forces,

$$\sum_i \int_1^2 \sum_{j\neq i} \mathbf{F}_{ij} \cdot \mathrm{d}\mathbf{r}_i \;=\; -\frac{1}{2}\sum_{i\neq j} \Delta U_{ij}, \tag{1.52}$$

where the factor 1/2 avoids double counting of pairs and ΔU_{ij} denotes the change in the pair potential between states 1 and 2. Hence,

$$W_{1\to2} \;=\; -\,\Delta V \;-\; \Delta U \;=\; T_2 - T_1, \tag{1.53}$$

which reproduces energy conservation (1.46):

$$T + V + U \;=\; \text{constant}. \tag{1.54}$$

Remark 1.3 – Explicit time dependence

If V or U depends explicitly on t (e.g., time-varying external fields), $\mathrm{d}(T + V + U) = \frac{\partial V}{\partial t}\mathrm{d}t + \frac{\partial U}{\partial t}\mathrm{d}t \neq 0$, and mechanical energy is not conserved.

1.3.6. Energy of Rigid Bodies

For a rigid body, the relative positions of particles are fixed, so if the internal potential depends only on particle separations, U is constant. The total mechanical energy reduces to

$$E = T + V. \tag{1.55}$$

For *pure rotation about a fixed axis* (through the center of mass or a fixed point),

$$T = \tfrac{1}{2} I \omega^2, \tag{1.56}$$

with I the moment of inertia about that axis and ω the angular velocity.

1.3.7. Galilean Transformations Between Frames

Let S be an inertial frame and S' another inertial frame moving at constant velocity $\mathbf{V}$ relative to S. If $\mathbf{b}$ is the position of the origin of S' measured in S at $t = 0$, then the Galilean transformation is

$$\mathbf{r}' = \mathbf{r} - \mathbf{V}t - \mathbf{b}, \tag{1.57}$$

so that

$$\mathbf{v}' = \mathbf{v} - \mathbf{V}, \qquad \mathbf{p}' = \mathbf{p} - m\,\mathbf{V}. \tag{1.58}$$

Equivalently, $\mathbf{v} = \mathbf{v}' + \mathbf{V}$ and $\mathbf{p} = \mathbf{p}' + m\mathbf{V}$. Since $\mathbf{V}$ is constant,

$$\dot{\mathbf{p}}' = \dot{\mathbf{p}} = \mathbf{F}, \tag{1.59}$$

so Newton's equations retain their form under Galilean transformations.
For a system of particles,

$$\mathbf{P}' = \mathbf{P} - M\,\mathbf{V}, \qquad \dot{\mathbf{P}}' = \dot{\mathbf{P}}, \tag{1.60}$$

equivalently $\mathbf{P} = \mathbf{P}' + M\,\mathbf{V}$.

1.3.8. The Center of Momentum Frame

The **center of momentum (CM) frame** is an inertial frame in which the total momentum vanishes *instantaneously*:

$$\mathbf{P}'(t_0) = \sum_a \mathbf{p}'_a(t_0) = \mathbf{0}. \tag{1.61}$$

Choosing the origin at the instantaneous center of mass, the center of mass positions satisfy

$$M\,\mathbf{R} = M\,\mathbf{R}' + M\,\mathbf{V}t + M\,\mathbf{b}, \tag{1.62}$$

and, differentiating twice,

$$M\,\ddot{\mathbf{R}} = M\,\ddot{\mathbf{R}}' = \mathbf{F}^{\text{ext}}. \tag{1.63}$$

Thus Newton's second law applies in the CM frame as usual. If $\mathbf{F}^{\text{ext}} = \mathbf{0}$, then $\ddot{\mathbf{R}}' = \mathbf{0}$ and, with the origin at the CM, one has

$$\mathbf{R}'(t) = \mathbf{0} \quad \text{for all } t. \tag{1.64}$$

(Otherwise, $\mathbf{R}' = \mathbf{0}$ only at the chosen instant t_0.)

1.3.9. Intrinsic Angular Momentum

The total angular momentum $\mathbf{L}$ and the external torque τ^{ext} are generally origin-dependent. Consider a shift of origin from O to O' with constant displacement $\mathbf{b}$:

$$\mathbf{r}_a = \mathbf{r}'_a + \mathbf{b}, \qquad \dot{\mathbf{r}}_a = \dot{\mathbf{r}}'_a, \qquad \mathbf{p}_a = \mathbf{p}'_a. \tag{1.65}$$

For a single particle,

$$\mathbf{L}_a = \mathbf{r}_a \times \mathbf{p}_a = \mathbf{r}'_a \times \mathbf{p}_a + \mathbf{b} \times \mathbf{p}_a = \mathbf{L}'_a + \mathbf{b} \times \mathbf{p}_a. \tag{1.66}$$

Summing over all particles,

$$\mathbf{L} = \sum_a \mathbf{r}_a \times \mathbf{p}_a = \sum_a \mathbf{r}'_a \times \mathbf{p}_a + \mathbf{b} \times \sum_a \mathbf{p}_a = \mathbf{L}' + \mathbf{b} \times \mathbf{P}, \tag{1.67}$$

where

$$\mathbf{L}' \equiv \sum_a \mathbf{r}'_a \times \mathbf{p}_a, \qquad \mathbf{P} \equiv \sum_a \mathbf{p}_a. \tag{1.68}$$

If O' is chosen as the *center of momentum* frame, then $\mathbf{P} = \mathbf{0}$ and $\mathbf{L} = \mathbf{L}'$. This shows the decomposition

- $\mathbf{L}'$: intrinsic (about O'),
- $\mathbf{b} \times \mathbf{P}$: orbital part due to motion of the center of mass relative to O.

For torques,

$$\tau^{\text{ext}} = \sum_a \mathbf{r}_a \times \mathbf{F}_a = \sum_a \mathbf{r}'_a \times \mathbf{F}_a + \mathbf{b} \times \sum_a \mathbf{F}_a = \tau'^{\text{ext}} + \mathbf{b} \times \mathbf{F}^{\text{ext}}. \tag{1.69}$$

If $\mathbf{b} = \mathbf{R}$ (the center of mass position), then

$$\tau^{\text{ext}} = \tau'^{\text{ext}} + \mathbf{R} \times \mathbf{F}^{\text{ext}}. \tag{1.70}$$

1.3.10. Transformation Law for Kinetic Energy

The kinetic energy in frame S is

$$T = \sum_a \tfrac{1}{2} m_a v_a^2. \tag{1.71}$$

With $\mathbf{v}_a = \mathbf{v}'_a + \mathbf{V}$ (where S' moves with constant $\mathbf{V}$ relative to S),

$$\begin{aligned} T &= \sum_a \tfrac{1}{2} m_a \big(\mathbf{v}'_a + \mathbf{V}\big)\cdot\big(\mathbf{v}'_a + \mathbf{V}\big) \\ &= \underbrace{\sum_a \tfrac{1}{2} m_a {v'_a}^2}_{T'} + \mathbf{V} \cdot \underbrace{\sum_a m_a \mathbf{v}'_a}_{\mathbf{P}'} + \tfrac{1}{2} M V^2. \end{aligned} \tag{1.72}$$

If S' is the center of momentum frame ($\mathbf{P}' = \mathbf{0}$),

$$T = T' + \tfrac{1}{2} M V^2. \tag{1.73}$$

For a rigid body,

$$T = \tfrac{1}{2} I\, \omega^2 + \tfrac{1}{2} M V^2, \tag{1.74}$$

where I is the moment of inertia about the rotation axis (typically through the center of mass) and $\mathbf{V}$ is the center-of-mass velocity.

1.3.11. Rotating Frames and Non-Inertial Dynamics

Let R be a frame rotating with angular velocity $\omega(t)$ relative to an inertial frame S, with coincident origins. The transport theorem for any vector $\mathbf{A}$ gives

$$\left(\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t}\right)_S = \left(\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t}\right)_R + \omega \times \mathbf{A}. \tag{1.75}$$

Applying this to $\mathbf{r}$ and $\mathbf{v}$:

$$\mathbf{v}_S = \mathbf{v}_R + \omega \times \mathbf{r}, \tag{1.76}$$

$$\mathbf{a}_S = \mathbf{a}_R + 2\,\omega \times \mathbf{v}_R + \dot{\omega} \times \mathbf{r} + \omega \times \big(\omega \times \mathbf{r}\big). \tag{1.77}$$

1.3.12. Equation of Motion in Rotating Frames

From Newton's second law $\mathbf{F} = m\,\mathbf{a}_S$ and (1.77),

$$m\,\mathbf{a}_R = \mathbf{F} - m\,\omega \times \big(\omega \times \mathbf{r}\big) - 2m\,\omega \times \mathbf{v}_R - m\,\dot{\omega} \times \mathbf{r}. \tag{1.78}$$

Identifying the fictitious forces,

$$\mathbf{F}_{\text{centrifugal}} = -\,m\,\omega \times \big(\omega \times \mathbf{r}\big), \tag{1.79}$$

$$\mathbf{F}_{\text{Coriolis}} = -\,2m\,\omega \times \mathbf{v}_R, \tag{1.80}$$

$$\mathbf{F}_{\text{Euler}} = -\,m\,\dot{\omega} \times \mathbf{r}, \tag{1.81}$$

the equation of motion becomes

$$m\,\mathbf{a}_R = \mathbf{F} + \mathbf{F}_{\text{centrifugal}} + \mathbf{F}_{\text{Coriolis}} + \mathbf{F}_{\text{Euler}}. \tag{1.82}$$

1.3.13. Energy in Rotating Frames

The inertial-frame kinetic energy expressed in rotating-frame variables is

$$T_S = \tfrac{1}{2}m\,\big\|\mathbf{v}_R + \omega \times \mathbf{r}\big\|^2 = \tfrac{1}{2}mv_R^2 + m\,(\omega \times \mathbf{r})\cdot\mathbf{v}_R + \tfrac{1}{2}m\,\big\|\omega \times \mathbf{r}\big\|^2. \tag{1.83}$$

Define the *relative* kinetic energy in the rotating frame

$$T \equiv \tfrac{1}{2}mv_R^2, \tag{1.84}$$

and (for constant ω) the effective potential

$$V_{\text{eff}} \equiv V(\mathbf{r}) - \tfrac{1}{2}m\,\big\|\omega \times \mathbf{r}\big\|^2 \;=\; V(\mathbf{r}) - \tfrac{1}{2}m\omega^2 r_\perp^2, \tag{1.85}$$

where $r_\perp$ is the distance to the rotation axis. Then the energy integral in the rotating frame (for conservative V and constant ω) is

$$E_{\text{rot}} \;=\; T + V_{\text{eff}}, \tag{1.86}$$

which is conserved if V has no explicit time dependence in the rotating frame.

Remark 1.4 – When does E_{rot} conserve?

Equation (1.86) follows from the rotating-frame Lagrangian $L' = \frac{1}{2}mv_R^2 + m(\omega \times \mathbf{r})\cdot\mathbf{v}_R - V(\mathbf{r})$, whose energy function is $E' = T + V - \frac{1}{2}m\|\omega \times \mathbf{r}\|^2 = T + V_{\text{eff}}$. It is a constant of motion when ω is constant and V is time independent.

1.3.14. Summary

- **Newton's Second Law**: In an inertial frame, the motion of each particle in a system obeys

$$m_a \ddot{\mathbf{r}}_a = \mathbf{F}_a, \qquad a = 1, 2, \dots, N. \tag{1.87}$$

  This yields $3N$ second-order ODEs which, in principle, determine the motion given initial positions/velocities and all forces. The forces may depend on positions, velocities, and possibly time (e.g., two particles connected by a spring: force depends on their relative position and any external driving).

- **Conservation Laws**:
  - **Linear Momentum**: If the weak form of Newton's third law holds ($\mathbf{F}_{ab} = -\mathbf{F}_{ba}$), then in the absence of external forces the total linear momentum is conserved:

$$\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} = \mathbf{F}^{\text{ext}}. \tag{1.88}$$

    (cf. (1.26))
  - **Angular Momentum**: If the strong form holds ($\mathbf{F}_{ab} = -\mathbf{F}_{ba} \parallel \mathbf{r}_{ab}$), the total angular momentum is conserved in the absence of external torques:

$$\frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t} = \tau^{\text{ext}}. \tag{1.89}$$

    (cf. (1.36))
  - **Energy**: If all forces (external and internal) are conservative, the total mechanical energy is conserved:

$$E \;=\; T + V + U \;=\; \text{constant}. \tag{1.90}$$

    (cf. (1.46))

- **Non-Inertial Frames**: Newton's second law extends to non-inertial frames by introducing fictitious forces that depend on frame motion (e.g., centrifugal and Coriolis; and Euler if $\dot{\omega} \neq \mathbf{0}$).

1.3.15. Limitations of Newtonian Mechanics

Despite its success, the Newtonian formulation has limitations that motivate more advanced approaches:

- **Vectorial Formulation and Coordinates**: Newtonian mechanics leans on vector equations that can become cumbersome in curvilinear coordinates. The Lagrangian formulation uses a scalar $\mathcal{L} = T - V$ and generalized coordinates $\{q_i\}$, handling non-Cartesian systems more naturally.

- **Constraints**: Many systems involve constraints (rigid bodies, inextensible strings, rolling without slipping). Newtonian treatments must explicitly model the associated forces (normal forces, tensions, constraint reactions), whereas Lagrangian methods incorporate constraints at the outset via generalized coordinates and, if needed, Lagrange multipliers.

- **Conservation Laws**: In the Newtonian framework, conservation of E, $\mathbf{P}$, and $\mathbf{L}$ typically appears as a consequence of specific force assumptions (e.g., weak/strong third law). The Lagrangian/Hamiltonian viewpoint reveals a deeper structure: conservation laws arise from symmetries (Noether's theorem)—time translation $\Rightarrow E$, space translation $\Rightarrow \mathbf{P}$, rotation $\Rightarrow \mathbf{L}$.

- **Relation to Modern Physics**:
  - **Special Relativity**: At high speeds ($v \sim c$), Newtonian mechanics is replaced by relativistic mechanics, accounting for Lorentz invariance and mass–energy equivalence.
  - **Quantum Mechanics**: At microscopic scales, quantum mechanics governs dynamics; the correspondence principle links classical and quantum descriptions in appropriate limits.

1.4. Degrees of Freedom and Constraints

1.4.1. Degrees of Freedom

The **degrees of freedom (DoF)** of a system are the minimum number of independent quantities required to completely specify the state of the system. For a system of N particles in 3D, one starts with $3N$ coordinates; constraints reduce the number of independent coordinates.

Example 1.1 – Pendulum

A simple pendulum consists of a mass m suspended from a fixed point by a rigid, massless rod of length L. The motion is constrained to a plane and to a circle of radius L, yielding a single DoF. A convenient generalized coordinate is the angle θ with the vertical. The Cartesian coordinates are

$$x = L\sin\theta, \qquad y = -L\cos\theta. \tag{1.91}$$

The generalized velocity is $\dot{\theta} = \frac{\mathrm{d}\theta}{\mathrm{d}t}$. Hence the dynamics are fully described by the single coordinate $\theta(t)$.

Example 1.2 – Rigid Body

A rigid body (fixed inter-particle distances) has six DoF in 3D:

1. **Three translational**: center-of-mass (COM) position $\mathbf{r}_{\mathrm{cm}} = (x, y, z)$.
2. **Three rotational**: orientation, e.g. Euler angles (ϕ, θ, ψ) (convention-dependent).

The translational motion is set by $\mathbf{r}_{\mathrm{cm}}$; the rotational motion can be described by (ϕ, θ, ψ) (e.g. ϕ: rotation about the fixed z-axis; θ: inclination; ψ: rotation about the body-fixed z'-axis).

If k independent holonomic constraints are applied to N particles, the DoF become

$$\mathrm{DoF} \;=\; 3N - k. \tag{1.92}$$

1.4.2. Constraints

Constraints reduce DoF by restricting motion. We focus on **holonomic constraints**, which give explicit relations among coordinates (and time). For contrast, we also mention **non-holonomic constraints**, which are inequalities or non-integrable differential relations.

Consider a particle with position $\mathbf{r}$ constrained to move on a fixed plane. A constraint force $\mathbf{f}^C$ acts normal to the plane, keeping the particle on it. For *virtual displacements* $\delta\mathbf{r}$ consistent with the constraints (at fixed time), the ideal constraint does no work because $\mathbf{f}^C \perp \delta\mathbf{r}$.

Remark 1.5 – Virtual work of ideal constraints

For ideal (workless) constraints,

$$\sum_i \mathbf{f}_i^C \cdot \delta\mathbf{r}_i \;=\; 0 \quad \text{for all virtual displacements consistent with the constraints.} \tag{1.93}$$

This does *not* preclude constraint forces from doing work over *actual* motions if the constraints (e.g. the plane) themselves move in time; we return to this when introducing D'Alembert's principle and generalized forces.

Holonomic Constraints

A constraint is **holonomic** if it can be written as an algebraic equation among the coordinates (and possibly time):

$$f(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N, t) = 0. \tag{1.94}$$

Example 1.3 – Rolling Without Slipping

A cylinder of radius a rolls without slipping along a straight line on a surface (see Fig. 1.2). The no-slip holonomic constraint along the direction of motion is

$$x' \;=\; a\,\theta, \tag{1.95}$$

where x' is the COM translation along the surface and θ is the cylinder's rotation angle. Consequently,

$$v_{\text{cm}} \;=\; \dot{x}' \;=\; a\,\dot{\theta}, \tag{1.96}$$

ensuring the contact point has zero velocity relative to the surface. One may choose a single generalized coordinate (e.g. θ or x') since (1.95) relates them.

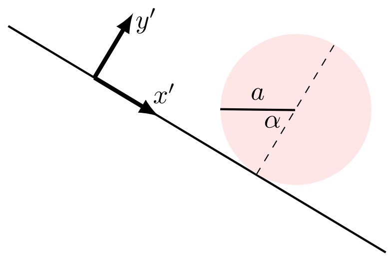


Figure 1.2: A cylinder of radius a rolling without slipping on an inclined plane. Here θ denotes the cylinder's rotation (used in (1.95)); α denotes the plane's inclination.

Non-Holonomic Constraints

A constraint is **non-holonomic** if it cannot be expressed as an algebraic relation like (1.94); common cases involve inequalities or non-integrable velocity relations.

Example 1.4 – Inequality Constraints

A particle confined to a rectangular box of width a and height b (Fig. 1.3) satisfies the inequalities

$$0 \le x \le a, \qquad 0 \le y \le b, \tag{1.97}$$

restricting its position to the interior. Such constraints bound the admissible region but do not reduce the configuration-space dimension algebraically (hence are non-holonomic).

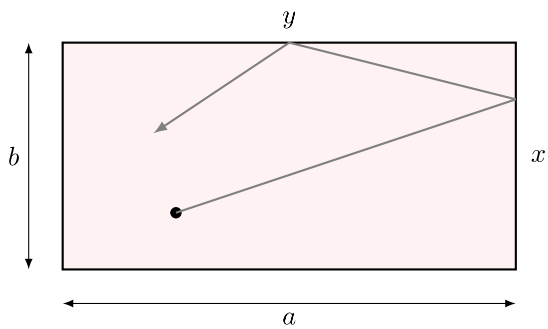


Figure 1.3: A particle confined within a rectangular box of width a and height b.

1.5. Generalized Quantities

1.5.1. Generalized Coordinates

The minimum number of independent coordinates required to describe a system's state are called **generalized coordinates**. These coordinates inherently account for all constraints in the system, reducing the complexity of the mathematical description.
For a system with $3N$ initial degrees of freedom and k holonomic constraints, we use

$$q_1, q_2, \ldots, q_{3N-k}.$$

The positions of all particles can then be expressed as functions of these coordinates:

$$\begin{aligned} \mathbf{r}_1 &= \mathbf{r}_1(q_1, \ldots, q_{3N-k}, t), \\ &\vdots \\ \mathbf{r}_i &= \mathbf{r}_i(q_1, \ldots, q_{3N-k}, t), \quad i = 1, 2, \ldots, N. \end{aligned} \tag{1.98}$$

Generalized coordinates can represent:

- distances (e.g., the length of a pendulum),
- angles (e.g., the orientation of a rigid body),
- combinations of distances and angles (e.g., cylindrical or spherical coordinates).

By using generalized coordinates, we incorporate the constraints into the formulation and simplify the system description.

1.5.2. Generalized Velocities

The **generalized velocities** are the time derivatives of the generalized coordinates:

$$\dot{q}_i = \frac{\mathrm{d}q_i}{\mathrm{d}t}, \qquad i = 1, \ldots, 3N - k. \tag{1.99}$$

They may represent translational rates, angular rates, or other rates depending on the chosen coordinates.
For example:

- In Cartesian coordinates, generalized velocities are v_x, v_y, v_z.
- For an angular coordinate θ, the generalized velocity is $\dot{\theta}$.

By the chain rule, particle velocities relate to $\dot{q}_i$ via

$$\dot{\mathbf{r}}_a = \sum_{i=1}^{3N-k} \frac{\partial \mathbf{r}_a}{\partial q_i} \dot{q}_i + \frac{\partial \mathbf{r}_a}{\partial t}, \qquad a = 1, \ldots, N. \tag{1.100}$$

1.5.3. Generalized Forces

The **generalized forces** Q_i are defined through the **virtual work** principle:

$$\delta W = \sum_{a=1}^{N} \mathbf{F}_a \cdot \delta \mathbf{r}_a = \sum_{i=1}^{3N-k} Q_i \, \delta q_i, \tag{1.101}$$

for virtual displacements $\delta \mathbf{r}_a$ consistent with the constraints at fixed time. Hence,

$$Q_i = \sum_{a=1}^{N} \mathbf{F}_a \cdot \frac{\partial \mathbf{r}_a}{\partial q_i} = \sum_{a=1}^{N} \sum_{\alpha \in \{x,y,z\}} F_{a,\alpha} \frac{\partial r_{a,\alpha}}{\partial q_i}. \tag{1.102}$$

If a potential $V(q,t)$ exists for the applied forces (and no velocity dependence), then

$$Q_i = -\frac{\partial V}{\partial q_i}. \tag{1.103}$$

Constraint forces for *ideal constraints* do no virtual work and thus do not contribute to Q_i.

Example 1.5 – Particle on a Circular Track

Consider a particle of mass m moving under gravity on a smooth circular track of radius a in the x-z plane, with the holonomic constraint

$$x^2 + z^2 = a^2. \tag{1.104}$$

A natural generalized coordinate is the angle θ along the circle:

$$x = a\cos\theta, \qquad z = a\sin\theta. \tag{1.105}$$

The **generalized velocity** is $\dot{\theta}$. Differentiating,

$$\dot{x} = -a\sin\theta\,\dot{\theta}, \qquad \dot{z} = \quad a\cos\theta\,\dot{\theta}, \tag{1.106}$$

so the speed is

$$v \;=\; \sqrt{\dot{x}^2 + \dot{z}^2} \;=\; a\,\dot{\theta}. \tag{1.107}$$

Forces:

- **Constraint force**: from the track, radial; it does no virtual work (ideal constraint).
- **Gravity**: $\mathbf{F} = (0, -mg)$ in (x, z)-components, i.e., $F_x = 0$, $F_z = -mg$.

Using virtual work for a virtual change $\delta\theta$,

$$\begin{aligned}\delta W &= \left[F_x\,\frac{\partial x}{\partial \theta} + F_z\,\frac{\partial z}{\partial \theta}\right]\delta\theta = [0\cdot(-a\sin\theta) + (-mg)\cdot(a\cos\theta)]\,\delta\theta \\ &= -\,mga\cos\theta\,\delta\theta.\end{aligned} \tag{1.108}$$

Thus the **generalized force** is

$$Q_\theta \;=\; -\,mga\cos\theta. \tag{1.109}$$

(Equivalently, with $V = mgz = mg\,a\sin\theta$, one has $Q_\theta = -\partial V/\partial\theta = -mga\cos\theta$.)

1.5.4. Functions of $\{q\}, \{\dot{q}\}, t$

Physical quantities (e.g. kinetic energy) can be expressed in terms of generalized coordinates $\{q\}$, generalized velocities $\{\dot{q}\}$, and time t. For N particles,

$$T \;=\; \frac{1}{2}\sum_{a=1}^{N} m_a\,\|\dot{\mathbf{r}}_a\|^2. \tag{1.110}$$

With the kinematic relations $\mathbf{r}_a = \mathbf{r}_a(q_1, \ldots, q_n, t)$, one has

$$\dot{\mathbf{r}}_a \;=\; \sum_{j=1}^{n} \frac{\partial \mathbf{r}_a}{\partial q_j}\,\dot{q}_j \;+\; \frac{\partial \mathbf{r}_a}{\partial t}. \tag{1.111}$$

Hence

$$\begin{aligned}T(q,\dot{q},t) &= \frac{1}{2}\sum_a m_a \left\|\sum_j \frac{\partial \mathbf{r}_a}{\partial q_j}\,\dot{q}_j \;+\; \frac{\partial \mathbf{r}_a}{\partial t}\right\|^2 \\ &= \frac{1}{2}\sum_{i,j=1}^{n} g_{ij}(q,t)\,\dot{q}_i\dot{q}_j \;+\; \sum_{i=1}^{n} A_i(q,t)\,\dot{q}_i \;+\; \frac{1}{2}\,\Phi(q,t),\end{aligned} \tag{1.112}$$

where

$$g_{ij}(q,t) \equiv \sum_a m_a \frac{\partial \mathbf{r}_a}{\partial q_i} \cdot \frac{\partial \mathbf{r}_a}{\partial q_j}, \qquad \text{(kinetic metric)} \tag{1.113}$$

$$A_i(q,t) \equiv \sum_a m_a \frac{\partial \mathbf{r}_a}{\partial q_i} \cdot \frac{\partial \mathbf{r}_a}{\partial t}, \tag{1.114}$$

$$\Phi(q,t) \equiv \sum_a m_a \left\| \frac{\partial \mathbf{r}_a}{\partial t} \right\|^2. \tag{1.115}$$

If the coordinates are *rheonomous* (explicitly time-dependent), A_i and Φ may be nonzero. For time-independent (scleronomous) constraints, $\frac{\partial \mathbf{r}_a}{\partial t} = \mathbf{0}$ and

$$T \;=\; \frac{1}{2} \sum_{i,j=1}^{n} g_{ij}(q)\, \dot{q}_i \dot{q}_j. \tag{1.116}$$

1.6. D'Alembert's Principle and Lagrange's Equations

Remark 1.6 – Historical note

D'Alembert's principle (1743) predates Lagrange's analytic formulation (1788). Hamilton's variational approach came later (1834–35). The principle of virtual work underlies the Lagrangian formalism.

A **virtual displacement** $\delta \mathbf{r}_i$ is an infinitesimal change of configuration at a fixed time t, consistent with the constraints (contrast with an actual displacement over a time interval $\mathrm{d}t$). D'Alembert's principle states that the total *virtual work of the impressed forces minus the inertial forces* vanishes:

$$\sum_{i=1}^{N} \big(\mathbf{F}_i - \dot{\mathbf{p}}_i\big) \cdot \delta \mathbf{r}_i \;=\; 0, \qquad \mathbf{p}_i \equiv m_i \dot{\mathbf{r}}_i. \tag{1.117}$$

Decompose $\mathbf{F}_i = \mathbf{F}_i^A + \mathbf{f}_i^C$ into applied and constraint forces. Introduce n independent generalized coordinates

$$\mathbf{r}_i = \mathbf{r}_i(q_1, \ldots, q_n, t), \qquad i = 1, \ldots, N, \tag{1.118}$$

with virtual displacements

$$\delta \mathbf{r}_i = \sum_{j=1}^{n} \frac{\partial \mathbf{r}_i}{\partial q_j}\, \delta q_j. \tag{1.119}$$

Particle velocities are

$$\mathbf{v}_i = \dot{\mathbf{r}}_i = \sum_{j=1}^{n} \frac{\partial \mathbf{r}_i}{\partial q_j}\, \dot{q}_j + \frac{\partial \mathbf{r}_i}{\partial t}. \tag{1.120}$$

A key identity (from the product rule) is

$$\sum_i m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_i m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \right) - \sum_i m_i \dot{\mathbf{r}}_i \cdot \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathbf{r}_i}{\partial q_j} \right). \tag{1.121}$$

Using $\frac{\partial \mathbf{v}_i}{\partial \dot{q}_j} = \frac{\partial \mathbf{r}_i}{\partial q_j}$ from (1.120), one obtains the standard kinetic-energy identity

$$\sum_i m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \;=\; \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j}. \tag{1.122}$$

Insert (1.119) into D'Alembert's principle (1.117) and group terms by δq_j:

$$\sum_{j=1}^{n}\left[\underbrace{\sum_i \mathbf{F}_i^A\cdot\frac{\partial \mathbf{r}_i}{\partial q_j}}_{Q_j^A}+\underbrace{\sum_i \mathbf{f}_i^C\cdot\frac{\partial \mathbf{r}_i}{\partial q_j}}_{Q_j^C}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot q_j}\right)+\frac{\partial T}{\partial q_j}\right]\delta q_j \;=\; 0. \tag{1.123}$$

For *ideal* constraints (no virtual work), $Q_j^C = 0$ (cf. §1.4.2), and because the δq_j are independent, we obtain Lagrange's equations in the kinetic-energy form:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot q_j}\right)-\frac{\partial T}{\partial q_j} \;=\; Q_j^A, \qquad j=1,\ldots,n. \tag{1.124}$$

If the applied forces are derivable from a potential $V(q,t)$ (no velocity dependence), $Q_j^A = -\frac{\partial V}{\partial q_j}$. Defining the Lagrangian $\mathcal{L}(q,\dot q,t)\equiv T-V$, we recover the Euler–Lagrange equations with possible nonconservative generalized forces $Q_j^{(\mathrm{nc})}$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot q_j}\right)-\frac{\partial \mathcal{L}}{\partial q_j} \;=\; Q_j^{(\mathrm{nc})}. \tag{1.125}$$

Remark 1.7 – Counting DoF with constraints

With $3N$ Cartesian coordinates and m independent holonomic constraints $g_k(\mathbf{r}_1,\ldots,\mathbf{r}_N,t)=0$, the number of generalized coordinates is $n = 3N - m$ (assuming no additional non-holonomic relations).

Remark 1.8 – Units of Q_j

Units of Q_j and q_j depend on the choice of coordinates; only the product $Q_j\,\delta q_j$ has the units of work (joules) in (1.123).

Substituting the results from Eqs. (**??**), (**??**), and (**??**) into the virtual work equation, Eq. (**??**), yields

$$\sum_{j=1}^{n}\left(\frac{d}{dt}\frac{\partial T}{\partial \dot q_j}-\frac{\partial T}{\partial q_j}\right)\delta q_j \;=\; \sum_{j=1}^{n} Q_j\,\delta q_j. \tag{1.126}$$

Since the q_j are independent,

$$\frac{d}{dt}\frac{\partial T}{\partial \dot q_j}-\frac{\partial T}{\partial q_j} \;=\; Q_j, \qquad j=1,2,\ldots,n. \tag{1.127}$$

If the applied forces are conservative, then

$$\mathbf{F}_i = -\nabla_i V, \qquad V = V(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N,t),$$

and

$$Q_j \;=\; \sum_{i=1}^{N}\mathbf{F}_i\cdot\frac{\partial \mathbf{r}_i}{\partial q_j} \;=\; -\sum_{i=1}^{N}\nabla_i V\cdot\frac{\partial \mathbf{r}_i}{\partial q_j}.$$

So that

$$Q_j \;=\; -\frac{\partial V}{\partial q_j}, \tag{1.128}$$

and the equation of motion becomes

$$\frac{d}{dt}\frac{\partial T}{\partial \dot q_j}-\frac{\partial T}{\partial q_j} \;=\; -\frac{\partial V}{\partial q_j}. \tag{1.129}$$

Since in general $T = T(q_i, \dot{q}_i, t)$ while (for conservative forces) $V = V(q_i, t)$, the above reduces to the Lagrange equations with $\mathcal{L} = T - V$:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \frac{\partial \mathcal{L}}{\partial q_j} \;=\; 0, \qquad j = 1, 2, \ldots, n, \qquad \mathcal{L} = T - V. \tag{1.130}$$

Remarks on the Lagrangian.

- Adding a total time derivative does not change the equations of motion:

$$\mathcal{L}'(q_i, \dot{q}_i, t) \;=\; \mathcal{L}(q_i, \dot{q}_i, t) + \frac{d}{dt} f(q_i, t). \tag{1.131}$$

- Scaling by a nonzero constant α leaves the equations unchanged: $\mathcal{L} \to \alpha\, \mathcal{L}$.

1.7. Velocity-Dependent Potentials and Dissipation Function

Looking back at the generalized equation

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} \;=\; Q_j, \qquad j = 1, 2, \ldots, n, \tag{1.132}$$

we can cast it into Lagrangian form

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \frac{\partial \mathcal{L}}{\partial q_j} \;=\; 0, \qquad j = 1, 2, \ldots, n, \tag{1.133}$$

if there exists a velocity-dependent generalized potential $V(q, \dot{q}, t)$ such that

$$Q_j \;=\; -\frac{\partial V}{\partial q_j} + \frac{d}{dt}\left(\frac{\partial V}{\partial \dot{q}_j}\right). \tag{1.134}$$

Then we define

$$\mathcal{L}(q, \dot{q}, t) \;=\; T(q, \dot{q}, t) - V(q, \dot{q}, t).$$

This generalizes the familiar single-particle case $\mathcal{L}(\mathbf{x}, \mathbf{v}, t) = \frac{1}{2}mv^2 - V(\mathbf{x})$, for which the Euler–Lagrange equation yields

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \mathbf{v}} - \frac{\partial \mathcal{L}}{\partial \mathbf{x}} = \mathbf{0} \;\Rightarrow\; m\,\frac{d^2\mathbf{x}}{dt^2} = -\,\nabla V(\mathbf{x}) \equiv \mathbf{F}.$$

Charged particle and the Lorentz force

For a particle of charge q and mass m in electromagnetic fields,

$$m\,\frac{d\mathbf{v}}{dt} = q\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right), \tag{1.135}$$

with potentials

$$\mathbf{E} = -\,\nabla\Phi - \frac{\partial \mathbf{A}}{\partial t}, \qquad \mathbf{B} = \nabla \times \mathbf{A}.$$

Equation (1.135) can be written as

$$m\,\frac{d\mathbf{v}}{dt} = q\left\{-\,\nabla\Phi - \frac{\partial \mathbf{A}}{\partial t} + \mathbf{v} \times (\nabla \times \mathbf{A})\right\}. \tag{1.136}$$

To match the generalized potential form

$$m\,\frac{d\mathbf{v}}{dt} = -\,\frac{\partial V}{\partial \mathbf{x}} + \frac{d}{dt}\left(\frac{\partial V}{\partial \mathbf{v}}\right), \tag{1.137}$$

use the identity (treating $\mathbf{v}$ as $\mathbf{x}$-independent in $\partial/\partial\mathbf{x}$)

$$\mathbf{v} \times (\nabla \times \mathbf{A}) = \nabla(\mathbf{A} \cdot \mathbf{v}) - (\mathbf{v} \cdot \nabla)\mathbf{A},$$

and choose the generalized potential

$$V(\mathbf{x},\mathbf{v},t) \;=\; q\,\Phi(\mathbf{x},t) - q\,\mathbf{A}(\mathbf{x},t)\cdot\mathbf{v}.$$

This yields the Lorentz force from Lagrange's equations exactly as required.

$$\mathbf{v}\times(\nabla\times\mathbf{A}) = \nabla(\mathbf{v}\cdot\mathbf{A}) - (\mathbf{v}\cdot\nabla)\mathbf{A}, \tag{1.138}$$

where the identity holds when $\mathbf{v}$ is treated as independent of $\mathbf{x}$ in the gradient.
Thus, Eq. (1.136) becomes

$$m\frac{d\mathbf{v}}{dt} = q\Big\{-\nabla\Phi - \frac{\partial\mathbf{A}}{\partial t} + \nabla(\mathbf{v}\cdot\mathbf{A}) - (\mathbf{v}\cdot\nabla)\mathbf{A}\Big\}. \tag{1.139}$$

Since $\Phi = \Phi(\mathbf{x},t)$ is independent of $\mathbf{v}$,

$$\frac{\partial}{\partial\mathbf{v}} = \mathbf{A}, \qquad \frac{\partial}{\partial\mathbf{v}} = \mathbf{A}.$$

Therefore Eq. (1.139) can be written in the generalized-potential form

$$m\frac{d\mathbf{v}}{dt} = -\frac{\partial V}{\partial\mathbf{x}} + \frac{d}{dt}\left(\frac{\partial V}{\partial\mathbf{v}}\right), \qquad V(\mathbf{x},\mathbf{v},t) = q\big(\Phi - \mathbf{v}\cdot\mathbf{A}\big), \tag{1.140}$$

i.e.

$$V = q(\Phi - \mathbf{v}\cdot\mathbf{A}).$$

We then identify the electromagnetic potential in the Lagrangian formalism as

$$V(\mathbf{x},\mathbf{v},t) = q\big(\Phi - \mathbf{v}\cdot\mathbf{A}\big), \tag{1.141}$$

so that the single-particle Lagrangian in $\mathbf{E},\mathbf{B}$ is

$$\mathcal{L}(\mathbf{x},\mathbf{v},t) = \frac{1}{2}m\mathbf{v}\cdot\mathbf{v} - q\,\Phi(\mathbf{x},t) + q\,\mathbf{A}(\mathbf{x},t)\cdot\mathbf{v}. \tag{1.142}$$

Its Euler–Lagrange equation is

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\mathbf{v}} - \frac{\partial\mathcal{L}}{\partial\mathbf{x}} = \mathbf{0}.$$

Expanding,

$$\frac{d}{dt}\Big(m\mathbf{v} + q\,\mathbf{A}(\mathbf{x},t)\Big) - q\,\nabla\Phi(\mathbf{x},t) + q\,\nabla\big(\mathbf{A}(\mathbf{x},t)\cdot\mathbf{v}\big) = \mathbf{0},$$

which gives

$$m\frac{d^2\mathbf{x}}{dt^2} = q\Big(-\frac{\partial\mathbf{A}}{\partial t} - (\mathbf{v}\cdot\nabla)\mathbf{A} - \nabla\Phi + \nabla(\mathbf{A}\cdot\mathbf{v})\Big).$$

Using $\nabla(\mathbf{A}\cdot\mathbf{v}) - (\mathbf{v}\cdot\nabla)\mathbf{A} = \mathbf{v}\times(\nabla\times\mathbf{A})$, we obtain

$$m\frac{d^2\mathbf{x}}{dt^2} = q\Big(-\nabla\Phi - \frac{\partial\mathbf{A}}{\partial t} + \mathbf{v}\times(\nabla\times\mathbf{A})\Big) = q\big(\mathbf{E} + \mathbf{v}\times\mathbf{B}\big). \tag{1.143}$$

Under a gauge transformation

$$\mathbf{A}\to\mathbf{A}+\nabla\chi, \qquad \Phi\to\Phi - \frac{\partial\chi}{\partial t}, \tag{1.144}$$

with scalar $\chi(\mathbf{x},t)$, Maxwell's equations and Lagrange's equations are invariant. The transformed Lagrangian is

$$\mathcal{L}'(\mathbf{x},\mathbf{v},t) = \frac{1}{2}m\mathbf{v}^2 - q\Big(\Phi - \frac{\partial\chi}{\partial t}\Big) + q\,\mathbf{v}\cdot\big(\mathbf{A}+\nabla\chi\big) = \mathcal{L} + q\left(\frac{\partial\chi}{\partial t} + \mathbf{v}\cdot\nabla\chi\right),$$

and since $\dfrac{d\chi}{dt} = \dfrac{\partial\chi}{\partial t} + \mathbf{v}\cdot\nabla\chi$, we have

$$\mathcal{L}'(\mathbf{x},\mathbf{v},t) = \mathcal{L}(\mathbf{x},\mathbf{v},t) + q\,\frac{d}{dt}\,\chi(\mathbf{x},t), \tag{1.145}$$

which leaves the equations of motion unchanged.
From the Lagrangian, the canonical momentum is

$$\mathbf{P} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{x}}} = m\mathbf{v} + q\mathbf{A}. \tag{1.146}$$

The Hamiltonian follows from the Legendre transform $H = \mathbf{P} \cdot \mathbf{v} - \mathcal{L}$. Using $\mathbf{v} = (\mathbf{P} - q\mathbf{A})/m$,

$$H(\mathbf{x}, \mathbf{P}) = \mathbf{P} \cdot \frac{\mathbf{P} - q\mathbf{A}}{m} - \left[\frac{1}{2}m\left(\frac{\mathbf{P} - q\mathbf{A}}{m}\right)^2 - q\Phi + q\,\mathbf{A} \cdot \frac{\mathbf{P} - q\mathbf{A}}{m}\right] = \frac{(\mathbf{P} - q\mathbf{A})^2}{2m} + q\,\Phi(\mathbf{x}, t),$$

i.e.

$$H(\mathbf{x}, \mathbf{P}) = \frac{(\mathbf{P} - q\mathbf{A})^2}{2m} + q\,\Phi(\mathbf{x}, t). \tag{1.147}$$

This exhibits the minimal-coupling prescription

$$\mathbf{P} \;\to\; \mathbf{P} - q\,\mathbf{A}, \tag{1.148}$$

plus the scalar term $q\,\Phi(\mathbf{x}, t)$.

1.8. Simple Applications of the Lagrangian Formulation

For a single particle,

$$\mathcal{L} = \frac{1}{2}m\mathbf{v} \cdot \mathbf{v} - V(\mathbf{x}, t), \tag{1.149}$$

we need the expression of T in various coordinate systems. Since

$$\mathbf{v} \cdot \mathbf{v} = v^2 = \left(\frac{ds}{dt}\right)^2, \tag{1.150}$$

and the line element is

$$\begin{aligned} d\mathbf{s} &= dx\,\hat{\imath} + dy\,\hat{\jmath} + dz\,\hat{k} && \text{(Cartesian)},\\ d\mathbf{s} &= dr\,\hat{\mathbf{r}} + r\,d\theta\,\hat{\theta} + r\sin\theta\,d\phi\,\hat{\phi} && \text{(Spherical)},\\ d\mathbf{s} &= dr\,\hat{\mathbf{r}} + r\,d\phi\,\hat{\phi} + dz\,\hat{\mathbf{z}} && \text{(Cylindrical)}, \end{aligned}$$

we obtain

$$T = \frac{1}{2}mv^2 = \begin{cases} \frac{1}{2}m\big(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\big), & \text{Cartesian},\\ \frac{1}{2}m\big(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\!\theta\;\dot{\phi}^2\big), & \text{Spherical},\\ \frac{1}{2}m\big(\dot{r}^2 + r^2\dot{\phi}^2 + \dot{z}^2\big), & \text{Cylindrical}. \end{cases} \tag{1.151}$$

1.8.1. Examples of Lagrangian Formulation

Example 1.6 – Single Particle in Cartesian Coordinates

A particle of mass m moving in a potential V. In Cartesian coordinates, the position is (x_1, x_2, x_3) and the velocity is $\mathbf{v} = (\dot{x}_1, \dot{x}_2, \dot{x}_3)$. The Lagrangian is

$$\mathcal{L} \;=\; \frac{1}{2}m\sum_{i=1}^{3}\dot{x}_i^2 \;-\; V(\{x_i\}, t). \tag{1.152}$$

The Euler–Lagrange equation for each x_j reads

$$\frac{\mathrm{d}}{\mathrm{dt}}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}_j}\right) - \frac{\partial \mathcal{L}}{\partial x_j} = 0. \tag{1.153}$$

Here

$$\frac{\partial \mathcal{L}}{\partial \dot{x}_j} = m\dot{x}_j, \qquad \frac{\mathrm{d}}{\mathrm{dt}}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}_j}\right) = m\ddot{x}_j, \qquad \frac{\partial \mathcal{L}}{\partial x_j} = -\frac{\partial V}{\partial x_j}. \tag{1.154}$$

Substituting into (1.153) gives the familiar Newton equation

$$m\ddot{x}_j = -\frac{\partial V}{\partial x_j}. \tag{1.155}$$

Example 1.7 – Single Particle in Polar Coordinates

A particle of mass m moving in a plane, with polar coordinates (r, θ). The speed components are $v_r = \dot{r}$ and $v_\theta = r\dot{\theta}$. The kinetic energy and Lagrangian are

$$T = \tfrac{1}{2}m\left(\dot{r}^{\,2} + r^2\dot{\theta}^2\right), \tag{1.156}$$

$$\mathcal{L} = \tfrac{1}{2}m\left(\dot{r}^{\,2} + r^2\dot{\theta}^2\right) - V(r,\theta,t). \tag{1.157}$$

Radial equation $(q_1 = r)$:

$$\frac{\partial \mathcal{L}}{\partial \dot{r}} = m\dot{r}, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{r}}\right) = m\ddot{r},$$

$$\frac{\partial \mathcal{L}}{\partial r} = mr\dot{\theta}^2 - \frac{\partial V}{\partial r}.$$

Hence

$$m\ddot{r} - mr\dot{\theta}^2 = -\frac{\partial V}{\partial r}. \tag{1.158}$$

Angular equation $(q_2 = \theta)$:

$$\frac{\partial \mathcal{L}}{\partial \dot{\theta}} = mr^2\dot{\theta}, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}}\right) = m\left(2r\dot{r}\dot{\theta} + r^2\ddot{\theta}\right),$$

$$\frac{\partial \mathcal{L}}{\partial \theta} = -\frac{\partial V}{\partial \theta}.$$

Therefore

$$m\left(2r\dot{r}\dot{\theta} + r^2\ddot{\theta}\right) = -\frac{\partial V}{\partial \theta}. \tag{1.159}$$

Equations (1.158)–(1.159) govern the radial and angular motion.

Example 1.8 – Atwood's Machine

Atwood's machine consists of a smooth pulley, a light inextensible string, and two masses M and m. The pulley is fixed; the string moves frictionlessly over it. The masses move vertically under gravity; see Fig. 1.4.

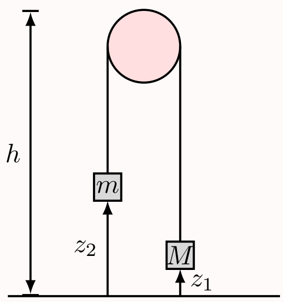


Figure 1.4: Schematic of Atwood's machine.

Let z_1 and z_2 be the vertical positions of M and m (upward positive), l the fixed string length, and h the height of the pulley above the ground. The constraint is

$$z_1 + z_2 + l = 2h. \tag{1.160}$$

Choose $q = z_1$ as the generalized coordinate; then

$$z_2 = 2h - l - z_1 \quad \Rightarrow \quad \dot{z}_2 = -\dot{z}_1. \tag{1.161}$$

The energies are

$$T = \tfrac{1}{2}M\dot{z}_1^2 + \tfrac{1}{2}m\dot{z}_2^2 = \tfrac{1}{2}(M+m)\,\dot{q}^{\,2}, \tag{1.162}$$
$$V = Mgz_1 + mgz_2 = (M-m)g\,q + \text{const.} \tag{1.163}$$

Hence the Lagrangian

$$\mathcal{L} = \tfrac{1}{2}(M+m)\,\dot{q}^{\,2} - (M-m)g\,q. \tag{1.164}$$

Euler–Lagrange gives

$$\frac{\mathrm{d}}{\mathrm{d}t}((M+m)\dot{q}) + (M-m)g = 0 \implies (M+m)\,\ddot{q} = -(M-m)g. \tag{1.165}$$

Thus the acceleration is

$$\ddot{q} = -\frac{(M-m)g}{M+m}. \tag{1.166}$$

Example 1.9 – Particle and Wedge

A particle of mass m slides without friction down a wedge of mass M, which is free to slide on a frictionless horizontal surface. The wedge makes an angle ϕ with the horizontal. The schematic is below.

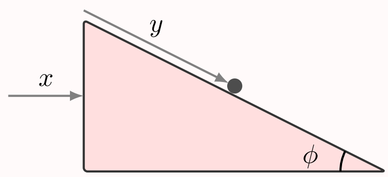


Figure 1.5: Particle of mass m on a movable wedge of mass M and angle ϕ.

We choose generalized coordinates x for the horizontal displacement of the wedge and y for the position of the particle along the inclined surface (measured up the plane). The kinetic energies are

$$T_{\text{wedge}} = \tfrac{1}{2}M\dot{x}^2, \qquad T_{\text{particle}} = \tfrac{1}{2}m\big(\dot{x}^2 + \dot{y}^2 + 2\dot{x}\dot{y}\cos\phi\big), \tag{1.167}$$

so the total kinetic energy is

$$T = \tfrac{1}{2}M\dot{x}^2 + \tfrac{1}{2}m\big(\dot{x}^2 + \dot{y}^2 + 2\dot{x}\dot{y}\cos\phi\big). \tag{1.168}$$

The vertical position of the particle is $y\sin\phi$; hence the potential energy (with an arbitrary zero) is

$$V = mg\,y\sin\phi. \tag{1.169}$$

The Lagrangian is

$$\mathcal{L} = T - V = \tfrac{1}{2}(M+m)\dot{x}^2 + m\dot{x}\dot{y}\cos\phi + \tfrac{1}{2}m\dot{y}^2 - mg\,y\sin\phi. \tag{1.170}$$

Applying Euler–Lagrange:

For x:

$$\frac{\partial\mathcal{L}}{\partial\dot{x}} = (M+m)\dot{x} + m\dot{y}\cos\phi, \quad \frac{\partial\mathcal{L}}{\partial x} = 0 \;\Rightarrow\; \frac{\mathrm{d}}{\mathrm{d}t}((M+m)\dot{x} + m\dot{y}\cos\phi) = 0. \tag{1.171}$$

For y:

$$\frac{\partial\mathcal{L}}{\partial\dot{y}} = m(\dot{y} + \dot{x}\cos\phi), \quad \frac{\partial\mathcal{L}}{\partial y} = -\,mg\sin\phi \;\Rightarrow\; \frac{\mathrm{d}}{\mathrm{d}t}(m(\dot{y} + \dot{x}\cos\phi)) = -\,mg\sin\phi. \tag{1.172}$$

Solving (1.171)–(1.172) for the accelerations yields the standard results

$$\ddot{x} = \frac{mg\sin\phi\cos\phi}{M + m\sin^2\phi}, \tag{1.173}$$

$$\ddot{y} = -\,\frac{(M+m)}{M + m\sin^2\phi}\,g\sin\phi. \tag{1.174}$$

Equation (1.171) reflects horizontal momentum conservation ($\partial\mathcal{L}/\partial x = 0$).

Example 1.10 – Simple Pendulum

Consider a simple pendulum of length b and bob mass m, displaced by an angle θ from the vertical. The geometry is shown below.

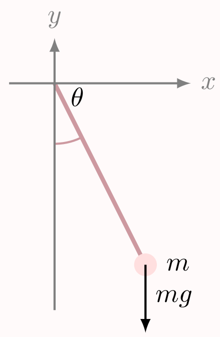


Figure 1.6: Simple pendulum of length b with bob mass m.

The Cartesian coordinates are

$$x = b\sin\theta, \qquad y = -\,b\cos\theta. \tag{1.175}$$

Thus

$$\dot{x} = b\dot{\theta}\cos\theta, \qquad \dot{y} = b\dot{\theta}\sin\theta. \tag{1.176}$$

The energies and Lagrangian are

$$T = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \tfrac{1}{2}mb^2\dot{\theta}^2, \tag{1.177}$$

$$U = mgy = -\,mgb\cos\theta, \tag{1.178}$$

$$\mathcal{L} = T - U = \tfrac{1}{2}mb^2\dot{\theta}^2 + mgb\cos\theta. \tag{1.179}$$

Euler–Lagrange for θ gives

$$\frac{\partial\mathcal{L}}{\partial\dot{\theta}} = mb^2\dot{\theta}, \quad \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\theta}}\right) = mb^2\ddot{\theta}, \quad \frac{\partial\mathcal{L}}{\partial\theta} = -\,mgb\sin\theta,$$

so

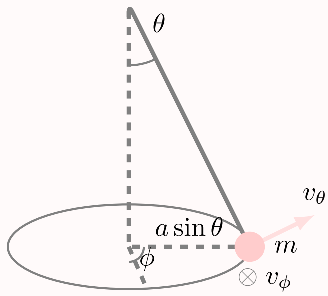
$$mb^2\ddot{\theta} + mgb\sin\theta = 0 \;\Longrightarrow\; \ddot{\theta} + \frac{g}{b}\sin\theta = 0. \tag{1.180}$$

For small oscillations ($\theta \ll 1$), $\sin\theta \simeq \theta$ and

$$\ddot{\theta} + \frac{g}{b}\theta = 0, \qquad \omega = \sqrt{\frac{g}{b}}, \qquad \theta(t) = \theta_0\cos(\omega t + \phi). \tag{1.181}$$

Example 1.11 – Spherical Pendulum

A spherical pendulum consists of a mass m swinging at a fixed length a from the origin. The motion is described by two angles: θ (from the vertical) and ϕ (azimuth). The system is illustrated in Fig.



Kinematics and kinetic energy:

$$v_\theta = a\,\dot{\theta}, \qquad v_\phi = a\sin\theta\,\dot{\phi}, \tag{1.182}$$

$$T = \tfrac{1}{2}m(v_\theta^2 + v_\phi^2) = \tfrac{1}{2}ma^2\dot{\theta}^2 + \tfrac{1}{2}ma^2\sin^2\theta\,\dot{\phi}^2. \tag{1.183}$$

Potential energy and Lagrangian:

$$V = -mga\cos\theta + \text{constant}, \tag{1.184}$$

$$\mathcal{L} = T - V = \tfrac{1}{2}ma^2\sin^2\theta\,\dot{\phi}^2 + \tfrac{1}{2}ma^2\dot{\theta}^2 + mga\cos\theta. \tag{1.185}$$

Euler–Lagrange equations:

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\theta}}\right) - \frac{\partial\mathcal{L}}{\partial\theta} = 0 \implies ma^2\ddot{\theta} = ma^2\sin\theta\cos\theta\,\dot{\phi}^2 - mga\sin\theta, \tag{1.186}$$

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\phi}}\right) - \frac{\partial\mathcal{L}}{\partial\phi} = 0 \implies \frac{d}{dt}\Big(ma^2\sin^2\theta\,\dot{\phi}\Big) = 0. \tag{1.187}$$

Hence the conserved vertical angular momentum is

$$L_z \equiv ma^2\sin^2\theta\,\dot{\phi} = \text{constant}. \tag{1.188}$$

Eliminating $\dot{\phi}$ in (1.186) using (1.188) gives

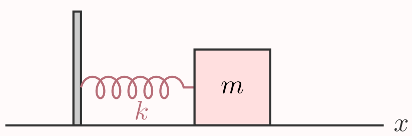
$$ma^2\ddot{\theta} = \frac{L_z^2\cos\theta}{ma^2\sin^3\theta} - mga\sin\theta. \tag{1.189}$$

Energy integral:

$$E = \tfrac{1}{2}ma^2\dot{\theta}^2 + \frac{L_z^2}{2ma^2\sin^2\theta} - mga\cos\theta. \tag{1.190}$$

Example 1.12 – Harmonic Oscillator

A mass m attached to a spring with spring constant k, resting on a flat surface. The spring is fixed to a vertical wall, and the mass can oscillate along the horizontal x-axis. The interaction between the spring's restoring force and the inertia of the mass results in simple harmonic motion.



Lagrangian and Euler–Lagrange:

$$\mathcal{L} = \tfrac{1}{2}m\dot{x}^2 - \tfrac{1}{2}kx^2, \tag{1.191}$$

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right) - \frac{\partial \mathcal{L}}{\partial x} = 0. \tag{1.192}$$

Derivatives:

$$\frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x}, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right) = m\ddot{x}, \qquad \frac{\partial \mathcal{L}}{\partial x} = -kx. \tag{1.193}$$

Equation of motion and solution:

$$m\ddot{x} + kx = 0, \qquad \omega = \sqrt{\frac{k}{m}}, \tag{1.194}$$

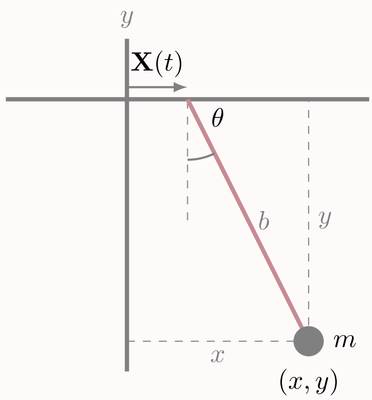
$$x(t) = A\cos(\omega t) + B\sin(\omega t), \tag{1.195}$$

with A, B fixed by initial conditions.

Example 1.13 – Driven Simple Pendulum

Consider a driven simple pendulum where the pivot moves horizontally according to $X(t)$. Let

$$X(t) = A\sin(\omega_0 t). \tag{1.196}$$



The coordinates of the bob are

$$x = X(t) + b\sin\theta, \tag{1.197}$$

$$y = -\,b\cos\theta. \tag{1.198}$$

Velocities:

$$\dot{x} = \dot{X} + b\dot{\theta}\cos\theta, \qquad \dot{y} = b\dot{\theta}\sin\theta. \tag{1.199}$$

Kinetic and potential energies:

$$T = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \tfrac{1}{2}m\left[\dot{X}^2 + 2b\dot{X}\dot{\theta}\cos\theta + b^2\dot{\theta}^2\right], \tag{1.200}$$

$$U = mgy = -\,mgb\cos\theta. \tag{1.201}$$

Lagrangian:

$$\mathcal{L} = T - U = \tfrac{1}{2}m\left[\dot{X}^2 + 2b\dot{X}\dot{\theta}\cos\theta + b^2\dot{\theta}^2\right] + mgb\cos\theta. \tag{1.202}$$

Euler–Lagrange for θ:

$$\frac{\partial \mathcal{L}}{\partial \dot\theta} = m\Big(b\dot X\cos\theta + b^2\dot\theta\Big), \tag{1.203}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot\theta}\right) = m\Big(b\ddot X\cos\theta - b\dot X\dot\theta\sin\theta + b^2\ddot\theta\Big), \tag{1.204}$$

$$\frac{\partial \mathcal{L}}{\partial \theta} = -\,mb\dot X\dot\theta\sin\theta - mgb\sin\theta. \tag{1.205}$$

Hence the equation of motion is

$$b\ddot\theta + g\sin\theta = -\,\ddot X\,\cos\theta. \tag{1.206}$$

With $X(t) = A\sin(\omega_0 t)$, $\ddot X = -A\omega_0^2\sin(\omega_0 t)$ and

$$b\ddot\theta + g\sin\theta = A\omega_0^2\sin(\omega_0 t)\,\cos\theta. \tag{1.207}$$

Chapter 2

Variational Principle and Lagrange's Equations

We have seen that Lagrange's equations are equivalent to Newton's equations of motion but provide a better formulation that works independently of the coordinate system used and avoids the concept of force, which can be very complicated. Lagrange's equations were derived in Chapter 1 using D'Alembert's virtual work principle; however, in this chapter we present a more general and elegant derivation via a variational principle. This mathematical principle is often regarded as a "Universal Fundamental Principle," called the least-action principle by physicists. It enables us to treat diverse physical phenomena on equal footing and in a unified way (see Ref. [3]).
Lagrange's equations follow from Hamilton's principle of least action: the motion of a given system between times t_i and t_f is such that the action

$$I \;=\; \int_{t_i}^{t_f} \mathcal{L}\Big(\vec{x},\, \dot{\vec{x}},\, t\Big)\, dt, \qquad \mathcal{L} = T - V, \tag{2.1}$$

is stationary for the actual path.

2.1. Variational Calculus

Variational calculus concerns extrema of quantities given as integrals. Given two fixed points $(x_1, y_1 = y(x_1))$ and $(x_2, y_2 = y(x_2))$, we seek a function $y(x)$ that extremizes the functional

$$I \;=\; \int_{x_1}^{x_2} F(x,\, y,\, y')\; dx, \qquad y' = \frac{\mathrm{d}y}{\mathrm{d}x}. \tag{2.2}$$

Example 2.1 – Fermat's Principle (Least Time)

Light travels between two fixed points along a path that makes the travel time stationary:

$$I \;=\; \int_{t_1}^{t_f} dt \;=\; \int_{r_1}^{r_f} \frac{ds}{v} \;=\; \frac{1}{c}\int_i^f n(\vec{r})\, ds, \tag{2.3}$$

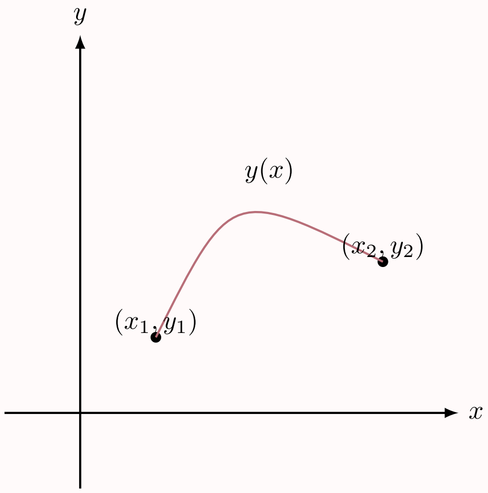
where $n(\vec{r})$ is the refractive index and c the vacuum speed of light.

Example 2.2 – Shortest Path Between Two Points

Find the curve $y(x)$ connecting (x_1, y_1) and (x_2, y_2) that minimizes arc length:

$$I \;=\; \int_{x_1}^{x_2} ds \;=\; \int_{x_1}^{x_2} \sqrt{dx^2 + dy^2} \;=\; \int_{x_1}^{x_2} \sqrt{1 + y'(x)^2}\, dx, \tag{2.4}$$

so here $F(x, y, y') = \sqrt{1 + y'^2}$.



Applying the Euler–Lagrange equation (derived below in Sec. 2.1), $\partial F/\partial y = 0$ implies

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) = 0 \implies \frac{\partial F}{\partial y'} = \frac{y'}{\sqrt{1+y'^2}} = C \implies y' = a = \text{const} \implies y(x) = ax + b, \tag{2.5}$$

i.e. the straight line through the two fixed points.

To find the general extremal of (2.2), consider a neighboring curve parametrized by α,

$$y(\alpha, x) \;=\; y(x) + \alpha\,\eta(x), \qquad \eta(x_1) = \eta(x_2) = 0, \tag{2.6}$$

with $y(0, x) = y(x)$. The functional becomes

$$I(\alpha) \;=\; \int_{x_1}^{x_2} F\big(x,\, y(\alpha, x),\, y'(\alpha, x)\big)\, dx. \tag{2.7}$$

Stationarity at $\alpha = 0$ requires

$$\left.\frac{\mathrm{d}I}{\mathrm{d}\alpha}\right|_{\alpha=0} = 0. \tag{2.8}$$

Using $\partial y/\partial\alpha = \eta$ and $\partial y'/\partial\alpha = d\eta/dx$,

$$\begin{aligned}\frac{\mathrm{d}I}{\mathrm{d}\alpha} &= \int_{x_1}^{x_2}\left(\frac{\partial F}{\partial y}\,\eta + \frac{\partial F}{\partial y'}\,\frac{\mathrm{d}\eta}{\mathrm{d}x}\right) dx \\ &= \int_{x_1}^{x_2}\left(\frac{\partial F}{\partial y}\,\eta - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right)\eta\right) dx \;+\; \eta\,\left.\frac{\partial F}{\partial y'}\right|_{x_1}^{x_2}.\end{aligned} \tag{2.9}$$

Since $\eta(x_1) = \eta(x_2) = 0$, the boundary term vanishes and, because η is arbitrary,

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) - \frac{\partial F}{\partial y} = 0. \tag{2.10}$$

Equation (2.10) is the Euler–Lagrange equation for the functional (2.2).

2.2. Lagrange's Equations from Hamilton's Principle

Euler's equation (2.10) mirrors the structure of Lagrange's equations (cf. Chapter 1), with the independent variable being time t:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right) - \frac{\partial \mathcal{L}}{\partial x} = 0, \qquad \mathcal{L} = T - V. \tag{2.11}$$

Hamilton's principle for conservative mechanical systems states that the actual motion renders the action

$$S = \int_{t_i}^{t_f} \mathcal{L}\Big(\vec{x}, \dot{\vec{x}}, t\Big)\, dt \tag{2.12}$$

stationary, which directly yields the Euler–Lagrange equations.
For n degrees of freedom with generalized coordinates $q_1, \ldots, q_n$,

$$S = \int_{t_i}^{t_f} \mathcal{L}(q_1, \ldots, q_n; \dot{q}_1, \ldots, \dot{q}_n; t)\, dt. \tag{2.13}$$

Varying with fixed endpoints $\delta q_i(t_i) = \delta q_i(t_f) = 0$,

$$\delta S = \int_{t_i}^{t_f} \sum_{i=1}^{n} \left(\frac{\partial \mathcal{L}}{\partial q_i}\, \delta q_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\, \delta \dot{q}_i \right) dt = \int_{t_i}^{t_f} \sum_{i=1}^{n} \left(\frac{\partial \mathcal{L}}{\partial q_i}\, \delta q_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\, \frac{\mathrm{d}}{\mathrm{d}t} \delta q_i \right) dt. \tag{2.14}$$

Integrating by parts the second term and using the vanishing endpoint variations,

$$\delta S = \int_{t_i}^{t_f} \sum_{i=1}^{n} \left[\frac{\partial \mathcal{L}}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \right] \delta q_i\, dt. \tag{2.15}$$

Because the δq_i are arbitrary and independent, stationarity $\delta S = 0$ implies the n Lagrange equations of motion:

$$\frac{\partial \mathcal{L}}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) = 0, \qquad i = 1, 2, \ldots, n. \tag{2.16}$$

Equations (2.16) are the Euler–Lagrange equations obtained from Hamilton's principle (2.12)–(2.13).

2.3. Lagrange's Equations in Presence of Constraints

For holonomic constraints, introduce constraint functions

$$f_\alpha(q_1, \ldots, q_n, t) = 0, \qquad \alpha = 1, 2, \ldots, k. \tag{2.17}$$

Virtual variations are taken at fixed time, so

$$\delta f_\alpha = \sum_{i=1}^{n} \frac{\partial f_\alpha}{\partial q_i}\, \delta q_i = 0, \qquad \alpha = 1, \ldots, k. \tag{2.18}$$

Introduce Lagrange multipliers $\lambda_\alpha(t)$ and define the augmented (constrained) Lagrangian

$$\mathcal{L}_c(q, \dot{q}, t, \lambda) = \mathcal{L}(q, \dot{q}, t) + \sum_{\alpha=1}^{k} \lambda_\alpha\, f_\alpha(q, t). \tag{2.19}$$

The corresponding action is

$$S_c = \int_{t_i}^{t_f} \left[\mathcal{L}(q, \dot{q}, t) + \sum_{\alpha=1}^{k} \lambda_\alpha\, f_\alpha(q, t) \right] dt. \tag{2.20}$$

Requiring $\delta S_c = 0$ with independent variations $\{\delta q_i\}$ and $\{\delta \lambda_\alpha\}$ (and fixed endpoints $\delta q_i(t_i) = \delta q_i(t_f) = 0$) gives

$$\delta S_c = \int_{t_i}^{t_f} \left\{ \sum_{i=1}^{n} \left[\frac{\partial \mathcal{L}}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \right] \delta q_i + \sum_{\alpha=1}^{k} \sum_{i=1}^{n} \lambda_\alpha \frac{\partial f_\alpha}{\partial q_i}\, \delta q_i + \sum_{\alpha=1}^{k} f_\alpha\, \delta \lambda_\alpha \right\} dt = 0. \tag{2.21}$$

Since δq_i and $\delta \lambda_\alpha$ are independent, we obtain the Euler–Lagrange equations with multipliers together with the constraints:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) - \frac{\partial \mathcal{L}}{\partial q_i} = \sum_{\alpha=1}^{k} \lambda_\alpha \frac{\partial f_\alpha}{\partial q_i}, \qquad i = 1, \ldots, n, \tag{2.22}$$

$$f_\alpha(q, t) = 0, \qquad \alpha = 1, \ldots, k. \tag{2.23}$$

(If constraints are specified as $f_\alpha = C_\alpha$, one may equivalently redefine $f_\alpha \to f_\alpha - C_\alpha$ to recover (2.23).) Equations (2.22)–(2.23) provide $n + k$ equations for the $n + k$ unknown functions $\{q_i(t)\}$ and $\{\lambda_\alpha(t)\}$.

Physical Meaning of the Lagrange Multipliers

For conservative systems $\mathcal{L} = T - V$, (2.22) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_i}\right) - \frac{\partial T}{\partial q_i} + \frac{\partial V}{\partial q_i} = \sum_{\alpha=1}^{k} \lambda_\alpha \frac{\partial f_\alpha}{\partial q_i} \equiv Q_i^{(\mathrm{C})}, \tag{2.24}$$

i.e., the right-hand side is the generalized force of constraint. Working directly with the $n-k$ independent generalized coordinates eliminates the explicit constraint forces from the equations.

Example 2.3 – Block on a Frictionless Incline via Multipliers

Consider a block of mass m on a fixed plane making an angle θ with the horizontal. Take Cartesian coordinates (x, y) with kinetic and potential energies

$$\mathcal{L} = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) - mg\,y. \tag{2.25}$$

The holonomic constraint is the line of the plane:

$$\frac{y}{x} = \tan\theta \iff f(x, y) = y - x\tan\theta = 0. \tag{2.26}$$

The augmented Lagrangian is

$$\mathcal{L}' = \mathcal{L} + \lambda f = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) - mg\,y + \lambda\,(y - x\tan\theta). \tag{2.27}$$

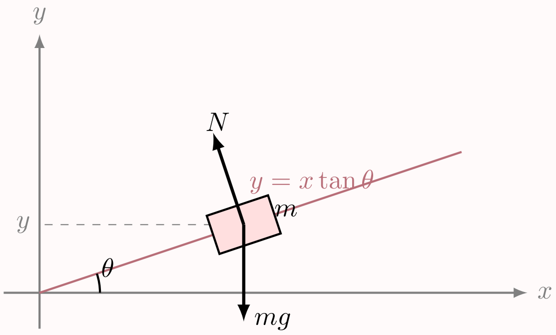


Figure 2.1: Block on a frictionless incline defined by $y = x\tan\theta$. The normal N is exactly perpendicular to the plane; gravity acts vertically.

Euler–Lagrange equations (plus $f = 0$) give

$$-\lambda\tan\theta - m\ddot{x} = 0, \tag{2.28}$$
$$\lambda - mg - m\ddot{y} = 0. \tag{2.29}$$

Comparing with Newton's form

$$m\ddot{x} = -N_x, \qquad m\ddot{y} = N_y - mg, \tag{2.30}$$

we identify

$$N_x = \lambda\tan\theta, \qquad N_y = \lambda. \tag{2.31}$$

Hence the normal reaction magnitude is

$$N = \sqrt{N_x^2 + N_y^2} = \lambda\sqrt{1 + \tan^2\theta} = \frac{\lambda}{\cos\theta} \implies \lambda = N\cos\theta. \tag{2.32}$$

2.4. Advantages of Variational Formulation

Because the Lagrangian is a scalar, the action and the resulting equations of motion are *coordinate-invariant*: any smooth change of generalized coordinates leaves Hamilton's principle unchanged. By working directly with the true degrees of freedom (either by choosing independent generalized coordinates or by enforcing constraints with multipliers), constraints are incorporated systematically. The variational approach is also broadly applicable beyond classical particle mechanics (e.g. circuits, optics, fields, relativity, quantum mechanics).

LC Circuit

Consider a series LC circuit with capacitor charge $q(t)$ and loop current $I(t) = \dot{q}$. Kirchhoff's loop law gives

$$L\,\dot{I} + \frac{q}{C} = 0 \quad \Longleftrightarrow \quad L\,\ddot{q} + \frac{1}{C}\,q = 0. \tag{2.33}$$

The mechanical analogy is $m\ddot{x} + kx = 0$ with

$$L \leftrightarrow m, \qquad \frac{1}{C} \leftrightarrow k. \tag{2.34}$$

The conservative Lagrangian (magnetic "kinetic" minus electric "potential") is

$$\mathcal{L}_{\mathrm{LC}}(q, \dot{q}) = \frac{1}{2} L\,\dot{q}^{\,2} - \frac{1}{2C}\,q^2, \tag{2.35}$$

whose Euler–Lagrange equation reproduces (2.33).

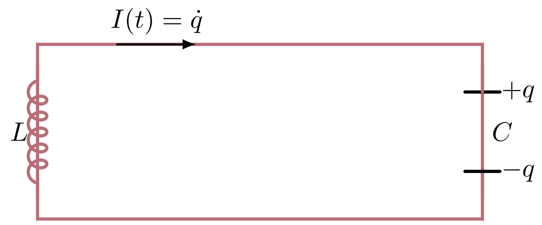


Figure 2.2: Series LC circuit with generalized coordinate q (capacitor charge) and $I = \dot{q}$.

Including Dissipation: Rayleigh Function

Resistive elements are non-conservative. They are incorporated via the *Rayleigh dissipation function* $\mathcal{F}$,

$$\mathcal{F}(\dot{q}) = \tfrac{1}{2} R\,\dot{q}^{\,2}, \qquad (I = \dot{q}), \tag{2.36}$$

which modifies the Euler–Lagrange equations to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}}\right) - \frac{\partial \mathcal{L}}{\partial q} + \frac{\partial \mathcal{F}}{\partial \dot{q}} = 0. \tag{2.37}$$

RC Circuit

For a series RC, KVL gives

$$R\,\dot{q} + \frac{1}{C}\,q = 0. \tag{2.38}$$

This follows from (2.37) using

$$\mathcal{L}_{\mathrm{RC}}(q, \dot{q}) = -\frac{1}{2C}\,q^2, \qquad \mathcal{F}(\dot{q}) = \tfrac{1}{2} R\,\dot{q}^{\,2}. \tag{2.39}$$

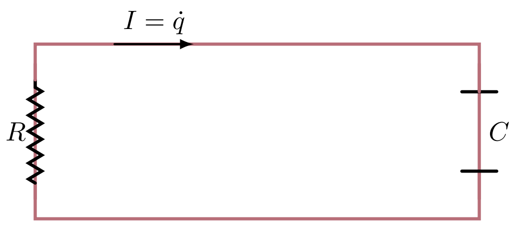


Figure 2.3: Series RC circuit. Dissipation enters through the Rayleigh function $\mathcal{F} = \frac{1}{2} R\dot{q}^2$.

RL Circuit

For a series RL,

$$L\,\dot{I} + R\,I = 0 \quad\Longleftrightarrow\quad L\,\ddot{q} + R\,\dot{q} = 0. \tag{2.40}$$

This follows from (2.37) with

$$\mathcal{L}_{\mathrm{RL}}(q,\dot{q}) = \tfrac{1}{2} L\,\dot{q}^2, \qquad \mathcal{F}(\dot{q}) = \tfrac{1}{2} R\,\dot{q}^2. \tag{2.41}$$

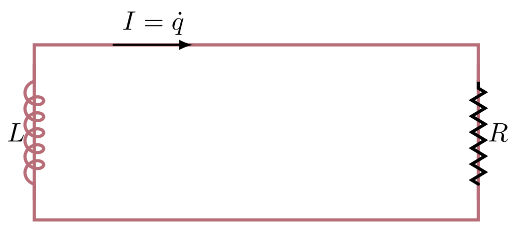


Figure 2.4: Series RL circuit.

RLC Circuit

A series RLC satisfies

$$L\,\ddot{q} + R\,\dot{q} + \frac{1}{C}\,q = 0, \tag{2.42}$$

which is the electrical analog of the damped oscillator $m\ddot{x} + \alpha\dot{x} + kx = 0$ under the correspondence (2.34) and $\alpha \leftrightarrow R$. Use

$$\mathcal{L}_{\mathrm{RLC}}(q,\dot{q}) = \tfrac{1}{2} L\,\dot{q}^2 - \tfrac{1}{2C}\,q^2, \qquad \mathcal{F}(\dot{q}) = \tfrac{1}{2} R\,\dot{q}^2, \tag{2.43}$$

in (2.37) to obtain (2.42).

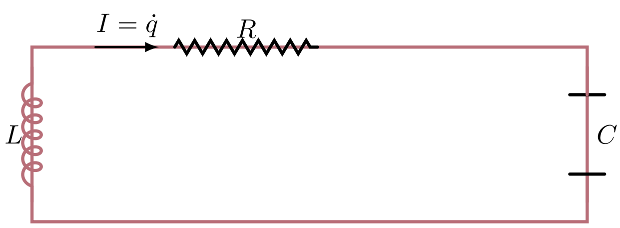


Figure 2.5: Series RLC circuit: L (magnetic inertia), R (dissipation), C (electric spring).

2.5. Symmetry and Conservation Laws

During the motion of a mechanical system, the $2N$ variables $\{q_i(t), \dot{q}_i(t)\}_{i=1}^N$ evolve in time. Certain functions of these variables remain constant along the motion and depend only on initial data; these are called *integrals of motion*. Even when the full solution of the Euler–Lagrange equations is not available, conserved quantities strongly constrain the dynamics.

Noether's Theorem

A *continuous symmetry* of the Lagrangian is a one-parameter family of transformations

$$q_i(t) \;\to\; q_i(t) + \epsilon\,\xi_i(q,t), \qquad \epsilon \ll 1, \tag{2.44}$$

$$\dot{q}_i(t) \;\to\; \dot{q}_i(t) + \epsilon\,\frac{\mathrm{d}}{\mathrm{d}t}\xi_i(q,t), \tag{2.45}$$

under which the Lagrangian is invariant (possibly up to a total time derivative).

Theorem 2.1 – Noether's theorem (simplest form)

If a transformation of the form (2.44)–(2.45) leaves the Lagrangian exactly invariant, $\delta\mathcal{L} = 0$, then the quantity

$$J \;=\; \sum_{i=1}^{N} \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\,\xi_i(q,t) \tag{2.46}$$

is conserved:

$$\frac{\mathrm{d}J}{\mathrm{d}t} = 0. \tag{2.47}$$

Proof. Using $\delta q_i = \epsilon \xi_i$ and $\delta \dot{q}_i = \epsilon\,\dot{\xi}_i$,

$$\begin{aligned}\delta\mathcal{L} &= \sum_i \left(\frac{\partial \mathcal{L}}{\partial q_i}\,\delta q_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\,\delta \dot{q}_i\right)\\ &= \epsilon \sum_i \left(\frac{\partial \mathcal{L}}{\partial q_i}\,\xi_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\,\dot{\xi}_i\right).\end{aligned} \tag{2.48}$$

Along true motions, the Euler–Lagrange equations give $\frac{\partial \mathcal{L}}{\partial q_i} = \frac{\mathrm{d}}{\mathrm{d}t}\big(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\big)$. Thus

$$\delta\mathcal{L} = \epsilon \sum_i \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right)\xi_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\,\dot{\xi}_i\right] \;=\; \epsilon\,\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\,\xi_i\right). \tag{2.49}$$

If $\delta\mathcal{L} = 0$, then $\frac{\mathrm{d}}{\mathrm{d}t}J = 0$ with J in (2.46). □
[General form] If the symmetry holds *up to a total time derivative*, $\delta\mathcal{L} = \epsilon\,\frac{\mathrm{d}F}{\mathrm{d}t}$, then $\frac{\mathrm{d}}{\mathrm{d}t}\big(J - F\big) = 0$. Many important symmetries (e.g. gauge transformations) take this form.

Examples

1. Cyclic (Ignorable) Coordinates

If a generalized coordinate q_s does not appear explicitly in $\mathcal{L}$, then the infinitesimal shift

$$q_s \to q_s + \epsilon, \qquad q_i \to q_i \quad (i \neq s) \tag{2.50}$$

is a symmetry, with $\xi_i = \delta_{is}$. Noether's charge (2.46) becomes

$$J = \frac{\partial \mathcal{L}}{\partial \dot{q}_s} \equiv p_s = \text{constant}, \tag{2.51}$$

i.e. the momentum conjugate to a cyclic coordinate is conserved. This also follows directly from the Euler–Lagrange equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_s}\right) - \frac{\partial \mathcal{L}}{\partial q_s} = 0 \implies \frac{\mathrm{d}p_s}{\mathrm{d}t} = 0 \quad (\partial\mathcal{L}/\partial q_s = 0). \tag{2.52}$$

2. Homogeneity of Space (Translations)

For a system of N particles with Cartesian coordinates $\vec{r}_i = (x_{i1}, x_{i2}, x_{i3})$, invariance under

$$\vec{r}_i \to \vec{r}_i + \epsilon\,\hat{n}, \qquad \text{equivalently} \quad x_{i\alpha} \to x_{i\alpha} + \epsilon\, n_\alpha \;(\alpha = 1, 2, 3), \tag{2.53}$$

implies the conserved quantity

$$J = \sum_{i=1}^{N} \sum_{\alpha=1}^{3} \frac{\partial \mathcal{L}}{\partial \dot{x}_{i\alpha}}\, n_\alpha = \sum_{\alpha=1}^{3} n_\alpha \left(\sum_{i=1}^{N} p_{i\alpha} \right) = \hat{n} \cdot \vec{P}, \tag{2.54}$$

where $p_{i\alpha} = \partial\mathcal{L}/\partial\dot{x}_{i\alpha}$ and $\vec{P} = \sum_{i=1}^{N} \vec{p}_i$ is the total momentum. Since $\hat{n}$ is arbitrary, all components of $\vec{P}$ are constant:

$$\vec{P} = \text{constant}. \tag{2.55}$$

3. Isotropy of Space

Isotropy of space means that the mechanical system does not change its properties when rotated as a whole about an arbitrary axis $\hat{n}$. An infinitesimal rotation by angle ϵ about $\hat{n}$ induces

$$\delta\vec{r}_\alpha \;=\; \epsilon\,\hat{n} \times \vec{r}_\alpha, \tag{2.56}$$

so that $\delta\vec{r}_\alpha \perp \hat{n}$ and $\delta\vec{r}_\alpha \perp \vec{r}_\alpha$. Its magnitude is

$$|\delta\vec{r}_\alpha| \;=\; r_\alpha \sin\theta\,\epsilon,$$

where θ is the angle between $\hat{n}$ and $\vec{r}_\alpha$.

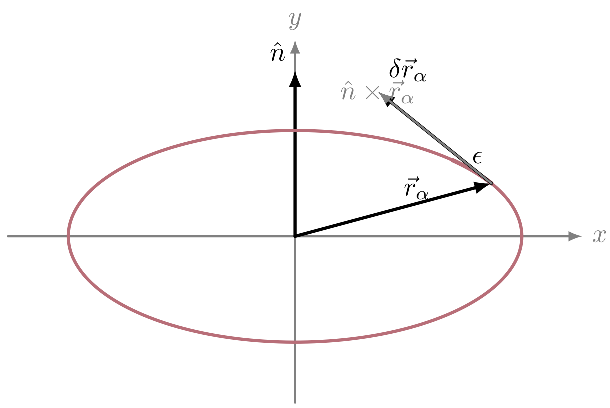


Figure 2.6: Isotropy of space: an infinitesimal rotation by ϵ about $\hat{n}$ moves the tip of $\vec{r}_\alpha$ along the orbit (ellipse in projection) with $\delta\vec{r}_\alpha = \epsilon\,\hat{n} \times \vec{r}_\alpha$.

Applying Noether's theorem with $\xi_\alpha = \hat{n} \times \vec{r}_\alpha$,

$$\begin{aligned} J &= \sum_{\alpha=1}^{N} \frac{\partial \mathcal{L}}{\partial \dot{\vec{r}}_\alpha} \cdot (\hat{n} \times \vec{r}_\alpha) = \sum_{\alpha=1}^{N} \sum_{i,j,k=1}^{3} p_{\alpha i}\, \epsilon_{ijk}\, n_j\, x_{\alpha k} \\ &= \sum_{\alpha=1}^{N} \sum_{j=1}^{3} n_j \left[\sum_{i,k=1}^{3} \epsilon_{ijk}\, x_{\alpha k} p_{\alpha i} \right] = \sum_{\alpha=1}^{N} \hat{n} \cdot (\vec{r}_\alpha \times \vec{p}_\alpha) \\ &= \hat{n} \cdot \vec{L} \;=\; \text{constant for all } \hat{n}, \end{aligned}$$

which implies conservation of total angular momentum

$$\vec{L} \;=\; \sum_{\alpha=1}^{N} \vec{r}_\alpha \times \vec{p}_\alpha \;=\; \text{constant}. \tag{2.57}$$

4. Homogeneity in Time

If the system is invariant under time translations,

$$t \;\to\; t+\epsilon, \tag{2.58}$$

then

$$\delta\mathcal{L} \;=\; \mathcal{L}(q,\dot{q},t+\epsilon)-\mathcal{L}(q,\dot{q},t) \;=\; \frac{\partial\mathcal{L}}{\partial t}\,\epsilon+\mathcal{O}(\epsilon^2). \tag{2.59}$$

Requiring invariance for arbitrary ϵ gives

$$\frac{\partial\mathcal{L}}{\partial t}=0. \tag{2.60}$$

Taking the total time derivative,

$$\begin{aligned}\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}t} &= \frac{\partial\mathcal{L}}{\partial t}+\sum_\alpha\frac{\partial\mathcal{L}}{\partial q_\alpha}\dot{q}_\alpha+\sum_\alpha\frac{\partial\mathcal{L}}{\partial\dot{q}_\alpha}\ddot{q}_\alpha \\ &= \sum_\alpha\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}_\alpha}\right)\dot{q}_\alpha+\sum_\alpha\frac{\partial\mathcal{L}}{\partial\dot{q}_\alpha}\ddot{q}_\alpha \;=\; \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_\alpha\frac{\partial\mathcal{L}}{\partial\dot{q}_\alpha}\dot{q}_\alpha\right),\end{aligned} \tag{2.61}$$

so

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_\alpha\dot{q}_\alpha\frac{\partial\mathcal{L}}{\partial\dot{q}_\alpha}-\mathcal{L}\right)=0,$$

i.e.

$$H \;=\; \sum_\alpha\dot{q}_\alpha\frac{\partial\mathcal{L}}{\partial\dot{q}_\alpha}-\mathcal{L} \;=\; \text{constant}. \tag{2.62}$$

For conservative systems $\mathcal{L}=T(q,\dot{q})-V(q)$ with T homogeneous of degree 2 in $\dot{q}$, Euler's theorem gives

$$\sum_\alpha\dot{q}_\alpha\frac{\partial T}{\partial\dot{q}_\alpha}=2T, \tag{2.63}$$

hence

$$H=2T-(T-V)=T+V, \tag{2.64}$$

the conserved total mechanical energy.

Example 2.4 – Example 1: Cartesian translation symmetry

For a particle in 3D with

$$\mathcal{L}=\tfrac{1}{2}m\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right)-V(z),$$

we have $\partial\mathcal{L}/\partial x=0$. Thus

$$p_x=\frac{\partial\mathcal{L}}{\partial\dot{x}}=m\dot{x}=\text{constant},$$

i.e. the x-component of linear momentum is conserved.

Example 2.5 – Example 2: Rotational symmetry in the plane

For a particle in plane polar coordinates with

$$\mathcal{L}=\tfrac{1}{2}m\left(\dot{r}^2+r^2\dot{\phi}^2\right)-V(r),$$

the coordinate ϕ is cyclic ($\partial\mathcal{L}/\partial\phi=0$), so

$$p_\phi=\frac{\partial\mathcal{L}}{\partial\dot{\phi}}=mr^2\dot{\phi}=\text{constant}\equiv L_z,$$

the conserved z-component of angular momentum (invariance under rotations about the z-axis).

Chapter 3

The Central Force Problem

In this chapter, we consider two interacting particles through a *central* potential, i.e., a potential that depends only on the distance between the two particles. The forces acting between them obey Newton's third law, either in its weak form,

$$\vec{F}_{ij} = -\vec{F}_{ji},$$

or in its strong form, in which the action–reaction forces are collinear with the line joining the particles.

3.1. Reduction to an Equivalent One-Body Central Problem

Consider two point particles with masses m_1 and m_2, interacting through a potential U that depends only on the relative position $\vec{r}_2 - \vec{r}_1$, the relative velocity $\dot{\vec{r}}_2 - \dot{\vec{r}}_1$, or higher time derivatives thereof. The system has six degrees of freedom. Introduce the center-of-mass and relative coordinates

$$\vec{R} = \frac{m_1\vec{r}_1 + m_2\vec{r}_2}{m_1 + m_2}, \qquad \vec{r} = \vec{r}_2 - \vec{r}_1. \tag{3.1}$$

With these, the Lagrangian becomes

$$\mathcal{L} = \frac{1}{2}(m_1 + m_2)\,\dot{\vec{R}}^2 + \frac{1}{2}\,\mu\,\dot{\vec{r}}^2 - U(\vec{r}, \dot{\vec{r}}, \ldots), \qquad \mu = \frac{m_1 m_2}{m_1 + m_2}, \tag{3.2}$$

i.e. a free center-of-mass plus a single particle of mass μ moving in the interaction potential.

3.1.1. Center of Mass

Using Newton's second and third laws, the equations for each particle are

$$m_\alpha \frac{d^2\vec{r}_\alpha}{dt^2} = \vec{f}^{\,\mathrm{ext}}_\alpha + \sum_\beta \vec{f}^{\,\mathrm{int}}_{\alpha\beta}, \tag{3.3}$$

where $\vec{f}^{\,\mathrm{ext}}_\alpha$ is the external force on m_α and $\vec{f}^{\,\mathrm{int}}_{\alpha\beta}$ is the internal force on α due to β. Summing over α,

$$\frac{d^2}{dt^2}\left(\sum_\alpha m_\alpha \vec{r}_\alpha\right) = \sum_\alpha \vec{f}^{\,\mathrm{ext}}_\alpha + \sum_{\alpha,\beta} \vec{f}^{\,\mathrm{int}}_{\alpha\beta} = \vec{F}_{\mathrm{ext}}, \tag{3.4}$$

and Newton's third law ($\vec{f}^{\,\mathrm{int}}_{\alpha\beta} = -\vec{f}^{\,\mathrm{int}}_{\beta\alpha}$) gives

$$\sum_{\substack{\alpha,\beta \\ \alpha\neq\beta}} \vec{f}^{\,\mathrm{int}}_{\alpha\beta} = \vec{0}, \tag{3.5}$$

so that

$$M \equiv \sum_\alpha m_\alpha, \qquad \vec{R} \equiv \frac{1}{M}\sum_\alpha m_\alpha \vec{r}_\alpha, \qquad M\,\ddot{\vec{R}} = \vec{F}_{\mathrm{ext}}. \tag{3.6}$$

Thus the center of mass responds only to external forces.

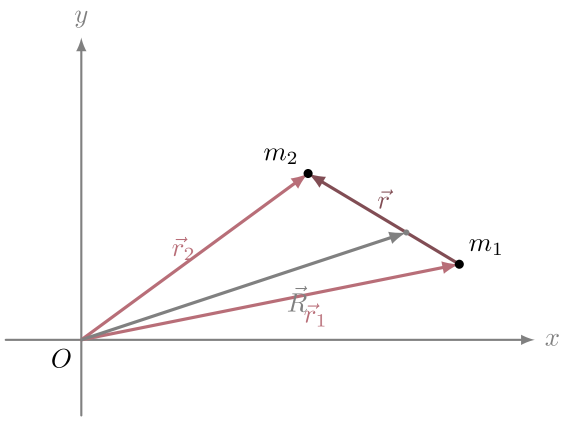


Figure 3.1: Center-of-mass and relative coordinates for a two-body system: $\vec{r} = \vec{r}_2 - \vec{r}_1$, $\vec{R} = \frac{m_1\vec{r}_1 + m_2\vec{r}_2}{m_1+m_2}$.

3.1.2. Continuous Mass Distribution

For a continuous mass distribution,

$$\vec{R} = \frac{1}{M}\int_V \vec{r}\,dm, \qquad M = \int_V dm, \tag{3.7}$$

with the integral over the occupied volume V. If the system is partitioned into sub-bodies M_i with volumes V_i and centers $\vec{R}_i$, then

$$\vec{R} = \frac{1}{M}\left[\sum_i \int_{V_i} \vec{r}\,dm\right] = \frac{\sum_i M_i\vec{R}_i}{M}, \qquad \vec{R}_i = \frac{1}{M_i}\int_{V_i} \vec{r}\,dm. \tag{3.8}$$

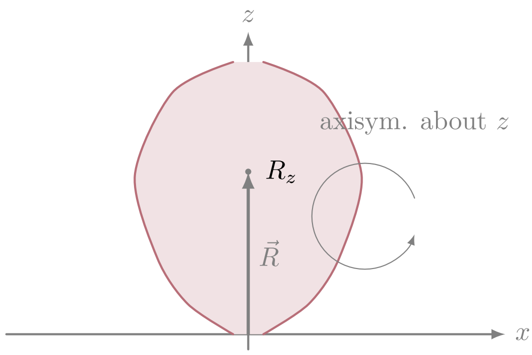


Figure 3.2: Center of mass of an axisymmetric continuous body. Rotational symmetry about z forces $\vec{R}$ to lie on the z-axis.

Theorem 3.1 – Center of Mass and Symmetry

Let $\rho(\vec{r})$ be a mass density in $\mathbb{R}^3$. If ρ is invariant under rotations about the z-axis, then the center of mass

$$\vec{R} = \frac{\displaystyle\int_V \vec{r}\,\rho(\vec{r})\,dV}{\displaystyle\int_V \rho(\vec{r})\,dV}$$

lies on that axis:

$$\int_V x\,\rho\,dV = 0, \qquad \int_V y\,\rho\,dV = 0, \qquad \Rightarrow \qquad \vec{R} = (0, 0, R_z).$$

3.1.3. Kinetic Energy

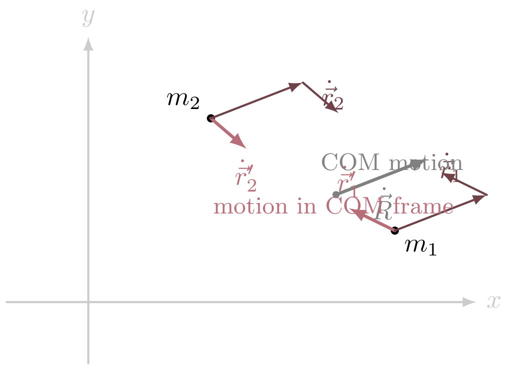


Figure 3.3: Kinetic-energy decomposition: each particle's velocity $\dot{\vec{r}}_\alpha$ splits into the center-of-mass motion $\dot{\vec{R}}$ plus the motion relative to the C.M., $\dot{\vec{r}}'_\alpha$.

The total kinetic energy of the system is

$$T = \frac{1}{2}\sum_\alpha m_\alpha\, \dot{\vec{r}}_\alpha^{\,2} = \frac{1}{2}\sum_\alpha m_\alpha\, \dot{\vec{r}}_\alpha\cdot\dot{\vec{r}}_\alpha.$$

Using

$$\dot{\vec{r}}_\alpha = \dot{\vec{R}} + \dot{\vec{r}}'_\alpha,$$

(where $\dot{\vec{r}}'_\alpha$ is the velocity of m_α in the C.M. frame) we get

$$\begin{aligned} T &= \frac{1}{2}\sum_\alpha m_\alpha \left(\dot{\vec{R}} + \dot{\vec{r}}'_\alpha\right)\cdot\left(\dot{\vec{R}} + \dot{\vec{r}}'_\alpha\right) \\ &= \frac{1}{2}\sum_\alpha m_\alpha \left[\dot{\vec{R}}^{\,2} + 2\,\dot{\vec{R}}\cdot\dot{\vec{r}}'_\alpha + \dot{\vec{r}}'^{\,2}_\alpha\right]. \end{aligned}$$

Since

$$\sum_\alpha m_\alpha \dot{\vec{r}}'_\alpha = \frac{d}{dt}\left(\sum_\alpha m_\alpha \vec{r}'_\alpha\right) = M\,\frac{d\vec{R}'}{dt} = \vec{0}$$

($\vec{R}'$ is the C.M. position relative to itself), the cross term vanishes. Hence

$$T = \frac{1}{2}M\,\dot{\vec{R}}^{\,2} + \frac{1}{2}\sum_\alpha m_\alpha\,\dot{\vec{r}}'^{\,2}_\alpha \;=\; T_{\text{C.M.}} + T_{\text{particles/C.M.}}, \tag{3.9}$$

with

$$T_{\text{C.M.}} = \tfrac{1}{2}M\,\dot{\vec{R}}^{\,2}, \qquad T'_{\text{particles/C.M.}} = \tfrac{1}{2}\sum_\alpha m_\alpha \dot{\vec{r}}'^{\,2}_\alpha.$$

For two particles,

$$\vec{r}_1 = \vec{R} + \vec{r}'_1, \quad \vec{r}_2 = \vec{R} + \vec{r}'_2, \quad \vec{r} = \vec{r}_2 - \vec{r}_1, \quad \vec{r}' = \vec{r}'_2 - \vec{r}'_1,$$

so that

$$\dot{\vec{r}}'_1 = -\frac{m_2}{m_1+m_2}\,\dot{\vec{r}}, \qquad \dot{\vec{r}}'_2 = \frac{m_1}{m_1+m_2}\,\dot{\vec{r}}.$$

Therefore,

$$\begin{aligned} m_1\dot{\vec{r}}'^{\,2}_1 + m_2\dot{\vec{r}}'^{\,2}_2 &= \left[m_1\left(\frac{m_2}{m_1+m_2}\right)^2 + m_2\left(\frac{m_1}{m_1+m_2}\right)^2\right]\dot{\vec{r}}^{\,2} \\ &= \frac{m_1 m_2}{m_1+m_2}\,\dot{\vec{r}}^{\,2} \;=\; \mu\,\dot{\vec{r}}^{\,2}, \end{aligned} \tag{3.10}$$

with the reduced mass $\mu = \dfrac{m_1 m_2}{m_1 + m_2}$. Putting this into (3.9) yields

$$T = \frac{1}{2} M \dot{\vec{R}}^2 + \frac{1}{2} \mu \dot{\vec{r}}^2, \tag{3.11}$$

and thus

$$\mathcal{L} = \frac{1}{2} M \dot{\vec{R}}^2 + \frac{1}{2} \mu \dot{\vec{r}}^2 - U(\vec{r}, \dot{\vec{r}}, \ldots). \tag{3.12}$$

Equivalently, starting from

$$\dot{\vec{r}}_1 = \dot{\vec{R}} - \frac{m_2}{m_1 + m_2} \dot{\vec{r}}, \qquad \dot{\vec{r}}_2 = \dot{\vec{R}} + \frac{m_1}{m_1 + m_2} \dot{\vec{r}},$$

we have

$$\begin{aligned} T &= \tfrac{1}{2} m_1 \dot{\vec{r}}_1^2 + \tfrac{1}{2} m_2 \dot{\vec{r}}_2^2 \\ &= \tfrac{1}{2} M \dot{\vec{R}}^2 + \tfrac{1}{2} \mu \dot{\vec{r}}^2, \end{aligned} \tag{3.13}$$

with $M = m_1 + m_2$ and $\mu = \frac{m_1 m_2}{m_1 + m_2}$.

3.2. Equations of Motion and First Integrals

The Lagrangian is

$$\mathcal{L} = \tfrac{1}{2} M \dot{\vec{R}}^2 + \tfrac{1}{2} \mu \dot{\vec{r}}^2 - U(\vec{r}, \dot{\vec{r}}, \ldots). \tag{3.14}$$

For conservative central forces, let $U = V(r)$. Then

$$\mathcal{L} = \tfrac{1}{2} M \dot{\vec{R}}^2 + \tfrac{1}{2} \mu \dot{\vec{r}}^2 - V(r). \tag{3.15}$$

Since $\vec{R}$ is cyclic (the Lagrangian does not depend on $\vec{R}$ explicitly),

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\vec{R}}}\right) = \vec{0} \;\;\Rightarrow\;\; \frac{\partial \mathcal{L}}{\partial \dot{\vec{R}}} = M \dot{\vec{R}} = \vec{P} = \text{constant}, \tag{3.16}$$

the conserved total momentum of the center of mass. Hence

$$\dot{\vec{R}} = \vec{V}_{\text{CM}} \quad \Longrightarrow \quad \vec{R}(t) = \vec{V}_{\text{CM}}\, t + \vec{R}_0. \tag{3.17}$$

This separates out the center-of-mass motion, allowing us to focus entirely on the relative dynamics. In the center of mass reference frame, the Lagrangian reduces to:

$$\mathcal{L} = \frac{1}{2} \mu \dot{\vec{r}}^2 - V(r). \tag{3.18}$$

For $V = V(r)$, the system is spherically symmetric, which implies angular momentum conservation:

$$\frac{\mathrm{d}\vec{\ell}}{\mathrm{d}t} = \vec{r} \times \vec{F} = \vec{r} \times (-\nabla V) = \vec{r} \times \left(-\frac{dV}{dr}\, \hat{r}\right) = \vec{0}. \tag{3.19}$$

Thus $\vec{\ell} = \vec{r} \times \vec{p}$ is constant. Hence $\vec{r}$ and $\vec{p}$ always lie in a plane orthogonal to $\vec{\ell}$; the motion is planar. Choosing the plane to be xy,

$$\dot{\vec{r}} = \dot{r}\, \hat{r} + r\, \dot{\theta}\, \hat{\theta}. \tag{3.20}$$

The Lagrangian becomes

$$\mathcal{L} = \frac{1}{2} \mu \left(\dot{r}^2 + r^2 \dot{\theta}^2\right) - V(r). \tag{3.21}$$

Since θ is cyclic,

$$p_\theta = \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = \mu r^2 \dot{\theta} = \text{const} \equiv \ell, \tag{3.22}$$

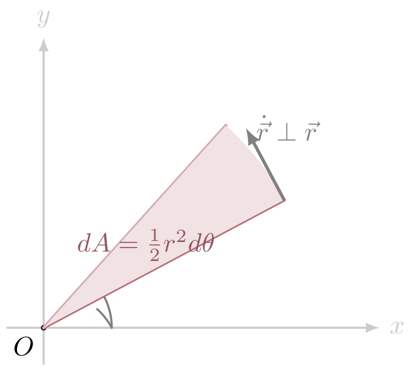


Figure 3.4: Kepler's second law: constant areal velocity $\dot{A} = \ell/(2\mu)$.

the conserved (scalar) angular momentum about $\hat{z}$. Indeed,

$$\vec{\ell} = \vec{r} \times \vec{p} = \mu\,\vec{r} \times \left(\dot{r}\,\hat{r} + r\,\dot{\theta}\,\hat{\theta}\right) = \mu r^2 \dot{\theta}\,\hat{z}.$$

The areal interpretation: the area swept in time dt is

$$dA = \frac{1}{2}\,|\vec{r} \times d\vec{r}| = \frac{1}{2}\,r^2\,d\theta \quad \Rightarrow \quad \frac{\mathrm{d}A}{\mathrm{d}t} = \frac{1}{2}\,r^2\dot{\theta} = \frac{\ell}{2\mu} = \text{const}.$$

The radial Euler–Lagrange equation,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{r}}\right) - \frac{\partial \mathcal{L}}{\partial r} = 0, \tag{3.23}$$

gives

$$\mu\,\ddot{r} - \mu\,r\,\dot{\theta}^2 + \frac{\mathrm{d}V}{\mathrm{d}r} = 0.$$

Using $\dot{\theta} = \ell/(\mu r^2)$ yields the standard radial equation

$$\mu\,\ddot{r} = -\frac{\mathrm{d}V}{\mathrm{d}r} + \frac{\ell^2}{\mu\,r^3} \;\equiv\; F_{\text{eff}}(r), \tag{3.24}$$

where the second term is the (outward) *centrifugal* contribution.
Since $\mathcal{L}$ in (3.21) has no explicit t-dependence, the energy is conserved:

$$E = T + V = \frac{1}{2}\,\mu\,\dot{r}^2 + \frac{\ell^2}{2\mu r^2} + V(r) = \frac{1}{2}\,\mu\,\dot{r}^2 + V_{\text{eff}}(r). \tag{3.25}$$

The *effective potential*

$$V_{\text{eff}}(r) = V(r) + \frac{\ell^2}{2\mu r^2} \tag{3.26}$$

collects the centrifugal term:

$$\frac{1}{2}\mu\,r^2\dot{\theta}^2 = \frac{1}{2}\mu\,r^2\left(\frac{\ell}{\mu r^2}\right)^2 = \frac{\ell^2}{2\mu r^2}.$$

We are left with a single DoF, $r(t)$, since

$$\dot{\theta}(t) = \frac{\ell}{\mu r^2(t)} \quad \Rightarrow \quad \theta(t) = \int \frac{\ell}{\mu r^2(t)}\,dt.$$

From (3.25),

$$\dot{r}^2 = \frac{2}{\mu}\left(E - V_{\text{eff}}(r)\right) \quad \Longrightarrow \quad dt = \frac{dr}{\sqrt{\frac{2}{\mu}\left(E - V_{\text{eff}}(r)\right)}}. \tag{3.27}$$

Equivalently,

$$t = \int_{r_0}^{r} \frac{dr}{\sqrt{\frac{2}{\mu}\left(E - V_{\text{eff}}(r)\right)}}.$$

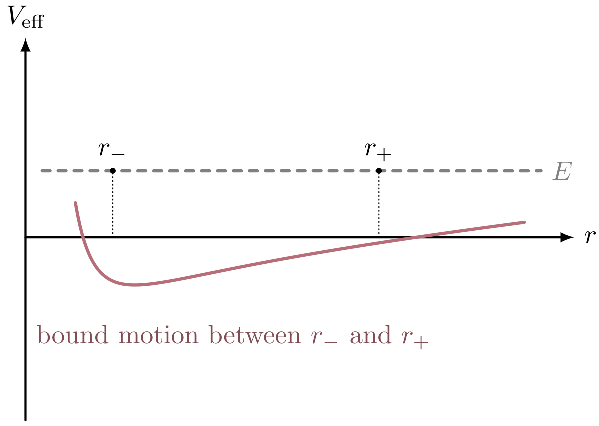


Figure 3.5: Qualitative "phase–space" (turning–point) analysis via the effective potential: for a given E, the allowed region satisfies $E \geq V_{\text{eff}}(r)$.

Because the problem reduces effectively to 1D motion in $V_{\text{eff}}(r)$, the *shape* of V_{eff} determines orbit types (bound, unbound) and turning points; see Fig. 3.5.
Full integration of (3.20) is possible for many power–law potentials $V(r) \propto r^n$ ($n = 1, 2, -1, \ldots$). The effective potential (3.26) dictates the orbit class (open/closed, periodic, etc.).

Coulomb potential. For

$$V(r) = -\frac{k}{r}, \qquad k > 0,$$

the effective potential is

$$V_{\text{eff}}(r) = -\frac{k}{r} + \frac{\ell^2}{2\mu r^2}. \tag{3.28}$$

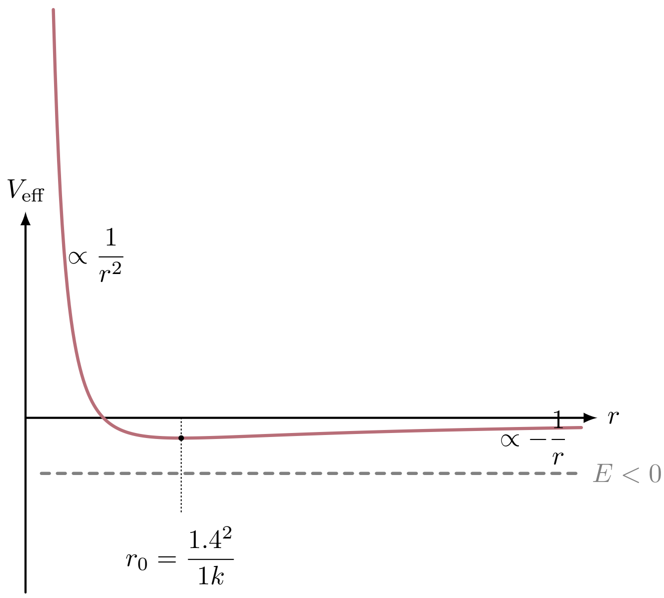


Figure 3.6: Effective potential $V_{\text{eff}}(r) = -k/r + \ell^2/(2\mu r^2)$ for the Coulomb problem. The minimum at $r_0 = \ell^2/(\mu k)$ corresponds to a circular orbit for $E = V_{\text{eff}}(r_0)$.

3.3. Isotropic Harmonic Oscillator

For the isotropic harmonic oscillator, the force and potential are

$$f(r) = -kr, \qquad V(r) = \frac{1}{2}kr^2.$$

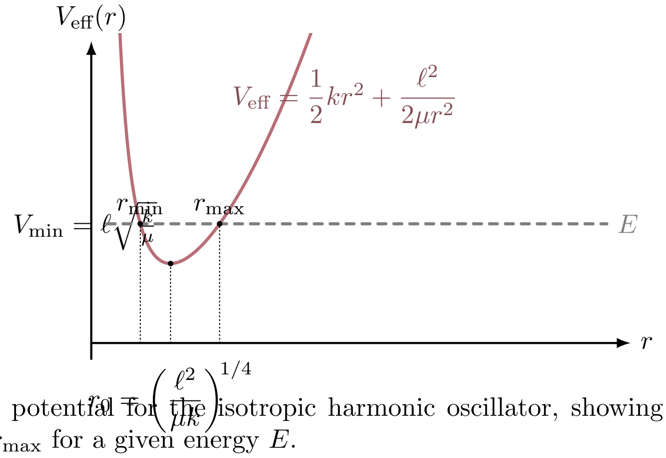


Figure 3.7: Effective potential for the isotropic harmonic oscillator, showing the minimum at r_0 and turning points $r_{\rm min}, r_{\rm max}$ for a given energy E.

The effective potential is

$$V_{\rm eff}(r) = \frac{1}{2}kr^2 + \frac{\ell^2}{2\mu r^2}. \tag{3.29}$$

A *circular* orbit occurs at an extremum of $V_{\rm eff}$, i.e.

$$\frac{\mathrm{d}V_{\rm eff}}{\mathrm{d}r} = 0 \implies kr - \frac{\ell^2}{\mu r^3} = 0 \implies r = r_0 = \left(\frac{\ell^2}{\mu k}\right)^{1/4}. \tag{3.30}$$

The *turning points* of the radial motion are obtained instead from the energy condition

$$E = V_{\rm eff}(r) \quad \text{at} \quad r = r_{\rm min},\ r_{\rm max}, \tag{3.31}$$

where $\dot{r} = 0$. The radial period can be written as

$$T = 2\int_{r_{\rm min}}^{r_{\rm max}} \frac{dr}{\sqrt{\frac{2}{\mu}\left(E - V_{\rm eff}(r)\right)}}. \tag{3.32}$$

Near the minimum, expand $V_{\rm eff}$ about r_0:

$$V_{\rm eff}(r) = V_{\rm min} + \underbrace{\left.\frac{\mathrm{d}V_{\rm eff}}{\mathrm{d}r}\right|_{r_0}}_{=0}(r - r_0) + \frac{1}{2}K\,(r - r_0)^2, \qquad K = \left.\frac{\mathrm{d}^2V_{\rm eff}}{\mathrm{d}r^2}\right|_{r_0}. \tag{3.33}$$

Hence

$$V_{\rm eff}(r) = V_{\rm min} + \frac{1}{2}K(r - r_0)^2, \qquad K = \left.\frac{\mathrm{d}^2V_{\rm eff}}{\mathrm{d}r^2}\right|_{r_0}, \tag{3.34}$$

and the small–oscillation period is

$$T = 2\pi\sqrt{\frac{\mu}{K}}, \tag{3.35}$$

valid for $E = V_{\rm min} + \epsilon$ with $\epsilon \ll 1$. *(For the isotropic HO one finds $K = k + 3\ell^2/(\mu r_0^4) = 4k$, so $T = \pi\sqrt{\mu/k}$.)*

3.4. Conditions for Closed Orbits and Stability Analysis

Classification of Orbits in a Central Potential

From earlier,

$$E = \frac{1}{2}\mu\dot{r}^2 + V_{\rm eff}(r), \qquad V_{\rm eff}(r) = V(r) + \frac{\ell^2}{2\mu r^2}. \tag{3.36}$$

The planar velocity is

$$\dot{\vec{r}} = \dot{r}\,\hat{r} + r\dot{\theta}\,\hat{\theta}. \tag{3.37}$$

When $\dot{r} = 0$ the motion is purely tangential. Allowed motion requires $E \geq V_{\text{eff}}(r)$; intersections of E with V_{eff} give turning points $r_{\min}, r_{\max}$. With fixed ℓ,

$$\dot{\theta} = \frac{\ell}{\mu r^2}, \tag{3.38}$$

which (by choosing the sign of ℓ) can be taken positive, so $\theta(t)$ is monotonic.
Equivalently,

$$\dot{r}^2 = \frac{2}{\mu}\big(E - V(r)\big) - \frac{\ell^2}{\mu^2 r^2}. \tag{3.39}$$

The total angular advance during one radial oscillation is

$$\Delta\theta = 2\int_{r_{\min}}^{r_{\max}} \frac{\dot{\theta}}{\dot{r}}\, dr = 2\int_{r_{\min}}^{r_{\max}} \frac{\ell\, dr}{\mu r^2 \sqrt{\frac{2}{\mu}\big(E - V_{\text{eff}}(r)\big)}} \tag{3.40}$$

$$= 2\int_{r_{\min}}^{r_{\max}} \frac{\ell\, dr}{\mu r^2 \sqrt{\frac{2}{\mu}\big(E - V(r)\big) - \dfrac{\ell^2}{\mu^2 r^2}}}. \tag{3.41}$$

The orbit is closed iff

$$\Delta\theta = 2\pi\,\frac{m}{n}, \qquad m, n \in \mathbb{Z}^+. \tag{3.42}$$

Otherwise, the trajectory is not commensurate and eventually fills an annulus.
The only central potentials for which *all* bound orbits are closed are

$$V(r) \propto \frac{1}{r} \quad \text{and} \quad V(r) \propto r^2.$$

The centrifugal term $\ell^2/(2\mu r^2) \to +\infty$ as $r \to 0$, preventing approach to the origin unless the attractive potential diverges sufficiently strongly, e.g.

$$V(r) = -\frac{\alpha}{r^n}, \quad \alpha > 0, \quad n \geq 2,$$

in which case fall to the center can occur (depending on parameters).

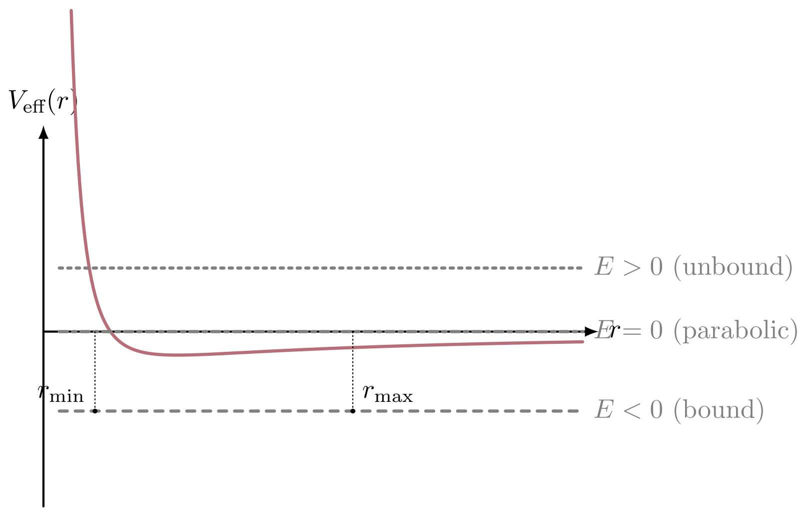


Figure 3.8: Typical effective potential with centrifugal barrier (e.g. attractive $1/r$). Energy levels determine orbit type: $E < 0$ bound (two turning points), $E = 0$ parabolic, $E > 0$ hyperbolic (one turning point).

Circular Orbits

Circular orbits are possible when V_{eff} has an extremum at $r = r_0$:

$$\left.\frac{\mathrm{d}V_{\text{eff}}}{\mathrm{d}r}\right|_{r_0} = \left.\frac{\mathrm{d}V}{\mathrm{d}r}\right|_{r_0} - \frac{\ell^2}{\mu r_0^3} = 0 \quad \Rightarrow \quad f(r_0) = -\left.\frac{\mathrm{d}V}{\mathrm{d}r}\right|_{r_0} = -\frac{\ell^2}{\mu r_0^3} < 0, \tag{3.43}$$

so the (centripetal) force must be attractive. Stability requires

$$\left.\frac{\mathrm{d}^2V_{\text{eff}}}{\mathrm{d}r^2}\right|_{r_0} = \left.\frac{\mathrm{d}^2V}{\mathrm{d}r^2}\right|_{r_0} + \frac{3\ell^2}{\mu r_0^4} > 0, \tag{3.44}$$

or, in terms of $f(r) = -dV/dr$,

$$-\left.\frac{\mathrm{d}f}{\mathrm{d}r}\right|_{r_0} - \frac{3}{r_0} f(r_0) > 0 \quad \Longleftrightarrow \quad \left.\frac{\mathrm{d}\ln f}{\mathrm{d}\ln r}\right|_{r_0} > -3.$$

Applying the condition in (3.4) for a central force model:

$$f(r) = -Kr^n, \quad K > 0, \tag{3.45}$$

gives:

$$\left.\frac{\partial f}{\partial r}\right|_{r=r_0} = -nKr_0^{n-1} < -\frac{3}{r_0} f(r_0) = 3Kr_0^{n-1}. \tag{3.46}$$

This leads to:

$$n > -3. \tag{3.47}$$

3.5. Stability of Circular Orbits

Circular orbits occur when $E = V_{\text{eff}}(r_0)$, and $V_{\text{eff}}(r)$ has a minimum. Then a small increase $E' = E + \epsilon$, $\epsilon \ll 1$, should result in small oscillatory motion about the circular orbit $r = r_0$.
Using the radial equation of motion:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{r}} - \frac{\partial \mathcal{L}}{\partial r} = 0, \tag{3.48}$$

we get

$$\mu\ddot{r} - \mu r\dot{\theta}^2 = -\frac{\partial V}{\partial r} \equiv f(r), \tag{3.49}$$

and substituting $\mu r^2\dot{\theta} = \ell$, this becomes

$$\mu\ddot{r} - \frac{\ell^2}{\mu r^3} = f(r). \tag{3.50}$$

Let

$$r = r_0 + \epsilon, \quad \epsilon \ll r_0, \tag{3.51}$$

so that we can expand in a Taylor series:

$$\frac{1}{r^3} = \frac{1}{(r_0+\epsilon)^3} = \frac{1}{r_0^3}\frac{1}{\left(1+\frac{\epsilon}{r_0}\right)^3}. \tag{3.52}$$

Using the binomial expansion for $(1+x)^{-3}$, we write:

$$\frac{1}{\left(1+\frac{\epsilon}{r_0}\right)^3} \approx 1 - 3\frac{\epsilon}{r_0}, \tag{3.53}$$

which gives:

$$\frac{1}{r^3} \approx \frac{1}{r_0^3}\left(1 - 3\frac{\epsilon}{r_0}\right) = \frac{1}{r_0^3} - \frac{3\epsilon}{r_0^4}. \tag{3.54}$$

Expanding $f(r)$ around r_0 using a Taylor series gives:

$$f(r) = f(r_0 + \epsilon) = f(r_0) + f'(r_0)\,\epsilon + \mathcal{O}(\epsilon^2). \tag{3.55}$$

Using (3.50), the radial equation becomes

$$\mu\ddot{\epsilon} - \frac{\ell^2}{\mu}\left(\frac{1}{r_0^3} - \frac{3\epsilon}{r_0^4}\right) = f(r_0) + f'(r_0)\,\epsilon.$$

At $r = r_0$, we have the circular–orbit condition

$$f(r_0) = -\frac{\ell^2}{\mu r_0^3}. \tag{3.56}$$

Substituting (3.56) and simplifying gives

$$\mu\ddot{\epsilon} + \frac{3\ell^2}{\mu r_0^4}\,\epsilon = f'(r_0)\,\epsilon. \tag{3.57}$$

Dividing through by μ, we find:

$$\ddot{\epsilon} + \epsilon\left(\frac{3\ell^2}{\mu^2 r_0^4} - \frac{f'(r_0)}{\mu}\right) = 0. \tag{3.58}$$

Thus, ϵ satisfies the ODE

$$\ddot{\epsilon} + \Omega^2\epsilon = 0, \qquad \Omega^2 = \frac{3\ell^2}{\mu^2 r_0^4} - \frac{f'(r_0)}{\mu} > 0. \tag{3.59}$$

For an attractive inverse–power force

$$f(r) = -\frac{K}{r^n} \quad (K > 0), \tag{3.60}$$

we have

$$f'(r) = n\,\frac{K}{r^{n+1}}. \tag{3.61}$$

Substituting $f'(r_0) = n\dfrac{K}{r_0^{n+1}}$ into (3.59), we get:

$$\Omega^2 = \frac{3\ell^2}{\mu^2 r_0^4} - n\,\frac{K}{\mu r_0^{n+1}} > 0. \tag{3.62}$$

From the equilibrium condition (3.56),

$$-\frac{K}{r_0^n} = -\frac{\ell^2}{\mu r_0^3} \quad\Rightarrow\quad K = \frac{\ell^2}{\mu}\,r_0^{n-3}. \tag{3.63}$$

Substituting this into (3.62) yields

$$\Omega^2 = \frac{3\ell^2}{\mu^2 r_0^4} - n\,\frac{\ell^2 r_0^{n-3}}{\mu^2 r_0^{n+1}} = \frac{\ell^2}{\mu^2 r_0^4}\,(3 - n).$$

For stability ($\Omega^2 > 0$) we require

$$n < 3. \tag{3.64}$$

This is the stability requirement for circular orbits under an attractive inverse–power force. The limiting case $n = 3$ should be treated separately.

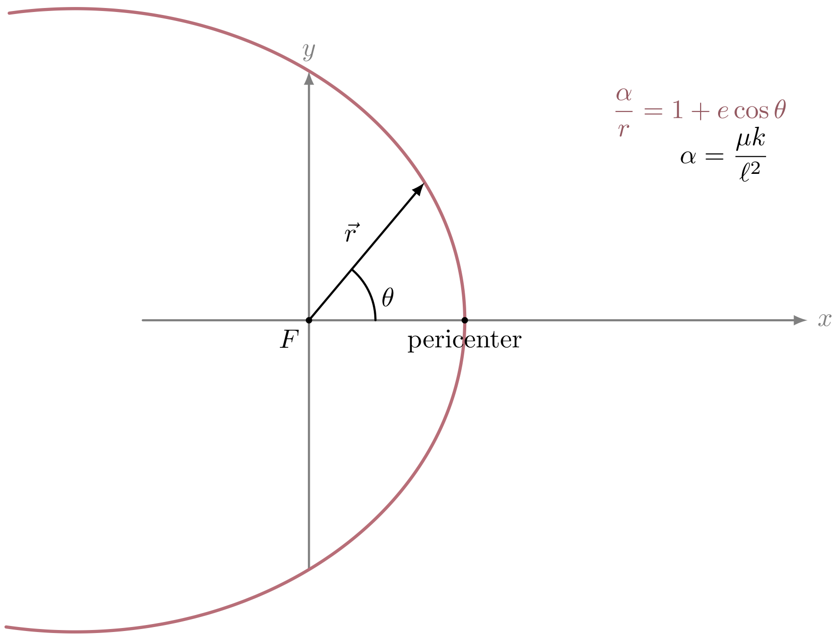


Figure 3.9: Typical Kepler orbit (ellipse, $e < 1$) with focus at the origin: $\alpha/r = 1 + e\cos\theta$.

3.6. The Kepler Problem: Inverse Square Law of Force

Historically, finding the orbits under a central force

$$f(r) = -\frac{k}{r^2}, \qquad V(r) = -\frac{k}{r}, \tag{3.65}$$

was one of the most important problems and deserved special attention because Kepler's three laws of planetary motion are based on it. The success of Newton's laws in explaining Kepler's assertions validated Newton's theory of planetary motion, which is based on the inverse–square law.
For the effective potential,

$$V_{\text{eff}}(r) = -\frac{k}{r} + \frac{\ell^2}{2\mu r^2}, \tag{3.66}$$

the path can be obtained by quadrature using conservation of angular momentum and energy. From

$$\ell = \mu r^2 \dot{\theta} \quad \Rightarrow \quad \theta(t) = \frac{\ell}{\mu}\int \frac{dt}{r^2(t)},$$

and

$$\frac{dr}{dt} = \frac{dr}{d\theta}\,\dot{\theta} = \frac{\ell}{\mu r^2}\,\frac{dr}{d\theta},$$

energy conservation gives

$$\frac{d\theta}{dr} = \frac{\mu r^2}{\ell}\,\frac{1}{\sqrt{\frac{2}{\mu}\big(E - V_{\text{eff}}(r)\big)}}.$$

Thus

$$\theta(r) = \int \frac{\ell\, dr/r^2}{\sqrt{\frac{2}{\mu}\,(E - V_{\text{eff}}(r))}}. \tag{3.67}$$

With the change of variable $u = 1/r$ (so $dr = -du/u^2$),

$$\theta(u) = \int \frac{du}{\sqrt{\dfrac{2\mu E}{\ell^2} + \dfrac{2\mu k}{\ell^2}u - u^2}}. \tag{3.68}$$

Completing the square shows the integral is elementary, leading to the standard conic form

$$u(\theta) \equiv \frac{1}{r(\theta)} = \alpha\big(1 + e\cos(\theta - \theta_0)\big), \qquad \alpha = \frac{\mu k}{\ell^2}, \quad e = \sqrt{1 + \frac{2E\ell^2}{\mu k^2}}, \tag{3.69}$$

where θ_0 is a constant set by initial conditions. Choosing the polar angle so that $\theta_0 = 0$ (pericenter at $\theta = 0$) yields

$$\frac{\alpha}{r} = 1 + e\cos\theta. \tag{3.70}$$

Orbit types. The trajectory depends on e (equivalently on the energy E):

$e > 1$	$E > 0$	Hyperbola
$e = 1$	$E = 0$	Parabola
$0 < e < 1$	$V_{\min} < E < 0$	Ellipse
$e = 0$	$E = V_{\min} = -\dfrac{\mu k^2}{2\ell^2}$	Circle

If instead of $r(\theta)$ we want $r(t)$ and $\theta(t)$, note that

$$\dot\theta = \frac{\ell}{\mu r^2} \quad \Rightarrow \quad dt = \frac{\mu r^2}{\ell}\, d\theta, \tag{3.71}$$

hence

$$t(\theta) = \frac{\mu}{\ell}\int r^2(\theta)\, d\theta. \tag{3.72}$$

(For bound ellipses, this yields Kepler's equation in terms of the eccentric anomaly.)

The Laplace–Runge–Lenz Vector

The Kepler problem admits an additional conserved vector besides $\vec\ell$. Using reduced mass μ and $\vec F = -k\,\hat r/r^2$,

$$\frac{d}{dt}\left(\vec p \times \vec\ell\right) = \dot{\vec p} \times \vec\ell = \vec F \times (\vec r \times \vec p) = \mu k\,\dot{\hat r}, \tag{3.73}$$

so

$$\frac{d}{dt}\left(\vec p \times \vec\ell - \mu k\,\hat r\right) = 0. \tag{3.74}$$

Define the Laplace–Runge–Lenz (LRL) vector

$$\vec A \equiv \vec p \times \vec\ell - \mu k\,\hat r, \qquad \frac{d\vec A}{dt} = 0. \tag{3.75}$$

Since $\vec A \perp \vec\ell$, it lies in the orbital plane and points toward pericenter. Taking the dot product with $\hat r$,

$$\vec A \cdot \hat r = (\vec p \times \vec\ell) \cdot \hat r - \mu k = \frac{\ell^2}{r} - \mu k, \tag{3.76}$$

where $(\vec p \times \vec\ell) \cdot \hat r = \ell^2/r$. Solving for $1/r$,

$$\frac{1}{r} = \frac{\mu k}{\ell^2}\left(1 + \frac{A}{\mu k}\cos\Theta\right), \tag{3.77}$$

where Θ is the angle between $\vec A$ and $\hat r$. Comparing with (3.70) shows the eccentricity is

$$e = \frac{A}{\mu k}. \tag{3.78}$$

Chapter 4

The Kinematics of Rigid Body Motion

In this chapter we introduce the kinematics of rigid bodies. Our emphasis is on (i) how to parameterize orientation, (ii) how vectors transform between a fixed (inertial) frame and a body–fixed frame, and (iii) the time–derivative relation that yields the angular–velocity vector $\boldsymbol{\omega}$. These tools are precisely what we will need in the next chapter to set up the dynamics via the Lagrangian formalism (building on the results already derived in Chs. 1–2; see, e.g., (1.141) for rotating frames).

4.1. Independent Coordinates of a Rigid Body

A rigid body is a set of points $\{\vec{r}_i\}_{i=1}^N$ with all pairwise distances fixed,

$$|\vec{r}_i - \vec{r}_j| = C_{ij} = \text{const}, \qquad \forall i, j. \tag{4.1}$$

Although there are $3N$ coordinates and $\frac{1}{2}N(N-1)$ constraints, those constraints are highly dependent. Fixing three non–collinear material points suffices to determine the entire body configuration; that is $3 \times 3 = 9$ coordinates. The three inter–point distances among those reference points are fixed,

$$r_{12} = C_{12}, \qquad r_{13} = C_{13}, \qquad r_{23} = C_{23}, \tag{4.2}$$

reducing the independent count by three. Hence a free rigid body in space has

$$9 - 3 = 6$$

degrees of freedom: three for translation of a chosen body origin, and three for orientation.

Example 4.1 – A second counting argument (geometric)

Choose a body origin P_1 (3 DoF). Given P_1, the second point P_2 lies on a sphere of radius C_{12} about P_1 (2 angles $\Rightarrow$ 2 DoF). Given P_1, P_2, the third point P_3 lies on the circle formed by intersection of two spheres (4.2); one angle φ on that circle fixes P_3 (1 DoF). Total $3 + 2 + 1 = 6$.

4.2. Direction Cosines and the Rotation (DCM) Matrix

Let $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$ be the fixed (space) basis and $\{\hat{\imath}', \hat{\jmath}', \hat{k}'\}$ a body–fixed basis rigidly attached to the body. Define the *direction cosines*

$$\cos\theta_{ij} \equiv \hat{e}'_i \cdot \hat{e}_j, \qquad i, j \in \{1, 2, 3\},$$

where $(\hat{e}_1, \hat{e}_2, \hat{e}_3) = (\hat{\imath}, \hat{\jmath}, \hat{k})$. Then

$$\hat{e}'_i = \sum_{j=1}^{3} \cos\theta_{ij}\, \hat{e}_j, \qquad i = 1, 2, 3. \tag{4.3}$$

Collect the $\cos\theta_{ij}$ into the 3×3 *direction–cosine matrix* (DCM)

$$R \;\equiv\; \big(\cos\theta_{ij}\big)_{i,j=1}^{3}.$$

For any vector $\vec{r}$, let $x = (x, y, z)^\top$ denote its components in the fixed frame and $x' = (x', y', z')^\top$ its components in the body frame. Using $\vec{r} = x_j \hat{e}_j = x'_i \hat{e}'_i$ and (4.3),

$$x' = R\,x, \qquad x = R^\top x'. \tag{4.4}$$

Because the two bases are orthonormal, R is orthogonal:

$$RR^\top = R^\top R = I, \qquad \det R = +1. \tag{4.5}$$

Elementwise, orthogonality gives the two families of identities

$$\sum_{j=1}^{3} \cos\theta_{ij}\cos\theta_{kj} = \delta_{ik}, \qquad \sum_{i=1}^{3} \cos\theta_{ij}\cos\theta_{ik} = \delta_{jk}, \tag{4.6}$$

which reduce the nine $\cos\theta_{ij}$ to only three independent parameters (e.g., Euler angles).

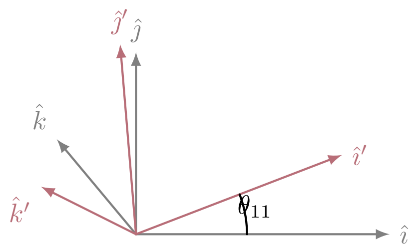


Figure 4.1: Fixed frame $(\hat{\imath}, \hat{\jmath}, \hat{k})$ and body frame $(\hat{\imath}', \hat{\jmath}', \hat{k}')$. The DCM R has entries $R_{ij} = \hat{e}'_i \cdot \hat{e}_j$ and maps fixed components to body components via (4.4).

4.3. Angular Velocity and Time Derivatives in the Two Frames

Because the body frame rotates, the direction cosines (hence $R(t)$) are time–dependent. Differentiate the identity $RR^\top = I$ to get

$$\dot{R}\,R^\top + R\,\dot{R}^\top = 0,$$

so both $\dot{R}\,R^\top$ and $\dot{R}^\top R$ are skew–symmetric matrices. Define the cross–product (hat) map $[\,\cdot\,]_\times$ by $[\,\omega\,]_\times v = \boldsymbol{\omega} \times v$. Then the (unique) angular–velocity vector satisfies

$$[\boldsymbol{\omega}]_\times^{(S)} = \dot{R}^\top R, \qquad [\boldsymbol{\omega}]_\times^{(B)} = R\,\dot{R}^\top, \tag{4.7}$$

where the superscripts indicate whether the components of $\boldsymbol{\omega}$ are taken in the fixed (S) or body (B) basis.[1]

Theorem 4.1 – Transport theorem (kinematic relation)

For any vector $\vec{A}$,

$$\left(\frac{d\vec{A}}{dt}\right)_S = \left(\frac{d\vec{A}}{dt}\right)_B + \boldsymbol{\omega} \times \vec{A}, \tag{4.8}$$

[1] These are equivalent: $[\omega]_\times^{(B)} = R\,[\omega]_\times^{(S)}\,R^\top$.

with $\boldsymbol{\omega}$ given by (4.7). In particular, if $\vec{A}$ is rigidly attached to the body (its body components are constant), then $(d\vec{A}/dt)_B = 0$ and

$$\left(\frac{d\vec{A}}{dt}\right)_S = \boldsymbol{\omega} \times \vec{A}.$$

Example 4.2 – Velocity field of a rigid body

Let O be a body–fixed point with space velocity $\vec{V}_O$. Any material point P has position $\vec{r}_P = \vec{r}_O + \vec{r}_{OP}$ with $\vec{r}_{OP}$ fixed in the body. By (4.8), $\dot{\vec{r}}_{OP}|_S = \boldsymbol{\omega} \times \vec{r}_{OP}$. Hence

$$\vec{v}_P = \vec{V}_O + \boldsymbol{\omega} \times \vec{r}_{OP}. \tag{4.9}$$

Taking another time derivative recovers the rotating–frame acceleration terms listed earlier in (1.141).

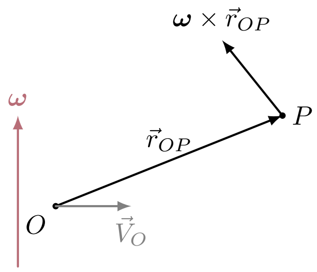


Figure 4.2: Rigid–body velocity field (4.9): $\vec{v}_P = \vec{V}_O + \boldsymbol{\omega} \times \vec{r}_{OP}$.

4.4. Euler Angles (3–1–3) and ω

Any proper rotation can be parameterized by three angles. Consistent with our earlier usage, we adopt the z–x'–z'' (3–1–3) Euler angles (ϕ, θ, ψ): rotate by ϕ about $\hat{k}$, then by θ about the new $\hat{\imath}'$, then by ψ about the final $\hat{k}'' = \hat{k}'$. The DCM $R(\phi, \theta, \psi)$ mapping fixed components to body components is orthogonal (explicit entries omitted here for brevity; they follow from the product of the elementary rotations).
The angular–velocity components in the *body* frame are

$$\boxed{\begin{aligned} \omega_1 &= \dot{\phi}\,\sin\theta\,\sin\psi + \dot{\theta}\,\cos\psi, \\ \omega_2 &= \dot{\phi}\,\sin\theta\,\cos\psi - \dot{\theta}\,\sin\psi, \\ \omega_3 &= \dot{\phi}\,\cos\theta + \dot{\psi}. \end{aligned}} \tag{4.10}$$

(These are obtained either by differentiating $R(\phi, \theta, \psi)$ and using (4.7), or by composing the instantaneous rotations about z, x', z'' and projecting onto the body axes.)

Remarks.

- The matrix R contains only orientation (no translation). Translation of a chosen point (e.g., center of mass) contributes the $\vec{V}_O$ term in (4.9).
- The pair $(\vec{V}_O, \boldsymbol{\omega})$ completely specifies the instantaneous motion of a rigid body (*Chasles' theorem*).
- For planar motion (say about $\hat{k}$), $\boldsymbol{\omega} = \dot{\theta}\,\hat{k}$ and (4.9) reduces to the familiar $v_P = v_O + \dot{\theta}\,\hat{k} \times r_{OP}$.

4.5. Orthogonal Transformations

We relabel the Cartesian components by $x\to x_1,\ y\to x_2,\ z\to x_3$ and write

$$a_{ij} \equiv \cos\theta_{ij}.$$

The linear transformation from fixed-frame components (x_1, x_2, x_3) to body-frame components (x'_1, x'_2, x'_3) is (cf. (4.4))

$$\begin{aligned} x'_1 &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ x'_2 &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ x'_3 &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3. \end{aligned} \tag{4.11}$$

In compact/matrix form,

$$\vec{x}' = A\,\vec{x}, \qquad A_{ij} = a_{ij}, \qquad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \tag{4.12}$$

Rotations preserve lengths:

$$\vec{x}\cdot\vec{x} = \vec{x}'\cdot\vec{x}' \qquad \Longleftrightarrow \qquad \sum_i x_i^2 = \sum_i {x'_i}^2. \tag{4.13}$$

Using (4.12),

$$\sum_i {x'_i}^2 = \sum_i \Big(\sum_j a_{ij}x_j\Big)\Big(\sum_k a_{ik}x_k\Big) = \sum_{j,k} x_j x_k \Big(\sum_i a_{ij}a_{ik}\Big) = \sum_j x_j^2,$$

hence we must have

$$\sum_i a_{ij}a_{ik} = \delta_{jk} \qquad \Longleftrightarrow \qquad A^T A = I, \tag{4.14}$$

i.e. A is an *orthogonal* matrix (and for proper rotations $\det A = +1$).

Rotation about the z-axis by an angle θ

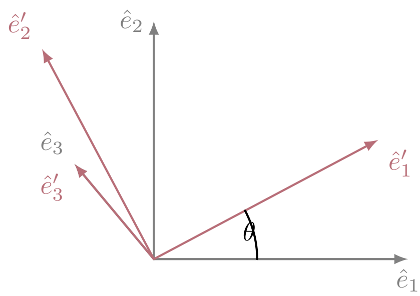


Figure 4.3: A right–hand rotation by θ about $\hat{e}_3$ carries $(\hat{e}_1, \hat{e}_2)$ to $(\hat{e}'_1, \hat{e}'_2)$ while $\hat{e}'_3 = \hat{e}_3$.

With our convention $\vec{x}' = A\vec{x}$ (fixed components $\to$ body components), the action on basis vectors is

$$A\hat{e}_3 = \hat{e}'_3 = \hat{e}_3, \qquad \hat{e}'_1 = \hat{e}_1\cos\theta + \hat{e}_2\sin\theta, \qquad \hat{e}'_2 = -\hat{e}_1\sin\theta + \hat{e}_2\cos\theta,$$

so

$$A = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Therefore the component relations are

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \Rightarrow \quad \begin{cases} x'_1 = \ \ x_1\cos\theta + x_2\sin\theta, \\ x'_2 = -x_1\sin\theta + x_2\cos\theta, \\ x'_3 = x_3. \end{cases}$$

(If you instead map body components to fixed components, you use A^T.)
Rotations are linear:

$$A(\alpha\vec{x}) = \alpha\, A\vec{x}, \qquad A(\vec{x}+\vec{y}) = A\vec{x} + A\vec{y}, \tag{4.15}$$

and are completely specified by their action on a basis,

$$A(\hat{e}_1), \qquad A(\hat{e}_2), \qquad A(\hat{e}_3). \tag{4.16}$$

Since $\vec{x}' = A\vec{x}$,

$$\vec{x}' \cdot \vec{x}' = (\vec{x}')^T(\vec{x}') = \vec{x}^{\,T}(A^T A)\,\vec{x}.$$

Because (4.13) holds for all $\vec{x}$, it follows that

$$A^T A = I, \tag{4.17}$$

i.e. A is orthogonal.

4.6. Formal Properties of the Transformation Matrix

Let $C = AB$. Then

$$C_{ij} = \sum_k A_{ik} B_{kj}, \tag{4.18}$$

and, in general, $AB \neq BA$ (non-commutative), while $(AB)C = A(BC)$ (associative). For addition,

$$C = A + B \quad\Longrightarrow\quad C_{ij} = A_{ij} + B_{ij}. \tag{4.19}$$

Moreover, orthogonal matrices satisfy $A^{-1} = A^T$ and, for proper rotations, $\det A = +1$.
Since by definition of an inverse matrix A^{-1}, we have:

$$A^{-1}A = AA^{-1} = I \tag{4.20}$$

If A is also orthogonal, like rotation matrices, then:

$$A^T A = AA^T = I \implies A^{-1} = A^T \tag{4.21}$$

This simple result has an important consequence since usually the computation of A^{-1} is very demanding. However, for orthogonal matrices, it has a very simple form:

$$A^{-1} = A^T \implies (A^{-1})_{ij} = (A^T)_{ij} = a_{ji} \tag{4.22}$$

Since $|A^T| = |A|$ for square $n \times n$ matrices, then:

$$A^T A = 1 \implies |A|^2 = 1 \implies |A| = \pm 1 \tag{4.23}$$

Therefore, the determinant of orthogonal matrices is ± 1.

Let $\hat{e}_i$ and $\hat{e}_i'$ be the unit vectors in the fixed and rotating frames, respectively. Then:

$$\vec{X}' = A\vec{X} \implies A^{-1}\vec{X}' = \vec{X} \quad \text{or} \quad \vec{X} = A^T\vec{X}' \tag{4.24}$$

with:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \tag{4.25}$$

Rows of the matrix A.

$$\hat{e}_1' = a_{11}\hat{e}_1 + a_{12}\hat{e}_2 + a_{13}\hat{e}_3, \qquad \hat{e}_2' = a_{21}\hat{e}_1 + a_{22}\hat{e}_2 + a_{23}\hat{e}_3, \qquad \hat{e}_3' = a_{31}\hat{e}_1 + a_{32}\hat{e}_2 + a_{33}\hat{e}_3.$$

Columns of the matrix A. Equivalently,

$$\hat{e}_1 = a_{11}\hat{e}_1' + a_{21}\hat{e}_2' + a_{31}\hat{e}_3', \quad \hat{e}_2 = a_{12}\hat{e}_1' + a_{22}\hat{e}_2' + a_{32}\hat{e}_3', \quad \hat{e}_3 = a_{13}\hat{e}_1' + a_{23}\hat{e}_2' + a_{33}\hat{e}_3'.$$

(Columns of A are the components of the fixed axes in the primed basis; rows are the components of the primed axes in the fixed basis. Since $A^T A = I$, $A^{-1} = A^T$.)

Orthogonality of the primed basis.

$$\hat{e}_i' \cdot \hat{e}_j' = \delta_{ij}$$

yields, for $i = j$,

$$a_{11}^2 + a_{12}^2 + a_{13}^2 = 1, \qquad a_{21}^2 + a_{22}^2 + a_{23}^2 = 1,$$
$$a_{31}^2 + a_{32}^2 + a_{33}^2 = 1, \tag{4.26}$$

and for $i \neq j$,

$$a_{11}a_{21} + a_{12}a_{22} + a_{13}a_{23} = 0, \qquad a_{11}a_{31} + a_{12}a_{32} + a_{13}a_{33} = 0,$$
$$a_{21}a_{31} + a_{22}a_{32} + a_{23}a_{33} = 0. \tag{4.27}$$

Unitary/orthogonal and inverse. A unitary matrix U satisfies

$$U^\dagger U = UU^\dagger = I \implies U^\dagger = U^{-1}. \tag{4.28}$$

For a real rotation matrix A (an orthogonal matrix),

$$A^\dagger = A^T = A^{-1}. \tag{4.29}$$

Remark. Orthogonal does <u>not</u> mean Hermitian. A generic 3D proper rotation has eigenvalues $\{1, e^{i\theta}, e^{-i\theta}\}$ (unit modulus), not all real. Only symmetric orthogonal matrices are Hermitian, and then their eigenvalues are ± 1.

Transpose of a product (fix).

$$(AB)^T_{ij} = (AB)_{ji} = \sum_k A_{jk}B_{ki} = \sum_k B^T_{ik}A^T_{kj} = (B^T A^T)_{ij},$$

hence

$$(AB)^T = B^T A^T. \tag{4.30}$$

4.7. Euler Angles

We specify the orientation via three successive rotations (our passive, component-mapping convention is the same used in (**??**)–(4.32)).

1) Rotation about the z-axis by ϕ

$$A_\phi = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \vec{X}_1' = A_\phi \vec{X}. \tag{4.31}$$

2) Rotation about the X'-axis by θ

We now rotate about $\hat{e}_1'$ by θ, taking $(\hat{e}_1', \hat{e}_2', \hat{e}_3')$ to $(\hat{e}_1'', \hat{e}_2'', \hat{e}_3'')$. The bases obey

$$\hat{e}_1' = \hat{e}_1'', \qquad \hat{e}_2' = \cos\theta\,\hat{e}_2'' - \sin\theta\,\hat{e}_3'', \qquad \hat{e}_3' = \sin\theta\,\hat{e}_2'' + \cos\theta\,\hat{e}_3'',$$

so the corresponding matrix (consistent with these relations) is

$$A_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \qquad \vec{X}_2'' = A_\theta \vec{X}_1'. \tag{4.32}$$

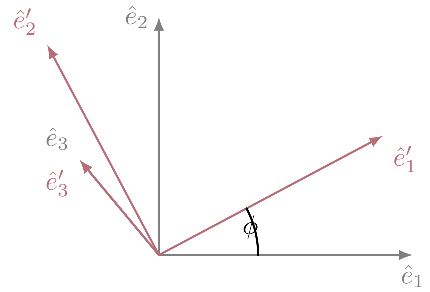


Figure 4.4: Rotation by ϕ about the z-axis: $(\hat{e}_1, \hat{e}_2) \mapsto (\hat{e}'_1, \hat{e}'_2)$, $\hat{e}'_3 = \hat{e}_3$.

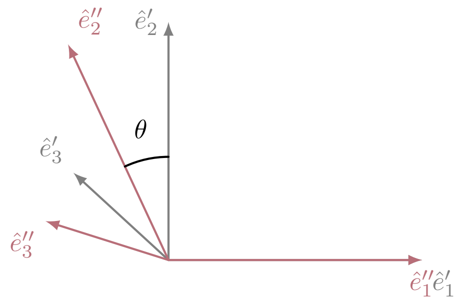


Figure 4.5: Rotation by θ about $\hat{e}'_1 = \hat{e}''_1$: the $(\hat{e}'_2, \hat{e}'_3)$ plane is rotated into $(\hat{e}''_2, \hat{e}''_3)$.

3) Rotation about the $\hat{e}''_3$ axis by ψ. The final counterclockwise rotation takes $(\hat{e}''_1, \hat{e}''_2, \hat{e}''_3)$ to $(\hat{e}'''_1, \hat{e}'''_2, \hat{e}'''_3)$ via

$$\hat{e}'''_3 = \hat{e}''_3, \qquad \hat{e}'''_1 = \cos\psi\, \hat{e}''_1 - \sin\psi\, \hat{e}''_2, \qquad \hat{e}'''_2 = \sin\psi\, \hat{e}''_1 + \cos\psi\, \hat{e}''_2,$$

which corresponds to the (passive) rotation matrix

$$A_\psi = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.33}$$

With $\vec{X}''_2 = A_\theta \vec{X}'_1$ from (4.32), we have

$$\vec{X}' = A_\psi\, \vec{X}''_2. \tag{4.34}$$

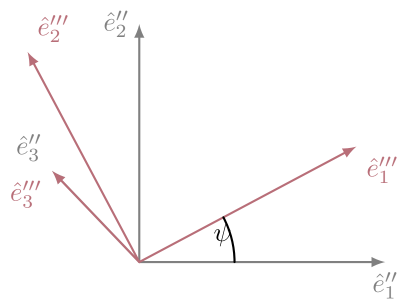


Figure 4.6: Final rotation by ψ about $\hat{e}''_3$.

Combining the three steps (4.31), (4.32), (4.34) gives the total mapping

$$\vec{X}' \;=\; A_\psi\, A_\theta\, A_\phi\, \vec{X},$$

hence

$$\vec{X}' = A\,\vec{X}, \qquad A = A_\psi A_\theta A_\phi, \tag{4.35}$$

with A_ϕ and A_θ as in (??), (??). Multiplying out yields

$$A = \begin{pmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & \sin\psi\sin\theta \\ -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & -\sin\psi\sin\phi + \cos\theta\cos\phi\cos\psi & \cos\psi\sin\theta \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{pmatrix}, \tag{4.36}$$

which is orthogonal by construction (cf. (4.14)) and satisfies $\det A = \det A_\psi \det A_\theta \det A_\phi = 1$.

Angular–velocity directions and the line of nodes. With the z–x'–z'' convention now fixed, the instantaneous rates are:

$$\dot\phi \text{ along } \hat e_3, \qquad \dot\theta \text{ along the line of nodes } \hat e_1', \qquad \dot\psi \text{ along the body axis } \hat e_3''.$$

Equivalently, the angular velocity can be written as the sum of these successive spins,

$$\boldsymbol{\omega} \;=\; \dot\phi\,\hat e_3 \;+\; \dot\theta\,\hat e_1' \;+\; \dot\psi\,\hat e_3'', \tag{4.37}$$

a form we will use immediately.

4.8. Euler's Theorem on the Motion of a Rigid Body

At any time t, the body's orientation is an orthogonal matrix $A(t)$ (e.g. $\in SO(3)$); if space and body frames coincide at $t = 0$, then $A(0) = I$ [cf. (??)].

Euler's Theorem. The motion about a fixed point is equivalent to a pure rotation about some axis through that point. In matrix terms, A admits a nonzero vector R such that

$$AR = R, \tag{4.38}$$

i.e. $\lambda = +1$ is an eigenvalue of A.
More generally, eigenvectors satisfy $AR = \lambda R$ with the characteristic equation

$$\det(A - \lambda I) = 0, \tag{4.39}$$

explicitly

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix} = 0. \tag{4.40}$$

Collecting eigenvectors as columns of X and eigenvalues in D, we have $AX = XD$ and $X^{-1}AX = D$ (see (??)).
To see that $\lambda = +1$ must occur for $A \in SO(3)$, use $A^T A = I$ and $\det A = +1$:

$$I - A^T = (A - I)A^T \implies \det\big(I - A^T\big) = \det(A - I)\det\big(A^T\big).$$

But $\det\big(A^T\big) = \det A = 1$ and $\det\big(I - A^T\big) = \det\big((I - A)^T\big) = \det(I - A)$. Since $I - A = -(A - I)$ and the matrix is 3×3, $\det(I - A) = (-1)^3 \det(A - I) = -\det(A - I)$. Equality then forces

$$\det(A - I) = 0, \tag{4.41}$$

which is precisely $\lambda = 1$ in (4.39). Let the eigenvalues be $\{e^{i\Phi}, e^{-i\Phi}, 1\}$ (the complex pair if $\Phi \neq 0, \pi$), so that

$$\mathrm{Tr}(A) = 1 + 2\cos\Phi. \tag{4.42}$$

The eigenvector for $\lambda = 1$ is the instantaneous rotation axis.

Axis–angle pair and checks

Given A, the rotation angle Φ follows from (4.42), and the unit axis $\hat n$ can be extracted (for $\Phi \neq 0, \pi$) via the antisymmetric part:

$$A - A^T \;=\; 2\sin\Phi \begin{pmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{pmatrix}.$$

This provides a direct check that the Euler–angle product (4.36) is indeed a proper rotation ($A \in SO(3)$).

4.9. Angular Velocity and the Transport of Vectors

We now relate time derivatives in the space frame and in the body frame—this is the key kinematic identity used in rigid-body dynamics.

Space vs. body components of $\boldsymbol{\omega}$

From (4.37), $\boldsymbol{\omega}$ is naturally expressed as a sum of spins about $\hat{e}_3$, the line of nodes $\hat{e}_1'$, and the body axis $\hat{e}_3''$. If we want its *space* components in the fixed basis $(\hat{e}_1, \hat{e}_2, \hat{e}_3)$, we can rewrite $\hat{e}_1'$ and $\hat{e}_3''$ using the intermediate rotations (via (**??**), (**??**)):

$$\hat{e}_1' = A_\phi\,\hat{e}_1, \qquad \hat{e}_3'' = A_\theta A_\phi\,\hat{e}_3,$$

so

$$\boldsymbol{\omega} = \dot{\phi}\,\hat{e}_3 + \dot{\theta}\,(A_\phi\,\hat{e}_1) + \dot{\psi}\,(A_\theta A_\phi\,\hat{e}_3). \tag{4.43}$$

Similarly, if body components are desired, expand $\hat{e}_3$ and $\hat{e}_1'$ in the body basis using $A^{-1} = A^T$ (cf. (4.14)).

Time derivative in moving vs. space frames (transport theorem)

Let $\vec{v}$ be any vector attached to the body. Its space components are $\vec{v} = (v_1, v_2, v_3)$ in $(\hat{e}_1, \hat{e}_2, \hat{e}_3)$; its body components are $\vec{v}' = (v_1', v_2', v_3')$ in $(\hat{e}_1', \hat{e}_2', \hat{e}_3')$, related by $\vec{v} = A^T\vec{v}'$ (passive view). Differentiating,

$$\left(\frac{d\vec{v}}{dt}\right)_{\text{space}} = \frac{d}{dt}(A^T\vec{v}') = \underbrace{(\dot{A}^T A)}_{\Omega}\,\vec{v} \;+\; A^T\left(\frac{d\vec{v}'}{dt}\right)_{\text{body}}.$$

The 3×3 matrix $\Omega = \dot{A}^T A$ is *skew-symmetric* (differentiate $A^T A = I$), so there exists a unique vector $\boldsymbol{\omega}$ with

$$\Omega\,\vec{x} = \boldsymbol{\omega}\times\vec{x} \quad \text{for all } \vec{x}.$$

Therefore, the rate of change seen in space is

$$\boxed{\left(\frac{d\vec{v}}{dt}\right)_{\text{space}} = \boldsymbol{\omega}\times\vec{v} \;+\; \left(\frac{d\vec{v}}{dt}\right)^{*}_{\text{body}}}, \tag{4.44}$$

where the starred term means "express the body-frame time derivative in the space frame" (multiplying by A^T). In words: *space-rate* = *spin* + *body-rate*. This is the transport theorem we will use in the dynamics chapter.

Small-angle generator (optional check)

For an infinitesimal δt, $A(t+\delta t) \approx (I + [\boldsymbol{\omega}]_\times \delta t)\,A(t)$, where

$$[\boldsymbol{\omega}]_\times = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix},$$

consistent with $\Omega = \dot{A}^T A$ and (4.46).

Theorem 4.2 – Chasles' Theorem

The most general motion of a rigid body can be considered as a combination of a translation of its center of mass plus a rotation about a suitable fixed point of the solid (often taken to be the center of mass).

Example 4.3 – Remarks

Rotations can be considered as finite (arbitrary angle) or infinitesimal (small angle). However, finite rotations are not commutative and hence cannot be represented by vectorial quantities. On the other hand, infinitesimal rotations are commutative and hence can be represented by vectorial quantities. In fact, we have seen that finite rotations are represented by orthogonal matrices and will not commute.

Example 4.4 – Infinitesimal Rotations

Consider an infinitesimal transformation in its most general form:

$$X' = (1+\epsilon)X \tag{4.45}$$

or in component form:

$$X'_i = \sum_j (\delta_{ij} + \epsilon_{ij}) X_j$$

The matrix elements ϵ_{ij} of the infinitesimal matrix ϵ are small so that we can always neglect ϵ^2 terms. If we apply two successive infinitesimal transformations, we end up with:

$$(1+\epsilon_1)(1+\epsilon_2) = 1 + \epsilon_1 + \epsilon_2 + \epsilon_1\epsilon_2$$

Since $\epsilon_1\epsilon_2$ can be neglected, we have:

$$(1+\epsilon_1)(1+\epsilon_2) = 1 + \epsilon_1 + \epsilon_2 \tag{4.46}$$

Hence, the commutativity of infinitesimal transformations.

Consider an infinitesimal rotation about the z-axis by $d\theta_z$, then:

$$A_z(d\theta_3) = \begin{pmatrix} 1 & d\theta_3 & 0 \\ -d\theta_3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 1 + d\theta_3 J_3 = 1 + \epsilon$$

$$\epsilon = d\theta_3 J_3, \quad J_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.47}$$

Where we used $\sin(d\theta) \approx d\theta$ and $\cos(d\theta) \approx 1 + \mathcal{O}(d\theta^2)$.
Similarly, for infinitesimal rotations about the x- and y-axes separately, we will have:

$$A_x(d\theta_1) \approx 1 + d\theta_1 J_1, \quad A_y(d\theta_2) \approx 1 + d\theta_2 J_2$$

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \tag{4.48}$$

J_1, J_2, J_3 are called the generators of infinitesimal rotations. We can show that:

$$[J_i, J_j] = \sum_k \epsilon_{ijk} J_k \tag{4.49}$$

We can generalize the above results for an infinitesimal rotation about an arbitrary axis $\hat{e}_\theta$:

$$A(d\theta) = 1 + d\theta \cdot \vec{J}, \quad d\theta = d\theta \hat{e}_\theta \tag{4.50}$$

For a finite rotation about the same axis $\hat{e}_0$, we can use the fact that infinitesimal rotations commute, so that:

$$A(\theta) = \lim_{n\to\infty} \left(1 + \frac{\theta \cdot \vec{J}}{n}\right)^n = e^{\theta \cdot \vec{J}} \quad ; \quad \vec{\theta} = \theta \hat{e}_\theta \tag{4.51}$$

Example 4.5 – Inverse and Structure of an Infinitesimal Rotation

Since $A = 1 + \epsilon$, then:

$$A^{-1} = (1+\epsilon)^{-1} = 1 - \epsilon \tag{4.52}$$

So that:

$$AA^{-1} = (1+\epsilon)(1-\epsilon) = 1 + \epsilon - \epsilon = 1$$

Furthermore, since A is an orthogonal matrix:

$$A^T = 1 + \epsilon^T = A^{-1} = 1 - \epsilon$$

$$\therefore \epsilon^T = -\epsilon \tag{4.53}$$

$$\implies (\epsilon^T)_{ii} = -\epsilon_{ii} = \epsilon_{ii} \implies \epsilon_{ii} = 0 \tag{4.54}$$

i.e., the diagonal elements of an antisymmetric matrix are necessarily zero. Using this information, we can write the general structure of the infinitesimal matrix ϵ,

$$\epsilon = \begin{pmatrix} 0 & d\Omega_3 & -d\Omega_2 \\ -d\Omega_3 & 0 & d\Omega_1 \\ d\Omega_2 & -d\Omega_1 & 0 \end{pmatrix} \tag{4.55}$$

which has only three independent infinitesimal parameters: $d\Omega_1, d\Omega_2, d\Omega_3$. Under the infinitesimal transformation, the vector $\vec{r}$ transforms as:

$$\vec{r}' = (1+\epsilon)\vec{r}$$

Thus, the change in $\vec{r}$ is:

$$d\vec{r} = \vec{r}' - \vec{r} = \epsilon\vec{r}$$

$$d\vec{r} = \begin{pmatrix} dx_1 \\ dx_2 \\ dx_3 \end{pmatrix} = \begin{pmatrix} 0 & d\Omega_3 & -d\Omega_2 \\ -d\Omega_3 & 0 & d\Omega_1 \\ d\Omega_2 & -d\Omega_1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \tag{4.56}$$

Thus, the differential change in the components of $\vec{r}$ are:

$$\begin{aligned} dx_1 &= x_2 d\Omega_3 - x_3 d\Omega_2 \\ dx_2 &= x_3 d\Omega_1 - x_1 d\Omega_3 \\ dx_3 &= x_1 d\Omega_2 - x_2 d\Omega_1 \end{aligned} \tag{4.57}$$

This can be written in vectorial form if we define the vector:

$$d\vec{\Omega} = \begin{pmatrix} d\Omega_1 \\ d\Omega_2 \\ d\Omega_3 \end{pmatrix} \tag{4.58}$$

Then, the differential change in $\vec{r}$ is:

$$d\vec{r} = \vec{r} \times d\vec{\Omega} \tag{4.59}$$

4.10. Rate of Change of a Vector

The infinitesimal-rotation result

$$d\vec{r} \;=\; \vec{r} \times d\vec{\Omega}, \qquad d\vec{\Omega} = \hat{n}\, d\phi$$

was established in Eq. (4.59). We now extend it to the time rate of change of an *arbitrary* vector $\vec{G}(t)$ when observed from a space-fixed (inertial) frame versus a body-fixed (rotating) frame.

Theorem 4.3 – Transport Theorem (Space vs. Body Derivative)

For any vector $\vec{G}(t)$,

$$d\vec{G}\big|_{\text{space}} \;=\; d\vec{G}\big|_{\text{body}} \;+\; d\vec{\Omega}\times\vec{G}, \tag{4.60}$$

and, dividing by dt with $\vec{\omega} = \dfrac{d\vec{\Omega}}{dt}$,

$$\left.\frac{d\vec{G}}{dt}\right|_{\text{space}} \;=\; \left.\frac{d\vec{G}}{dt}\right|_{\text{body}} \;+\; \vec{\omega}\times\vec{G}. \tag{4.61}$$

Equivalently, as an operator identity,

$$\left.\frac{d}{dt}\right|_{\text{space}} \;=\; \left.\frac{d}{dt}\right|_{\text{body}} \;+\; \vec{\omega}\times\;. \tag{4.62}$$

Component/basis proof (keeps the style): Let $\{\hat{e}_i\}$ be a fixed orthonormal basis and $\{\hat{e}'_i(t)\}$ the body-fixed basis (same origin). For any $\vec{A}$,

$$\vec{A} \;=\; \sum_{i=1}^{3} A_i\,\hat{e}_i \;=\; \sum_{i=1}^{3} A'_i\,\hat{e}'_i,$$

with A_i and A'_i the fixed- and body-frame components. Differentiating in the fixed frame and using $\dot{\hat{e}}'_i\big|_F = \vec{\omega}\times\hat{e}'_i$,

$$\left.\frac{d\vec{A}}{dt}\right|_F = \sum_i \dot{A}'_i\,\hat{e}'_i + \sum_i A'_i\,\dot{\hat{e}}'_i\big|_F = \left.\frac{d\vec{A}}{dt}\right|_R + \vec{\omega}\times\vec{A},$$

which is Eq. (4.61). (Here the subscript R emphasizes "rotating".)

Skew form (useful later): Introduce the skew-symmetric "cross" matrix

$$[\vec{\omega}]_\times \;=\; \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}, \qquad [\vec{\omega}]_\times\vec{v} = \vec{\omega}\times\vec{v}.$$

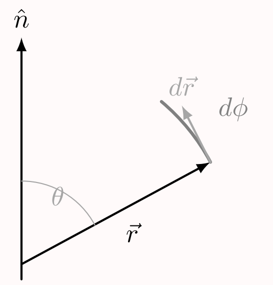
Then (4.61) reads

$$\left(\frac{d}{dt}\right)_{\text{space}} \;=\; \left(\frac{d}{dt}\right)_{\text{body}} + [\vec{\omega}]_\times.$$

If $A(t)$ is the rotation taking body coordinates to space coordinates, the body basis columns form A and one has $\dot{A} = [\vec{\omega}]_\times A$.

Example 4.6 – Graphical Interpretation of $d\vec{r}$

Let $\hat{n}$ be the instantaneous axis and θ the angle between $\hat{n}$ and $\vec{r}$. A rotation by $d\phi$ sweeps the tip of $\vec{r}$ along an arc of length $|d\vec{r}| = r\sin\theta\,d\phi$, directed perpendicular to the plane spanned by $\hat{n}$ and $\vec{r}$. Since $|\hat{n}\times\vec{r}| = r\sin\theta$, the vector form follows immediately.



4.11. The Coriolis Effect

We now apply Eqs. (4.61)–(4.62) to kinematics and dynamics in a frame rotating with angular velocity $\vec{\omega}(t)$ about the *same origin* as the inertial frame.[2]

Velocity and Acceleration Transformations

With $\vec{r}(t)$ the position of a particle,

$$\left.\frac{d\vec{r}}{dt}\right|_F = \left.\frac{d\vec{r}}{dt}\right|_R + \vec{\omega}\times\vec{r}, \tag{4.63}$$

i.e.

$$\vec{V}_F = \vec{V}_R + \vec{\omega}\times\vec{r}. \tag{4.64}$$

Differentiate (4.64) in the inertial frame and use (4.61) carefully:

$$\begin{aligned}\left.\frac{d\vec{V}_F}{dt}\right|_F &= \left.\frac{d\vec{V}_R}{dt}\right|_F + \left.\frac{d}{dt}\right|_F(\vec{\omega}\times\vec{r})\\ &= \left(\left.\frac{d\vec{V}_R}{dt}\right|_R + \vec{\omega}\times\vec{V}_R\right) + \left(\dot{\vec{\omega}}\times\vec{r} + \vec{\omega}\times\left.\frac{d\vec{r}}{dt}\right|_F\right)\\ &= \left.\frac{d\vec{V}_R}{dt}\right|_R + 2\,\vec{\omega}\times\vec{V}_R + \dot{\vec{\omega}}\times\vec{r} + \vec{\omega}\times(\vec{\omega}\times\vec{r}). \end{aligned} \tag{4.65}$$

Thus the inertial acceleration is

$$\vec{a}_F = \vec{a}_R + 2\,\vec{\omega}\times\vec{V}_R + \dot{\vec{\omega}}\times\vec{r} + \vec{\omega}\times(\vec{\omega}\times\vec{r}).$$

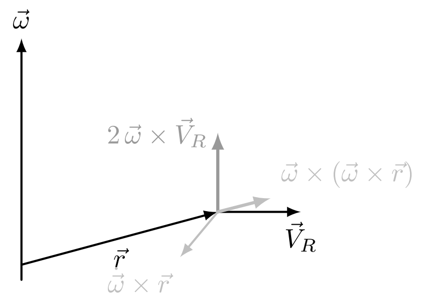
Newton's Second Law in a Rotating Frame

Newton's law holds in the inertial frame:

$$\vec{F}_{\text{net}} = m\,\vec{a}_F. \tag{4.66}$$

Combining (4.65) with (4.66) gives the equation of motion *as seen in the rotating frame*:

$$m\,\vec{a}_R = \vec{F}_{\text{net}} - \underbrace{2m\,\vec{\omega}\times\vec{V}_R}_{\text{Coriolis}} - \underbrace{m\,\dot{\vec{\omega}}\times\vec{r}}_{\text{Euler}} - \underbrace{m\,\vec{\omega}\times(\vec{\omega}\times\vec{r})}_{\text{centrifugal}}. \tag{4.67}$$



Orders of Magnitude on Earth

For Earth, $|\vec{\omega}| \simeq 7.292\times10^{-5}\,\mathrm{s}^{-1}$; on human timescales $\dot{\vec{\omega}} \approx \vec{0}$ so the Euler term is negligible. The centrifugal term has magnitude $m\omega^2 r_\perp$ with $r_\perp = r\sin\theta$ and amounts to about $3\times10^{-3}\,mg$ at the equator (it slightly reduces apparent weight). The Coriolis term $-2m\,\vec{\omega}\times\vec{V}_R$ produces the familiar sideways deflection of moving bodies in the rotating (Earth-fixed) frame.

[2] If the origins differ by $\vec{R}(t)$, add the standard translational terms $\dot{\vec{R}}, \ddot{\vec{R}}$.

Example 4.7 – Quick Checks (consistent with our earlier results)

- A vector $\vec{G}$ rigidly attached to the body: $\frac{d\vec{G}}{dt}\big|_{\text{body}} = 0$. Then from (4.61) its change in space is purely rotational: $\frac{d\vec{G}}{dt}\big|_{\text{space}} = \vec{\omega} \times \vec{G}$.
- Purely tangential motion when $\dot{r} = 0$: with $\dot{\theta} = \ell/(\mu r^2)$ (see Eq. (3.22)), the kinetic energy splits exactly as in (3.21), and the "centrifugal" term in V_{eff} (3.26) matches $\frac{1}{2}\mu r^2\dot{\theta}^2$.

Summary of Chapter 4

1. **Degrees of freedom of a rigid body:**

$$6 \;=\; 3\,(\text{C.M. position}) \;+\; 3\,(\text{orientation via Euler angles } \phi, \theta, \psi).$$

General rigid motion can be viewed as translation of the C.M. plus a rotation about a suitable fixed point (*Chasles' theorem*; often written "Charles' theorem" in some texts).

2. **Describing orientation:**

   - By *direction cosines* $a_{ij} = \cos\theta_{ij}$ (direction-cosine matrix) relating body and space bases, see Eqs. (4.7)–(4.12).
   - Equivalently by the three *Euler angles* (ϕ, θ, ψ).

3. **Rotations are orthogonal transformations:**

$$A^T A \;=\; AA^T \;=\; I \quad (\text{orthogonality, Eq. (4.14)}),$$

and for proper rotations $|A| = +1$ (so $A \in \mathrm{SO}(3)$).

4. **Know the elementary rotation matrices and their composition:**

$$A_\phi = R_z(\phi)\;\; (\text{Eq. (??)}), \qquad A_\theta = R_{x'}(\theta)\;\; (\text{Eq. (??)}), \qquad A_\psi = R_{z''}(\psi)\;\; (\text{Eq. (4.33)}),$$

with the overall body-to-space transformation

$$A \;=\; A_\psi\, A_\theta\, A_\phi \quad (\text{Eqs. (4.35)–(4.36)}).$$

5. **Euler's theorem on rigid body rotation:**

$$\exists\, \hat{n} \text{ such that } A\,\hat{n} \;=\; \hat{n} \quad (\text{eigenvalue } \lambda_1 = 1).$$

The three eigenvalues are $\{1, e^{i\Phi}, e^{-i\Phi}\}$ with

$$\mathrm{Tr}(A) = 1 + 2\cos\Phi \quad (\text{Eq. (4.42)}), \qquad |A| = 1.$$

6. **Infinitesimal rotations and generators:**

$$A \;=\; I + \epsilon P + \mathcal{O}(\epsilon^2), \qquad P^T = -P,$$

so P is antisymmetric with three independent parameters,

$$P = \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}, \qquad P_{ji} = -P_{ij}, \;\; P_{ii} = 0.$$

In axis–angle form $A(\vec{\theta}) = \exp\Big([\vec{\theta}]_\times\Big)$ (Eq. (4.51)); the Cartesian generators J_i satisfy the usual commutators (Eq. (4.49)) and the first-order composition rule (Eq. (4.46)).

7. **Transport theorem (rate of change of a vector):**

$$\left.\frac{d\vec{G}}{dt}\right|_{\text{space}} \;=\; \left.\frac{d\vec{G}}{dt}\right|_{\text{body}} \;+\; \vec{\omega} \times \vec{G} \quad (\text{Eqs. (4.60)–(4.62)}).$$

8. **Dynamics in a rotating frame (same origin):**

$$m\,\vec{a}_R \;=\; \vec{F}_{\text{net}} \;-\; 2m\,\vec{\omega} \times \vec{V}_R \;-\; m\,\dot{\vec{\omega}} \times \vec{r} \;-\; m\,\vec{\omega} \times (\vec{\omega} \times \vec{r}) \quad (\text{Coriolis, Euler, centrifugal; Eq. (4.67)}).$$

Chapter 5

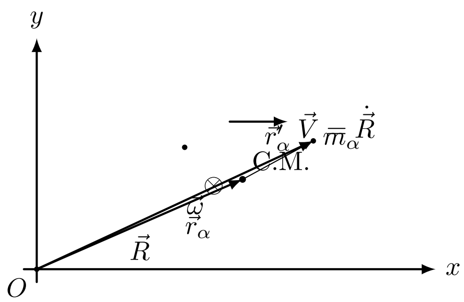
Rigid Body Equations of Motion

By *Chasles' (Charles') theorem* (see Chapter 4), any general motion of a rigid body can be represented as the combination of (i) a translation of a chosen fixed point of the body (we will take it to be the center of mass), plus (ii) a pure rotation about that fixed point. As we saw in Chapter 1, both the kinetic energy and the angular momentum take especially transparent forms once they are expressed as a translational piece of the C.M. and a piece relative to the C.M.
Consider a rigid body made of N point masses m_α ($\alpha = 1, \dots, N$). Let

$\vec{r}_\alpha$ be the position of m_α in a fixed inertial frame, $\vec{R}$ the C.M. position,

and let $\vec{r}'_\alpha$ denote the position of m_α relative to the axes through the C.M. Then

$$\vec{r}_\alpha = \vec{R} + \vec{r}'_\alpha. \tag{5.1}$$



Geometry of $\vec{r}_\alpha = \vec{R} + \vec{r}'_\alpha$.

If the body rotates with instantaneous angular velocity $\vec{\omega}$ about an axis through the C.M., then by the fixed–vs–rotating derivative relation (cf. (4.61)–(4.62)) the velocity of particle m_α in the fixed frame is

$$\vec{v}_\alpha \equiv \dot{\vec{r}}_\alpha\big|_{\text{fixed}} = \dot{\vec{R}}\big|_{\text{fixed}} + \dot{\vec{r}}'_\alpha\big|_{\text{fixed}} = \vec{V} + \vec{\omega} \times \vec{r}'_\alpha. \tag{5.2}$$

5.1. Kinetic Energy Split

The kinetic energy is

$$\begin{aligned} T &= \frac{1}{2}\sum_\alpha m_\alpha\, \vec{v}_\alpha^2 = \frac{1}{2}\sum_\alpha m_\alpha \left(\vec{V} + \vec{\omega} \times \vec{r}'_\alpha\right)\cdot\left(\vec{V} + \vec{\omega} \times \vec{r}'_\alpha\right) \\ &= \frac{1}{2}\sum_\alpha m_\alpha\, \vec{V}^{\,2} + \sum_\alpha m_\alpha\, \vec{V}\cdot(\vec{\omega} \times \vec{r}'_\alpha) + \frac{1}{2}\sum_\alpha m_\alpha\, (\vec{\omega} \times \vec{r}'_\alpha)^2. \end{aligned} \tag{5.3}$$

Using $\sum_\alpha m_\alpha \vec{r}'_\alpha = \vec{0}$ (definition of C.M.), the cross term vanishes:

$$\sum_\alpha m_\alpha\, \vec{V}\cdot(\vec{\omega} \times \vec{r}'_\alpha) = \vec{V}\cdot\Big(\vec{\omega} \times \sum_\alpha m_\alpha \vec{r}'_\alpha\Big) = 0.$$

Hence

$$T = \frac{1}{2} M V^2 + \frac{1}{2} \sum_\alpha m_\alpha (\vec{\omega} \times \vec{r}'_\alpha)^2, \qquad M = \sum_\alpha m_\alpha. \tag{5.4}$$

We will often refer to

$$T_{\text{C.M.}} = \frac{1}{2} M V^2, \qquad T_{\text{rot}} = \frac{1}{2} \sum_\alpha m_\alpha (\vec{\omega} \times \vec{r}'_\alpha)^2,$$

so that $T = T_{\text{C.M.}} + T_{\text{rot}}$. (Only external forces contribute to the potential energy: internal forces do no work on—nor change the internal potential of—the rigid body as a whole.)

5.2. Total Angular Momentum About a Fixed Origin

The total angular momentum about the fixed origin O is

$$\vec{L}_O = \sum_\alpha \vec{r}_\alpha \times \vec{p}_\alpha = \sum_\alpha m_\alpha \vec{r}_\alpha \times \vec{v}_\alpha. \tag{5.5}$$

Insert (5.12) and (5.13):

$$\begin{aligned} \vec{L}_O &= \sum_\alpha m_\alpha (\vec{R} + \vec{r}'_\alpha) \times (\vec{V} + \vec{\omega} \times \vec{r}'_\alpha) \\ &= \underbrace{\sum_\alpha m_\alpha \vec{R} \times \vec{V}}_{= M \vec{R} \times \vec{V}} + \underbrace{\sum_\alpha m_\alpha \vec{R} \times (\vec{\omega} \times \vec{r}'_\alpha)}_{= \vec{R} \times [\vec{\omega} \times \sum m_\alpha \vec{r}'_\alpha] = \vec{0}} + \underbrace{\sum_\alpha m_\alpha \vec{r}'_\alpha \times \vec{V}}_{= (\sum m_\alpha \vec{r}'_\alpha) \times \vec{V} = \vec{0}} \\ &\quad + \sum_\alpha m_\alpha \vec{r}'_\alpha \times (\vec{\omega} \times \vec{r}'_\alpha). \end{aligned} \tag{5.6}$$

Therefore

$$\boxed{\vec{L}_O = M \vec{R} \times \vec{V} \; + \; \sum_\alpha m_\alpha \vec{r}'_\alpha \times (\vec{\omega} \times \vec{r}'_\alpha)} \quad \equiv \quad \vec{L}_{\text{C.M.}} + \vec{L}_{\text{rot}}. \tag{5.7}$$

5.3. Angular Momentum and Kinetic Energy in the C.M. Frame

For the rest of this section we work in the C.M.-fixed frame, so $\vec{R} = \vec{0}$ and $\vec{V} = \vec{0}$. We therefore drop the primes and the "rot" subscripts, with the understanding that all position vectors are measured from the C.M. and all quantities below are *purely rotational*. Then

$$\vec{L} = \sum_\alpha m_\alpha \vec{r}_\alpha \times (\vec{\omega} \times \vec{r}_\alpha) = \sum_\alpha m_\alpha \left[\, (\vec{r}_\alpha \cdot \vec{r}_\alpha) \, \vec{\omega} - (\vec{r}_\alpha \cdot \vec{\omega}) \, \vec{r}_\alpha \, \right] \tag{5.8}$$

(using $\vec{a} \times (\vec{b} \times \vec{c}) = \vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b})$).
In components (with $r_{\alpha i}$ the i-th component of $\vec{r}_\alpha$ and ω_i that of $\vec{\omega}$),

$$\begin{aligned} L_i &= \sum_\alpha m_\alpha \Big[\big(\sum_k r_{\alpha k}^2 \big) \, \omega_i - r_{\alpha i} \, \big(\sum_j \omega_j r_{\alpha j} \big) \Big] \\ &= \sum_j \underbrace{\left(\sum_\alpha m_\alpha \Big[\Big(\sum_k r_{\alpha k}^2 \Big) \delta_{ij} - r_{\alpha i} r_{\alpha j} \Big] \right)}_{I_{ij}} \omega_j. \end{aligned} \tag{5.9}$$

This identifies the **inertia tensor**

$$\boxed{I_{ij} \; = \; \sum_\alpha m_\alpha \Big(r_\alpha^2 \, \delta_{ij} - r_{\alpha i} \, r_{\alpha j} \Big)} \qquad \Rightarrow \qquad \boxed{L_i = \sum_j I_{ij} \, \omega_j} \quad \text{or} \quad \boxed{\vec{L} = \mathbf{I} \, \vec{\omega}}. \tag{5.10}$$

Consequently, the rotational kinetic energy can be written as

$$T_{\text{rot}} = \frac{1}{2} \sum_\alpha m_\alpha (\vec{\omega} \times \vec{r}_\alpha)^2 = \frac{1}{2} \vec{\omega} \cdot \vec{L} = \frac{1}{2} \sum_{i,j} \omega_i \, I_{ij} \, \omega_j. \tag{5.11}$$

Remarks.

- Equations (5.4) and (5.15) give the clean split $T = T_{\text{C.M.}} + T_{\text{rot}}$ and $\vec{L}_O = \vec{L}_{\text{C.M.}} + \vec{L}_{\text{rot}}$.
- In the C.M. frame all motion is rotational, and (5.17)–(5.18) define the standard inertia tensor representation $\vec{L} = \mathbf{I}\,\vec{\omega}$ and $T_{\text{rot}} = \frac{1}{2}\vec{\omega}\cdot\mathbf{I}\,\vec{\omega}$.

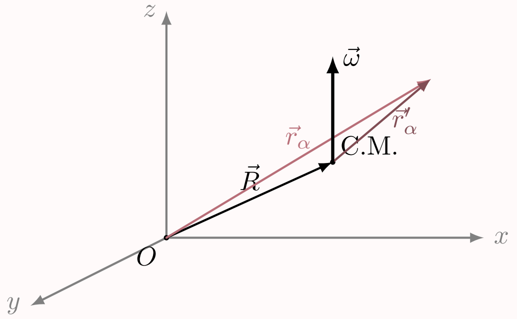
Example 5.1 – Rigid Body Equations of Motion (transcribed from notes)

Translation–rotation split about the C.M. Consider a rigid body of particles m_α with

$$\vec{r}_\alpha = \vec{R} + \vec{r}^{\,\prime}_\alpha, \qquad \sum_\alpha m_\alpha \vec{r}^{\,\prime}_\alpha = \vec{0}. \tag{5.12}$$

If the body rotates with angular velocity $\vec{\omega}$ about an axis through the C.M., then (using Chap. 4, Eqs. (4.59)–(4.62))

$$\vec{v}_\alpha = \dot{\vec{r}}_\alpha = \dot{\vec{R}} + \left.\frac{d\vec{r}^{\,\prime}_\alpha}{dt}\right|_{\text{space}} = \vec{V} + \vec{\omega}\times\vec{r}^{\,\prime}_\alpha. \tag{5.13}$$



Kinetic energy split.

$$\begin{aligned} T &= \frac{1}{2}\sum_\alpha m_\alpha \big(\vec{V} + \vec{\omega}\times\vec{r}^{\,\prime}_\alpha\big)^2 \\ &= \frac{1}{2}MV^2 + \frac{1}{2}\sum_\alpha m_\alpha |\vec{\omega}\times\vec{r}^{\,\prime}_\alpha|^2 \equiv T_{\text{C.M.}} + T_{\text{rot}}, \qquad M = \sum_\alpha m_\alpha. \end{aligned} \tag{5.14}$$

Angular momentum about an inertial origin O.

$$\vec{L}_O = \sum_\alpha \vec{r}_\alpha \times (m_\alpha \vec{v}_\alpha) = \underbrace{\vec{R}\times(M\vec{V})}_{\vec{L}_{\text{C.M.}}} + \underbrace{\sum_\alpha m_\alpha\,\vec{r}^{\,\prime}_\alpha \times (\vec{\omega}\times\vec{r}^{\,\prime}_\alpha)}_{\vec{L}_{\text{rot}}}. \tag{5.15}$$

Inertia tensor from $\vec{L}_{\text{rot}}$. Dropping primes ($\vec{r}_\alpha \equiv$ C.M.-relative),

$$\vec{L} \equiv \vec{L}_{\text{rot}} = \sum_\alpha m_\alpha\,\vec{r}_\alpha \times (\vec{\omega}\times\vec{r}_\alpha) \quad \Rightarrow \quad L_i = \sum_j I_{ij}\,\omega_j, \tag{5.16}$$

with

$$I_{ij} = \sum_\alpha m_\alpha \big(r_\alpha^2 \delta_{ij} - x_{i\alpha}x_{j\alpha}\big), \qquad \vec{L} = \mathbf{I}\,\vec{\omega}. \tag{5.17}$$

Diagonal entries (moments) and off–diagonal entries (products) are

$$\begin{aligned} &I_{xx} = \sum m_\alpha(y_\alpha^2 + z_\alpha^2), \quad I_{yy} = \sum m_\alpha(x_\alpha^2 + z_\alpha^2), \quad I_{zz} = \sum m_\alpha(x_\alpha^2 + y_\alpha^2), \\ &I_{xy} = I_{yx} = -\sum m_\alpha x_\alpha y_\alpha, \quad I_{xz} = I_{zx} = -\sum m_\alpha x_\alpha z_\alpha, \quad I_{yz} = I_{zy} = -\sum m_\alpha y_\alpha z_\alpha. \end{aligned} \tag{5.18}$$

Continuous limit:

$$I_{ij} = \int_V \rho(\vec{r})\Big(r^2\delta_{ij} - x_i x_j\Big)\, d^3r. \tag{5.19}$$

Tensor character and transformation law. Under a proper rotation A (Chap. 4),

$$I'_{ij} = \sum_{h,k} A_{ih}\, I_{hk}\, A_{jk} \quad \Longrightarrow \quad \mathbf{I}' = A\,\mathbf{I}\,A^T. \tag{5.20}$$

Rotational kinetic energy.

$$T_{\text{rot}} = \frac{1}{2}\sum_\alpha m_\alpha |\vec{\omega}\times\vec{r}_\alpha|^2 = \frac{1}{2}\sum_{i,j}\omega_i I_{ij}\omega_j = \frac{1}{2}\vec{\omega}\cdot\mathbf{I}\,\vec{\omega}. \tag{5.21}$$

Alternative: since $\vec{v}_\alpha = \vec{\omega}\times\vec{r}_\alpha$,

$$T_{\text{rot}} = \tfrac{1}{2}\sum m_\alpha \vec{v}_\alpha\cdot\vec{v}_\alpha = \tfrac{1}{2}\sum m_\alpha\vec{\omega}\cdot(\vec{r}_\alpha\times\vec{v}_\alpha) = \tfrac{1}{2}\vec{\omega}\cdot\vec{L} = \tfrac{1}{2}\vec{\omega}\cdot\mathbf{I}\,\vec{\omega}.$$

Equivalent Levi–Civita form and a useful remark.

$$I_{ij} = \sum_\alpha m_\alpha\, \varepsilon_{ik\ell}\,\varepsilon_{jm\ell}\, x_{k\alpha}x_{m\alpha}. \tag{5.22}$$

Even if the body rotates about z ($\vec{\omega} = \omega\hat{z}$), $\vec{L} = \mathbf{I}\vec{\omega}$ need *not* be parallel to $\vec{\omega}$ unless the products I_{xz}, I_{yz} vanish. Symmetry under $z \to -z$ forces $I_{xz} = I_{yz} = 0$; three orthogonal symmetry planes diagonalize $\mathbf{I}$.

Example 5.2 – Planar Mass Distribution and the Perpendicular–Axis Theorem

Let the mass be entirely in the xy-plane, so every particle has $z_\alpha = 0$. Starting from the discrete inertia tensor

$$I_{ij} = \sum_\alpha m_\alpha\big(r'^2_\alpha\delta_{ij} - x'_{\alpha i}x'_{\alpha j}\big)\,, \qquad r'^2_\alpha = x'^2_\alpha + y'^2_\alpha + z'^2_\alpha,$$

we obtain, term by term,

$$\begin{aligned}
I_{xx} &= \sum_\alpha m_\alpha(y'^2_\alpha + z'^2_\alpha) = \sum_\alpha m_\alpha y'^2_\alpha,\\
I_{yy} &= \sum_\alpha m_\alpha(x'^2_\alpha + z'^2_\alpha) = \sum_\alpha m_\alpha x'^2_\alpha,\\
I_{zz} &= \sum_\alpha m_\alpha(x'^2_\alpha + y'^2_\alpha) = I_{xx} + I_{yy},\\
I_{xz} &= I_{yz} = 0, \qquad I_{xy} = I_{yx} = -\sum_\alpha m_\alpha x'_\alpha y'_\alpha.
\end{aligned}$$

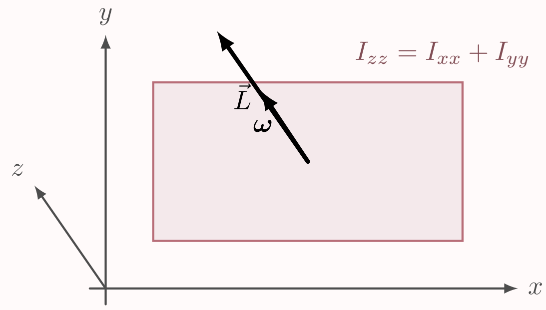
Hence the tensor is block–diagonal:

$$\boxed{\mathbf{I} = \begin{pmatrix} I_{xx} & I_{xy} & 0 \\ I_{xy} & I_{yy} & 0 \\ 0 & 0 & I_{xx}+I_{yy} \end{pmatrix}}$$

If the lamina rotates about $\hat{z}$ with $\boldsymbol{\omega} = (0,0,\omega)$, then

$$\vec{L} = \mathbf{I}\,\boldsymbol{\omega} = \big(0,\,0,\,(I_{xx}+I_{yy})\omega\big) = I_{zz}\omega\,\hat{z},$$

so $\vec{L} \parallel \boldsymbol{\omega}$ and $\hat{z}$ is a principal axis.



Example 5.3 – Uniform Solid Cube about a Corner: I

A homogeneous cube of side a and density ρ occupies $0 \le x, y, z \le a$, with origin O at the corner $(0,0,0)$. Its mass is $M = \rho a^3$.
Diagonal moments about O.

$$I_{xx} = \int_V \rho\,(y^2+z^2)\,dV = \rho\left[\int_0^a\int_0^a\int_0^a y^2\,dz\,dy\,dx + \int_0^a\int_0^a\int_0^a z^2\,dz\,dy\,dx\right]$$
$$= \rho\left[a\cdot\frac{a^3}{3}\cdot a + a\cdot a\cdot\frac{a^3}{3}\right] = \frac{2}{3}\rho a^5 = \boxed{\frac{2}{3}Ma^2}.$$

By symmetry, $I_{yy} = I_{zz} = \dfrac{2}{3}Ma^2$.
Products of inertia about O.

$$I_{xy} = -\int_V \rho\,xy\,dV = -\rho\left[\int_0^a x\,dx\right]\left[\int_0^a y\,dy\right]\left[\int_0^a dz\right] = -\rho\left(\tfrac{a^2}{2}\right)\left(\tfrac{a^2}{2}\right)a = \boxed{-\tfrac{1}{4}Ma^2},$$
$$I_{xz} = I_{yz} = \boxed{-\tfrac{1}{4}Ma^2}.$$

Inertia tensor about the corner O.

$$\boxed{\mathbf{I}_O = \frac{Ma^2}{12}\begin{pmatrix} 8 & -3 & -3 \\ -3 & 8 & -3 \\ -3 & -3 & 8 \end{pmatrix}}$$

Rotation about the z-edge: $\boldsymbol{\omega} = (0,0,\omega)$.

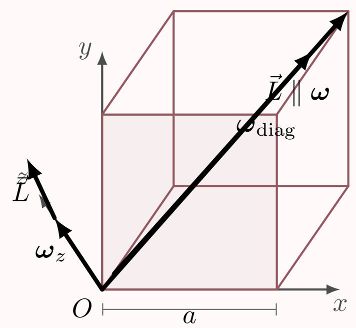
$$\vec{L} = \mathbf{I}_O\,\boldsymbol{\omega} = \frac{Ma^2\omega}{12}\begin{pmatrix} -3 \\ -3 \\ 8 \end{pmatrix},$$

so $\vec{L}$ is *not* parallel to $\boldsymbol{\omega}$; the edge Oz is *not* a principal axis (about the corner).
Rotation about the body diagonal through O. Take the unit vector $\hat{u} = \frac{1}{\sqrt{3}}(1,1,1)$ and $\boldsymbol{\omega} = \omega\,\hat{u}$. Since $\mathbf{I}_O(1,1,1)^T = \frac{Ma^2}{12}(2,2,2)^T$,

$$\vec{L} = \mathbf{I}_O\,\boldsymbol{\omega} = \left(\frac{Ma^2}{6}\right)\omega\,\hat{u} \quad\Rightarrow\quad \boxed{\vec{L}\parallel\boldsymbol{\omega},\quad I_{\text{diag@corner}} = \frac{Ma^2}{6}}.$$

(The other two principal moments—eigenvalues for any unit vector orthogonal to $(1,1,1)$—are both $I_\perp = \dfrac{11}{12}Ma^2$.)
Rotational energy (C.M. or any frame). With $\vec{L} = \mathbf{I}\boldsymbol{\omega}$ and $(\vec{A}\times\vec{B})^2 = A^2B^2 - (\vec{A}\cdot\vec{B})^2$,

$$T_{\text{rot}} = \frac{1}{2}\sum_\alpha m_\alpha\big(\boldsymbol{\omega}\times\vec{r}'_\alpha\big)^2 = \frac{1}{2}\,\boldsymbol{\omega}\cdot\mathbf{I}\,\boldsymbol{\omega} = \boxed{\tfrac{1}{2}\,\boldsymbol{\omega}\cdot\vec{L}}.$$



Example 5.4 – Planar Mass Distribution in the xy–Plane

Consider a rigid body whose entire mass lies in the plane $z = 0$ (so $z_\alpha = 0$ for all particles). The products of inertia that involve z vanish:

$$I_{13} = I_{31} = I_{23} = I_{32} = 0.$$

Moreover,

$$I_{33} = \sum_\alpha m_\alpha (x_\alpha^2 + y_\alpha^2), \qquad I_{11} = \sum_\alpha m_\alpha (y_\alpha^2), \qquad I_{22} = \sum_\alpha m_\alpha (x_\alpha^2).$$

Hence the *correct* identity for a planar distribution is

$$I_{11} + I_{22} = \sum_\alpha m_\alpha (x_\alpha^2 + y_\alpha^2) = I_{33}.$$

Thus, in axes $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ with $\hat{e}_3 \parallel \hat{z}$,

$$\mathbf{I} = \begin{pmatrix} I_{11} & I_{12} & 0 \\ I_{21} & I_{22} & 0 \\ 0 & 0 & I_{33} \end{pmatrix}, \qquad I_{21} = I_{12}.$$

If the body rotates about the z–axis with $\boldsymbol{\omega} = (0, 0, \omega)$, then

$$\mathbf{L} = \mathbf{I}\,\boldsymbol{\omega} = (0, 0, I_{33}\,\omega),$$

so $\mathbf{L} \parallel \boldsymbol{\omega}$ in this case.

Example 5.5 – Uniform Solid Cube — Corner Axes vs. Center–of–Mass Axes

Let a uniform solid cube of side a and total mass M have density $\rho = M/a^3$.
(a) Axes through a *corner* (edges along x, y, z; integrate $x, y, z \in [0, a]$).
Moments about the coordinate axes (through the corner) are

$$I_{xx} = \rho \int_0^a \int_0^a \int_0^a (y^2 + z^2)\, dy\, dz\, dx = \frac{2}{3} M a^2,$$

and cyclically $I_{yy} = I_{zz} = \frac{2}{3} M a^2$. Products of inertia (about the corner) are

$$I_{xy} = -\rho \int_0^a \int_0^a \int_0^a xy\, dx\, dy\, dz = -\frac{1}{4} M a^2,$$

and cyclically $I_{xz} = I_{yz} = -\frac{1}{4}Ma^2$. Therefore

$$\boxed{\mathbf{I}_{\text{corner}} = \frac{Ma^2}{12}\begin{pmatrix} 8 & -3 & -3 \\ -3 & 8 & -3 \\ -3 & -3 & 8 \end{pmatrix}.}$$

Two instructive checks:

- If $\boldsymbol{\omega} = (0, 0, \omega)$, then $\mathbf{L} = \mathbf{I}_{\text{corner}}\,\boldsymbol{\omega} = \frac{Ma^2\omega}{12}(-3, -3, 8)$, so $\mathbf{L} \nparallel \boldsymbol{\omega}$.
- If $\boldsymbol{\omega} = \frac{\omega}{\sqrt{3}}(1, 1, 1)$ (rotation about the body diagonal), then

$$\mathbf{L} = \frac{Ma^2}{12}\begin{pmatrix} 8 & -3 & -3 \\ -3 & 8 & -3 \\ -3 & -3 & 8 \end{pmatrix}\frac{\omega}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{Ma^2\omega}{6\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

so $\mathbf{L} \parallel \boldsymbol{\omega}$ and the diagonal is a proper (principal) axis.

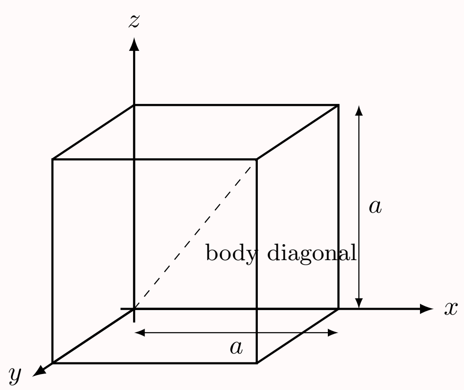
(b) Axes through the *center of mass* (set $x, y, z \in [-a/2, a/2]$).
By symmetry, products of inertia vanish and all three moments are equal. Explicitly,

$$I_{xx} = \rho\int_{-a/2}^{a/2}\int_{-a/2}^{a/2}\int_{-a/2}^{a/2}(y^2 + z^2)\,dy\,dz\,dx = \frac{1}{6}Ma^2,$$

and $I_{yy} = I_{zz} = \frac{1}{6}Ma^2$, $I_{xy} = I_{xz} = I_{yz} = 0$. Hence

$$\boxed{\mathbf{I}_{\text{CM}} = \frac{Ma^2}{6}\,\mathbf{1}_{3\times3}}$$

and the XYZ axes are principal axes.



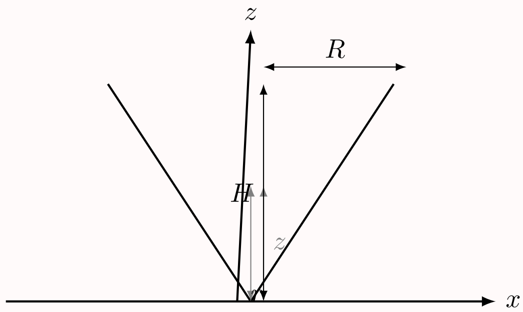
Example 5.6 – Inertia Tensor of a Solid Right Circular Cone (apex at O)

Consider a homogeneous solid cone of height H and base radius R, with the z–axis along its symmetry axis and apex at the origin. Use cylindrical coordinates (r, ϕ, z) with

$$0 \le z \le H, \qquad 0 \le \phi < 2\pi, \qquad 0 \le r \le r_{\max}(z) = \frac{R}{H}\,z.$$

Density ρ is constant, so the mass

$$M = \rho\,\text{Vol} = \rho\,\frac{\pi R^2 H}{3}.$$



Moment about the symmetry axis.

$$I_{zz} = \int \rho(x^2 + y^2)\, dV = \int \rho\, r^2 (r\, dr\, d\phi\, dz) = \rho \int_0^H \int_0^{2\pi} \int_0^{(R/H)z} r^3\, dr\, d\phi\, dz.$$

Compute

$$\int_0^{(R/H)z} r^3\, dr = \frac{1}{4}\left(\frac{R}{H}\right)^4 z^4, \qquad \int_0^{2\pi} d\phi = 2\pi, \qquad \int_0^H z^4\, dz = \frac{H^5}{5},$$

so

$$I_{zz} = \rho\, \frac{2\pi}{4} \left(\frac{R}{H}\right)^4 \frac{H^5}{5} = \rho\, \frac{\pi R^4 H}{10} = \frac{3}{10} M R^2.$$

Moments about x and y. By symmetry $I_{xx} = I_{yy}$ and

$$I_{xx} + I_{yy} = I_{zz} + 2 \int \rho z^2\, dV.$$

The auxiliary integral is

$$\int \rho z^2\, dV = \rho \int_0^H \int_0^{2\pi} \int_0^{(R/H)z} z^2 (r\, dr\, d\phi\, dz) = \rho\, 2\pi \int_0^H z^2 \left[\frac{1}{2}\left(\frac{R}{H}\right)^2 z^2\right] dz = \rho\, \frac{\pi R^2 H^3}{5} = \frac{3}{5} M H^2.$$

Hence

$$I_{xx} + I_{yy} = \frac{3}{10} M R^2 + 2 \cdot \frac{3}{5} M H^2 = \frac{3}{10} M \left(R^2 + 4H^2\right),$$

and therefore

$$\boxed{I_{xx} = I_{yy} = \frac{3}{20} M \left(R^2 + 4H^2\right), \qquad I_{zz} = \frac{3}{10} M R^2.}$$

Planes $x = 0$ and $y = 0$ are symmetry planes $\Rightarrow I_{xy} = I_{xz} = I_{yz} = 0$, so the inertia tensor in these axes is

$$\boxed{\mathbf{I}_{\text{cone}} = \begin{pmatrix} \frac{3}{20} M(R^2 + 4H^2) & 0 & 0 \\ 0 & \frac{3}{20} M(R^2 + 4H^2) & 0 \\ 0 & 0 & \frac{3}{10} M R^2 \end{pmatrix}.}$$

Rotation about z. For $\boldsymbol{\omega} = (0, 0, \omega)$, $\mathbf{L} = \mathbf{I}_{\text{cone}} \boldsymbol{\omega} = (0, 0, I_{zz}\omega)$, so $\mathbf{L} \parallel \boldsymbol{\omega}$ and z is a proper axis.

5.4. Foundations: I, symmetry, and first worked examples

From $\vec{L} = \sum m\, \vec{r} \times (\vec{\omega} \times \vec{r})$ to the inertia tensor

Starting from the C.M.–fixed viewpoint (so all positions are measured from the C.M.),

$$\vec{L} = \sum_\alpha m_\alpha \Big[(\vec{r}_\alpha \cdot \vec{r}_\alpha)\, \vec{\omega} - (\vec{r}_\alpha \cdot \vec{\omega})\, \vec{r}_\alpha \Big] \quad \Longrightarrow \quad L_i = \sum_j I_{ij}\, \omega_j, \tag{5.23}$$

with the *inertia tensor*

$$I_{ij} = \sum_\alpha m_\alpha\Big(r_\alpha^2\,\delta_{ij} - x_{\alpha i}x_{\alpha j}\Big), \qquad \vec{L} = \mathbf{I}\,\vec{\omega}. \tag{5.24}$$

In the continuum limit ($\rho(\vec{r})$ mass density),

$$I_{ij} = \int_V \rho(\vec{r})\Big(r^2\delta_{ij} - x_i x_j\Big)\,d^3r. \tag{5.25}$$

The matrix I_{ij} is real and *symmetric* ($I_{ij} = I_{ji}$), hence diagonalizable by an orthogonal matrix; only six independent components are needed.

Components.

$$I_{xx} = \sum m_\alpha(y_\alpha^2 + z_\alpha^2), \qquad I_{yy} = \sum m_\alpha(x_\alpha^2 + z_\alpha^2), \qquad I_{zz} = \sum m_\alpha(x_\alpha^2 + y_\alpha^2), \tag{5.26}$$

$$I_{xy} = I_{yx} = -\sum m_\alpha x_\alpha y_\alpha, \quad I_{xz} = I_{zx} = -\sum m_\alpha x_\alpha z_\alpha, \quad I_{yz} = I_{zy} = -\sum m_\alpha y_\alpha z_\alpha. \tag{5.27}$$

Two identities we will use repeatedly.

$$(\vec{A}\times\vec{B})^2 = A^2B^2 - (\vec{A}\cdot\vec{B})^2, \qquad \varepsilon_{ik\ell}\varepsilon_{jm\ell} = \delta_{ij}\delta_{km} - \delta_{im}\delta_{jk}. \tag{5.28}$$

Using (5.28), one also writes I_{ij} in Levi–Civita form:

$$I_{ij} = \sum_\alpha m_\alpha\,\varepsilon_{ik\ell}\varepsilon_{jm\ell}\,x_{\alpha k}x_{\alpha m}. \tag{5.29}$$

Rotational kinetic energy in terms of I. From $\vec{v}_\alpha = \vec{\omega}\times\vec{r}_\alpha$ and (5.28),

$$\begin{aligned} T_{\text{rot}} &= \tfrac12\sum_\alpha m_\alpha(\vec{\omega}\times\vec{r}_\alpha)^2 = \tfrac12\sum_\alpha m_\alpha\Big(\omega^2 r_\alpha^2 - (\vec{\omega}\cdot\vec{r}_\alpha)^2\Big) = \tfrac12\,\vec{\omega}\cdot\Big(\sum_\alpha m_\alpha\,\vec{r}_\alpha\times(\vec{\omega}\times\vec{r}_\alpha)\Big) \\ &= \boxed{\tfrac12\,\vec{\omega}\cdot\vec{L} = \tfrac12\,\vec{\omega}^T\mathbf{I}\,\vec{\omega}}. \end{aligned} \tag{5.30}$$

Transformation under a rotation of axes. If $\vec{L}' = A\vec{L}$ and $\vec{\omega}' = A\vec{\omega}$ with $A \in SO(3)$, then

$$\boxed{\mathbf{I}' = A\,\mathbf{I}\,A^T}, \qquad \operatorname{Tr}\mathbf{I} = \sum_a I_a, \quad \det\mathbf{I} = \prod_a I_a \ \text{(invariants)}. \tag{5.31}$$

Moment about an arbitrary axis. For a unit vector $\hat{n} = (\alpha,\beta,\gamma)^T$,

$$\boxed{I(\hat{n}) = \hat{n}^T\mathbf{I}\,\hat{n} = \alpha^2 I_{xx} + \beta^2 I_{yy} + \gamma^2 I_{zz} + 2\alpha\beta I_{xy} + 2\alpha\gamma I_{xz} + 2\beta\gamma I_{yz}}. \tag{5.32}$$

Symmetry tests you should always try first

- A plane of symmetry perpendicular to $\hat{x}$ forces $I_{xy} = I_{xz} = 0$; similarly for the other axes.
- Three mutually orthogonal symmetry planes *diagonalize* **I**.
- Any symmetry axis through the origin is a *principal axis* (eigenvector of **I**).
- For a lamina in $z = 0$ (all $z_\alpha = 0$):

$$I_{xz} = I_{yz} = 0, \qquad I_{zz} = I_{xx} + I_{yy} \quad \text{(perpendicular–axis theorem)}. \tag{5.33}$$

Example 5.7 – Planar mass in the xy–plane

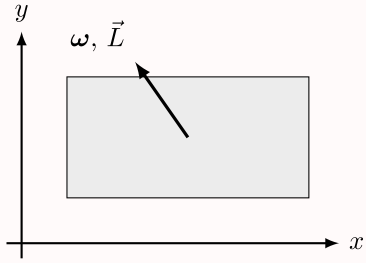
With $z_\alpha = 0$ for all particles,

$$I_{xx} = \sum m_\alpha y_\alpha^2, \quad I_{yy} = \sum m_\alpha x_\alpha^2, \quad I_{zz} = \sum m_\alpha(x_\alpha^2 + y_\alpha^2) = I_{xx} + I_{yy}, \quad I_{xz} = I_{yz} = 0.$$

Thus

$$\mathbf{I} = \begin{pmatrix} I_{xx} & I_{xy} & 0 \\ I_{xy} & I_{yy} & 0 \\ 0 & 0 & I_{xx}+I_{yy} \end{pmatrix}.$$

If the lamina rotates about $\hat{z}$, then $\vec{L} = \mathbf{I}\,\boldsymbol{\omega} = I_{zz}\omega\,\hat{z}$, so $\vec{L} \parallel \boldsymbol{\omega}$ and $\hat{z}$ is a principal axis.



Example 5.8 – Uniform solid cube about a corner *vs.* about its C.M.

Let a cube of side a occupy $0 \le x, y, z \le a$ (origin at a corner), with density $\rho = M/a^3$.
(i) About the corner $(0,0,0)$.

$$I_{xx} = \rho \int_0^a \int_0^a \int_0^a (y^2+z^2)\,dy\,dz\,dx = \frac{2}{3}Ma^2,$$

and cyclically $I_{yy} = I_{zz} = \frac{2}{3}Ma^2$. Products of inertia:

$$I_{xy} = -\rho \int_0^a \int_0^a \int_0^a xy\,dx\,dy\,dz = -\frac{1}{4}Ma^2, \qquad I_{xz} = I_{yz} = -\frac{1}{4}Ma^2.$$

Hence

$$\boxed{\mathbf{I}_{\text{corner}} = \frac{Ma^2}{12}\begin{pmatrix} 8 & -3 & -3 \\ -3 & 8 & -3 \\ -3 & -3 & 8 \end{pmatrix}}.$$

Two instructive checks:

$$\boldsymbol{\omega} = (0,0,\omega) \;\Rightarrow\; \vec{L} = \tfrac{Ma^2\omega}{12}(-3,-3,8) \;\nparallel\; \boldsymbol{\omega},$$

so the z–edge is *not* a principal axis about the corner. But for the body diagonal $\hat{u} = \frac{1}{\sqrt{3}}(1,1,1)$,

$$\mathbf{I}_{\text{corner}}\hat{u} = \frac{Ma^2}{12}(2,2,2)^T \;\Rightarrow\; I_{\text{diag(corner)}} = \frac{Ma^2}{6} \quad \text{and} \quad \vec{L} \parallel \boldsymbol{\omega}.$$

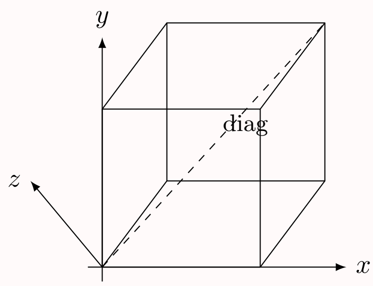
(ii) About the C.M. Shift the origin to the center; symmetry planes $x=0, y=0, z=0$ kill all products, and the three moments are equal:

$$I_{xx} = I_{yy} = I_{zz} = \rho \int_{-a/2}^{a/2}\int_{-a/2}^{a/2}\int_{-a/2}^{a/2} (y^2+z^2)\,dy\,dz\,dx = \frac{1}{6}Ma^2.$$

Thus

$$\boxed{\mathbf{I}_{\text{CM}} = \frac{Ma^2}{6}\,\mathbf{1}_{3\times 3}}$$

and every axis through the C.M. is principal (full cubic symmetry).



Example 5.9 – Solid right circular cone (apex at O

Homogeneous cone of height H and base radius R. In cylindrical coordinates,

$$0 \le z \le H, \quad 0 \le \phi < 2\pi, \quad 0 \le r \le (R/H)z, \qquad M = \rho\,\frac{\pi R^2 H}{3}.$$

Axial moment I_{zz}.

$$I_{zz} = \int \rho(x^2+y^2)\,dV = \rho\int_0^H\int_0^{2\pi}\int_0^{(R/H)z} r^2(r\,dr\,d\phi\,dz) = \rho\,\frac{\pi R^4 H}{10} = \boxed{\frac{3}{10}MR^2}.$$

Transverse moments $I_{xx} = I_{yy}$. Use $I_{xx}+I_{yy} = I_{zz} + 2\int \rho z^2\,dV$:

$$\int \rho z^2\,dV = \rho\int_0^H\int_0^{2\pi}\int_0^{(R/H)z} z^2(r\,dr\,d\phi\,dz) = \rho\,2\pi\int_0^H z^2\frac{1}{2}\left(\frac{R}{H}\right)^2 z^2\,dz = \rho\,\frac{\pi R^2 H^3}{5} = \frac{3}{5}MH^2.$$

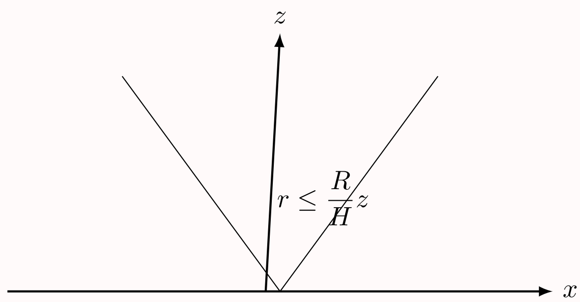
Therefore

$$I_{xx}+I_{yy} = \frac{3}{10}MR^2 + 2\cdot\frac{3}{5}MH^2 = \frac{3}{10}M\big(R^2+4H^2\big) \;\Rightarrow\; \boxed{I_{xx} = I_{yy} = \frac{3}{20}M\big(R^2+4H^2\big)}.$$

Planes $x=0$ and $y=0$ are symmetry planes $\Rightarrow I_{xy} = I_{xz} = I_{yz} = 0$, so

$$\boxed{\mathbf{I}_{\text{cone}} = \begin{pmatrix} \frac{3}{20}M(R^2+4H^2) & 0 & 0 \\ 0 & \frac{3}{20}M(R^2+4H^2) & 0 \\ 0 & 0 & \frac{3}{10}MR^2 \end{pmatrix}},$$

and the z–axis is a principal axis (indeed, an axis of symmetry).



Remarks tying these results to what comes next

- The symmetry/diagonalization tricks above are exactly what we will need to set up the Euler equations in a *body frame* aligned with principal axes, where $\mathbf{I} = \mathrm{diag}(I_1, I_2, I_3)$ is constant in time.
- The identity $T_{\text{rot}} = \frac{1}{2}\vec{\omega}\cdot\vec{L}$ (5.30) will be our main tool to expose conservation laws in torque–free motion and to interpret precession geometrically.
- In the next part we formalize principal axes via the eigenvalue problem for $\mathbf{I}$, derive the tensor form of the parallel–axis theorem, and introduce the inertia ellipsoid—then proceed to Euler's equations and their integrals.

5.5. Principal Axes of Inertia

Eigenvalue problem and diagonalization

The inertia tensor $\mathbf{I}$ defined in (5.24) is real and symmetric, so it admits an orthonormal eigenbasis. Principal axes are the directions along which $\vec{L}$ is parallel to $\vec{\omega}$:

$$\mathbf{I}\,\hat{v} = \lambda\,\hat{v} \quad \Longleftrightarrow \quad \vec{L} = \lambda\,\vec{\omega} \text{ whenever } \vec{\omega} \parallel \hat{v}. \tag{5.34}$$

Nontrivial solutions require the characteristic equation

$$\det(\mathbf{I} - \lambda\mathbf{1}) = 0, \tag{5.35}$$

which yields three (real) eigenvalues $I_1 \le I_2 \le I_3$ and orthonormal eigenvectors $\hat{v}_1, \hat{v}_2, \hat{v}_3$ with

$$\mathbf{I}\,\hat{v}_a = I_a\,\hat{v}_a, \qquad \hat{v}_a \cdot \hat{v}_b = \delta_{ab}, \qquad a, b \in \{1, 2, 3\}. \tag{5.36}$$

Let $U = (\,\hat{v}_1\ \hat{v}_2\ \hat{v}_3\,)$, then $U^T U = \mathbf{1}$ and

$$U^T \mathbf{I}\, U = \operatorname{diag}(I_1, I_2, I_3). \tag{5.37}$$

Two invariants that are useful checks in computations are

$$\operatorname{Tr}\mathbf{I} = I_1 + I_2 + I_3, \qquad \det\mathbf{I} = I_1 I_2 I_3. \tag{5.38}$$

Positive–definiteness. For any unit vector $\hat{n}$,

$$\hat{n}^T \mathbf{I}\,\hat{n} = \sum_\alpha m_\alpha \big(r_\alpha^2 - (\hat{n}\cdot\vec{r}_\alpha)^2\big) = \sum_\alpha m_\alpha\, d_\alpha^2 \ \ge 0, \tag{5.39}$$

the sum of squared distances d_α from the axis $\hat{n}$. Thus $\mathbf{I}$ is positive–definite if the body is not confined to the axis.

Theorem 5.1 – Axis of symmetry ⇒ principal axis

Let $\hat{n}$ be an axis of rotational symmetry through the origin. Then $R(\varphi)\,\mathbf{I}\,R(\varphi)^T = \mathbf{I}$ for every rotation $R(\varphi)$ about $\hat{n}$. Applying $R(\varphi)$ to $\mathbf{I}\,\hat{n}$ shows $\mathbf{I}\,\hat{n}$ is invariant under all such rotations, hence must be parallel to $\hat{n}$; i.e. $\hat{n}$ is an eigenvector of $\mathbf{I}$.

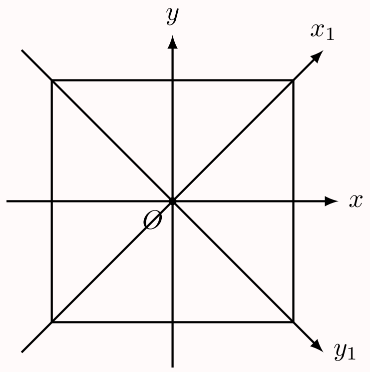
Degeneracy and nonuniqueness. If $I_1 = I_2 \ne I_3$ (axial symmetry), any orthonormal pair within the plane orthogonal to $\hat{v}_3$ is also principal; there are infinitely many principal triads. A plate with $I_{xx} = I_{yy}$ (about the C.M.) is a familiar example: rotating the in–plane axes by any angle leaves $\mathbf{I}$ diagonal with the same entries.

Example 5.10 – Square lamina: many principal triads

For a uniform square plate (side a, mass M) lying in $z = 0$,

$$\mathbf{I} = \frac{Ma^2}{12}\operatorname{diag}(1, 1, 2) \quad \text{in } (x, y, z)\text{, but equally in } (x_1, y_1, z),$$

for any in–plane rotation (x_1, y_1) by angle θ. Here $I_{xx} = I_{yy}$ makes every choice of in–plane orthonormal axes principal; z is the symmetry axis.



Moment of inertia about an arbitrary axis and the inertia ellipsoid

Let $\hat{n} = (\alpha, \beta, \gamma)$ be the direction cosines of an axis through the origin. The *axis moment* is the quadratic form

$$I(\hat{n}) = \hat{n}^T \mathbf{I}\, \hat{n} = \alpha^2 I_{xx} + \beta^2 I_{yy} + \gamma^2 I_{zz} + 2\alpha\beta I_{xy} + 2\alpha\gamma I_{xz} + 2\beta\gamma I_{yz}. \tag{5.40}$$

In principal axes, (5.40) reduces to $I(\hat{n}') = I_1\alpha'^2 + I_2\beta'^2 + I_3\gamma'^2$. The *inertia ellipsoid* is the locus

$$\hat{\rho}^T \mathbf{I}\, \hat{\rho} = 1 \quad \iff \quad I_1(\rho_1')^2 + I_2(\rho_2')^2 + I_3(\rho_3')^2 = 1 \text{ in principal axes,} \tag{5.41}$$

i.e. the quadratic surface associated with the positive–definite form $\mathbf{I}$. Diagonalizing this ellipsoid (*via* the orthogonal change of basis U) is exactly the principal–axes construction above.

How I changes when the coordinate frame changes

We need two basic operations: rotations of axes and shifts of origin.

Rotation of axes. If the new frame is obtained by a proper rotation $A \in SO(3)$, then vectors transform as $\vec{L}' = A\vec{L}$, $\vec{\omega}' = A\vec{\omega}$. From $\vec{L} = \mathbf{I}\vec{\omega}$,

$$\boxed{\mathbf{I}' = A\, \mathbf{I}\, A^T}. \tag{5.42}$$

Shift of the origin (parallel–axis theorem, tensor form). Let (x, y, z) be C.M.–based, and (X, Y, Z) be translated by $\vec{a}$ (no rotation). For a mass element, $\vec{R} = \vec{a} + \vec{r}$. The inertia tensor about the shifted origin is

$$\begin{aligned} J_{ij} &= \sum_\alpha m_\alpha \Big(\delta_{ij} R_\alpha^2 - R_{\alpha i} R_{\alpha j} \Big) \\ &= \sum_\alpha m_\alpha \Big(\delta_{ij} (r_\alpha^2 + 2\vec{a}\cdot\vec{r}_\alpha + a^2) - (r_{\alpha i} + a_i)(r_{\alpha j} + a_j) \Big). \end{aligned} \tag{5.43}$$

Using $\sum_\alpha m_\alpha \vec{r}_\alpha = \vec{0}$ (C.M. at the origin), the mixed terms drop and we get

$$\boxed{J_{ij} = I_{ij} + M\Big(a^2 \delta_{ij} - a_i a_j\Big)}, \qquad M = \sum_\alpha m_\alpha. \tag{5.44}$$

Particularly, if $\vec{a} \perp \hat{z}$ with $d = \|\vec{a}\|$, $J_{zz} = I_{zz} + Md^2$ (the familiar scalar parallel–axis theorem).

Combined change (rotate *and* shift). If the new frame is first shifted by $\vec{a}$ from the C.M. and then rotated by A,

$$\boxed{\mathbf{J}' = A\Big(\mathbf{I} + M(a^2\mathbf{1} - \vec{a}\,\vec{a}^T)\Big)A^T}. \tag{5.45}$$

Theorem 5.2 – Choosing the body frame sensibly

Because I_{ij} depends on where the origin is and on the orientation of axes, we *always* choose the body (rotating) frame to be C.M.–based and aligned with the principal axes. In that frame all mass elements have fixed coordinates and $\mathbf{I}$ is constant in time, which is ideal for dynamics.

Preparation for Euler's equations (next part)

We now collect the pieces needed for the equations of motion in a rotating body frame. First, the fixed–vs–rotating derivative relation (Chap. 4) applied to any vector $\vec{G}$:

$$\left.\frac{d\vec{G}}{dt}\right|_{\text{inertial}} = \left.\frac{d\vec{G}}{dt}\right|_{\text{body}} + \vec{\omega} \times \vec{G}. \tag{5.46}$$

With $\vec{G} = \vec{L}$ and $\vec{N}$ the total external torque about the C.M. (in the inertial frame),

$$\left.\frac{d\vec{L}}{dt}\right|_{\text{inertial}} = \vec{N} \quad \Longrightarrow \quad \boxed{\left.\frac{d\vec{L}}{dt}\right|_{\text{body}} + \vec{\omega} \times \vec{L} = \vec{N}}. \tag{5.47}$$

In a principal–axes body frame,

$$\vec{L} = \mathbf{I}\,\vec{\omega}, \qquad \mathbf{I} = \operatorname{diag}(I_1, I_2, I_3), \tag{5.48}$$

and (5.47) becomes Euler's equations (derived explicitly in Part 3).

Translational ⊕ rotational split of the Lagrangian (summary)

With the origin at the C.M. and a body–fixed principal frame,

$$T = \tfrac{1}{2} M\,\dot{\vec{R}}^2 + \tfrac{1}{2}\,\vec{\omega}^T \mathbf{I}\,\vec{\omega}, \qquad \mathcal{L}(\vec{R}, \dot{\vec{R}}; \text{orientation}, \vec{\omega}) = T - U, \tag{5.49}$$

so Newton's law governs $\vec{R}(t)$, $M\ddot{\vec{R}} = \vec{F}_{\text{ext}}$, while the rotational dynamics follows from (5.47). In Part 3 we turn (5.47) into the component form of Euler's equations, extract the conserved quantities for $\vec{N} = \vec{0}$, and study stability about principal axes.

5.6. Euler Equations in a Principal–Axes Body Frame

With origin at the C.M. and the body axes aligned with the principal directions,

$$\vec{L} = \mathbf{I}\,\vec{\omega}, \qquad \mathbf{I} = \operatorname{diag}(I_1, I_2, I_3), \tag{5.50}$$

and from the fixed–vs–rotating derivative relation (cf. Part 2, (5.47))

$$\left.\frac{d\vec{L}}{dt}\right|_{\text{body}} + \vec{\omega} \times \vec{L} = \vec{N}. \tag{5.51}$$

Taking components in the principal frame yields the *Euler equations*

$$\boxed{\begin{aligned} I_1\,\dot{\omega}_1 - (I_2 - I_3)\,\omega_2\omega_3 &= N_1, \\ I_2\,\dot{\omega}_2 - (I_3 - I_1)\,\omega_3\omega_1 &= N_2, \\ I_3\,\dot{\omega}_3 - (I_1 - I_2)\,\omega_1\omega_2 &= N_3. \end{aligned}} \tag{5.52}$$

Compact Levi–Civita form (no sum on i on the lhs):

$$I_i\,\dot{\omega}_i + \sum_{j,k} \epsilon_{ijk}\,\omega_j\,L_k = N_i, \qquad L_k = I_k\,\omega_k. \tag{5.53}$$

Torque–free motion and first integrals

Set $\vec{N} = \vec{0}$ in (5.52):

$$\begin{aligned} I_1\dot{\omega}_1 &= (I_2 - I_3)\,\omega_2\omega_3, \\ I_2\dot{\omega}_2 &= (I_3 - I_1)\,\omega_3\omega_1, \\ I_3\dot{\omega}_3 &= (I_1 - I_2)\,\omega_1\omega_2. \end{aligned} \tag{5.54}$$

Theorem 5.3 – Conserved quantities for $\vec{N}$

For torque–free motion:

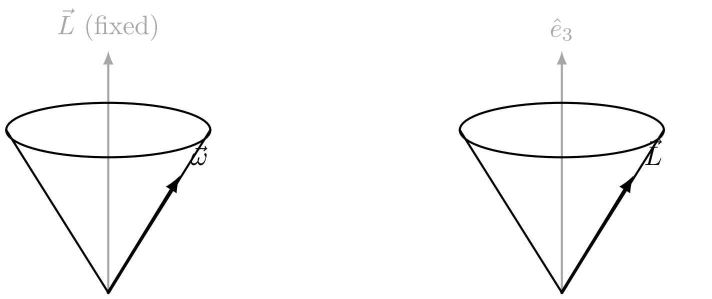
$$T_{\text{rot}} = \tfrac{1}{2}\,\vec{\omega}\cdot\mathbf{I}\,\vec{\omega} = \frac{1}{2}\left(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2\right) = \text{const}, \tag{5.55}$$

$$\left.\frac{d\vec{L}}{dt}\right|_{\text{inertial}} = \vec{0} \quad\Longrightarrow\quad \vec{L} = \text{const in magnitude and direction (in inertial space)}, \tag{5.56}$$

$$\vec{\omega}\cdot\vec{L} = 2T_{\text{rot}} = \text{const}. \tag{5.57}$$

Sketch. Multiply (5.54) componentwise by ω_i and sum to get $\frac{d}{dt}\left(\frac{1}{2}\sum I_i\omega_i^2\right) = 0$. Equation (5.56) is Newton's second law for rotation in the inertial frame with zero external torque. The last identity follows from (5.55) and $\vec{L} = \mathbf{I}\vec{\omega}$.

Geometric picture (Poinsot construction). At fixed energy T_{rot} the vector $\vec{\omega}$ lies on the *energy ellipsoid* $\vec{\omega}^{\,T}\mathbf{I}\,\vec{\omega} = 2T_{\text{rot}}$ in body space; at fixed $\vec{L}$ it must also satisfy $\vec{L} = \mathbf{I}\,\vec{\omega}$ with $\|\vec{L}\|$ constant in space. The motion is the rolling (without slipping) of the inertia ellipsoid on the *invariable plane* perpendicular to $\vec{L}$; equivalently, $\vec{\omega}$ precesses about $\vec{L}$ on a *space cone*, while in the body frame $\vec{L}$ precesses about a body principal axis on a *body cone*.



space cone (of $\vec{\omega}$ about $\vec{L}$) body cone (of $\vec{L}$ about $\hat{e}_3$)

Steady rotations (relative equilibria) and their stability

A *steady rotation* is a solution of (5.52) with constant $\vec{\omega}$ in the body frame. For $\vec{N} = \vec{0}$, the only such solutions are

$$\vec{\omega} = \omega_a\,\hat{e}_a \quad\text{with}\quad a \in \{1, 2, 3\}, \tag{5.58}$$

i.e. spins about a principal axis.

Example 5.11 – Linear stability about a principal–axis steady spin

Let $\vec{\omega}(t) = \omega_1\,\hat{e}_1 + \varepsilon_2(t)\,\hat{e}_2 + \varepsilon_3(t)\,\hat{e}_3$ with $|\varepsilon_{2,3}| \ll |\omega_1|$ and $\vec{N} = \vec{0}$. From (5.54),

$$\dot{\omega}_1 = \mathcal{O}(\varepsilon_2\varepsilon_3) \approx 0, \qquad \begin{cases} I_2\dot{\varepsilon}_2 = (I_3 - I_1)\,\omega_1\,\varepsilon_3, \\ I_3\dot{\varepsilon}_3 = (I_1 - I_2)\,\omega_1\,\varepsilon_2. \end{cases}$$

Differentiate the first and eliminate ε_3 to obtain

$$\ddot{\varepsilon}_2 + \lambda_1^2\,\varepsilon_2 = 0, \qquad \boxed{\lambda_1^2 = \omega_1^2\,\frac{(I_1 - I_2)(I_1 - I_3)}{I_2 I_3}}, \tag{5.59}$$

and similarly for ε_3. Therefore:

- If I_1 is the *smallest* or *largest* principal moment, then $(I_1 - I_2)(I_1 - I_3) > 0$ and $\lambda_1^2 > 0 \Rightarrow$ stable small oscillations.
- If I_1 is the *intermediate* moment, then $(I_1 - I_2)(I_1 - I_3) < 0$ and $\lambda_1^2 < 0 \Rightarrow$ exponential growth (instability).

By cyclic permutation,

$$\boxed{\begin{aligned} \Omega_1 &= |\omega_1|\sqrt{\frac{(I_1 - I_2)(I_1 - I_3)}{I_2 I_3}}, \\ \Omega_2 &= |\omega_2|\sqrt{\frac{(I_2 - I_3)(I_2 - I_1)}{I_3 I_1}}, \\ \Omega_3 &= |\omega_3|\sqrt{\frac{(I_3 - I_1)(I_3 - I_2)}{I_1 I_2}}. \end{aligned}} \tag{5.60}$$

Thus the **intermediate–axis theorem**: steady rotation is unstable about the axis with the middle principal moment, but stable about the axes with the largest and the smallest moments.

Energy viewpoint (alternative stability test). At fixed $\|\vec{L}\|$, the steady spins (5.58) are stationary points of $T_{\rm rot} = \frac{1}{2}\vec{\omega}\cdot\vec{L}$ subject to $\vec{L}$ fixed (Lagrange multipliers). The smallest and largest principal axes are local minimum/maximum of $T_{\rm rot}$ on the constraint, hence Lyapunov–stable; the intermediate axis is a saddle, hence unstable.

A convenient complex form for planar components (used later)

It will be useful (especially for symmetric tops) to combine the in–plane components as

$$\eta(t) \equiv \omega_1(t) + i\,\omega_2(t), \tag{5.61}$$

so that, when $I_1 = I_2 \equiv I$ (axial symmetry) and $\omega_3 = \text{const}$ (torque–free), the pair (5.54) reduces to

$$\dot{\eta} - i\,\Omega\,\eta = 0, \qquad \Omega \equiv \frac{I_3 - I}{I}\,\omega_3, \tag{5.62}$$

with solution $\eta(t) = A\,e^{i(\Omega t + \delta)}$. We will exploit this in Part 4.

5.7. Symmetric Top with One Point Fixed ($I_1 = I_2 \equiv I \neq I_3$)

We attach body axes $(\hat{e}_1, \hat{e}_2, \hat{e}_3)$ to the top, with $\hat{e}_3$ the symmetry axis. The fixed point (pivot) is at the origin O; the C.M. lies a distance ℓ along $\hat{e}_3$. We use the ZXZ Euler angles from Chap. 4 (Sec. 4.3): rotate by ϕ about space z, by θ about the new x', and by ψ about the new z'' (see (**??**)–(4.33), (4.35)).

Angular velocity components in the body frame. From Chap. 4 (cf. the column form of the rotation matrix $A = A_\psi A_\theta A_\phi$), the body–frame components are

$$\boxed{\omega_1 = \dot{\phi}\,\sin\theta\,\sin\psi + \dot{\theta}\,\cos\psi, \qquad \omega_2 = \dot{\phi}\,\sin\theta\,\cos\psi - \dot{\theta}\,\sin\psi, \qquad \omega_3 = \dot{\phi}\,\cos\theta + \dot{\psi}.} \tag{5.63}$$

Kinetic and potential energies. Axial symmetry ($I_1 = I_2 = I$) gives

$$T = \tfrac{1}{2}I(\omega_1^2+\omega_2^2)+\tfrac{1}{2}I_3\omega_3^2 = \frac{I}{2}\Big(\dot\theta^2+\dot\phi^2\sin^2\theta\Big)+\frac{I_3}{2}\Big(\dot\psi+\dot\phi\cos\theta\Big)^2, \tag{5.64}$$

$$U = Mg\ell\,\cos\theta \quad \text{(pivot at } O\text{, C.M. along } \hat e_3). \tag{5.65}$$

Thus the Lagrangian $\mathcal{L} = T - U$ is

$$\boxed{\mathcal{L}(\phi,\theta,\psi;\dot\phi,\dot\theta,\dot\psi) = \frac{I}{2}\Big(\dot\theta^2+\dot\phi^2\sin^2\theta\Big)+\frac{I_3}{2}\Big(\dot\psi+\dot\phi\cos\theta\Big)^2 - Mg\ell\,\cos\theta.} \tag{5.66}$$

Two cyclic coordinates and first integrals

Angles ϕ and ψ do not appear explicitly in $\mathcal{L}$, hence their conjugate momenta are constants:

$$p_\psi = \frac{\partial\mathcal{L}}{\partial\dot\psi} = I_3\Big(\dot\psi+\dot\phi\cos\theta\Big) = I_3\,\omega_3 \;\equiv\; \underbrace{I_3 a}_{\text{const}}, \tag{5.67}$$

$$p_\phi = \frac{\partial\mathcal{L}}{\partial\dot\phi} = I\sin^2\theta\,\dot\phi + I_3\cos\theta\Big(\dot\psi+\dot\phi\cos\theta\Big) = I\sin^2\theta\,\dot\phi + I_3 a\,\cos\theta \;\equiv\; \underbrace{b}_{\text{const}}. \tag{5.68}$$

Physical meaning: $p_\psi = \vec L\cdot\hat e_3$ (body–axis spin component), and $p_\phi = \vec L\cdot\hat z$ (space–vertical component). We will use

$$\boxed{\dot\phi = \frac{b - I_3 a\cos\theta}{I\,\sin^2\theta}, \qquad \dot\psi = a - \dot\phi\cos\theta = a - \frac{(b-I_3 a\cos\theta)\cos\theta}{I\,\sin^2\theta}.} \tag{5.69}$$

Energy and effective one–dimensional problem in $\theta(t)$

The total energy

$$E = T + U = \frac{I}{2}\dot\theta^2 + \underbrace{\frac{(b-I_3 a\cos\theta)^2}{2I\,\sin^2\theta} + Mg\ell\,\cos\theta}_{V_{\text{eff}}(\theta)} + \frac{I_3 a^2}{2}. \tag{5.70}$$

The constant $\frac{1}{2}I_3 a^2$ can be dropped by defining $E' \equiv E - \frac{1}{2}I_3 a^2$.

$$\boxed{E' = \frac{I}{2}\,\dot\theta^2 + V_{\text{eff}}(\theta), \qquad V_{\text{eff}}(\theta) = \frac{(b-I_3 a\cos\theta)^2}{2I\,\sin^2\theta} + Mg\ell\,\cos\theta.} \tag{5.71}$$

Hence

$$\dot\theta^2 = \frac{2}{I}\Big(E' - V_{\text{eff}}(\theta)\Big), \qquad t - t_0 = \int_{\theta(t_0)}^{\theta(t)} \frac{d\theta}{\sqrt{\frac{2}{I}\big(E' - V_{\text{eff}}(\theta)\big)}}. \tag{5.72}$$

Dimensionless normalization and the $u = \cos\theta$ form. Define (as in Part 2)

$$\alpha \equiv \frac{2E'}{I}, \qquad \beta \equiv \frac{2Mg\ell}{I}, \qquad \tilde a \equiv \frac{I_3}{I}a = \frac{p_\psi}{I}, \qquad \tilde b \equiv \frac{b}{I} = \frac{p_\phi}{I}, \tag{5.73}$$

so that

$$\alpha = \dot\theta^2 + \frac{(\tilde b - \tilde a\cos\theta)^2}{\sin^2\theta} + \beta\cos\theta. \tag{5.74}$$

With $u = \cos\theta$ and $\dot u = -\dot\theta\sin\theta$,

$$\boxed{\dot u^{\,2} = f(u) \equiv (1-u^2)(\alpha-\beta u) - (\tilde b - \tilde a\,u)^2, \qquad u\in[-1,1].} \tag{5.75}$$

Expanding,

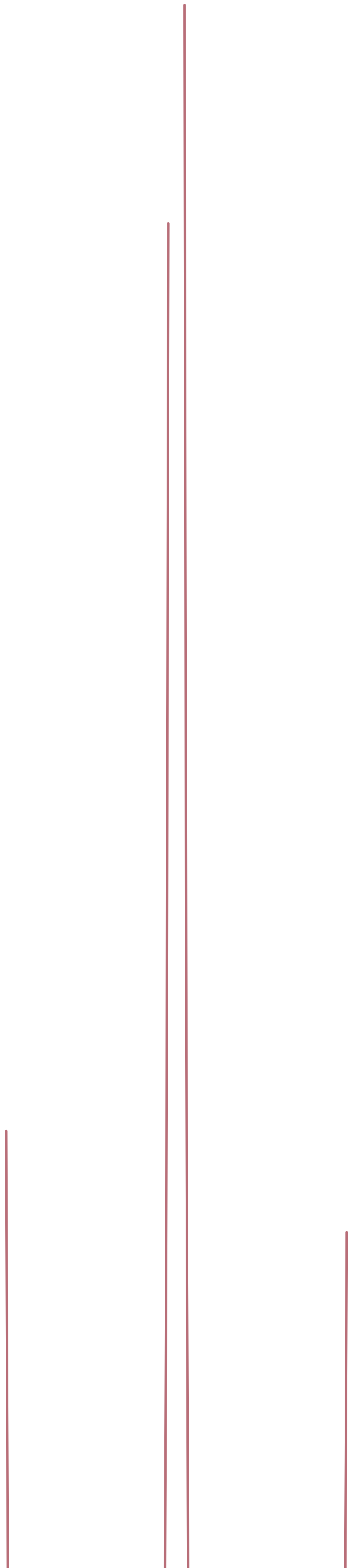
$$f(u) = \beta\,u^3 - (\alpha+\tilde a^{\,2})u^2 + (2\tilde a\tilde b - \beta)u + (\alpha - \tilde b^{\,2}). \tag{5.76}$$

Turning points of the nutation are the real roots $u_1 < u_2 < u_3$ of $f(u) = 0$ that lie in the physical interval $[-1, 1]$. Motion is allowed where $f(u) \geq 0$. The azimuthal rate has definite sign according to

$$\boxed{\dot{\phi} = \frac{\tilde{b} - \tilde{a}u}{1 - u^2} \quad \Longrightarrow \quad \operatorname{sign}(\dot{\phi}) = \operatorname{sign}(\tilde{b} - \tilde{a}u).} \tag{5.77}$$

Consequently:

- **Monotonic precession** if $\tilde{b}/\tilde{a} < u_1$ or $\tilde{b}/\tilde{a} > u_2$.
- **Reversing precession** if $u_1 < \tilde{b}/\tilde{a} < u_2$ (the sign flips once per half–nutation).
- **Cusp** if $\tilde{b}/\tilde{a} = u_1$ or u_2: here $\dot{\theta} = 0 = \dot{\phi}$.



Equation of motion for θ and steady precession

Differentiating $E' = \frac{1}{2} I\dot{\theta}^2 + V_{\text{eff}}(\theta)$ gives $I\ddot{\theta} = -\,dV_{\text{eff}}/d\theta$. A compact, useful form (using (5.69)) is

$$\boxed{I\,\ddot{\theta} = \frac{\big(b - I_3 a\cos\theta\big)\,\big(b\cos\theta - I_3 a\big)}{I\,\sin^3\theta} + Mg\ell\,\sin\theta.} \tag{5.78}$$

A *steady (regular) precession* has constant tilt $\theta(t) \equiv \theta_0$ (so $\dot{\theta} = \ddot{\theta} = 0$). Then (5.78) reduces to the quadratic for $\dot{\phi}$ (use $b = I\sin^2\theta_0\,\dot{\phi} + I_3 a\cos\theta_0$):

$$\boxed{I\,\cos\theta_0\,\dot{\phi}^{\,2} - I_3 a\,\dot{\phi} + Mg\ell = 0.} \tag{5.79}$$

Hence the two standard branches

$$\boxed{\dot{\phi}_{\pm} = \frac{I_3 a \;\pm\; \sqrt{I_3^2 a^2 - 4I\,Mg\ell\,\cos\theta_0}}{2I\cos\theta_0}, \qquad \dot{\psi}_{\pm} = a - \dot{\phi}_{\pm}\cos\theta_0.} \tag{5.80}$$

Existence condition:

$$\boxed{I_3^2 a^2 \;\geq\; 4I\,Mg\ell\,\cos\theta_0.} \tag{5.81}$$

For *rapid spin* ($I_3|a| \gg \sqrt{IMg\ell}$) and moderate θ_0,

$$\dot{\phi}_{-} \approx \frac{Mg\ell}{I_3 a} \quad \text{(slow precession)}, \qquad \dot{\phi}_{+} \approx \frac{I_3}{I}\,\frac{a}{\cos\theta_0} \quad \text{(fast precession)}. \tag{5.82}$$

Example 5.12 – Sleeping (upright) top and critical spin

For the upright steady orientation $\theta_0 = 0$ (symmetry axis along the vertical), (5.81) gives the *critical spin*:

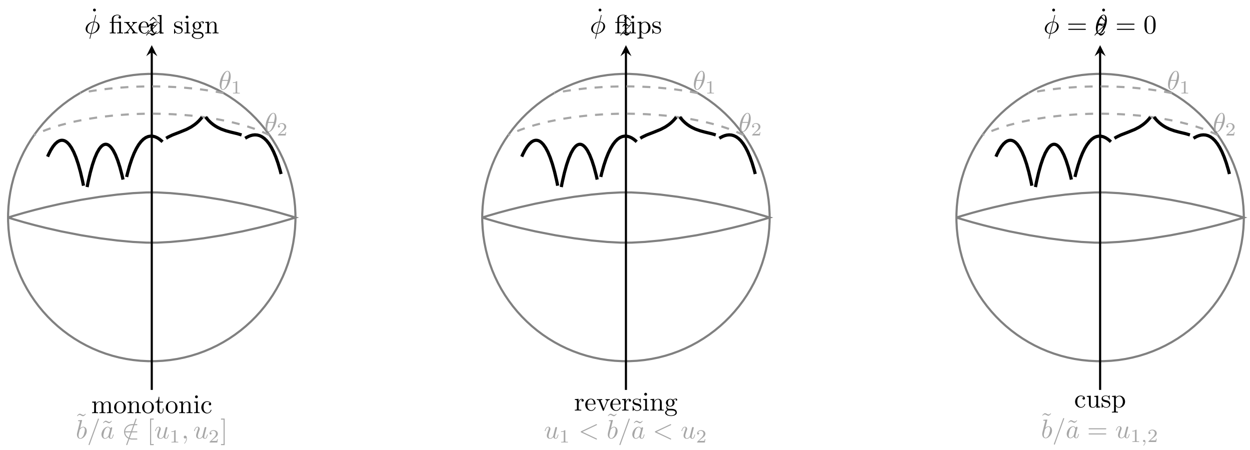
$$\boxed{|a| = |\omega_3| \;\geq\; \omega_{3,\text{crit}} = \frac{2}{I_3}\sqrt{I\,Mg\ell}\,.} \tag{5.83}$$

If $|\omega_3| > \omega_{3,\text{crit}}$, the vertical ("sleeping") top is dynamically stable and precesses slowly with $\dot{\phi} \simeq Mg\ell/(I_3\omega_3)$ (from (5.82)). At or below threshold, no real $\dot{\phi}$ solves (5.79); the top must nutate.

Nutation band, turning points, and sense (review with $u = \cos\theta$)

From (5.75)–(5.76), the two physical turning points $u_1 < u_2$ in $[0,1]$ (typical heavy top above the support plane) are the roots of $f(u) = 0$ with $f(u) \geq 0$ for $u \in [u_1, u_2]$. The azimuthal drift obeys $\text{sign}(\dot{\phi}) = \text{sign}(\tilde{b} - \tilde{a}u)$ ((5.77)):

- If $\tilde{b}/\tilde{a} < u_1$ or $\tilde{b}/\tilde{a} > u_2$: monotonic precession during nutation.
- If $u_1 < \tilde{b}/\tilde{a} < u_2$: precession reverses once each half–nutation.
- If $\tilde{b}/\tilde{a} = u_1$ or u_2: a cusp ($\dot{\phi} = 0 = \dot{\theta}$) occurs on the spherical trace.



Small nutations about a steady precession (optional)

Let $\theta(t) = \theta_0 + \varepsilon(t)$, with θ_0 solving $dV_{\rm eff}/d\theta = 0$. Expand $V_{\rm eff}(\theta) = V_{\rm eff}(\theta_0) + \frac{1}{2}K\,\varepsilon^2 + \dots$. Then from $E' = \frac{1}{2}I\dot\theta^2 + V_{\rm eff}$,

$$\ddot\varepsilon + \Omega_\theta^2\,\varepsilon = 0, \qquad \boxed{\Omega_\theta^2 = \frac{K}{I} = \frac{1}{I}\,\frac{d^2V_{\rm eff}}{d\theta^2}\bigg|_{\theta_0}.} \tag{5.84}$$

Inserting the explicit $V_{\rm eff}$ from (5.71) yields a (lengthy but straightforward) formula for K in terms of a, b, I, I_3, θ_0. For the upright case $\theta_0 = 0$, $\Omega_\theta^2 > 0$ reduces precisely to the critical–spin condition (5.83).

Full quadratures for the Lagrange (symmetric) top

With $I_1 = I_2 \equiv I$, $I_3 \neq I$, and one point fixed, we already reduced the motion to

$$E' = \frac{I}{2}\dot\theta^{\,2} + V_{\rm eff}(\theta), \qquad V_{\rm eff}(\theta) = \frac{(b - I_3 a\cos\theta)^2}{2I\sin^2\theta} + Mg\ell\cos\theta,$$

and introduced $u = \cos\theta$ so that

$$\dot u^{\,2} = f(u), \qquad f(u) = (1-u^2)(\alpha - \beta u) - (\tilde b - \tilde a\,u)^2, \tag{5.85}$$

with the dimensionless constants (cf. (5.73))

$$\alpha = \frac{2E'}{I}, \qquad \beta = \frac{2Mg\ell}{I}, \qquad \tilde a = \frac{I_3}{I}\,a = \frac{p_\psi}{I}, \qquad \tilde b = \frac{b}{I} = \frac{p_\phi}{I}.$$

Let the three real roots of $f(u) = 0$ be $u_1 < u_2 < u_3$ (generically). The physically allowed nutation interval is the subsegment within $[-1, 1]$ where $f(u) \geq 0$ (typically $u \in [u_1, u_2] \subset [0, 1]$ for a heavy top above the support plane).

Time as an elliptic integral. Since $dt = du/\dot u$, the time between two u-values is

$$t - t_0 = \int_{u(t_0)}^{u(t)} \frac{du}{\sqrt{f(u)}}. \tag{5.86}$$

One period of nutation (from u_1 to u_2 and back) is

$$T_\theta = 2\int_{u_1}^{u_2} \frac{du}{\sqrt{f(u)}}. \tag{5.87}$$

The integral is of Legendre's type after an affine change of variable mapping $[u_1, u_2]$ to $[0, 1]$.

Longitude and spin angles as quadratures. From (5.69) and $u = \cos\theta$,

$$\dot\phi = \frac{\tilde b - \tilde a\,u}{1 - u^2}, \qquad \dot\psi = a - \dot\phi\,u. \tag{5.88}$$

Therefore,

$$\phi(t) - \phi_0 = \int_{t_0}^{t} \frac{\tilde b - \tilde a\,u(t')}{1 - u(t')^2}\,dt' = \int_{u(t_0)}^{u(t)} \frac{\tilde b - \tilde a\,u}{(1-u^2)\sqrt{f(u)}}\,du, \tag{5.89}$$

and likewise

$$\psi(t) - \psi_0 = \int_{u(t_0)}^{u(t)} \left[\frac{a(1-u^2) - u(\tilde b - \tilde a\,u)}{(1-u^2)\sqrt{f(u)}}\right] du. \tag{5.90}$$

Both are elliptic integrals of the third kind in general.

Average precession during one nutation

Define the net azimuthal advance and time over a full θ-cycle:

$$\Delta\phi \equiv 2\int_{u_1}^{u_2} \frac{\tilde{b} - \tilde{a}\,u}{(1-u^2)\sqrt{f(u)}}\,du, \qquad T_\theta \equiv 2\int_{u_1}^{u_2} \frac{du}{\sqrt{f(u)}}.$$

The *mean precession rate* is

$$\boxed{\langle\dot\phi\rangle \;=\; \frac{\Delta\phi}{T_\theta} = \frac{\displaystyle\int_{u_1}^{u_2} \frac{\tilde{b} - \tilde{a}\,u}{(1-u^2)\sqrt{f(u)}}\,du}{\displaystyle\int_{u_1}^{u_2} \frac{du}{\sqrt{f(u)}}}}. \tag{5.91}$$

Similarly one defines $\langle\dot\psi\rangle$ by replacing the numerator integrand with the one in (5.90).

Example 5.13 – Fast top: asymptotics of the mean precession

For a rapid spinner with large $|a| = |\omega_3|$ one has a narrow nutation band around some θ_0 and $I_3|a| \gg \sqrt{IMg\ell}$. Expanding V_{eff} to quadratic order around θ_0 (Part 4) shows that

$$\theta(t) \approx \theta_0 + A\cos(\Omega_\theta t + \delta), \qquad \Omega_\theta^2 = \frac{1}{I}\,V''_{\text{eff}}(\theta_0).$$

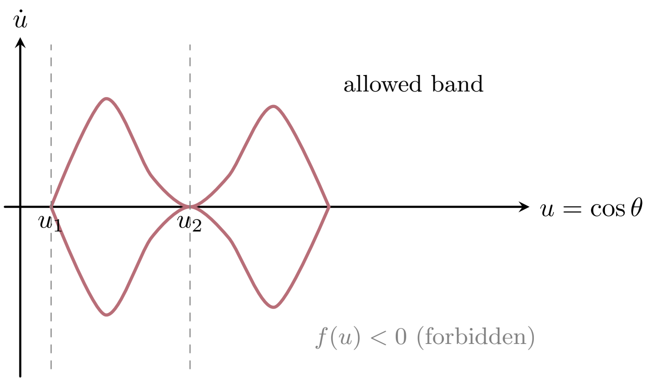
Averaging $\dot\phi = (\tilde{b} - \tilde{a}\cos\theta)/\sin^2\theta$ over one small oscillation yields, to leading order,

$$\boxed{\langle\dot\phi\rangle \;\simeq\; \frac{Mg\ell}{I_3 a} \quad \text{(slow precession branch)}, \qquad \langle\dot\psi\rangle \;\simeq\; a - \langle\dot\phi\rangle\cos\theta_0.} \tag{5.92}$$

Higher–order corrections are $\mathcal{O}(A^2)$ and depend on the small nutation amplitude.

Phase portrait in $(u, \dot u)$ and turning–point geometry

It is often helpful to picture the reduced dynamics as motion with "energy" α in the 1D effective landscape $f(u)$. Since $\dot u^2 = f(u)$, the phase curve is mirrored about $\dot u = 0$; the intercepts at $\dot u = 0$ mark the turning points $u_{1,2}$.



Special cases, limiting regimes, and checks

(i) Pure spin about the symmetry axis. Take $\theta(t) \equiv \theta_0$ and set $Mg\ell = 0$ (no gravity) or make ω_3 so large that the gravitational term is negligible over the time of interest. Then

$$\dot\phi = \frac{\tilde{b} - \tilde{a}\cos\theta_0}{\sin^2\theta_0}, \qquad \dot\psi = a - \dot\phi\cos\theta_0$$

are constants; the body spins about $\hat e_3$ while $\hat e_3$ precesses uniformly about the fixed $\vec{L}$ (Poinsot picture; see Part 3).

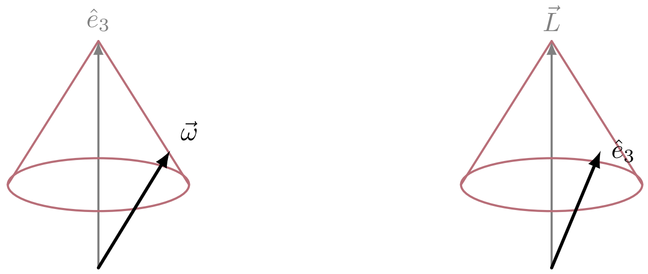
(ii) Sleeping top. With $\theta_0 = 0$, eq. (5.79) gives the quadratic in $\dot\phi$ and the existence condition $I_3^2 a^2 \geq 4IMg\ell$ (Part 4, (5.83)). The slow branch is $\dot\phi \simeq Mg\ell/(I_3 a)$, and small–nutation stability is equivalent to $|a| \geq \omega_{3,\text{crit}}$.

(iii) Weak gravity (or light top). If $\beta \ll \tilde a^2$, then $f(u) \approx (1-u^2)(\alpha) - (\tilde b - \tilde a\, u)^2$. The turning points are close and the nutation frequency Ω_θ is high compared with $\langle\dot\phi\rangle$; precession is slow.

(iv) Near–horizontal top. For $\theta \approx \pi/2$ ($u \approx 0$), the denominator $1-u^2$ is near unity whereas the gravity term is linear in u; small tilts away from $\pi/2$ are strongly influenced by the balance between $\tilde b$ and $\tilde a$.

Geometric view: cones revisited (space vs. body)

During torque–free motion (Part 3) the vector $\vec L$ is fixed in inertial space and $\vec\omega$ draws the *space cone* about $\vec L$; in the body the tip of $\vec\omega$ traces the *polhode* on the inertia ellipsoid. For the heavy (Lagrange) top, $\vec L$ is no longer constant in the body frame but $\vec L \cdot \hat z = p_\phi$ and $\vec L \cdot \hat e_3 = p_\psi$ remain constant. The symmetry axis $\hat e_3$ traces a latitude band on the unit sphere between θ_2 and θ_1; the longitudinal drift is determined by the sign rule $\text{sign}(\dot\phi) = \text{sign}(\tilde b - \tilde a u)$ (Part 4).



body cone　　space cone

Worked example: admissible band, reversals, and mean precession (symbolic route)

Given physical data I, I_3, M, g, ℓ and initial conditions $(\theta_0, \phi_0, \psi_0; \dot\theta_0, \dot\phi_0, \dot\psi_0)$:

1. Compute the first integrals $p_\psi = I_3(\dot\psi_0 + \dot\phi_0 \cos\theta_0)$, $p_\phi = I \sin^2\theta_0\, \dot\phi_0 + I_3 \cos\theta_0(\dot\psi_0 + \dot\phi_0\cos\theta_0)$, and $E = T + U$ using (5.64)–(5.65). Form $\alpha, \beta, \tilde a, \tilde b$.
2. Form $f(u)$ from (5.85)–(5.76) and find its real roots $u_1 < u_2 < u_3$. The physical nutation range is where $f(u) \geq 0$ within $[-1, 1]$.
3. Check the *precession sense* on this band using $\text{sign}(\dot\phi) = \text{sign}(\tilde b - \tilde a u)$. If $\tilde b/\tilde a \in (u_1, u_2)$, the precession reverses once per half–nutation; otherwise it is monotonic.
4. Evaluate T_θ by (5.87) and the mean $\langle\dot\phi\rangle$ by (5.91) (either numerically or by reducing to standard elliptic integrals).

Example 5.14 – Steady precession band and onset of nutation

Suppose the top is released with $\dot\theta_0 = 0$ at $\theta = \theta_0$ and with the steady–precession relation (5.79) satisfied by $\dot\phi_0$ (choose the slow branch). Then $dV_{\text{eff}}/d\theta|_{\theta_0} = 0$ and the motion is initially steady. If the actual spin $|a|$ is decreased *below* the threshold set by (5.81), the steady solution ceases to exist and the motion must nutate between turning angles $\theta_2 < \theta < \theta_1$, determined by $E' = V_{\text{eff}}(\theta)$.

Chapter 6

Coupled Oscillations

There is a large number of physical systems that exhibit oscillatory motion about a stable equilibrium and are treated elegantly using the Lagrangian formalism developed earlier. Small oscillations appear in molecular vibrations, coupled mechanical systems, electrical networks, etc. When the deviations from equilibrium are sufficiently small, the dynamics are well described by a system of *coupled linear harmonic oscillators*.

6.1. Formulation of the Problem

Consider a system with n degrees of freedom described by the Lagrangian

$$\mathcal{L}(q_1, \ldots, q_n; \dot{q}_1, \ldots, \dot{q}_n; t).$$

Assume a *stable* equilibrium at $\vec{q}_0 = (q_{01}, \ldots, q_{0n})$, characterized by the vanishing generalized forces

$$Q_i \;=\; -\left.\frac{\partial V}{\partial q_i}\right|_0 \;=\; 0, \qquad i = 1, \ldots, n, \tag{6.1}$$

and zero velocities at equilibrium. In a neighborhood of $\vec{q}_0$ set

$$q_i \;=\; q_{0i} + \eta_i, \qquad |\eta_i| \text{ small}, \tag{6.2}$$

and expand V in a Taylor series:

$$V(q_1, \ldots, q_n) = V(\vec{q}_0) + \sum_i \left.\frac{\partial V}{\partial q_i}\right|_0 \eta_i + \frac{1}{2}\sum_{i,j} \left.\frac{\partial^2 V}{\partial q_i \partial q_j}\right|_0 \eta_i\eta_j + \cdots . \tag{6.3}$$

Since the linear term vanishes at a minimum (Eq. (6.1)), and dropping the constant $V(\vec{q}_0)$, the potential to quadratic order is

$$V(\boldsymbol{\eta}) = \frac{1}{2}\sum_{i,j=1}^{n} V_{ij}\,\eta_i\eta_j, \qquad V_{ij} \equiv \left.\frac{\partial^2 V}{\partial q_i \partial q_j}\right|_0 = V_{ji}. \tag{6.4}$$

Likewise, to quadratic order in velocities the kinetic energy can be written

$$T = \frac{1}{2}\sum_{i,j=1}^{n} T_{ij}\,\dot{\eta}_i\dot{\eta}_j, \qquad T_{ij} = T_{ji} > 0, \tag{6.5}$$

where T_{ij} are the (constant) elements of the *mass* (kinetic) matrix evaluated at the equilibrium configuration. Hence

$$\mathcal{L} = \frac{1}{2}\sum_{i,j}\Big(T_{ij}\dot{\eta}_i\dot{\eta}_j - V_{ij}\eta_i\eta_j\Big). \tag{6.6}$$

The Euler–Lagrange equations give n coupled *linear, second-order* ODEs:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\eta}_k} - \frac{\partial \mathcal{L}}{\partial \eta_k} = 0, \qquad k = 1, \ldots, n. \tag{6.7}$$

Since

$$\frac{\partial \mathcal{L}}{\partial \dot{\eta}_k} = \frac{1}{2} \sum_{i,j} T_{ij} \left(\delta_{ik} \dot{\eta}_j + \delta_{jk} \dot{\eta}_i \right) = \sum_j T_{kj}\, \dot{\eta}_j \quad (\text{using } T_{ij} = T_{ji}),$$

and

$$\frac{\partial \mathcal{L}}{\partial \eta_k} = -\sum_j V_{kj}\, \eta_j,$$

we arrive at

$$\sum_{j=1}^{n} \Big(T_{kj}\, \ddot{\eta}_j + V_{kj}\, \eta_j \Big) = 0, \qquad k = 1, \dots, n. \tag{6.8}$$

In all problems of interest here T and V are real symmetric matrices. For a stable equilibrium, T is positive definite and V is positive definite (no zero modes) at $\vec{q}_0$.

6.2. The Eigenvalue Equation and Principal–Axes Transformation

Because the system is conservative, small deviations execute harmonic motion. Seek normal-mode solutions

$$\eta_i(t) = A_i\, e^{i\omega t}, \tag{6.9}$$

for which Eq. (6.8) becomes

$$\sum_{j=1}^{n} \left(V_{kj} - \omega^2 T_{kj} \right) A_j = 0, \qquad k = 1, \dots, n. \tag{6.10}$$

In matrix form,

$$\left(V - \omega^2 T \right) \mathbf{A} = \mathbf{0}, \qquad \mathbf{A} = \begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix}. \tag{6.11}$$

Non-trivial solutions require the characteristic (secular) equation

$$\det\!\left(V - \omega^2 T \right) = 0, \tag{6.12}$$

a polynomial of degree n in ω^2. Its n (real, non-negative) roots ω_r^2 are the *eigenfrequencies*. For each ω_r, the corresponding eigenvector $\mathbf{A}_r$ is obtained from (6.11) up to an overall scale.
Equivalently, we may write the generalized eigenvalue problem

$$V\, \mathbf{A} = \lambda\, T\, \mathbf{A}, \qquad \lambda = \omega^2. \tag{6.13}$$

Multiplying on the left by $\mathbf{A}^T$ and using symmetry,

$$\omega^2 = \frac{\mathbf{A}^T V\, \mathbf{A}}{\mathbf{A}^T T\, \mathbf{A}} \;\geq\; 0, \tag{6.14}$$

(the Rayleigh quotient), which is real and positive for $T > 0$, $V > 0$.

Orthogonality of Normal Modes

Let ω_r^2 and $\mathbf{A}_r$ $(r = 1, \dots, n)$ solve

$$V\, \mathbf{A}_r = \omega_r^2\, T\, \mathbf{A}_r. \tag{6.15}$$

Similarly,

$$V\, \mathbf{A}_s = \omega_s^2\, T\, \mathbf{A}_s. \tag{6.16}$$

Left-multiply (6.15) by $\mathbf{A}_s^T$ and (6.16) by $\mathbf{A}_r^T$:

$$\mathbf{A}_s^T V\, \mathbf{A}_r = \omega_r^2\, \mathbf{A}_s^T T\, \mathbf{A}_r, \qquad \mathbf{A}_r^T V\, \mathbf{A}_s = \omega_s^2\, \mathbf{A}_r^T T\, \mathbf{A}_s.$$

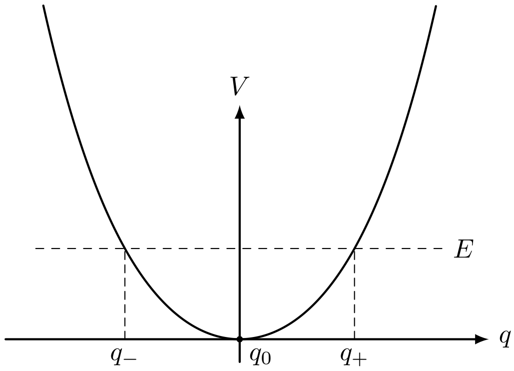


Figure 6.1: Potential near a stable equilibrium q_0. For a fixed energy E, the motion is confined between the turning points $q_\pm$.

Using symmetry ($T^T = T$, $V^T = V$) so that $\mathbf{A}_s^T T\,\mathbf{A}_r = \mathbf{A}_r^T T\,\mathbf{A}_s$ and $\mathbf{A}_s^T V\,\mathbf{A}_r = \mathbf{A}_r^T V\,\mathbf{A}_s$, subtraction yields

$$(\omega_r^2 - \omega_s^2)\,\mathbf{A}_r^T T\,\mathbf{A}_s = 0. \tag{6.17}$$

Thus for $\omega_r \neq \omega_s$,

$$\mathbf{A}_r^T T\,\mathbf{A}_s = 0, \tag{6.18}$$

i.e. distinct normal modes are orthogonal with respect to the T–inner product. Within a degenerate subspace one can choose the eigenvectors T–orthonormal via Gram–Schmidt. We adopt the normalization

$$\mathbf{A}_s^T T\,\mathbf{A}_s = 1, \tag{6.19}$$

so that in general

$$\boxed{\mathbf{A}_r^T T\,\mathbf{A}_s = \delta_{rs}} \tag{6.20}$$

or in components,

$$\sum_{j,k=1}^{n} T_{jk}\,a_{jr}\,a_{ks} = \delta_{rs}. \tag{6.21}$$

Orthogonality of eigenvectors and modal expansion (cleaned)

From the generalized eigenvalue equations already stated,

$$V\,\mathbf{A}_r = \omega_r^2\,T\,\mathbf{A}_r, \qquad V\,\mathbf{A}_s = \omega_s^2\,T\,\mathbf{A}_s, \tag{6.15 \& 6.16}$$

left–multiply the first by $\mathbf{A}_s^T$ and the second by $\mathbf{A}_r^T$:

$$\mathbf{A}_s^T V\,\mathbf{A}_r = \omega_r^2\,\mathbf{A}_s^T T\,\mathbf{A}_r, \qquad \mathbf{A}_r^T V\,\mathbf{A}_s = \omega_s^2\,\mathbf{A}_r^T T\,\mathbf{A}_s.$$

Using $T^T = T$, $V^T = V$ gives

$$\mathbf{A}_s^T T\,\mathbf{A}_r = \mathbf{A}_r^T T\,\mathbf{A}_s, \qquad \mathbf{A}_s^T V\,\mathbf{A}_r = \mathbf{A}_r^T V\,\mathbf{A}_s \tag{??}$$

and subtraction yields

$$(\omega_r^2 - \omega_s^2)\,\mathbf{A}_r^T T\,\mathbf{A}_s = 0. \tag{6.17}$$

Hence for $\omega_r \neq \omega_s$,

$$\mathbf{A}_r^T T\,\mathbf{A}_s = 0, \qquad (T\text{–orthogonality; cf. (6.18)}) \tag{6.22}$$

and with the normalization choice (possible since $T > 0$)

$$\mathbf{A}_s^T T\,\mathbf{A}_s = 1, \tag{6.19}$$

we have the T–orthonormal set

$$\boxed{\mathbf{A}_r^T T\,\mathbf{A}_s = \delta_{rs}} \tag{6.20}$$

or, in components,

$$\sum_{j,k=1}^{n} T_{jk}\,a_{jr}\,a_{ks} = \delta_{rs}. \tag{6.21}$$

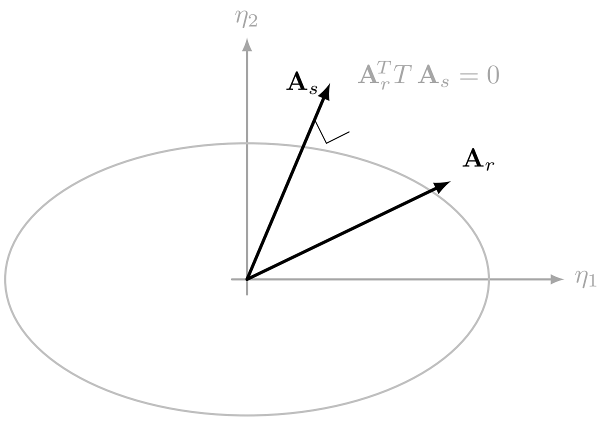


Figure 6.2: T–orthogonality of distinct normal modes: $\mathbf{A}_r^T T\,\mathbf{A}_s = 0$ for $r \neq s$.

Modal (normal–coordinate) expansion. Let the $n \times n$ matrix A collect the mode vectors as columns, $A = [\,\mathbf{A}_1 \;\; \cdots \;\; \mathbf{A}_n\,]$. Any small deviation $\boldsymbol{\eta}(t)$ can be expanded as

$$\boldsymbol{\eta}(t) = \sum_{r=1}^{n} \mathbf{A}_r\,\xi_r(t), \qquad \text{i.e.} \qquad \eta_j(t) = \sum_{r=1}^{n} a_{jr}\,\xi_r(t), \tag{6.23}$$

which defines the *normal coordinates* $\xi_r(t)$. Projecting with $\mathbf{A}_s^T T$ and using (6.20) gives the inversion formulas

$$\xi_s(t) = \mathbf{A}_s^T T\,\boldsymbol{\eta}(t), \qquad \dot{\xi}_s(t) = \mathbf{A}_s^T T\,\dot{\boldsymbol{\eta}}(t). \tag{6.24}$$

Each ξ_r obeys a single uncoupled oscillator equation:

$$\ddot{\xi}_r + \omega_r^2\,\xi_r = 0 \quad \Longrightarrow \quad \xi_r(t) = C_r\cos(\omega_r t - \delta_r) = \mathrm{Re}\big(c_r\,e^{i\omega_r t}\big), \tag{6.25}$$

with real amplitudes $C_r \geq 0$ and phases δ_r, or complex constants c_r.

Initial–value determination of modal amplitudes. Given initial data at $t = 0$,

$$\boldsymbol{\eta}(0), \qquad \dot{\boldsymbol{\eta}}(0),$$

the T–orthonormality (6.20) and the inversion (6.24) yield

$$\xi_r(0) = \mathbf{A}_r^T T\,\boldsymbol{\eta}(0), \qquad \dot{\xi}_r(0) = \mathbf{A}_r^T T\,\dot{\boldsymbol{\eta}}(0), \tag{6.26}$$

and hence (real form)

$$\xi_r(t) = \xi_r(0)\cos\omega_r t + \frac{\dot{\xi}_r(0)}{\omega_r}\,\sin\omega_r t. \tag{6.27}$$

Lagrangian in normal coordinates

With $\boldsymbol{\eta}(t) = \sum_r \mathbf{A}_r\,\xi_r(t)$, one has

$$\dot{\boldsymbol{\eta}} = \sum_r \mathbf{A}_r\,\dot{\xi}_r.$$

Using (6.20),

$$T = \frac{1}{2}\,\dot{\boldsymbol{\eta}}^T T\,\dot{\boldsymbol{\eta}} = \frac{1}{2}\sum_{r,s}\dot{\xi}_r\,\dot{\xi}_s\,\mathbf{A}_r^T T\,\mathbf{A}_s = \frac{1}{2}\sum_r \dot{\xi}_r^{\,2}, \tag{6.28}$$

$$V = \frac{1}{2}\,\boldsymbol{\eta}^T V\,\boldsymbol{\eta} = \frac{1}{2}\sum_{r,s}\xi_r\,\xi_s\,\mathbf{A}_r^T V\,\mathbf{A}_s = \frac{1}{2}\sum_r \omega_r^2\,\xi_r^{\,2}, \tag{6.29}$$

since $\mathbf{A}_r^T V\,\mathbf{A}_s = \omega_s^2\,\mathbf{A}_r^T T\,\mathbf{A}_s = \omega_r^2\,\delta_{rs}$. Therefore

$$\boxed{\mathcal{L} = T - V = \frac{1}{2}\sum_{r=1}^{n}\Big(\dot{\xi}_r^{\,2} - \omega_r^2\,\xi_r^{\,2}\Big)}\,, \tag{6.30}$$

i.e. a sum of n independent harmonic oscillators.

Recipe: eigenfrequencies, eigenvectors, and normal modes

1. Write the quadratic forms

$$T = \frac{1}{2}\sum_{j,k} T_{jk}\,\dot\eta_j\dot\eta_k, \qquad V = \frac{1}{2}\sum_{j,k} V_{jk}\,\eta_j\eta_k, \tag{6.31}$$

   identify the symmetric matrices T (positive definite) and V.

2. Solve the secular equation for the n eigenfrequencies:

$$\det\big(V - \omega^2 T\big) = 0 \quad \Rightarrow \quad \{\omega_r^2\}_{r=1}^n. \tag{6.32}$$

3. For each ω_r, solve the homogeneous system

$$\big(V - \omega_r^2 T\big)\,\mathbf{A}_r = \mathbf{0} \quad \Leftrightarrow \quad \sum_{j=1}^n (V_{kj} - \omega_r^2 T_{kj})\,a_{jr} = 0,\ \ k = 1,\dots,n, \tag{6.33}$$

   and normalize the eigenvectors by

$$\mathbf{A}_r^T T\,\mathbf{A}_s = \delta_{rs}. \tag{6.34}$$

4. Form the modal expansion $\boldsymbol{\eta}(t) = \sum_r \mathbf{A}_r\,\xi_r(t)$ and solve $\ddot\xi_r + \omega_r^2\xi_r = 0$ for each r with initial data projected by (6.26).

Example 6.1 – Two Coupled Masses and Springs — Normal Modes

Consider two equal masses m attached to fixed walls by identical springs of stiffness k, and coupled by a spring of stiffness k'. Let x_1, x_2 be the small longitudinal displacements (to the right) from the common equilibrium configuration.

Lagrangian and equations of motion.

$$\mathcal{L} = \frac{m}{2}\big(\dot x_1^2 + \dot x_2^2\big) - \frac{k}{2}\big(x_1^2 + x_2^2\big) - \frac{k'}{2}\,(x_2 - x_1)^2. \tag{6.35}$$

Euler–Lagrange equations

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot x_i} - \frac{\partial \mathcal{L}}{\partial x_i} = 0, \qquad i = 1, 2, \tag{6.36}$$

lead to

$$m\ddot x_1 + kx_1 + k'(x_1 - x_2) = 0, \qquad m\ddot x_2 + kx_2 + k'(x_2 - x_1) = 0.$$

Define

$$\omega_0^2 = \frac{k}{m}, \qquad \Omega^2 = \frac{k'}{m}. \tag{6.37}$$

With the center/relative coordinates

$$x_C = \frac{x_1 + x_2}{2}, \qquad x_R = \frac{x_1 - x_2}{2},$$

the system decouples:

$$\ddot x_C + \omega_0^2\,x_C = 0, \qquad \ddot x_R + \big(\omega_0^2 + 2\Omega^2\big)\,x_R = 0. \tag{6.38}$$

Hence the normal–mode frequencies are

$$\omega_C = \omega_0 = \sqrt{\frac{k}{m}}, \qquad \omega_R = \sqrt{\omega_0^2 + 2\Omega^2} = \sqrt{\frac{k + 2k'}{m}}.$$

General solutions:

$$x_C(t) = A_0\cos(\omega_0 t + \delta_0), \qquad x_R(t) = A_1\cos(\omega_R t + \delta_1). \tag{6.39}$$

Transforming back,

$$\begin{aligned} x_1(t) &= x_C(t) + x_R(t) = A_0\cos(\omega_0 t + \delta_0) + A_1\cos(\omega_R t + \delta_1),\\ x_2(t) &= x_C(t) - x_R(t) = A_0\cos(\omega_0 t + \delta_0) - A_1\cos(\omega_R t + \delta_1), \end{aligned}$$

i.e. each coordinate is a superposition of the two normal modes.

Matrix form and normal shapes (compare Eqs. (6.11)–(6.14)). With $\mathbf{x} = (x_1, x_2)^T$,

$$T = \tfrac{1}{2}\dot{\mathbf{x}}^T(m\mathbf{1})\,\dot{\mathbf{x}}, \qquad V = \tfrac{1}{2}\,\mathbf{x}^T\begin{pmatrix} k+k' & -k' \\ -k' & k+k' \end{pmatrix}\mathbf{x}. \tag{6.40}$$

The generalized eigenproblem $(V - \omega^2 T)\mathbf{A} = \mathbf{0}$ gives

$$\det\begin{pmatrix} k+k'-m\omega^2 & -k' \\ -k' & k+k'-m\omega^2 \end{pmatrix} = 0 \implies \omega^2 \in \left\{\frac{k}{m}, \frac{k+2k'}{m}\right\}. \tag{6.41}$$

Associated (unnormalized) eigenvectors:

$$\omega = \omega_0: \quad \mathbf{A}_C = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ (in–phase)}, \qquad \omega = \omega_R: \quad \mathbf{A}_R = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ (out–of–phase)}. \tag{6.42}$$

Since $T = m\mathbf{1}$, the T–orthonormal modes are

$$\widehat{\mathbf{A}}_C = \frac{1}{\sqrt{2m}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \widehat{\mathbf{A}}_R = \frac{1}{\sqrt{2m}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \Rightarrow \quad \widehat{\mathbf{A}}_\alpha^T T\,\widehat{\mathbf{A}}_\beta = \delta_{\alpha\beta}.$$

Normal coordinates (cf. (6.23)–(6.30)). Collecting modes as columns $A = [\,\mathbf{A}_C\ \mathbf{A}_R\,]$, set

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A\begin{pmatrix} \xi_C \\ \xi_R \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} \xi_C \\ \xi_R \end{pmatrix} \iff \begin{pmatrix} \xi_C \\ \xi_R \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_C \\ x_R \end{pmatrix}. \tag{6.43}$$

In (ξ_C, ξ_R) the Lagrangian is diagonal,

$$\mathcal{L} = \tfrac{1}{2}(\dot{\xi}_C^2 - \omega_0^2\xi_C^2) + \tfrac{1}{2}(\dot{\xi}_R^2 - \omega_R^2\xi_R^2),$$

and initial data project onto modes via

$$\xi_\alpha(0) = \widehat{\mathbf{A}}_\alpha^T T\,\mathbf{x}(0), \qquad \dot{\xi}_\alpha(0) = \widehat{\mathbf{A}}_\alpha^T T\,\dot{\mathbf{x}}(0), \qquad \alpha \in \{C, R\}. \tag{6.44}$$

6.2.1. Standard Approach (matrix method)

Example 6.2 – Two–mass system via the matrix eigenproblem (see Fig. ??)

With $\omega_0^2 = \frac{k}{m}$, $\Omega^2 = \frac{k'}{m}$ (cf. (6.37)), the equations of motion can be written as

$$\frac{d^2}{dt^2}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} \omega_0^2+\Omega^2 & -\Omega^2 \\ -\Omega^2 & \omega_0^2+\Omega^2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{6.45}$$

Normal–mode ansatz. Seek solutions of the form

$$\mathbf{X}(t) \equiv \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \mathbf{a}_n\,e^{i\omega_n t}, \qquad \mathbf{a}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \end{pmatrix}. \tag{6.46}$$

Substituting (9.60) into (9.59) gives the algebraic system

$$\underbrace{\begin{pmatrix} \omega_0^2+\Omega^2-\omega_n^2 & -\Omega^2 \\ -\Omega^2 & \omega_0^2+\Omega^2-\omega_n^2 \end{pmatrix}}_{\equiv K(\omega_n^2)}\begin{pmatrix} a_{1n} \\ a_{2n} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{6.47}$$

Characteristic equation and eigenfrequencies. Non-trivial $\mathbf{a}_n$ exist iff

$$\det K(\omega_n^2) = \det\begin{pmatrix} \omega_0^2+\Omega^2-\omega_n^2 & -\Omega^2 \\ -\Omega^2 & \omega_0^2+\Omega^2-\omega_n^2 \end{pmatrix} = 0. \tag{6.48}$$

Expanding,

$$(\omega_0^2+\Omega^2-\omega_n^2)^2-\Omega^4=0 \quad\Longrightarrow\quad \omega_0^2+\Omega^2-\omega_n^2=\pm\Omega^2.$$

Hence

$$\boxed{\omega_0=\sqrt{\tfrac{k}{m}}\,,\qquad \omega_1=\sqrt{\omega_0^2+2\Omega^2}=\sqrt{\tfrac{k+2k'}{m}}}\,. \tag{6.49}$$

Eigenvectors (mode shapes).

- For $\omega_n=\omega_0$: $K(\omega_0^2)=\begin{pmatrix}\Omega^2 & -\Omega^2\\ -\Omega^2 & \Omega^2\end{pmatrix}$, so $(a_{10}-a_{20})=0$ and we may choose $\mathbf{a}_0=\begin{pmatrix}1\\1\end{pmatrix}$ (in-phase).
- For $\omega_n=\omega_1$: $K(\omega_1^2)=\begin{pmatrix}-\Omega^2 & -\Omega^2\\ -\Omega^2 & -\Omega^2\end{pmatrix}$, so $(a_{11}+a_{21})=0$ and we may choose $\mathbf{a}_1=\begin{pmatrix}1\\-1\end{pmatrix}$ (out-of-phase).

With $T=m\mathbf{1}$, the T–orthonormal forms are

$$\widehat{\mathbf{a}}_0=\frac{1}{\sqrt{2m}}\begin{pmatrix}1\\1\end{pmatrix},\qquad \widehat{\mathbf{a}}_1=\frac{1}{\sqrt{2m}}\begin{pmatrix}1\\-1\end{pmatrix},\qquad\Rightarrow\qquad \widehat{\mathbf{a}}_\alpha^T\,T\,\widehat{\mathbf{a}}_\beta=\delta_{\alpha\beta}.$$

General solution and its real form. The complex normal–mode superposition is

$$\mathbf{X}(t)=A_0\begin{pmatrix}1\\1\end{pmatrix}e^{i\omega_0 t}+A_1\begin{pmatrix}1\\-1\end{pmatrix}e^{i\omega_1 t},\qquad A_\alpha\in\mathbb{C}. \tag{6.50}$$

Writing $A_\alpha=|A_\alpha|e^{i\delta_\alpha}$,

$$\mathbf{X}(t)=|A_0|\begin{pmatrix}1\\1\end{pmatrix}e^{i(\omega_0 t+\delta_0)}+|A_1|\begin{pmatrix}1\\-1\end{pmatrix}e^{i(\omega_1 t+\delta_1)}. \tag{6.51}$$

Taking the real part gives the physical displacements:

$$\begin{aligned} x_1(t)&=|A_0|\cos(\omega_0 t+\delta_0)+|A_1|\cos(\omega_1 t+\delta_1),\\ x_2(t)&=|A_0|\cos(\omega_0 t+\delta_0)-|A_1|\cos(\omega_1 t+\delta_1). \end{aligned} \tag{6.52}$$

Initial–data projection (useful in practice). Equivalently, introduce normal coordinates $x_C=\frac{1}{2}(x_1+x_2)$, $x_R=\frac{1}{2}(x_1-x_2)$ (cf. (6.43)). Then

$$x_C(t)=x_C(0)\cos\omega_0 t+\frac{v_C(0)}{\omega_0}\sin\omega_0 t,\quad x_R(t)=x_R(0)\cos\omega_1 t+\frac{v_R(0)}{\omega_1}\sin\omega_1 t,$$

with $v_C(0)=\frac{1}{2}(\dot{x}_1(0)+\dot{x}_2(0))$, $v_R(0)=\frac{1}{2}(\dot{x}_1(0)-\dot{x}_2(0))$, and $x_1=x_C+x_R,\ x_2=x_C-x_R$.

Coupled Pendula

Example 6.3 – Two identical pendula coupled by a light spring (small angles)

Consider two identical simple pendula (mass m, length b), whose pivots are separated by a fixed horizontal distance a_0. The bobs are connected by a light horizontal spring of constant k whose natural length equals the pivot separation a_0, so the spring is unstrained at equilibrium. Let θ_1,θ_2 be the small angular deflections from the downward vertical, measured positive to the right. To first order,

$$x_i\simeq b\,\theta_i,\qquad y_i\simeq -b+\mathcal{O}(\theta_i^2),$$

so the spring extension is $\Delta\ell\simeq(x_2-x_1)-a_0-(a_0-0)=b(\theta_2-\theta_1)$.

Kinetic energy. Each bob has speed $v_i^2 = \dot{x}_i^2 + \dot{y}_i^2 \simeq b^2\dot{\theta}_i^2$, hence

$$T = \frac{1}{2}mb^2\big(\dot{\theta}_1^2 + \dot{\theta}_2^2\big). \tag{6.53}$$

Gravitational potential. Choosing the zero at the lowest point ($\theta_i = 0$), one has

$$U_g = mg\big[(b - b\cos\theta_1) + (b - b\cos\theta_2)\big] = mgb\,(2 - \cos\theta_1 - \cos\theta_2).$$

Up to an irrelevant constant and to quadratic order,

$$U_g \simeq \frac{1}{2}mgb\big(\theta_1^2 + \theta_2^2\big). \tag{6.54}$$

Spring energy. With $\Delta\ell \simeq b(\theta_2 - \theta_1)$,

$$U_s = \frac{1}{2}k\,\Delta\ell^2 = \frac{1}{2}kb^2(\theta_2 - \theta_1)^2.$$

Total potential and Lagrangian. Thus

$$U = U_g + U_s \;=\; \frac{1}{2}mgb\big(\theta_1^2 + \theta_2^2\big) + \frac{1}{2}kb^2(\theta_2 - \theta_1)^2, \tag{6.55}$$

and

$$\mathcal{L} = T - U = \frac{1}{2}mb^2\big(\dot{\theta}_1^2 + \dot{\theta}_2^2\big) - \frac{1}{2}kb^2(\theta_2 - \theta_1)^2 - \frac{1}{2}mgb\big(\theta_1^2 + \theta_2^2\big). \tag{6.56}$$

Euler–Lagrange equations. Using $\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\theta}_i} - \frac{\partial\mathcal{L}}{\partial\theta_i} = 0$ (cf. Chap. 5),

$$mb^2\ddot{\theta}_1 + mgb\,\theta_1 + kb^2(\theta_1 - \theta_2) = 0, \qquad mb^2\ddot{\theta}_2 + mgb\,\theta_2 + kb^2(\theta_2 - \theta_1) = 0.$$

Matrix form and normal modes. Introduce $\omega_0^2 \equiv \dfrac{g}{b}$, $\omega_1^2 \equiv \dfrac{k}{m}$. With $\boldsymbol{\theta} = (\theta_1, \theta_2)^T$, the equations read

$$\ddot{\boldsymbol{\theta}} + \begin{pmatrix} \omega_0^2 + \omega_1^2 & -\omega_1^2 \\ -\omega_1^2 & \omega_0^2 + \omega_1^2 \end{pmatrix}\boldsymbol{\theta} = \vec{0}. \tag{6.57}$$

Seek $\boldsymbol{\theta}(t) = \mathbf{a}\,e^{i\omega t}$ to obtain

$$\begin{pmatrix} \omega_0^2 + \omega_1^2 - \omega^2 & -\omega_1^2 \\ -\omega_1^2 & \omega_0^2 + \omega_1^2 - \omega^2 \end{pmatrix}\mathbf{a} = \vec{0},$$

whose characteristic equation is

$$\det\begin{pmatrix} \omega_0^2 + \omega_1^2 - \omega^2 & -\omega_1^2 \\ -\omega_1^2 & \omega_0^2 + \omega_1^2 - \omega^2 \end{pmatrix} = 0. \tag{6.58}$$

It yields

$$(\omega_0^2 + \omega_1^2 - \omega^2)^2 - \omega_1^4 = 0 \implies \boxed{\omega_S = \omega_0, \quad \omega_F = \sqrt{\omega_0^2 + 2\omega_1^2}}.$$

The normalized eigenvectors (with respect to the mass metric $M = mb^2\,\mathbf{1}$) are

$$\mathbf{v}_S = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \quad \text{(symmetric, in phase)}, \qquad \mathbf{v}_F = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix} \quad \text{(antisymmetric, out of phase)}. \tag{6.59}$$

Hence the real solution is the superposition

$$\begin{pmatrix}\theta_1\\\theta_2\end{pmatrix} = A_S\begin{pmatrix}1\\1\end{pmatrix}\cos(\omega_S t + \delta_S) + A_F\begin{pmatrix}1\\-1\end{pmatrix}\cos(\omega_F t + \delta_F),$$

where the constants follow from initial data. In the decoupled limit $k \to 0$, $\omega_F \to \omega_S = \sqrt{g/b}$ as expected.

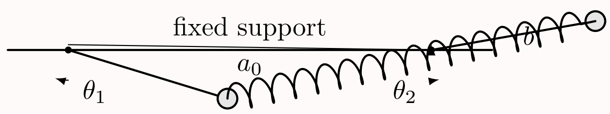
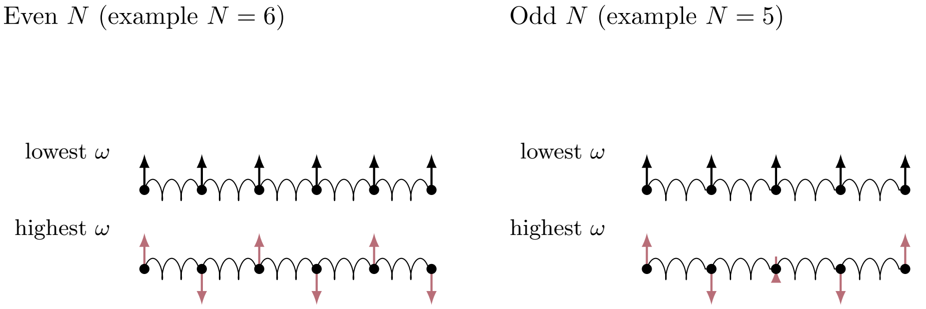


Figure 6.4: Two identical pendula (length b, mass m) coupled by a spring of constant k. Small angles: $x_i \simeq b\theta_i$, so $\Delta\ell \simeq b(\theta_2 - \theta_1)$.

Theorem 6.1 – Symmetry vs. frequency in uniform coupled systems

For a system of N identical oscillators with uniform nearest–neighbour couplings (positive spring constants) and equal masses, the *most symmetric* normal mode has the *lowest* frequency, while the *least symmetric* (alternating–sign) mode has the *highest* frequency.

Sketch. In mass–normalized coordinates the squared frequency is the Rayleigh quotient $\omega^2(\mathbf{q}) = \dfrac{\mathbf{q}^T K\,\mathbf{q}}{\mathbf{q}^T\mathbf{q}}$ (cf. (6.14) with $T = \mathbf{1}$), where K penalizes differences of neighbouring coordinates. The uniform vector $\mathbf{q} \propto (1, 1, \ldots, 1)$ minimizes $\mathbf{q}^T K\mathbf{q}$ because all differences vanish; the alternating vector $(1, -1, 1, -1, \ldots)$ maximizes those differences, hence maximizes ω^2.



Typical vibrational bands for identical masses/springs: symmetry $\Rightarrow$ low ω; alternating pattern $\Rightarrow$ high ω. Molecular vibrational frequencies lie mainly in the infrared: $10^{12}\,\text{Hz} < f_{\text{IR}} < 10^{14}\,\text{Hz}$.

6.2.2. Infinite Number of Coupled Oscillators and Transition to a Continuous String

Example 6.4 – Discrete transverse oscillations → dispersion → wave equation

Consider an infinite chain of equal point masses m separated by a distance d along the horizontal axis. The masses are joined by taut, light segments that sustain a constant tension τ. Horizontal motion is suppressed; let $y_j(t)$ be the small vertical displacement of the j-th mass from equilibrium. The net vertical force on mass j is the difference of the vertical components of the tensions in the adjacent segments:

$$F_j = \tau\sin\theta_{j+1} - \tau\sin\theta_j, \tag{6.60}$$

where θ_j is the (small) angle the j-th segment makes with the horizontal.
Using the small–angle approximation,

$$\sin\theta_{j+1} \simeq \frac{y_{j+1} - y_j}{d}, \qquad \sin\theta_j \simeq \frac{y_j - y_{j-1}}{d}, \tag{6.61}$$

Newton's second law gives the standard three–point stencil (discrete Laplacian)

$$m\,\ddot{y}_j = \frac{\tau}{d}\Big(y_{j-1} - 2y_j + y_{j+1}\Big). \tag{6.62}$$

Discrete Lagrangian. The dynamics follow from the quadratic Lagrangian

$$\mathcal{L} = \frac{1}{2}\sum_i d\left[\frac{m}{d}\,\dot{y}_i^2 - \tau\left(\frac{y_{i+1} - y_i}{d}\right)^2\right], \tag{6.63}$$

since the Euler–Lagrange equations

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{y}_j} - \frac{\partial\mathcal{L}}{\partial y_j} = 0 \qquad (j \in \mathbb{Z}) \tag{6.64}$$

reproduce (6.62). For later use note the linear system form

$$M\,\ddot{\mathbf{Y}} + K\,\mathbf{Y} = \mathbf{0}, \qquad \mathbf{Y} = (\ldots, y_{j-1}, y_j, y_{j+1}, \ldots)^T, \tag{6.65}$$

noting the crucial *plus* sign: $\ddot{\mathbf{Y}} = -M^{-1}K\,\mathbf{Y}$.

Translation invariance and plane waves. Define the discrete translation operator $(Ty)_j := y_{j+1}$. The linear difference operator

$$(\mathsf{L}y)_j := m\,\ddot{y}_j - \frac{\tau}{d}\Big(y_{j-1} - 2y_j + y_{j+1}\Big) \tag{6.66}$$

commutes with T: $[\mathsf{L}, T] = 0$. Hence we can choose simultaneous eigenvectors of L and T. Let $y_j(t) = a_j\, e^{i\omega t}$ and impose

$$T a_j = \beta\, a_j \;\Rightarrow\; a_{j+1} = \beta\, a_j \;\Rightarrow\; a_j = \beta^{\,j} a_0. \tag{6.67}$$

To avoid blow–up as $|j| \to \infty$ we require $|\beta| = 1$, so we set

$$\beta = e^{ikd} \quad (k \in \mathbb{R}), \qquad a_{j\pm1} = e^{\pm ikd}\, a_j. \tag{6.68}$$

Substituting $y_j = a_j e^{i\omega t}$ into (6.62) gives the **dispersion relation**

$$\omega^2(k) = \frac{2\tau}{md}\Big(1 - \cos kd\Big) = \frac{4\tau}{md}\,\sin^2\frac{kd}{2}. \tag{6.69}$$

Thus $\omega(k) = \omega(-k)$ and the $\pm k$ plane waves are degenerate. The corresponding eigenvector is

$$a_j = a_0\, e^{ijkd}. \tag{6.70}$$

General solution; fixed ends. Superposing the $\pm k$ waves,

$$y_j(t) = \mathrm{Re}\Big\{e^{i(\omega t+\delta)}\big(C_+ e^{ijkd} + C_- e^{-ijkd}\big)\Big\}. \tag{6.71}$$

Identifying $x = jd$ suggests the continuous form

$$y(x,t) = \mathrm{Re}\Big\{e^{i\delta}\big(C_+ e^{i(\omega t + kx)} + C_- e^{-i(\omega t+kx)}\big)\Big\}. \tag{6.72}$$

For a finite chain of N interior masses with fixed ends $y_0(t) = y_{N+1}(t) = 0$,

$$C_+ = -C_-, \qquad k_n = \frac{n\pi}{(N+1)d}, \quad n = 1, 2, \ldots, N, \tag{6.73}$$

and the discrete normal modes are $a_j^{(n)} \propto \sin(jk_n d)$ with

$$\omega_n = 2\sqrt{\frac{\tau}{md}}\,\sin\frac{n\pi}{2(N+1)}.$$

Long–wavelength limit and the wave equation. Let $d \to 0$ while keeping the length $L = (N+1)d$ fixed and the linear density

$$\mu = \frac{m}{d}$$

finite. The central difference approximates $y_{j-1} - 2y_j + y_{j+1} \simeq d^2\,\partial_{xx} y(x_j, t)$. Dividing (6.62) by d gives, in the limit,

$$\mu\,\frac{\partial^2 y}{\partial t^2} = \tau\,\frac{\partial^2 y}{\partial x^2}, \qquad c^2 := \frac{\tau}{\mu}. \tag{6.74}$$

Thus the transverse displacement obeys the **wave equation** $\partial_{tt} y = c^2\,\partial_{xx} y$ with wave speed $c = \sqrt{\tau/\mu}$.
From (6.69), for $kd \ll 1$ one has $\omega(k) \simeq c\,|k|$ (since $\sin(kd/2) \simeq kd/2$), so the discrete dispersion becomes linear and non-dispersive in the continuum limit.

Phase and group velocities (discrete chain). With $\omega(k) = 2\sqrt{\frac{\tau}{md}}\,\sin\frac{kd}{2}$,

$$v_p = \frac{\omega}{k} = \frac{2}{k}\sqrt{\frac{\tau}{md}}\,\sin\frac{kd}{2}, \qquad v_g = \frac{d\omega}{dk} = \sqrt{\frac{\tau d}{m}}\,\cos\frac{kd}{2} = c\,\cos\frac{kd}{2}.$$

Hence $v_g \to c$ as $kd \to 0$ and $v_g \to 0$ as $kd \to \pi$.

Continuous string with fixed ends. For a string of length L with $y(0,t) = y(L,t) = 0$, the normal modes are

$$y_n(x,t) = A_n \sin\frac{n\pi x}{L}\cos\big(\omega_n t + \delta_n\big), \qquad \omega_n = \frac{n\pi c}{L}, \quad n = 1, 2, \dots$$

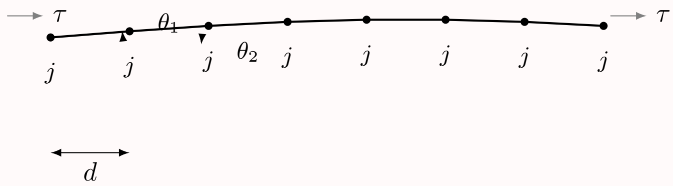


Figure 6.5: Discrete string: small transverse displacements $y_j(t)$, spacing d, constant tension τ. For long wavelengths ($kd \ll 1$) the dispersion $\omega^2 = \frac{4\tau}{md}\sin^2\frac{kd}{2}$ tends to $\omega = c|k|$ with $c = \sqrt{\tau/\mu}$ and $\mu = m/d$.

6.2.3. Transition from Discrete to Continuous Model

Taylor expansion, continuum limit, and d'Alembert solution
Starting from the discrete equation of motion derived above,

$$m\,\ddot{y}_j(t) = \frac{\tau}{d}\Big(y_{j-1}(t) - 2y_j(t) + y_{j+1}(t)\Big), \tag{6.75}$$

identify the lattice sites with points on the x–axis:

$$y_j(t) = y(jd,t) \equiv y(x,t), \qquad y_{j\pm1}(t) = y(x \pm d, t). \tag{6.76}$$

Then (6.75) reads

$$m\,\ddot{y}(x,t) = \frac{\tau}{d}\Big[y(x-d,t) - 2y(x,t) + y(x+d,t)\Big]. \tag{6.77}$$

Second–difference via Taylor expansion. For a smooth function F,

$$F(x \pm d) = F(x) \pm dF'(x) + \frac{d^2}{2}F''(x) \pm \frac{d^3}{6}F^{(3)}(x) + \frac{d^4}{24}F^{(4)}(x) + \mathcal{O}(d^5). \tag{6.78}$$

Hence

$$F(x+d) - 2F(x) + F(x-d) = d^2F''(x) + \frac{d^4}{12}F^{(4)}(x) + \mathcal{O}(d^6).$$

Applying this to $F(\cdot) = y(\cdot,t)$ and dividing (6.77) by d gives

$$\frac{m}{d}\,\ddot{y}(x,t) = \frac{\tau}{d^2}\left[d^2 y_{xx}(x,t) + \frac{d^4}{12}y_{xxxx}(x,t) + \mathcal{O}(d^6)\right] = \tau\,y_{xx}(x,t) + \mathcal{O}(d^2). \tag{6.79}$$

Defining the *linear mass density* $\mu = \dfrac{m}{d}$ and letting $d \to 0$ at fixed μ yields the **wave equation**

$$\mu\,y_{tt}(x,t) = \tau\,y_{xx}(x,t) \quad\Longleftrightarrow\quad y_{tt} - v_p^2 y_{xx} = 0, \qquad v_p := \sqrt{\frac{\tau}{\mu}}. \tag{6.80}$$

d'Alembert general solution. Introduce the characteristic variables $\xi = x - v_p t$ and $\zeta = x + v_p t$. If $f = f(\xi)$ then

$$\partial_x f = f'(\xi), \quad \partial_{xx} f = f''(\xi), \quad \partial_t f = -v_p f'(\xi), \quad \partial_{tt} f = v_p^2 f''(\xi),$$

so $\partial_{xx} f - \frac{1}{v_p^2}\partial_{tt} f = 0$. The same holds for $g(\zeta)$. Therefore the most general solution of (6.80) is the superposition of a right–moving and a left–moving wave:

$$y(x,t) = f(x - v_p t) + g(x + v_p t). \tag{6.81}$$

(Any equivalent notation—e.g. swapping f and g or the signs—represents the same family of solutions.)

Two counter-propagating pulses (superposition). Because (6.80) is linear, pulses pass through one another without distortion. At $t = 0$ choose, for example,

$$f(\xi) = \frac{A}{1 + \left(\xi - \xi_0\right)^2}, \qquad g(\zeta) = \frac{B}{1 + \left(\zeta - \zeta_0\right)^2},$$

so that initially one pulse is centered near $x = \xi_0$ and the other near $x = \zeta_0$. Their subsequent motion is $y(x,t) = f(x - v_p t) + g(x + v_p t)$. (Note: each shape must be a function of a *single* traveling argument, not of x and t independently.)

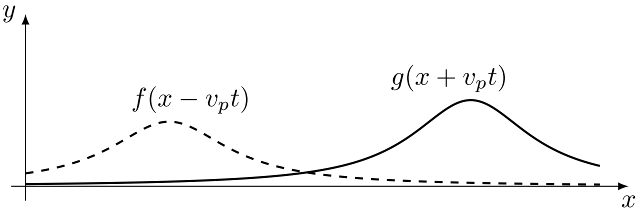


Figure 6.6: Two pulses at an instant $t > 0$: a right–mover $f(x - v_p t)$ (dashed) and a left–mover $g(x + v_p t)$ (solid). Superposition gives the exact solution.

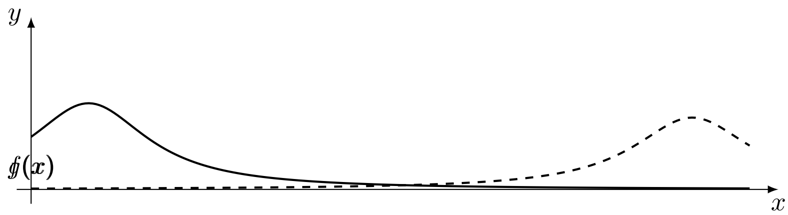


Figure 6.7: Example of well–separated initial pulses at $t = 0$. Under (6.80), the dashed pulse moves right with speed v_p, the solid one moves left with the same speed.

Remark 6.1 – Consistency with the discrete dispersion

From the lattice dispersion $\omega^2(k) = \dfrac{4\tau}{md} \sin^2 \dfrac{kd}{2}$ (see (6.69)), the long–wavelength limit $kd \ll 1$ gives $\omega \simeq v_p |k|$ with $v_p = \sqrt{\tau/\mu}$, in agreement with the continuum wave equation (6.80).

6.2.4. Separation of Variables

Example 6.5 – Separation of variables for a stretched string with fixed ends

We seek the normal modes of the wave equation (cf. (6.80))

$$\left(\frac{\partial^2}{\partial x^2} - \frac{1}{v_p^2} \frac{\partial^2}{\partial t^2} \right) \Psi(x,t) = 0, \tag{6.82}$$

as a basis for expanding general solutions like (6.81). Assume a product form

$$\Psi(x,t) = F(x)\, G(t).$$

Substitute into (6.82):

$$G(t)\, F''(x) \;=\; \frac{F(x)}{v_p^2}\, G''(t) \quad \Longrightarrow \quad \frac{F''(x)}{F(x)} = \frac{1}{v_p^2} \frac{G''(t)}{G(t)} = -k^2,$$

where the separation constant is chosen as $-k^2$ ($k \geq 0$) so that bounded spatial solutions can satisfy fixed–end boundary conditions.

Spatial and temporal factors. The separated ODEs are

$$F'' + k^2 F = 0, \qquad G'' + \omega^2 G = 0, \qquad \omega = v_p k.$$

Thus,

$$F(x) = A\cos(kx) + B\sin(kx), \tag{6.83}$$

$$G(t) = C\cos(\omega t) + D\sin(\omega t), \qquad \omega = v_p k. \tag{6.84}$$

Fixed–end boundary conditions. For a string of length L with both ends fixed,

$$\Psi(0,t) = 0, \qquad \Psi(L,t) = 0. \tag{6.85}$$

From $\Psi(0,t) = F(0)G(t) = 0$ for all t we must have $F(0) = 0 \Rightarrow A = 0$, so $F(x) = B\sin(kx)$. The second condition $\Psi(L,t) = B\sin(kL)G(t) = 0$ for all t requires

$$\sin(kL) = 0 \implies k_n = \frac{n\pi}{L}, \qquad n = 1, 2, 3, \ldots$$

(hence no $n = 0$ mode for fixed–fixed ends). The corresponding frequencies are

$$\omega_n = v_p k_n = \frac{n\pi v_p}{L}.$$

Normal modes and general expansion. The n-th normal mode (standing wave) is

$$\Psi_n(x,t) = \sin(k_n x)\left[a_n\cos(\omega_n t) + b_n\sin(\omega_n t)\right].$$

By superposition, the general solution is

$$\Psi(x,t) = \sum_{n=1}^{\infty} \sin(k_n x)\left[a_n\cos(\omega_n t) + b_n\sin(\omega_n t)\right]. \tag{6.86}$$

Equivalently, using a single–phase form,

$$\Psi(x,t) = \sum_{n=1}^{\infty} A_n \sin(k_n x)\,\sin(\omega_n t + \beta_n), \tag{6.87}$$

with the relation $A_n\sin\beta_n = a_n$ and $A_n\cos\beta_n = b_n$.

Orthogonality and determination from initial data. The sine modes satisfy

$$\int_0^L \sin(k_n x)\sin(k_m x)\,dx = \frac{L}{2}\,\delta_{mn}. \tag{6.88}$$

Given initial displacement $\Psi(x,0)$ and initial velocity $\dot{\Psi}(x,0)$,

$$\Psi(x,0) = \sum_{n=1}^{\infty} a_n\,\sin(k_n x), \tag{6.89}$$

$$\dot{\Psi}(x,0) = \sum_{n=1}^{\infty} \omega_n b_n\,\sin(k_n x). \tag{6.90}$$

Projecting with (6.88) yields the coefficients

$$a_m = \frac{2}{L}\int_0^L \Psi(x,0)\,\sin(k_m x)\,dx, \tag{6.91}$$

$$b_m = \frac{2}{L\,\omega_m}\int_0^L \dot{\Psi}(x,0)\,\sin(k_m x)\,dx. \tag{6.92}$$

If $\dot{\Psi}(x,0) = 0$, then $b_m = 0 \Rightarrow \beta_m = \frac{\pi}{2}$, so $\sin\beta_m = 1$ and

$$A_m = \frac{2}{L}\int_0^L \Psi(x,0)\,\sin(k_m x)\,dx, \tag{6.93}$$

in agreement with the phase form (6.87).

Remark 6.2 – Traveling–wave form

The separated factors (6.83)–(6.84) can be written with exponentials,

$$F(x) = Ae^{ikx} + Be^{-ikx}, \qquad G(t) = Ce^{i\omega t} + De^{-i\omega t}, \tag{6.94}$$

so a product term behaves like

$$\Psi(x,t) \sim e^{\pm ikx}\, e^{\pm i\omega t} = e^{i\frac{\omega}{v_p}(x \pm v_p t)},$$

i.e. a right– or left–moving traveling wave. Imposing fixed–end boundary conditions enforces standing–wave superpositions, which the sine basis $\{\sin(k_n x)\}$ provides.

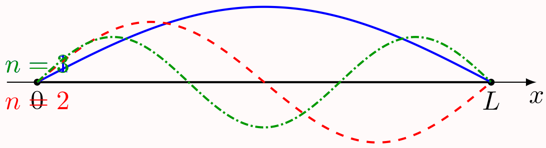


Figure 6.8: First three mode shapes $\sin(k_n x)$ with $k_n = n\pi/L$ for a string fixed at $x = 0, L$. Amplitudes are scaled for visibility.

Chapter 7

The Hamiltonian Equations of Motion

Classical mechanics admits several equivalent formulations. Newton's laws use forces and accelerations; the Lagrangian formulation (Hamilton's principle) uses generalized coordinates and eliminates constraint forces explicitly. The Hamiltonian formulation introduces no new physics, but reorganizes dynamics on *phase space* (q, p). Its geometric structure leads naturally to Liouville's theorem, Poisson brackets, canonical transformations, and underpins both statistical mechanics and quantum mechanics (cf. (7.1)).

Remark 7.1 – Hamiltonians in quantum and statistical mechanics

In quantum theory the Hamiltonian generates time evolution and spectra; in statistical mechanics it controls ensembles:

$$\rho = \frac{e^{-\beta H}}{Z}, \qquad \hat{H}\Psi = E\Psi, \qquad i\hbar\,\frac{dA}{dt} = [H, A], \ \dots \tag{7.1}$$

7.1. Hamiltonian as a Legendre Transform of the Lagrangian

Let $\mathcal{L}(q, \dot{q}, t)$ be a regular Lagrangian (precise regularity stated below). Define the *canonical momenta*

$$p_i \;\equiv\; \frac{\partial \mathcal{L}}{\partial \dot{q}_i}, \tag{7.2}$$

and the *Hamiltonian* (Legendre transform in the velocities)

$$\mathcal{H}(q, p, t) \;\equiv\; \sum_i p_i \dot{q}_i - \mathcal{L}(q, \dot{q}, t), \qquad \dot{q} = \dot{q}(q, p, t) \text{ via (7.2)}. \tag{7.3}$$

A direct differential yields

$$d\mathcal{H} = \sum_i \dot{q}_i\, dp_i - \sum_i \frac{\partial \mathcal{L}}{\partial q_i}\, dq_i - \frac{\partial \mathcal{L}}{\partial t}\, dt, \tag{7.4}$$

hence, as partial derivatives at fixed (q, p, t),

$$\frac{\partial \mathcal{H}}{\partial p_i} = \dot{q}_i, \qquad \frac{\partial \mathcal{H}}{\partial q_i} = -\frac{\partial \mathcal{L}}{\partial q_i}, \qquad \frac{\partial \mathcal{H}}{\partial t} = -\frac{\partial \mathcal{L}}{\partial t}. \tag{7.5}$$

Theorem 7.1 – Regularity of the Legendre transform

The velocity–momentum map $\dot{q} \mapsto p = \partial\mathcal{L}/\partial\dot{q}$ is invertible iff the Hessian

$$\mathbf{W}_{ij}(q, \dot{q}, t) \;\equiv\; \frac{\partial^2 \mathcal{L}}{\partial \dot{q}_i\, \partial \dot{q}_j} \tag{7.6}$$

is nonsingular. For natural systems ($\mathcal{L} = T(\dot{q}, q) - V(q, t)$ with T positive definite quadratic in $\dot{q}$), $\mathbf{W}$ is positive definite and the transform exists globally.

Example 7.1 – Natural Lagrangians

If $\mathcal{L} = \frac{1}{2}\,\dot{q}^T\mathbf{M}(q)\dot{q} - V(q, t)$ (cf. the quadratic forms used in the coupled–oscillations chapter), then

$$p = \mathbf{M}(q)\dot{q}, \qquad \mathcal{H}(q, p, t) = \tfrac{1}{2}\,p^T\mathbf{M}^{-1}(q)p + V(q, t).$$

Remark 7.2 – Energy function

For Lagrangians of the form $T(\dot{q}, q) - V(q, t)$ with T homogeneous of degree 2 in $\dot{q}$, the Legendre transform gives $\mathcal{H} = T + V$. If $\partial\mathcal{L}/\partial t = 0$, then $\partial\mathcal{H}/\partial t = 0$ and $\mathcal{H}$ is conserved (see (7.8)).

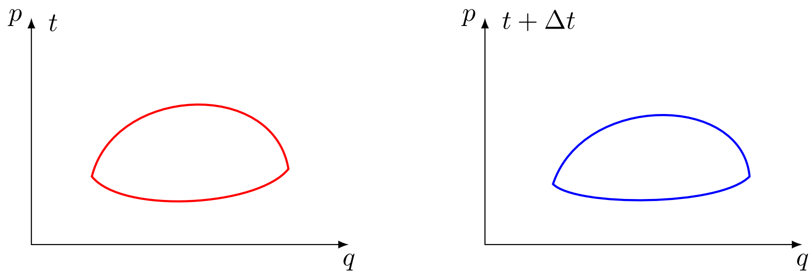


Figure 7.1: Phase–space (q, p) regions are carried into regions of equal volume under Hamiltonian evolution (Liouville's theorem).

7.2. Hamilton's Equations

Theorem 7.2 – Hamilton's canonical equations

Let $\mathcal{H}(q, p, t)$ be defined by (7.3). Then the equations of motion are

$$\dot{q}_i = \frac{\partial\mathcal{H}}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial\mathcal{H}}{\partial q_i}, \qquad i = 1, \ldots, n. \tag{7.7}$$

Proof. Using (7.5) and the Lagrange equations $\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\right) - \frac{\partial\mathcal{L}}{\partial q_i} = 0$, we identify $\partial\mathcal{H}/\partial p_i = \dot{q}_i$ and, *on-shell*, $\partial\mathcal{H}/\partial q_i = -\dot{p}_i$.

Remark 7.3 – Conservation laws (cyclic variables and time invariance)

If $\mathcal{H}$ (equivalently $\mathcal{L}$) is independent of q_k, then $\dot{p}_k = 0$ (p_k is conserved). If $\partial\mathcal{H}/\partial t = 0$, then

$$\frac{d\mathcal{H}}{dt} = \frac{\partial\mathcal{H}}{\partial t} = 0, \tag{7.8}$$

so $\mathcal{H}$ is a constant of motion (the energy for natural systems).

Example 7.2 – 1D harmonic oscillator

With $\mathcal{L} = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}kq^2$:

$$p = m\dot{q}, \qquad \mathcal{H} = \frac{p^2}{2m} + \frac{kq^2}{2}.$$

Hamilton's equations $\dot{q} = p/m$, $\dot{p} = -kq$ give $\ddot{q} + \omega^2 q = 0$, $\omega = \sqrt{k/m}$.

Example 7.3 – Charged particle in an electromagnetic field

For $\mathcal{L} = \frac{1}{2}m\dot{\mathbf{r}}^2 + \frac{q}{c}\dot{\mathbf{r}}\cdot\mathbf{A}(\mathbf{r},t) - q\phi(\mathbf{r},t)$,

$$\mathbf{p} = \frac{\partial\mathcal{L}}{\partial\dot{\mathbf{r}}} = m\dot{\mathbf{r}} + \frac{q}{c}\mathbf{A}, \quad \mathcal{H} = \frac{1}{2m}\Big(\mathbf{p} - \frac{q}{c}\mathbf{A}\Big)^2 + q\phi.$$

The canonical momentum $\mathbf{p}$ differs from the mechanical momentum $m\dot{\mathbf{r}}$.

7.3. Poisson Brackets and Phase–Space Flow

Theorem 7.3 – Poisson bracket and time evolution

For any C^1 function $F(q,p,t)$ on phase space,

$$\{F,G\} \equiv \sum_{i=1}^{n}\left(\frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i}\right), \tag{7.9}$$

and Hamilton's equations imply

$$\frac{dF}{dt} = \{F,\mathcal{H}\} + \frac{\partial F}{\partial t}. \tag{7.10}$$

In particular, $\dot{q}_i = \{q_i,\mathcal{H}\},\ \dot{p}_i = \{p_i,\mathcal{H}\}$.

Remark 7.4 – Algebraic properties

The Poisson bracket is bilinear, antisymmetric, satisfies the Leibniz rule, and the Jacobi identity. Canonical coordinates obey $\{q_i,q_j\} = 0,\ \{p_i,p_j\} = 0,\ \{q_i,p_j\} = \delta_{ij}$.

Theorem 7.4 – Liouville's theorem (phase–space volume conservation)

Let $\mathbf{z} = (q_1,\dots,q_n,p_1,\dots,p_n)$ and $\dot{\mathbf{z}} = \mathbf{X}_H(\mathbf{z})$ be the Hamiltonian vector field. The flow is divergence–free:

$$\nabla_{\mathbf{z}}\cdot\mathbf{X}_H = \sum_{i=1}^{n}\left(\frac{\partial\dot{q}_i}{\partial q_i} + \frac{\partial\dot{p}_i}{\partial p_i}\right) = \sum_{i=1}^{n}\left(\frac{\partial^2\mathcal{H}}{\partial p_i\,\partial q_i} - \frac{\partial^2\mathcal{H}}{\partial q_i\,\partial p_i}\right) = 0, \tag{7.11}$$

so any phase–space volume $V = \int_\Omega d^nq\,d^np$ is preserved:

$$\frac{dV}{dt} = 0. \tag{7.12}$$

7.4. Legendre Transforms: Generalities and a Thermodynamic Aside

Example 7.4 – Single-variable Legendre transform

For a smooth $f(x,y)$ with $u \equiv \partial f/\partial x$, define

$$F(u,y) = f(x,y) - ux, \qquad dF = v\,dy - x\,du, \quad v \equiv \frac{\partial f}{\partial y}. \tag{7.13}$$

Then $v = \partial F/\partial y,\ x = -\partial F/\partial u$.

Remark 7.5 – Thermodynamic potentials as Legendre transforms

From the first law $dE = T\,dS - p\,dV$, the Helmholtz free energy $F = E - TS$ obeys

$$dF = -S\,dT - p\,dV, \qquad -S = \frac{\partial F}{\partial T}, \quad -p = \frac{\partial F}{\partial V}, \tag{7.14}$$

and the enthalpy $H = E + pV$ satisfies $dH = T\,dS + V\,dP$. These mirror (7.3)–(7.5) with $(q, \dot{q}) \leftrightarrow (S, V)$ etc.

7.5. Cyclic Coordinates and Central Examples

Theorem 7.5 – Cyclic coordinate

If $\mathcal{L}$ (equivalently $\mathcal{H}$) does not depend on q_k, then p_k is conserved:

$$\frac{\partial \mathcal{H}}{\partial q_k} = 0 \implies \dot{p}_k = 0. \tag{7.15}$$

Example 7.5 – Planar central force

With polar coordinates (r, θ), $\mathcal{L} = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - V(r)$. Then

$$p_r = m\dot{r}, \quad p_\theta = mr^2\dot{\theta}, \quad \mathcal{H} = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + V(r),$$

and θ is cyclic $\Rightarrow$ p_θ is conserved (angular momentum).

Remark 7.6 – Substitution and invertibility

Do not think of the replacement $\dot{q}_i \mapsto p_i$ as a mere rescaling ($p_i = \mu\,\dot{q}_i$). In general,

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = p_i(q, \dot{q}, t),$$

with each p_i depending on *all* velocities. Passing to Hamiltonian form requires inverting the velocity–momentum map $\dot{q} \mapsto p(q, \dot{q}, t)$ to obtain $\dot{q} = \dot{q}(q, p, t)$. This inversion is guaranteed locally only when the Hessian $\mathbf{W}_{ij} = \partial^2\mathcal{L}/\partial\dot{q}_i\partial\dot{q}_j$ is nonsingular (cf. regularity, Eq. (7.6)).

Remark 7.7 – From second order to first order (phase space)

Newton's and Lagrange's equations are second order in time, e.g.

$$\ddot{x} + f(\dot{x}, x, t) = 0. \tag{7.16}$$

Introduce phase–space variables $q \equiv x$ and $v \equiv \dot{x}$. Then (7.16) is equivalent to the first–order system

$$\begin{cases} \dot{q} = v, \\ \dot{v} = -\,f(v, q, t). \end{cases} \tag{7.17}$$

In Hamiltonian mechanics (q, v) is replaced by (q, p), with (q, p) evolving by Hamilton's equations (7.7).

Remark 7.8 – Energy and the Hamiltonian

From Hamilton's equations,

$$\frac{d\mathcal{H}}{dt} = \sum_i \left(\frac{\partial \mathcal{H}}{\partial q_i} \dot{q}_i + \frac{\partial \mathcal{H}}{\partial p_i} \dot{p}_i \right) + \frac{\partial \mathcal{H}}{\partial t} = \sum_i (-\dot{p}_i\, \dot{q}_i + \dot{q}_i\, \dot{p}_i) + \frac{\partial \mathcal{H}}{\partial t} = \frac{\partial \mathcal{H}}{\partial t}. \tag{7.18}$$

Thus, if $\partial\mathcal{H}/\partial t = 0$,

$$\frac{d\mathcal{H}}{dt} = 0 \quad \Longrightarrow \quad \mathcal{H} = E = \text{constant}. \tag{7.19}$$

For a natural (conservative) system with

$$\mathcal{L}(q, \dot{q}) = T(\dot{q}, q) - V(q), \qquad T = \tfrac{1}{2} \sum_{i,j} g_{ij}(q)\, \dot{q}_i \dot{q}_j,$$

the canonical momenta are

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \sum_j g_{ij}(q)\, \dot{q}_j. \tag{7.20}$$

If the kinetic metric $\mathbf{g}(q) = [g_{ij}(q)]$ is invertible, denote its inverse by $\mathbf{g}^{-1}(q) = [g^{ij}(q)]$, i.e.

$$\sum_k g_{ik}(q)\, g^{kj}(q) = \delta_i^{\ j}. \tag{7.21}$$

Then

$$\dot{q}_i = \sum_j g^{ij}(q)\, p_j, \tag{7.22}$$

and substituting into $E = T + V$ gives the Hamiltonian in (q, p):

$$\boxed{\mathcal{H}(q, p) = \tfrac{1}{2} \sum_{i,j} g^{ij}(q)\, p_i p_j + V(q)}. \tag{7.23}$$

This coincides with the Legendre transform definition (Eq. (7.3)) and is time–independent when V is.

7.6. Construction of the Hamiltonian

Given $\mathcal{L}(q, \dot{q}, t) = T(\dot{q}, q) - V(q, t)$:

1. **Choose coordinates.** Pick generalized coordinates q_i and write $T = \frac{1}{2}\, \dot{q}^T \mathbf{g}(q) \dot{q}$, identifying the kinetic metric $\mathbf{g}(q)$ (cf. the mass matrix in the coupled–oscillations chapter).
2. **Define momenta.** Set $p_i = \partial\mathcal{L}/\partial\dot{q}_i = \sum_j g_{ij}(q)\dot{q}_j$ (7.20).
3. **Invert the metric.** Compute $g^{ij}(q)$ from (7.21) and solve for velocities: $\dot{q}_i = \sum_j g^{ij}(q) p_j$ (7.22).
4. **Legendre transform.** Substitute $\dot{q}(q, p, t)$ into

$$\mathcal{H}(q, p, t) = \sum_i p_i \dot{q}_i - \mathcal{L}(q, \dot{q}, t),$$

   or equivalently use (7.24) for natural systems to obtain $\mathcal{H}(q, p, t)$.

Theorem 7.6 – Hamiltonian for natural Lagrangians

For any natural system with positive–definite kinetic metric $\mathbf{g}(q)$,

$$\mathcal{H}(q, p, t) = \tfrac{1}{2}\, p^T \mathbf{g}^{-1}(q)\, p + V(q, t),$$

and (q, p) evolve via Hamilton's equations (7.7). If $\partial V/\partial t = 0$, then $\mathcal{H}$ is a constant of motion.

Remark 7.9 – Caveat: singular Lagrangians

If the Hessian $\mathbf{W} = \partial^2\mathcal{L}/\partial\dot{q}^2$ is singular, the Legendre transform is not invertible. Such constrained systems require Dirac's theory of constraints (primary/secondary constraints and extended Hamiltonians).

7.7. General Form of the Lagrangian and Hamiltonian

We collect a useful class of Lagrangians that are at most quadratic in velocities and may contain a term linear in $\dot{q}$:

$$\boxed{\mathcal{L}(q,\dot{q},t) = \mathcal{L}_0(q,t) \;+\; \dot{q}^T a(q,t) \;+\; \tfrac{1}{2}\,\dot{q}^T\,\mathcal{T}(q,t)\,\dot{q}} \tag{7.24}$$

Here $q \in \mathbb{R}^n$ is a column vector, $\mathcal{T} = \mathcal{T}^T$ is a symmetric (typically positive–definite) matrix field, and $a(q,t)$ is a column vector field. In Cartesian coordinates for a single particle of mass m,

$$\tfrac{1}{2}\,\dot{q}^T\mathcal{T}\dot{q} = \tfrac{1}{2}m(\dot{x}^2+\dot{y}^2+\dot{z}^2), \qquad \dot{q}^T a = \dot{\vec{r}}\cdot\vec{a}. \tag{7.25}$$

The canonical momentum and Hamiltonian follow from the Legendre transform:

$$p \equiv \frac{\partial\mathcal{L}}{\partial\dot{q}} = \mathcal{T}\,\dot{q} + a \quad\Longrightarrow\quad \dot{q} = \mathcal{T}^{-1}(p-a), \tag{7.26}$$

$$\mathcal{H} = p^T\dot{q} - \mathcal{L} = \tfrac{1}{2}\,\dot{q}^T\mathcal{T}\dot{q} - \mathcal{L}_0(q,t) = \tfrac{1}{2}\,(p-a)^T\mathcal{T}^{-1}(p-a) - \mathcal{L}_0(q,t). \tag{7.27}$$

For natural systems, $\mathcal{L}_0 = -V(q,t)$, hence

$$\boxed{\mathcal{H}(q,p,t) = \tfrac{1}{2}\,(p-a)^T\mathcal{T}^{-1}(q,t)\,(p-a) + V(q,t)}. \tag{7.28}$$

Example 7.6 – Minimal Coupling: Charged Particle in Electromagnetic Fields

For a particle of charge q and mass m,

$$\mathcal{L}(\vec{x},\dot{\vec{x}},t) = \tfrac{1}{2}m\,\dot{\vec{x}}^{\,2} + q\,\vec{A}(\vec{x},t)\cdot\dot{\vec{x}} - q\,\Phi(\vec{x},t), \tag{7.29}$$

so $a = q\vec{A}$, $\mathcal{T} = m\,\mathbf{1}$. Then

$$\vec{p} = \frac{\partial\mathcal{L}}{\partial\dot{\vec{x}}} = m\dot{\vec{x}} + q\vec{A}, \qquad \dot{\vec{x}} = \frac{1}{m}\big(\vec{p} - q\vec{A}\big),$$

and

$$\boxed{\mathcal{H}(\vec{x},\vec{p},t) = \frac{\big(\vec{p} - q\vec{A}(\vec{x},t)\big)^2}{2m} + q\,\Phi(\vec{x},t)}. \tag{7.30}$$

Remark 7.10 – Canonical vs. kinetic momentum; gauge

In EM, the canonical momentum $\vec{p}$ differs from the kinetic momentum $m\vec{v}$ by $q\vec{A}$: $m\vec{v} = \vec{p} - q\vec{A}$. Under the gauge change $\vec{A}\to\vec{A}+\nabla\chi$, $\Phi\to\Phi-\partial_t\chi$, $\mathcal{L}$ shifts by a total time derivative $q\,\frac{d\chi}{dt}$ (no effect on equations of motion), and $\mathcal{H}$ retains the form (7.30).

7.8. Geometrical Aspect of the Hamiltonian Approach

The kinetic energy of a particle moving in a configuration space with metric $g_{ij}(q)$ is

$$ds^2 = g_{ij}(q)\,dq^i dq^j, \qquad T = \tfrac{1}{2}\,m\,g_{ij}(q)\,\dot{q}^i\dot{q}^j. \tag{7.31}$$

Thus a natural Lagrangian is

$$\mathcal{L}(q, \dot{q}, t) = \tfrac{1}{2}\, m\, g_{ij}(q)\, \dot{q}^{\,i} \dot{q}^{\,j} - V(q, t). \tag{7.32}$$

The canonical momenta are

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}^{\,i}} = m\, g_{ij}(q)\, \dot{q}^{\,j}, \qquad \dot{q}^{\,i} = \frac{1}{m}\, g^{ij}(q)\, p_j, \tag{7.33}$$

with g^{ij} the inverse metric. Hence

$$\boxed{\mathcal{H}(q, p, t) = \tfrac{1}{2m}\, g^{ij}(q)\, p_i p_j + V(q, t)}. \tag{7.34}$$

Example 7.7 – Free Particle in Plane Polar Coordinates

Let $q = (r, \theta)$, $ds^2 = dr^2 + r^2 d\theta^2 \Rightarrow g_{ij} = \mathrm{diag}(1, r^2)$. Then $p_r = m\dot{r}$, $p_\theta = mr^2\dot{\theta}$, and

$$\mathcal{H}(r, \theta; p_r, p_\theta) = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + V(r, \theta).$$

For a central potential $V = V(r)$, θ is cyclic, so p_θ is conserved.

7.9. Liouville's Theorem

Theorem 7.7 – Liouville

Let (q_i, p_i) evolve under Hamilton's equations $\dot{q}_i = \partial\mathcal{H}/\partial p_i$, $\dot{p}_i = -\partial\mathcal{H}/\partial q_i$. Then the Hamiltonian phase–flow is incompressible:

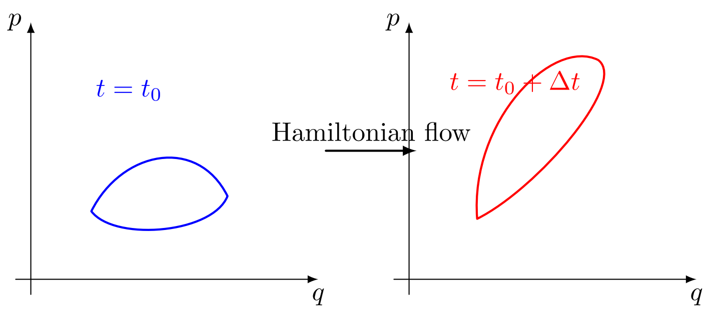
$$\sum_{i=1}^{n} \left(\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right) = 0, \qquad \det\!\left(\frac{\partial(q(t), p(t))}{\partial(q(0), p(0))} \right) = 1. \tag{7.35}$$

Equivalently, if $\rho(q, p, t)$ is the density of representative points in phase space, it satisfies the continuity equation

$$\boxed{\frac{\partial \rho}{\partial t} + \sum_{i=1}^{n} \frac{\partial}{\partial q_i}(\rho\, \dot{q}_i) + \sum_{i=1}^{n} \frac{\partial}{\partial p_i}(\rho\, \dot{p}_i) = 0}. \tag{7.36}$$

Since the divergence of the Hamiltonian vector field vanishes, the material derivative along trajectories obeys $\frac{d\rho}{dt} = \partial_t \rho + \sum_i (\dot{q}_i \partial_{q_i} \rho + \dot{p}_i \partial_{p_i} \rho) = 0$.

Using Hamilton's equations, $\partial\dot{q}_i/\partial q_i = \partial^2\mathcal{H}/\partial q_i \partial p_i$ and $\partial\dot{p}_i/\partial p_i = -\partial^2\mathcal{H}/\partial p_i \partial q_i$, which cancel pairwise in the sum. The Jacobian statement follows by integrating the divergence–free condition.



Phase–space volume is preserved while shapes deform (Liouville).

7.10. Routh's Procedure

When some coordinates are cyclic, it is economical to perform a *partial* Legendre transform—Hamiltonian treatment for the cyclic ones, Lagrangian for the rest. Let the configuration split as

$$q = (q_1, \dots, q_s;\ q_{s+1}, \dots, q_n) \equiv (q_a;\ q_\alpha),$$

with q_α $(\alpha = s+1, \dots, n)$ cyclic in $\mathcal{L}(q, \dot{q}, t)$. Define the conjugate momenta $p_\alpha \equiv \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}$ and *eliminate* $\dot{q}_\alpha$ in favor of p_α. The **Routhian** is the partial Legendre transform

$$\boxed{R\big(q_a, \dot{q}_a;\ p_\alpha;\ t\big) = \sum_{\alpha=s+1}^{n} p_\alpha\, \dot{q}_\alpha \ -\ \mathcal{L}(q_a, q_\alpha;\ \dot{q}_a, \dot{q}_\alpha;\ t) \quad \text{with} \quad \dot{q}_\alpha = \dot{q}_\alpha(q, p, t).} \tag{7.37}$$

It obeys mixed Hamilton–Lagrange equations:

$$\text{(Hamilton for cyclic)} \qquad \frac{\partial R}{\partial q_\alpha} = -\dot{p}_\alpha, \qquad \frac{\partial R}{\partial p_\alpha} = \dot{q}_\alpha, \tag{7.38}$$

$$\text{(Lagrange for noncyclic)} \qquad \frac{d}{dt}\left(\frac{\partial R}{\partial \dot{q}_a}\right) - \frac{\partial R}{\partial q_a} = 0. \tag{7.39}$$

For genuinely cyclic q_α, $\partial \mathcal{L}/\partial q_\alpha = 0 \Rightarrow \dot{p}_\alpha = 0$ by (7.38), i.e. each p_α is a constant of motion.

Remark 7.11 – Partial Legendre transform

Equation (7.37) is a *partial* transform (over α-indices only). The identities (7.38)–(7.39) hold *after* expressing $\dot{q}_\alpha$ in terms of (q, p, t) using $p_\alpha = \partial \mathcal{L}/\partial \dot{q}_\alpha$. This requires the block $\partial^2 \mathcal{L}/\partial \dot{q}_\alpha \partial \dot{q}_\beta$ to be invertible.

Example 7.8 – Central Force with One Cyclic Angle

Take plane polar coordinates (r, θ) for a particle of mass m in a central potential $V(r)$:

$$\mathcal{L}(r, \theta; \dot{r}, \dot{\theta}) = \tfrac{1}{2} m\big(\dot{r}^{\,2} + r^2 \dot{\theta}^{\,2}\big) - V(r).$$

The angle θ is cyclic; its momentum $p_\theta = \partial \mathcal{L}/\partial \dot{\theta} = m r^2 \dot{\theta} \equiv \ell$ is constant. Eliminating $\dot{\theta} = \ell/(m r^2)$ in (7.37) gives the Routhian

$$\boxed{R(r, \dot{r}; \ell) = \ell\, \dot{\theta} - \mathcal{L} = -\tfrac{1}{2} m\, \dot{r}^{\,2} + \tfrac{\ell^2}{2 m r^2} + V(r).} \tag{7.40}$$

The Euler–Lagrange equation (7.39) for r yields

$$\frac{d}{dt}\big(-m\dot{r}\big) - \left(-\frac{\ell^2}{m r^3} + \frac{dV}{dr}\right) = 0 \implies m\,\ddot{r} = \frac{\ell^2}{m r^3} - \frac{dV}{dr}, \tag{7.41}$$

which is exactly the radial equation obtained from the full Lagrange equations. *Special case:* $V(r) = \kappa/r^2 \Rightarrow \frac{dV}{dr} = -2\kappa/r^3$, so $m\ddot{r} = \ell^2/(m r^3) + 2\kappa/r^3$.

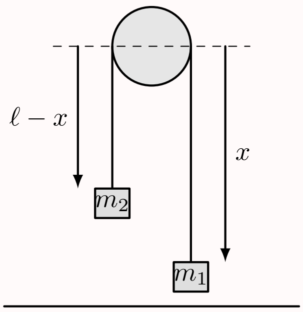
Theorem 7.8 – Can all coordinates be made cyclic? (Foreshadowing HJ)

Not in general. If a Hamiltonian system with n degrees of freedom is *Liouville–integrable* (admits n independent, globally defined integrals in involution), then there exists a canonical transformation to *action–angle* variables (α_k, φ_k) such that

$$\mathcal{H} = \mathcal{H}(\alpha_1, \dots, \alpha_n), \qquad \dot{\alpha}_k = 0, \quad \dot{\varphi}_k = \omega_k(\alpha),$$

so the φ_k are cyclic and the solution is $\varphi_k(t) = \varphi_k(0) + \omega_k t$. Hamilton–Jacobi theory (next chapter) constructs these variables by quadrature when integrability holds.

Example 7.9 – Atwood Machine (One Degree of Freedom



Let x be the downward displacement of m_1 from the pulley; the inextensible string imposes $x_2 = \ell - x$ for mass m_2. With downward positive, the gravitational potential is $U = -m_1 g x - m_2 g(\ell - x) = -(m_1 - m_2) g\, x - m_2 g\ell$. Discarding the additive constant, the one–coordinate Lagrangian is

$$\mathcal{L} = \tfrac{1}{2}(m_1 + m_2)\,\dot{x}^{\,2} + (m_1 - m_2)\,g\,x.$$

The canonical momentum $p = \partial\mathcal{L}/\partial\dot{x} = (m_1 + m_2)\dot{x}$, and the Hamiltonian (energy) is

$$\boxed{\mathcal{H}(x,p) = \frac{p^2}{2(m_1 + m_2)} - (m_1 - m_2)\,g\,x}. \tag{7.42}$$

Hamilton's equations give $\dot{x} = \partial\mathcal{H}/\partial p = p/(m_1 + m_2)$ and $\dot{p} = -\partial\mathcal{H}/\partial x = (m_1 - m_2)g$, hence

$$\ddot{x} = \frac{\dot{p}}{m_1 + m_2} = \frac{m_1 - m_2}{m_1 + m_2}\,g,$$

in agreement with the standard result.

Example 7.10 – Point Particle on a Cone

A point mass m moves on a right circular cone with opening parameter

$$\frac{r}{z} = \tan\alpha \equiv C, \qquad z > 0.$$

The holonomic constraint gives

$$r = Cz, \qquad \dot{r} = C\,\dot{z}.$$

Lagrangian. With cylindrical coordinates (r, ϕ, z) and gravity along $+z$,

$$T = \tfrac{1}{2}m\big(\dot{r}^{\,2} + r^2\dot{\phi}^{\,2} + \dot{z}^{\,2}\big) = \tfrac{1}{2}m\Big[(1 + C^2)\dot{z}^{\,2} + C^2 z^2\dot{\phi}^{\,2}\Big], \qquad V = mgz,$$

so

$$\mathcal{L} = \tfrac{1}{2}m\Big[(1 + C^2)\dot{z}^{\,2} + C^2 z^2\dot{\phi}^{\,2}\Big] - mgz. \tag{7.43}$$

Canonical momenta.

$$p_z = \frac{\partial\mathcal{L}}{\partial\dot{z}} = m(1 + C^2)\dot{z}, \qquad p_\phi = \frac{\partial\mathcal{L}}{\partial\dot{\phi}} = mC^2 z^2\dot{\phi}. \tag{7.44}$$

The angle ϕ is cyclic $\Rightarrow \dot{p}_\phi = 0$; we may write $p_\phi = \ell$ (constant).
Hamiltonian. Using $\mathcal{H} = \dot{z}\,p_z + \dot{\phi}\,p_\phi - \mathcal{L}$ and (7.44),

$$\boxed{\mathcal{H}(z, \phi; p_z, p_\phi) = \frac{p_z^2}{2m(1 + C^2)} + \frac{p_\phi^2}{2mC^2 z^2} + mgz.} \tag{7.45}$$

Hamilton's equations.

$$\dot{z} = \frac{\partial\mathcal{H}}{\partial p_z} = \frac{p_z}{m(1 + C^2)}, \qquad \dot{p}_z = -\frac{\partial\mathcal{H}}{\partial z} = \frac{p_\phi^2}{mC^2 z^3} - mg,$$

$$\dot{\phi} = \frac{\partial \mathcal{H}}{\partial p_\phi} = \frac{p_\phi}{mC^2 z^2}, \qquad \dot{p}_\phi = -\frac{\partial \mathcal{H}}{\partial \phi} = 0.$$

Effective one–dimensional motion in z. With $p_\phi = \ell$ fixed,

$$U_{\text{eff}}(z) = \frac{\ell^2}{2mC^2 z^2} + mgz, \qquad m_{\text{eff}}(z\text{-kinetic}) = m(1 + C^2). \tag{7.46}$$

The equilibrium z_0 satisfies $U'_{\text{eff}}(z_0) = 0$:

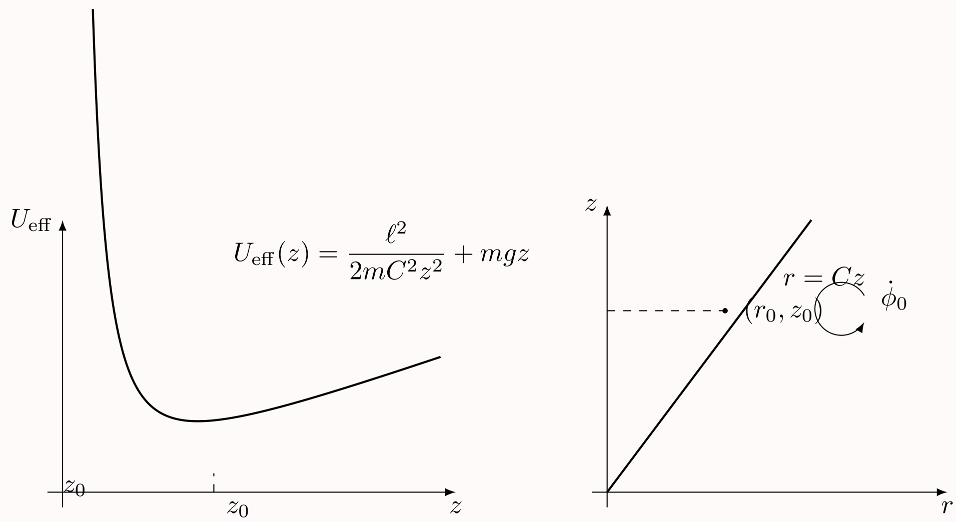
$$-\frac{\ell^2}{mC^2 z_0^3} + mg = 0 \implies \boxed{z_0 = \left(\frac{\ell^2}{m^2 g\, C^2}\right)^{1/3}.} \tag{7.47}$$

Equivalently, $\ell = \pm m\sqrt{g\, C^2\, z_0^3}$.
Circular motion at the minimum. At $z = z_0$ the height is constant and

$$\dot{\phi}_0 = \frac{\ell}{mC^2 z_0^2} = \sqrt{\frac{g}{C^2 z_0}}, \qquad r_0 = C z_0,$$

i.e. uniform rotation around the cone at fixed z_0.
Small vertical oscillations about z_0. Since $U''_{\text{eff}}(z) = \frac{3\ell^2}{mC^2 z^4}$, the small–oscillation frequency is

$$\omega_z^2 = \frac{U''_{\text{eff}}(z_0)}{m(1 + C^2)} = \frac{3g}{(1 + C^2)\, z_0}. \tag{7.48}$$



Left: Effective potential with a minimum at z_0. *Right:* Cone cross–section $r = Cz$; at z_0 the particle circles the axis with angular speed $\dot{\phi}_0$.

Chapter 8

Canonical Transformations

Remark 8.1 – Notation and the p vs. P Convention

Throughout this chapter (q_i, p_i) denote the *old* canonical coordinates and momenta; after a transformation we write (Q_i, P_i) for the *new* canonical pair. Thus P_i are not constants by notation—they are just the new canonical momenta. When we perform a *point transformation* (e.g. to polar coordinates) but do not introduce a new canonical set, we continue to use lower-case symbols, e.g. p_r, p_θ.

The advantage of the Hamiltonian formalism lies not only in computation but in reorganizing a problem by a good choice of variables. In central-force motion, a point transformation to polar coordinates exposes a cyclic coordinate and simplifies the dynamics.
If the generalized coordinates q_i are cyclic, then

$$\dot{p}_i \;=\; -\frac{\partial \mathcal{H}}{\partial q_i} \;=\; 0 \;\;\Longrightarrow\;\; p_i = \text{constant} \equiv \alpha_i, \tag{8.1}$$

and the Hamiltonian reduces to

$$\mathcal{H} = \mathcal{H}(\alpha_1, \alpha_2, \ldots, \alpha_n). \tag{8.2}$$

Consequently

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i} \equiv \omega_i(\boldsymbol{\alpha}) = \text{constant} \quad\Longrightarrow\quad q_i(t) = \omega_i\, t + \delta_i. \tag{8.3}$$

Example 8.1 – Point Transformation for a Central Potential

Cartesian to polar:

$$x = r\cos\theta, \qquad y = r\sin\theta.$$

With a central potential $V(r)$,

$$\mathcal{L}(x, y, \dot{x}, \dot{y}) = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) - V\big(\sqrt{x^2 + y^2}\big), \tag{8.4}$$

becomes

$$\mathcal{L}(r, \theta, \dot{r}, \dot{\theta}) = \tfrac{1}{2}m\big(\dot{r}^2 + r^2\dot{\theta}^2\big) - V(r). \tag{8.5}$$

Here θ is cyclic, so $p_\theta = \partial\mathcal{L}/\partial\dot{\theta} = mr^2\dot{\theta}$ is conserved and the radial motion follows from an effective one-dimensional problem at fixed p_θ.

Remark 8.2 – On “Reduction” of Degrees of Freedom

The configuration space remains two-dimensional (r, θ), but once p_θ is fixed by symmetry, the equations reduce to a single nontrivial radial equation—an *effective* one-degree-of-freedom dynamics.

More generally, a point transformation on configuration space is

$$Q_i = Q_i(q,t). \tag{8.6}$$

In the Hamiltonian setting we will consider phase–space transformations from one canonical set (q,p) to another (Q,P):

$$Q_i = Q_i(q,p,t), \qquad P_i = P_i(q,p,t). \tag{8.7}$$

8.1. Canonical Transformations

A transformation $(q,p) \mapsto (Q,P)$ is **canonical** if there exists a (possibly new) Hamiltonian $K(Q,P,t)$ such that Hamilton's equations retain their form,

$$\dot{Q}_i = \frac{\partial K}{\partial P_i}, \qquad \dot{P}_i = -\frac{\partial K}{\partial Q_i}. \tag{8.8}$$

8.1.1. Requirements for Canonicity

Variational Approach

The old variables obey Hamilton's variational principle

$$\delta \int \Big(\sum_i p_i \dot{q}_i - \mathcal{H}(q,p,t) \Big)\, dt = 0, \tag{8.9}$$

and the new variables must satisfy the analogous principle

$$\delta \int \Big(\sum_i P_i \dot{Q}_i - K(Q,P,t) \Big)\, dt = 0. \tag{8.10}$$

The two actions differ by a total time derivative:

$$\sum_i p_i \dot{q}_i - \mathcal{H} \;=\; \sum_i P_i \dot{Q}_i - K \;+\; \frac{dF}{dt}, \tag{8.11}$$

for some auxiliary function F (its detailed use will be taken up next).

Remark 8.3 – Invariance Under Total Derivatives and Scalings

Adding a total time derivative to the Lagrangian one–form or multiplying the entire action by a nonzero constant does not change the equations of motion. Relation (8.11) is the precise statement of this invariance in the canonical setting.

Generating Functions

With $\lambda = 1$ the equality of canonical one–forms up to a total derivative is

$$\sum_i p_i \dot{q}_i \;=\; \sum_i P_i \dot{Q}_i \;-\; K \;+\; \frac{dF}{dt}, \tag{8.12}$$

which implies

$$dF \;=\; \sum_i p_i \, dq_i \;-\; \sum_i P_i \, dQ_i \;+\; \big(K - \mathcal{H}\big)\, dt. \tag{8.13}$$

Remark 8.4 – Endpoint conditions and total time derivatives

The variational principles for (q,p) and (Q,P) differ by a boundary term. Equality of the equations of motion requires the actions to differ by a total derivative dF/dt; endpoint variations are treated in the usual way (fixed endpoints or vanishing variations).

Remark 8.5 – Old vs. new variables

We reserve lower–case p_i for the *old* momenta and upper–case P_i for the *new* momenta throughout. Likewise q_i vs. Q_i.

Type $F_1(q, Q, t)$. Treat $F = F_1(q, Q, t)$ as a function of (q, Q, t). Comparing differentials in (8.13) gives

$$p_i = \frac{\partial F_1}{\partial q_i}, \qquad P_i = -\frac{\partial F_1}{\partial Q_i}, \qquad K = \mathcal{H} + \frac{\partial F_1}{\partial t}. \tag{8.14}$$

Local invertibility. The map $(q, Q) \mapsto (p, P)$ defined by (8.14) is locally invertible if the mixed Hessian $\left[\partial^2 F_1/\partial q_i \partial Q_j\right]$ is nonsingular.

Type $F_2(q, P, t)$ (Legendre transform in Q). Use $\sum_i P_i dQ_i = d(\sum_i P_i Q_i) - \sum_i Q_i dP_i$ in (8.13):

$$d\left(F_1 + \sum_i P_i Q_i\right) = \sum_i p_i\, dq_i \;-\; \sum_i Q_i\, dP_i \;+\; (K - \mathcal{H})\, dt.$$

Define

$$F_2(q, P, t) \;\equiv\; F_1(q, Q, t) + \sum_i P_i Q_i, \tag{8.15}$$

so that matching coefficients yields

$$p_i = \frac{\partial F_2}{\partial q_i}, \qquad Q_i = \frac{\partial F_2}{\partial P_i}, \qquad K = \mathcal{H} + \frac{\partial F_2}{\partial t}. \tag{8.16}$$

Local invertibility. The matrix $\left[\partial^2 F_2/\partial q_i \partial P_j\right]$ must be nonsingular.

Type $F_3(p, Q, t)$ (Legendre transform in q). Use $\sum_i p_i dq_i = d(\sum_i p_i q_i) - \sum_i q_i dp_i$ in (8.13):

$$d\left(F_1 - \sum_i p_i q_i\right) = -\sum_i q_i\, dp_i \;-\; \sum_i P_i\, dQ_i \;+\; (K - \mathcal{H})\, dt.$$

Define

$$F_3(p, Q, t) \;\equiv\; F_1(q, Q, t) - \sum_i p_i q_i, \tag{8.17}$$

to obtain

$$q_i = -\frac{\partial F_3}{\partial p_i}, \qquad P_i = -\frac{\partial F_3}{\partial Q_i}, \qquad K = \mathcal{H} + \frac{\partial F_3}{\partial t}. \tag{8.18}$$

Type $F_4(p, P, t)$ (Legendre in q and Q). Combining the two identities gives

$$d\left(F_1 - \sum_i p_i q_i + \sum_i P_i Q_i\right) = -\sum_i q_i\, dp_i \;+\; \sum_i Q_i\, dP_i \;+\; (K - \mathcal{H})\, dt.$$

Define

$$F_4(p, P, t) \;\equiv\; F_1(q, Q, t) - \sum_i p_i q_i + \sum_i P_i Q_i, \tag{8.19}$$

leading to

$$q_i = -\frac{\partial F_4}{\partial p_i}, \qquad Q_i = \frac{\partial F_4}{\partial P_i}, \qquad K = \mathcal{H} + \frac{\partial F_4}{\partial t}. \tag{8.20}$$

Theorem 8.1 – Generating Function ¿ Canonicity

Any transformation defined by one of $\{F_1, F_2, F_3, F_4\}$ via Eqs. (8.14), (8.16), (8.18), (8.20) is canonical. Moreover, time–dependence of F shifts the Hamiltonian as $K = \mathcal{H} + \partial F/\partial t$.

Remark 8.6 – Regularity (no hidden singularities)

The mixed Hessian required to solve for the new variables (e.g. $\partial^2 F_2/\partial q\,\partial P$ for F_2) must be invertible. At points where it is singular the transformation is not locally canonical/invertible.

Explicit Canonical Transformations

Example 8.2 – Identity (Type F_2)

Let $F_2(q,P) = \sum_i q_i P_i$. Then

$$p_i = \frac{\partial F_2}{\partial q_i} = P_i, \qquad Q_i = \frac{\partial F_2}{\partial P_i} = q_i, \qquad K = \mathcal{H}.$$

Thus $(Q,P) = (q,p)$.

Example 8.3 – Point Transformation Q

Take $F_2(q,P,t) = \sum_i f_i(q,t)P_i + g(q,t)$. Then

$$Q_i = \frac{\partial F_2}{\partial P_i} = f_i(q,t), \qquad p_j = \frac{\partial F_2}{\partial q_j} = \sum_i \frac{\partial f_i}{\partial q_j} P_i + \frac{\partial g}{\partial q_j},$$

and $K = \mathcal{H} + \partial F_2/\partial t$. Hence every point transformation is canonical; g produces the familiar momentum shift.

Example 8.4 – Space Translation (Type F_2)

In one d.o.f., set $F_2(q,P) = (q+a)P$. Then

$$Q = \frac{\partial F_2}{\partial P} = q + a, \qquad p = \frac{\partial F_2}{\partial q} = P, \qquad K = \mathcal{H}.$$

This is a canonical shift $q \mapsto q + a$ with unchanged momentum.

Example 8.5 – Momentum Shift (Gauge–like)

Let $F_2(q,P,t) = qP + g(q,t)$. Then $Q = q$ and

$$p = \frac{\partial F_2}{\partial q} = P + \frac{\partial g}{\partial q}, \qquad K = \mathcal{H} + \frac{\partial g}{\partial t}.$$

This is the standard momentum shift (e.g. minimal coupling in EM appears this way under gauge change).

Example 8.6 – Phase–Space Rotation by pi/2

In one d.o.f., take $F_4(p,P) = pP$. Then

$$q = -\frac{\partial F_4}{\partial p} = -P, \qquad Q = \frac{\partial F_4}{\partial P} = p, \qquad K = \mathcal{H}.$$

Thus $(Q,P) = (p,-q)$: a 90° symplectic rotation.

Example 8.7 – Symplectic Scaling

Let $F_2(q,P) = \frac{1}{\lambda}\sum_i q_i P_i$ with $\lambda \neq 0$. Then

$$Q_i = \frac{q_i}{\lambda}, \qquad p_i = \frac{P_i}{\lambda}, \qquad \Rightarrow \quad (q,p) = (\lambda Q,\, P/\lambda).$$

The phase–space area $dq \wedge dp$ is preserved: $dq \wedge dp = d(\lambda Q) \wedge d(P/\lambda) = dQ \wedge dP$.

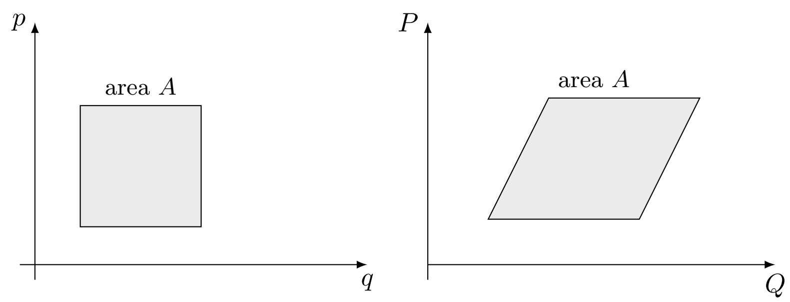


Figure 8.1: Area preservation under a canonical map: the shape may distort, but symplectic area is invariant.

8.1.2. Symplectic (Jacobian) Criterion, Worked Out

Stack variables

$$\eta = \begin{pmatrix} q \\ p \end{pmatrix}, \qquad \xi = \begin{pmatrix} Q \\ P \end{pmatrix}, \qquad J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}.$$

Hamilton's equations are $\dot\eta = J\,\partial\mathcal{H}/\partial\eta$. For a time–independent change $\xi = \xi(\eta)$ with Jacobian

$$M \;\equiv\; \frac{\partial \xi}{\partial \eta} = \begin{pmatrix} \frac{\partial Q}{\partial q} & \frac{\partial Q}{\partial p} \\ \frac{\partial P}{\partial q} & \frac{\partial P}{\partial p} \end{pmatrix},$$

the transformation is canonical iff

$$M^T J\, M \;=\; J. \tag{8.21}$$

Theorem 8.2 – is Symplectic

For a transformation defined by $F_2(q,P,t)$ via (8.16), its Jacobian satisfies (8.21).

Proof (block form). Write

$$Q_q = \frac{\partial^2 F_2}{\partial q\, \partial P}, \quad Q_P = \frac{\partial^2 F_2}{\partial P\, \partial P}, \quad p_q = \frac{\partial^2 F_2}{\partial q\, \partial q}, \quad p_P = \frac{\partial^2 F_2}{\partial P\, \partial q} = Q_q^T.$$

Then

$$M = \begin{pmatrix} Q_q & Q_P \frac{\partial P}{\partial p} \\ p_q & p_P \frac{\partial P}{\partial p} \end{pmatrix}.$$

Using the implicit relation $Q = \partial F_2/\partial P$ one finds $\partial P/\partial p = (Q_q)^{-1}$ and the blocks simplify to

$$M = \begin{pmatrix} Q_q & Q_P\,(Q_q)^{-1} \\ p_q & Q_q^{T\,-1}(p_q^T Q_P - p_q Q_q^T)(Q_q)^{-1} \end{pmatrix}.$$

A direct (but routine) block multiplication shows $M^T J M = J$ thanks to equality of mixed partials. □

Remark 8.7 – Time–dependent case

For time–dependent generating functions the instantaneous Jacobian at fixed t still satisfies (8.21). The only change is the shift $K = \mathcal{H} + \partial F/\partial t$ already obtained in (8.14), (8.16), (8.18), (8.20).

Example 8.8 – One–form (Poincaré–Cartan) check

For $F_4(p, P) = pP$ we found $(Q, P) = (p, -q)$. Then

$$\sum_i p_i dq_i \;=\; \sum_i Q_i\, d(-P_i) \;=\; \sum_i P_i\, dQ_i \;-\; d\Bigl(\sum_i P_i Q_i\Bigr),$$

so the canonical one–forms differ by a total derivative as required by (8.12).

Example 8.9 – Area preservation in 1D

Any map generated by $F_2(q, P)$ preserves $dq \wedge dp$:

$$dq \wedge dp \;=\; dQ \wedge dP \qquad \Longleftrightarrow \qquad \det\begin{pmatrix} Q_q & Q_p \\ P_q & P_p \end{pmatrix} = 1,$$

which is equivalent to (8.21) in one degree of freedom.

8.2. Poisson Brackets and Other Canonical Invariants

We define the Poisson bracket of two functions $U(q, p)$ and $V(q, p)$ by

$$\{U, V\}_{q,p} = \sum_{i=1}^{n} \left(\frac{\partial U}{\partial q_i}\frac{\partial V}{\partial p_i} - \frac{\partial U}{\partial p_i}\frac{\partial V}{\partial q_i} \right). \tag{8.22}$$

Remark 8.8 – Matrix form and compact notation

Let $\eta = \begin{pmatrix} q \\ p \end{pmatrix} \in \mathbb{R}^{2n}$ and $J = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix}$ (the standard symplectic matrix). Then

$$\{U, V\}_{q,p} = (\nabla_\eta U)^T\, J\, (\nabla_\eta V). \tag{8.23}$$

We will often omit the subscript $\{\,\cdot\,,\,\cdot\,\}_{q,p}$ when the canonical pair is clear from context.

Fundamental brackets

From (8.22) it follows immediately that

$$\{q_i, q_j\} = 0, \qquad \{p_i, p_j\} = 0, \qquad \{q_i, p_j\} = \delta_{ij}. \tag{8.24}$$

Introduce the matrix of brackets of the canonical coordinates,

$$\{\eta, \eta\}_{lm} = \{\eta_l, \eta_m\}, \qquad \eta = \begin{pmatrix} q \\ p \end{pmatrix}.$$

Then (8.24) is summarized by

$$\{\eta, \eta\}_\eta = J. \tag{8.25}$$

Time evolution and constants of motion

For any (possibly time–dependent) observable $F(q, p, t)$,

$$\frac{dF}{dt} = \{F, \mathcal{H}\} + \frac{\partial F}{\partial t}. \tag{8.26}$$

In particular, F is a constant of motion iff $\{F, \mathcal{H}\} + \partial F/\partial t = 0$. Hamilton's equations are the special cases $F = q_i$ and $F = p_i$:

$$\dot{q}_i = \{q_i, \mathcal{H}\} = \frac{\partial \mathcal{H}}{\partial p_i}, \qquad \dot{p}_i = \{p_i, \mathcal{H}\} = -\frac{\partial \mathcal{H}}{\partial q_i}.$$

Algebraic properties

For smooth U, V, W and scalars a, b:

$$\{U, V\} = -\{V, U\} \quad \text{(antisymmetry)}, \tag{8.27}$$
$$\{aU + bV, W\} = a\{U, W\} + b\{V, W\} \quad \text{(linearity)}, \tag{8.28}$$
$$\{UV, W\} = U\{V, W\} + V\{U, W\} \quad \text{(Leibniz rule)}, \tag{8.29}$$
$$\{U, \{V, W\}\} + \{V, \{W, U\}\} + \{W, \{U, V\}\} = 0 \quad \text{(Jacobi identity)}. \tag{8.30}$$

The Jacobi identity is equivalent to the statement that the Poisson bracket defines a Lie algebra structure on observables.

Canonicity in bracket form

Let a change of variables $\eta \mapsto \xi$ with $\xi = \begin{pmatrix} Q \\ P \end{pmatrix}$ have Jacobian

$$M = \frac{\partial \xi}{\partial \eta} \in \mathbb{R}^{2n \times 2n}.$$

Using (8.23) one finds

$$\{\xi, \xi\}_\eta = M\, J\, M^T. \tag{8.31}$$

Hence the following are equivalent:

$$\text{(i) } \{\xi, \xi\}_\eta = J \iff \text{(ii) } M\, J\, M^T = J \iff \text{(iii) } M^T J\, M = J. \tag{8.32}$$

Any such transformation is called *canonical* (or *symplectic*). In that case, for all observables U, V,

$$\{U, V\}_{q,p} = \{U, V\}_{Q,P}, \tag{8.33}$$

i.e. the Poisson bracket is form–invariant.

Remark 8.9 – Lowercase p versus uppercase P

Throughout: p denotes the *old* canonical momentum (paired with q); P denotes the *new* canonical momentum (paired with Q) after a canonical transformation. Formulas comparing the two (e.g. $\{q, p\}$ versus $\{Q, P\}$) always respect this distinction.

Phase–space volume and Jacobian

With $d\eta = dq\, dp$ and $d\xi = dQ\, dP$,

$$\iint dq\, dp = \iint \left| \frac{\partial(q, p)}{\partial(Q, P)} \right| dQ\, dP = \iint |\det M|^{-1}\, dQ\, dP. \tag{8.34}$$

From $M^T J M = J$ we obtain $(\det M)^2 = 1 \Rightarrow |\det M| = 1$. For transformations connected continuously to the identity, $\det M = 1$. Thus the canonical measure $dq\, dp$ is invariant (Liouville measure).

Theorem 8.3 – Canonicity criteria (equivalent forms)

For a differentiable change of variables $\eta \mapsto \xi$, the following are equivalent:

1. $\{Q_i, Q_j\} = 0$, $\{P_i, P_j\} = 0$, $\{Q_i, P_j\} = \delta_{ij}$.
2. $M^T J\, M = J$.

3. There exists a generating function F (of any type) yielding (Q, P) from (q, p) as in the canonical formalism (see previous section).
4. The 2–form $\sum_i dP_i \wedge dQ_i$ equals $\sum_i dp_i \wedge dq_i$.

Other canonical invariants

Theorem 8.4 – Poincaré integral invariants

Let γ be a closed curve in phase space and Σ a surface with $\partial\Sigma = \gamma$. Then the following quantities are invariant under canonical transformations and under Hamiltonian time evolution:

$$I_1(\gamma) = \oint_\gamma p \cdot dq, \tag{8.35}$$

$$I_2(\Sigma) = \iint_\Sigma \sum_{i=1}^{n} dp_i \wedge dq_i. \tag{8.36}$$

Remark 8.10 – Hamiltonian flow is canonical

The time evolution map $\Phi_t : (q(0), p(0)) \mapsto (q(t), p(t))$ generated by $\mathcal{H}$ satisfies $D\Phi_t^T J\, D\Phi_t = J$. Hence time evolution preserves Poisson brackets and phase–space volume (Liouville's theorem, proved earlier in this chapter).

Worked checks and canonical tests

Example 8.10 – Free particle and brackets with $\mathcal{H}$

For $\mathcal{H} = p^2/2m$ in 1D,

$$\dot{q} = \{q, \mathcal{H}\} = \frac{\partial \mathcal{H}}{\partial p} = \frac{p}{m}, \qquad \dot{p} = \{p, \mathcal{H}\} = -\frac{\partial \mathcal{H}}{\partial q} = 0,$$

and for any $F(q, p)$, $\dot{F} = \{F, \mathcal{H}\}$. For instance, $\{q^2, \mathcal{H}\} = 2q\,\dot{q}$, consistent with (8.29).

P

Example 8.11 – Scaling map: Q

The Jacobian is $M = \begin{pmatrix} a\mathbf{1} & 0 \\ 0 & \frac{1}{a}\mathbf{1} \end{pmatrix}$. Then

$$M^T J\, M = \begin{pmatrix} a\mathbf{1} & 0 \\ 0 & \frac{1}{a}\mathbf{1} \end{pmatrix} \begin{pmatrix} 0 & \mathbf{1} \\ -\mathbf{1} & 0 \end{pmatrix} \begin{pmatrix} a\mathbf{1} & 0 \\ 0 & \frac{1}{a}\mathbf{1} \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{1} \\ -\mathbf{1} & 0 \end{pmatrix} = J.$$

Hence the map is canonical and $\{Q_i, P_j\} = \delta_{ij}$.

Example 8.12 – Translation in momentum by a function of q: Q

Compute $\{Q_i, Q_j\} = 0$. Next, $\{Q_i, P_j\} = \{q_i, p_j + f_j(q)\} = \delta_{ij} + \{q_i, f_j(q)\} = \delta_{ij}$. Also, $\{P_i, P_j\} = \{p_i + f_i(q), p_j + f_j(q)\} = \partial f_j/\partial q_i - \partial f_i/\partial q_j = 0$ iff $\partial f_j/\partial q_i = \partial f_i/\partial q_j$, i.e. $f = \nabla_q g$ is a gradient. This is the canonical "gauge shift" generated by $F_2(q, P, t) = q \cdot P - g(q, t)$.

Example 8.13 – Planar polar coordinates and canonical momenta

Let $x = r\cos\theta,\ \ y = r\sin\theta$ with Cartesian canonical pair $(x, p_x), (y, p_y)$. Define

$$p_r = \cos\theta\, p_x + \sin\theta\, p_y, \qquad p_\theta = -r\sin\theta\, p_x + r\cos\theta\, p_y.$$

Then, using (8.22) with $(x, y; p_x, p_y)$ as the underlying canonical set, one finds

$$\{r, p_r\} = 1, \quad \{\theta, p_\theta\} = 1, \quad \{r, \theta\} = 0, \quad \{r, p_\theta\} = 0, \quad \{\theta, p_r\} = 0, \quad \{p_r, p_\theta\} = 0.$$

Thus $(r, \theta; p_r, p_\theta)$ is a canonical set, in agreement with the point transformation plus appropriate momenta.

Consistency checks for students

When presented with a proposed change of variables $(q, p) \mapsto (Q, P)$, any one of the following suffices to certify canonicity; in practice it is useful to do two:

1. Verify all fundamental brackets in (8.24) with (Q, P) in place of (q, p).
2. Compute $M = \partial(Q, P)/\partial(q, p)$ and check $M^T J\, M = J$.
3. Exhibit a generating function F producing (Q, P) from (q, p).

Theorem 8.5 – Canonical transformation and Poisson bracket preservation

A transformation

$$\eta = \begin{pmatrix} q \\ p \end{pmatrix} \longrightarrow \xi = \begin{pmatrix} Q \\ P \end{pmatrix}$$

is canonical iff it preserves the fundamental brackets, equivalently iff $M^T J\, M = J$. In that case the Poisson bracket of *any* pair of observables is invariant in form,

$$\{U, V\}_{q,p} = (\nabla_\eta U)^T J (\nabla_\eta V) = (\nabla_\xi U)^T J (\nabla_\xi V) = \{U, V\}_{Q,P},$$

and the phase–space volume element $dq\, dp$ is preserved.

Relationship Between Time Derivatives and Poisson Brackets

Let $U(q, p, t)$ be a smooth observable. Its total time derivative is

$$\frac{dU}{dt} = \sum_i \frac{\partial U}{\partial q_i}\dot q_i + \sum_i \frac{\partial U}{\partial p_i}\dot p_i + \frac{\partial U}{\partial t}.$$

Using Hamilton's equations $\dot q_i = \partial\mathcal{H}/\partial p_i$, $\dot p_i = -\partial\mathcal{H}/\partial q_i$,

$$\frac{dU}{dt} = \sum_i \left(\frac{\partial U}{\partial q_i}\frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial U}{\partial p_i}\frac{\partial \mathcal{H}}{\partial q_i} \right) + \frac{\partial U}{\partial t} = \{U, \mathcal{H}\}_{q,p} + \frac{\partial U}{\partial t}.$$

Hence

$$\boxed{\frac{dU}{dt} = \{U, \mathcal{H}\} + \frac{\partial U}{\partial t}} \tag{8.37}$$

and, in symplectic notation with $\eta = (q, p)$ and J as in (8.23),

$$\frac{dU}{dt} = (\nabla_\eta U)^T J\, \nabla_\eta \mathcal{H} + \frac{\partial U}{\partial t}.$$

Example 8.14 – Checks on (8.37)

Taking $U = q_i$ gives $\dot{q}_i = \{q_i, \mathcal{H}\} = \partial\mathcal{H}/\partial p_i$. Taking $U = p_i$ gives $\dot{p}_i = \{p_i, \mathcal{H}\} = -\partial\mathcal{H}/\partial q_i$. Taking $U = \mathcal{H}$ gives $\dot{\mathcal{H}} = \{\mathcal{H}, \mathcal{H}\} + \partial\mathcal{H}/\partial t = \partial\mathcal{H}/\partial t$.

Theorem 8.6 – Conserved quantities and Poisson brackets

(i) U is a constant of motion iff $\dfrac{dU}{dt} = 0$, i.e.

$$\{U, \mathcal{H}\} + \frac{\partial U}{\partial t} = 0.$$

If U has no explicit time dependence, this reduces to $\{U, \mathcal{H}\} = 0$. Equivalently, $\{\mathcal{H}, U\} = \partial U/\partial t$.
(ii) If U and V are constants of motion, then so is their Poisson bracket:

$$\{\mathcal{H}, \{U, V\}\} = 0.$$

Reason: Jacobi identity with $(U, V, \mathcal{H})$ and $\{U, \mathcal{H}\} = \{V, \mathcal{H}\} = 0$.

Remark 8.11 – Geometric reading

Equation (8.37) says the Hamiltonian vector field $X_{\mathcal{H}} = J\nabla_\eta\mathcal{H}$ advects observables: $\dot{U} = \mathcal{L}_{X_{\mathcal{H}}}U + \partial U/\partial t$. Constants of motion are those annihilated by this flow (up to explicit time dependence).

Main Message and Bridge to Hamilton–Jacobi Theory

A canonical transformation

$$(q, p) \xrightarrow{F} (Q, P)$$

maps the Hamiltonian according to

$$K(Q, P, t) = \mathcal{H}(q, p, t) + \frac{\partial F}{\partial t}. \tag{8.38}$$

The ideal goal is to choose F so that the transformed Hamiltonian is as simple as possible. The extreme case is

$$K(Q, P, t) = 0, \tag{8.39}$$

which yields

$$\dot{Q}_i = \frac{\partial K}{\partial P_i} = 0, \qquad \dot{P}_i = -\frac{\partial K}{\partial Q_i} = 0 \quad \Rightarrow \quad Q_i(t) = Q_i(0), \quad P_i(t) = P_i(0). \tag{8.40}$$

Writing the generating function as Hamilton's principal function $S(q, \alpha, t)$ (with constants α identified with the new momenta), condition (8.39) becomes the first–order PDE

$$\boxed{\mathcal{H}\left(q, \frac{\partial S}{\partial q}, t\right) + \frac{\partial S}{\partial t} = 0} \tag{8.41}$$

—the Hamilton–Jacobi equation, which is the subject of the next chapter.
Q

Example 8.15 – Q1 — Show that P

Poisson–bracket test.

$$\frac{\partial P}{\partial p} = p, \quad \frac{\partial P}{\partial q} = q, \quad \frac{\partial Q}{\partial q} = \frac{p}{p^2 + q^2}, \quad \frac{\partial Q}{\partial p} = -\frac{q}{p^2 + q^2}.$$

Hence

$$\{Q,P\}_{q,p} = \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial P}{\partial q} = \frac{p\cdot p + q\cdot q}{p^2+q^2} = 1,$$

and clearly $\{Q,Q\} = \{P,P\} = 0$. The transformation is canonical.

One–form check (gives a generating function). Since $dQ = \dfrac{p\,dq - q\,dp}{p^2+q^2}$, we have

$$P\,dQ = \frac{1}{2}\,(p\,dq - q\,dp) \quad\Rightarrow\quad p\,dq - P\,dQ = \frac{1}{2}\,(p\,dq + q\,dp) = d\left(\frac{1}{2}pq\right).$$

Thus $F(q,p) = \frac{1}{2}pq$ satisfies $p\,dq - P\,dQ = dF$, confirming canonicity.

P

Example 8.16 – Q2 — Show that Q

Poisson–bracket test.

$$\frac{\partial Q}{\partial q} = \tan p, \qquad \frac{\partial Q}{\partial p} = q\,\sec^2 p, \qquad \frac{\partial P}{\partial q} = 0, \qquad \frac{\partial P}{\partial p} = \cot p.$$

Therefore

$$\{Q,P\}_{q,p} = \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial P}{\partial q} = \tan p\cdot\cot p - (q\sec^2 p)\cdot 0 = 1,$$

and $\{Q,Q\} = \{P,P\} = 0$.

Generating function of type $F_3(p,Q)$. With $F_3(p,Q) = -Q\,\ln(\sin p)$, the F_3 relations

$$q = -\frac{\partial F_3}{\partial p}, \qquad P = -\frac{\partial F_3}{\partial Q}$$

give

$$q = Q\,\cot p \iff Q = q\,\tan p, \qquad P = \ln(\sin p),$$

exactly the stated canonical map.

Example 8.17 – Q3 — Harmonic oscillator via $S(q$

For $\mathcal{H}(q,p) = \dfrac{p^2}{2m} + \dfrac{k}{2}q^2$ with $\omega = \sqrt{k/m}$, take the type-F_1 generating function

$$S(q,\Phi) = \frac{1}{2}\,m\omega\,q^2\cot\Phi.$$

From the F_1 rules

$$p = \frac{\partial S}{\partial q} = m\omega\,q\,\cot\Phi, \qquad J \equiv P = -\frac{\partial S}{\partial \Phi} = \frac{1}{2}\,m\omega\,q^2\,\csc^2\Phi.$$

Solve for (q,p):

$$q = \sqrt{\frac{2J}{m\omega}}\,\sin\Phi, \qquad p = \sqrt{2J\,m\omega}\,\cos\Phi.$$

Substitute in $\mathcal{H}$:

$$\mathcal{H} = \frac{(2J\,m\omega)\cos^2\Phi}{2m} + \frac{k}{2}\left(\frac{2J}{m\omega}\right)\sin^2\Phi = J\,\omega.$$

Thus $K(J) = \omega J$, so

$$\dot\Phi = \frac{\partial K}{\partial J} = \omega, \qquad \dot J = -\frac{\partial K}{\partial \Phi} = 0,$$

i.e. $J = \text{const}$ and $\Phi(t) = \omega t + \Phi_0$ (action–angle variables).

Example 8.18 – Q4 — Canonical map from $F_2(q$

Generating function (type F_2).

$$F_2(q,P)=\frac{1}{3}\,q^3+qP,\qquad p=\frac{\partial F_2}{\partial q}=q^2+P,\qquad Q=\frac{\partial F_2}{\partial P}=q.$$

Hence the inverse relations are

$$q=Q,\qquad P=p-q^2=p-Q^2,\qquad p=P+Q^2.$$

Since F_2 is time–independent, the new Hamiltonian is $K(Q,P)=H(q,p)$ with the substitutions above.

Canonicity check (Poisson bracket). From $Q=q$ and $P=p-q^2$ we have

$$\frac{\partial Q}{\partial q}=1,\ \frac{\partial Q}{\partial p}=0,\quad \frac{\partial P}{\partial q}=-2q,\ \frac{\partial P}{\partial p}=1,$$

so

$$\{Q,P\}_{q,p}=\frac{\partial Q}{\partial q}\frac{\partial P}{\partial p}-\frac{\partial Q}{\partial p}\frac{\partial P}{\partial q}=1\cdot 1-0\cdot(-2q)=1,$$

and $\{Q,Q\}=\{P,P\}=0$. The transformation is canonical.

Transforming the harmonic oscillator. For

$$H(q,p)=\frac{p^2}{2m}+\frac{1}{2}\,m\omega^2q^2,$$

insert $q=Q$ and $p=P+Q^2$:

$$K(Q,P)=H(Q,P)=\frac{(P+Q^2)^2}{2m}+\frac{1}{2}\,m\omega^2Q^2=\frac{P^2}{2m}+\frac{Q^4}{2m}+Q^2\left(\frac{P}{m}+\frac{1}{2}m\omega^2\right).$$

Hamilton's equations in (Q,P).

$$\dot Q=\frac{\partial K}{\partial P}=\frac{P+Q^2}{m},\qquad \dot P=-\frac{\partial K}{\partial Q}=-\left(\frac{2Q^3}{m}+\frac{2PQ}{m}+m\omega^2Q\right).$$

Consistency with the original variables follows from $p=P+Q^2$:

$$\dot q=\dot Q=\frac{P+Q^2}{m}=\frac{p}{m},\qquad \dot p=\dot P+2Q\dot Q=-m\omega^2Q=-m\omega^2q.$$

Thus the canonical map generated by F_2 correctly reproduces the oscillator dynamics in the new variables.

Example 8.19 – Q5 — Two Degrees of Freedom: Canonical Transformation and Linearized Hamiltonian

Data.

$$Q_1=q_1^2,\qquad Q_2=q_1+q_2.$$

Find $P_1(q_1,q_2,p_1,p_2)$, $P_2(q_1,q_2,p_1,p_2)$ so that the transformation $(q_i,p_i)\mapsto(Q_i,P_i)$ is canonical. Then, for

$$H(q,p)=\left(\frac{p_1-p_2}{2q_1}\right)^2+p_2+(q_1+q_2)^2, \tag{8.42}$$

choose the free functions so that H becomes $K(Q,P)=P_1^2+P_2$, and solve the motion.

(a) Determine P_1,P_2 from Poisson–bracket conditions. Canonicity requires

$$\{Q_i,Q_j\}=0,\qquad \{P_i,P_j\}=0,\qquad \{Q_i,P_j\}=\delta_{ij}.$$

Because Q_1, Q_2 depend on q only, $\{Q_1, Q_2\} = 0$ holds identically. Use the ansatz

$$P_1 = \frac{p_1 - p_2}{2q_1} + g(q_1, q_2), \qquad P_2 = p_2 + h(q_1, q_2), \tag{8.43}$$

with g, h to be fixed.
Brackets with $Q_1 = q_1^2$.

$$\{Q_1, P_1\} = \frac{\partial Q_1}{\partial q_1}\frac{\partial P_1}{\partial p_1} = 2q_1\left(\frac{1}{2q_1}\right) = 1, \qquad \{Q_1, P_2\} = \frac{\partial Q_1}{\partial q_1}\frac{\partial P_2}{\partial p_1} = 0.$$

So far consistent; P_2 must be independent of p_1, as in (8.43).
Brackets with $Q_2 = q_1 + q_2$.

$$\{Q_2, P_2\} = \frac{\partial Q_2}{\partial q_2}\frac{\partial P_2}{\partial p_2} = 1 \cdot 1 = 1,$$

$$\{Q_2, P_1\} = \frac{\partial Q_2}{\partial q_1}\frac{\partial P_1}{\partial p_1} + \frac{\partial Q_2}{\partial q_2}\frac{\partial P_1}{\partial p_2} = 1 \cdot \frac{1}{2q_1} + 1 \cdot \left(-\frac{1}{2q_1}\right) = 0.$$

Bracket $\{P_1, P_2\} = 0$. Using

$$\frac{\partial P_1}{\partial q_1} = -\frac{p_1 - p_2}{2q_1^2} + \frac{\partial g}{\partial q_1}, \quad \frac{\partial P_1}{\partial q_2} = \frac{\partial g}{\partial q_2}, \quad \frac{\partial P_1}{\partial p_1} = \frac{1}{2q_1}, \quad \frac{\partial P_1}{\partial p_2} = -\frac{1}{2q_1},$$

and $\partial P_2/\partial p_1 = 0$, $\partial P_2/\partial p_2 = 1$, we get

$$\begin{aligned}\{P_1, P_2\} &= \left(\frac{\partial P_1}{\partial q_1}\frac{\partial P_2}{\partial p_1} - \frac{\partial P_1}{\partial p_1}\frac{\partial P_2}{\partial q_1}\right) + \left(\frac{\partial P_1}{\partial q_2}\frac{\partial P_2}{\partial p_2} - \frac{\partial P_1}{\partial p_2}\frac{\partial P_2}{\partial q_2}\right) \\ &= 0 - \frac{1}{2q_1}\frac{\partial h}{\partial q_1} + \frac{\partial g}{\partial q_2} + \frac{1}{2q_1}\frac{\partial h}{\partial q_2} = 0.\end{aligned} \tag{8.44}$$

Equivalently,

$$\boxed{\frac{\partial h}{\partial q_1} - \frac{\partial h}{\partial q_2} = 2q_1 \frac{\partial g}{\partial q_2}}. \tag{8.45}$$

Any g, h obeying (8.45) give a canonical pair (Q, P).

Remark 8.12 – Invertibility domain

Since $Q_1 = q_1^2$, take a branch with $q_1 \neq 0$ (e.g. $q_1 > 0$) so $q_1 = \sqrt{Q_1}$ is single valued.

(b) Choose g, h to obtain $K(Q, P) = P_1^2 + P_2$. From (8.43),

$$\frac{p_1 - p_2}{2q_1} = P_1 - g, \qquad p_2 = P_2 - h.$$

Substituting into (8.42),

$$H = (P_1 - g)^2 + (P_2 - h) + (q_1 + q_2)^2. \tag{8.46}$$

To make H equal to $K = P_1^2 + P_2$, choose

$$\boxed{g(q_1, q_2) = 0, \qquad h(q_1, q_2) = (q_1 + q_2)^2}. \tag{8.47}$$

This satisfies (8.45) because $\partial h/\partial q_1 = \partial h/\partial q_2 = 2(q_1 + q_2)$. Then (8.46) gives

$$K(Q, P) = H = P_1^2 + P_2,$$

which is independent of Q.

Equations of motion and solution. From $K = P_1^2 + P_2$,

$$\dot Q_1 = \frac{\partial K}{\partial P_1} = 2P_1, \quad \dot P_1 = -\frac{\partial K}{\partial Q_1} = 0, \qquad \dot Q_2 = \frac{\partial K}{\partial P_2} = 1, \quad \dot P_2 = -\frac{\partial K}{\partial Q_2} = 0.$$

Hence

$$P_1(t) = \alpha_1, \quad P_2(t) = \alpha_2, \qquad Q_1(t) = 2\alpha_1 t + \beta_1, \quad Q_2(t) = t + \beta_2,$$

with constants set by initial data. Inverting,

$$q_1(t) = \sqrt{Q_1(t)} = \sqrt{2\alpha_1 t + \beta_1}\,, \qquad q_2(t) = Q_2(t) - q_1(t) = t + \beta_2 - \sqrt{2\alpha_1 t + \beta_1}\,.$$

Remark 8.13 – Old momenta (optional)

With (8.47):

$$p_2(t) = P_2 - h = \alpha_2 - (Q_2)^2 = \alpha_2 - (t+\beta_2)^2, \qquad p_1(t) = p_2 + 2q_1 P_1 = \alpha_2 - (t+\beta_2)^2 + 2\alpha_1\sqrt{2\alpha_1 t + \beta_1}\,.$$

Example 8.20 – Q6 — Solve Hamilton's Equations

Given

$$H(q_1, q_2, p_1, p_2) = \frac{p_1^2}{2m} + \frac{(p_2 - k\,q_1)^2}{2m}. \tag{8.48}$$

Here p_i are the (old) canonical momenta; no capital P_i appear.
Hamilton's equations.

$$\dot q_1 = \frac{\partial H}{\partial p_1} = \frac{p_1}{m}, \qquad \dot q_2 = \frac{\partial H}{\partial p_2} = \frac{p_2 - kq_1}{m},$$

$$\dot p_1 = -\frac{\partial H}{\partial q_1} = -\frac{1}{m}(p_2 - kq_1)(-k) = \frac{k}{m}(p_2 - kq_1), \qquad \dot p_2 = -\frac{\partial H}{\partial q_2} = 0.$$

Hence $p_2(t) = \alpha_2$ (constant).
Equation for q_1. Using $\dot q_1 = p_1/m$ and $\dot p_1$ above,

$$\ddot q_1 = \frac{\dot p_1}{m} = \frac{k}{m^2}\big(\alpha_2 - kq_1\big) \quad \Longrightarrow \quad \ddot q_1 + \omega_0^2 q_1 = \frac{k\alpha_2}{m^2}, \qquad \omega_0 \equiv \frac{k}{m}.$$

Let $q_1(0) = q_{10}$, $\dot q_1(0) = v_{10}$. The solution is

$$q_1(t) = q_{10}\cos\omega_0 t + \frac{v_{10}}{\omega_0}\sin\omega_0 t + \frac{\alpha_2}{k}\big(1 - \cos\omega_0 t\big). \tag{8.49}$$

Equation for q_2. From $\dot q_2 = (\alpha_2 - kq_1)/m = -\omega_0\big(q_1 - \alpha_2/k\big)$ and (8.49),

$$\dot q_2 = -\omega_0\left[\big(q_{10} - \alpha_2/k\big)\cos\omega_0 t + \frac{v_{10}}{\omega_0}\sin\omega_0 t\right].$$

Integrating and fixing the constant with $q_2(0) = q_{20}$,

$$q_2(t) = q_{20} + \frac{v_{10}}{\omega_0}\big(\cos\omega_0 t - 1\big) - \big(q_{10} - \tfrac{\alpha_2}{k}\big)\sin\omega_0 t. \tag{8.50}$$

Momenta (optional).

$$p_2(t) = \alpha_2, \qquad p_1(t) = m\dot q_1(t) = -m\omega_0\big(q_{10} - \tfrac{\alpha_2}{k}\big)\sin\omega_0 t + mv_{10}\cos\omega_0 t.$$

Equations (8.49)–(8.50) solve the motion for arbitrary initial data.

Chapter 9

Hamilton–Jacobi Theory and Action–Angle Variables

We have already seen that the power of canonical transformations lies in turning dynamics into geometry. The aim here is to find a special canonical transformation that maps (q_i, p_i) at time t to new variables (Q_i, P_i) that are *constants of motion*. Those constants can then be related directly to the initial conditions at $t = 0$.

9.1. Goal and Set–Up

We seek a canonical transformation

$$(q, p) \xrightarrow{F} (Q, P)$$

generated by a suitable function F such that the transformed Hamiltonian K (see (8.38)) satisfies

$$K(Q, P, t) = \mathcal{H}(q, p, t) + \frac{\partial F}{\partial t}, \tag{9.1}$$

and the new variables obey

$$\dot{Q}_i = \frac{\partial K}{\partial P_i}, \qquad \dot{P}_i = -\frac{\partial K}{\partial Q_i}. \tag{9.2}$$

We want the special case where Q_i and P_i are constants:

$$Q_i(t) = \alpha_i, \qquad P_i(t) = \beta_i.$$

By (9.2), this requirement is equivalent to making K independent of both Q and P:

$$\frac{\partial K}{\partial Q_i} = 0, \qquad \frac{\partial K}{\partial P_i} = 0 \quad \Rightarrow \quad K(Q, P, t) = K(t). \tag{9.3}$$

(Adding a function of time to the Hamiltonian does not change the equations of motion, so we may take $K(t) \equiv 0$ without loss of generality; cf. (8.39).)

9.2. Choosing the Generating Function

Recall the four standard types of generating functions (see the previous chapter). For the construction we need here, the most convenient choice is $F_2(q, P, t)$. We impose the constancy of the new momenta by *naming* them:

$$P_i = \beta_i \quad \text{(constants)}.$$

With $F_2 = F_2(q, \beta, t)$ the transformation equations are

$$p_i = \frac{\partial F_2}{\partial q_i}(q, \beta, t), \qquad Q_i = \frac{\partial F_2}{\partial P_i}(q, P, t)\Big|_{P=\beta} = \frac{\partial F_2}{\partial \beta_i}(q, \beta, t). \tag{9.4}$$

To also make Q_i constant we require

$$\frac{\partial F_2}{\partial \beta_i}(q, \beta, t) = \alpha_i \quad \text{(constants)}. \tag{9.5}$$

Integrating (9.5) with respect to β gives

$$F_2(q, \beta, t) = \sum_i \alpha_i \beta_i + S(q, \alpha, t), \tag{9.6}$$

where $S(q, \alpha, t)$ is independent of β. Then (9.4) reduces to

$$p_i = \frac{\partial S}{\partial q_i}(q, \alpha, t), \qquad Q_i = \frac{\partial S}{\partial \alpha_i}(q, \alpha, t) = \alpha_i \ \ \text{(constant)}. \tag{9.7}$$

Thus, once S is known, the original variables $q_i(t), p_i(t)$ are obtained by solving (9.7) for q_i (and then p_i), with α fixed by initial data; the constants $\beta_i = Q_i(0)$ provide the second set of integration constants. Using (9.1) with $K \equiv 0$ and $F \equiv F_2$ now yields the central equation:

$$\boxed{\mathcal{H}\left(q, \frac{\partial S}{\partial q}, t\right) + \frac{\partial S}{\partial t} = 0} \quad \text{(Hamilton–Jacobi equation)}. \tag{9.8}$$

The function $S(q, \alpha, t)$ is Hamilton's principal function. A complete integral S depending on n independent constants $\alpha_1, \dots, \alpha_n$ produces the canonical transformation to constants.

Remark 9.1 – Momentum Notation

Throughout this chapter, p_i denotes the *old* momenta (conjugate to q_i) and P_i the *new* momenta (conjugate to Q_i). In the special construction above we set $P_i = \beta_i$ (constants) and work with $F_2(q, \beta, t)$ or, equivalently, with $S(q, \alpha, t)$ via (9.6). This keeps the p/P distinction unambiguous.

9.3. Remarks on the Four Types of Generating Functions

The Hamilton–Jacobi equation can be written for any of the four standard generating functions. The form of the PDE—and which set (α or β) emerges as constants of motion—depends on the type.

Remark 9.2 – HJ Equation for Different F

Let F be a generating function producing a canonical map to constants $Q_i = \alpha_i$, $P_i = \beta_i$.

- For $F_1(q, Q, t)$ or $F_2(q, P, t)$ one has

$$\mathcal{H}\left(q, \frac{\partial F}{\partial q}, t\right) + \frac{\partial F}{\partial t} = 0. \tag{9.9}$$

  In the construction above, taking $F_2(q, \beta, t) = \sum_i \alpha_i \beta_i + S(q, \alpha, t)$ leads to (9.8) with α playing the role of constants of motion and $Q_i = \partial S/\partial \alpha_i = \alpha_i$.

- For $F_3(p, Q, t)$ or $F_4(p, P, t)$ one has

$$\mathcal{H}\left(\frac{\partial F}{\partial p}, p, t\right) + \frac{\partial F}{\partial t} = 0. \tag{9.10}$$

  In these cases the constants are identified with either Q_i or P_i according to the type chosen.

In all cases the requirement $Q_i = \alpha_i$, $P_i = \beta_i$ is equivalent to K being independent of Q, P (cf. (9.3)), and the transformation to constants is encoded in the corresponding F.

Remark 9.3 – What Solving the HJ Equation Gives You

A complete integral $S(q, \alpha, t)$ provides the canonical transformation that renders the dynamics trivial in (Q, P)-space. The map back to (q, p) is obtained from (9.7); the parameters α (and the values $Q(0) = \beta$) are fixed by the initial conditions. The whole problem reduces to finding this special canonical transformation.

9.4. The Hamilton–Jacobi Equation for the Principal Generating Function

Our objective is to choose the *best* generating function F that yields a canonical transformation making the transformed Hamiltonian K cyclic in all new coordinates and momenta. A simple way to enforce

$$\dot{Q}_i = 0, \qquad \dot{P}_i = 0 \tag{9.11}$$

is to pick the new Hamiltonian to vanish up to a function of time:

$$K = \mathcal{H}(q, p, t) + \frac{\partial F}{\partial t} = 0. \tag{9.12}$$

Then

$$\frac{\partial K}{\partial P_i} = 0, \qquad \frac{\partial K}{\partial Q_i} = 0, \tag{9.13}$$

so Q_i and P_i are constants of motion.
Using the type-F_2 relations from (9.4) with $F_2 = F_2(q, P, t)$,

$$p_i = \frac{\partial F_2}{\partial q_i}, \tag{9.14}$$

and inserting (9.12) gives the **Hamilton–Jacobi (HJ) equation**

$$\mathcal{H}\left(q, \frac{\partial F_2}{\partial q}, t\right) + \frac{\partial F_2}{\partial t} = 0. \tag{9.15}$$

It is convenient to write F_2 as Hamilton's *principal function*

$$F_2 = S = S(q_1, \ldots, q_n;\, \alpha_1, \ldots, \alpha_n;\, t), \tag{9.16}$$

where the α_i are independent constants of integration. Since only $\partial S/\partial q_i$ and $\partial S/\partial t$ appear,

$$\mathcal{H}\left(q, \frac{\partial S}{\partial q}, t\right) + \frac{\partial S}{\partial t} = 0, \tag{9.17}$$

and an additive constant in S is immaterial. We therefore keep the shorthand

$$S = S(q, \alpha, t), \tag{9.18}$$

with $\alpha = (\alpha_1, \ldots, \alpha_n)$.

Remark 9.4 – Additive Constant in S

Only the derivatives of S enter the transformation formulas (9.4), hence $S \mapsto S + \text{const}$ leaves the dynamics unchanged.

We identify the constants with the *new* conserved momenta:

$$P_i = \alpha_i, \qquad i = 1, \ldots, n. \tag{9.19}$$

Using the F_2 map (9.4) with $F_2 = S$ gives, for all t,

$$p_i = \frac{\partial S}{\partial q_i}(q, \alpha, t). \quad \text{At } t = 0: \quad p_i(0) = \frac{\partial S}{\partial q_i}\big(q(0), \alpha, 0\big), \tag{9.20}$$

which fixes $\alpha = \alpha\big(q(0), p(0)\big)$ from the initial data.
Similarly,

$$Q_i = \frac{\partial S}{\partial \alpha_i}(q, \alpha, t) = \beta_i \quad \text{(constants)}. \quad \text{At } t = 0: \quad \beta_i = \frac{\partial S}{\partial \alpha_i}\big(q(0), \alpha, 0\big). \tag{9.21}$$

Inverting (9.21) yields the complete solution

$$q_i = q_i(\alpha, \beta, t), \qquad p_i = p_i(\alpha, \beta, t). \tag{9.22}$$

Thus, solving the PDE (9.17) is equivalent to integrating the equations of motion.

Remark 9.5 – Freedom in the Constants

Any convenient combinations of the n constants may be used, e.g.

$$\gamma_0 = \gamma_0(\alpha_1, \ldots, \alpha_n), \tag{9.23}$$

provided the transformation $(\alpha) \leftrightarrow (\gamma_0, \ldots)$ is invertible.

Principal Function as the Action

Because $S = S(q, \alpha, t)$,

$$\frac{dS}{dt} = \sum_i \frac{\partial S}{\partial q_i}\dot{q}_i + \frac{\partial S}{\partial t} = \sum_i p_i \dot{q}_i + \frac{\partial S}{\partial t}.$$

Using (9.17) we find

$$\frac{dS}{dt} = \sum_i p_i \dot{q}_i - \mathcal{H} = \mathcal{L}, \tag{9.24}$$

hence

$$S = \int \mathcal{L}\, dt + \text{const.} \tag{9.25}$$

Remark 9.6 – Time–Independent Hamiltonian

If $\mathcal{H}$ does not depend explicitly on t, set

$$S(q, \alpha, t) = S_0(q, \alpha) + S_1(t). \tag{9.26}$$

Equation (9.17) gives $-\dot{S}_1 = \mathcal{H}(q, \partial S_0/\partial q) = a$ (constant). Writing $a = E$ and integrating,

$$S(q, \alpha, t) = W(q, \alpha) - Et, \qquad \mathcal{H}\left(q, \frac{\partial W}{\partial q}\right) = E, \tag{9.27}$$

where W is Hamilton's characteristic function.

9.5. Application to Harmonic Oscillators

Example 9.1 – One–Dimensional Harmonic Oscillator

Consider

$$\mathcal{H}(q, p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2 = E, \qquad \omega^2 = \frac{k}{m}. \tag{9.28}$$

The Hamilton–Jacobi equation reads

$$\frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 + \frac{1}{2}m\omega^2 q^2 + \frac{\partial S}{\partial t} = 0. \tag{9.29}$$

Separated ansatz. Seek a separable solution

$$S(q, \alpha, t) = S_1(q, \alpha) + S_2(t, \alpha). \tag{9.30}$$

Substituting (9.30) into (9.29) gives

$$\frac{1}{2m}\left(\frac{\partial S_1}{\partial q}\right)^2 + \frac{1}{2}m\omega^2 q^2 = -\frac{dS_2}{dt} = \alpha, \tag{9.31}$$

so that

$$S_2(t,\alpha) = -\alpha t, \qquad \left(\frac{\partial S_1}{\partial q}\right)^2 = 2m\alpha - m^2\omega^2 q^2. \tag{9.32}$$

Choose the positive root (the other choice flips the sign of p):

$$\frac{\partial S_1}{\partial q} = \sqrt{2m\alpha}\,\sqrt{1-\frac{m\omega^2 q^2}{2\alpha}}. \tag{9.33}$$

Formally,

$$S_1(q,\alpha) = \sqrt{2m\alpha}\int dq\,\sqrt{1-\frac{m\omega^2 q^2}{2\alpha}}. \tag{9.34}$$

Hence

$$S(q,\alpha,t) = \sqrt{2m\alpha}\int dq\,\sqrt{1-\frac{m\omega^2 q^2}{2\alpha}} - \alpha t. \tag{9.35}$$

Canonical map from S. Only the derivatives of S are needed. The new coordinate

$$Q = \frac{\partial S}{\partial \alpha} = \sqrt{\frac{m}{2\alpha}}\int \frac{dq}{\sqrt{1-\frac{m\omega^2 q^2}{2\alpha}}} - t \equiv \beta, \tag{9.36}$$

is obtained by differentiating under the integral sign. With the change $u = \sqrt{\frac{m\omega^2}{2\alpha}}\,q$,

$$\int \frac{dq}{\sqrt{1-\frac{m\omega^2 q^2}{2\alpha}}} = \frac{1}{\omega}\sin^{-1}\left(\sqrt{\frac{m\omega^2}{2\alpha}}\,q\right),$$

and therefore

$$\beta = \frac{1}{\omega}\sin^{-1}\left(\sqrt{\frac{m\omega^2}{2\alpha}}\,q\right) - t. \tag{9.37}$$

Solution $(q(t), p(t))$. Solving (9.37) for q gives

$$q(t) = \sqrt{\frac{2\alpha}{m\omega^2}}\,\sin(\omega t+\beta). \tag{9.38}$$

Using $p = \dfrac{\partial S}{\partial q}$ with (9.33) and (9.38),

$$p(t) = \sqrt{2m\alpha}\,\cos(\omega t+\beta). \tag{9.39}$$

Energy and phase. From (9.38)–(9.39),

$$\frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2 = \alpha = E,$$

so the separation constant α is the energy, while $Q=\beta$ is the phase of the oscillation.

Constants from initial data. At $t=0$,

$$q_0 = \sqrt{\frac{2\alpha}{m\omega^2}}\sin\beta, \qquad p_0 = \sqrt{2m\alpha}\cos\beta, \tag{9.40}$$

so that

$$\tan\beta = \frac{m\omega\,q_0}{p_0}, \qquad \alpha = \frac{p_0^2}{2m} + \frac{1}{2}m\omega^2 q_0^2. \tag{9.41}$$

Thus S generates the canonical transformation from (q,p) to the phase–energy pair (β,α).

Remark 9.7 – Separation of Variables for Time–Independent Hamiltonians

If $\mathcal{H}$ has no explicit time dependence, the standard choice is

$$S(q, \alpha, t) = W(q, \alpha) - \alpha\, t. \tag{9.42}$$

The Hamilton–Jacobi equation reduces to the stationary form

$$\mathcal{H}\left(q, \frac{\partial W}{\partial q}\right) = \alpha, \tag{9.43}$$

and the remaining constants follow from $Q_i = \dfrac{\partial W}{\partial \alpha_i}$.

Example 9.2 – Two–Dimensional Anisotropic Harmonic Oscillator

Consider the separable Hamiltonian

$$\mathcal{H}(x, y, p_x, p_y) = \frac{p_x^2 + p_y^2}{2m} + \frac{1}{2} m\omega_x^2 x^2 + \frac{1}{2} m\omega_y^2 y^2, \qquad \omega_x^2 = \frac{k_x}{m}, \quad \omega_y^2 = \frac{k_y}{m}. \tag{9.44}$$

Using the stationary ansatz $S = W(x, y; \alpha_x, \alpha_y) - (\alpha_x + \alpha_y)t$ and the additivity $\mathcal{H} = \mathcal{H}_x + \mathcal{H}_y$, write

$$W(x, y; \alpha_x, \alpha_y) = W_x(x; \alpha_x) + W_y(y; \alpha_y). \tag{9.45}$$

The stationary HJ equation separates into

$$\frac{1}{2m}\left(\frac{dW_x}{dx}\right)^2 + \frac{1}{2} m\omega_x^2 x^2 = \alpha_x, \qquad \frac{1}{2m}\left(\frac{dW_y}{dy}\right)^2 + \frac{1}{2} m\omega_y^2 y^2 = \alpha_y, \tag{9.46}$$

with total energy $\alpha = \alpha_x + \alpha_y = E$. By the 1D result, the solution is

$$\begin{aligned} x(t) &= \sqrt{\frac{2\alpha_x}{m\omega_x^2}}\, \sin(\omega_x t + \beta_x), \qquad p_x(t) = \sqrt{2m\alpha_x}\, \cos(\omega_x t + \beta_x), \\ y(t) &= \sqrt{\frac{2\alpha_y}{m\omega_y^2}}\, \sin(\omega_y t + \beta_y), \qquad p_y(t) = \sqrt{2m\alpha_y}\, \cos(\omega_y t + \beta_y), \end{aligned} \tag{9.47}$$

where (β_x, β_y) are phases fixed by $x(0), y(0), p_x(0), p_y(0)$.

9.6. Hamilton–Jacobi Equation for the Characteristic Function

Remark 9.8 – Two Procedures You Will Use in Practice

We summarize two standard routes for solving Hamilton–Jacobi problems.

(A) General, time–dependent case $\mathcal{H}(q, p, t)$. Seek a canonical transformation to variables that are *all* constants of motion:

$$(q, p) \xrightarrow{S} (Q, P), \qquad Q_i = \beta_i, \quad P_i = \gamma_i \text{ (constants)}.$$

It is sufficient to set the transformed Hamiltonian to zero,

$$K \equiv 0 \quad \Rightarrow \quad \dot{Q}_i = \frac{\partial K}{\partial P_i} = 0, \qquad \dot{P}_i = -\frac{\partial K}{\partial Q_i} = 0.$$

The generating function is Hamilton's principal function $S(q, \beta, t)$ that satisfies

$$\boxed{\mathcal{H}\left(q, \frac{\partial S}{\partial q}, t\right) + \frac{\partial S}{\partial t} = 0}. \tag{9.48}$$

A complete solution contains n nontrivial constants $\alpha_1, \ldots, \alpha_n$. The new momenta can be chosen as *any* invertible combination of these:

$$P_i = \gamma_i(\alpha_1, \ldots, \alpha_n).$$

The transformation equations are

$$p_i = \frac{\partial S}{\partial q_i}, \qquad Q_i = \frac{\partial S}{\partial \beta_i} = \beta_i,$$

and the remaining n constants are fixed from the initial conditions.

(B) Time–independent case $\mathcal{H}(q, p)$. Ask only that the *new momenta* are constants by demanding that the new Hamiltonian be cyclic in all Q_i:

$$K = K(P) \quad \Rightarrow \quad \dot{P}_i = 0, \qquad \dot{Q}_i = \frac{\partial K}{\partial P_i} \equiv v_i(P) = \text{const.}$$

Hence

$$P_i = \beta_i, \qquad Q_i(t) = v_i(\beta)\, t + \beta_i', \tag{9.49}$$

where β_i, β_i' are constants. The generating function can be taken *time independent*, the characteristic function $W(q, \beta)$, with

$$\boxed{\mathcal{H}\left(q, \frac{\partial W}{\partial q}\right) = K(\beta)}. \tag{9.50}$$

Here $K(\beta)$ provides one of the constants (often the energy), and the other $n-1$ constants come from $\beta_2, \ldots, \beta_n$, making n independent parameters in total. The mapping is

$$p_i = \frac{\partial W}{\partial q_i}, \qquad Q_i = \frac{\partial W}{\partial \beta_i} = v_i(\beta)\, t + \beta_i'.$$

These last equations are solved for $q_i(t)$ in terms of t and the $2n$ constants $\{\beta_i, \beta_i'\}$. The constants are then evaluated from the initial data $(q_i(0), p_i(0))$.

9.7. Separation of Variables in the Hamilton–Jacobi Equation

Remark 9.9 – When and How Separation Works

The HJ equation reduces dynamics to a first–order PDE; it becomes easy if variables separate. A coordinate q_i is called *separable* if the principal function admits an additive split

$$S(q_1, \ldots, q_n; \alpha, t) = S^{(i)}(q_i; \alpha, t) + S^{(\neg i)}(q_1, \ldots, \widehat{q_i}, \ldots, q_n; \alpha, t). \tag{9.51}$$

Two ingredients determine success:

1) Separability of the Hamiltonian. Start from the full HJ equation

$$\mathcal{H}\left(q_1, \ldots, q_n;\ \frac{\partial S}{\partial q_1}, \ldots, \frac{\partial S}{\partial q_n};\ t\right) + \frac{\partial S}{\partial t} = 0. \tag{9.52}$$

If $\mathcal{H}$ is autonomous, energy is conserved and we set

$$S(q, \alpha, t) = W(q, \alpha) - Et, \qquad E = -\frac{\partial S}{\partial t}. \tag{9.53}$$

A common separable form is

$$\mathcal{H}(q,p)=\sum_{i=1}^{n}T_i(p_i)+\sum_{i=1}^{n}V_i(q_i), \tag{9.54}$$

which turns the stationary HJ equation into

$$\sum_{i=1}^{n}\Big[T_i\Big(\frac{\partial W}{\partial q_i}\Big)+V_i(q_i)\Big]=E=\sum_{i=1}^{n}\alpha_i. \tag{9.55}$$

This suggests the separable ansatz

$$W(q;\alpha)=\sum_{i=1}^{n}W_i(q_i;\alpha_i), \tag{9.56}$$

and yields n decoupled first–order ODEs,

$$T_i\Big(\frac{dW_i}{dq_i}\Big)+V_i(q_i)=\alpha_i, \qquad i=1,\dots,n. \tag{9.57}$$

2) Symmetries / conservation laws. Additional integrals of motion (from symmetries) provide the α_i and ensure decoupling, as in the oscillator examples.

9.7.1. Separation of Variables for Non-Conservative Systems

When the Hamiltonian depends explicitly on time, the Hamilton–Jacobi (HJ) equation is

$$H\Big(q_1,\dots,q_n,\ \frac{\partial S}{\partial q_1},\dots,\frac{\partial S}{\partial q_n},\ t\Big)+\frac{\partial S}{\partial t}=0. \tag{9.58}$$

A convenient ansatz separates the q– and t–dependence:

$$S(q_1,\dots,q_n,t)=\sum_{i=1}^{n}W_i(q_i)-F(t), \tag{9.59}$$

so that

$$H\Big(q_1,\dots,q_n,\ \frac{\partial W}{\partial q_1},\dots,\frac{\partial W}{\partial q_n},\ t\Big)-\frac{dF}{dt}=0, \qquad \frac{\partial W}{\partial q}\equiv\Big(\tfrac{dW_1}{dq_1},\dots,\tfrac{dW_n}{dq_n}\Big). \tag{9.60}$$

Assume H is additively separable as

$$H(q,p,t)=\sum_{i=1}^{n}\Big[T_i(p_i)+V_i(q_i)\Big]+V_t(t). \tag{9.61}$$

Then (9.60) becomes

$$\sum_{i=1}^{n}\Big[T_i\Big(\tfrac{dW_i}{dq_i}\Big)+V_i(q_i)\Big]+V_t(t)-\frac{dF}{dt}=0, \tag{9.62}$$

which separates into

$$\boxed{T_i\Big(\tfrac{dW_i}{dq_i}\Big)+V_i(q_i)=\alpha_i, \qquad \frac{dF}{dt}-V_t(t)=\alpha, \qquad \alpha=\sum_{i=1}^{n}\alpha_i} \tag{9.63}$$

Remark 9.10 – Condition for separability with explicit t

Equations (9.61)–(9.63) imply the structure

$$T(p)=\sum_{i=1}^{n}T_i(p_i),\qquad V(q,t)=\sum_{i=1}^{n}V_0(q_i)+V_t(t), \tag{9.64}$$

as the natural condition for separability in the presence of explicit time dependence.

9.7.2. Most General Additive Structure

We may also take

$$S(q,t)=\sum_{i=1}^{n}S_i(q_i,t), \tag{9.65}$$

with an additively separable Hamiltonian

$$H(q,p,t)=\sum_{i=1}^{n}H_i(q_i,p_i,t). \tag{9.66}$$

The HJ equation (9.58) then yields

$$\sum_{i=1}^{n}\left[H_i\left(q_i,\frac{\partial S_i}{\partial q_i},t\right)+\frac{\partial S_i}{\partial t}\right]=0. \tag{9.67}$$

Hence each term must be constant:

$$H_i\left(q_i,\frac{\partial S_i}{\partial q_i},t\right)+\frac{\partial S_i}{\partial t}=\beta_i,\qquad \sum_{i=1}^{n}\beta_i=0. \tag{9.68}$$

Time-independent case. If H is time-independent, use the stationary ansatz

$$S(q,\alpha,t)=W(q,\alpha)-\alpha\,t. \tag{9.69}$$

Let $E=\alpha=\sum_i\alpha_i$ denote the conserved energy. Then (9.68) becomes

$$H_i\left(q_i,\frac{dW_i}{dq_i};\alpha_1,\ldots,\alpha_n\right)=\gamma_i,\qquad \gamma_i=\beta_i+\alpha_i,\qquad \sum_{i=1}^{n}\gamma_i=E. \tag{9.70}$$

9.7.3. Ignorable Coordinate and Reduction (Kepler-type set-up)

If q_1 is cyclic (ignorable), the stationary HJ equation is

$$H\left(q_1,\ldots,q_n,\frac{\partial W}{\partial q_1},\ldots,\frac{\partial W}{\partial q_n}\right)=E. \tag{9.71}$$

Since $p_1=\dfrac{\partial W}{\partial q_1}=\alpha_1$ is constant, write

$$W(q,\alpha)=\alpha_1q_1+W'(q_2,\ldots,q_n;\alpha_1), \tag{9.72}$$

and (9.71) reduces to

$$H\left(q_2,\ldots,q_n,\frac{\partial W'}{\partial q_2},\ldots,\frac{\partial W'}{\partial q_n};\ \alpha_1\right)=E. \tag{9.73}$$

If r coordinates are cyclic, the characteristic function separates as

$$W(q,\alpha)=\sum_{\text{non-cyclic } i}W_i(q_i;\alpha)\ +\sum_{\text{cyclic } j}\alpha_jq_j, \tag{9.74}$$

and each non-cyclic W_i satisfies a one-coordinate stationary HJ equation of the form

$$H\left(q_i,\frac{dW_i}{dq_i};\text{parameters }\alpha\right)=\alpha_i. \tag{9.75}$$

9.7.4. A Practical Criterion for Separability of a Single Coordinate

A coordinate q_i is *separable* if q_i and its conjugate momentum $p_i = \dfrac{\partial W}{\partial q_i}$ enter the Hamiltonian through a single combination $f(q_i, p_i)$. Then one can write the stationary HJ equation as

$$H\left(q_{\neq i}, \frac{\partial W'}{\partial q_{\neq i}}, f\left(q_i, \frac{dW_i}{dq_i}\right)\right) = \alpha, \qquad W = W_i(q_i;\alpha_i) + W'(q_{\neq i};\alpha), \tag{9.76}$$

i.e.

$$H\left(q_{\neq i}, \frac{\partial W'}{\partial q_{\neq i}}, f\left(q_i, \frac{dW_i}{dq_i}\right)\right) = \alpha. \tag{9.77}$$

If only q_i varies, the left-hand side can remain constant for arbitrary q_i only if

$$\boxed{f\left(q_i, \frac{dW_i}{dq_i}\right) = \alpha_i} \tag{9.78}$$

so that $W_i = W_i(q_i;\alpha_i)$. Substituting back gives the reduced equation

$$H\left(q_{\neq i}, \frac{\partial W'}{\partial q_{\neq i}}; \alpha_i\right) = \alpha, \qquad W' = W'(q_{\neq i};\alpha,\alpha_i). \tag{9.79}$$

An equivalent inversion viewpoint writes

$$f\left(q_i, \frac{dW_i}{dq_i}\right) = g\left(q_{\neq i}, \frac{\partial W'}{\partial q_{\neq i}}; \alpha\right), \tag{9.80}$$

where the left side depends only on q_i while the right depends on the remaining coordinates. Hence both sides must be constants; in particular,

$$f\left(q_i, \frac{dW_i}{dq_i}\right) = \alpha_i, \qquad g\left(q_{\neq i}, \frac{\partial W'}{\partial q_{\neq i}}; \alpha\right) = \alpha_i', \tag{9.81}$$

which again reproduces (9.78) and a reduced equation of the form (9.79).

Remark 9.11 – On the choice of coordinates

Separation in the HJ equation depends not only on the physical system but crucially on the choice of generalized coordinates. The criterion (9.76)–(9.81) is a practical test you can apply in concrete problems.

9.7.5. Kepler Central–Force Problem: stationary characteristic function

Example 9.3 – Planar Kepler problem; stationary W

For motion in a central potential we work in polar coordinates (r, ϕ). The Hamiltonian and the effective potential are

$$\mathcal{H}(r,\phi;p_r,p_\phi) = \frac{p_r^2}{2m} + \frac{p_\phi^2}{2mr^2} + V(r) = \frac{p_r^2}{2m} + V_{\text{eff}}(r), \tag{9.82}$$

$$V_{\text{eff}}(r) = \frac{p_\phi^2}{2mr^2} + V(r). \tag{9.83}$$

Since ϕ is cyclic, p_ϕ = constant. For the *stationary* Hamilton–Jacobi (HJ) method we write the characteristic function as

$$W(r,\phi) = W_1(r) + \alpha_\phi\,\phi, \qquad \alpha_\phi = \text{constant} = p_\phi. \tag{9.84}$$

Energy is conserved, so the stationary HJ equation is

$$\mathcal{H}\left(r, \frac{dW_1}{dr}, \alpha_\phi\right) = \alpha_1 = E = \text{constant}, \tag{9.85}$$

i.e.

$$\frac{1}{2m}\left(\frac{dW_1}{dr}\right)^2 + \frac{\alpha_\phi^2}{2mr^2} + V(r) = \alpha_1. \tag{9.86}$$

Hence

$$\frac{dW_1}{dr} = \sqrt{2m[\alpha_1 - V(r)] - \frac{\alpha_\phi^2}{r^2}}, \tag{9.87}$$

and the characteristic function is

$$\boxed{W(r, \phi; \alpha_\phi, \alpha_1) = \int dr \sqrt{2m[\alpha_1 - V(r)] - \frac{\alpha_\phi^2}{r^2}} \; + \; \alpha_\phi \, \phi \, .} \tag{9.88}$$

Remark 9.12 – About the four generating functions

In the HJ construction we choose the generating function so that the *new* variables are constants. Whether those constants are identified with the new coordinates Q or the new momenta P depends on the type F_1, F_2, F_3 or F_4, but the stationary HJ equation for W has the form used above.

9.7.6. Kepler Central–Force Problem: time–dependent principal function S

Example 9.4 – Full separation

For the Coulomb/Kepler potential $V(r) = -\dfrac{\kappa}{r}$ the Hamiltonian is

$$\mathcal{H}(r, \phi; p_r, p_\phi) = \frac{p_r^2}{2m} + \frac{p_\phi^2}{2mr^2} - \frac{\kappa}{r}. \tag{9.89}$$

The time–dependent HJ equation

$$\mathcal{H}\left(r, \phi; \frac{\partial S}{\partial r}, \frac{\partial S}{\partial \phi}\right) + \frac{\partial S}{\partial t} = 0 \tag{9.90}$$

reads explicitly

$$\frac{\partial S}{\partial t} + \frac{1}{2m}\left[\left(\frac{\partial S}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial S}{\partial \phi}\right)^2\right] - \frac{\kappa}{r} = 0. \tag{9.91}$$

We separate variables as

$$S(r, \phi, t) = S_1(r) + S_2(\phi) + S_3(t). \tag{9.92}$$

Substituting (9.92) into (9.91) gives

$$\frac{1}{2m}\left[\left(\frac{dS_1}{dr}\right)^2 + \frac{1}{r^2}\left(\frac{dS_2}{d\phi}\right)^2\right] - \frac{\kappa}{r} = -\frac{dS_3}{dt}. \tag{9.93}$$

Since the left–hand side depends only on (r, ϕ) and the right–hand side only on t, both are constants. Set

$$-\frac{dS_3}{dt} = \beta_3 = E = \text{constant} \quad \Rightarrow \quad S_3(t) = -\beta_3 t. \tag{9.94}$$

Cyclicity of ϕ implies

$$\frac{dS_2}{d\phi} = p_\phi = \beta_2 = \text{constant} \quad \Rightarrow \quad S_2(\phi) = \beta_2 \, \phi. \tag{9.95}$$

With (9.94)–(9.95), the radial equation from (9.93) is

$$\left(\frac{dS_1}{dr}\right)^2 = 2m\beta_3 + \frac{2m\kappa}{r} - \frac{\beta_2^2}{r^2}, \tag{9.96}$$

so that

$$\frac{dS_1}{dr} = \sqrt{2m\beta_3 + \frac{2m\kappa}{r} - \frac{\beta_2^2}{r^2}}\,. \tag{9.97}$$

Therefore the (time–dependent) principal function is

$$\boxed{S(r,\phi,t;\beta_2,\beta_3) = \int dr\,\sqrt{2m\beta_3 + \frac{2m\kappa}{r} - \frac{\beta_2^2}{r^2}} \;+\; \beta_2\,\phi \;-\; \beta_3\,t\,.} \tag{9.98}$$

Here $\beta_2 = p_\phi$ (constant angular momentum) and $\beta_3 = E$ (constant energy). The corresponding new coordinates $Q_i = \partial S/\partial\beta_i$ are fixed by the initial conditions.

9.7.7. Kepler Central–Force Problem

For plane polar coordinates (r, φ),

$$H(r,\varphi;p_r,p_\varphi) = \frac{1}{2m}\left(p_r^2 + \frac{p_\varphi^2}{r^2}\right) + V(r) = \frac{p_r^2}{2m} + V_{\text{eff}}(r), \qquad V_{\text{eff}}(r) = \frac{p_\varphi^2}{2mr^2} + V(r). \tag{9.99}$$

For the Kepler potential $V(r) = -K/r$ we have

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\varphi^2}{r^2}\right) - \frac{K}{r}, \qquad p_\varphi = mr^2\dot{\varphi} = \text{const.}$$

Because φ is cyclic, the stationary (time–independent) characteristic function can be separated as

$$W(r,\varphi) = W_1(r) + \alpha_\varphi\,\varphi, \qquad \alpha_\varphi = \text{const} = p_\varphi. \tag{9.100}$$

The stationary HJ equation $H(r, \partial W/\partial r; \alpha_\varphi) = \alpha_1 = E$ gives

$$\left(\frac{dW_1}{dr}\right)^2 + \frac{\alpha_\varphi^2}{r^2} + 2m\,V(r) = 2m\,\alpha_1 \quad\Rightarrow\quad \frac{dW_1}{dr} = \sqrt{2m(\alpha_1 - V(r)) - \frac{\alpha_\varphi^2}{r^2}}\,. \tag{9.101}$$

Hence

$$W_1(r) = \int dr\,\sqrt{2m(\alpha_1 - V(r)) - \frac{\alpha_\varphi^2}{r^2}}, \qquad W(r,\varphi) = \int dr\,\sqrt{2m(\alpha_1 - V(r)) - \frac{\alpha_\varphi^2}{r^2}} \;+\; \alpha_\varphi\,\varphi. \tag{9.102}$$

Passing to the time–dependent principal function $S = W - \beta_3 t$ and naming the constants

$$\beta_2 \equiv p_\varphi, \qquad \beta_3 \equiv E,$$

we get (for $V(r) = -K/r$)

$$S(r,\varphi,t;\beta_2,\beta_3) = \int dr\,\sqrt{2m\beta_3 + \frac{2mK}{r} - \frac{\beta_2^2}{r^2}} \;+\; \beta_2\,\varphi \;-\; \beta_3\,t. \tag{9.103}$$

New coordinates from S. By construction $Q_i = \partial S/\partial\beta_i = \alpha_i$ are constants. Thus

$$Q_r = \frac{\partial S}{\partial\beta_3} = \int \frac{m\,dr}{\sqrt{2m\beta_3 + \frac{2mK}{r} - \frac{\beta_2^2}{r^2}}} - t = \alpha_3, \tag{9.104}$$

$$Q_\varphi = \frac{\partial S}{\partial\beta_2} = \varphi - \int \frac{\beta_2\,dr}{r^2\sqrt{2m\beta_3 + \frac{2mK}{r} - \frac{\beta_2^2}{r^2}}} = \alpha_2. \tag{9.105}$$

These quadratures determine $r(t)$ and $\varphi(t)$.

Remark 9.13 – Physical Interpretation of the Principal Function S

Let $S = S(q, \alpha, t)$ with Q, P constants. Then

$$\frac{dS}{dt} = \sum_i \frac{\partial S}{\partial q_i}\dot{q}_i + \frac{\partial S}{\partial t} = \sum_i p_i\dot{q}_i + \frac{\partial S}{\partial t}.$$

Using $H = \sum_i p_i\dot{q}_i - L$ and $H + \partial S/\partial t = 0$,

$$\frac{dS}{dt} = L \quad \Longrightarrow \quad S = \int L(q, \dot{q}, t)\, dt. \tag{9.106}$$

So S coincides with the classical action.

Remark 9.14 – Using F3 or F4

If instead $S = F_3(p, Q, t)$ or $S = F_4(p, P, t)$ with $Q = \alpha$ and $P = \beta$ constants, then

$$\frac{dS}{dt} = -\sum_i q_i\dot{p}_i + \frac{\partial S}{\partial t} = -\frac{d}{dt}\left(\sum_i p_i q_i\right) + L.$$

Hence, up to endpoint terms,

$$S(t_f) - S(t_0) = \int_{t_0}^{t_f} L\, dt - \Big[\sum_i p_i q_i\Big]_{t_0}^{t_f}, \qquad \delta S = \int_{t_0}^{t_f} \delta L\, dt, \tag{9.107}$$

so S retains the interpretation as the action.

9.8. Action–Angle Variables for One Degree of Freedom

Periodic motion appears in many contexts. A convenient treatment proceeds via a canonical change to *action–angle* variables.
Consider a conservative 1D system with

$$H(q, p) = \alpha_1 = \text{const}.$$

Solving for p gives an orbit $p = P(q; \alpha_1)$ in the (q, p) phase plane.

Two kinds of periodic motion.

1. **Libration (bound orbits).** The trajectory in (q, p) is a closed curve; $q(t)$ and $p(t)$ are periodic with the same frequency.
2. **Rotation (unbounded in q).** The motion is periodic in the sense of a variable that advances by 2π per cycle (e.g., an angle), while p is periodic in that angle.

9.9. Action–Angle Variables for One Degree of Freedom

For a conservative 1D system with Hamiltonian $H(q, p) = E$, the momentum along the energy curve is $p = p(q; E)$. The **action** is defined by

$$J \;=\; \oint p(q; E)\, dq, \tag{9.108}$$

where the line integral runs over one full cycle (a closed loop for libration, or a 2π advance for rotation).
In one degree of freedom $J = J(E)$ can be inverted to $E = E(J)$.
Let $W = W(q; J)$ be Hamilton's characteristic function with $p = \partial W/\partial q$. The **angle** variable is defined by

$$w \;=\; \frac{\partial W}{\partial J}, \qquad \dot{w} \;=\; \frac{\partial H}{\partial J} \equiv \nu(J), \qquad w(t) = \nu\, t + \beta. \tag{9.109}$$

Remark 9.15 – Convention

With the present convention $J = \oint p\,dq$ (no $1/2\pi$), the ordinary frequency is $\nu = \omega/2\pi$ when the angular frequency is ω.

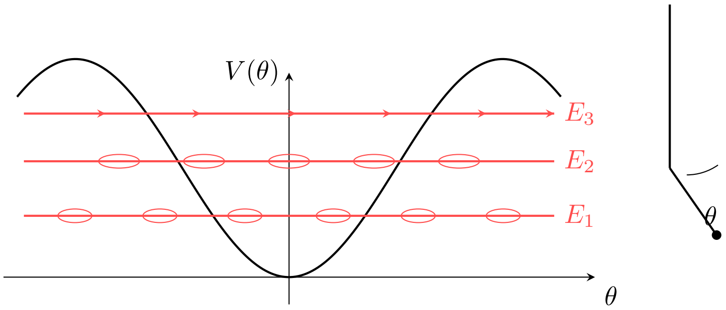


Figure 9.1: Simple pendulum: $V(\theta)$ with typical energies. E_1, E_2: librations (closed oscillations). E_3: rotation.

9.9.1. Harmonic Oscillator

Hamiltonian and energy curve.

$$H(x,p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 = E. \tag{9.110}$$

The energy shell in (x,p) is an ellipse

$$\frac{x^2}{a^2} + \frac{p^2}{b^2} = 1, \qquad a = \sqrt{\frac{2E}{m\omega^2}}, \quad b = \sqrt{2mE}.$$

Its area gives the action:

$$J = \oint p\,dx = \pi ab = \pi\sqrt{\frac{2E}{m\omega^2}}\sqrt{2mE} = \frac{2\pi E}{\omega}. \tag{9.111}$$

Therefore

$$E = \frac{\omega}{2\pi}J, \qquad \nu = \frac{\partial E}{\partial J} = \frac{\omega}{2\pi}, \qquad w(t) = \nu t + \beta. \tag{9.112}$$

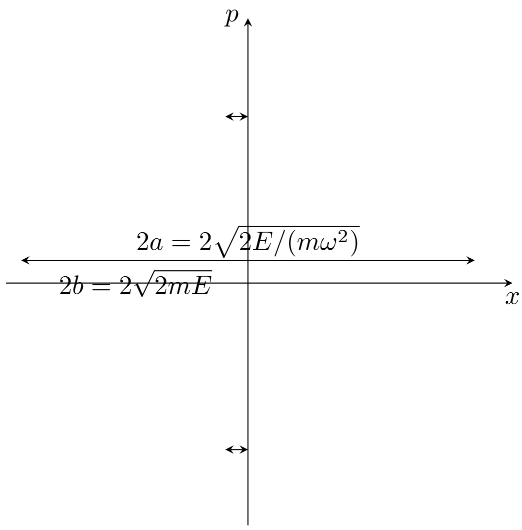


Figure 9.2: Harmonic oscillator in phase space. The action is the area $J = \pi ab = 2\pi E/\omega$.

Direct integral (same result). With $p(x) = \sqrt{2m\bigl(E - \frac{1}{2}m\omega^2 x^2\bigr)}$ and $a = \sqrt{2E/(m\omega^2)}$,

$$J = 4\int_0^a \sqrt{2m\bigl(E - \tfrac{1}{2}m\omega^2 x^2\bigr)}\,dx = 4\sqrt{2mE}\int_0^a \sqrt{1 - \frac{m\omega^2 x^2}{2E}}\,dx$$
$$\xrightarrow{t=\sqrt{\frac{m\omega^2}{2E}}\,x} 4\sqrt{2mE}\,a\int_0^1 \sqrt{1-t^2}\,dt = 4\frac{2E}{\omega}\left(\frac{\pi}{4}\right) = \frac{2\pi E}{\omega}.$$

The motion is

$$x(t) = \sqrt{\frac{2E}{m\omega^2}}\sin(\omega t + \beta), \qquad p(t) = \sqrt{2mE}\cos(\omega t + \beta). \tag{9.113}$$

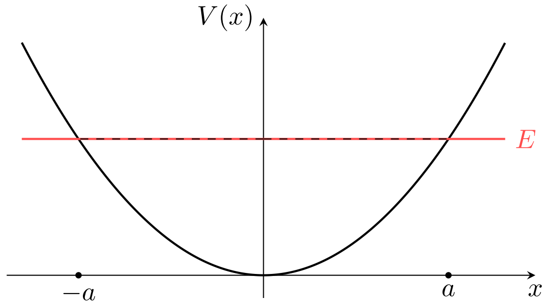


Figure 9.3: Harmonic potential and turning points $\pm a$, $a = \sqrt{2E/(m\omega^2)}$.

9.9.2. Free Particle in a Rigid Box

For a particle in $0 < x < L$ with elastic reflections,

$$H = \frac{p^2}{2m} = E, \qquad p_0 = \pm\sqrt{2mE}.$$

The closed loop in (x, p) is a rectangle of width L and height $2p_0$, hence

$$J = \oint p\,dx = 2Lp_0 = 2L\sqrt{2mE}, \qquad E = \frac{J^2}{8mL^2}, \qquad \nu = \frac{\partial E}{\partial J} = \frac{J}{4mL^2}. \tag{9.114}$$

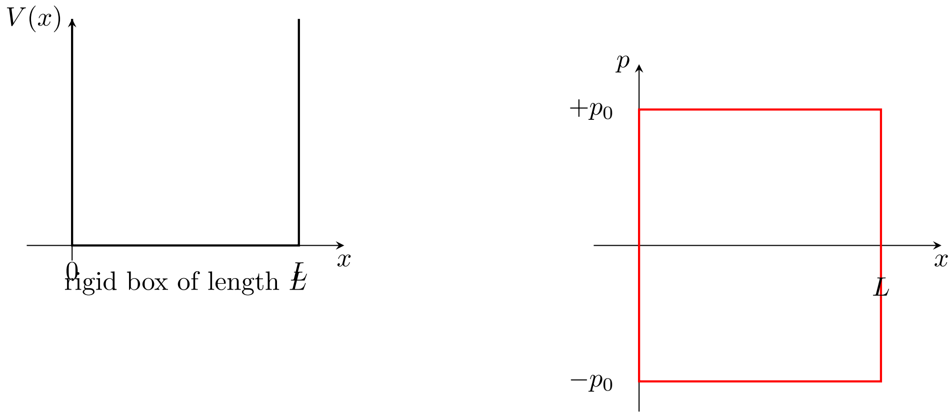


Figure 9.4: Left: infinite square well. Right: phase–space loop is a rectangle; area $= J = 2Lp_0$.

Remark 9.16 – Summary for the HO (all steps)

From (9.110) and $x_{\max} = a = \sqrt{2E/(m\omega^2)}$,

$$J = 4\int_0^a \sqrt{2m\left(E - \tfrac{1}{2}m\omega^2 x^2\right)}\, dx \xrightarrow{t=\sqrt{m\omega^2/(2E)}\, x} \frac{2\pi E}{\omega}.$$

Hence $E = \dfrac{\omega}{2\pi}J$, $\nu = \dfrac{\omega}{2\pi}$, and $x(t), p(t)$ in (9.113) follow.

Sum Up: Action–Angle Variables

We define the action variables on closed curves C_α in each (q_α, p_α) phase plane:

$$J_\alpha \equiv \oint_{C_\alpha} p_\alpha\, dq_\alpha = \oint_{C_\alpha} \frac{\partial S}{\partial q_\alpha}\, dq_\alpha = J_\alpha(\beta_1, \beta_2, \dots, \beta_n), \qquad \alpha = 1, \dots, n. \tag{9.115}$$

Take a complete integral (Hamilton's characteristic function)

$$S(q, \beta) = \sum_{\alpha=1}^{n} S_\alpha\big(q_\alpha; \beta_1, \dots, \beta_n\big), \tag{9.116}$$

which depends on n constants β_i. Since $J_\alpha = J_\alpha(\beta)$, we can (locally) invert to obtain

$$\beta_\alpha = \beta_\alpha(J_1, \dots, J_n), \qquad S = S(q, J_1, \dots, J_n). \tag{9.117}$$

Define the angle variables by

$$p_\alpha = \frac{\partial S}{\partial q_\alpha}, \qquad w_\alpha \equiv \frac{\partial S}{\partial J_\alpha}. \tag{9.118}$$

Then the new canonical coordinates and momenta are

$$Q_\alpha = w_\alpha, \qquad P_\alpha = J_\alpha.$$

Hamilton's equations in these variables read

$$\dot{J}_\alpha = -\frac{\partial H}{\partial w_\alpha}, \qquad \dot{w}_\alpha = \frac{\partial H}{\partial J_\alpha} \equiv f_\alpha(J), \tag{9.119}$$

so that J_α are constants of motion (because $H = H(J)$ is independent of w) and

$$w_\alpha(t) = f_\alpha(J)\, t + C_\alpha. \tag{9.120}$$

Over one period T_α of the α-th motion,

$$\Delta w_\alpha = \int_0^{T_\alpha} \dot{w}_\alpha\, dt = f_\alpha(J)\, T_\alpha. \tag{9.121}$$

On the other hand,

$$\Delta w_\alpha = \oint_{C_\alpha} dw_\alpha = \oint_{C_\alpha} \sum_{i=1}^{n} \frac{\partial}{\partial q_i}\left(\frac{\partial S}{\partial J_\alpha}\right) dq_i = \frac{\partial}{\partial J_\alpha} \oint_{C_\alpha} p_\alpha\, dq_\alpha = \frac{\partial J_\alpha}{\partial J_\alpha} = 1. \tag{9.122}$$

Therefore,

$$f_\alpha(J) = \frac{1}{T_\alpha}, \qquad \boxed{\dot{w}_\alpha = \frac{\partial H}{\partial J_\alpha} = \frac{1}{T_\alpha}} \tag{9.123}$$

i.e. $\dot{w}_\alpha$ is the physical frequency of the α-th oscillation.

Bibliography

[1] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*, 3rd ed., Addison–Wesley, 2002.

[2] J. R. Taylor, *Classical Mechanics*, University Science Books, 2005.

[3] J. Coopersmith, *The Lazy Universe: An Introduction to the Principle of Least Action*, Oxford University Press, 2017.