Lectures on Theoretical Mechanics

Dietrich Belitz
Department of Physics
University of Oregon

Contents

Acknowledgments

These lectures are based in part on notes taken in lectures delivered by Prof. Wolfgang Götze at the Technical University Munich between 1976 and 1985. In their current form they have been taught as a 15-week course for first-year graduate students at the University of Oregon. The original handwritten notes were typeset by Spencer Alexander during the academic year 2013/14. Julian Smith helped with the first round of proofreading. Thanks are due to everyone who has alerted me to typos since. All remaining mistakes and misconceptions are, of course, the author's responsibility.

Disclaimer

These notes are still a work in progress. If you notice any mistakes, whether it's trivial typos or conceptual problems, please send email to dbelitz@uoregon.edu.

Chapter 1

Mathematical principles of mechanics

§ 1 Philosophical comments

§ 1.1 The role of theoretical physics

The diagram below is meant to give a rough idea of how theoretical physics ("Theory") is interconnected with some related subjects.

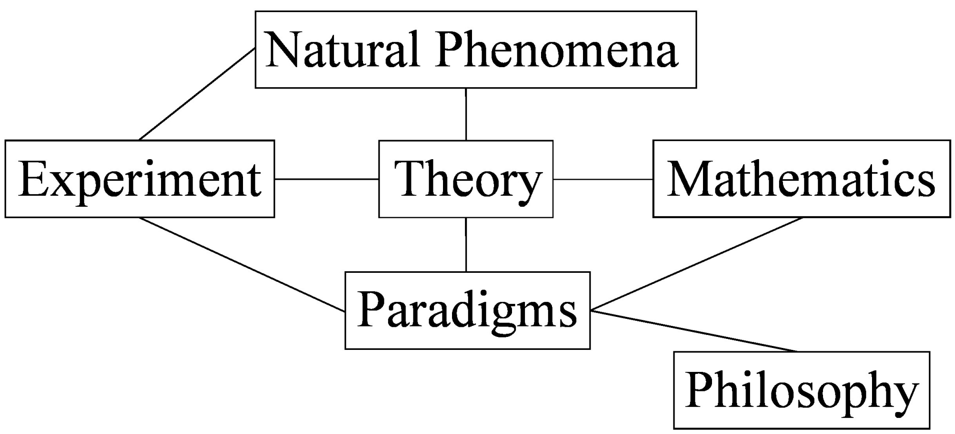

Figure 1.1: The role of theory.

Examples:

Natural phenomena: Planetary motion, atomic spectra, galaxy formation, etc.

Experiment: Measure speed of light, drop objects from leaning tower in Pisa, etc.

Mathematics: ODEs, Calculus of Variations, Group Theory, etc.

Paradigms: Balls and springs, fields, strings, etc.

§ 1.2 The structure of a physical theory

The flowchart below is meant to describe the logical process behind the construction of physical theories.

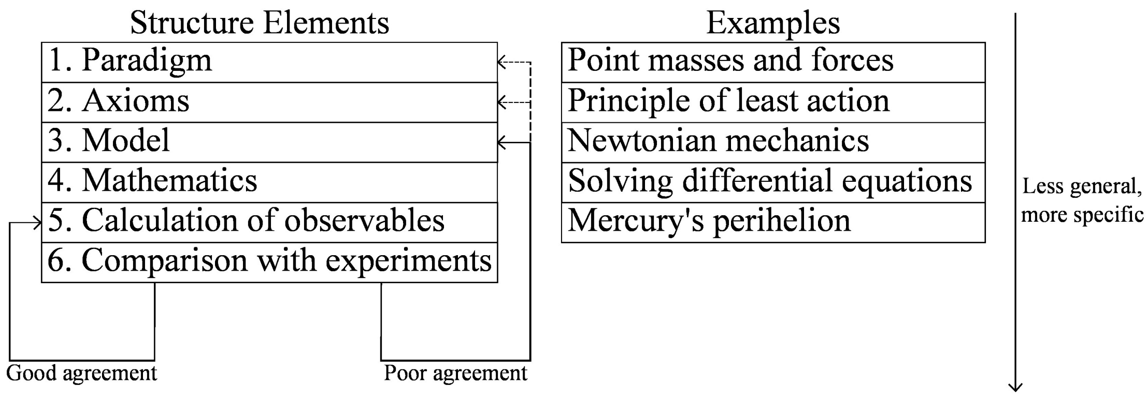

Figure 1.2: The structure of a physical theory.

Remark 1: Occasionally, physical problems spawn the invention of new mathematics (e.g., calculus, differential geometry).

Remark 2: Most theorists deal with steps 4 - 6 in Figure 1.2. Some very good theorists deal with step 3 (e.g., Dirac, Landau). Very few exceptional theorists deal with steps 1 and 2 (e.g., Bohr, Einstein). Occasional theorists deal with setp 4 in an innovative way (e.g., Newton, Witten).

Remark 3: This course explains steps 2 through 5, using the specific example of classical mechanics.

Remark 4: Guidelines for building axioms and models often include intuition, beauty, simplicity etc.; they may or may not include direct experimental input. (Famous example: GR did not!)

§ 2 Review of basic mathematical concepts and notation

§ 2.1 Number sets and functions

We will often use the following familiar **number sets**:

The **natural numbers**,

$$\mathbb{N} = \{1, 2, 3, \ldots\} . \tag{1.1}$$

The **integers**,

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\} . \tag{1.2}$$

The **rational numbers**,

$$\mathbb{Q} = \left\{ \dots, -1, -\frac{1}{2}, -\frac{1}{5}, \dots, 0, \dots, \frac{1}{3}, \frac{1}{2}, 1, \dots \right\}. \tag{1.3}$$

Remark 1*:* Care must be taken to eliminate equivalent fractions ($1/2 = 2/4$, etc.). See Math books for details.

The **real numbers**,

$$\mathbb{R} = \mathbb{Q} \cup \{\text{a suitable completion}\} \tag{1.4}$$

Remark 2*:* A rigorous definition of $\mathbb{R}$ is quite involved, see a suitable math book.

We will also use the **Cartesian product set**

$$\mathbb{R}^n \equiv \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ times}} = \{(x_1, \dots, x_n) \mid x_i \in \mathbb{R}\} \ , \tag{1.5}$$

defined as the set of all ordered real n-tuplets.

Remark 3*:* Elements of $\mathbb{R}^n$ are denoted by $\boldsymbol{x} = (x_1, \dots, x_n)$, $x_i \in \mathbb{R}$, and are called n**-vectors**, or simply **vectors**. 1-vectors are called **scalars**. This is actually a special case of a much more general concept; see any book on vector spaces.

Consider functions $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^m$. We say $\boldsymbol{f}$ is an (m-vector-valued) function of n real variables and write

$$\boldsymbol{f}(\boldsymbol{x}) \equiv \boldsymbol{f}(x_1, \dots, x_n) = \boldsymbol{y} \quad , \quad \boldsymbol{x} \in \mathbb{R}^n \quad , \quad \boldsymbol{y} \in \mathbb{R}^m. \tag{1.6}$$

Example 1: $\boldsymbol{f} : \mathbb{R}^3 \to \mathbb{R} \quad , \quad f(\boldsymbol{x}) = |\boldsymbol{x}| \equiv \left(x_1^2 + x_2^2 + x_3^2\right)^{1/2}$ is a real-valued function of $\boldsymbol{x} \in \mathbb{R}^3$ called the **norm** of $\boldsymbol{x}$.

Remark 4*:* For $m = 1$, we write f instead of $\boldsymbol{f}$.

§ 2.2 Differentiation

Definition 1:

a) Let $n = m = 1$. We define the **derivative** of f,

$$f' \equiv \frac{df}{dx} : \mathbb{R} \to \mathbb{R}, \tag{1.7}$$

as

$$f'(x) \equiv \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[f(x + \varepsilon) - f(x)\right]. \tag{1.8}$$

provided the limit exists.

b) For $n > 1$, $m = 1$ we define **partial derivatives** of f,

$$\frac{\partial f}{\partial x_i} \equiv \partial_i f : \mathbb{R}^n \to \mathbb{R}, \tag{1.9}$$

as

$$\partial_i f \equiv \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[f\left(x_1, \ldots, x_i + \varepsilon, \ldots, x_n\right) - f\left(x_1, \ldots, x_i, \ldots, x_n\right) \right]. \tag{1.10}$$

We define the **gradient** of f,

$$\frac{\partial f}{\partial \boldsymbol{x}} \equiv \nabla f : \mathbb{R}^n \to \mathbb{R}^n \tag{1.11}$$

as

$$\nabla f \equiv \left(\partial_1 f, \ldots, \partial_n f\right). \tag{1.12}$$

c) For $n = 1$, $m > 1$, we define the **derivative** of $\boldsymbol{f}$,

$$\frac{d\boldsymbol{f}}{dx} : \mathbb{R} \to \mathbb{R}^m, \tag{1.13}$$

as

$$\frac{d\boldsymbol{f}}{dx} \equiv \left(\frac{df_1}{dx}, \ldots \frac{df_m}{dx} \right). \tag{1.14}$$

d) For $n = m$, we define the **divergence** of $\boldsymbol{f}$,

$$\operatorname{div} \boldsymbol{f} \equiv \nabla \cdot \boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}, \tag{1.15}$$

as

$$\begin{aligned} \nabla \cdot \boldsymbol{f} &\equiv \sum_{i=1}^{n} \partial_i f_i \\ &\equiv \partial^i f_i, \end{aligned} \tag{1.16}$$

where the last line uses the common **summation convention**.

e) For $n = m = 3$ we define the **curl** of $\boldsymbol{f}$,

$$\operatorname{curl} \boldsymbol{f} \equiv \nabla \times \boldsymbol{f} : \mathbb{R}^3 \to \mathbb{R}^3, \tag{1.17}$$

as

$$\left(\nabla \times \boldsymbol{f}\right)_i \equiv \varepsilon_{ijk} \partial_j f_k, \tag{1.18}$$

where

$$\varepsilon_{ijk} = \begin{cases} +1 & \text{if } (ijk) \text{ is an even permutation of } (1,2,3), \\ -1 & \text{if } (ijk) \text{ is an odd permutation of } (1,2,3), \\ 0 & \text{if } (ijk) \text{ is not a permutation of } (1,2,3), \end{cases} \tag{1.19}$$

is the **Levi-Civita tensor**, or the **completely antisymmetric tensor of rank** 3.

Remark 1*:* We will define tensors in general, and deal with them in detail, in Chapter 3.

Definition 2: Let $I = [t_0, t_1] \subset \mathbb{R}$ be a real interval, and let $\boldsymbol{x} : I \to \mathbb{R}^n$ be a function of $t \in I$. Let $f : \mathbb{R}^n \times I \to \mathbb{R}$ be a real-valued function of $\boldsymbol{x}$ and t. Then the **total derivative** of f with respect to t is the function

$$\frac{df}{dt} : I \to \mathbb{R}, \tag{1.20}$$

defined by

$$\frac{df}{dt}\left(t^*\right) \equiv \partial_t f\left(\boldsymbol{x}\left(t^*\right), t^*\right) + \partial_i f\left(\boldsymbol{x}\left(t^*\right), t^*\right) \frac{dx_i}{dt}\left(t^*\right). \tag{1.21}$$

Example 1: Let $f : \mathbb{R}^2 \to \mathbb{R}$, $f(\boldsymbol{x}) = |\boldsymbol{x}| = \sqrt{x_1^2 + x_2^2}$ be the norm of the vector $\boldsymbol{x}$. Then,

$$\frac{\partial f}{\partial x_i} = \frac{x_i}{|\boldsymbol{x}|} \quad , \quad \nabla f = \frac{\boldsymbol{x}}{|\boldsymbol{x}|}, \tag{1.22}$$

so ∇f is the normalized vector. Let $\boldsymbol{x} : [0, 2\pi] \to \mathbb{R}$, $\boldsymbol{x}(t) = (a \cos t, b \sin t)$, with $a, b \in \mathbb{R}$. Then,

$$\begin{aligned} \frac{df}{dt} &= \partial_i f \frac{dx_i}{dt} \\ &= -\frac{x_1}{|\boldsymbol{x}|} a \sin t + \frac{x_2}{|\boldsymbol{x}|} b \cos t \\ &= \frac{1}{|\boldsymbol{x}|} \left(b^2 - a^2\right) \sin t \cos t \\ &= \frac{1}{2} \left(b^2 - a^2\right) \frac{\sin 2t}{\sqrt{a^2 \cos^2 t + b^2 \sin^2 t}}. \end{aligned} \tag{1.23}$$

Problem 1: Total derivative

Consider Example 1 above: Let $\boldsymbol{x} : \mathbb{R} \to \mathbb{R}^2$ be a function defined by

$$\boldsymbol{x}(t) = (a \cos t, b \sin t) \quad , \quad a, b \in \mathbb{R},$$

and let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function defined by

$$f(\boldsymbol{x}) = \sqrt{x_1^2 + x_2^2}.$$

Discuss the behavior of the total derivative, df/dt, and give a geometric interpretation of the result.

Hint: First determine the geometric figure in $\mathbb{R}^2$ that $\boldsymbol{x}$ provides a parametric representation of, then consider the geometric meaning of $|\boldsymbol{x}(t)| = f(\boldsymbol{x}(t))$.

---

§ 2.3 Stationary points

Definition 1: Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. f is called stationary at $\boldsymbol{x}_0 \in \mathbb{R}^n$ if

$$\partial_i f(\boldsymbol{x}_0) = 0 \quad , \quad (i = 1, \ldots, n). \tag{1.24}$$

Theorem 1: A necessary condition for f to have an extremum in the point $\boldsymbol{x}_0$ is that f is stationary at $\boldsymbol{x}_0$.

Proof: MATH 281 or equivalent. □

Remark 1: While necessary, f being stationary at $\boldsymbol{x}_0$ is not a sufficient condition for f to have an extremum at $\boldsymbol{x}_0$, due to the possibility of saddle points.

Corollary 1: Let the arguments $\boldsymbol{x} = (x_1, \ldots, x_n)$ of f be constrained to a set $S \subset \mathbb{R}^n$ defined by N constraints

$$g_j(x_1, \ldots x_n) \equiv c_j \quad , \quad j = 1, \ldots, N \quad , \quad c_j \in \mathbb{R}, \tag{1.25}$$

with N functions $g_j : \mathbb{R}^n \to \mathbb{R}$. Then a necessary condition for f to have an extremum at $\boldsymbol{x}^0 \in S$ is

$$\partial_i \tilde{f}\left(\boldsymbol{x}^0\right) = 0, \tag{1.26}$$

where

$$\tilde{f}(\boldsymbol{x}) = f(\boldsymbol{x}) + \sum_{j=1}^{N} \lambda_j g_j(\boldsymbol{x}), \tag{1.27}$$

and the $n+N$ unknowns $x_1^0, \ldots, x_n^0; \lambda_1, \ldots, \lambda_N$ are to be determined from the $n+N$ conditions given in Eqs. (1.25) and (1.26).

Proof: MATH 281 or equivalent. □

Remark 2: The constants λ_i are called **Lagrange multipliers**.

Example 1: Let $f(x, y, z) = x - y + z$ be defined on the 2-sphere

$$S_2 = \left\{(x, y, z) \mid x^2 + y^2 + z^2 = 1\right\}. \tag{1.28}$$

Then,

$$\tilde{f}(x, yz) = x - y + z + \lambda\left(x^2 + y^2 + z^2\right). \tag{1.29}$$

A necessary condition for extrema is

$$1 + 2\lambda x = 0 \quad , \quad -1 + 2\lambda y = 0 \quad , \quad 1 + 2\lambda z = 0, \tag{1.30}$$

and

$$x^2 + y^2 + z^2 = 1. \tag{1.31}$$

Solving this system of four equations, we find that candidates for extrema are

$$(x, y, z) = \pm\frac{1}{\sqrt{3}}(1, -1, 1). \tag{1.32}$$

Problem 2: Extrema subject to constraints

Consider Example 1 above: Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a function defined by

$$f(x, y, z) = x - y + z.$$

Let $S_2 = \left\{(x, y, z) \mid x^2 + y^2 + z^2 = 1\right\}$ be the 2-sphere.

a) Show that the extremal points for f on S_2 are $(1, -1, 1)/\sqrt{3}$, and $(-1, 1, -1)/\sqrt{3}$, as claimed in the lecture.

b) Determine which of these extremal points, if any, are a maximum or a minimum, and determine the corresponding extremal values of f on S_2.

Problem 3: Minimal distance on the 2-sphere

Consider the 2-sphere, $S_2 = \left\{(x,y,z) \mid x^2+y^2+z^2=1\right\}$ embedded in $\mathbb{R}^3$. Find the point on S_2 that is closest to the point $(1,1,1) \in \mathbb{R}^3$, and determine the distance between the two points.

Proposition 1: (Taylor expansion)

A function $f : \mathbb{R}^n \to \mathbb{R}$ that is n-times differentiable at $\boldsymbol{x}$ can, in a neighborhood of $\boldsymbol{x}$, be represented by a power series

$$\begin{aligned} f\left(x_1+\varepsilon, x_2, \ldots, x_n\right) &= f\left(x_1, \ldots, x_n\right) + \varepsilon \frac{\partial f}{\partial x_1}\left(x_1, \ldots, x_n\right) + \ldots \\ &= \sum_{m=0}^{M} \frac{1}{m!} \varepsilon^m \frac{\partial^m f}{\partial x_1^m}\left(x_1, \ldots, x_n\right) + R_M, \end{aligned} \tag{1.33}$$

with R_M a remainder, and analogously for the other variables $x_2, \ldots, x_n$.

Proof: MATH 251-3 or equivalent. □

Remark 3: Taylor's theorem gives an explicit bound for the remainder R_M; see your favorite calculus book for details.

§ 2.4 Ordinary differential equations

Definition 1:

a) Let $I = [t_-, t_+] \subset \mathbb{R}$ be a real interval, and let $\boldsymbol{y} : I \to \mathbb{R}^n$ be a function of t. Let $\boldsymbol{f} : \mathbb{R}^n \times I \to \mathbb{R}^n$ be a function of $\boldsymbol{y}$ and t. Then

$$\frac{d\boldsymbol{y}}{dt} = f\left(\boldsymbol{y}, t\right), \tag{1.34}$$

or, in components,

$$\frac{dy_i}{dt} = f_i\left(\boldsymbol{y}, t\right) \quad , \quad (i = 1, \ldots, n) \tag{1.35}$$

is called a system of n **ordinary differential equations**, or **ODEs**, of first order.

b) Let $t_0 \in I$, $\boldsymbol{y}_0 \in \mathbb{R}^n$. A function $\boldsymbol{y} : I \to \mathbb{R}^n$ that obeys equation (1.34) and has the property that

$$\boldsymbol{y}\left(t_0\right) = \boldsymbol{y}_0 \tag{1.36}$$

is said to solve equation (1.34) under the **initial condition** given by equation (1.36).

Theorem 1: Let $\boldsymbol{f}\left(\boldsymbol{y}, t\right)$ and its partial derivatives with respect to each y_i be bounded and continuous in a neighborhood of $N\left(\boldsymbol{y}_0, t_0\right)$ of $\left(\boldsymbol{y}_0, t_0\right) \in \mathbb{R}^{n+1} \times I$. Then there exists a unique solution $\boldsymbol{y}\left(t\right)$ to equation (1.34) under the initial condition given by equation (1.36).

Proof: MATH 256 □

Example 1: The system of ODEs

$$\frac{dx_1}{dt} = 2x_1 + 4x_2 + 2 \quad , \quad \frac{dx_2}{dt} = x_1 - x_2 + 4, \tag{1.37}$$

with

$$x_1\,(t=0) = -8 \quad , \quad x_2\,(t=0) = 1 \tag{1.38}$$

is uniquely solved by

$$x_1\,(t) = e^{-2t} - 4e^{3t} + 3 \quad , \quad x_2\,(t) = e^{-2t} - 4e^{3t} + 1. \tag{1.39}$$

Remark 1*:* The initial condition uniquely determines the solution of a first-order ODE.

Problem 4: System of ODEs

Solve the system of first order ODEs considered in Example 1 above:

$$\begin{aligned} \dot{x} &= 2x + 4y + 2, \\ \dot{y} &= x - y + 4. \end{aligned}$$

Corollary 1: Let $\boldsymbol{y} : I \to \mathbb{R}^n$ be a function of t, and let $\dot{\boldsymbol{y}} = d\boldsymbol{y}/dt$ be its derivative. Let $\boldsymbol{f}\,(\boldsymbol{y}, \dot{\boldsymbol{y}}, t) : \mathbb{R}^n \times \mathbb{R}^n \times I \to \mathbb{R}^n$ be bounded and continuously differentiable in a neigborhood of $(\boldsymbol{y}_0, \dot{\boldsymbol{y}}_0, t_0) \in \mathbb{R}^n \times \mathbb{R}^n \times I$. Then there exists a unique function $\boldsymbol{y}$ such that

$$\boldsymbol{y}\,(t_0) = \boldsymbol{y}_0 \quad , \quad \dot{\boldsymbol{y}}\,(t_0) = \dot{\boldsymbol{y}}_0 \quad , \quad \ddot{\boldsymbol{y}}\,(t) = \boldsymbol{f}\,(\boldsymbol{y}, \dot{\boldsymbol{y}}, t)\,. \tag{1.40}$$

Proof: MATH 256. □

Remark 2*:* $\ddot{\boldsymbol{y}} = \boldsymbol{f}\,(\boldsymbol{y}, \dot{\boldsymbol{y}}, t)$ is a system of second-order ODEs.

Remark 3*:* The initial point $\boldsymbol{y}_0$ and the initial tangent vector $\dot{\boldsymbol{y}}_0$ uniquely determine the solution of a second-order ODE.

Example 2: $\ddot{y}\,(t) = -y$ has the general solution $y\,(t) = c_1 \cos t + c_2 \sin t$. The initial conditions $y\,(t=0) = 1$ and $\dot{y}\,(t=0) = 0$ determine $c_1 = 1$, $c_2 = 0$.

§ 2.5 Integration

Definition 1: Let $f : I \to \mathbb{R}$ be a scalar function of t. Then the **Riemann integral** is defined as the limit of a sum,

$$\begin{aligned} F &= \int_{t_-}^{t_+} dt\, f\,(t) \\ &\equiv \lim_{N\to\infty} \sum_{i=1}^{N-1} (t_{i+1} - t_i)\, f\,(t_i)\,. \end{aligned} \tag{1.41}$$

Remark 1*:* F exists if f is bounded and continuous on I.

Remark 2*:* The generalization for $f : I \times J \to \mathbb{R}$, $F = \int_{t_-}^{t_+} dt \int_{u_-}^{u_+} du\, f\,(t, u)$ is straightforward.

Remark 3: Let $B = \{f : I \to \mathbb{R} \mid f \text{ is bounded and continuous}\}$. Then, $F : B \to \mathbb{R}$ maps functions onto numbers. Such mappings are called **functionals**.

Proposition 1 (Integration by parts)**:** An integral of the form $\int_{t_-}^{t_+} dt\, f(t) \frac{d}{dt} g(t)$ can be rewritten as

$$\begin{aligned} \int_{t_-}^{t_+} dt\, f(t) \frac{d}{dt} g(t) &= -\int_{t_-}^{t_+} dt \left(\frac{d}{dt} f(t)\right) g(t) + f(t_+)\, g(t_+) - f(t_-)\, g(t_-) \\ &\equiv -\int_{t_-}^{t_+} dt \left(\frac{d}{dt} f(t)\right) g(t) + f(t)\, g(t) \Big|_{t_-}^{t_+}. \end{aligned} \tag{1.42}$$

Proof: MATH 251-3. □

§ 2.6 Paths and path functionals

Definition 1:

a) Let $I = [t_-, t_+] \subset \mathbb{R}$ and let $\boldsymbol{q} : I \to \mathbb{R}^n$ be continuously differentiable. Then the set $\mathcal{C} \equiv \{\boldsymbol{q}(t) \,|t \in I\} \subset \mathbb{R}^n$ is called a **path**, or **curve**, in $\mathbb{R}^n$, and $\boldsymbol{q}(t)$ is called a **parameterization** of $\mathcal{C}$ with parameter t.

b) $\mathcal{C}$ inherits an order from the $\gtrless$ order defined on I as follows: $\boldsymbol{q}(t_1) < \boldsymbol{q}(t_2)$ is defined to be true if and only if $t_1 < t_2$.

c) $\boldsymbol{q}_\pm \equiv \boldsymbol{q}(t_\pm)$ are the **start** and **end points** of $\mathcal{C}$.

d) The **tangent vector** $\boldsymbol{\tau}(t)$ to $\mathcal{C}$ in the point $\boldsymbol{q}(t)$ is defined as

$$\boldsymbol{\tau}(t) \equiv \frac{d}{dt} \boldsymbol{q}(t) \equiv \dot{\boldsymbol{q}}(t). \tag{1.43}$$

d) If $\dot{\boldsymbol{q}}(t) \neq 0 \quad \forall t \in I$, then the curve is called **smooth**.

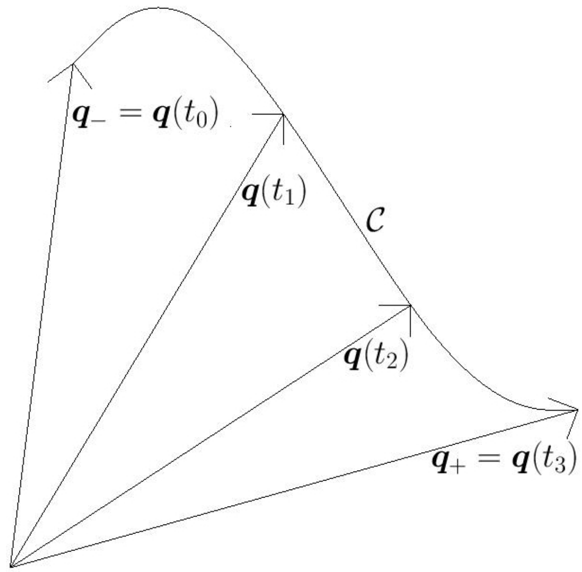

Figure 1.3: A parameterized path.

Definition 2: Let $L : \mathbb{R}^n \times \mathbb{R}^n \times I \to \mathbb{R}$ be a function of $\boldsymbol{q}$, $\dot{\boldsymbol{q}}$, and t. Let L be twice continuously differentiable with respect to all arguments. Let $\mathcal{C}$ be a path with parameterization $\boldsymbol{q}(t)$. Then we define a functional $S_L(\mathcal{C})$ as

$$S_L(\mathcal{C}) := \int_{t_-}^{t_+} dt\, L(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t), t)\,. \tag{1.44}$$

Remark 1*:* For a given L, $S_L(\mathcal{C})$ is characteristic of the path $\mathcal{C}$.

Remark 2*:* If $\mathcal{C}$ changes slightly, $\mathcal{C} \to \mathcal{C} + \delta\mathcal{C}$, then $S_L(\mathcal{C})$ changes slightly as well, $S_L \to S_L + \delta S_L$.

Example 1: Let $n = 2$ and let $\mathcal{C}$ be a path with parameterization $\boldsymbol{q}(t)$. The length $l_{\mathcal{C}}$ of $\mathcal{C}$ is given by

$$\begin{aligned} l_{\mathcal{C}} &= \lim_{N\to\infty} \sum_{i=1}^{N-1} \sqrt{[\boldsymbol{q}(t_{i+1}) - \boldsymbol{q}(t_i)]^2} \\ &= \lim_{N\to\infty} \sum_{i=1}^{N-1} (t_{i+1} - t_i) \sqrt{\frac{[\boldsymbol{q}(t_{i+1}) - \boldsymbol{q}(t_i)]^2}{[t_{i+1} - t_i]^2}} \\ &= \lim_{N\to\infty} \sum_{i=1}^{N-1} \Delta t \sqrt{(\Delta \boldsymbol{q}/\Delta t)^2} \\ &= \int_{t_-}^{t_+} dt\, \sqrt{\dot{\boldsymbol{q}}^2(t)}. \end{aligned} \tag{1.45}$$

That is, the choice $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = \sqrt{\dot{\boldsymbol{q}}^2} = |\dot{\boldsymbol{q}}|$ yields $S_L(\mathcal{C}) = l_{\mathcal{C}}$.

Example 2: Let $n = 3$. With t representing time, let a point mass m move on a trajectory $\boldsymbol{x}(t)$ in a potential $U(\boldsymbol{x})$. The kinetic energy is given by $E_{\text{kin}} = \dot{\boldsymbol{x}}^2/2m$. Consider $L(\boldsymbol{x}, \dot{\boldsymbol{x}}, t) = E_{\text{kin}}(\dot{\boldsymbol{x}}) - U(\boldsymbol{x})$. Then,

$$S = \int_{t_-}^{t_+} dt\, L(\boldsymbol{x}, \dot{\boldsymbol{x}}) \tag{1.46}$$

is called the **action** for the trajectory $\boldsymbol{x}(t)$.

Problem 5: Passage time

Consider a path $\mathcal{C}$ in $\mathbb{R}^2$ with a parameterization $\boldsymbol{q}(t)$, and a point mass moving along $\mathcal{C}$ with speed $v(\boldsymbol{q})$. Let $T(\mathcal{C})$ be the passage time of the particle from $\boldsymbol{q}_-$ to $\boldsymbol{q}_+$. Find the function $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ such that the functional $S_L(\mathcal{C})$ is equal to $T(\mathcal{C})$.

Problem 6: Get by with a little help from your Friend$^{\text{TM}}$

Spring-loaded camming devices, also known as Friends^{TM} or Camalots^{TM}, depending on the manufacturer, are used by rock climbers to protect the climber in case of a fall. The devices consist of four metal wedges that pairwise rotate against one another so that the outside edge of each pair of wedges moves along a curve. The cam is placed in a crack with parallel walls, where the springs hold it in place. The *camming angle* α is defined as the angle between the line from the center of rotation to the contact point with the rock and the tangent to the curve in the contact point. The figure below shows the camming angle for an almost-extended cam in a wide crack, and a largely retracted cam in a narrow crack.

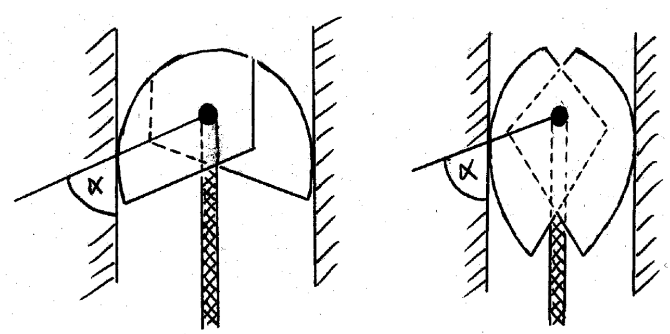

Figure 1.4

If you fall, you want to get as much help from you Friend^{TM} as the laws of physics let you. To ensure that, you want the camming angle to be the same irrespective of the width of the crack (see the figure). Determine the shape of the curve the cam surfaces must form to ensure that is the case.

a) Parametrize the curve, using polar coordinates, as $(r(t), \varphi(t))$. Find the tangent vector in the point $P = (r, \varphi)$ in Cartesian coordinates.

b) Define the angle β between the tangent in P and the line through P that is perpendicular to the radius vector from the origin to P. How is β related to α? Show that

$$\tan\beta = (dr/d\varphi)/r.$$

c) Solve the differential equation that results from requiring that $\beta = \text{constant}$ along the curve. Discuss the solution, which is the desired shape of the cam.

---

§ 2.7 Surfaces

Definition 1:

a) Let $I_t = [t_-, t_+] \subset \mathbb{R}$ and let $I_u = [u_-, u_+] \subset \mathbb{R}$ be intervals. Let $\boldsymbol{r} : I_t \times I_u \to \mathbb{R}^3$ be a continuously differentiable function $\boldsymbol{r}(t, u)$. Then,

$$S \equiv \{\boldsymbol{r}(t, u) \mid (t, u) \in I_t \times I_u\} \tag{1.47}$$

is called a **surface** in $\mathbb{R}^3$ with parameterization $\mathbf{r}(t, u)$.

b) The **standard normal vector** $\boldsymbol{n}(t, u)$ of S in $\boldsymbol{r}(t, u)$ is defined as

$$\boldsymbol{n}(t, u) \equiv \partial_t \boldsymbol{r}(t, u) \times \partial_u \boldsymbol{r}(t, u) . \tag{1.48}$$

c) If $\boldsymbol{n}$ is nonzero everywhere, then the surface is called **smooth**.

Example 1: Let $I_\varphi = [0, 2\pi]$, $I_\theta = [0, 2\pi]$, and consider

$$\boldsymbol{r}(\varphi, \theta) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta) \tag{1.49}$$

which is a parameterization of S_2 in $\mathbb{R}^3$. Then

$$\partial_\varphi \boldsymbol{r} = (-\sin\theta\sin\varphi, \sin\theta\cos\varphi, 0) , \tag{1.50}$$

$$\partial_\theta \boldsymbol{r} = (\cos\theta\cos\varphi, \cos\theta\sin\varphi, -\sin\theta) , \tag{1.51}$$

Thus

$$\boldsymbol{n}(\varphi, \theta) = \left(-\sin^2\theta\cos\varphi, -\sin^2\theta\sin\varphi, -\sin\theta\cos\theta\right) . \tag{1.52}$$

Remark 1*:* $|\boldsymbol{n}(\varphi, \theta)| = |\sin\theta|$.

Remark 2*:* S_2 is not smooth but S_2 minus its poles is smooth.

§ 2.8 Line and surface integrals

Definition 1:

a) Let $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^n$ be a function, and let $\mathcal{C}$ be a smooth path in $\mathbb{R}^n$ with parameterization $\boldsymbol{q}(t)$ and tangent vector $\boldsymbol{\tau}(t) = \dot{\boldsymbol{q}}(t)$. Then the **line integral** of $\boldsymbol{f}$ over $\mathcal{C}$ is defined as

$$\int_{\mathcal{C}} d\boldsymbol{l} \cdot \boldsymbol{f} := \int_{t_-}^{t_+} dt\, \boldsymbol{\tau}(t) \cdot \boldsymbol{f}(\boldsymbol{q}(t)) , \tag{1.53}$$

with **measure** $d\boldsymbol{l} \equiv \boldsymbol{\tau}(t)\, dt$.

b) The **length** of $\mathcal{C}$ is defined as

$$l(\mathcal{C}) := \int_{t_-}^{t_+} dt\ |\boldsymbol{\tau}(t)| . \tag{1.54}$$

Remark 1: This is consistent with the intuitive concept of the length of a curve, see §2.6 Example 1.

Remark 2: Integration over a closed curve $\mathcal{C}$ is denoted by $\oint_{\mathcal{C}} d\boldsymbol{l}$.

Definition 2:

a) Let $\boldsymbol{f} : \mathbb{R}^3 \to \mathbb{R}^3$ be a function and let S be a smooth surface in $\mathbb{R}^3$ with parameterization $\boldsymbol{r}(t,u)$ and standard normal vector $\boldsymbol{n}(t,u)$. The **surface integral** of $\boldsymbol{f}$ over S is defined as

$$\int_S d\boldsymbol{\sigma} \cdot \boldsymbol{f} := \int_{t_-}^{t_+} dt \int_{u_-}^{u_+} du\, \boldsymbol{n}(t,u) \cdot \boldsymbol{f}(\boldsymbol{r}(t,u)), \tag{1.55}$$

with measure $d\boldsymbol{\sigma} \equiv \boldsymbol{n}(t,u)\, dtdu$.

b) The **area** of S is defined as

$$A(S) := \int_{t_-}^{t_+} dt \int_{u_-}^{u_+} du\, |\boldsymbol{n}(t,u)|. \tag{1.56}$$

Example 1: Let S be the 2-sphere S_2. From §2.7, $|\boldsymbol{n}(\theta,\varphi)| = |\sin\theta|$, so that

$$\begin{aligned} A(S_2) &= \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta\, |\sin\theta| \\ &= 2\pi \int_{-1}^{1} d\cos\theta \\ &= 4\pi. \end{aligned} \tag{1.57}$$

Remark 3: A flat surface can be parameterized by the Cartesian coordinates of its points (see Figure 1.5):

$$\begin{aligned} & \boldsymbol{r}(x,y,z) = (x,y,0) \\ \Longrightarrow\quad & \boldsymbol{n}(x,y) = \begin{pmatrix}1\\0\\0\end{pmatrix} \times \begin{pmatrix}0\\1\\0\end{pmatrix} = \begin{pmatrix}0\\0\\1\end{pmatrix} \\ \Longrightarrow\quad & A(S) = \int_S dxdy. \end{aligned} \tag{1.58}$$

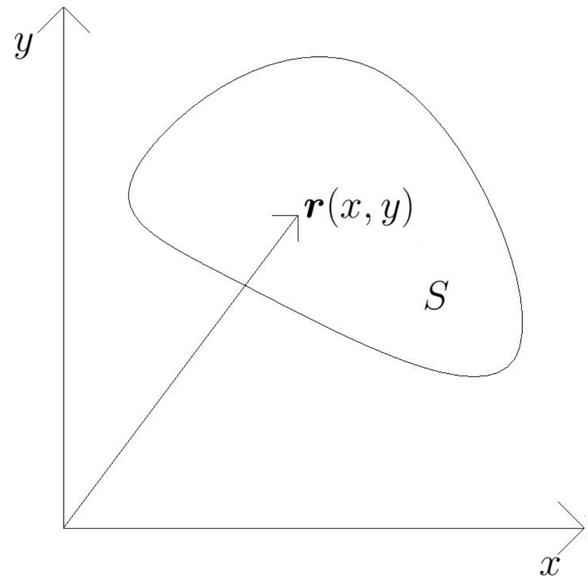

Figure 1.5: A parameterized flat surface.

Problem 7: Enclosed area

Consider a closed curve $\mathcal{C}$ in $\mathbb{R}^2$ with parameterization $\boldsymbol{q}(t) = (x(t), y(t))$. Show that the area A enclosed by $\mathcal{C}$ can be written

$$A = \frac{1}{2}\int_{\mathcal{C}} dt\,[x(t)\,\dot{y}(t) - y(t)\,\dot{x}(t)]\,.$$

Hint: Start from §1.6 Remark 5. Find a function $\boldsymbol{f} : \mathbb{R}^3 \to \mathbb{R}^3$ with the property $\boldsymbol{n}\cdot(\nabla\times\boldsymbol{f}) \equiv 1$, where $\boldsymbol{n}$ is the normal vector for the area enclosed by $\mathcal{C}$. Then use Stokes's theorem.

Theorem 1 (**Gauss's theorem**)**:** Let $V \subset \mathbb{R}^3$ be a volume with smooth surface (V), and let $\boldsymbol{f} : \mathbb{R}^3 \to \mathbb{R}^3$ be a function. Then

$$\int_V dV\,\nabla\cdot\boldsymbol{f} = \int_{(V)} d\boldsymbol{\sigma}\cdot\boldsymbol{f}, \tag{1.59}$$

with $dV = dxdydz$ the measure of the volume integral.

Proof: MATH 281,2 or equivalent □

Theorem 2 (**Stokes's theorem**)**:** Let S be a surface in $\mathbb{R}^3$ bounded by a smooth curve (S). Then,

$$\int_S d\boldsymbol{\sigma}\cdot(\nabla\times\boldsymbol{f}) = \oint_{(S)} d\boldsymbol{l}\cdot\boldsymbol{f}. \tag{1.60}$$

Proof: MATH 281,2 or equivalent □

§ 2.9 The implicit function theorem

Suppose we want to define a function $y(x)$ locally by an implicit relation $F(x, y) = 0$.

Theorem 1: For $i = 1, \ldots, n$, let each $F^i : \mathbb{R}^{m+n} \to \mathbb{R}^n$ be a function of $m + n$ variables

$$F^i(\boldsymbol{x}, \boldsymbol{y}) := F^i(x_1, \ldots, x_m, y_1, \ldots, y_n) \tag{1.61}$$

such that

a) F^i is continuously differentiable in neighborhoods $N^m(\boldsymbol{x}_0) \subset \mathbb{R}^m$, $N^n(\boldsymbol{y}_0) \subset \mathbb{R}^n$,

b) $F^i(\boldsymbol{x}_0, \boldsymbol{y}_0) = 0 \quad \forall i$, and

c) $\det(F_{ij})|_{\boldsymbol{x}_0, \boldsymbol{y}_0} \neq 0$ with $F_{ij} \equiv \partial F^i / \partial y_j$ an $n \times n$ matrix.

Then there exist n functions $f^i : \mathbb{R}^m \to \mathbb{R}$

$$f^i(x_1, \ldots, x_m) \equiv f^i(\boldsymbol{x}) \tag{1.62}$$

such that

a) f^i is unique and continuously differentiable in some neighborhood $\tilde{N}^m(\boldsymbol{x}_0) \neq \emptyset$,

b) $f^i(\boldsymbol{x}_0) = y_i^0 \quad \forall i$, and

c) $F^i(x_1, \ldots, x_m, f_1(x_1, \ldots, x_m), \ldots, f_n(x_1, \ldots, x_m)) = 0 \quad \forall \boldsymbol{x} \in \tilde{N}^m(\boldsymbol{x}_0)$.

Proof: The idea of the proof is as follows: If for a given $\boldsymbol{x}$ there is a $\boldsymbol{y}$ such that $F(\boldsymbol{x}, \boldsymbol{y}) = 0$, then for $\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{\epsilon}$ one can find a $\boldsymbol{y}' = \boldsymbol{y} + \boldsymbol{\delta}$ such that $F(\boldsymbol{x}', \boldsymbol{y}') = 0$ provided that everying is sufficiently well-behaved. See an analysis course for how to implement this idea. □

Example 1: Let $F(x, y) = x - y^2$.

a) If $x_0 = y_0 = 1$, then $\partial F / \partial y|_{x_0, y_0} = -2 \neq 0$, so $f(x) = \sqrt{x}$ is unique and continuously differentiable for $x > 0$.

b) If $x_0 = y_0 = 0$, then $\partial F / \partial y|_{x_0, y_0} = 0$, so $f(x) = \pm\sqrt{x}$ is neither unique nor continuously differentiable.

Corollary 1: With $i = 1, \ldots, n$, let each $F^i(\boldsymbol{q})$ be a function of $n+m$ variables $\boldsymbol{q} = (q_1, \ldots, q_{n+m})$ such that

a) F^i is continuously differentiable for $\boldsymbol{q} \in N^{n+m}(\boldsymbol{q}_0)$,

b) $F^i(\boldsymbol{q}_0) = 0 \quad \forall i$, and

c) $\text{rank}\left(\partial F^i / \partial q_j\right)\Big|_{\boldsymbol{q}_0} = n$.

Then it is possible to choose m variables

$$x_1 \equiv q_{i_1}, \ldots, x_m \equiv q_{i_m} \tag{1.63}$$

and find n continuously differentiable functions $f^j(\boldsymbol{x})$ such that

1. $F(\boldsymbol{x}, \boldsymbol{f}(\boldsymbol{x})) = 0$, and
2. $f^1(\boldsymbol{x}_0) = q^0_{j_1}, \ldots, f^n(\boldsymbol{x}_0) = q^0_{j_n}$,

where $(i_1, \ldots, i_m, j_1, \ldots j_n)$ is a permutation of $(1, \ldots, m+n)$.

Proof: Since rank $\left(\partial F^i / \partial q_j\right) = n$, $\boldsymbol{x}$ and $\boldsymbol{y}$ can be chosen such that Theorem 1 applies. □

Remark 1: consider $n+m$ variables $q_1, \ldots, q_{n+m}$ with n independent constraints. Then n variables can be eliminated using $y_1 = q_{j_1}, \ldots, y_n = q_{j_n}$ to obtain m independent variables $x_1 = q_{i_1}, \ldots, x_m = q_{j_m}$.

Example 2: Let $n = 1$, $m = 2$, $F(q_1, q_2, q_3) = q_1^2 + q_2^2 + q_3^2 - 1$ (which is a 2-sphere, see §2.3 Example 2). Consider the point $(0, 0, 1)$. We have

$$F(0,0,1) = 0, \tag{1.64}$$

$$\left.\frac{\partial F}{\partial q_1}\right|_{(0,0,1)} = \left.\frac{\partial F}{\partial q_2}\right|_{(0,0,1)} = 0, \tag{1.65}$$

$$\left.\frac{\partial F}{\partial q_3}\right|_{(0,0,1)} = 2 \neq 0. \tag{1.66}$$

Chose

$$x = q_1 \quad , \quad y = q_2, \tag{1.67}$$

and a function

$$f(x,y) = \sqrt{1 - x^2 - y^2}. \tag{1.68}$$

$f(x,y)$ is continuously differentiable for $x^2 + y^2 < 1$, so $z = q_3$ can be eliminated and expressed in terms of x and y.

§ 3 Calculus of variations

§ 3.1 Three classic problems

Consider the following three classic problems:

a) **The brachistochrone problem** (John Bernoulli, 1696)
A massive particle moves under the force of gravity from A to B along a curve $\mathcal{C}$. Which curve $\mathcal{C}$ yields the shortes passage time?

b) **The geodesic problem** (John Bernoulli 1697)
Two points A and B on a 2-sphere S_2 (or any manifold) are connected by a curve $\mathcal{C} \subset S_2$. Which curve $\mathcal{C}$ has the shortest length?

c) **The isoperimetric problem** (Pappus of Alexandria, 2nd century CE, and Jacob Bernoulli)
Consider a closed curve $\mathcal{C} \subset \mathbb{R}^2$ with fixed length l. For which curve $\mathcal{C}$ will that curve enclose the largest area?

Remark 1: Each classic problem above is an extremal value problem that ask for the extremum of a functional under the variation of a function.

Remark 2: Problem (c) involves a constraint.

Remark 3: Extrema of functions under variation of a number was discussed eariler, in §2.3. Extrema of functionals under variation of a function requires a new mathematical apparatus, the **calculus of variations**.

§ 3.2 Path neighborhoods

Consider intervals $I_t = [t_-, t_+] \subset \mathbb{R}$, $I_\varepsilon = [\varepsilon_-, \varepsilon_+] \subset \mathbb{R}$, and a function $\boldsymbol{q} : I_\varepsilon \times I_t \to \mathbb{R}^n$, with $\boldsymbol{q}_\varepsilon(t) = (q_1^\varepsilon(t), \ldots, q_2^\varepsilon(t))$ such that

i) $\boldsymbol{q}_\varepsilon(t)$ parametrizes a path $\mathcal{C}_\varepsilon$ for any fixed $\varepsilon \in I_\varepsilon$.

ii) $\boldsymbol{q}_{\varepsilon=0}(t) \equiv \boldsymbol{q}(t)$ parametrizes a path $\mathcal{C} \equiv \mathcal{C}_{\varepsilon=0.}$

iii) $\boldsymbol{q}_\varepsilon(t_\pm) = \boldsymbol{q}(t_\pm) \quad \forall \varepsilon \in I_\varepsilon$.

iv) $\boldsymbol{q}_\varepsilon(t)$ is continuously differentiable with respect to ε for any fixed $t \in I_t$.

Definition 1: Under the above conditions, the set of paths $\{\mathcal{C}_\varepsilon | \varepsilon \in I_\varepsilon\}$ forms an ε-neighborhood of the path $\mathcal{C}$, and $\boldsymbol{q}_\varepsilon(t)$ parametrizes that neighborhood.

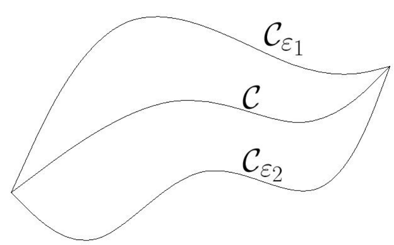

Figure 1.6: Path neighborhoods.

Definition 2:

a) Let L be a function and let $\mathcal{S}_L$ be a functional according to §2.6 Definition 2. Then, $\mathcal{S}_L(\mathcal{C})$ is called **stationary**, and $\mathcal{C}$ is called an **extremal** of $\mathcal{S}_L$, if

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[\mathcal{S}_L(\mathcal{C}_\varepsilon) - \mathcal{S}_L(\mathcal{C})\right] = 0. \tag{1.69}$$

b) $\mathcal{S}_L$ is called a **minimal** (or **maximal**) under variations of the path $\mathcal{C}$ if there exists an $\varepsilon > 0$ such that $\mathcal{S}_L(\mathcal{C}_\varepsilon) > \mathcal{S}_L(\mathcal{C})$ (or $\mathcal{S}_L(\mathcal{C}_\varepsilon) < \mathcal{S}_L(\mathcal{C})$) for all paths $\mathcal{C}_\varepsilon$ in an ε-neighborhood of $\mathcal{C}$.

Proposition 1: Stationarity of $\mathcal{S}_L$ is a necessary condition for $\mathcal{S}_L$ to be either minimal or maximal.

Proof: Once the neighborhood has been parameterized, $\mathcal{S}_L(\mathcal{C})$ can be considered a function $f : \mathbb{R} \to \mathbb{R}$ of ε. Equation (1.69) can then be written as $df/d\varepsilon = 0$, and the assertion then follows from §2.3. □

§ 3.3 A fundamental lemma

Lemma 1: Let $I = [t_-, t_+] \subset \mathbb{R}$, and let $f : I \to \mathbb{R}$ be a continuous function. If

$$\int_{t_-}^{t_+} dt\, \eta(t)\, f(t) = 0 \tag{1.70}$$

for every function η that is continuously differentiable on I and vanishes at $t = t_\mp$, then

$$f(t) = 0 \quad \forall\, t \in I. \tag{1.71}$$

Proof: Suppose that Eq. (1.71) did not hold, so that $\exists\, t^* \in I : f(t^*) \neq 0$. Then, either $f(t^*) > 0$ or $f(t^*) < 0$.

Case 1. $f(t^*) > 0$. Continuity requires that $\exists \left[t_-^*, t_+^*\right] = I^* \subset I$ with $t^* \in I^*$ such that $f(t > 0)\, \forall t \in I^*$. Define

$$\eta(t) = \begin{cases} \left(t - t_-^*\right)^2 \left(t_+^* - t\right)^2, & t \in I^* \\ 0 & t \notin I^* \end{cases}. \tag{1.72}$$

which fulfills the requirements stated in the Lemma. Then

$$\begin{aligned} \int_{t_-}^{t_+} dt\, \eta(t)\, f(t) &= \int_{t_-^*}^{t_+^*} dt\, \underbrace{\eta(t)\, f(t)}_{>0} \\ &> 0, \end{aligned} \tag{1.73}$$

so Eq. (1.70) does not hold. The supposition was thus false, and hence Eq. (1.71) must hold.

Case 2. $f(t^*) < 0$. The set of arguments for this case mirror that of Case 1 exactly. □

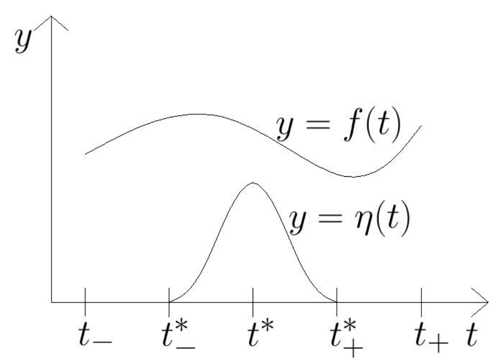

Figure 1.7: Proving Lemma 1.

Corollary 1: Let $\boldsymbol{f} : I \to \mathbb{R}^n$ be a continuous function, and let

$$\begin{aligned} \int_{t_-}^{t_+} dt\, \boldsymbol{\eta}(t) \cdot \boldsymbol{f}(t) &= \int_{t_-}^{t_+} dt\, \eta_i(t)\, f^i(t) \\ &= 0 \end{aligned} \tag{1.74}$$

for every continuously differentiable function $\boldsymbol{\eta} : I \to \mathbb{R}^n$ that vanishes at $t = t_{\mp}$. Then,

$$\boldsymbol{f}(t) = \boldsymbol{0} \quad \forall t \in I. \tag{1.75}$$

Proof: Problem 8. □

Remark 1: The lemma and corollary above remain true if we require η (or $\boldsymbol{\eta}$) to be n times continuously differentiable. See Problem 9.

§ 3.4 The Euler-Lagrange equations

Theorem 1: Let $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ be a function and let $\boldsymbol{q}(t)$ be a parameterization of the path $\mathcal{C}$. Then the functional

$$\mathcal{S}_L(\mathcal{C}) = \int_{t_-}^{t_+} dt\, L(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t), t) \tag{1.76}$$

is stationary, i.e., $\mathcal{C}$ is an extremal path of $\mathcal{S}_L$, if and only if $\boldsymbol{q}(t)$ obeys the **Euler-Lagrange equations**,

$$\boxed{\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i} \quad , \quad i = (1, \ldots, n).} \tag{1.77}$$

Proof: Let $\{\mathcal{C}_\varepsilon\}$ be a neighborhood of $\mathcal{C}$, and let $\boldsymbol{q}_\varepsilon(t)$ be a parameterization of $\mathcal{C}_\varepsilon$. Then,

$$\begin{aligned}
\frac{1}{\varepsilon}\left[\mathcal{S}_L(\mathcal{C}_\varepsilon) - \mathcal{S}_L(\mathcal{C})\right] &= \frac{1}{\varepsilon}\int_{t_-}^{t_+} dt \left[L(\boldsymbol{q}_\varepsilon(t), \dot{\boldsymbol{q}}_\varepsilon(t), t) - L(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t), t)\right] \\
&= \frac{1}{\varepsilon}\int_{t_-}^{t_+} dt \left[L(\boldsymbol{q}_{\varepsilon=0}(t), \dot{\boldsymbol{q}}_{\varepsilon=0}(t), t) + \sum_{i=1}^{n} \left.\frac{\partial L}{\partial q_i}\right|_{\varepsilon=0} \left(q^i_\varepsilon(t) - q^i(t)\right)\right. \\
&\quad \left. + \sum_{i=1}^{n} \left.\frac{\partial L}{\partial \dot{q}^i_\varepsilon}\right|_{\varepsilon=0} \left(\dot{q}^i_\varepsilon(t) - \dot{q}^i(t)\right) + O\left(\varepsilon^2\right) - L(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t), t)\right] \\
&= \int_{t_-}^{t_+} dt \sum_{i=1}^{n} \left[\frac{\partial L}{\partial q_i} \underbrace{\frac{1}{\varepsilon}\left(q^i_\varepsilon(t) - q^i(t)\right)}_{\eta^i_\varepsilon(t) \xrightarrow{\varepsilon\to 0} \eta_i(t)} + \frac{\partial L}{\partial \dot{q}_i} \underbrace{\frac{1}{\varepsilon}\left(\dot{q}^i_\varepsilon(t) - \dot{q}^i(t)\right)}_{\dot{\eta}^i_\varepsilon(t) \xrightarrow{\varepsilon\to 0} \dot{\eta}_i(t)}\right] + O(\varepsilon) \\
&= \int_{t_-}^{t_+} dt \sum_{i=1}^{n} \left[\frac{\partial L}{\partial q_i}\eta_i(t) + \frac{\partial L}{\partial \dot{q}_i}\dot{\eta}_i(t)\right] + O(\varepsilon) \\
&= \int_{t_-}^{t_+} dt \sum_{i=1}^{n} \left[\frac{\partial L}{\partial q_i}\eta_i(t) - \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right)\eta_i(t)\right] + \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i}\eta_i(t)\Big|_{t_-}^{t_+} + O(\varepsilon) \\
&= \int_{t_-}^{t_+} dt \sum_{i=1}^{n} \left[-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} + \frac{\partial L}{\partial q_i}\right]\eta_i(t) + O(\varepsilon).
\end{aligned} \tag{1.78}$$

where the second line follows from Taylor-expanding about $\varepsilon = 0$ (see §2.3). In the limit

$\varepsilon \to 0$, which must be 0 since $\mathcal{C}_\varepsilon$ approaches $\mathcal{C}$, note that

$$\begin{aligned} & \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \left[\mathcal{S}_L\left(\mathcal{C}_\varepsilon\right) - \mathcal{S}_L\left(\mathcal{C}\right)\right] = 0 \\ \Longleftrightarrow \quad & \int_{t_-}^{t_+} dt \sum_{i=1}^{n} \left[-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} + \frac{\partial L}{\partial q_i}\right] \eta_i\left(t\right) = 0. \end{aligned} \tag{1.79}$$

But the path variations $\boldsymbol{q}_\varepsilon$ are arbitrary, so that the function $\boldsymbol{\eta}$ is arbitrary. Then, from the §3.3 corollary, we must have

$$-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} + \frac{\partial L}{\partial q_i} = 0 \quad \forall i, \tag{1.80}$$

as desired. □

Remark 1*:* Equation 1.77 gives a set of n coupled ODEs for the components of $\boldsymbol{q}$.

Remark 2*:* Using the chain rule to write equation 1.77 more explicitly gives

$$\sum_{j=1}^{n} \frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j}\ddot{q}_j + \sum_{j=1}^{n} \frac{\partial^2 L}{\partial \dot{q}_i \partial q_j}\dot{q}_j + \frac{\partial^2 L}{\partial \dot{q}_i \partial t} = \frac{\partial L}{\partial q_i}. \tag{1.81}$$

Defining a symmetric matrix M as

$$\begin{aligned} M_{ij}\left(\boldsymbol{q}, \dot{\boldsymbol{q}}, t\right) &\equiv \frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j} \\ &= M_{ji}, \end{aligned} \tag{1.82}$$

and defining a function

$$F_i\left(\boldsymbol{q}\dot{\boldsymbol{q}}, t\right) \equiv \frac{\partial L}{\partial q_i} - \sum_{j=1}^{n} \frac{\partial^2 L}{\partial \dot{q}_i \partial q_j}\dot{q}_j - \frac{\partial^2 L}{\partial \dot{q}_i \partial t}, \tag{1.83}$$

Equation 1.77 can be rewritten as

$$\sum_{j=1}^{n} M_{ij}\ddot{q}_j = F_i. \tag{1.84}$$

Written in this form, the equation has the form of Newton's second Law.

Remark 3*:* In general, second-order ODEs are not integrable in closed form. It is worth-while to study special cases that are integrable, but most books on mechanics give the impression that there are only these special cases.

Example 1: We can find the paths $\mathcal{C} \subset \mathbb{R}^2$ between two points (x_-, y_-) and (x_+, y_+) that make the length of $\mathcal{C}$ stationary, i.e., find the **geodesics** in $\mathbb{R}^2$. From §2.6 Example 1, the length of the path is given by

$$\begin{aligned} l_{\mathcal{C}} &= \int_{t_-}^{t_+} dt\, \sqrt{\dot{x}^2(t) + \dot{y}^2(t)} \\ &= \int_{x_-}^{x_+} dx\, \sqrt{1 + (dy/dx)^2} \\ &\equiv \int_{x_-}^{x_+} dx\, \sqrt{1 + y'^2} \\ &= \int_{x_-}^{x_+} dx\, L(y')\,, \end{aligned} \tag{1.85}$$

where $\mathrm{L}(y, y', x) = L(y') = \sqrt{1+y'^2}$. From equation 1.77,

$$\begin{aligned} & \frac{d}{dx}\frac{\partial L}{\partial y'} = 0 \\ \Longrightarrow\quad & \frac{\partial L}{\partial y'} = \text{constant} = \frac{y'}{\sqrt{1+y'^2}} \\ \Longrightarrow\quad & y' = \text{constant} \\ \Longrightarrow\quad & y(x) = c_1 x + c_2, \end{aligned} \tag{1.86}$$

i.e., the extremal paths being sought are straight lines. Given boundary condtions

$$y(x_\mp) = y_\mp, \tag{1.87}$$

the values of the constants c_1 and c_2 can be determined as

$$c_1 = \frac{y_+ - y_-}{x_+ - x_-} \quad , \quad c_2 = \frac{x_+ y_- - x_- y_+}{x_+ - x_-}. \tag{1.88}$$

Remark 4: In Example 1 above, it was tacitly assumed that the extremals can be represented with y a single-valued function of x. For a more general proof, see problem 13.

§ 4 General discussion of the Euler-Lagrange equations

§ 4.1 First integrals of the Euler-Lagrange equations

Proposition 1: Let $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t)$ be independent of q_i for a given $1 \leq i \leq n$. Then,

$$p_i(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \equiv \frac{\partial L}{\partial \dot{q}_i}(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \tag{1.89}$$

is conserved along any extremal path, i.e.,

$$p_i(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t), t) \equiv c_i(\mathcal{C}) \tag{1.90}$$

if $\mathcal{C}$ is extremal, with $c_i(\mathcal{C})$ a constant that is characteristic of $\mathcal{C}$.

Proof:

$$\begin{aligned}\frac{d}{dt}p_i &= \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\\ &= \frac{\partial L}{\partial q_i}\\ &= 0,\end{aligned} \tag{1.91}$$

where the second line follows from the Euler-Lagrange equations. □

Remark 1*:* A variable q_i on which L does not depend is called **cyclic**.

Remark 2*:* $p_i = \partial L / \partial \dot{q}_i$ is called **[the momentum] conjugate** to q_i.

Proposition 2: Let $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = L(\boldsymbol{q}, \dot{\boldsymbol{q}})$ be independent of t. Then,

$$\begin{aligned}H(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) &\equiv \dot{q}_i p_i(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) - L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)\\ &= \dot{q}_i p_i(\boldsymbol{q}, \dot{\boldsymbol{q}}) - L(\boldsymbol{q}, \dot{\boldsymbol{q}})\\ &= H(\boldsymbol{q}, \dot{\boldsymbol{q}})\end{aligned} \tag{1.92}$$

is constant along any extremal path, i.e.,

$$H(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t)) \equiv E(\mathcal{C}) \tag{1.93}$$

if $\mathcal{C}$ is extremal, with $E(\mathcal{C})$ a constant that is characteristic of $\mathcal{C}$.

Proof:

$$\begin{aligned}\frac{dH}{dt} &= \ddot{q}_i p_i + \dot{q}_i \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i - \frac{\partial L}{\partial q_i}\dot{q}_i\\ &= \dot{q}_i\left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i}\right)\\ &= \dot{q}_i\left(\frac{\partial L}{\partial q_i} - \frac{\partial L}{\partial q_i}\right)\\ &= 0,\end{aligned} \tag{1.94}$$

where the third line follows from the Euler-Lagrange equations. □

Remark 3*:* If L is independent of t, then the variational problem is called **autonomous**.

Remark 4*:* H, as defined in the first line of equation 1.92, is called **Jacobi's integral**.

§ 4.2 Example: The brachistochrone problem

Consider the following statement of the **brachistochrone problem**:

> A point mass slides withouth friction on an inclined plane, with inclination angle α, from point P_1 to point P_2. Which path yields the shortest passage time?

The problem can be solved as follows: Recall from §2.6 Example 1 that the length $l_{\mathcal{C}}$ of a path $\mathcal{C}$ parameterized by $\boldsymbol{q}(t) = (x(t), y(t))$ is given by

$$\begin{aligned} l_{\mathcal{C}} &= \int_{t_-}^{t_+} dt \sqrt{\dot{\boldsymbol{q}}^2(t)} \\ &= \int_{t_-}^{t_+} dt \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}. \end{aligned} \tag{1.95}$$

Then, the passage time T is given by

$$T = \int_{t_-}^{t_+} dt \frac{\sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}}{v(x(t), y(t))}. \tag{1.96}$$

From the fact that the point mass is on an inclined plane,

$$z = y \sin \alpha, \tag{1.97}$$

with the coordinates as in Figure 1.8.

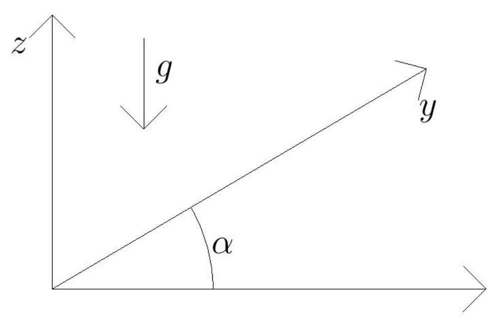

Figure 1.8: The inclined plane for the brachistochrone problem.

Energy conservation can be written as (choosing $y_1 = z_1 = 0$)

$$\begin{aligned} \frac{m}{2} v^2 &= -mgz \\ &= -mgy \sin \alpha, \end{aligned} \tag{1.98}$$

so that

$$\begin{aligned} v(x, y) &= v(y) \\ &= \sqrt{2g \sin \alpha} \sqrt{-y} \\ &= \sqrt{-ay}, \end{aligned} \tag{1.99}$$

where

$$a \equiv 2g \sin \alpha > 0. \tag{1.100}$$

Using x as the parameter, with $y' = dy/dx$,

$$\begin{aligned} T &= \int_{x_1}^{x_2} dx \, \frac{\sqrt{1+y'^2}}{v(y)} \\ &= \int_{x_1}^{x_2} dx \, L(y, y'), \end{aligned} \tag{1.101}$$

where

$$L(y, y') = \frac{\sqrt{1+y'^2}}{v(y)}. \tag{1.102}$$

The problem is autonomous, so that by §2.5 Proposition 2,

$$\begin{aligned} H(y, y') &= y' \frac{\partial L}{\partial y'} - L \\ &= \frac{y'^2}{v(y)\sqrt{1+y'^2}} - \frac{\sqrt{1+y'^2}}{v(y)} \\ &= \text{constant}, \end{aligned} \tag{1.103}$$

so that

$$\begin{aligned} & \frac{1}{v(y)\sqrt{1+y'^2}} = \text{constant} \\ \Longrightarrow \quad & \sqrt{-ay}\sqrt{1+y'^2} = \text{constant} \\ \Longrightarrow \quad & y\left(1+y'^2\right) = \text{constant} \equiv c_1 < 0. \end{aligned} \tag{1.104}$$

Substituting

$$y' = \cot t \quad , \quad t = \cot^{-1}(y') \tag{1.105}$$

gives

$$\begin{aligned} & y = \frac{c_1}{1+\cot^2 t} = c_1 \sin^2 t = \frac{1}{2} c_1 (1 - \cos 2t) \\ \Longrightarrow \quad & dx = \frac{dy}{y'} = \frac{2c_1 \sin t \cos t dt}{\cot t} = 2c_1 \sin^2 t dt = c_1 (1 - \cos 2t) \, dt \\ \Longrightarrow \quad & x = c_2 + c_1 t - \frac{1}{2} c_1 \sin 2t, \end{aligned} \tag{1.106}$$

or

$$x = \frac{1}{2} c_1 (2t - \sin 2t) + c_2. \tag{1.107}$$

With initial conditions

$$x_1 = x(t=0) = 0 \quad , \quad y_1 = y(t=0) = 0, \tag{1.108}$$

equations 1.106 and 1.107 give

$$x(t) = c(2t - \sin 2t) \quad , \quad y(t) = c(1 - \cos 2t), \tag{1.109}$$

where $c = c_1/2 < 0$.

Remark 1*:* Equation 1.109 is a parametric equation for a cycloid with $\gamma = 1$. To see this more clearly, a reparameterization can be performed as $\tau = 2t - \gamma$, $\tilde{x} = -x + \gamma c$, to give

$$\begin{aligned} \tilde{x}(\tau) &= (-c)(\tau + \pi - \sin(\tau + \pi)) + \pi c \\ &= (-c)(\tau + \sin\tau), \end{aligned} \tag{1.110}$$

$$\begin{aligned} y(\tau) &= c(1 - \cos(\tau + \pi)) \\ &= c(1 + \cos\tau) \\ &= (-c)(-1 - \cos\tau). \end{aligned} \tag{1.111}$$

Remark 2*:* The constant c is a scalar factor that can be chosen such that the cycloid contains P_2.

Remark 3*:* If $x_2 > 0$, then we can think of t as being less than 0.

Remark 4*:* For further discussion, see Problem 12.

§ 4.3 Covariance of the Euler-Lagrange equations

Consider a coordinate transformation of the form (the inverse of φ_i below is assumed to exist)

$$Q_i(t) = \varphi_i(\boldsymbol{q}(t), t) \quad , \quad q_i(t) = \varphi_i^{-1}(\boldsymbol{Q}(t), t). \tag{1.112}$$

For instance, a transformation satisfying equation 1.112 in $\mathbb{R}^2$ is

$$\begin{aligned} & x = r\sin\varphi \quad , \quad y = r\cos\varphi \\ \Longleftrightarrow \quad & r = \sqrt{x^2 + y^2} \quad , \quad \varphi = \tan^{-1}(x/y). \end{aligned} \tag{1.113}$$

Question. Are the Euler-Lagrange equations invariant under a coordinate transformation of the form given in equation 1.112?

Consider a path parameterization

$$\boldsymbol{q}(t) = \varphi^{-1}(\boldsymbol{Q}(t), t), \tag{1.114}$$

with tangent vector $\dot{\boldsymbol{q}}(t)$. The componets of the tangent vector are

$$\begin{aligned} \dot{q}_i(t) &= \frac{\partial}{\partial Q_j}\varphi_i^{-1}(\boldsymbol{Q}, t)\dot{Q}_j + \partial_t\varphi_i^{-1}(\boldsymbol{Q}, t) \\ &\equiv \varphi_i^{-1}\left(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, t\right). \end{aligned} \tag{1.115}$$

Defining

$$\Lambda\left(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, t\right) \equiv L\left(\boldsymbol{\varphi}^{-1}(\boldsymbol{Q}, t), \dot{\boldsymbol{\varphi}}^{-1}\left(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, t\right), t\right) \tag{1.116}$$

gives

$$\begin{aligned} \mathcal{S}(\mathcal{C}) &= \int_{t_-}^{t_+} dt\, L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \\ &= \int_{t_-}^{t_+} dt\, \Lambda\left(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, t\right). \end{aligned} \tag{1.117}$$

Recall from §3.4 that

$$
\begin{aligned}
& \mathcal{S} \text{ is stationary} \\
\iff \quad & \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i} \\
\iff \quad & \frac{d}{dt}\frac{\partial \Lambda}{\partial \dot{Q}_i} = \frac{\partial \Lambda}{\partial Q_i}.
\end{aligned} \tag{1.118}
$$

Theorem 1: The Euler-Lagrange equations are **covariant**, i.e., are invariable under coordinate transformations.

Remark 1*:* We don't need to distinguish between L and Λ, so we can write $L\left(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, t\right)$ instead of $\Lambda\left(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, t\right)$.

§ 4.4 Gauge invariance

Definition 1: Let $\mathcal{C}$ be a path with parameterization $\boldsymbol{q}(t)$, and let $G(\boldsymbol{q}, t) \in \mathbb{R}$ be a continuously differentiable function. Then the transformation

$$
\boxed{L_1(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \to L_2(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = L_1(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) + \frac{d}{dt} G(\boldsymbol{q}, t)} \tag{1.119}
$$

is called a **gauge transformation**, with **gauge function** G.

Theorem 1: If L_1 and L_2 are related by a gauge transformation, then any extremal of L_1 is also an extremal of L_2.

Proof:

$$
\begin{aligned}
\mathcal{S}_2 &= \int_{t_-}^{t_+} dt\, L_2(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \\
&= \int_{t_-}^{t_+} dt\, \left[L_1(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) + \frac{d}{dt} G(\boldsymbol{q}, t)\right] \\
&= \int_{t_-}^{t_+} dt\, L_1(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) + \int_{t_-}^{t_+} dt\, \frac{d}{dt} G(\boldsymbol{q}, t) \\
&= \mathcal{S}_1 + G(\boldsymbol{q}(t_+), t_+) - G(\boldsymbol{q}(t_-), t_-) \\
&\equiv \mathcal{S}_1 + \Delta G.
\end{aligned} \tag{1.120}
$$

Under a variation of the path, $t_\pm$ does not change, nor does $\boldsymbol{q}(t_\pm)$. Then, ΔG does not change under a variation of the path. Then, $\mathcal{S}_1 + \Delta G$ is stationary exactly when $\mathcal{S}_1$ is stationary, i.e., $\mathcal{S}_2$ is stationary if and only if $\mathcal{S}_1$ is stationary. □

Remark 1*:* The converse of the theorem is not true.

Remark 2*:* If equation 1.119 holds, then L_1 and L_2 are called **equivalent**.

§ 4.5 Nöther's theorem

Consider a family of coordinate transformations as were described in §4.3,

$$Q^i = \varphi_\alpha^i \left(\boldsymbol{q}, t\right), \tag{1.121}$$

with $\alpha \in \mathbb{R}$ a continuous parameter. Let

$$Q^i = q^i + \alpha f^i \left(\boldsymbol{q}, t\right) + O\left(\alpha^2\right) \quad , \quad \text{for } \alpha \to 0. \tag{1.122}$$

Remark 1*:* We know that $\boldsymbol{\varphi}_\alpha \left(\boldsymbol{q}, t\right)$ is differentiable with respect to α at $\alpha = 0$ - see Taylor's theorem in §2.3.

Theorem 1 (Nöther's theorem)**:** Let $L\left(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, t\right)$ and $L\left(\boldsymbol{q}, \dot{\boldsymbol{q}}, t\right)$ be equivalent except for terms of $O\left(\alpha^2\right)$, i.e., let there be a gauge function $G\left(\boldsymbol{q}, t\right)$ such that

$$\boxed{L\left(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, t\right) = L\left(\boldsymbol{q}, \dot{\boldsymbol{q}}, t\right) + \alpha \frac{d}{dt} G\left(\boldsymbol{q}, t\right) + O\left(\alpha^2\right).} \tag{1.123}$$

Then, the quantity

$$\boxed{j\left(\boldsymbol{q}, \dot{\boldsymbol{q}}, t\right) \equiv p_i\left(t\right) f_i\left(\boldsymbol{q}, t\right) - G\left(\boldsymbol{q}, t\right),} \tag{1.124}$$

with $p_i = \partial L / \partial \dot{q}_i$ conjugate to q_i according to §4.1, is constant for all extremal paths, i.e.,

$$\begin{aligned} j\left(\boldsymbol{q}, \dot{\boldsymbol{q}}, t\right) &\equiv j\left(\mathcal{C}\right) \\ &= \text{constant} \end{aligned} \tag{1.125}$$

if $\mathcal{C}$ is extremal.

Proof:

$$\begin{aligned} L\left(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, t\right) &= L\left(\boldsymbol{q} + \alpha \boldsymbol{f}, \dot{\boldsymbol{q}} + \alpha \dot{\boldsymbol{f}}, t\right) + O\left(\alpha^2\right) \\ &= L\left(\boldsymbol{q}, \dot{\boldsymbol{q}}, t\right) + \alpha \left(\frac{\partial L}{\partial q_i} f_i + \frac{\partial L}{\partial \dot{q}_i} \dot{f}_i\right) + O\left(\alpha^2\right) \\ &= L\left(\boldsymbol{q}, \dot{\boldsymbol{q}}, t\right) + \alpha \frac{d}{dt} G\left(\boldsymbol{q}, t\right) + O\left(\alpha^2\right), \end{aligned} \tag{1.126}$$

where the second line follows from Taylor-expanding the first about $\alpha = 0$, and the third line follows from equation 1.123. Comparing like powers of α (specifically, the α^1 terms),

$$\begin{aligned} \frac{d}{dt} G\left(\boldsymbol{q}, t\right) &= \frac{\partial L}{\partial q_i} f_i + \frac{\partial L}{\partial \dot{q}_i} \dot{f}_i \\ &= \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i}\right) f_i + \frac{\partial L}{\partial \dot{q}_i} \frac{d}{dt} f_i \\ &= \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} f_i\right) \\ &= \frac{d}{dt} \left(p_i f_i\right), \end{aligned} \tag{1.127}$$

where the second line follows from the Euler-Lagrange equations and the fact that $\dot{f}_i \equiv df_i/dt$. Subtracting the LHS from each side,

$$\begin{aligned} 0 &= \frac{d}{dt}\left(p_i f_i - G\right) \\ \Longrightarrow \quad j &\equiv p_i f_i - G = \text{constant}, \end{aligned} \tag{1.128}$$

as desired. □

Remark 2: If equation 1.123 holds, we say that the coordinate transformation $\boldsymbol{Q} = \varphi_\alpha\left(\boldsymbol{q}, t\right)$ represents a **continuous symmetry** of L.

Remark 3: A continuous symmetry implies the existence of an invariant quantity j.

Example 1: Let q_{i_0} be cyclic and consider the transformation

$$Q_i = q_i + \alpha\delta_{ii_0}. \tag{1.129}$$

Then,

$$\begin{aligned} L\left(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, t\right) &= L\left(q_1, \ldots, q_{i_0} + \alpha, \ldots, q_n, \dot{\boldsymbol{q}}, t\right) \\ &= L\left(\boldsymbol{q}, \dot{\boldsymbol{q}}, t\right) + \alpha\frac{\partial L}{\partial q_{i_0}} \\ &= L\left(\boldsymbol{q}, \dot{\boldsymbol{q}}, t\right), \end{aligned} \tag{1.130}$$

so that equation 1.123 applies, with

$$G\left(\boldsymbol{q}, t\right) = 0 \quad , \quad f_i\left(\boldsymbol{q}, t\right) = \delta_{ii_0}. \tag{1.131}$$

Then, the quantity

$$j = p_{i_0} \tag{1.132}$$

is conserved along any extremal path. Note that this result was derived directly from the Euler-Lagrange equations in §4.1 as well.

Example 2: Let $\boldsymbol{q} \in \mathbb{R}^3$, and consider a rotation about the q_3-axis, so that the coordinate transformation is

$$\begin{aligned} Q_1 &= q_1\cos\alpha + q_2\sin\alpha \\ &= q_1 + \alpha q_2 + O\left(\alpha^2\right), \end{aligned} \tag{1.133}$$

$$\begin{aligned} Q_2 &= q_1\sin\alpha + q_2\cos\alpha \\ &= q_2 - \alpha q_2 + O\left(\alpha^2\right). \end{aligned} \tag{1.134}$$

Let L be invariant under such rotations,

$$L\left(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, t\right) = L\left(\boldsymbol{q}, \dot{\boldsymbol{q}}, t\right). \tag{1.135}$$

Then, equation 1.123 is satisfied with

$$G(\boldsymbol{q}, t) = 0 \quad , \quad \boldsymbol{f}(\boldsymbol{q}, t) = (q_2, -q_1, 0). \tag{1.136}$$

Then, the quantity

$$\begin{aligned} j &= p_1 q_2 - p_2 q_1 \\ &= (\boldsymbol{p} \times \boldsymbol{q})_3 \end{aligned} \tag{1.137}$$

is conserved along any extremal path.

Remark 4: In Example 1, the continuous symmetry is translational invariance in the i_0-direction, as given in equation 1.129.

Remark 5: In Example 2, it was found that $(\boldsymbol{p} \times \boldsymbol{q})_3$ is conserved if the Lagrangian is invariant under rotations about the q_3-axis. Similarly, $(\boldsymbol{p} \times \boldsymbol{q})_i$ is conserved if the Lagrangian is invariant under rotations about the q_1-axis. Then, if the Lagrangian is invariant under rotations about all three axes, then the quantity

$$\boldsymbol{l} \equiv -\boldsymbol{p} \times \boldsymbol{q} \tag{1.138}$$

is constant.

Remark 6: Remark 2 also follows directly from the Euler-Lagrange equations. Let

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = L(|\boldsymbol{q}|, |\dot{\boldsymbol{q}}|, t). \tag{1.139}$$

Then,

$$\begin{aligned} & \frac{\partial L}{\partial q_i} = \frac{\partial L}{\partial |\boldsymbol{q}|} \frac{q_i}{|\boldsymbol{q}|} \\ \Longrightarrow \quad & \frac{\partial L}{\partial \boldsymbol{q}} \propto \boldsymbol{q}, \end{aligned} \tag{1.140}$$

$$\begin{aligned} & \frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial |\boldsymbol{q}|} \frac{\dot{q}_i}{|\boldsymbol{q}|} \\ \Longrightarrow \quad & \frac{\partial L}{\partial \dot{\boldsymbol{q}}} \propto \dot{\boldsymbol{q}}. \end{aligned} \tag{1.141}$$

Then,

$$\begin{aligned} \frac{d}{dt}\boldsymbol{l} &= \dot{\boldsymbol{q}} \times \boldsymbol{p} + \boldsymbol{q} \times \dot{\boldsymbol{p}} \\ &= \dot{\boldsymbol{q}} \times \frac{\partial L}{\partial \dot{\boldsymbol{q}}} + \boldsymbol{q} \times \frac{d}{dt}\frac{\partial L}{\partial \dot{\boldsymbol{q}}} \\ &= \dot{\boldsymbol{q}} \times \frac{\partial L}{\partial \dot{\boldsymbol{q}}} + \boldsymbol{q} \times \frac{\partial L}{\partial \boldsymbol{q}} \\ &= 0, \end{aligned} \tag{1.142}$$

where the third line follows from the Euler-Lagrange equations, and the last line follows from the fact that both terms are the cross products of parallel vectors, which can be seen from equations 1.140 and 1.141.

§ 5 The principle of least action

§ 5.1 The axioms of classical mechanics

Definition 1: A system of point masses whose positions are completely determined by specify f real numbers is called a **mechanical system with f degrees of freedom**.

Example 1: One free point mass in $\mathbb{R}^3$. For this system, $f = 3$.

Example 2: A mathematical pendulum that can swing in a plane. For this system, $f = 1$.

Axiom 1. A mechanical system with f degrees of freedom is described by a path in $\mathbb{R}^f$, writeable with parameterization

$$\boldsymbol{q}(t) = (q_1(t), \ldots, q_f(t)). \tag{1.143}$$

- Each q_i is called the **i'th generalized coordinate**.
- Each $\dot{q}_i$ is called the **i'th generalized velocity**.
- The path parameter, t, is called **time**.

Remark 1: Having the path be in $\mathbb{R}^f$ is actually too restrictive, in general. For instance, the appropriate space for a planar pendulum (Example 2) is the 1-sphere S_1, not $\mathbb{R}$. We will ignore such topological coordinates.

Remark 2: Note that the mathematical structures postulated are paths (as opposed to, for example, fields).

Remark 3: This is an *axiom* - questions like, "what is time?" are out of order.

Axiom 2. Every mechanical system is characterized by a function $L(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t), t)$, called the **Lagrangian**. The physical paths minimize the functional

$$\mathcal{S} = \int_{t_-}^{t_+} dt\, L(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t), t), \tag{1.144}$$

where $\mathcal{S}$ is called the **action**.

Remark 4: Axiom 2 is known as the **principle of least action**, or **Hamilton's principle**.

Remark 5: From §3.4 and Axiom 2, a necessary condition for a physical path is that it satisfies the f ODEs

$$\boxed{\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t), t) = \frac{\partial L}{\partial q_i}(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t), t),} \tag{1.145}$$

i.e., that it satisfies the **Euler-Lagrange equations**.

Remark 6*:* The **[generalized] momentum** and **[generalized] force**, respectively, are given by

$$\begin{aligned} \boldsymbol{p}(t) &= (p_1, \ldots, p_f) \\ &\equiv \frac{\partial L}{\partial \dot{\boldsymbol{q}}}, \end{aligned} \tag{1.146}$$

and

$$\boldsymbol{F}(t) = \frac{d}{dt}\boldsymbol{p}(t). \tag{1.147}$$

Then, equation 1.145 can be rewritten as

$$\boldsymbol{F}(t) = \frac{d}{dt}\boldsymbol{p}(t), \tag{1.148}$$

which is **Newton's second law**.

Remark 7*:* From §2.4 Remark 3, the initial coordinates $\boldsymbol{q}(t_-)$ and the initial velocities $\dot{\boldsymbol{q}}(t_-)$ completely determine the path.

§ 5.2 Conservation laws

Definition 1: A quantity that is independent of time along the physical path is called a **constant of motion**, or a **conserved quantity**.

Proposition 1: If the Lagrangian is independent of a coordinate q_i, then that coordinate's conjugate momentum p_i is conserved.

Proof: From Newton's second law,

$$\begin{aligned} & F_i = \frac{\partial L}{\partial q_i} = 0 \\ \Longrightarrow \quad & \frac{d}{dt}p_i = 0. \end{aligned} \tag{1.149}$$

□

Remark 1*:* From §4.4,

$$\begin{aligned} q_i \text{ is cyclic} \quad &\Longrightarrow \quad \text{translational invariance} \\ &\Longrightarrow \quad p_i \text{ is conserved.} \end{aligned} \tag{1.150}$$

This is an example of Nöther's theorem.

Remark 2*:* The fact that q_i is cyclic is sufficient, but not necessary, for p_i to be constant. Nöther's theorem is more general, and deeper, than the first integral version of the statement above and in §4.1.

Remark 3*:* In the above proof, **Newton's first law**, or the **law of inertia**, was effectively found. It is writeable as

$$\begin{aligned} & F_i = 0 \\ \Longrightarrow \quad & p_i = \text{constant.} \end{aligned} \tag{1.151}$$

Definition 2:

a) A system whose Lagrangian has no explicit time dependence, so that

$$L\left(\boldsymbol{q}, \dot{\boldsymbol{q}}, t\right) = L\left(\boldsymbol{q}, \dot{\boldsymbol{q}}\right), \tag{1.152}$$

is called **conservative**.

b) In mechanics, Jacobi's integral is called **energy**.

Proposition 2: For any conservative system, energy is conserved.

Proof: The proof is given in §4.1 Proposition 2. □

Remark 4*:* Time translational invariance implies energy conservation. This is another example of Nöther's theorem, or rather, a generalization of Nöther's theorem (see problem 14).

Definition 3: Consider $f = 3$ and Cartesian coordinates,

$$\begin{aligned} \boldsymbol{q} &\equiv \boldsymbol{x} \\ &= \left(x_1, x_2, x_3\right). \end{aligned} \tag{1.153}$$

Then,

$$\boldsymbol{l}\left(t\right) \equiv \boldsymbol{x}\left(t\right) \times \boldsymbol{p}\left(t\right) \tag{1.154}$$

is called the **angular momentum**.

Proposition 3: For a system with $f = 3$ that is invariant under rotations of $\boldsymbol{x}$, angular momentum is conserved.

Proof: The proof is given in §4.4 Example 2. □

§ 6 Problems for Chapter 1

1. Total derivative

Consider §2.2 Example 1: Let $\boldsymbol{x} : \mathbb{R} \to \mathbb{R}^2$ be a function defined by

$$\boldsymbol{x}\left(t\right) = \left(a\cos t, b\sin t\right) \quad , \quad a, b \in \mathbb{R}, \tag{1.155}$$

and let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function defined by

$$f\left(\boldsymbol{x}\right) = \sqrt{x_1^2 + x_2^2}. \tag{1.156}$$

Discuss the behavior of the total derivative, df/dt, and give a geometric interpretation of the result. *Hint:* First determine the geometric figure in $\mathbb{R}^2$ that $\boldsymbol{x}$ provides a parametric representation of, then consider the geometric meaning of $\left|\boldsymbol{x}\left(t\right)\right| = f\left(\boldsymbol{x}\left(t\right)\right)$.

2. Extrema subject to constraints

Consider §2.3 Example 1: Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a function defined by

$$f(x, y, z) = x - y + z. \tag{1.157}$$

Let $S_2 = \left\{(x, y, z) \,|x^2 + y^2 + z^2 = 1\right\}$ be the 2-sphere.

a) Show that the extremal points for f on S_2 are $(1, -1, 1)/\sqrt{3}$, and $(-1, 1, -1)/\sqrt{3}$, as claimed in the lecture.

b) Prove which of these extremal points, if any, are a maximum or a minimum, and determine the respective extremal values of f on S_2.

3. System of ODEs

Solve the system of first order ODEs considered in §2.4 Example 1:

$$\dot{x} = 2x + 4y + 2, \tag{1.158}$$

$$\dot{y} = x - y + 4. \tag{1.159}$$

4. Minimal distance

Consider the 2-sphere, $S_2 = \left\{(x, y, z) \,|x^2 + y^2 + z^2 = 1\right\}$ embedded in $\mathbb{R}^3$. Find the point on S_2 that is closest to the point $(1, 1, 1) \in \mathbb{R}^3$, and determine the distance between the two points.

5. Passage time

Consider a path $\mathcal{C}$ in $\mathbb{R}^2$ with a parameterization $\boldsymbol{q}(t)$, and a point mass moving along $\mathcal{C}$ with speed $v(\boldsymbol{q})$. Let $T(\mathcal{C})$ be the passage time of the particle from $\boldsymbol{q}_-$ to $\boldsymbol{q}_+$. Find the function $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ such that the functional $S_L(\mathcal{C})$ is equal to $T(\mathcal{C})$.

6. Get by with a little help from your Friend™

Spring-loaded camming devices, also known as Friends™ or Camalots™, depending on the manufacturer, are used by rock climbers to protect the climber in case of a fall. The devices consist of four metal wedges that pairwise rotate against one another so that the outside edge of each pair of wedges moves along a curve. The cam is placed in a crack with parallel walls, where the springs hold it in place. The *camming angle* α is defined as the angle between the line from the center of rotation to the contact point with the rock and the tangent to the curve in the contact point. The figure below shows the camming angle for an almost-extended cam in a wide crack, and a largely retracted cam in a narrow crack.

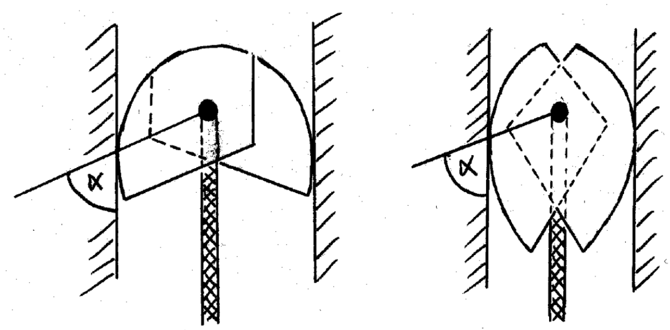

Figure 1.9

If you fall, you want to get as much help from you Friend^{TM} as the laws of physics let you. To ensure that, you want the camming angle to be the same irrespective of the width of the crack (see the figure). Determine the shape of the curve the cam surfaces must form to ensure that is the case.

a) Parametrize the curve, using polar coordinates, as $(r(t), \varphi(t))$. Find the tangent vector in the point $P = (r, \varphi)$ in Cartesian coordinates.

b) Define the angle β between the tangent in P and the line through P that is perpendicular to the radius vector from the origin to P. How is β related to α? Show that

$$\tan \beta = (dr/d\varphi)/r. \tag{1.160}$$

c) Solve the differential equation that results from requiring that $\beta =$ constant along the curve. Discuss the solution, which is the desired shape of the cam.

7. Enclosed area

Consider a closed curve $\mathcal{C}$ in $\mathbb{R}^2$ with parameterization $\boldsymbol{q}(t) = (x(t), y(t))$. Show that the area A enclosed by $\mathcal{C}$ can be written

$$A = \frac{1}{2} \int_{\mathcal{C}} dt \left[x(t)\, \dot{y}(t) - y(t)\, \dot{x}(t) \right]. \tag{1.161}$$

Hint: Start from §1.6 Remark 5. Find a function $\boldsymbol{f} : \mathbb{R}^3 \to \mathbb{R}^3$ with the property $\boldsymbol{n} \cdot (\nabla \times \boldsymbol{f}) \equiv 1$, where $\boldsymbol{n}$ is the normal vector for the area enclosed by $\mathcal{C}$. Then use Stokes's theorem.

8. Corollary to the Basic Lemma

Prove §3.3 Corollary 1: Let $\boldsymbol{f} : I \to \mathbb{R}^n$ be a continuously differentiable function, and

$$\int_{t_-}^{t_+} dt\, \boldsymbol{\eta}(t) \cdot \boldsymbol{f}(t) = 0 \tag{1.162}$$

for every continuously differentiable function $\boldsymbol{\eta} : I \to \mathbb{R}^n$ that obeys $\boldsymbol{\eta}(t_-) = \boldsymbol{\eta}(t_+) = 0$. Then

$$\boldsymbol{f}(t) = \boldsymbol{0} \quad \forall t \in I. \tag{1.163}$$

9. The Basic Lemma revisited

Prove the following stronger version of §3.3 Lemma 1: Let $I = [t_-, t_+]$ be a real interval, and let $f : I \to \mathbb{R}$ be a continuous function. If

$$\int_{t_-}^{t_+} dt \eta(t) f(t) = 0 \tag{1.164}$$

for every function $\eta(t)$ that is n-times continuously differentiable on I and obeys $\eta(t_-) = \eta(t_+) = 0$, then

$$f(t) = 0 \quad \forall t \in I. \tag{1.165}$$

10. Variational problem with a constraint

The functional

$$S_1(\mathcal{C}) = \int_{t_-}^{t_+} dt\, L_1(q, \dot{q}, t) \tag{1.166}$$

is called stationary under variations of the path $\mathcal{C}$ with constraint

$$S_2(\mathcal{C}) = \int_{t_-}^{t_+} dt\, L_2(q, \dot{q}, t) = \text{constant} \tag{1.167}$$

if $\lim_{\epsilon\to 0} [S_1(\mathcal{C}_\epsilon) - S_1(\mathcal{C})]/\epsilon = 0$ for all paths $\mathcal{C}_\epsilon$ in an ϵ-neighborhood of $\mathcal{C}$ that fulfills the condition given by equation 1.167.

Prove the following theorem: If $\mathcal{C}$ is an extremal of S_1 which satisfies the constraint given by equation 1.167, and if $\mathcal{C}$ is not also an extremal of S_2, then there exists a constant λ (called a Lagrange multiplier) such that $\mathcal{C}$ is an extremal of the functional

$$S_3(\mathcal{C}) = \int_{t_-}^{t_+} dt\, [L_1(q, \dot{q}, t) + \lambda L_2(q, \dot{q}, t)], \tag{1.168}$$

and λ is determined by equation 1.167.
Hint: Consider a set of paths $q^i_{\epsilon_1\epsilon_2}(t) = q^i(t) + \epsilon_1 \eta^i_1(t) + \epsilon_2 \eta^i_2(t)$ with a suitably chosen $\epsilon_2 = f(\epsilon_1)$.

11. Dido's problem

Let an area A in the xy-plane be enclosed by a straight line between two points O and P that are a distance d apart and a continuously differentiable curve $\mathcal{C}$ with end points O and P and fixed length $l > d$. Determine the curve $\mathcal{C}$ that maximizes A.

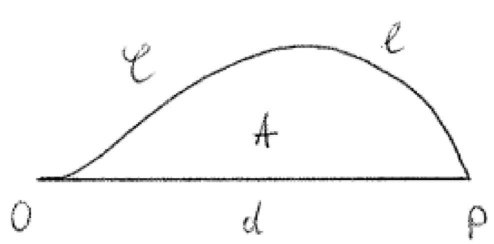

Figure 1.10

Hint: Use the results of Problems 7 and 10.
Note: Pappus of Alexandria knew the answer some 1,700 years ago.

12. Brachistochrone

A ski race is held on an inclined plane with inclination angle α. The skiers glide frictionaless from point P_1 to point P_2. Choose a coordinate system such that the y-axis is along the fall line. The skiers try to find the path of least passage time, or brachistochrone.

a) A skier who has taken a course in theoretical mechanics chooses the brachistochrone for her path. Determine her passage time as a function of her y-velocity at point P_2, $y_2' = (dy/dx)_2$.

b) A mathematically challenged competitor chooses the shortest path from P_1 to P_2. Determine this skier's passage time as a function of the winner's y_2'.

c) Discuss the ratio of the two passage times as a function of the winner's y_2'.

Hint: The parameter value t_2 for the brachistochrone at the end of point P_2 is a known function of y_2'. It therefore is sufficient to discuss the passage times as functions of t_2.

13. Geodesics in $\mathbb{R}^2$ revisited

In §3.4 we found the geodesics in $\mathbb{R}^2$ under the assumption that they can be represented in the form $y = y(x)$. Show that the solution we found also solves the Euler-Lagrange equations that result from the more general parameterization $(x(t), y(t))$.

14. More about Nöther's theorem

a) Prove the following generalization of Nöther's theorem:
Let $\alpha \in \mathbb{R}$ be a continuous parameter for the following set of transformations:

$$Q^i = q^i + \alpha f^i(\boldsymbol{q}, t) + O\left(\alpha^2\right), \tag{1.169}$$

$$T = t + \alpha f^0(t) + O\left(\alpha^2\right). \tag{1.170}$$

Let $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ behave under this transformation as follows:

$$L\left(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, T\right) = L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)\, dt/dT + O\left(\alpha^2\right), \tag{1.171}$$

with $\dot{\boldsymbol{q}} = d\boldsymbol{q}/dt$ and $\dot{\boldsymbol{Q}} = d\boldsymbol{Q}/dT$. Then,

$$J = \sum_{i=1}^{n} \left(\frac{\partial L}{\partial \dot{q}_i}\right) f^i - f^0 \left[\sum_{i=1}^{n} \left(\frac{\partial L}{\partial \dot{q}_i}\right) \dot{q}_i - L\right] \tag{1.172}$$

is constant for all extremal paths.

b) Consider a system that is time time-translationally-invariant, i.e., $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t + t_1) = L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ for all $t_1 \in \mathbb{R}$. Show that this implies the invariance of Jacobi's integral.

15. The principles of Fermat and Jacobi

The following axioms of geometrical optics are known as Fermat's principle:

1) The propagation of light is described by paths in $\mathbb{R}^3$, called *light rays.*

2) An isotropic optical medium is characterized by a real-valued function $n(\boldsymbol{x}) = n(|\boldsymbol{x}|) > 0$. The physical light rays are the paths that minimize the functional

$$S = \int_{\tau_-}^{\tau_+} d\tau \, n(\boldsymbol{x}(\tau)) \sqrt{(\dot{\boldsymbol{x}}(\tau))^2}, \tag{1.173}$$

with $\boldsymbol{x}(\tau)$ a parameterization of the path.

a) Consider two isotropic and homogeneous optical media with indices of refraction equal to n_1 and n_2, respectively, and whose interface is given by the plane $x_3 \equiv z = 0$. Suppose a light ray goes from medium 1 into medium 2. Show that Fermat's principle yields Snell's law.

b) For the same situation as in part (a), suppose that a light ray is reflected at the interface. Show that the angle of reflection is equal to the angle of incidence.

c) Consider a flat earth whose surface is the plane $z = 0$. In the lower layers of the atmosphere, and under certain atmospheric conditions, the expression

$$n(z) = \sqrt{a - bz} \quad , \quad (a, b > 0) \tag{1.174}$$

is a reasonable approximation for the atmospheric index of refraction. Determine and discuss the light rays according to this model, and explain the phenomenon known as mirage.

d) Show that the conservative Lagrangian from §1.5,

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \frac{1}{2} g_{ij}(\boldsymbol{q}) \, \dot{q}_i \dot{q}_j - U(\boldsymbol{q}) \equiv T(\boldsymbol{q}, \dot{\boldsymbol{q}}) - U(\boldsymbol{q}), \tag{1.175}$$

has the same extremals as a Lagrangian

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}) = n(\boldsymbol{q}) \sqrt{T(\boldsymbol{q}, \dot{\boldsymbol{q}})} \tag{1.176}$$

with

$$n(\boldsymbol{q}) = \sqrt{E - U(\boldsymbol{q})}. \tag{1.177}$$

e) Show that the result of part (d) can be rewritten as

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \sqrt{G_{ij}(\boldsymbol{q}) \, \dot{q}_i \dot{q}_j}, \tag{1.178}$$

which is the Lagrangian for geodetics in a Riemannian space with metric G_{ij}. Determine G_{ij}. This statement about the geometrization of mechanics is known as Jacobi's principle.

16. Minimum surface of revolution and the catenary

Consider a curve $\mathcal{C}$ in $\mathbb{R}^2$ that links two points P_1 and P_2. Form a surface of revolution by rotating $\mathcal{C}$ about the x-axis.

a) Determine all smooth curves of the form $y = f(x)$ that make the surface area stationary.
Hint: The resulting curves are known as *catenaries*.

b) Discuss the number of stationary solutions for

$$P_1 = (-x_0, y_0) \quad , \quad P_2 = (x_0, y_0) \tag{1.179}$$

as a function of $\alpha = y_0/x_0$.

Now model a hanging chain as a curve $\mathcal{C}$ in $\mathbb{R}^2$ with length l and a constant linear mass density μ (i.e., the mass of the chain is μl). Suppose that the chain is suspended between two points P_1 and P_2 in a homogeneous gravitational field.

c) Parametrize $\mathcal{C}$ as $(x(t), y(t))$ with Cartesian coordinates x and y, and find the potential energy of the chain as a functional of $\mathcal{C}$.
Hint: First consider an infinitesimal piece of the chain, then integrate.

d) Keeping in mind that the length of the chain is fixed, formulate the isoperimetric problem that minimizes the potential energy.

e) By introducing suitable variables, show that the problem maps onto the problem considered in part (a) and give the general solution. State the boundary conditions that determine the constants of integration in the general solution.

Chapter 2

Mechanics of point masses

§ 1 Point masses in potentials

§ 1.1 A single free particle

Definition 1:

a) A **point mass** or **particle** is a mechanical system with $f = 3$ degrees of freedom.

b) A **free particle** is a particle whose motion is independent of any other system.

Remark 1*:* “Point mass” is an idealized concept. Billiard balls, planets, stars, galaxies, etc. can all be approximated as point masses in appropriate contexts.

Remark 2*:* Recall from Ch.1 §5.1 that it is necessary to postulate the Lagrangian of a free particle.

Axiom 1. There exist coordinate systems and time scales such that the Lagrangian L_0 of a free particle is a function of $\boldsymbol{v}^2 \equiv |\dot{\boldsymbol{x}}|^2$ only:

$$L_0(\boldsymbol{x}, \boldsymbol{v}, t) = L_0(\boldsymbol{v}^2), \tag{2.1}$$

(apart from equivalencies described in Ch.1 §4.3), with $\boldsymbol{v} \equiv \dot{\boldsymbol{x}}$ the **velocity**.

Definition 2: Any such coordinate systems plus the corresponding time scale is called an **inertial system**, abbreviated **IS**.

Remark 3*:* Some plausibility arguments for the existence of inertial systems:

a) A free particle is the only object in an otherwise empty universe.

b) Empty space is translationally invariant, so that $L_0(\boldsymbol{x}, \boldsymbol{v}, t) = L_0(\boldsymbol{x} + \boldsymbol{a}, \boldsymbol{v}, t)\ \forall \boldsymbol{a} \in \mathbb{R}^3$. Choosing $\boldsymbol{a} = -\boldsymbol{x}$ gives $L_0(\boldsymbol{x}, \boldsymbol{v}, t) = L_0(\boldsymbol{v}, t)$.

c) Time in an empty universe is homogeneous, so that $L_0(\boldsymbol{v}, t + t_0) = L_0(\boldsymbol{v}, t) \quad \forall t_0$, so that $L_0(\boldsymbol{v}, t) = L_0(\boldsymbol{v})$.

d) Empty space is isotropic, so that $L_0(\boldsymbol{v}) = L_0(\boldsymbol{v}^2)$.

Remark 4: Emperical evience suggests that these assumptions are reasonable on everyday length and time scales.

Definition 3:

$$m\left(\boldsymbol{v}^2\right) := 2\frac{dL_0}{d\boldsymbol{v}^2} \tag{2.2}$$

is called the **mass** of the free particle.

Axiom 2. The mass is positive-definite,

$$m\left(\boldsymbol{v}^2\right) > 0. \tag{2.3}$$

Proposition 1: The momentum of a free particle is related to its mass and velocity by

$$\boldsymbol{p} = m\left(\boldsymbol{v}^2\right)\boldsymbol{v}. \tag{2.4}$$

Proof:

$$\begin{aligned} \boldsymbol{p} &= \frac{\partial L_0}{\partial \dot{\boldsymbol{x}}} \\ &= \frac{\partial L_0}{\partial \boldsymbol{v}} \\ &= \frac{\partial L_0}{\partial \boldsymbol{v}^2} 2\boldsymbol{v} \\ &= m\left(\boldsymbol{v}^2\right)\boldsymbol{v}. \end{aligned} \tag{2.5}$$

□

Theorem 1 (Newton's first law)**:** The physical paths of free particles in an inertial system are straight lines:

$$\boldsymbol{x}(t) = \boldsymbol{x}_0 + \boldsymbol{v}t, \tag{2.6}$$

with $\boldsymbol{v} = \dot{\boldsymbol{x}} =$ constant.

Proof: $\boldsymbol{x}$ is cyclic, so $\boldsymbol{p} =$ constant. Then, $\boldsymbol{v} = \boldsymbol{p}/m\left(\boldsymbol{v}^2\right) =$ constant. Then, $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{v}t$ with $\boldsymbol{v}$ constant. □

Remark 5: Newton's first law is independent of the functional form of $L_0\left(\boldsymbol{v}^2\right)$.

Remark 6: Newton's first law follows directly from Axiom 1. This demonstrates that Axiom 1 is nontrivial, since coordinate systems exist for which Newton's first law does *not* hold (e.g., a cartesian system fixed to a moving car).

Axiom 3 (Galilean mechanics)**.** The mass of a free particle is a constant, independent of $\mathbf{v}^2$,

$$m_G(\boldsymbol{v}^2) \equiv m = \text{constant}. \tag{2.7}$$

Axiom 3′ (Einsteinian mechanics, or Special Relativity)**.** The mass of a free particle has the form

$$m_E\left(\boldsymbol{v}^2\right) \equiv \frac{m}{\sqrt{1 - \boldsymbol{v}^2/c^2}}, \tag{2.8}$$

with m the Galilean mass and c the speed of light in vacuum.

Proposition 2: The Galilean and Einsteinian models are completely determined by their respective Lagrangians:

Galilean mechanics:

$$L_0^G\left(\boldsymbol{v}^2\right) = \frac{m}{2}\boldsymbol{v}^2. \tag{2.9}$$

Special relativity:

$$L_0^E\left(\boldsymbol{v}^2\right) = -mc^2\sqrt{1-\boldsymbol{v}^2/c^2}. \tag{2.10}$$

Proof: Beginning with the forms of the Lagrangians in the proposition, the form of $m\left(\boldsymbol{v}^2\right)$ in Axiom 3 or $3'$ can be recovered:

$$2\frac{dL_0^G}{d\boldsymbol{v}^2} = m \quad , \quad 2\frac{dL_0^E}{d\boldsymbol{v}^2} = \frac{m}{\sqrt{1-\boldsymbol{v}^2/c^2}} = m_E\left(\boldsymbol{v}^2\right), \tag{2.11}$$

i.e., $L_0\left(\boldsymbol{v}^2\right)$ uniquely determines $m\left(\boldsymbol{v}^2\right)$. □

Remark 7*:* Consider the case $\boldsymbol{v}^2 \ll c$. From equation 2.10,

$$\begin{aligned} L_0^E\left(\boldsymbol{v}^2\right) &= -mc^2\sqrt{1-\boldsymbol{v}^2/c^2} \\ &= -mc^2\left[1-\frac{1}{2}\frac{\boldsymbol{v}^2}{c^2}+O\left(\boldsymbol{v}^4/c^4\right)\right] \\ &= -mc^2+\frac{1}{2}m\boldsymbol{v}^2+O\left(\boldsymbol{v}^4/c^4\right) \\ &= L_0^G\left(\boldsymbol{v}^2\right)-mc^2+O\left(\boldsymbol{v}^4/c^2\right). \end{aligned} \tag{2.12}$$

To lowest order in $\boldsymbol{v}^2/c^2$, L_0^G and L_0^E are equivalent, so that for $\boldsymbol{v}^2 \ll c^2$, the **nonrelativistic limit**, special relativity reduces to Galilean mechanics.

Remark 8*:* Relative to everyday scales, c is a large velocity, so that the nonrelativistic limit suffices for many purposes.

Remark 9*:* In special relativity, m is called the **rest mass**, and $E_0 \equiv mc^2$ is called the **rest energy** of a particle.

Remark 10*:* We will consider Galilean mechanics unless otherwise noted.

§ 1.2 Galileo's principle of relativity

Definition 1: Let $(\boldsymbol{x}, t)$ be an IS. The coordinate transformation

$$\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{V}t \quad , \quad t' = t, \tag{2.13}$$

with $\boldsymbol{V} =$ constant, is called a **Galileo transformation**.

Theorem 1: If $(\boldsymbol{x}, t)$ is an IS and $(\boldsymbol{x}', t')$ is related to $(\boldsymbol{x}, t)$ by a Galileo tranformation, then $(\boldsymbol{x}', t')$ is also an IS.

Proof:

$$\boldsymbol{v}' = \dot{\boldsymbol{x}}' = \boldsymbol{v} + \boldsymbol{V}, \tag{2.14}$$

so that

$$\begin{aligned} L_0^{G\prime} &= \frac{m}{2}\boldsymbol{v}'^2 \\ &= \frac{m}{2}\boldsymbol{v}^2 + m\boldsymbol{v}\cdot\boldsymbol{V} + \frac{m}{2}\boldsymbol{V}^2 \\ & \frac{m}{2}\boldsymbol{v}^2 + \frac{d}{dt}\underbrace{\left(m\boldsymbol{x}\cdot\boldsymbol{V} + \frac{m}{2}\boldsymbol{V}^2 t\right)}_{G(\boldsymbol{x},t)} \\ &= \frac{m}{2}\boldsymbol{v}^2 + \frac{d}{dt}G(\boldsymbol{x},t) \\ &= L_0^G + \frac{d}{dt}G(\boldsymbol{x},t), \end{aligned} \tag{2.15}$$

so that $L_0^{G\prime}$ and L_0^G are equivalent. □

Axiom 4 (Galileo's principle of relativity)**.** The laws of motion

a) are the same in all IS's, and

b) are invariant under Galileo tranformations.

Remark 1: From Axiom 4a, there is no absolute reference frame - all non-accelerated coordinate systemes are postulated to be equivalent.

Remark 2: Galilean mechanics obey's Galileo's principle of relativity.

Remark 3: Maxwell's E&M does *not* obey Galileo's principle of relativity, which created a problem that Einstein solved. More generally, even mechanics is found to be at odds with Galileo's principle of relativity for $\neg\left(\boldsymbol{v}^2 \ll c^2\right)$.

Remark 4: One often does not think of the principle of relativity in the context of nonrelativistic mechanics, but that is a mistake.

§ 1.3 Potentials

Axiom 5. The influence of the environment on a particle is characterized by

a) a function $U(\boldsymbol{x},t) : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}$, called a **scalar potential**, and

b) a function $\boldsymbol{V}(\boldsymbol{x},t) : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$, called a **vector potential**,

such that the Lagrangian takes the form

$$L(\boldsymbol{x},\boldsymbol{v},t) = L_0\left(\boldsymbol{v}^2\right) - U(\boldsymbol{x},t) + \boldsymbol{v}\cdot\boldsymbol{V}(\boldsymbol{x},t). \tag{2.16}$$

Remark 1: In equation 2.16, L_0 can be either L_0^G or L_0^E.

Remark 2*:* The scalar and vector potentials U and $\boldsymbol{V}$ are determined either by experiment, or by another theory, external to mechanics.

Remark 3*:* In the case that $\boldsymbol{V}(\boldsymbol{x},t)=\boldsymbol{0}$:

a) Equation 2.16 becomes

$$\begin{aligned} L(\boldsymbol{x},\boldsymbol{v},t) &= L_0\left(\boldsymbol{v}^2\right)-U(\boldsymbol{x},t) \\ &= T\left(\boldsymbol{v}^2\right)-U(\boldsymbol{x},t)\,, \end{aligned} \tag{2.17}$$

with $T\equiv L_0$ the **kinetic energy**[1].

b) Recall from equation 12, $L_0^E=-mc^2+L_0^G\left(\boldsymbol{v}^2\right)+O\left(\boldsymbol{v}^4/c^2\right)=-mc^2+\frac{m}{2}\boldsymbol{v}^2+O\left(\boldsymbol{v}^4/c^2\right)$. So, the kinetic energy in the Einsteinian case, with $\boldsymbol{V}=\boldsymbol{0}$, is effectively defined as the energy due to motion (i.e., energy dependent on $\boldsymbol{v}$), minus the rest energy mc^2.

Example 1 (Charged particles in an electromagnetic field)**:** U and $\boldsymbol{V}$ are determined by a theory external to mechanics, namely, Maxwell's electromagnetism. They take the form

$$U(\boldsymbol{x},t)=e\varphi(\boldsymbol{x},t) \quad , \quad \boldsymbol{V}(\boldsymbol{x},t)=\frac{e}{c}\boldsymbol{A}(\boldsymbol{x},t)\,, \tag{2.18}$$

where e is the charge of the particle, which has the role of a coupling constant, and φ and $\boldsymbol{A}$ determine the electromagnetic field via

$$\boldsymbol{E}(\boldsymbol{x},t)=-\nabla\varphi(\boldsymbol{x},t)-\frac{1}{c}\partial_t\boldsymbol{A}(\boldsymbol{x},t) \quad , \quad \boldsymbol{B}(\boldsymbol{x},t)=\nabla\times\boldsymbol{A}(\boldsymbol{x},t)\,. \tag{2.19}$$

Special case 1: Homogeneous electric field $\boldsymbol{E}=(0,0,E)$.
This is realized by

$$\varphi(\boldsymbol{x},t)=-Ez \quad , \quad \boldsymbol{A}(\boldsymbol{x},t)=0, \tag{2.20}$$

so that

$$U(\boldsymbol{x},t)=-eEz \quad , \quad \boldsymbol{V}(\boldsymbol{x},t)=0. \tag{2.21}$$

Special case 2: Homogeneous magnetic field $\boldsymbol{B}=(0,0,B)$.
This is realized by

$$\boldsymbol{A}(\boldsymbol{x},t)=\frac{1}{2}\boldsymbol{B}\times\boldsymbol{x} \quad , \quad \varphi(\boldsymbol{x},t)=0, \tag{2.22}$$

so that

$$U(\boldsymbol{x},t)=0 \quad , \quad \boldsymbol{V}(\boldsymbol{x},t)=\frac{e}{2c}\boldsymbol{B}\times\boldsymbol{x}. \tag{2.23}$$

Example 2 (Particle in a gravitational field)**:** U and $\boldsymbol{V}$ are determined by experiment (e.g., Kepler laws) or by an external theory (e.g., general relativity). An environment consisting of a massive spherical body of mass M, centered at the orign $\boldsymbol{x}=0$ produces potentials

$$U(\boldsymbol{x},t)=U(\boldsymbol{x})=-\frac{GMm}{|\boldsymbol{x}|} \quad , \quad \boldsymbol{V}(\boldsymbol{x},t)=0, \tag{2.24}$$

[1] If $\boldsymbol{V}\neq\boldsymbol{0}$, then the Lagrangian in equation 2.16 has a term dependent on both $\boldsymbol{x}$ and $\boldsymbol{v}$, so cannot be written as the sum of a term dependent on only $\boldsymbol{x}$ and a term dependent on only $\boldsymbol{v}$. In such a situation, the concept of kinetic energy, which for the $\boldsymbol{V}=\boldsymbol{0}$ case is the portion of the Lagrangian dependent only on velocity, loses its meaning.

with m the mass of the particle, which has the role of a coupling constant.

Special case: Near the surface of the earth, $\boldsymbol{x} = (0, 0, z)$, $|\boldsymbol{x}| \backsimeq R_E$ = radius of Earth. Then,

$$\nabla U \backsimeq \frac{GMm}{R_E^2} \frac{\boldsymbol{x}}{R_E} = gm\hat{\boldsymbol{z}}, \tag{2.25}$$

with

$$g = GM/R_E^2. \tag{2.26}$$

Then,

$$U(\boldsymbol{x}, t) \backsimeq mgz. \tag{2.27}$$

This is called a **homogeneous gravitational potential**.

Remark 4: In Example 2, there is a priori no reason for the coupling constant m, the "theory mass", to be the same as the parameter in L_0, the "inertial mass". However, experimentally the two are found to be the same to tremendous accuracy. This is known as the **equivalence principle**.

Remark 5: A charged particle in a homogeneous electric field and a particle in a homogeneous gravitational field are mathematically equivalent.

Definition 1: Let $G : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}$ be a differentiable function. Then the transformation of the potentials given by

$$U(\boldsymbol{x}, t) \to U'(\boldsymbol{x}, t) = U(\boldsymbol{x}, t) - \partial_t G(\boldsymbol{x}, t), \tag{2.28}$$

$$\boldsymbol{V}(\boldsymbol{x}, t) \to \boldsymbol{V}'(\boldsymbol{x}, t) = \boldsymbol{V}(\boldsymbol{x}, t) + \nabla G(\boldsymbol{x}, t) \tag{2.29}$$

is called a **gauge transformation**.

Theorem 1: Any gauge transformation of the potentials leads to an equivalent Lagrangian.

Proof: Let $L = L_0 - U + \boldsymbol{v} \cdot \boldsymbol{V}$, as in equation 2.16. Then, the transformation gives

$$\begin{aligned} L' &= L_0 - U' + \boldsymbol{v} \cdot \boldsymbol{V}' \\ &= L_0 - U + \boldsymbol{v} \cdot \boldsymbol{V} + \partial_t G + \boldsymbol{v} \cdot \nabla G \\ &= L_0 - U + \boldsymbol{v} \cdot \boldsymbol{V} + \frac{d}{dt} G(\boldsymbol{x}, t). \end{aligned} \tag{2.30}$$

□

Remark 6: Potentials are unique only up to gauge transformations.

Remark 7: Quantities that do not change under gauge transformations are called **gauge-invariant**.

§ 1.4 The equations of motion

Definition 1: For a particle in a vector potential we call

$$\frac{\partial L_0}{\partial \boldsymbol{v}} = \boldsymbol{p} \tag{2.31}$$

the **momentum** (as in Ch.1§5.1 Remark 6), and

$$\frac{\partial L}{\partial \boldsymbol{v}} = \boldsymbol{p} + \boldsymbol{V} \equiv \boldsymbol{\pi} \tag{2.32}$$

the **generalized momentum**.

Remark 1*:* Note that $\boldsymbol{p}$ is gauge invariant, whereas $\boldsymbol{\pi}$ is not.

Proposition 1 (Newton's second law)**:** For a particle subject to a scalar potential plus a vector potential the equations of motion take the form of **Newton's second law**:

$$\frac{d}{dt}\boldsymbol{p}(\boldsymbol{x},t) = \boldsymbol{F}^{(1)}(\boldsymbol{x},t) + \boldsymbol{F}^{(2)}(\boldsymbol{x},\boldsymbol{v},t), \tag{2.33}$$

with

$$\boldsymbol{F}^{(1)}(\boldsymbol{x},t) = -\nabla U(\boldsymbol{x},t) - \partial_t \boldsymbol{V}(\boldsymbol{x},t) \tag{2.34}$$

a velocity-independent force, and

$$\boldsymbol{F}^{(2)}(\boldsymbol{x},\boldsymbol{v},t) = \boldsymbol{v} \times (\nabla \times \boldsymbol{V}(\boldsymbol{x},t)) \tag{2.35}$$

a velocity-dependent force.

Proof: The Euler-Lagrange equations, with the definitions in equations 2.31 and 2.32, give (with L writeable as in equation 2.16)

$$\begin{aligned} \frac{d}{dt}\frac{\partial L}{\partial v_i} &= \frac{d}{dt}\pi_i \\ &= \frac{d}{dt}p_i + \frac{d}{dt}V_i \\ &= \frac{\partial L}{\partial x_i} \\ &= -\partial_i U + v_j \partial_i V_j. \end{aligned} \tag{2.36}$$

Comparing lines 2 and 4,

$$\begin{aligned} \frac{d}{dt}p_i &= -\partial_i U + v_j \partial_i V_j - \frac{d}{dt}V_i \\ &= -\partial_i U + v_j \partial_i V_j - \partial_t V_i - \partial_j v_j V_i \\ &= \underbrace{-\partial_i U - \partial_t V_i}_{F_i^{(1)}} + \underbrace{v_j \partial_i V_j - v_j \partial_j V_i}_{F_i^{(2)}} \end{aligned} \tag{2.37}$$

where the second term in the last line can be identified as $F_i^{(2)}$ as follows:

$$\begin{aligned} \boldsymbol{F}_i^{(2)} &= (\boldsymbol{v} \times (\nabla \times \boldsymbol{V}(\boldsymbol{x},t)))_i \\ &= \varepsilon_{ijl} v_j \varepsilon_{klm} \partial_l V_m \\ &= \varepsilon_{kij} \varepsilon_{klm} v_j \partial_l V_i \\ &= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\, v_j \partial_l V_m \\ &= v_j \partial_i V_j - v_j \partial_j V_i \\ &= F_i^{(2)}. \end{aligned} \tag{2.38}$$

□

Remark 2*:* The forces $\boldsymbol{F}^{(1)}$ and $\boldsymbol{F}^{(2)}$ are separately gauge-invariant - see problem 16.

§ 1.5 Systems of point masses

Axiom 6. In a system of N point masses that are subject to potentials and interact with one another, the interaction between the particles is characterized by a function, the **interaction potential**,

$$U_{\text{int}} : \mathbb{R}^{3N} \to \mathbb{R} \tag{2.39}$$

such that the Lagrangian takes the form

$$\begin{aligned} & L\left(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(N)}, \boldsymbol{v}^{(1)}, \ldots, \boldsymbol{v}^{(N)}, t\right) \\ = \quad & \sum_{\alpha=1}^{N} L_0\left(\boldsymbol{v}^{(\alpha)2}\right) - \sum_{\alpha=1}^{N} U\left(\boldsymbol{x}^{(\alpha)}, t\right) \\ & + \sum_{\alpha=1}^{N} \boldsymbol{v}^{(\alpha)} \cdot \boldsymbol{V}\left(\boldsymbol{x}^{(\alpha)}, t\right) + U_{\text{int}}\left(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(N)}\right). \end{aligned} \tag{2.40}$$

Special Cases:

1. For N non-interacting particles (i.e., $U_{\text{int}} = 0$), L is the sum of N independent particles,

$$L = \sum_{\alpha=1}^{N} L\left(\boldsymbol{x}^{(\alpha)}, \boldsymbol{v}^{(\alpha)}, t\right), \tag{2.41}$$

   where each L in the summation includes the L_0, U, and$\boldsymbol{V}$ terms.

2. For N non-interacting particles with with $\boldsymbol{V} = \boldsymbol{0}$ and $U(\boldsymbol{x}, t) = U(\boldsymbol{x})$,

$$L = \sum_{\alpha=1}^{N} \left[T\left(\boldsymbol{v}^{(\alpha)2}\right) - U\left(\boldsymbol{x}^{(\alpha)}\right)\right]. \tag{2.42}$$

Remark 1*:* Special case 2 is the case most often encountered in simple applications.

Proposition 1: For general coordinates, rather than Cartesian coordinates, the Lagrangian for special case 2 in Galilean mechanics takes the form

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \frac{1}{2} \sum_{i,j=1}^{f} g_{ij}(\boldsymbol{q})\, \dot{q}_i \dot{q}_j - U(\boldsymbol{q}). \tag{2.43}$$

Proof: In Galilean mechanics, equation 2.42 can be rewritten as

$$L = \sum_{\alpha=1}^{N} \frac{m^{(\alpha)}}{2} \left(\boldsymbol{v}^{(\alpha)}\right)^2 - \sum_{\alpha=1}^{N} U\left(\boldsymbol{x}^{(\alpha)}\right). \tag{2.44}$$

The coordinates can be relabelled as

$$x^{(1)}_{1,2,3} \equiv x_{1,2,3} \quad , \quad x^{(2)}_{1,2,3} \equiv x_{4,5,6} \quad , \cdots , \quad x^{(\alpha)}_{1,2,3} \equiv x_{3\alpha+1,3\alpha+2,3\alpha+3} \quad , \cdots . \tag{2.45}$$

The coordinate transformation from general coordinates $\boldsymbol{q}$ to Cartesian coordinates $\mathbf{x}$ can be written as

$$x_i = f_i \left(q_i, \ldots, q_f \right) \quad , \quad \dot{x}_i = \partial_j f_i \left(q_i, \ldots, q_f \right) \dot{q}_j , \tag{2.46}$$

with $f = 3N$.

$$\begin{aligned} L &= \sum_{i=1}^{f} \frac{1}{2} m_i \dot{x}_i^2 - U\left(x_1, x_2, x_3\right) - \ldots - U\left(x_{f-2}, x_{f-1}, x_f\right) \\ &= \sum_{i=1}^{f} \frac{m_i}{2} \sum_{j,k=1}^{f} \partial_j f_i\left(\boldsymbol{q}\right) \partial_k f_i\left(\boldsymbol{q}\right) \dot{q}_j \dot{q}_k \\ &\quad \underbrace{-U\left[f_1\left(\boldsymbol{q}\right), f_2\left(\boldsymbol{q}\right), f_3\left(\boldsymbol{q}\right)\right] - \ldots - U\left[f_{f-2}\left(\boldsymbol{q}\right), f_{f-1}\left(\boldsymbol{q}\right), f_f\left(\boldsymbol{q}\right)\right]}_{-U(\boldsymbol{q})} \\ &= \frac{1}{2} \sum_{i,j=1}^{f} \left(\sum_{k=1}^{f} m_k \partial_i f_k\left(\boldsymbol{q}\right) \partial_j f_j\left(\boldsymbol{q}\right) \right) \dot{q}_i \dot{q}_j - U\left(\boldsymbol{q}\right), \end{aligned} \tag{2.47}$$

where $m_1 = m_2 = m_3 = m^{(1)}$, $m_4 = m_5 = m_6 = m^{(2)}$, etc. Letting

$$g_{ij}\left(\boldsymbol{q}\right) \equiv \sum_{k=1}^{f} m_k \partial_i f_k\left(\boldsymbol{q}\right) \partial_j f_j\left(\boldsymbol{q}\right), \tag{2.48}$$

equation 2.47 becomes

$$L = \frac{1}{2} \sum_{i,j=1}^{f} g_{ij}\left(\boldsymbol{q}\right) \dot{q}_i \dot{q}_j - U\left(\boldsymbol{q}\right), \tag{2.49}$$

as desired. □

Example 1 (Planar double pendulum)**:** There are $f = 2$ degrees of freedom. The coordinates can be labelled as

- φ_1, the angle between the vertical -axis (y) and the line connecting the origin to the first point mass m_1, and

- φ_2, the angle between a line parallel to the vertical axis and the line connecting the point mass m_1 and the second point mass m_2.

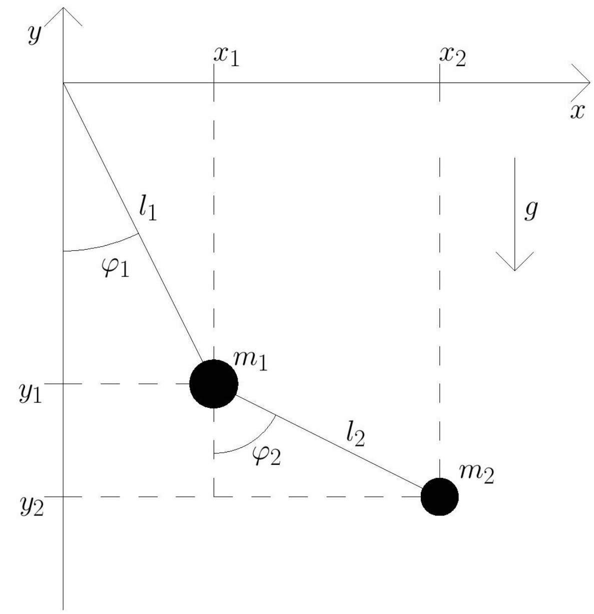

Figure 2.1: Planar double pendulum.

With lengths l_1 and l_2 defined as indicated in Figure 2.1, the coordinates of the first point mass are

$$x_1 = l_1 \sin\varphi_1 \quad , \quad y_1 = -l_1 \cos\varphi_1, \tag{2.50}$$

so that

$$\begin{aligned} T_1 &= \frac{m_1}{2}\left(\dot{x}_1^2 + \dot{y}_1^2\right) \\ &= \frac{m_1}{2} l_1^2 \dot{\varphi}_1^2, \end{aligned} \tag{2.51}$$

$$U_1 = -m_1 g l_1 \cos\varphi_1. \tag{2.52}$$

The coordinates of the second particle are

$$x_2 = l_1 \sin\varphi_1 + l_2 \sin\varphi_2 \quad , \quad y_2 = -l_1 \cos\varphi_1 - l_2 \cos\varphi_2, \tag{2.53}$$

so that

$$\begin{aligned} T_2 &= \frac{m_2}{2}\left(\dot{x}_2^2 + \dot{y}_2^2\right) \\ &= \frac{m_2}{2}\Big(l_1^2\dot{\varphi}_1^2\cos^2\varphi_1 + l_2^2\dot{\varphi}_2^2\cos^2\varphi_2 + 2l_1l_2\dot{\varphi}_1\dot{\varphi}_2\cos\varphi_1\cos\varphi_2 \\ &\qquad + l_1^2\dot{\varphi}_1^2\sin^2\varphi_1 + l_2^2\dot{\varphi}_2^2\sin^2\varphi_2 + 2l_1l_2\dot{\varphi}_1\dot{\varphi}_2\sin\varphi_1\sin\varphi_2\Big) \\ &= \frac{m_2}{2}\left(l_1^2\dot{\varphi}_1^2 + l_2^2\dot{\varphi}_2^2 + 2l_1l_2\cos\left(\varphi_1+\varphi_2\right)\dot{\varphi}_1\dot{\varphi}_2\right), \end{aligned} \tag{2.54}$$

$$U_2 = -m_2 g l_1 \cos\varphi_1 - m_2 g l_2 \cos\varphi_2. \tag{2.55}$$

Then,

$$\begin{aligned} L(\varphi_1, \varphi_2, \dot{\varphi}_1, \dot{\varphi}_2) &= T_1 + T_2 - U_1 - U_2 \\ &= \frac{1}{2}(m_1 + m_2) l_1^2 \dot{\varphi}_1^2 + \frac{1}{2} m_2 l_2^2 \dot{\varphi}_2^2 \\ &\quad + m_2 l_1 l_2 \cos(\varphi_1 - \varphi_2) \dot{\varphi}_1 \dot{\varphi}_2 \\ &\quad + (m_1 + m_2) g l_1 \cos\varphi_1 + m_2 g l_2 \cos\varphi_2. \end{aligned} \tag{2.56}$$

Remark 2*:* The Lagrangian has the form of equation 2.43 with

$$(g_{ij}(\varphi_1, \varphi_2)) = \begin{pmatrix} (m_1 + m_2) l_1^2 & l_1 l_2 \cos(\varphi_1 - \varphi_2) \\ l_1 l_2 \cos(\varphi_1 - \varphi_2) & m_2 l_2^2 \end{pmatrix}, \tag{2.57}$$

$$U(\varphi_1, \varphi_2) = -(m_1 + m_2) g l_1 \cos\varphi_1 - m_2 g l_2 \cos\varphi_2. \tag{2.58}$$

§ 2 Simple examples for the motion of point masses

§ 2.1 Galileo's law of falling bodies

Consider a particle in a homogenous gravitational potential. From §1.4, the kinetic energy is given by

$$\begin{aligned} T &= L_0(\boldsymbol{v}^2) \\ &= \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2), \end{aligned} \tag{2.59}$$

and from §1.3 Example 2, the potential energy is given by

$$U = mgz. \tag{2.60}$$

Then, the Lagrangian for the particle is

$$L(z, \dot{x}, \dot{y}, \dot{z}) = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - mgz. \tag{2.61}$$

Method 1 (Direct integration of the Euler-Lagrange equations)**.** In the Lagrangian, x and y are cyclic. Then,

$$\begin{aligned} & \frac{\partial L}{\partial \dot{x}} = m\dot{x} = \text{constant} \\ \Longrightarrow \quad & x(t) = a + bt, \end{aligned} \tag{2.62}$$

$$\begin{aligned} & \frac{\partial L}{\partial \dot{y}} = m\dot{y} = \text{constant} \\ \Longrightarrow \quad & y(t) = c + dt, \end{aligned} \tag{2.63}$$

where a, b, c, and d are constants. Choosing a coordinate system such that

$$\begin{aligned} & x\,(t=0) = y\,(t=0) = 0 \\ \Longrightarrow \quad & a = c = 0, \end{aligned} \tag{2.64}$$

$$\begin{aligned} & \dot{y}\,(t=0) = 0 \\ \Longrightarrow \quad & d = 0, \end{aligned} \tag{2.65}$$

and letting $v_x \equiv b$, gives

$$x\,(t) = v_x t \quad , \quad y\,(t) = 0, \tag{2.66}$$

so that the path lies in the xz-plane.

The remaining Euler-Lagrange equation is for z,

$$\begin{aligned} & \frac{d}{dt}\frac{\partial L}{\partial \dot{z}} = m\ddot{z} = \frac{\partial L}{\partial z} = -mg \\ \Longrightarrow \quad & \dot{z} = v_z^0 - gt, \end{aligned} \tag{2.67}$$

so that, letting $z_0 = z\,(t=0) =$ constant and $v_z^0 = \dot{z}\,(t=0) =$ constant,

$$z\,(t) = z_0 - v_z^0 t - \frac{1}{2}gt^2. \tag{2.68}$$

Remark 1*:* We have determined the path of the particle, i.e., its position as a function of time, by solving the Euler-Lagrange equations.

Continuing, $z\,(t)$ can be re-parameterized, using x as the parameter. Rewriting $z\,(t)$ as

$$z\,(t) = -\frac{1}{2}g\left[t^2 - 2\left(\frac{v_z^0}{g}\right)t + \left(\frac{v_z^0}{g}\right)^2\right] + \frac{1}{2}\frac{\left(v_z^0\right)^2}{g} + z_0. \tag{2.69}$$

Chossing, without loss of generality, $z_0 = -\frac{1}{2}\left(v_z^0\right)^2/g$, and defining $t^* \equiv -v_z^0/g$, gives

$$z\,(t) = -\frac{1}{2}g\,(t+t^*)^2\,. \tag{2.70}$$

Solving equation 2.66 for t gives

$$t = x/v_x. \tag{2.71}$$

Substituting into equation 2.70,

$$\begin{aligned} z &= -\frac{1}{2}g\,(x/v_x + t^*)^2 \\ &= -\frac{g}{2v_x^2}\,(x + v_x t^*)^2\,, \end{aligned} \tag{2.72}$$

so that

$$z\,(x) = -\frac{1}{2}g\,(x+x^*)^2 \quad , \quad x^* \equiv v_x t^* = -v_x v_z^0/g. \tag{2.73}$$

Remark 2*:* This is the **orbit** of the particle, i.e., the curve in $\mathbb{R}^3$ that it moves on. The orbit does not give information about where the particle is a a given time.

Remark 3*:* From equation 2.73, the orbit of a free-falling Galilean particle is a parabola.

Method 2 (Conservation of energy)**.** The equation 2.61 Lagrangian has no explicit time-dependence, so energy is conserved (Ch.1§5.2). From Ch.1§4.1,

$$\begin{aligned} E &= \dot{x}_i \frac{\partial L}{\partial \dot{x}_i} - L \\ &= m\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) - L \\ &= \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) + mgz \\ &= \text{constant.} \end{aligned} \tag{2.74}$$

Since x and y are still cyclic, equations 2.62 through 2.66 still apply. Then,

$$\begin{aligned} & E = \frac{m}{2}\dot{z}^2 + \frac{m}{2}v_x^2 + mgz \\ \Longrightarrow\quad & \frac{m}{2}\dot{z}^2 = -mgz + E - \frac{m}{2}v_x^2 \equiv -mgz + \frac{m}{2}\xi^2 \\ \Longrightarrow\quad & \dot{z}^2 = \xi^2 - 2gz \\ \Longrightarrow\quad & \frac{dz}{dt} = \sqrt{\xi^2 - 2gz}. \end{aligned} \tag{2.75}$$

The last line of equation 2.75 is a separable ODE, so we can write

$$\begin{aligned} & \int_{z_0}^{z} \frac{d\tilde{z}}{\sqrt{\xi^2 - 2g\tilde{z}}} = \int_{t_0}^{t} dt \\ \Longrightarrow\quad & -\frac{1}{g}\left(\sqrt{\xi^2 - 2gz} - \sqrt{\xi^2 - 2gz_0}\right) = t - t_0 \\ \Longrightarrow\quad & \xi^2 - 2gz = \left[g\left(t - t_0\right) - \sqrt{\xi^2 - 2gz_0}\right]^2, \end{aligned} \tag{2.76}$$

so that

$$\begin{aligned} z\left(t\right) &= \frac{1}{2g}\left\{\xi^2 - \left[g\left(t - t_0\right) - \sqrt{\xi^2 - 2gz_0}\right]^2\right\} \\ &= -\frac{1}{2}g\left(t + \text{constant}\right)^2, \end{aligned} \tag{2.77}$$

as in equation 2.70.

Method 3 (Gauge transformation)**.** Define

$$\begin{aligned} G\left(z, t\right) &\equiv U\left(z\right)t \\ &= mgzt. \end{aligned} \tag{2.78}$$

Then,

$$\frac{d}{dt}G = mgz + mg\dot{z}t \tag{2.79}$$

From §4.3,

$$\begin{aligned} L' &= L + \frac{d}{dt}G \\ &= \frac{m}{2}\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right) - mgz + mgz + mg\dot{z}t \\ &= \frac{m}{2}\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right) + mg\dot{z}t \end{aligned} \tag{2.80}$$

is equivalent to L. In L', x and y are cyclic, so equations 2.62 through 2.66 still apply. Additionally, z is cyclic in L', so

$$\begin{aligned} &\frac{\partial L'}{\partial \dot{z}} = m\dot{z} + mgt = \text{constant} \\ \Longrightarrow \quad &\dot{z} = -gt + \text{constant}, \end{aligned} \tag{2.81}$$

so that

$$z(t) = -\frac{1}{2}gt^2 + v_z^0 t + z_0, \tag{2.82}$$

as in equation 2.68.

Method 4 (Nöther's theorem)**.** Under the coordinate transformation

$$\begin{aligned} &z \to z' = z + \alpha \\ \Longrightarrow \quad &\dot{z}' = \dot{z}, \end{aligned} \tag{2.83}$$

we can write the transformed Lagrangian as

$$\begin{aligned} L\left(z', \dot{x}, \dot{y}, \dot{z}'\right) &= \frac{m}{2}\left(\dot{x}^2+\dot{y}^2+\dot{z}'^2\right) - mgz' \\ &= \frac{m}{2}\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right) - mgz - mg\alpha \\ &= L\left(z, \dot{x}, \dot{y}, \dot{z}\right) - mg\alpha \\ &= L\left(z, \dot{x}, \dot{y}, \dot{z}\right) - \alpha\frac{d}{dt}\left(mgt\right). \end{aligned} \tag{2.84}$$

Then, from Ch.1 §4.5, there is a conserved quantity,

$$\begin{aligned} \frac{\partial L}{\partial \dot{z}} + mgt &= m\dot{z} + mgt, \\ &= \text{constant} \end{aligned} \tag{2.85}$$

so that

$$z(t) = -\frac{1}{2}gt^2 + v_z^0 t + z_0, \tag{2.86}$$

which again matches equation 2.68.

Remark 4*:* All four methods yield the same results, as they should.

Remark 5*:* Utilization of cyclic variables is crucial within each method.

§ 2.2 Particle on an inclined plane

This problem can be treated as the problem of a free-falling particle with the following constraint:

$$z = (x - \xi) \tan \alpha. \tag{2.87}$$

Choosing $\xi = 0$ without loss of generality, we can write

$$x = \kappa z \quad , \quad \kappa \equiv \cot \alpha. \tag{2.88}$$

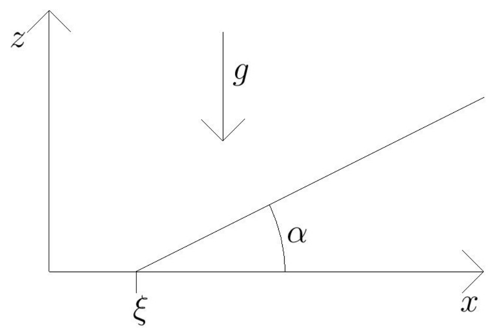

Figure 2.2: Particle on an inclined plane.

Method 1 (Lagrange multipliers)**.** Writing the Lagrangian as the Lagrangian for the particle with the Lagrangian for the constraint,

$$L = \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) - mgz + \lambda\left(x - \kappa z\right). \tag{2.89}$$

Since y is cyclic in the above Lagrangian, we have

$$\dot{y} = \text{constant}. \tag{2.90}$$

The remaining Euler-Lagrange equations are

$$m\ddot{x} = \lambda. \tag{2.91}$$

$$m\ddot{z} = -mg - \lambda\kappa, \tag{2.92}$$

The constraint given in equation 2.88, combined with equation 2.91 and 2.92, give

$$\begin{aligned}
& \ddot{x} = \kappa\ddot{z}, \\
\Longrightarrow \quad & \lambda = m\kappa\ddot{z} \\
\Longrightarrow \quad & \ddot{z}\left(1 - \kappa^2\right) = -g \\
\Longrightarrow \quad & \ddot{z} = -\frac{g}{1+\kappa^2} = -g\sin^2\alpha = -g_{\text{eff}}.
\end{aligned} \tag{2.93}$$

Solving equations 2.90 and 2.93,

$$x\left(t\right) = \kappa z\left(t\right), \tag{2.94}$$

$$y(t) = y_0 + v_y^0 t, \tag{2.95}$$

$$z(t) = z_0 + v_z^0 t - \frac{1}{2} g_{\text{eff}} t^2. \tag{2.96}$$

Remark 1: Motion in the z-direction is equivalent to that for a free-falling body, but with the acceleration g replaced with the acceleration g_{eff}.

Remark 2: Here, we have treated the problem with $f = 3 + 1$ constraint.

Method 2 (Direct substitution of the constraint)**.** From the constraint in equation 2.88,

$$\dot{x} = \kappa \dot{z}. \tag{2.97}$$

Then, the Lagrangian can be written as

$$L(z, \dot{y}, \dot{z}) = \frac{m}{2}\left[\left(1 + \kappa^2\right)\dot{z}^2 + \dot{y}^2\right] - mgz. \tag{2.98}$$

Since y is cyclic, equations 2.90 and 2.95 still apply. From equation 2.98, the Euler-Lagrange equations give

$$\begin{aligned} & m\left(1 + \kappa^2\right)\ddot{z} - mg \\ \Longrightarrow \quad & \ddot{z} = -g_{\text{eff}}, \end{aligned} \tag{2.99}$$

i.e., equation 2.93 is reproduced. From here, $x(t)$ and $z(t)$ can be found as was done in Method 1.

Remark 3: Here, we have used the constraint to reduce the number of degrees of freedom from $f = 3$ to $f = 2$.

§ 2.3 Particle on a rotating pole

Consider a particle moving without friction on a pole that rotates with a constant angular velocity ω. The constraint can be written in polar coordinates as

$$\varphi = \omega t. \tag{2.100}$$

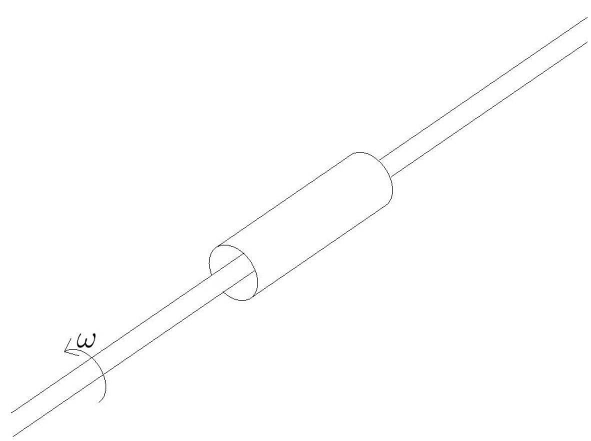

Figure 2.3: Particle on a rotating pole.

Choosing the xy-plane to be the plane in which the pole is rotating, the Lagrangian for free motion in that plane is writeable as

$$\begin{aligned} L &= \frac{m}{2}\left(\dot{x}^2+\dot{y}^2\right) \\ &= \frac{m}{2}\left(\dot{r}^2+r^2\dot{\varphi}^2\right). \end{aligned} \tag{2.101}$$

Substituting the constraint gives a Lagrangian with $f=1$ degrees of freedom,

$$L\left(r,\dot{r}\right)=\frac{m}{2}\left(\dot{r}^2+\omega^2r^2\right). \tag{2.102}$$

Remark 1*:* Elimination of the constraint maps the problem onto one with Lagrangian

$$L=T-U \quad , \quad T=\frac{m}{2}\dot{r}^2 \quad , \quad U=-\frac{m}{2}\omega^2r^2, \tag{2.103}$$

so that a free particle with $f=2+1$ constraint is equivalent to a particle with $f=1$ in a potential.

From equation 2.103, the Euler-Lagrange equations give

$$\begin{aligned} & m\ddot{r}=m\omega^2 r \\ \Longrightarrow \quad & r\left(t\right)=a\cosh\omega t+b\sinh\omega t, \end{aligned} \tag{2.104}$$

with a, b constant. Suppose that there are the initial conditions

$$r\left(t=0\right)=r_0>0 \quad , \quad \dot{r}\left(t=0\right)=0. \tag{2.105}$$

Then, we must have $a=r_0$, $b=0$, so that

$$r\left(t\right)=r_0\cosh\omega t. \tag{2.106}$$

Remark 2*:* From equation 2.106, the distance of the particle from the origin grows exponentially with time.

Remark 3*:* Enforcing the constraint requires a force that grows exponentially with time.

§ 2.4 The harmonic oscillator

Consider a particle with $f=1$ degrees of freedom in a harmonic potential given by

$$U\left(q\right)=\frac{1}{2}kq^2. \tag{2.107}$$

Such a system is called a **harmonic oscillator**, and its Lagrangian is writeable as

$$L\left(q,\dot{q}\right)=\frac{m}{2}\dot{q}^2-\frac{k}{2}q^2, \tag{2.108}$$

so that the Euler-Lagrange equations give

$$m\ddot{q}=-kq. \tag{2.109}$$

Defining the quantity ω^2 as

$$\omega^2 \equiv k/m \tag{2.110}$$

gives

$$\ddot{q} = -\omega^2 q, \tag{2.111}$$

so that

$$q(t) = a\cos\omega t + b\sin\omega t, \tag{2.112}$$

with a, b constant. Suppose that there are initial conditions

$$q(t=0) = q_0 \quad , \quad \dot{q}(t=0) = \dot{q}_0. \tag{2.113}$$

Then, we must have $a = q_0$, $b\omega = \dot{q}_0$, so that

$$q(t) = q_0\cos\omega t + (\dot{q}_0/\omega)\sin\omega t. \tag{2.114}$$

Remark 1*:* Note that

$$\begin{aligned} & \cos(\omega t + \varphi) = \cos\omega t\cos\varphi - \sin\omega t\sin\varphi \\ \Longrightarrow\quad & q(t) = a\cos(\omega t + \varphi) = a\cos\omega t\cos\varphi - a\sin\omega t\sin\varphi \\ \Longrightarrow\quad & a\cos\varphi = q_0 \quad , \quad -a\sin\varphi = \dot{q}_0/\omega \\ \Longrightarrow\quad & \tan\varphi = -\dot{q}_0/q_0\omega, \end{aligned} \tag{2.115}$$

so that

$$a = \frac{q_0}{\cos\left[\tan^{-1}(\dot{q}_0/q_0\omega)\right]}. \tag{2.116}$$

Then, the motion is bounded with $|q(t)| \leq a$, where a is called the **amplitude**.

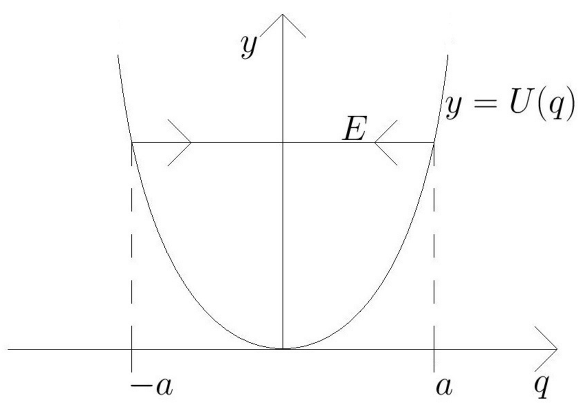

Figure 2.4: The harmonic oscillator.

Remark 2*:* The problem is conservative, i.e., energy is conserved. The energy is writeable as

$$\begin{aligned} E &= \frac{m}{2}\dot{q}^2 + \frac{m}{2}\omega^2 q^2 \\ &= \frac{m}{2}\omega^2 a^2. \end{aligned} \tag{2.117}$$

§ 3 One-dimensional conservative systems

§ 3.1 Definition of the model

Consider a conservative system with $f = 1$ (i.e., a **one-dimensional system**). From §1.3, the most general Lagrangian can be written as

$$L(q, \dot{q}) = \frac{m}{2} g(q) \dot{q}^2 - U(q), \tag{2.118}$$

with $g(q) > 0$, otherwise arbitrary.

Remark 1*:* Above, we have factored out the mass from $g(q)$.

Proposition 1: It is always possible to find a coordinate transformation such that

$$L\left(Q, \dot{Q}\right) = \frac{m}{2} \dot{Q}^2 - U(Q). \tag{2.119}$$

Proof: Let

$$Q = f(q) \equiv \int_{q_0}^{q} d\tilde{q} \sqrt{g(\tilde{q})}. \tag{2.120}$$

Then,

$$\dot{Q} = \sqrt{g(q)} \dot{q}, \tag{2.121}$$

so that

$$\dot{Q}^2 = g(q) \dot{q}^2. \tag{2.122}$$

□

Remark 2*:* It therefore suffices to study

$$L = \frac{m}{2} \dot{q}^2 - U(q) \tag{2.123}$$

with q a suitably chosen generalized coordinate and $m > 0$.

§ 3.2 Solution to the equations of motion

The problem represented by equation 2.123 (or equation 2.118) is conservative (i.e., the Lagrangian has no explicit time-dependence):

$$\frac{m}{2} \dot{q}^2 + U(q) = E, \tag{2.124}$$

a constant. Rearranging,

$$\begin{aligned} \dot{q}^2 &\equiv v^2(q) \\ &= \frac{2}{m} [E - U(q)]. \end{aligned} \tag{2.125}$$

Our coordinate space can therefore be decomposed into three distinct regions:

Case 1. $v^2(q) < 0 \iff E < U(q)$. There is no physical solution since $\dot{q}^2$ must be positive. We refer to this as the **forbidden region**.

Case 2. $v^2(q) > 0 \iff E > U(q)$. This is the **allowed region**, where $\dot{q} = \pm v(q)$ with

$$v(q) = \sqrt{2/m}\sqrt{E - U(q)} > 0. \tag{2.126}$$

There are thus two possibilities for $\dot{q}$:

a) $\dot{q} = v(q) > 0$: particle moves to q-right.

b) $\dot{q} = -v(q) < 0$: particle moves to q-left.

Case 3. $v^2(q) = 0 \iff E = U(q)$. Here $\dot{q} = v(q) = 0$; the particle is at rest.

Remark 1: Usually, $E = U(q)$ only for isolated points $q_1, q_2, \ldots$. These points are called **turning points**.

Now consider an interval $I = [q_0, q]$ in the allowed region that does not contain any turning points. In this interval,

$$\frac{dq}{dt} = \pm v(q). \tag{2.127}$$

Using separation of variables,

$$\boxed{t - t_0 = \pm \int_{q_0}^{q} \frac{dx}{v(x)},} \tag{2.128}$$

with $\pm$ for motion to the q-right or q-left, respectively.

Remark 2: Equation 2.128 solves the most general one-dimensional conservative problem in closed form. All that remains is to evaluate the integral in equation 2.128 and invert the resulting function to obtain $q(t)$.

Remark 3: A graphical description of the cases discussed above is provided in Figure 2.5 below.

§ 3.3 Unbounded motion

Let

$$U_{\pm\infty} \equiv \lim_{q\to\pm\infty} U(q) \quad , \quad U_{\pm\infty} < \infty. \tag{2.129}$$

Case 1. $E > U(q) \quad \forall q$.
This implies that the allowed region is all possible values of $q \in \mathbb{R}$. There are no turning points, so that the only two possiblities are

a) $\dot{q} = v(q) > 0 \quad \forall q$, or

b) $\dot{q} = v(q) < 0 \quad \forall q$.

Then, $q(t)$ increases monotonically or decreases monotonically[2] in cases (a) and (b), respectively. Choosing, without loss of generality,

$$t_0 = q(t_0) = 0, \tag{2.130}$$

[2] A **monotonic function** is a function between ordered sets that preserves the given order. To say that $q(t)$ increases/decreases monotonically simply means that it increases or decreases across the whole interval; if it increases monotonically, then $\frac{d}{dt}q > 0$ across the whole interval, and if it decreases monotonically, then $\frac{d}{dt}q < 0$ across the whole interval.

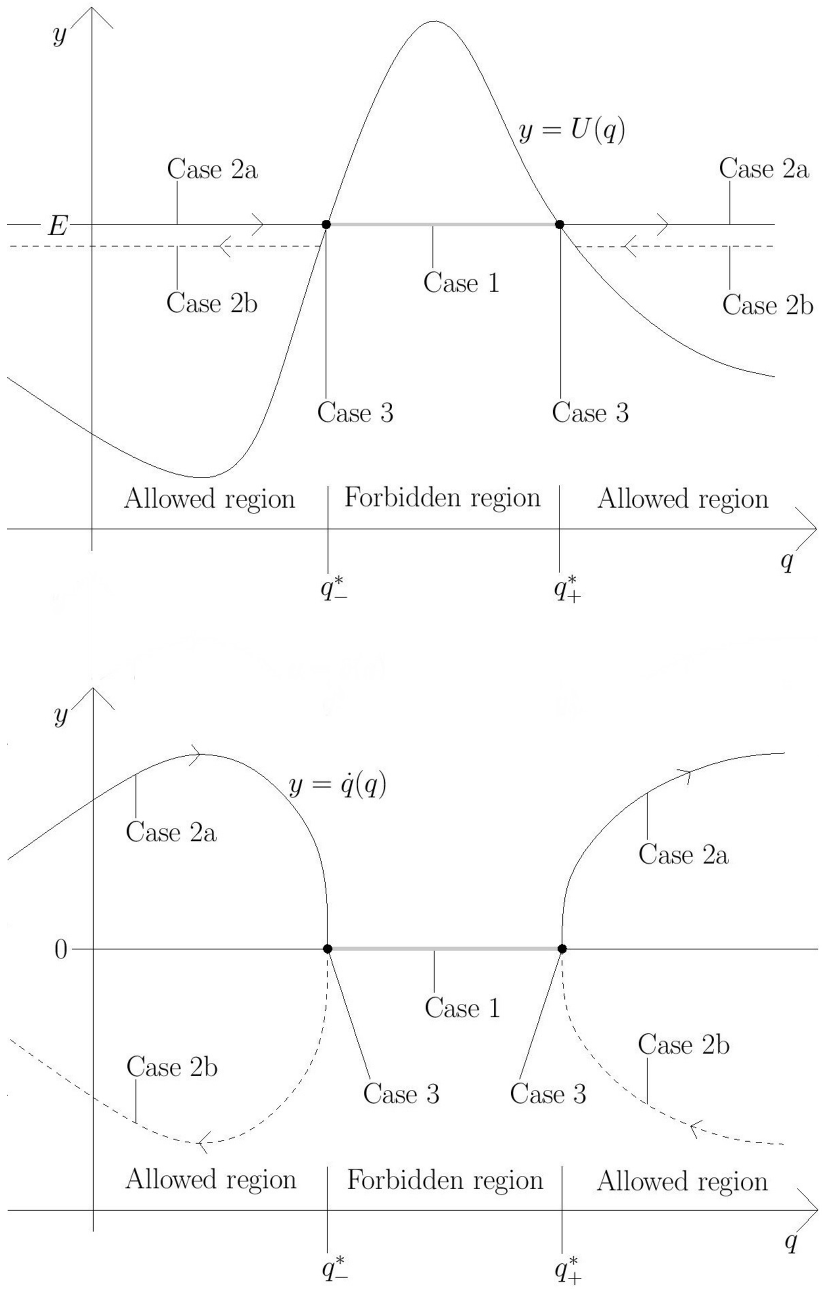

Figure 2.5: Graphical description of the forbidden region, turning points, and allowed region. Case 1: Forbidden region, $v^2(q) < 0 \iff E < U(q)$. Case 2: Turning points, $v^2(q) = 0 \iff E = U(q)$. Case 3: Allowed region, $v^2(q) > 0 \iff E > U(q)$.

applying the solution to the one-dimensional conservative problem given by equation 2.128, and solving for q gives

$$q\,(t \to \pm\infty) = \pm\infty. \tag{2.131}$$

Additionally, from equation 2.126,

$$\dot{q}\,(t \to \pm\infty) = \pm\sqrt{2/m}\sqrt{E - U_{\pm\infty}}, \tag{2.132}$$

with + and −corresponding to cases (a) and (b), respectively, in equation 2.132.

Case 2. $\exists q^* \quad E = U\,(q^*)$, and $E > U\,(q) \quad \forall q > q^*$.
This implies that the allowed region is all possible values of $q \in \mathbb{R}^{>q^*}$, and that there is a turning point at $q = q^*$. Choosing, without loss of generality,

$$t_0 = 0 \quad , \quad q\,(t_0) = q^*, \tag{2.133}$$

gives, by equation 2.128,

$$t = \pm\left|\int_{q^*}^{q} \frac{dx}{v\,(x)}\right| \text{ for } t \gtreqless 0. \tag{2.134}$$

Solving for q, and from equation 2.126,

$$q\,(t \to \pm\infty) = +\infty, \tag{2.135}$$

$$\dot{q}\,(t \to +\infty) = \pm\sqrt{2/m}\sqrt{E - U_{\pm\infty}} \text{ for } t \gtreqless 0. \tag{2.136}$$

Remark 1*:* The motion is symmetric with respect to the turning point, i.e., with the chosen $t_0 = 0$,

$$q\,(-t) = q\,(t)\,. \tag{2.137}$$

Case 3. $\exists q^* \quad E = U\,(q^*)$, and $E > U\,(q) \quad \forall q < q^*$.
This mirrors Case 2 exactly. The allowed region is all possible values of $q \in \mathbb{R}^{<q^*}$. Equation 2.133 can again be chosen without loss of generality, and in place of equations 2.134-2.136, we have

$$t = \mp\left|\int_{q^*}^{q} \frac{dx}{v\,(x)}\right| \text{ for } t \gtreqless 0, \tag{2.138}$$

$$q\,(t \to \pm\infty) = -\infty, \tag{2.139}$$

$$\dot{q}\,(t \to +\infty) = \mp\sqrt{2/m}\sqrt{E - U_{\pm\infty}} \text{ for } t \gtreqless 0. \tag{2.140}$$

The Case 2 remark still applies.

The remarks below apply to all three cases.

Remark 2*:* ...the motion is **unbounded**, i.e., $\lim_{t\to\pm\infty} |q\,(t)| = \infty$.

Remark 3*:* ...the particle visits each point along the physical path at most twice.

Remark 4*:* ...the particle spends a finite amount of time in any finite interval $I \subseteq \mathbb{R}$.

Remark 5*:* A graphical description of the cases discussed above, as well as Case 4 (discussed in §3.4), is provided in Figure 2.6 below.

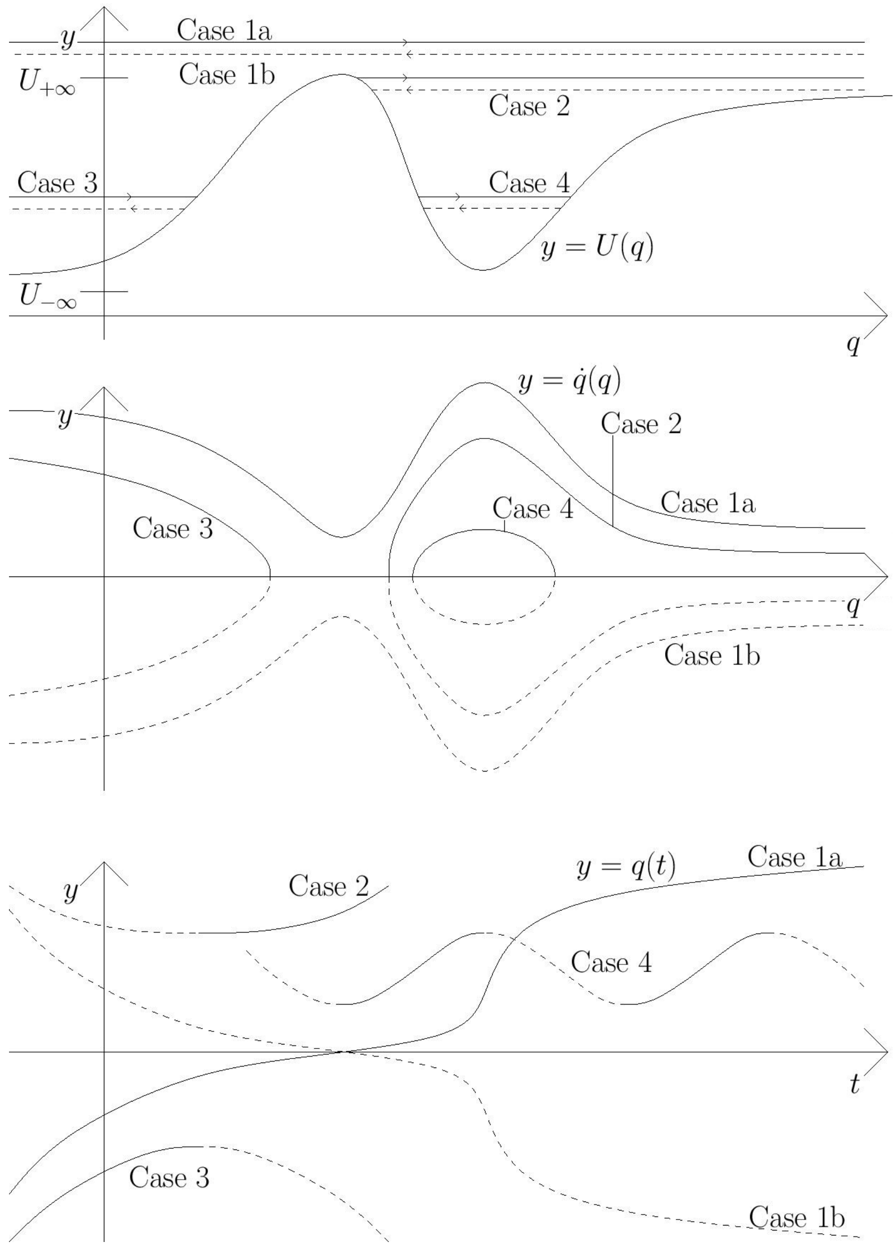

Figure 2.6: Graphical description of unbounded and bounded motion. Case 1: Unbounded motion, $\forall q \quad E > U(q)$, allowed region is all q. Case 2: Unbounded motion, $\exists q^* \quad E = U(q^*)$, and $\forall q > q^* \quad E > U(q)$, allowed region is all possible values of $q > q^*$. Case 3: Unbounded motion, $\exists q^* \quad E = U(q^*)$, and $\forall q < q^* \quad E > U(q)$, allowed region is all $q < q^*$. Case 4: Bounded motion, $\exists q_+, q_- \quad E = U(q_\pm)$ and $E > U(q)$ for $q_- < q < q_+$, allowed region is $q_- < q < q_+$.

§ 3.4 Bounded motion

Case 4. $\exists q_+^*, q_-^* \quad E = U\left(q_\pm^*\right)$ and $E > U(q)$ for $q_-^* < q < q_+^*$.
The turning points can be relabelled as $q_\pm^* \equiv q_\pm$. The allowed region is $q_- < q < q_+$, so that the motion is **bounded**. Defining, without loss of generality,

$$t_0 = 0 \quad , \quad q(t_0) = q_- \tag{2.141}$$

gives, from equation 2.128,

$$t = \pm \int_{q_-}^{q} \frac{dx}{v(x)}, \tag{2.142}$$

with $+$ for $t \geq 0$ and $-$ for $t \leq 0$.

A time T for bounded motion can also be defined. Let $q(T/2) = q_+$. Then, from equations 2.126 and 2.142,

$$\begin{aligned} T &= T(E) \\ &\equiv \sqrt{2m} \int_{q_-}^{q_+} \frac{dq}{\sqrt{E - U(q)}}. \end{aligned} \tag{2.143}$$

Remark 1: From equations 2.141-2.143, the motion is symmetric with respect to the turning point, i.e.,

$$q(-t) = q(t) \tag{2.144}$$

for $-\frac{T}{2} \leq t \leq \frac{T}{2}$.

Proposition 1: Bounded motion is T-periodic, i.e.,

$$q(t + nT) = q(t) \quad \forall n \in \mathbb{Z}. \tag{2.145}$$

Proof: Consider $Q(t) \equiv q(t - T)$, with $T/2 \leq t \leq 3T/2$. Then,

$$\begin{aligned} m\ddot{Q}(t) &= m\ddot{q}(t - T) \\ &= F(q(t - T)) \\ &= F(Q(t)). \end{aligned} \tag{2.146}$$

Then, $Q(T)$ and $q(t)$ obey the same ODE. Furthermore,

$$\begin{aligned} Q(T/2) &= q(-T/2) \\ &= q(T/2), \end{aligned} \tag{2.147}$$

and

$$\begin{aligned} \dot{Q}(T/2) &= \dot{q}(-T/2) \\ &= 0 \\ &= \dot{q}(T/2). \end{aligned} \tag{2.148}$$

These three facts, together with Ch.1 §2.4 Theorem 1, give

$$Q(t) = q(t), \tag{2.149}$$

An identical argument can be used with $Q(t) \equiv q(t+T)$, so that

$$q(t \pm T) = q(t), \tag{2.150}$$

which can be automatically extended to give

$$q(t+nT) = q(t) \quad \forall n \in \mathbb{Z}. \tag{2.151}$$

□

Remark 2: Bounded motion is also referred to as **oscillation**. Systems that display bounded motion are called **oscillatory**.

Remark 3: T is called the **period** of the oscillation, with

$$a \equiv \frac{1}{2}(q_+ - q_-) \tag{2.152}$$

called the **amplitude**, and

$$\begin{aligned} \omega &= \omega(E) \\ &\equiv 2\pi/T \end{aligned} \tag{2.153}$$

called the **frequency** of the oscillation.

§ 3.5 Equilibrium positions

Definition 1: A point x_0 is called an **equilibrium position** if it has the property that:

$$x(t_0) = x_0 \quad , \quad v(t_0) = 0 \qquad \Longrightarrow \qquad x(t) = x_0 \quad , \quad v(t) = 0. \tag{2.154}$$

Theorem 1: x_0 is an equilibrium position iff x_0 is a stationary point of $U(x)$.

Proof:

1. Assume that x_0 is an equilibrium position. Then, $\dot{x}|_{x_0} = 0 \implies \dot{x}(t) = 0 \forall t$. Then, $\dot{x}|_{x_0} = 0 \implies \ddot{x}(t_0) = 0$. Multiplying by m and using Newton's second law (§1.4), $\dot{x}|_{x_0} = 0 \implies 0 = m\ddot{x}(t_0) = -\frac{dU}{dx}|_{x_0}$, i.e., x_0 is a stationary point.
2. Assume that x_0 is a stationary point. Then, $\frac{dU}{dx}|_{x_0} = 0$. By Newton's second law, $m\ddot{x}(t_0) = 0$. If $\dot{x}(t_0) = 0$, then $\dot{x}(t_0 + \delta t) = \dot{x}(t_0) + \delta t \ddot{x}(t_0) = 0 + 0 = 0$. Then, $\dot{x}(t) = 0$ and $x(t) = x_0$. Then, equation 2.154 is satisfied, so x_0 is an equilibrium position.

□

Classification of equilibrium positions:

For $f = 1$, there are exactly three possibilities.

Case 1. x_0 is a local minimum of $U(x)$.

A small perturbation $E \to E + \delta E$ with $\delta E > 0$ (note that $\delta E < 0$ is impossible, see §3.2) leads to oscillatory motion with $x_- \le x \le x_+$. Such an equilibrium position is called **stable**, since it has the property that $\delta E \to 0 \implies x_\pm \to x_0$.

Case 2. x_0 is a local maximum of $\underline{U(x)}$.
A small perturbation $E \to E+\delta E$ leads to locally unbounded motion. Such an equilibrium position is called **unstable**.

Case 3. x_0 is a point of inflection of $U(x)$.
This is qualitatively the same as Case 2. Such an equilibrium position is unstable.

§ 4 Motion in a central field

§ 4.1 Reduction to a one-dimensional problem

Consider a conservative system with $f = 1$ and a spherically symmetric potential

$$U(\boldsymbol{x}) = U(r) \quad , \quad r \equiv |\boldsymbol{x}| . \tag{2.155}$$

The Lagrangian for such a system is writeable as

$$L = \frac{m}{2}\boldsymbol{v}^2 - U(\boldsymbol{r}) . \tag{2.156}$$

A switch to spherical coordinates can be performed as

$$x = r\sin\theta\cos\varphi \quad , \quad y = r\sin\theta\sin\varphi \quad , \quad z = r\cos\varphi, \tag{2.157}$$

so that

$$\begin{aligned} \boldsymbol{v}^2 &= \dot{x}^2 + \dot{y}^2 + \dot{z}^2 \\ &= \dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\varphi}^2. \end{aligned} \tag{2.158}$$

Substituting into the Lagrangian,

$$L = \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\varphi}^2\right) - U(r) . \tag{2.159}$$

To continue, various symmetries can be used.

1. **Rotational invariance I**: L is invariant under rotations. From Ch.1§4.4 Remark 2, Nöther's theorem gives that

$$\begin{aligned} \boldsymbol{x} \times \boldsymbol{p} &= \boldsymbol{l} \\ &= \text{constant}. \end{aligned} \tag{2.160}$$

   By the definition of $\boldsymbol{l}$ as the cross product of $\boldsymbol{x}$ and $\boldsymbol{p}$, $\boldsymbol{l}$ must be perpendicular to $\boldsymbol{x}$, so $\boldsymbol{x} \cdot \boldsymbol{l} = 0$. Then, the path must lie in a plane that is perpendicular to $\boldsymbol{l}$ (which is constant, i.e., both its magnitude and direction are constant). Our coordinate system can then be chosen with $\hat{\boldsymbol{z}}_{||}\boldsymbol{l}$ and $z = 0$ on the plane on which the path can be found, so that effectively

$$\theta \equiv \pi/2. \tag{2.161}$$

2. **Rotational invariance II**: φ is cyclic in the Lagrangian, so

$$\begin{aligned}\frac{\partial L}{\partial\dot{\varphi}} &= mr^2\sin^2\theta\dot{\varphi}\\ &= mr^2\dot{\varphi}\\ &= l,\end{aligned}\tag{2.162}$$

which is constant[3]. Solving the last two lines for $\dot{\varphi}$,

$$\dot{\varphi} = \frac{l}{mr^2}.\tag{2.163}$$

3. **Conservation of energy**: t is not in the Lagrangian, so the system is conservative, so that

$$\begin{aligned}E &= \frac{\partial L}{\partial\dot{q}_i}\dot{q}_i - L\\ &= \frac{\partial L}{\partial\dot{r}}\dot{r} + \frac{\partial L}{\partial\dot{\theta}}\dot{\theta} + \frac{\partial L}{\partial\dot{\varphi}}\dot{\varphi} - L\\ &= \frac{m}{2}\dot{r}^2 + U(r) + \frac{l^2}{2mr^2},\end{aligned}\tag{2.164}$$

where the last line follows from equation 2.163 and §4.2.

Remark 1: The three-dimensional problem represented by the Lagrangian in equation 2.159 has been reduced to a one-dimensional problem, represented by the energy in equation 2.164.

Remark 2: In §4.1, the most general one-dimensional conservative problem was solved up to performing an integral and inverting a function (equation 2.128). Since this problem has been reduced to falling in that category, it has been solved as well.

§ 4.2 Kepler's second law

From §4.1, $\boldsymbol{x}\cdot\boldsymbol{l} = 0$, and a coordinate system can be chosen such that $\hat{\boldsymbol{z}}_{\|}\boldsymbol{l}$ and $\theta = \pi/2$, as in Figure 2.7.

[3] See §4.2 for the relation between $\boldsymbol{l}$ and l.

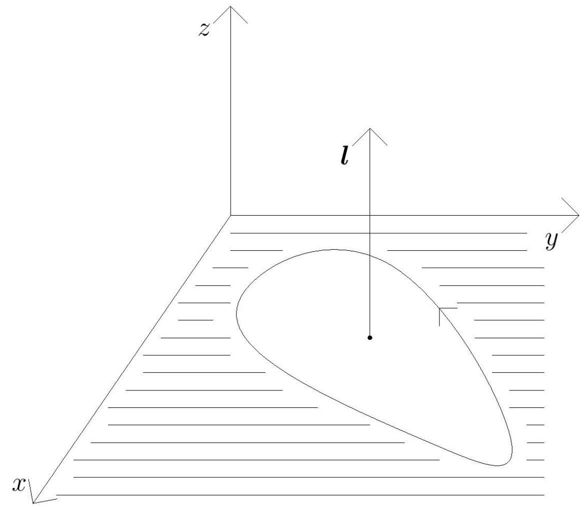

Figure 2.7: From §4.1, the path lies in a plane perpendicular to $\boldsymbol{l}$.

Because $\hat{\boldsymbol{z}}_{\|}\boldsymbol{l}$, equation 2.160 gives

$$\begin{aligned} l &= l_z \\ &= m\left(x\dot{y} - y\dot{x}\right) \\ &= m\left(r^2\cos^2\varphi\dot{\varphi} + r^2\sin^2\varphi\dot{\varphi}\right) \\ &= mr^2\dot{\varphi}. \end{aligned} \tag{2.165}$$

Remark 1*:* l is the length of $\boldsymbol{l}$.

Remark 2*:* $l \gtreqless 0$ for $\dot{\varphi} \gtreqless 0$.

Case 1. $l = 0$.

From equation 2.165, $\dot{\varphi} = 0$ as well, so that the orbit is a straight line, with

$$\varphi\left(t\right) = \varphi_0. \tag{2.166}$$

Case 2. $l \neq 0$.

Consider the area swept by the radius vector during an infinitesimal time:

$$dt = \frac{d\varphi}{\dot{\varphi}} \quad , \quad dA = \frac{1}{2}r^2 d\varphi. \tag{2.167}$$

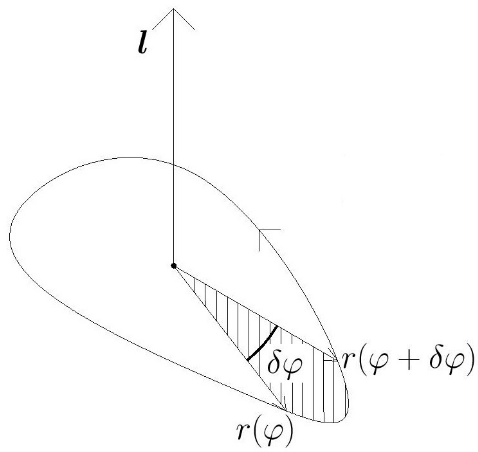

Figure 2.8: The area swept by the radius vector during an infinitesimal time.

Then, we have

$$\begin{aligned} \frac{d}{dt} A(t, t_0) &= \frac{1}{2} r^2 (\varphi(t)) \dot{\varphi}(t) \\ &= \frac{1}{2} \frac{l}{m}, \end{aligned} \tag{2.168}$$

where the last line follows from Eq. (2.163) in §4.1.

Theorem 1 (Kepler's second law)**:** The radius vector from the force center to the point mass sweeps equal areas in equal times.

Proof: This has already been proven - it is effectively stated in Eq. (2.168). □

Remark 3*:* This holds irrespective of the functional form of $U(r)$.

Remark 4*:* It is a consequence of angular momentum conservation.

Remark 5*:* Kepler found it emperically by analyzing Tycho Brahe's data on the orbit of Mars.

Remark 6*:* The **sectorial velocity** dA/dt is a measure of l.

Remark 7*:* Recall that angular momentum is conserved if and only if the system is rotationally invariant (see Ch.1§4.4). This leads to the orbits being planar and to Kepler's second law.

§ 4.3 Radial motion

From Eq. (2.164) in §4.1,

$$E = \frac{m}{2} \dot{r}^2 + U_{\text{eff}}(r), \tag{2.169}$$

where an **effective potential** $U_{\text{eff}}(r)$ has been defined as

$$U_{\text{eff}}(r) \equiv U(r) + \frac{l^2}{2mr^2}. \tag{2.170}$$

Remark 1: The three-dimensional motion in a central potential $U(r)$ has been reduced to a one-dimensional motion in an effective potential $U_{\text{eff}}(r)$. The quantity $l^2/2mr^2$ is called the **centrifugal potential**, or the **centrifugal barrier**.

Remark 2: For $l \neq 0$, angular momentum conservation makes it energetically costly to approach the center.

Applying equation 2.128 from §3.2, the solution (again up to an integral and inversion) is

$$t - t_0 = \pm\sqrt{\frac{m}{2}} \int_{r_0}^{r} \frac{dx}{\sqrt{E - U_{\text{eff}}(x)}}, \tag{2.171}$$

with $\pm$ for $\dot{r} \gtrless 0$ and with $r_0 \equiv r(t_0)$.

Remark 3: For a given E and l, $r(t)$ is determined by a single one-dimensional integral, after which$\varphi(t)$ can be determined by integrating equation 2.163 from §4.1.

Remark 4: From §3.2, it is $U_{\text{eff}}(r)$, rather than $U(r)$, that can be used to directly determine the turning points of the r-motion, since equation 2.169 represents the energy for which the problem is a one-dimensional conservative problem, rather than a three dimensional conservative problem.

§ 4.4 Equation of the orbit

From equation 2.163 in §4.1 and equation 2.169 in §4.3,

$$\begin{aligned} \dot{\varphi} &= \frac{l}{mr^2} \\ &= \frac{d\varphi}{dr}\dot{r} \\ &= \frac{d\varphi}{dr}\left[\pm\sqrt{\frac{2}{m}}\sqrt{E - U_{\text{eff}}(r)}\right]. \end{aligned} \tag{2.172}$$

Then,

$$\frac{d\varphi}{dr} = \pm\frac{l}{mr^2}\sqrt{\frac{m}{2}}\frac{1}{\sqrt{E - U_{\text{eff}}(r)}}. \tag{2.173}$$

Integrating,

$$\boxed{\varphi - \varphi_0 = \pm\frac{l}{\sqrt{2m}} \int_{r_0}^{r} \frac{dx}{x^2\sqrt{E - U_{\text{eff}}(x)}},} \tag{2.174}$$

with $\pm$ for $\frac{d\varphi}{dr} \gtrless 0$, and with $r_0 = r(\varphi_0)$. Equation 2.174 is called the **equation of orbit**.

Remark 1: The orbit $\varphi(r)$ (or alternatively, $r(\varphi)$), is determined by a single one-dimensional integral, plus the inversion of a function.

Remark 2*:* The path $r(t)$ (and similarly, $\varphi(t) = \varphi(r(t))$) is then determined by one more one-dimensional integral, plus the inversion of a function.

Remark 3*:* The central field problem is therefore solved *completely* in terms of two one-dimensional integrals, given by equations 2.128 and 2.174.

§ 4.5 Classification of orbits

Recall from §3.3 and §3.4 that the motion for a a one-dimensional conservative system depends on the relationship between the energy E and the one-dimensional potential U. Depending on this relationship, the motion can be bounded or unbounded. Then, in the case of the central field problem with potential $U = U(r)$, which was reduced to a one-dimensional conservative system with $U = U_{\text{eff}}(r)$, the motion depends on the relative values of E and $U_{\text{eff}}(r)$.

Example 1 (Attractive Coulomb potential: $U(r) = -\alpha/r$, with $\alpha > 0$)**:** Recalling that $U_{\text{eff}}(r) = U(r) + l^2/2mr^2$, $U_{\text{eff}}(r)$ will have the following properties:

a) $\lim_{r\to 0} U_{\text{eff}}(r) = +\infty$

b) $\lim_{r\to\infty} U_{\text{eff}}(r) = 0$

c) $U_{\text{eff}}(r)$ has a single, absolute minimum below 0.

Bounded motion will occur when $E > U_{\text{eff}}(r)$ for a finite region and unbounded motion will occur when $E > U_{\text{eff}}(r)$ for an infinite region, so for the attractive Coulomb potential, unbounded motion occurs for $E \geq 0$, and bounded motion occurs for $E < 0$. See Figure 2.9.

Example 2 (Repulsive Coulomb potential: $U(r) = \alpha/r$, with $\alpha > 0$)**:** Properties (a) and (b) from the previous example still apply. However, property (c) does not - instead, $U_{\text{eff}}(r)$ will have no minimum, so that it can never go below 0. Then, every region for which $E > U_{\text{eff}}(r)$ must be infinite, so only unbounded motion can occur for the repulsive Coulomb potential. See Figure 2.9.

Let r_- be the r-leftmost turning point for motion in a Coulomb potential. If the motion is bounded, let r_+ be the r-rightmost turning point.

Remark 1*:* r_- is called the **pericenter**, and r_+ is called the **apocenter**.

Remark 2*:* $r_\pm$ are the turning points for radial motion.

Equations 2.137 and 2.144 apply to radial motion, so that the r-motion is symmetric with respect to the turning points, i.e.,

$$r(t_\pm + t) = r(t_\pm - t)\,. \tag{2.175}$$

As a result, equation 2.163 gives $\dot{\varphi}(t)$ is symmetric with respect to the turning points, so that

$$\varphi(t_\pm + t) - \varphi(t_\pm) = -\left[\varphi(t_\pm - t) - \varphi(t_\pm)\right], \tag{2.176}$$

that is, φ-motion is antisymmetric with respect to the turning points.

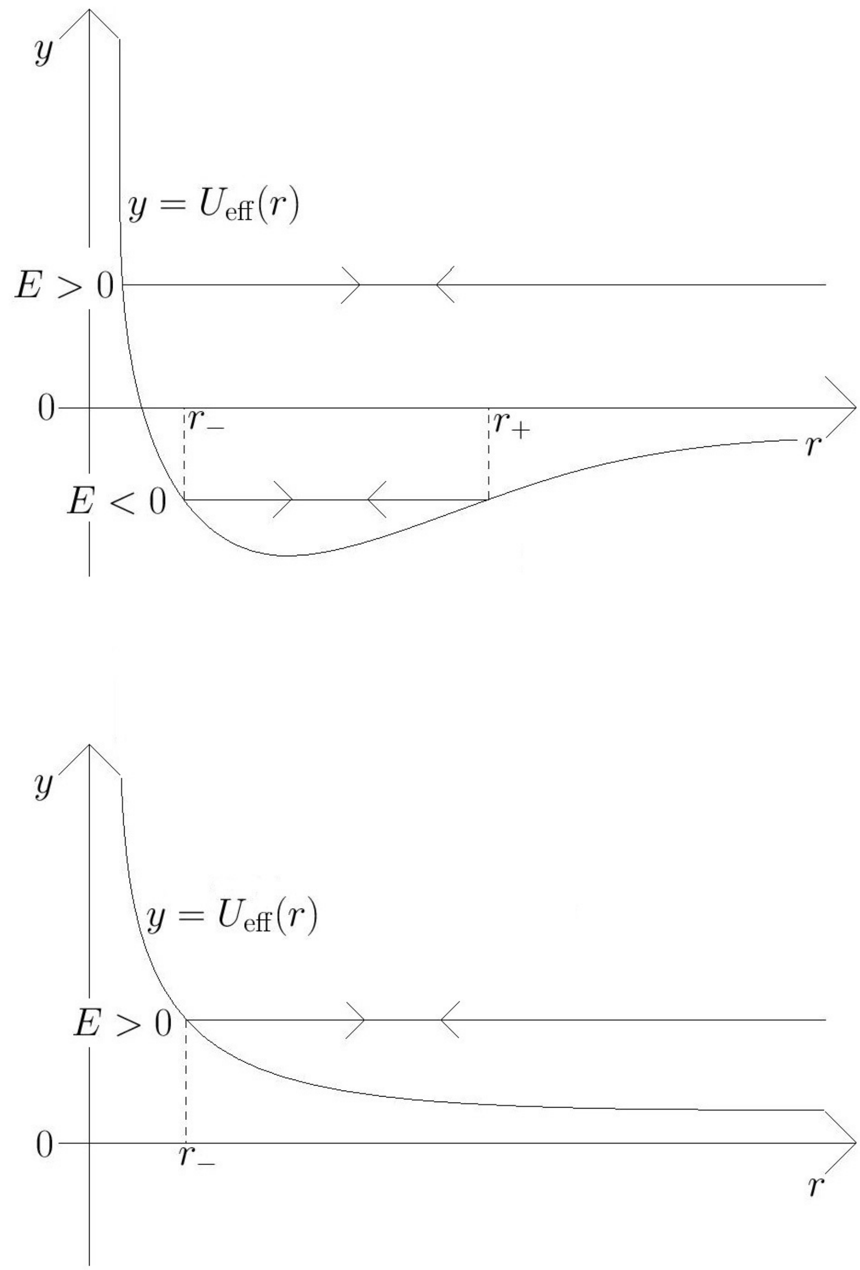

Figure 2.9: Graphical description of §4.5 Examples 1 (top) and 2 (bottom).

Case 1. Unbounded motion
If the motion is unbounded, the only turning point can be the pericenter at r_-. Choosing without loss of generality $\varphi(r_-) = 0$ (which is not, in general, the same as φ_-), equation 2.174 gives

$$\begin{aligned} \varphi(r \to \infty) &= \pm \frac{l}{\sqrt{2m}} \int_{r_-}^{\infty} \frac{dx}{x^2 \sqrt{E - U_{\text{eff}}(x)}} \\ &\equiv \varphi_\pm, \end{aligned} \tag{2.177}$$

so that the asymptotic orbit is a straight line. Since $\varphi_- = -\varphi_+$, the incoming and outgoing asymptotic orbits must be at equal and opposite angles with respect to $\varphi(r_-)$.

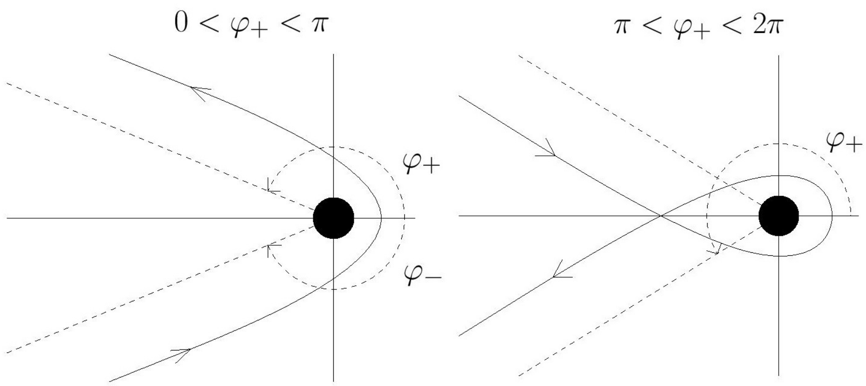

Figure 2.10: Unbounded motion in a central potential for different values of φ_+.

Remark 3: Which values of φ_+ are realizable depends on $U(r)$, and hence so does the allowed motion - for instance, the orbit can cross itself only if $\varphi_+ \geq \pi$ is allowed. More on this in §4.6.

Case 2. Bounded motion
If the motion is bounded, there are turning points at both the pericenter r_- and the apocenter r_+. From equation 2.151, the r-motion is periodic,

$$r(t) = r(t + nT). \tag{2.178}$$

Then, from equations 2.174 and 2.176,

$$\begin{aligned} \varphi(T/2) &= \frac{l}{\sqrt{2m}} \int_{r_-}^{r_+} \frac{dx}{x^2 \sqrt{E - U_{\text{eff}}(x)}} \\ &\equiv \frac{1}{2}\Phi, \end{aligned} \tag{2.179}$$

with Φ the rotation angle of the pericenter (or apocenter) during one period (see Figure 2.11):

$$\varphi(t + nT) = n\Phi + \varphi(t). \tag{2.180}$$

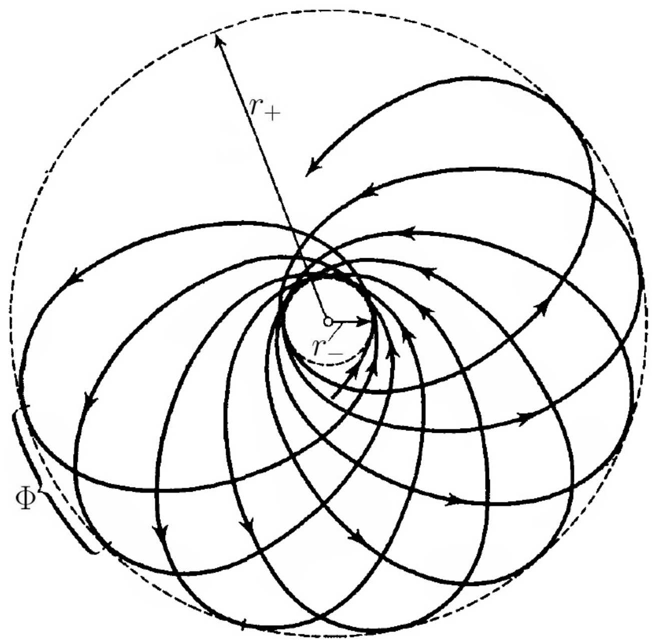

Figure 2.11: Bounded motion in a central potential. Adapted from Landau and Lifschitz Third Edition, Figure 9.

Remark 4: The orbit is closed if and only if $\Phi/2\pi$ is rational; otherwise, the orbit is dense on the annulus.

Remark 5: It can be shown that bounded orbits are closed only for $U(r) \propto r^{\alpha}$, with $\alpha = -1$ or $\alpha = 2$ only. This is known as **Bertrand's theorem**.

Remark 6: From experiment, planetary orbits are almost closed, with $U(r) \propto r^{-1}$. The possibility of$U(r) \propto r^2$ can be ruled out, since it is clear from experiment that the force on planetary bodies does not increase with distance.

Remark 7: The pericenter r_- may be 0 if $U(r \to 0) \propto r^{-\alpha}$ with $\alpha \geq 2$. See problem 27.

§ 4.6 Kepler's problem

Consider the "$1/r$-potential",

$$U(r) = -\frac{\alpha}{r}, \tag{2.181}$$

where α is the strength of the potential. The force is given by

$$\begin{aligned} \boldsymbol{F} &= -\nabla U \\ &= -\frac{\alpha}{r^2}\frac{\boldsymbol{r}}{r}, \end{aligned} \tag{2.182}$$

which is the **inverse-square law**.

Remark 1: $\alpha > 0 \implies U(r) < 0$, $F \propto -\hat{\boldsymbol{r}}/r^2$ - the potential is attractive.
$\alpha < 0 \implies U(r) > 0$, $F \propto +\hat{\boldsymbol{r}}/r^2$ - the potential is repulsive.

Example 1 (Gravity)**:** $\alpha = Gm_1m_2 > 0$.

Example 2 (Coulomb)**:** $\alpha = -e_1e_2 > 0$ for $\operatorname{sign}(e_1) = \operatorname{sign}(e_2)$,
< 0 for $\operatorname{sign}(e_1) \neq \operatorname{sign}(e_2)$.

In equation 2.181, $U = U(r)$, so this reduces to a one-dimensional conservative problem as described throughout §4. From equation 2.170,

$$U_{\text{eff}}(r) = -\frac{\alpha}{r} + \frac{l^2}{2mr^2}. \tag{2.183}$$

Define a parameter p, known simply as the **parameter** in this context, as

$$p \equiv \frac{l^2}{m\alpha}. \tag{2.184}$$

Then,

$$\begin{aligned} U_{\text{eff}}(r) &= -\frac{\alpha}{r} + \frac{1}{2}\frac{\alpha}{r^2}p \\ &= \frac{\alpha}{p}\left[\frac{1}{2}\left(\frac{p}{r}\right)^2 - \frac{p}{r}\right]. \end{aligned} \tag{2.185}$$

Discussion of $U_{\text{eff}}(r)$:

The square-bracket term of equation 2.185 is of the form $\left(\frac{1}{2}x^2 - x\right)$. The extremal points of U_{eff} can be found by setting

$$\begin{aligned} 0 &= \frac{d}{dx}\left(\frac{1}{2}x^2 - x\right) \\ &= x - 1, \end{aligned} \tag{2.186}$$

i.e., $x = p/r = 1$. Then,

$$\begin{aligned} \min\left(U_{\text{eff}}(r)\right) &= U_{\text{eff}}(r = p) \\ &= -\frac{\alpha}{2p}. \end{aligned} \tag{2.187}$$

Other useful pieces of information from equation 2.185 are

$$U_{\text{eff}}(r \to 0) = \frac{\alpha}{2p}\left(\frac{p}{r}\right)^2, \tag{2.188}$$

$$U_{\text{eff}}(r \to \infty) = -\frac{\alpha}{p}\frac{p}{r}. \tag{2.189}$$

Discussion of the types of motion:

Case 1. $E \geq 0$: Unbounded motion.

Case 2. $E < 0$: Bounded motion

a) The turning points can be found at

$$\begin{aligned}
& E = U_{\text{eff}}\left(r\right) = \frac{\alpha}{2p}\left(\frac{p}{r_\pm}\right)^2 - \frac{\alpha}{p}\left(\frac{p}{r_\pm}\right) \\
\Longrightarrow \quad & \left(\frac{p}{r_\pm}\right)^2 - 2\left(\frac{p}{r_\pm}\right) - 2p\frac{E}{\alpha} = 0 \\
\Longrightarrow \quad & \frac{p}{r_\pm} = \frac{1}{2}\left[2 \mp \sqrt{4 + 8pE/\alpha}\right] = 1 \mp \sqrt{1 + 2pE/\alpha}.
\end{aligned} \tag{2.190}$$

Define the **eccentricity** as

$$\begin{aligned}
e &\equiv \sqrt{1 + 2pE/\alpha} \\
&= \sqrt{1 + 2l^2E/m\alpha^2}.
\end{aligned} \tag{2.191}$$

Remark 2: $E < 0 \iff e < 1 \iff$ bounded motion,
$E > 0 \iff e > 1 \iff$ unbounded motion.

Knowing the above remark, the pericenter r_- and, if appropriate, apocenter r_+ are

$$r_\mp = \frac{p}{1 \pm e}. \tag{2.192}$$

Remark 3: In the context of planetary motion, $r_\mp$ is called the **perihelion** or **apohelion**, respectively.

Remark 4: For a repulsive potential, unbounded motion is seen only.

b) The equation of orbit can be found from equation 2.174. Choosing without loss of generality $\varphi\left(r = r_-\right) = 0$ and φ increasing with r, and applying equation 2.185,

$$\begin{aligned}
\varphi\left(r\right) &= \frac{l}{\sqrt{2m}}\int_{r_-}^{r}\frac{dx}{x^2\sqrt{E - U_{\text{eff}}\left(x\right)}} \\
&= \frac{l}{\sqrt{2m}}\int_{r_-}^{r}\frac{dx}{x^2\sqrt{E - \frac{\alpha}{2p}\left(\frac{p}{x}\right)^2 + \frac{\alpha}{p}\left(\frac{p}{x}\right)}} \\
&= \frac{l}{\sqrt{m}}\frac{\sqrt{p}}{\sqrt{\alpha}}\int_{r_-}^{r}\frac{dx}{x^2\sqrt{2pE/\alpha + 2\left(\frac{p}{x}\right) - \left(\frac{p}{x}\right)^2}}.
\end{aligned} \tag{2.193}$$

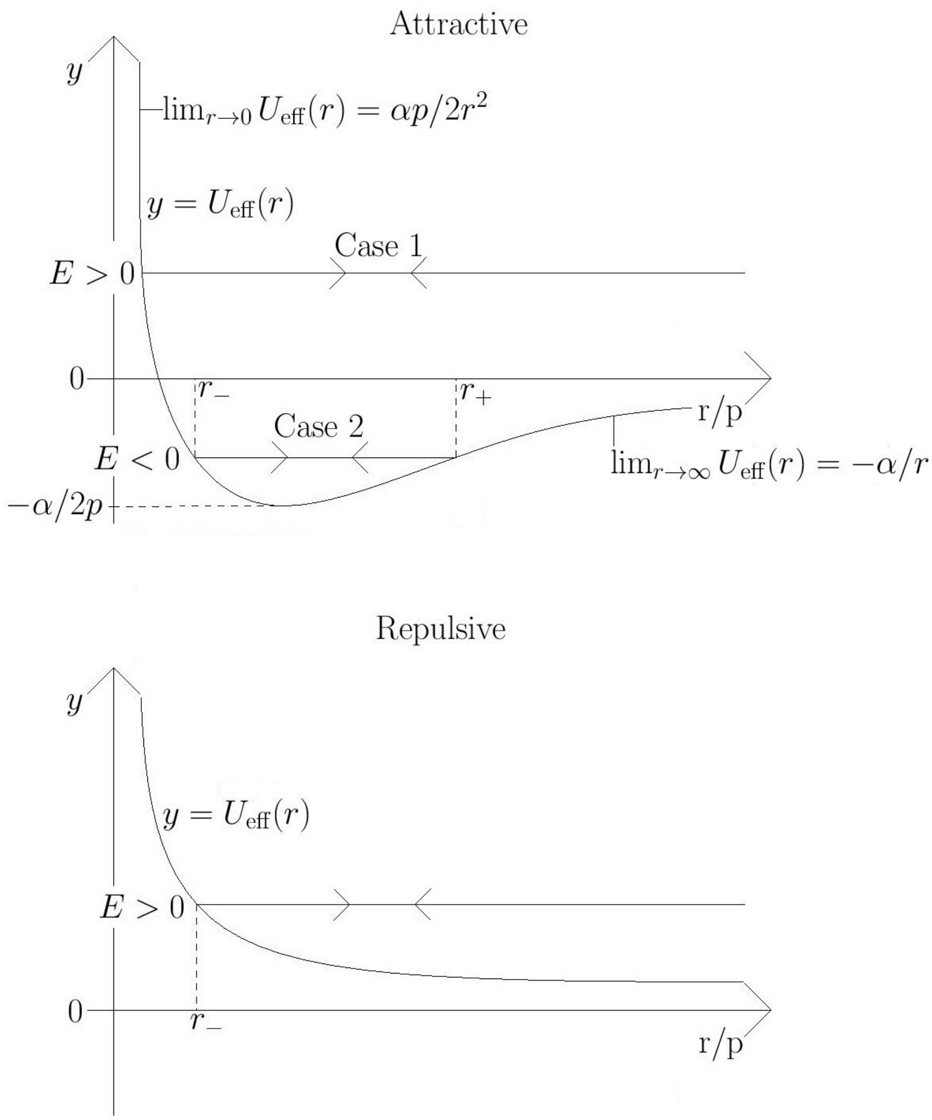

Figure 2.12: Graphical description of a $-\alpha/r$ potential. For $\alpha > 0$, the potential is attractive, resulting in Kepler's problem. Case 1: $E \geq 0$, unbounded motion. Case 2: $E < 0$, bounded motion. For $\alpha < 0$, the potential is repulsive.

Let $z \equiv p/x$, so that $dz = -p dx/x^2$. Then, $dx/x^2 = -\frac{1}{p} dz$, so that

$$\begin{aligned}
\varphi(r) &= -\frac{l}{\sqrt{pm\alpha}} \int_{p/r_-}^{p/r} \frac{dz}{\sqrt{2pE/\alpha + 2z - z^2}} \\
&= -\frac{l}{\sqrt{pm\alpha}} \int_{1+e}^{p/r} \frac{dz}{\sqrt{e^2 - 1 + 2z - z^2}} \\
&= -\frac{l}{\sqrt{pm\alpha}} \int_{1+e}^{p/r} \frac{dz}{\sqrt{e^2 - (z-1)^2}} \\
&= -\frac{l}{\sqrt{pm\alpha}} \int_{e}^{p/r-1} \frac{dx}{\sqrt{e^2 - x^2}} \\
&= -\frac{l}{\sqrt{pm\alpha}} \left[\sin^{-1} \frac{p/r-1}{e} - \sin^{-1} 1\right] \\
&= -\sin^{-1}\left[\frac{1}{e}\left(\frac{p}{r} - 1\right)\right] + \frac{\pi}{2},
\end{aligned} \tag{2.194}$$

where in the last line $l^2 = pm\alpha$ was used. Inverting,

$$\begin{aligned}
\frac{1}{e}\left(\frac{p}{r} - 1\right) &= -\sin\left[\varphi(r) - \frac{\pi}{2}\right] \\
&= \cos\left[\varphi(r)\right],
\end{aligned} \tag{2.195}$$

so that the **equation of orbit** is

$$\frac{p}{r} = 1 + e\cos\left(\varphi(r)\right), \tag{2.196}$$

Remark 5*:* The orbits are conic sections.

Remark 6*:* Let $e > 1$, so that the motion is unbounded. Then, letting $\varphi_+ \equiv \varphi(r \to +\infty)$, we have from equation 2.196 that $1 + e\cos\varphi_+ = 0$. Then, $\cos\varphi_+ < 0$, and $|\cos\varphi_+| < 1$. Continuity implies

$$\frac{\pi}{2} < \varphi_+ < \pi, \tag{2.197}$$

so that φ_+ is bounded from below and above. Thus, while orbits with $\frac{\pi}{2} < \varphi_+ < \pi$ are allowed, orbits with $\varphi_+ > \pi$ are not. This changes qualitatively in the framework of special relativity (see problem 28).

c) Classification of orbits: In Cartesian coordinates,

$$x = r\sin\varphi \quad , \quad y = r\cos\varphi \quad , \quad x^2 + y^2 = r^2. \tag{2.198}$$

Then,

$$p = r + ey, \tag{2.199}$$

so that

$$\begin{aligned}
x^2 + y^2 &= (p - ey)^2 \\
&= e^2y^2 - 2epy + p^2,
\end{aligned} \tag{2.200}$$

so that

$$x^2 + \left(1 - e^2\right) y^2 + 2epy - p^2 = 0. \tag{2.201}$$

Two useful **invariants** can be defined as:

$$\begin{aligned} \Delta &\equiv \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1-e^2 & ep \\ 0 & ep & -p^2 \end{vmatrix} \\ &= -p^2 \\ &\leq 0, \end{aligned} \tag{2.202}$$

so that the conic section is non-degenerate for $p \neq 0 \iff l \neq 0$.

$$\begin{aligned} \delta &= \begin{vmatrix} 1 & 0 \\ 0 & 1-e^2 \end{vmatrix} \\ &= 1 - e^2. \end{aligned} \tag{2.203}$$

The possible orbits are given in Figure 2.13, as well as described, below.

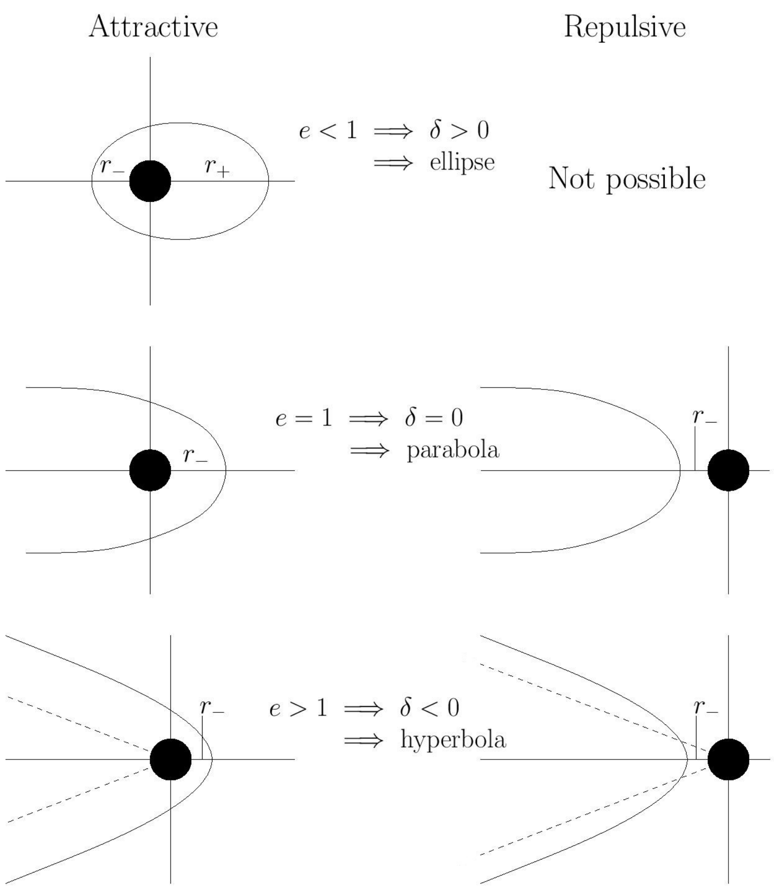

Figure 2.13: Classification of orbits in a $-\alpha/r$ potential.

1. $e < 1 \implies \delta > 0$, an **ellipse**. This is possible only for the attractive $1/r$ potential.
2. $e = 1 \implies \delta = 0$, a **parabola**. This is possible for both the attractive and repulsive $1/r$ potentials.
3. $e > 1 \implies \delta < 0$, a **hyperbola**. This is possible for both the attractive and repulsive $1/r$ potentials.

§ 4.7 Kepler's laws of planetary motion

Consider bounded motion in an attractive potential: $\alpha > 0$, $E < 0$. From §4.6,

$$p > 0 \quad , \quad 0 \le e < 1. \tag{2.204}$$

From Eq. (2.201),

$$\left(1 - e^2\right)\left[y^2 + \frac{2ep}{1-e^2}y + \left(\frac{ep}{1-e^2}\right)^2\right] - \frac{e^2p^2}{1-e^2} - p^2 + x^2 = 0, \tag{2.205}$$

so that, with $y^* \equiv ep/\left(1 - e^2\right)$,

$$\begin{aligned} x^2 + \left(1-e^2\right)\left(y+y^*\right)^2 &= p^2\left(1 + \frac{e^2}{1-e^2}\right) \\ &= \frac{p^2}{1-e^2} \\ &= a^2\left(1-e^2\right), \end{aligned} \tag{2.206}$$

so that

$$\frac{x^2}{b^2} + \frac{\left(y+y^*\right)^2}{a^2} = 1, \tag{2.207}$$

with a and b the **semi-major axis** and **semi-minor axis**, respectively:

$$\begin{aligned} a &= \frac{p}{1-e^2} \\ &= -\frac{p}{2pE/\alpha} \\ &= -\frac{\alpha}{2E}, \end{aligned} \tag{2.208}$$

$$\begin{aligned} b &= \frac{p}{\sqrt{1-e^2}} \\ &= \sqrt{p}\sqrt{-\alpha/2E} \\ &= \sqrt{ap}. \end{aligned} \tag{2.209}$$

Equation (2.207) represents **Kepler's first law**: Orbits are ellipses with the center of force at one focal point.

Remark 1: $a = p\left[1 + O\left(e^2\right)\right]$, $b = p\left[1 + O\left(e^2\right)\right]$, $y^* = ep\left[1 + Om(e^2\right] = ep + O\left(e^2\right)$. Apart from terms of $O\left(e^2\right)$, orbits are circles whose centers are shifted from the center of force by ea (Tycho Brahe).

Kepler's second law was already derived from the law of areas, see §4.2.

Kepler's third law: Ratios of periods equal ratios of semi-major axes to the 3/2-power.

Proof: From the law of areas,

$$\frac{dA}{dt} = \frac{l}{2m}, \tag{2.210}$$

so that

$$\begin{aligned} \pi ab &= \oint dt \frac{dA}{dt} \\ &= \frac{l}{2m} T. \end{aligned} \tag{2.211}$$

Then,

$$\begin{aligned} T &= \frac{2\pi m}{l} ab \\ &- \frac{2\pi m}{l} \sqrt{p} a^{3/2} \\ &= 2\pi \sqrt{\frac{m}{\alpha}} a^{3/2}. \end{aligned} \tag{2.212}$$

□

Remark 2: Kepler's third law is a corollary of Kepler's second law.

Remark 3: The second law holds for any central potential, whereas the first and third laws hold only for $1/r$ potentials.

§ 5 Introduction to perturbation theory

§ 5.1 General concept

Let L be the Lagrangian for a realistic system which is not soluble in closed form. Let L_0 be the Lagrangian for an idealized system, with

$$\boxed{L = L_0 + \delta L,} \tag{2.213}$$

where

$$\boxed{\begin{aligned} \delta L &= L - L_0 \\ &\equiv \varepsilon V, \end{aligned}} \tag{2.214}$$

is a small **perturbation**, with

$$\varepsilon \ll 1 \tag{2.215}$$

a **small parameter**. The idea of perturbation theory is to expand the solution $q(t)$ for the realistic system with Lagrangian L in powers of ε,

$$\boxed{q(t) = q_0(t) + \varepsilon q_1(t) + \varepsilon^2 q_2(t) + \ldots,} \tag{2.216}$$

where each $q_i(t)$ is the i'th order correction to the solution for the idealized system with Lagrangian L_0. Note that $q_0(t)$ is the uncorrected solution for the idealized system.

Example 1: Let L be the Lagrangian for the solar system. Then, L_0 could be the Lagrangian for a single planet and the Sun, so that δL is the correction to the Lagrangian that represents the other planets. Then, $\varepsilon = m_{\text{planets}}/m_{\text{Sun}}$.

Example 2: Let L be the Lagrangian for a single planet orbiting an oblate sun. Then, L_0 could be the Lagrangian for the planet with a spherical sun, so that δL is the correction to the Lagrangian that represents the sun's oblateness. Then, $\varepsilon =$ the quadrapole moment of the oblate sun.

Remark 1*:* Very few interesting problems are exactly soluble. Physics is perturbation theory (of the 0'th order).

Remark 2*:* Depending on your point of view, $q_0(t)$ is either the exact solution of the idealized system, or an approximate solution of the realistic system.

Remark 3*:* The hunt for ε sometimes produces unexpected choices, e.g,. $\varepsilon = 4 - d$.

Remark 4*:* Problem: Convergence.

§ 5.2 Digression - Fourier series

Let

$$I \equiv [-\pi, \pi], \tag{2.217}$$

and consider a set of functions on I defined by

$$\boxed{\varphi_n(x) \equiv \frac{1}{\sqrt{2\pi}} e^{inx} \quad , \quad n \in \mathbb{Z}.} \tag{2.218}$$

Proposition 1: Each φ_n obeys the **orthogonality relation**,

$$\int_I dx\, \varphi_n^*(x)\, \varphi_m(x) = \delta_{mn}. \tag{2.219}$$

Proof: Using Euler's formula,

$$\begin{aligned}
\int_I dx\, \varphi_n^*(x)\, \varphi_m(x) &= \int_{-\pi}^{\pi} dx\, \frac{e^{i(m-n)x}}{2\pi} \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} dx\, \{\cos[(m-n)x] + i \sin[(m-n)x]\} \\
&= \begin{cases} \int_{-\pi}^{\pi} dx\, \frac{e^0}{2\pi} = 1, & m-n=0 \\ \frac{1}{2\pi(m-n)} \{\sin[(m-n)x] - i\cos[(m-n)x]\} \Big|_{x=-\pi}^{\pi} = 0, & m-n \neq 0 \end{cases} \\
&= \delta_{mn}. \qquad (2.220)
\end{aligned}$$

□

Definition 1: The function $f : I \to \mathbb{C}$ defined by

$$\boxed{f(x) = \sum_n f_n \varphi_n(x) \quad , \quad f_n \in \mathbb{C}} \qquad (2.221)$$

is called a **Fourier series**, with **Fourier coefficients** f_n.

Lemma 1: The Fourier coefficients are uniquely determined by the function $f(x)$ as

$$\boxed{f_n = \int_I dx\, \varphi_n^*(x)\, f(x).} \qquad (2.222)$$

Proof:

$$\begin{aligned}
\int_I dx\, \varphi_n^*(x)\, f(x) &= \int_{-\pi}^{\pi} dx\, \varphi_n^*(x) \sum_m f_m \varphi_m(x) \\
&= \sum_m f_m \int dx\, \varphi_n^*(x)\, \varphi_m(x) \\
&= \sum_m f_m \delta_{mn} \\
&= f_n, \qquad (2.223)
\end{aligned}$$

where the second line follows from Proposition 1. □

Proposition 2: Fourier series have the following properties:

a) Let $\{f_n\}$ and $\{g_n\}$ determine the functions $f(x)$ and $g(x)$, respectively. Then, $\{\alpha f_n + \beta g_n\}$ determines the function $\alpha f(x) + \beta g(x)\ \forall \alpha, \beta \in \mathbb{C}$. That is, the Fourier series satisfy **linearity**.

b) Let $\{f_n\}$ determine $f(x)$. Then, $\left\{f_{-n}^*\right\}$ determines $f^*(x)$.

c) Let $\{f_n\}$ determine $f(x)$. Then, $\{inf_n\}$ determines $f'(x)$.

Proof:

a)

$$\begin{aligned}\alpha f(x)+\beta g(x) &= \alpha\sum_n f_n\varphi_n(x)+\beta\sum_n g_n\varphi_n(x)\\ &= \sum_n(\alpha f_n+\beta g_n)\varphi_n(x)\,.\end{aligned}\qquad(2.224)$$

b)

$$\begin{aligned}f^*(x) &= \sum_n[f_n\varphi_n(x)]^*\\ &= \sum_n f_n^*\varphi_{-n}(x)\\ &= \sum_n f_{-n}^*\varphi_n(x)\,,\end{aligned}\qquad(2.225)$$

where the second line follows from the definition of $\varphi(x)$ given in equation 2.218.

c)

$$\begin{aligned}f'(x) &= \frac{d}{dx}\sum_n f_n\varphi_n(x)\\ &= \sum_n f_n in\varphi_n(x)\,,\end{aligned}\qquad(2.226)$$

which again follows from the definition of $\varphi(x)$ given in equation 2.218.

□

Remark 1*:* From part (b) of Proposition 2,

$$f(x)\in\mathbb{R}\iff f_n=f_{-n}^*.\qquad(2.227)$$

Remark 2*:* From part (c) of Proposition 2, differentiation in real space corresponds to multiplication in Fourier space.

Question. Which functions can be written as Fourier series?

Theorem 1: Let $f(x)$ be a function on I that has been periodically continued onto $\mathbb{R}$, i.e., $f:\mathbb{R}\to\mathbb{C}$ with

$$f(x+2\pi)=f(x)\quad\forall x\in\mathbb{R},\qquad(2.228)$$

and let f be piecewise continuous and continuoulsy differentiable. Then, $f(x)$ can be written as a Fourier series,

$$f(x)=\sum_n f_n\varphi_n(x)\,,\qquad(2.229)$$

with

$$f_n=\int_{-\pi}^{\pi}dx\,\varphi_n^*(x)\,f(x)\,.\qquad(2.230)$$

Proof: Analysis course. □

Remark 3: Any function that satisfies equation 2.228 can be written as a Fourier series.

Remark 4: Theorem 1 can easily be generalized to periodic functions with a period other than 2π using a change of variables. This is done in equations 2.231 and 2.232 in §5.3 below.

§ 5.3 Oscillations and Fourier series

Recall from §3.4 that bounded motion is periodic,

$$q(t) = q(t + nT) \quad , \quad n \in \mathbb{Z}. \tag{2.231}$$

Define

$$\begin{aligned} & x \equiv \frac{2\pi}{T} t = \omega t \\ \Longrightarrow \quad & q(x) = q\left(x + n\frac{2\pi}{T}T\right) = q(x + 2\pi n). \end{aligned} \tag{2.232}$$

Equation 2.232 is equivalent to the condition given in equation 2.228. Then, from §5.2, $q(t)$ can be written as

$$\begin{aligned} q(t) &= \sum_n q_n e^{in\omega t} \quad , \quad (\text{with } n = 0, \pm 1, \pm 2, \ldots) \\ &= q_0 + \sum_{n=1}^{\infty} \left(q_n e^{in\omega t} + q_{-n} e^{-in\omega t}\right) \\ &= q_0 + \sum_{n=1}^{\infty} \left(q_n e^{in\omega t} + q_n^* e^{-in\omega t}\right) \\ &= q_0 + 2\sum_{n=1}^{\infty} \mathrm{Re}\left(q_n e^{in\omega t}\right). \end{aligned} \tag{2.233}$$

In the above equation, line 3 follows from the fact that $q(t) \in \mathbb{R}$, so that equation 2.227 applies, and line 4 follows from Euler's formula, as follows:

$$\begin{aligned} q_n e^{in\omega t} + q_n^* e^{-in\omega t} &= q_n \left[\cos(n\omega t) + i \sin(n\omega t)\right] + q_n^* \left[\cos(-n\omega t) + i\sin(-n\omega t)\right] \\ &= \left[q_n + q_n^*\right] \cos(n\omega t) + \left[q_n - q_n^*\right] i \sin(n\omega t) \\ &= \left[\mathrm{Re}(q_n) + \mathrm{Im}(q_n) + \mathrm{Re}(q_n^*) + \mathrm{Im}(q_n^*)\right] \cos(n\omega t) \\ &\quad + \left[\mathrm{Re}(q_n) + \mathrm{Im}(q_n) - \mathrm{Re}(q_n^*) - \mathrm{Im}(q_n^*)\right] i \sin(n\omega t) \\ &= \left[\mathrm{Re}(q_n) + \mathrm{Im}(q_n) + \mathrm{Re}(q_n) - \mathrm{Im}(q_n)\right] \cos(n\omega t) \\ &\quad + \left[\mathrm{Re}(q_n) + \mathrm{Im}(q_n) - \mathrm{Re}(q_n) + \mathrm{Im}(q_n)\right] i \sin(n\omega t) \\ &= \underbrace{2\mathrm{Re}(q_n)}_{\in\mathbb{R}} \cos(n\omega t) + \underbrace{2i\mathrm{Im}(q_n)}_{\in\mathbb{R}} \sin(n\omega t) \\ &= 2\mathrm{Re}\left(q_n e^{in\omega t}\right). \end{aligned} \tag{2.234}$$

Since each $q_n \in \mathbb{C}$, we can write each Fourier coefficient in for $q(t)$ as

$$q_n = \frac{1}{2} a_n e^{-i\theta_n} \quad , \quad \begin{matrix} a_n \geq 0 \\ -\pi \leq \theta_n \leq \pi \end{matrix}. \tag{2.235}$$

Substituting into equation 2.233,

$$\boxed{\begin{aligned} q(t) &= q_0 + \sum_{n=1}^{\infty} \mathrm{Re}\left[a_n e^{i(n\omega t - \theta_n)}\right] \\ &= q_0 + \sum_{n=1}^{\infty} a_n \cos(n\omega t - \theta_n) \\ &= \underbrace{q_0}_{\text{average position}} + \underbrace{a_1 \cos(\omega t - \theta_1)}_{\text{fundamental oscillation}} + \underbrace{a_2 \cos(2\omega t - \theta_2)}_{2'\text{nd harmonic}} + \underbrace{\cdots}_{\text{higher harmonics}} \quad . \end{aligned}} \tag{2.236}$$

In equation 2.236, the **average position** q_0 is also referred to as the **0'th harmonic**, and the **fundamental oscillation** is also referred to as the **1'st harmonic**. Similarly, the n'th term is referred to as the **n'th harmonic**.

Remark 1*:* Each a_n is called the **amplitude** of the n'th harmonic.

Remark 2*:* The complete solution for $q(t)$ is known if and only if the frequency ω and each of the amplitudes $\{q_0, a_n\}$ are known.

Remark 3*:* From §2.4, the harmonic oscillator has no higher harmonics.

Remark 4*:* From problem 29, the pendulum has all odd harmonics.

Remark 5*:* The language convention used is that :

- If there is a given $a_{n>1} \neq 0$, the oscillation is **anharmonic**. Otherwise, it is **harmonic**.
- If $\omega = \omega(E) \neq$ constant, the oscillation is **nonlinear**. Otherwise, it is **linear**.

Remark 6*:* The harmonic oscillator (§2.4) is both linear and harmonic. The pendulum (problem 29) is both nonlinear and anharmonic. Some textbooks (such as Landau and Lifschitz) do not distinguish between nonlinearity and anharmonicity - this is a bad use of language.

§ 5.4 Summary thus far, and the anharmonic oscillator

Summary thus far:

Before continuing, it is worth summarizing our scheme developed to this point. In §5.1, our premises for perturbation theory to apply to a given system were given as follows:

1. The Lagrangian for the system is writeable as the sum of the Lagrangian L_0 for an idealized, soluble system, and a perturbation δL:

$$L = L_0 + \delta L \quad , \quad \delta L = \varepsilon V \quad , \quad \varepsilon \ll 1. \tag{2.237}$$

2. The solution $q(t)$ for the perturbed system is writeable as a power series in ε,

$$\begin{aligned} q(t) &= \sum_{j=0}^{\infty} \varepsilon^j q_j(t) \\ &\equiv \sum_{j=0}^{\infty} q^{(j)}(t). \end{aligned} \tag{2.238}$$

where $q_0(t)$ is the solution for the system with Lagrangian L_0. Suggestively, we can rewrite the above equation with $q_n(t) \equiv q(t)$, giving

$$\begin{aligned} q_n(t) &= \sum_{j=0}^{\infty} \varepsilon^j q_{n,j}(t) \\ &\equiv \sum_{j=0}^{\infty} q_n^{(j)}(t). \end{aligned} \tag{2.239}$$

This version will come in handy later in this section.

In §5.2, properties of Fourier series were developed. These were applied in §5.3 to write the solution for any system with bounded motion as

$$q(t) = q_0 + \sum_{n=1}^{\infty} a_n \cos(n\omega t - \theta_n) \tag{2.240}$$

To this point, we have not applied perturbation theory in any way to solve our bounded-motion system. To continue, we will show that because any system with bounded motion must have a potential minimum somewhere, the Lagrangian for any system with bounded motion can be written as $L = L_0 + \delta L$ with $L_0 = L_{\text{H.O.}}$, so that our first perturbation theory premise is satisfied for any system with bounded motion if we can find an appropriate small parameter ε. Rather than attempting to apply the perturbation expansion given in equation 2.238 to $q(t)$ directly, the Euler-Lagrange equations and a Fourier expansion can be applied to our system, resulting in an equation for each Fourier coefficient in terms of all other Fourier coefficients.

A suitable choice for a small parameter ε will then be revealed, so that, finally applying a perturbation expansion, each Fourier coefficient q_n can be expanded using equation 2.238 (or 2.239). We can then determine the order in ε of each term in the Fourier coefficient equation mentioned at the end of the last paragraph. Then, by choosing to find our solution to an arbtirary finite order in ε, we can eliminate almost all terms in the Fourier coefficient equation, so that each Fourier coefficient can be found to a given $O\left(\varepsilon^i\right)$ in terms of a finite number of other coefficients that can be found to $O\left(\varepsilon^i\right)$ as well. We will thus have a solution for each Fourier coefficient to $O\left(\varepsilon^i\right)$, which can be substituted into equation 2.240 to give the solution for $q(t)$ to $O\left(\varepsilon^i\right)$. It will turn out that solution for the frequency ω to $O\left(\varepsilon^i\right)$ will depend only on the solution to $q(t)$ to $O\left(\varepsilon^{j<i}\right)$, so that the Fourier coefficient equation is sufficient to solve for both $q(t)$ and ω to $O\left(\varepsilon^i\right)$ by ignoring all terms of $O\left(\varepsilon^{k>i}\right)$.

The anharmonic oscillator:

Consider a potential $U(x)$ with a minimum at x_0. Expanding the potential in powers of $q = x - x_0$ gives

$$\begin{aligned} U(x) &= U(x_0 + q) \\ &= U(x_0) + q\underbrace{U'(x_0)}_{=0} + \frac{1}{2}q^2 U''(x_0) + \frac{1}{3!}q^3 U'''(x_0) + \ldots, \end{aligned} \tag{2.241}$$

where $U'(x_0)$ is 0 because $U(x)$ has a minimum at x_0. The Lagrangian for a system with such a potential is

$$L = \frac{m}{2}\dot{q}^2 - \frac{m}{2}\omega_0^2 q^2 - \frac{m\alpha_3}{3}q^3 - \frac{m\alpha_4}{4}q^4 + \ldots, \tag{2.242}$$

where

$$\alpha_n = \frac{U^{(n)}(x_0)}{(n-1)!m} \quad , \quad \omega_0^2 = \alpha_2 = \frac{U''(x_0)}{m}. \tag{2.243}$$

The Lagrangian has the form of the harmonic oscillator Lagrangian $L_{\text{H.O.}}$ with a perturbation δL. We can write

$$L = L_{\text{H.O.}} + \delta L, \tag{2.244}$$

with

$$L_{\text{H.O.}} = \frac{m}{2}\dot{q}^2 - \frac{m}{2}\omega_0^2 q^2 \quad , \quad \delta L = -m\sum_{n=2}^{\infty}\frac{1}{n}\alpha_n q^n. \tag{2.245}$$

Remark 1*:* As $q \to 0$, $\delta L \to 0$. However, q is not dimensionless, and so cannot be the small parameter ε needed for perturbation theory as described in §5.1.

Remark 2*:* Different powers of q are linearly independent, so we are assured that there exists a value of q for which (constants) $q^{n_1} =$ (constants) $q^{n_2 \neq n_1}$. Then, a length scale l can be defined as

$$\left.\frac{m}{2}\omega_0^2 q^2\right|_{q=l_n} = \left.\frac{m\alpha_n}{n}q^n\right|_{q=l_n} \quad , \quad l \equiv \min l_n. \tag{2.246}$$

Then, $\delta L \ll L_0$ provided that $q \ll l$, and a possible choice for the small parameter ε is

$$\varepsilon = \frac{a_1}{l}. \tag{2.247}$$

From equation 2.242, the equation of motion for the system is (applying the Euler-Lagrange equation, adding $\omega_0^2 q$ to each side, and multiplying by $-1/m$)

$$-\ddot{q} - \omega_0^2 q = \alpha_3 q^2 + \alpha_4 q^3 + \ldots. \tag{2.248}$$

From §5.3, $q = q(t)$ can be written as a Fourier series,

$$q(t) = \sum_n q_n e^{in\omega t}. \tag{2.249}$$

Substituting into equation 2.248, the $-\ddot{q}$ term can immediately be replaced with with $(n\omega)^2 q$, to give

$$\begin{aligned}\sum_n \left[(n\omega)^2 - \omega_0^2\right] q_n e^{in\omega t} &= \alpha_3 \sum_{n_1 n_2} q_{n_1} q_{n_2} e^{i(n_1+n_2)\omega t} \\ &\quad + \alpha_4 \sum_{n_1 n_2 n_3} q_{n_1} q_{n_2} q_{n_3} e^{i(n_1+n_2+n_3)\omega t} + \ldots . \end{aligned} \tag{2.250}$$

Consider that, letting $\tau = \omega t$,

$$\begin{aligned}\int_0^{T=2\pi/\omega} dt\, e^{-im\omega t} e^{in\omega t} &= \frac{1}{\omega}\int_0^{2\pi} d\tau\, e^{i(n-m)\tau} \\ &= T\delta_{mn}. \end{aligned} \tag{2.251}$$

Then, multiplying equation 2.250 by $\frac{1}{T}\int_0^T dt\, e^{-im\omega t}$ yields

$$\boxed{\begin{aligned}\left[(m\omega)^2 - \omega_0^2\right] q_m &= \alpha_3 \sum_{n_1 n_2} \delta_{m,n_1+n_2} q_{n_1} q_{n_2} \\ &\quad + \alpha_4 \sum_{n_1 n_2 n_3} \delta_{m,n_1+n_2+n_3} q_{n_1} q_{n_2} q_{n_3} + \ldots \\ &= \sum_{i=3}^{\infty} \alpha_i \sum_{n_1,\cdots,n_{i-1}} \delta_{m,\sum_{j=1}^{i-1} n_j} \prod_{j=1}^{i-1} q_{n_j}. \end{aligned}} \tag{2.252}$$

Lemma 1: In the equation of motion as presented in equation 2.252,

$$q_{-m} = q_m \quad \forall m. \tag{2.253}$$

Proof: From equation 2.252,

$$\begin{aligned}\left[(m\omega)^2 - \omega_0^2\right] q_{-m} &= \alpha_3 \sum_{n_1 n_2} \delta_{-m,n_1+n_2} q_{n_1} q_{n_2} + \ldots \\ &= \alpha_3 \sum_{n_1 n_2} \delta_{m,n_1+n_2} q_{-n_1} q_{-n_2} + \ldots \\ &= \alpha_3 \sum_{n_1 n_2} \delta_{m,n_1+n_2} q_{n_1} q_{n_2} + \ldots \\ &= \left[(m\omega)^2 - \omega_0^2\right] q_m \end{aligned} \tag{2.254}$$

where line 3 follows from the fact that all of the summations are between $-\infty$ and ∞. □

Remark 3: Energy E is conserved (see equations 2.244 and 2.245), and $q_{\pm 1} \equiv a/2$ depends on E. It is convenient fo fix a instead of E, and to consider E to be a function of a.

Remark 4: In the case that $\delta L = 0$, we have

$$q(t) = a\cos(\omega_0 t + \varphi_0), \tag{2.255}$$

so that perturbation theory amounts to an expansion in powers of a, with the small parameter ε given by (see Remark 2)

$$\varepsilon = \frac{a}{l}. \tag{2.256}$$

Note that $a = a_1$ in equation 2.235, provided that $\delta L = 0$.

We can now apply the perturbation expansion discussed in §5.1 to each Fourier coefficient q_n, to give

$$q_n = \frac{1}{2} a \delta_{n,\pm 1} + q_n^{(2)} + q_n^{(3)} + \dots, \tag{2.257}$$

where, from equations 2.239and 2.256,

$$\begin{aligned} q_n^{(\nu)} &\backsimeq O\left(\varepsilon^{\nu}\right) \\ &\backsimeq O\left(a^{\nu}\right). \end{aligned} \tag{2.258}$$

Lemma 2: The Fourier coefficients $q_{n\geq 0}$ are of order

$$q_0 \simeq O\left(a^2\right) \quad , \quad q_{n\geq 1} \simeq O\left(a^n\right). \tag{2.259}$$

Proof: From equations 2.253 and 2.257, we trivially have

$$\begin{aligned} q_1 &= q_{-1} \\ &\simeq O\left(a\right). \end{aligned} \tag{2.260}$$

From equation 2.258, we also have that

$$q_n^{(\nu\geq 2)} \geq O\left(a^2\right). \tag{2.261}$$

From equation 2.257, the only Fourier coefficients for which perturbative corrections of order lower than $\nu = 2$ contribute are q_1 and q_{-1}. Then, from equation 2.261,

$$\begin{aligned} q_{n\neq\pm 1} &\geq O\left(q_n^{(\nu\geq 2)}\right) \\ &\geq O\left(a^2\right). \end{aligned} \tag{2.262}$$

Consider the Fourier coefficient q_0. From equation 2.252, the first Kronecker delta that is not necessarily equal to 0 is the first one in the equation that appears, δ_{0,n_1+n_2}, with $n_1 = \pm 1$ and $n_2 = \mp 1$. Substituting into equation 2.252 (and rearranging slightly), we have

$$\begin{aligned} q_0 &\propto \delta_{0,1-1} q_1 q_{-1} + \dots \\ &= q_1^2 + \dots \\ &\leq O\left(a^2\right), \end{aligned} \tag{2.263}$$

where lines 2 and 3 both follow from equation 2.260. Applying equation 2.262, we have

$$q_0 \simeq O\left(a^2\right). \tag{2.264}$$

By nearly identical logic, rearranging equation 2.252 to find the order of the Fourier coefficient q_2 gives

$$\begin{aligned} q_2 &\propto \delta_{0,1+1} q_1 q_1 + \ldots \\ &= q_1^2 + \ldots \\ &\leq O\left(a^2\right). \end{aligned} \tag{2.265}$$

Applying equation 2.262, we have

$$q_2 \simeq O\left(a^2\right). \tag{2.266}$$

From equations 2.260, 2.262, and 2.266, we have that

$$O\left(q_{i+1}\right) \simeq O\left(a q_i\right) \quad , \quad O\left(q_{l>i}\right) > O\left(q_i\right) \quad , \quad (i=1). \tag{2.267}$$

If we can also show that

$$O\left(q_{i+1}\right) \quad \simeq \quad O\left(a q_i\right) \quad , \quad O\left(q_{l>i}\right) > O\left(q_i\right) \quad , \quad (i=1,\ldots j-1) \tag{2.268}$$

implies that

$$O\left(q_{j+1}\right) \simeq O\left(a q_j\right) \quad , \quad O\left(q_{l>j}\right) > O\left(q_j\right), \tag{2.269}$$

then the proof of the lemma will have been completed as a proof by induction.

A rearranged version of equation 2.252 that will be convenient for the rest of this proof is written below:

$$q_{j+1} = \beta_1 \sum_{n_1 n_2} \delta_{j+1, n_1+n_2} q_{n_1} q_{n_2} + \beta_2 \sum_{n_1 n_2 n_3} \delta_{j+1, n_1+n_2+n_3} q_{n_1} q_{n_2} q_{n_3} + \ldots, \tag{2.270}$$

where each β_i is some constant. Assume equation 2.268, and consider the β_1 term in equation 2.270. Clearly, one possible value of (n_1, n_2) for which the β_1 Kronecker delta is nonzero is

$$(n_1, n_2) = (1, j). \tag{2.271}$$

From equations 2.260 and 2.268, the order of the resulting product of Fourier coefficients is trivially

$$\begin{aligned} O\left(q_1 q_j\right) &\simeq O\left(a^1 a^j\right) \\ &\simeq O\left(a^{j+1}\right). \end{aligned} \tag{2.272}$$

All other values of (n_1, n_2) for which the β_1 Kronecker delta is nonzero can be found by incrementing n_1 by a number k , decrementing n_2 by k, and finding permutations the resulting numbers. Again from equations 2.260 and 2.268, the order of the resulting product of Fourier coefficients is trivially[4]

$$\begin{aligned} O\left(q_{1+k} q_{j-k}\right) &\simeq O\left(a^{1+k} a^{j-k}\right) \\ &\simeq O\left(a^{j+1}\right). \end{aligned} \tag{2.273}$$

[4]Here, if we have $k=j$, we will have terms of $O\left(q_{j+1} q_0\right) \simeq O\left(q_{j+1} a^2\right)$ from equation 2.264. If we have $k>j$, then from equation 2.253 and the second part of equation 2.268, we will have terms of at least $O\left(q_{j+1} a\right)$. In both of these cases, we have contributions to q_{j+1} that are clearly of greater order than $O\left(q_{j+1}\right)$, so we can ignore these cases in determining $O\left(q_{j+1}\right)$. Similar comments apply to the $\beta_{n>1}$ Kronecker deltas referred to in the rest of this proof.

So, the lowest order contribution to q_{j+1} that can be attained from the β_1 Kronecker delta is of order $O\left(a^{j+1}\right)$. If we go to the β_2 Kronecker delta, we can take any (n_1, n_2) for which the β_1 Kronecker delta was nonzero, decrement n_1 or n_2, and add an $n_3 = 1$ term to the tuplet to assure that the β_2 Kronecker delta is nonzero, since if $n_1 + n_2 = j + 1$, then clearly $1 + n_1 + (n_2 - 1) = j + 1$. We can denote this as

$$(n_1, n_2) \to (1, n_1, n_2 - 1)\,. \tag{2.274}$$

For $(n_1, n_2) = (1, j)$, this is writeable as

$$(1, j) \to (1, 1, j - 1)\,. \tag{2.275}$$

This can be repeated if we move to the β_{1+m} Kronecker delta, to give

$$(1, j) \to \big(1, \underbrace{1, \ldots, 1}_{m \text{ times}}, j - m\big). \tag{2.276}$$

Other terms for which the β_{1+m} Kronecker deltas are nonzero can be found by incrementing any number of the terms in in equation 2.276 any number of times while decrementing any other the same number of times, and then permuting the terms. By logic analogous to that seen in equation 2.273, this will not change the order of the resulting term (except for cases analogous to those discussed in footnote 4, which result in terms of order higher than $O\left(q_{j+1}\right)$). Then, For $m < j$, we have

$$\begin{aligned} O\left(q_{j+1}\right) &\simeq O\Bigg(q_1 \underbrace{q_1 \ldots q_1}_{m \text{ times}} q_{j-m}\Bigg) \\ &\simeq O\Bigg(a^1 \underbrace{a^1 \ldots a^1}_{m \text{ times}} a^{j-m}\Bigg) \\ &\simeq O\left(a^{j+1}\right), \end{aligned} \tag{2.277}$$

so that the lowest-order contributions that can be attained from the Kronecker deltas corresponding to any of β_1 through β_j are of order $O\left(a^{j+1}\right)$. We can ignore the $m \geq j$ case in determining $O\left(q_{j+1}\right)$, since

$$\begin{aligned} O\Bigg(q_1 \underbrace{q_1 \ldots q_1}_{m \text{ times}} q_{j-m}\Bigg) &\simeq O\left(a^{j+1} a^{m-j} q_{m-j}\right) \\ &\simeq O\left(a^{m+1} q_{m-j}\right) \\ &> O\left(a^{j+1}\right), \end{aligned} \tag{2.278}$$

where line 3 follows from the fact that since $m \geq j$, $O\left(a^{m+1}\right) \geq O\left(a^{j+1}\right)$, and that from equation 2.262, $O\left(q_{m-j}\right) \geq O\left(a\right)$.

Observing equations 2.277 and 2.278, we have that

$$O\left(q_{j+1}\right) \simeq O\left(a q_j\right), \tag{2.279}$$

which is the first part of equation 2.269. The second part of equation 2.269,

$$O\left(q_{l>j}\right) > O\left(q_j\right), \tag{2.280}$$

can be shown by repeating the logic between equations 2.270 and 2.278 almost exactly, but with each j replaced with $j + j_{\text{extra}}$, where $j_{\text{extra}} \geq 1$. Then, in place of equation 2.279, we will have

$$O\left(q_{j+j_{\text{extra}}+1}\right) \geq O\left(a q_j\right), \tag{2.281}$$

which is equivalent to equation 2.280. □

Corollary 1: The Fourier coefficients $q_{n\leq -1}$ are of order

$$q_{n\leq -1} \simeq O\left(a^{-n}\right). \tag{2.282}$$

Proof: The corollary follows directly from equations 2.253 and 2.259. □

Equations 2.258 and 2.259, together with equation 2.252, are sufficient to solve for each Fourier coefficient, as well as the frequency ω, to arbitrary finite order in ε or a. After these are found, they can be substituted into equations 2.235 and 2.236 to find the solution to the system $q\left(t\right)$ to that order.

The rest of this subsection is devoted to finding low-order solutions for the Fourier coefficients and frequency. From equation 2.252, the $n = 1$ equation is given by

$$\begin{aligned} &\left(\omega^2 - \omega_0^2\right)\frac{a}{2} = O\left(a^2\right) \\ \Longrightarrow \quad &\omega^2 = \omega_0^2 + O\left(a\right). \end{aligned} \tag{2.283}$$

Similarly, equation 2.252 also gives that for $n \neq 1$,

$$q_n = \frac{1}{n^2\omega^2 - \omega_0^2}\left[\alpha_3 \sum_{n_1 n_2} \delta_{n,n_1+n_2} \underbrace{q_{n_1} q_{n_2}}_{O(a^2)} + \alpha_4 \sum_{n_1 n_2 n_3} \delta_{n,n_1+n_2+n_3} \underbrace{q_{n_1} q_{n_2} q_{n_3}}_{O(a^3)} + \ldots\right]. \tag{2.284}$$

With equation 2.284, we can solve for each q_n up to second order:

$$\begin{aligned} q_0 &= -\frac{1}{\omega_0^2}\left[2\alpha_3 q_1 q_{-1} + O\left(a^3\right)\right] \\ &= -\alpha_3 \frac{a^2}{2\omega_0^2} + O\left(a^3\right), \end{aligned} \tag{2.285}$$

$$\begin{aligned} q_2 &= \frac{1}{4\omega^2 - \omega_0^2}\left[\alpha_3 q_1^2 + O\left(a^3\right)\right] \\ &= \frac{1}{3\omega_0^2}\alpha_3 \frac{a^2}{4} + O\left(a^3\right) \\ &= \alpha_3 \frac{a^2}{12\omega_0^2} + O\left(a^3\right), \end{aligned} \tag{2.286}$$

$$q_{n\geq 3} = O\left(a^3\right), \tag{2.287}$$

where the second line of equation 2.286 follows from the second line of equation 2.283.

From equation 2.236, as well as equations 2.285-2.287, the solution for $q(t)$ to second order is

$$q(t) = -\frac{\alpha_3}{2\omega_0^2}a^2 + a\cos\omega t + \frac{\alpha_3}{6\omega_0^2}a^2\cos 2\omega t + O\left(a^3\right). \tag{2.288}$$

The frequency can also be solved to second order. Substituting the q_1 term into the first line of the $n=1$ equation 2.283, and using equation 2.235, gives

$$\left(\omega^2 - \omega_0^2\right) q_1 = \alpha_3\left[2q_0q_1 + 2q_2q_{-1} + O\left(a^5\right)\right] + \alpha_4\left[3q_1^2q_{-1}^2 + O\left(a^5\right)\right]. \tag{2.289}$$

Applying equation 2.253,

$$\begin{aligned}\omega^2 - \omega_0^2 &= 2\alpha_3\left(q_0 + q_2\right) + 3\alpha_4 q_1^2 + O\left(a^4\right)\\ &= 2\alpha_3\left(-\alpha_3\frac{a^2}{2\omega_0^2} + \alpha_3\frac{a^2}{12\omega_0^2}\right) + 3\alpha_4\frac{a^2}{4} + O\left(a^4\right)\\ &= -\frac{5}{6}\left(\frac{\alpha_3 a}{\omega_0}\right)^2 + \frac{3}{4}\alpha_4 a^2 + O\left(a^4\right).\end{aligned} \tag{2.290}$$

Solving for ω^2,

$$\begin{aligned}\omega^2 &= \omega_0^2 + a^2\left(\frac{3}{4}\alpha_4 - \frac{5}{6}\frac{\alpha_3^2}{\omega_0^2}\right) + O\left(a^2\right)\\ &= \omega_0^2\left[1 + a^2\left(\frac{3}{4}\frac{\alpha_4}{\omega_0^2} - \frac{5}{6}\frac{\alpha_3^2}{\omega_0^4}\right) + O\left(a^4\right)\right],\end{aligned} \tag{2.291}$$

so that

$$\omega = \omega_0 + a^2\left(\frac{3}{8}\frac{\alpha_4}{\omega_0} - \frac{5}{6}\frac{\alpha_3^2}{\omega_0^3}\right) + O\left(a^4\right). \tag{2.292}$$

Example 1: Let $\alpha_4 = 0$, $\alpha_3 > 0$. Then, $q_0 < 0$, i.e., the average position is shifted to the left of that for the harmonic oscillator. Graphing the potential, it becomes flatter then that of the harmonic oscillator, so that ω decreases.

Example 2: Let $\alpha_4 = 0$, $\alpha_3 < 0$. Then, $q_0 > 0$, i.e., the average position is shifted to the right of that for the harmonic oscillator. The graph of the potential is again flatter then that of the harmonic oscillator, so that ω decreases.

Example 3: Let $\alpha_4 \gtrless 0$, $\alpha_3 = 0$. Then, $q_0 = 0$, i.e., the average position is not shifted.The potential is steeper or flatter, respectively, so that ω increases or decreases, respectively.

Remark 5*:* The method presented in this subsection allows for the systematic calculation of $q(t)$ and ω up to any desired order in a. See problem 30.

§ 5.5 Perturbation theory for the central field problem

Consider bounded motion in a central potential

$$U_\varepsilon(r) = U(r) + \varepsilon V(r) \quad , \quad \varepsilon \ll 1. \tag{2.293}$$

This is a problem that has been solved for $\varepsilon = 0$. For example, in the case that $U(r) \propto 1/r$, the equation of orbit was found in §4.4 (see equation 2.174) to be

$$\begin{aligned} \varphi - \varphi_0 &= \frac{l}{\sqrt{2m}} \int_{r_0}^{r} \frac{dx}{x^2 \sqrt{E - U_{\text{eff}}(x)}} \\ &= \frac{l}{\sqrt{2m}} \int_{r_0}^{r} \frac{dx}{x^2 \sqrt{E - l^2/2mx^2 - U(x)}}. \end{aligned} \tag{2.294}$$

In the case that we have $U_\varepsilon(r)$ instead of $U(r)$, the same equation of orbit will apply with $U_\varepsilon(r)$,

$$\varphi - \varphi_0 = \frac{l}{\sqrt{2m}} \int_{r_0}^{r} \frac{dx}{x^2 \sqrt{E - l^2/2mx^2 - U_\varepsilon(x)}}. \tag{2.295}$$

Let $r_\pm$ be the turning points of the radial motion. From §4.5, during one period, the pericenter (as well as the apocenter) rotates by an angle

$$\Phi = 2\left[\varphi(r_+) - \varphi(r_-)\right], \tag{2.296}$$

the **perihelion advance**.

Remark 1*:* In Kepler's problem, $\Phi = 2\pi$. Letting $\delta\Phi$ denote the first-order correction to Φ,

$$\begin{aligned} \Phi(\varepsilon) &= \frac{2l}{\sqrt{2m}} \int_{r_-(\varepsilon)}^{r_+(\varepsilon)} \frac{dx}{x^2 \sqrt{E - l^2/2mx^2 - U(x) - \varepsilon V(x)}} \\ &= \Phi + \varepsilon\delta\Phi + O\left(\varepsilon^2\right). \end{aligned} \tag{2.297}$$

Remark 2*:* Note that r_+ and r_- depend on ε.

It is desirable to determine $\delta\Phi$. It is convenient to use (each $\sqrt{\cdots}$ below represents the square root on the LHS of the equation).

$$\begin{aligned} \frac{\partial}{\partial l} \int_{r_-}^{r_+} dx \sqrt{E - l^2/2mx^2 - U_\varepsilon(x)} &= \frac{\partial r_+}{\partial l} \underbrace{\sqrt{\cdots}\Big|_{x=r_+}}_{=0} - \frac{\partial r_-}{\partial l} \underbrace{\sqrt{\cdots}\Big|_{x=r_-}}_{=0} \\ &\quad - \frac{l}{2m} \int_{r_-}^{r_+} \frac{dx}{x^2 \sqrt{\cdots}}, \end{aligned} \tag{2.298}$$

where each of the indicated quantities are 0 by the definition of $r_\pm$ as being located where $U(r) =$

$U_{\text{eff}}(r)$. Using this result in conjunction with equation 2.297,

$$\begin{aligned}
\Phi(\varepsilon) &= -2\sqrt{2m}\frac{\partial}{\partial l}\int_{r_-}^{r_+} dx\,\sqrt{E - l^2/2mx^2 - U(x) - \varepsilon V(x)} \\
&= -2\sqrt{2m}\frac{\partial}{\partial l}\left[\int_{r_-}^{r_+} dx\sqrt{\cdots}\bigg|_{\varepsilon=0} + \varepsilon\frac{d}{d\varepsilon}\bigg|_{\varepsilon=0}\int_{r_-}^{r_+} dx\,V\sqrt{\cdots} + O\left(\varepsilon^2\right)\right] \\
&= \Phi - \varepsilon 2\sqrt{2m}\frac{\partial}{\partial l}\left[\frac{\partial r_+}{\partial l}\underbrace{\sqrt{\cdots}\bigg|_{x=r_+}}_{=0} - \frac{\partial r_-}{\partial \varepsilon}\underbrace{\sqrt{\cdots}\bigg|_{x=r_-}}_{=0}\right. \\
&\quad \left. -\frac{1}{2}\int_{r_-}^{r_+} dx\,V(x)/\sqrt{\cdots}\bigg|_{\varepsilon=0}\right] + O\left(\varepsilon^2\right) \\
&= \Phi + \varepsilon\sqrt{2m}\frac{\partial}{\partial l}\int_{r_-}^{r_+} dx\,\frac{V(x)}{\sqrt{E - l^2/2mx^2 - U(x)}} + O\left(\varepsilon^2\right).
\end{aligned} \tag{2.299}$$

Identifying the ε coefficient as $\delta\Phi$,

$$\delta\Phi = \sqrt{2m}\frac{\partial}{\partial l}\int_{r_-}^{r_+} dx\,\frac{V(x)}{\sqrt{E - l^2/2mx^2 - U(x)}}. \tag{2.300}$$

Equation 2.295 gives

$$\begin{aligned}
\frac{d\varphi}{dr} &= \frac{l}{\sqrt{2m}}\frac{1}{r^2\sqrt{E - U_{\text{eff}}}} \\
\Longrightarrow \quad \frac{dr}{\sqrt{E - U_{\text{eff}}(r)}} &= \frac{\sqrt{2m}}{l}r^2 d\varphi \\
\Longrightarrow \quad \delta\Phi &= 2m\frac{\partial}{\partial l}\frac{1}{l}\int_0^{\Phi/2} d\varphi\, r^2(\varphi)\, V(r(\varphi)),
\end{aligned} \tag{2.301}$$

where equation 2.300 was used in the last line. Note that $r(\varphi)$ is the orbit for the unperturbed case. We have therefore found the perihelion advance to $O(\varepsilon)$.

Remark 3*:* In equation 2.301, $\delta\Phi$ is expressed in terms of the purturbing potential V and the orbit $r(\varphi)$ of the unperturbed problem.

§ 5.6 Perturbations of Kepler's problem

Recall Kepler's problem and its solution from §4,

$$\begin{aligned}
U(r) &= -\frac{\alpha}{r} \\
\Longrightarrow \quad \Phi = 2\pi \quad &, \quad r(\varphi) = \frac{p}{1 + e\cos\varphi},
\end{aligned} \tag{2.302}$$

with

$$p = \frac{l^2}{m\alpha} \quad , \quad e = \sqrt{1 + 2l^2E/m\alpha^2}. \tag{2.303}$$

Let a perturbing potential $V(r)$ be of the form

$$V(r) = \frac{\beta}{r^{n+1}}, \tag{2.304}$$

i.e., the perturbing potential is a power-law correction to the $1/r$-potential. From §5.5,

$$\begin{aligned}\delta\Phi &= 2m\frac{\partial}{\partial l}\frac{1}{l}\int_0^\pi d\varphi\, r^2(\varphi)\,\beta\,(r(\varphi))^{-n-1} \\ &= 2m\beta\frac{\partial}{\partial l}\frac{1}{l}p^{1-n}\int_0^\pi d\varphi\,(1+e\cos\varphi)^{n-1}.\end{aligned} \tag{2.305}$$

Remark 1*:* Checking the case for which $n = 0$, we have

$$\begin{aligned}\delta\Phi &\propto \frac{\partial}{\partial l}l\int_0^\pi \frac{d\varphi}{1+e\cos\varphi} \\ &= \int_0^\pi \frac{d\varphi}{1+e\cos\varphi} - l\frac{1}{2e}\frac{4lE}{m\alpha^2}\int_0^\pi d\varphi\,\frac{\cos\varphi}{(1+e\cos\varphi)^2} \\ &= \int_0^\pi \frac{d\varphi}{1+e\cos\varphi} - \frac{1}{e}\left(e^2-1\right)\int_0^\pi d\varphi\,\frac{\cos\varphi}{(1+e\cos\varphi)^2} \\ &= \frac{1}{e}\int_0^\pi d\varphi\,\frac{\left(1-e^2\right)\cos\varphi + e + e^2\cos\varphi}{(1+e\cos\varphi)^2} \\ &= \frac{1}{e}\int_0^\pi d\varphi\,\frac{e+\cos\varphi}{(1+e\cos\varphi)^2} \\ &= 0,\end{aligned} \tag{2.306}$$

where the last line follows from theory external to mechanics, and can be found in texts such as Gradshteyn and Ryzhik.

Remark 2*:* Because p and e are elementary functions of l, the problem was reducible to a one-dimensional integral above.

Example 1: Let $n = 1$ and $\beta = -1$, so that the power-law correction to the $1/r$-potential is an attractive $1/r^2$-potential. Then (see Figure 2.14 below),

$$\begin{aligned}\delta\Phi &= -2m\frac{\partial}{\partial l}\frac{1}{l}\pi \\ &= 2\pi m/l^2 \\ &> 0.\end{aligned} \tag{2.307}$$

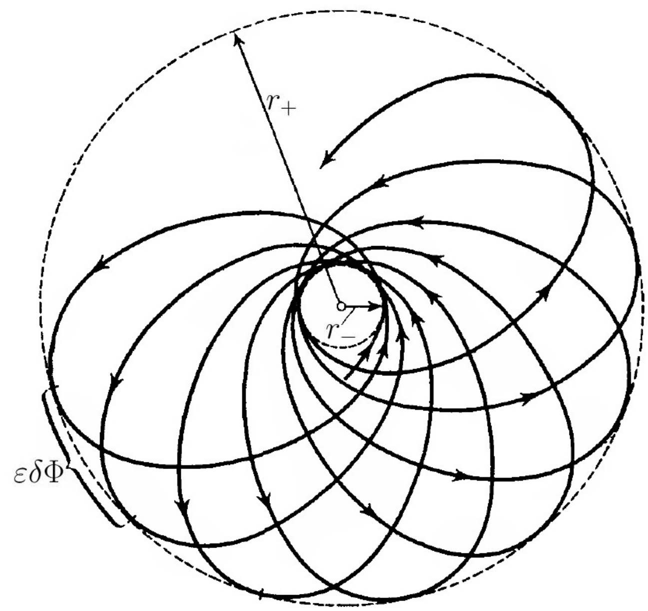

Figure 2.14: Orbit for a Kepler potential with a $-1/r^2$ perturbation, as in Example 1. Adapted from Landau and Lifschitz Third Edition, Figure 9.

§ 5.7 Digression - introduction to potential theory

§ 5.7.1 Poisson's equation

Let $\rho(\boldsymbol{y})$ be a mass density distribution.

Proposition 1: The potential $U(\boldsymbol{x})$ felt by a test mass m at point $\boldsymbol{x}$ due to the mass density distribution $\rho(\boldsymbol{y})$ is

$$U(\boldsymbol{x}) = -Gm \int d\boldsymbol{y}\, \frac{\rho(\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|}. \tag{2.308}$$

Proof: If there were a mass M at a point $\boldsymbol{y}$, we would have

$$U_{\text{single}}(\boldsymbol{x}) = -Gm \frac{M}{|\boldsymbol{x}-\boldsymbol{y}|}. \tag{2.309}$$

Similarly, with N point masses M_i at points $\boldsymbol{y}_i$, we would have

$$U_{\text{discrete}}(\boldsymbol{x}) = -Gm \sum_i \frac{M_i}{|\boldsymbol{x}_i-\boldsymbol{y}_i|}. \tag{2.310}$$

Taking the continuum limit to go from discrete masses at discrete positions to a mass density distribution,

$$\begin{aligned}
& U_{\text{discrete}}(\boldsymbol{x}) = -Gm \sum_i \frac{M_i}{|\boldsymbol{x}_i - \boldsymbol{y}_i|} \\
\implies \quad & U(\boldsymbol{x}) = -Gm \int d\boldsymbol{y}\, \frac{\rho(\boldsymbol{y})}{|\boldsymbol{x} - \boldsymbol{y}|},
\end{aligned} \tag{2.311}$$

which is equation 2.308. □

Remark 1: Equation 2.308 is called **Poisson's equation**. It is important in electricity and magnetism as well.

Remark 2: The principle underlying its derivation above is that linear superposition can be applied to potentials. This is a nontrivial assumption.

§ 5.7.2 Spherically symetric density distributions

Consider a mass density distribution that is writeable as

$$\rho(\boldsymbol{x}) = \rho(r) \quad , \quad r = |\boldsymbol{x}| . \tag{2.312}$$

Coordinates can be chosen, without loss of generality, such that we can write $\boldsymbol{x}$ and $\boldsymbol{y}$ with spherical coordinates, though still using a Cartesian basis, as

$$\boldsymbol{x} = (0, 0, r) \quad , \quad \boldsymbol{y} = y(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta), \tag{2.313}$$

so that

$$\begin{aligned}
|\boldsymbol{x} - \boldsymbol{y}| &= \sqrt{\boldsymbol{x}^2 + \boldsymbol{y}^2 - 2\boldsymbol{x}\cdot\boldsymbol{y}} \\
&= \sqrt{r^2 + y^2 - 2ry\cos\theta}.
\end{aligned} \tag{2.314}$$

Then, the potential felt by a test mass m at point $\boldsymbol{x}$ resulting from the mass density distribution is writeable as

$$\begin{aligned}
U(\boldsymbol{x}) &= U(r) \\
&= -Gm \int_0^{2\pi} d\varphi \int_{-1}^{1} d\cos\theta \int_0^\infty dy\, y^2 \frac{\rho(y)}{\sqrt{r^2 + y^2 - 2ry\cos\theta}} \\
&= -2\pi Gm \int_0^\infty dy\, y^2 \rho(y) \int_{-1}^{1} d\eta\, \frac{1}{\sqrt{r^2 + y^2 - 2ry\eta}} \\
&= -2\pi Gm \int_0^\infty dy\, y^2 \rho(y) \frac{-2}{2ry} \sqrt{r^2 + y^2 - 2ry\eta}\Big|_{-1}^{1} \\
&= 2\pi \frac{Gm}{r} \int_0^\infty dy\, y\rho(y) \left(|r - y| - |r + y|\right) \\
&= 2\pi \frac{Gm}{r} \left[\int_0^r dy\, y\eta(y)(-2y) + \int_r^\infty dy\, y\eta(y)(-2r)\right] \\
&= -4\pi \frac{Gm}{r} \left[\int_0^r dy\, y^2\eta(y) + r\int_r^\infty dy\, y\eta(y)\right],
\end{aligned} \tag{2.315}$$

or, equivalently,

$$U(r) = -Gm\left[\frac{1}{r}Q(r) + P(r)\right], \tag{2.316}$$

with

$$Q(r) = \int d\boldsymbol{y}\,\eta(y)\,\Theta(r-y) \quad , \quad P(r) = \int d\boldsymbol{y}\,\frac{1}{y}\eta(y)\,\theta(y-r)\,. \tag{2.317}$$

Above, $Q(r)$ is the potential at r produced by whatever mass is inside a sphere of radius r, and $P(r)$ is the potential at r produced by whatever mass is outside a sphere of radius r.

Remark 1*:* Let $\rho(y) = 0$ for $y > R$. Then,

$$\begin{aligned} & P(r>R) = 0 \quad , \quad Q(r>R) = \int d\boldsymbol{y}\,\rho(\boldsymbol{y}) \equiv M, \\ \Longrightarrow \quad & U(r>R) = -\frac{GmM}{r}. \end{aligned} \tag{2.318}$$

This is the same potential as is produced by a point mass M.

§ 5.7.3 Multipole expansion

Consider a bounded mass density distribution,

$$\rho(\boldsymbol{y}) = 0\,\text{for}\ |\boldsymbol{y}| \gg R. \tag{2.319}$$

Question. What is $U(\boldsymbol{x})$ for $|\boldsymbol{x}| \equiv r > R$?

Noting that

$$\boldsymbol{x} \simeq O(r) \quad , \quad \boldsymbol{y} \simeq O(R)\,, \tag{2.320}$$

the poential $U(\boldsymbol{x})$ can be found for $|\boldsymbol{x}| \equiv r \gg R$ (implying that $U(\boldsymbol{x})$ will be approximately spherically symmetric) as follows:

$$\begin{aligned} \frac{1}{|\boldsymbol{x}-\boldsymbol{y}|} &\approx \frac{1}{\sqrt{r^2 - 2\boldsymbol{x}\cdot\boldsymbol{y} + \boldsymbol{y}^2}} \\ &= \frac{1}{r}\left[1 - \frac{2\boldsymbol{x}\cdot\boldsymbol{y}}{r^2} + \frac{\boldsymbol{y}^2}{r^2}\right]^{-1/2} \\ &= \frac{1}{r}\left[1 + \frac{\boldsymbol{x}\cdot\boldsymbol{y}}{r^2} - \frac{1}{2}\frac{\boldsymbol{y}^2}{r^2} + \frac{3}{8}\frac{(2\boldsymbol{x}\cdot\boldsymbol{y})^2}{r^4} + O\left(\frac{R^3}{r^3}\right)\right] \\ &= \frac{1}{r}\left[1 + \frac{1}{r}\frac{\boldsymbol{x}\cdot\boldsymbol{y}}{r} + \frac{1}{r^2}\left(\frac{3}{2}\frac{(\boldsymbol{x}\cdot\boldsymbol{y})^2}{r^2} - \frac{1}{2}\boldsymbol{y}^2\right) + O\left(\frac{R^3}{r^3}\right)\right]. \end{aligned} \tag{2.321}$$

Applying equation 2.308,

$$
\begin{aligned}
-\frac{1}{Gm}U(\boldsymbol{x}) &= \int d\boldsymbol{y}\frac{\rho(\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|} \\
&= \frac{1}{r}\int d\boldsymbol{y}\,\rho(\boldsymbol{y}) + \frac{1}{r^2}\sum_{i=1}^{3}\frac{x_i}{r}\int d\boldsymbol{y}\,y_i\rho(\boldsymbol{y}) \\
&\quad + \frac{1}{r^2}\sum_{i,j=1}^{3}\left[\frac{3}{2}\frac{x_ix_j}{r^2}-\frac{1}{2}\delta_{ij}\right]\int d\boldsymbol{y}\,y_iy_j\rho(\boldsymbol{y}) + O\left(r^{-4}\right),
\end{aligned} \tag{2.322}
$$

so that

$$
U(\boldsymbol{x}) = -Gm\left[\frac{M}{r}+\frac{1}{r^2}\sum_{i=1}^{3}\frac{x_i}{r}D_i+\frac{1}{r^3}\sum_{i,j=1}^{3}\frac{x_ix_j}{r^2}Q_{ij}+O\left(r^{-4}\right)\right], \tag{2.323}
$$

where

$$
M \equiv \int d\boldsymbol{y}\,\rho(\boldsymbol{y}) \tag{2.324}
$$

is the **monopole moment** (which is equal to the **total mass**),

$$
D_i \equiv \int d\boldsymbol{y}\,y_i\rho(\boldsymbol{y}) \tag{2.325}
$$

is the i'th component of the **dipole moment**, and

$$
Q_{ij} = \frac{1}{2}\int d\boldsymbol{y}\,\left(3y_iy_j-\delta_{ij}y^2\right)\rho(\boldsymbol{y}). \tag{2.326}
$$

is the ij'th component of the **quadrupole moment**.

Remark 1*:* At a large distance from the source, the potential is given by a multipole expansion.

Remark 2*:* The fact that $Q_{ij} = Q_{ji}$ implies that the matrix Q can be diagonalized by an orthogonal transformation. Then, without loss of generality, coordinates can be chosen such that

$$
Q = \begin{pmatrix} Q_{11} & 0 & 0 \\ 0 & Q_{22} & 0 \\ 0 & 0 & Q_{33} \end{pmatrix}. \tag{2.327}
$$

Furthermore, consider that

$$
\begin{aligned}
& \operatorname{tr}(Q) = \sum_i Q_{ii} = \frac{1}{2}\int d\boldsymbol{y}\,\left(3\boldsymbol{y}^2-3\boldsymbol{y}^2\right)\rho(\boldsymbol{y}) = 0 \\
\Longrightarrow\quad & Q_{33} = -Q_{11}-Q_{22}.
\end{aligned} \tag{2.328}
$$

Then, Q is fully characterized by two numbers.

Remark 3*:* Note that the monopole moment is a scalar, the dipole moment is a vector (with components D_i), and the quadrupole moment is a matrix (with components Q_{ij}).

§ 5.7.4 The quadrupole moment of a spheroid

Consider a spheroid (i.e., an ellipsoid with two equal axes) of uniform mass density. The spheroid is given by the equation

$$\left(\frac{x}{R_1}\right)^2+\left(\frac{y}{R_1}\right)^2+\left(\frac{z}{R_3}\right)^2=1 \tag{2.329}$$

in Cartesian coordinates, or equivalently

$$\left(\frac{r}{R_1}\right)^2+\left(\frac{z}{R_3}\right)^2=1 \tag{2.330}$$

in cylindrical coordinates, and the equation for the charge density is

$$\begin{aligned}\rho\left(\boldsymbol{y}\right) &= \rho\left(-\boldsymbol{y}\right)\\ &= \begin{cases}\rho, & \boldsymbol{y}\in\text{spheroid}\\ 0, & \boldsymbol{y}\notin\text{spheroid}\end{cases}.\end{aligned} \tag{2.331}$$

The mass distribution's monopole moment is

$$\begin{aligned}M &= \rho V_{\text{spheroid}}\\ &= \frac{4\pi}{3}\rho R_1^2 R_3,\end{aligned} \tag{2.332}$$

and its dipole moment is

$$\begin{aligned}\boldsymbol{D} &= \int d\boldsymbol{y}\,\boldsymbol{y}\rho\left(\boldsymbol{y}\right)\\ &= \int d\boldsymbol{y}\,\left(-\boldsymbol{y}\right)\rho\left(-\boldsymbol{y}\right)\\ &= -\int d\boldsymbol{y}\,\boldsymbol{y}\rho\left(\boldsymbol{y}\right)\\ &= 0,\end{aligned} \tag{2.333}$$

where line 3 follows from equation 2.331, and line 4 follows from comparing lines 1 and 3.

The quadrupole moment of the mass distribution can be found as follows: We have

$$\rho\left(\boldsymbol{y}\right)=\rho\left(r,z\right), \tag{2.334}$$

with

$$y_1=r\cos\varphi \quad , \quad y_2=r\sin\varphi \quad , \quad y_3=z. \tag{2.335}$$

From equation 2.326,

$$\begin{aligned}Q_{ij} &= \frac{1}{2}\int d\boldsymbol{y}\,\left(3y_iy_j-\delta_{ij}y^2\right)\rho\left(\boldsymbol{y}\right)\\ &= \frac{1}{2}\int_0^{2\pi}d\varphi\int_0^{\infty}dr\,r\int_{-\infty}^{\infty}dz\,\left(3y_iy_j-\delta_{ij}y^2\right)\rho\left(r,z\right),\end{aligned} \tag{2.336}$$

so that

$$Q_{ij} \propto \begin{cases} \int_0^{2\pi} d\varphi \, \cos\varphi \sin\varphi = 0, & i,j = 1,2 \\ \int_0^{2\pi} d\varphi \, \cos\varphi = 0, & i,j = 1,3 \\ \int_0^{2\pi} d\varphi \, \sin\varphi = 0, & i,j = 2,3 \\ \int_0^{2\pi} d\varphi \, \cos^2\varphi = \int_0^{2\pi} d\varphi \, \sin^2\varphi \propto Q_{22} & i,j = 1,1 \end{cases} . \tag{2.337}$$

Recalling that the matrix (Q) is symmetric and equation 2.328 from §5.7.3 Remark 2, we have that the (Q) is characterized by a single number, Q, given by

$$Q = Q_{11} = Q_{22} = -\frac{1}{2} Q_{33}. \tag{2.338}$$

Then, from equations 2.330, 2.331, and 2.336,

$$\begin{aligned} Q &= -\frac{1}{2} Q_{33} \\ &= -\frac{1}{4} 2\pi \int_0^{\infty} dr\, r \int_{-\infty}^{\infty} dz \, \left[3z^2 - \left(r^2 + z^2\right)\right] \rho \Theta \left[-\left(\frac{r}{R_1}\right)^2 + \left(\frac{z}{R_3}\right)^2 + 1 \right] \\ &= -\frac{\pi}{2} \rho \int_0^{R_1} dr\, r 2 \int_0^{R_3 \sqrt{1 - r^2/R_1^2}} dz \, \left(2z^2 - r^2\right) \\ &= -\pi \rho R_1^2 R_3 \int_0^1 dr\, r \int_0^{\sqrt{1-r^2}} dz \, \left(2R_3^2 z^2 - R_1^2 r^2\right) \\ &= -\pi \rho R_1^2 R_3 \int_0^1 dr\, r \left[\frac{2}{3} R_3^2 \left(1 - r^2\right)^{3/2} - R_1^2 r^2 \left(1 - r^2\right)^{1/2} \right] \\ &= -\frac{\pi}{2} \rho R_1^2 R_3 \int_0^1 dx \, \left[\frac{2}{3} R_3^2 \left(1 - x\right)^{3/2} - R_1^2 x \left(1 - x\right)^{1/2} \right] \\ &= -\frac{3}{8} M \left[\frac{2}{3} R_3^2 \frac{2}{5} - R_1^2 \frac{4}{5} \right] \\ &= M \frac{1}{8} \frac{4}{5} \left(R_1^2 - R_3^2\right) \\ &= \frac{M}{10} \left(R_1 + R_3\right) \left(R_1 - R_3\right). \end{aligned} \tag{2.339}$$

Defining **oblateness** η as

$$\eta \equiv \left(R_1 - R_3\right) / R_3 \tag{2.340}$$

allows equation 2.339 to be written as

$$\begin{aligned} Q &= \frac{M}{10} \eta R_3 \left(R_1 + R_3\right) \\ &= \frac{M}{5} R^2 \eta + O\left(\eta^2\right), \end{aligned} \tag{2.341}$$

with

$$\begin{aligned} R_1 &= R_3 + O\left(\eta\right) \\ &\equiv R + O\left(\eta\right). \end{aligned} \tag{2.342}$$

§ 5.8 Mercury's perihelion advance due to an oblate sun

Assume that the Sun is an oblate spheroid with oblateness

$$\eta = \frac{\delta R}{R_3} \ll 1 \quad , \quad (\delta R = R_1 - R_3 > 0) \, . \tag{2.343}$$

From §5.7, the quadropole moment is

$$(Q) = \begin{pmatrix} Q & 0 & 0 \\ 0 & Q & 0 \\ 0 & 0 & -2Q \end{pmatrix} \quad , \quad Q = \frac{M}{5} R^2 \eta + O\left(\eta^2\right) . \tag{2.344}$$

Recall from §5.7.3,

$$U(\boldsymbol{x}) = -Gm \left[\frac{M}{r} + \frac{1}{r^2} \sum_{ij} \frac{x_i x_j}{r^2} Q_{ij} + O\left(r^{-4}\right) \right] . \tag{2.345}$$

Assume that the the symmetry axis for the spheroid is perpendicular to the orbital plane. Note that this is a very nontrivial assumption. If it holds true, it simplifies the problem greatly, since it implies the the coordinates in which the quadrupole moment is simplest are also those for which the orbit lies in the $x_1 x_2$-plane. Under this assumption, we have that in the orbital plane (x_1, x_2),

$$Q_{ij} = \delta_{ij} Q \quad , \quad (i, j = 1, 2) \, , \tag{2.346}$$

$$x_1^2 + x_2^2 = r^2. \tag{2.347}$$

Then, the potential on the orbital plane is given by

$$\begin{aligned} U(\boldsymbol{x}) &= U(r) \\ &= -\frac{GmM}{r} - \frac{GmQ}{r^3} + O\left(r^{-4}\right) \\ &= -\frac{\alpha}{r} - \frac{\alpha}{M} \frac{M}{\rho} \eta \frac{R^2}{r^3} + O\left(\eta^2/r^2, r^{-4}\right) , \end{aligned} \tag{2.348}$$

where $\alpha = GmM$ is the gravitational potential prefactor, as in §4.6. Then,

$$U(r) = -\frac{\alpha}{r} + \varepsilon \frac{\beta}{r^{n+1}}, \tag{2.349}$$

with

$$\varepsilon = \eta \quad , \quad n = 2 \quad , \quad \beta = -\frac{\alpha R^2}{5}. \tag{2.350}$$

Substituting into equation 2.305 from §5.6,

$$\begin{aligned} \delta\Phi &= 2m\beta \frac{\partial}{\partial l} \frac{1}{l} \frac{1}{p} \int_0^\pi d\varphi \, \left(1 + e\cos\varphi\right) \\ &= 2\pi m\beta \frac{\partial}{\partial l} \frac{m\alpha}{l^3} \\ &= -6\pi m\beta \frac{m\alpha}{l^4} \\ &= \frac{6\pi}{5} \left(\frac{R}{p}\right)^2 , \end{aligned} \tag{2.351}$$

where the second line follows from the fact that $\int_0^\pi d\varphi\, e\cos\varphi = 0$. Continuing, the correction to the perihelion advance can be found as

$$\begin{aligned} \delta p &= \eta\delta\Phi \\ &= \frac{6\pi}{5}\left(\frac{R}{p}\right)^2 \eta + O\left(\eta^2\right). \end{aligned} \tag{2.352}$$

Discussion of Mercury's anomalous perihelion advance:

1. The radius of the Sun is $R = 6.96 \times 10^{10}\,\mathrm{cm}$. The orbit of Mercury has parameter $p = a\left(1 - e^2\right) = 5.55 \times 10^{12}\,\mathrm{cm}$ and period $T = 0.241\,\mathrm{years}$. Then,

$$\begin{aligned} \delta p &= 5.92 \times 10^{-4}\eta\,\frac{\mathrm{rad.}}{\mathrm{orbit}} \\ &= \eta \times 5.92\frac{360 \times 3,600\,\mathrm{s}}{2\pi}\frac{100}{T/\mathrm{yrs.}}\frac{''}{\mathrm{century}} \\ &= \eta \times 5.07 \times 10^4\frac{''}{\mathrm{century}}. \end{aligned} \tag{2.353}$$

2. By measuring δp and subtracting perturbations from all other planets, we can account for the **anomalous perihelion advance** of Mercury. Experiment gives

$$\delta p_{\mathrm{exp}} = 43.1 \pm 0.5\frac{''}{\mathrm{century}}. \tag{2.354}$$

This is to be compared with theory. General relativity and the Brans-Dicke theory of gravity give, respectively,

$$\delta p_{\mathrm{GR}} = 43.03\frac{''}{\mathrm{century}} \quad , \quad \delta p_{\mathrm{BD}} = 39.5\frac{''}{\mathrm{century}}. \tag{2.355}$$

3. Suppose that $\eta = 6.9 \times 10^{-5}$. Then,

$$\begin{aligned} &\delta p_\eta = 3.5\frac{''}{\mathrm{century}} \\ \Longrightarrow\quad &\delta p_{\mathrm{BD}} + \delta p_\eta = 43\frac{''}{\mathrm{century}} = \delta p_{\mathrm{exp}}. \end{aligned} \tag{2.356}$$

Comparing the above equations in this discussion indicates that without knowing the Sun's oblateness η, it is not clear which theory of gravity Mercury's perihelion advance favors.

4. To settle this, it is necessary to measure the Sun's oblateness. This is tricky - the core of the Sun is not the same as its atmosphere, and results from previous experiments have not quite been conclusive. This experiment did not invalidate the Brans-Dicke theory of gravity - however, other experiments, such as atomic clocks flown on satellites, did.

§ 6 Scattering theory

§ 6.1 Scattering experiments

Consider the unbounded motion of a point mass m in a potential $U(\boldsymbol{x})$. Let

$$\lim_{r\to\infty} U(r) = 0 \quad , \quad r = |\boldsymbol{x}| \,. \tag{2.357}$$

Applying conservation of energy, we have

$$\begin{aligned} & E = \frac{m}{2}v^2 + U(\boldsymbol{x}) \\ \Longrightarrow \quad & \lim_{r\to\infty} v = \sqrt{\frac{2E}{m}} \equiv v_0 \\ \Longrightarrow \quad & \lim_{t\to\pm\infty} v(t) = v_0. \end{aligned} \tag{2.358}$$

It is convenient to label the system in the limit $t \to -\infty$ as the **initial state**, and in the limit $t \to \infty$ as the **final state**. Without loss of generality, we can choose our IS such that in the initial state, we have

$$\begin{aligned} \boldsymbol{v}_i &\equiv \lim_{t\to-\infty} v(t) \\ &\equiv (0, 0, v_0) \,, \end{aligned} \tag{2.359}$$

and

$$\begin{aligned} \boldsymbol{x}_i(t) &\equiv \lim_{t\to-\infty} \boldsymbol{x}(t) \\ &\equiv (b\cos\psi, b\sin\psi, v_0 t) \,, \end{aligned} \tag{2.360}$$

and in the final state, we have

$$\begin{aligned} \boldsymbol{v}_f &\equiv \lim_{t\to+\infty} v(t) \\ &\equiv v_0(\sin\chi\cos\alpha, \sin\chi\sin\alpha, \cos\chi) \,, \end{aligned} \tag{2.361}$$

and

$$\begin{aligned} \boldsymbol{x}_f(t) &\equiv \lim_{t\to+\infty} \boldsymbol{x}(t) \\ &\equiv \boldsymbol{x}_0 + \boldsymbol{v}_f t. \end{aligned} \tag{2.362}$$

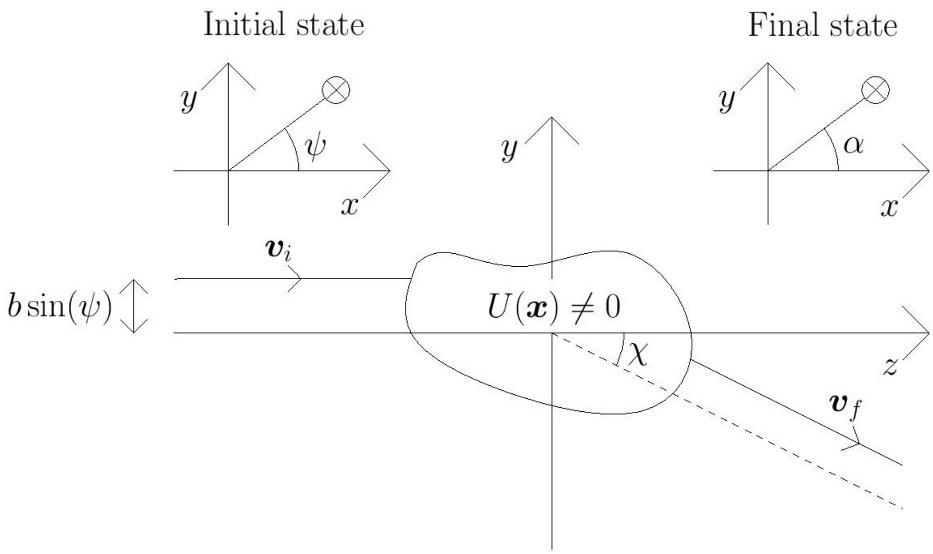

Figure 2.15: Setup for scattering theory.

Remark 1: The plane of motion in the final state is determined by α. The angular deviation of the direction of motion in the final state from that in the initial state is determined by χ.

Definition 1: The solid angle $\Omega = (\alpha, \chi)$ is called the **scattering angle**.

The idea of scattering theory is to obtain information about the potential $U(\boldsymbol{x})$ to which a particle with unbounded motion is exposed by observing that how the particle scatters (i.e., how its final state differs from its initial state) as a result of that exposure.

In scattering experiments, the **initial information** known regarding the particle is its **incident intensity**, I, defined as

$$I \equiv \frac{\text{number of particles}}{\text{unit time} \times \text{unit area}} \text{ traveling in the } z\text{−direction with velocity } v_0. \tag{2.363}$$

The **final measurement** made on the particle to gain information about the scattering potential is of the **scattering intensity**, S_Ω, defined as

$$S_\Omega \equiv \frac{\text{number of particles}}{\text{unit time}} \text{ in a given solid angle } \Omega, \tag{2.364}$$

where the given solid angle Ω can be described in terms of χ and α as

$$\Omega = \Omega\left(\chi_0 \leq \chi \leq \chi_1, \alpha_0 \leq \alpha \leq \alpha_1\right). \tag{2.365}$$

Definition 2: The observable

$$\Sigma_\Omega \equiv \frac{S_\Omega}{I} \tag{2.366}$$

is called the **scattering cross section**.

Remark 2: The scattering cross section Σ_Ω has units of area.

Remark 3: In Rutherford's scattering experiments, measurements made on scales of 10 cm revealed information about potentials on the scales of 10^{-13} cm.

Remark 4: High-energy experimentalists make a living doing scattering experiments.

§ 6.2 Classical theory for the scattering cross section

Assume that we have a homogeneous incident beam with a density n particles per unit volume. Then, the incident intensity of the beam is given by

$$I = nv_0. \tag{2.367}$$

Remark 1: The position of the incident beam with respect to the origin is characterized by the **impact parameter** b and the angle ψ.

Remark 2: At this point in our scattering theory, we have no information about the potential $U(\boldsymbol{x})$. To continue, we need axioms which ensure that nothing "weird" happens inside of the region where $U(\boldsymbol{x}) \neq 0$ (which is effectively a black box).

Axiom 1. There exists a piecewise, continuously differentiable function F that maps every incident beam (characterized by the impact parameter b and the angle ψ) onto a scattered beam (characterized by the solid angle Ω):

$$F : (b, \psi) \mapsto (\chi, \alpha) \equiv \Omega. \tag{2.368}$$

Remark 3: The function F maps from the xy-plane to the surface of a sphere.

Axiom 2. The function F is an n-to-1 mapping - that is, F maps at most $n < \infty$ incident beams to a single scattered beam.

From the above two axioms, we can immediately note that the scattering intensity S_Ω will be given by

$$\begin{aligned} S_\Omega &= nv_0 F^{-1}(\Omega) \\ &= nv_0 F^{-1}(\chi_0 \leq \chi \leq \chi_1, \alpha_0 \leq \alpha \leq \alpha_1) \\ &= nv_0 \sum_{l=1}^{n} \int db\, b_l \int d\psi \\ &= nv_0 \sum_{l=1}^{\infty} \int d\chi d\alpha \left| \frac{\partial (b_l, \psi_l)}{\partial (\chi, \alpha)} \right| b_l \\ &= I \sum_{l=1}^{n} \int d\Omega \frac{b_l}{\sin\chi} \left| \frac{\partial (b_l, \psi_l)}{\partial (\chi, \alpha)} \right|. \end{aligned} \tag{2.369}$$

In equation 2.369, the integrals in line 3 refer to the l'th incident beam contributing to the scattered beam with solid angle Ω (see Axiom 2). Line 4 follows from equation 2.367, and from the fact that

$$d\Omega = \sin\chi d\chi d\alpha. \tag{2.370}$$

Substituting into equation 2.366,

$$\Sigma_\Omega = \int d\Omega \sigma\left(\Omega\right), \tag{2.371}$$

where $\sigma\left(\Omega\right)$ is given by

$$\boxed{\begin{aligned} \sigma\left(\Omega\right) &\equiv \frac{d\Sigma_\Omega}{d\Omega} \\ &= \sum_{l=1}^{n} b_l\left(\chi,\alpha\right)\left|\frac{\partial\left(b_l,\psi_l\right)}{\partial\left(\chi,\alpha\right)}\right|\frac{1}{\sin\chi}. \end{aligned} \tag{2.372}}$$

Remark 4: The quantity $\sigma\left(\Omega\right)$ is called the **differential cross section**. It is effectively defined as

$$\sigma\left(\Omega\right) \equiv \frac{\text{number of scattered particles}/\left(\text{unit time}\times\text{unit solid angle}\right)}{\text{number of incident particles}/\left(\text{unit time}\times\text{unit area}\right)}. \tag{2.373}$$

Remark 5: The **total cross section**, σ, is defined as

$$\sigma \equiv \int_{4\pi} d\Omega\sigma\left(\Omega\right), \tag{2.374}$$

i.e.,

$$\sigma \equiv \frac{\text{number of scattered particles}/\left(\text{unit time}\right)}{\text{number of incident particles}/\left(\text{unit time}\times\text{unit area}\right)}. \tag{2.375}$$

§ 6.3 Scattering by a central potential

Recall from §4.1 that the orbit of a particle in a central potential lies in a plane. Then, for a central potential, we must have that

$$\alpha = \psi, \tag{2.376}$$

so that , substituting into equation 2.372, the differential cross section can immediately be simplified to

$$\boxed{\begin{aligned} \sigma\left(\Omega\right) &= \sigma\left(\chi\right) \\ &= \sum_{l=1}^{n} b_l\left(\chi\right)\left|\frac{db_l}{d\chi}\right|\frac{1}{\sin\chi}. \end{aligned} \tag{2.377}}$$

Additionally, recall from §4.5 that for unbounded motion in a central potential, the incoming and outgoing asymptotic orbits must be at equal and opposite angles with respect to the angle of the pericenter. Then, we have

$$\begin{aligned} \chi &= \pi - 2\left(\pi - \varphi_+\right) \\ &= 2\varphi_+ - \pi, \end{aligned} \tag{2.378}$$

i.e., χ is determined by the asymptotic orbit. We also have from §4.5 that the asymptotic orbit $\varphi_+$is a function of the energy E and angular momentum l of the system,

$$\varphi_+ = \varphi_+\left(E,l\right), \tag{2.379}$$

where

$$\begin{aligned} l &= \lim_{t\to-\infty} |\boldsymbol{x}\times m\boldsymbol{v}| \\ &= mv_0 \left| \begin{pmatrix} b\cos\psi \\ b\sin\psi \\ v_0 t \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right| \\ &= mv_0 \left| \begin{pmatrix} b\sin\psi \\ -b\cos\psi \\ 0 \end{pmatrix} \right| \\ &= mv_0 b \end{aligned} \tag{2.380}$$

is a constant. Then, φ_+ and χ can be written as functions of the impact parameter and the initial velocity,

$$\begin{aligned} & \varphi_+ = \varphi_+ (b, v_0) \\ \Longrightarrow \quad & \chi = \chi (b, v_0) \,. \end{aligned} \tag{2.381}$$

Inverting the second line of equation 2.381, we have

$$b = b(\chi)\,, \tag{2.382}$$

which can be substituted into equation 2.377 to find the differential cross section $\sigma(\Omega)$.

§ 6.4 Rutherford scattering

Let

$$U(r) = -\frac{\alpha}{r}. \tag{2.383}$$

This is a **Coulomb potential**, which is a conservative potential, so that far from the potential we have $\boldsymbol{v} = (0, 0, v_0)$, with

$$v_0 = \sqrt{\frac{2E}{m}}. \tag{2.384}$$

From §4.6 Remark 6 that

$$\begin{aligned} & 1 + e\cos\varphi_+ = 0 \\ \Longrightarrow \quad & -\frac{1}{e} = \cos\varphi_+. \end{aligned} \tag{2.385}$$

Substituting equation 2.378,

$$\begin{aligned} -\frac{1}{e} &= \cos\left(\frac{\chi}{2} + \frac{\pi}{2}\right) \\ &= -\sin\frac{\chi}{2}. \end{aligned} \tag{2.386}$$

Recalling the definition of the eccentricity from equation in §4.6,

$$e = \sqrt{1 + 2l^2 E/m\alpha^2}, \tag{2.387}$$

we have

$$\begin{aligned}-\frac{1}{e} &= -\frac{1}{\sqrt{1+2l^2E/m\alpha^2}} \\ &= -\frac{1}{\sqrt{1+2m^2v_0^2b^2E/m\alpha^2}} \\ &= -\frac{1}{\sqrt{1+4E^2b^2/\alpha^2}},\end{aligned} \tag{2.388}$$

where equation 2.380 was substituted between lines 1 and 2, and equation 2.384 was substituted between lines 2 and 3. Substituting into equation 2.386, multiplying by -1, squaring, and taking the reciprocal of each side gives

$$\begin{aligned}1+4\frac{E^2}{\alpha^2}b^2 &= \frac{1}{\sin^2(\chi/2)} \\ \implies \quad b &= \frac{\alpha}{2E}\sqrt{1-\frac{1}{\sin^2(\chi/2)}} = \frac{\alpha}{2E}\frac{\cos(\chi/2)}{\sin(\chi/2)},\end{aligned} \tag{2.389}$$

so that we can write,

$$b(\chi) = \frac{\alpha}{2E}\cot\frac{\chi}{2}. \tag{2.390}$$

Continuing, we have

$$\begin{aligned}b\frac{db}{d\chi} &= -\frac{1}{2}\left(\frac{\alpha}{2E}\right)^2\frac{\cos(\chi/2)}{\sin(\chi/2)}\frac{1}{\sin^2(\chi/2)} \\ &= -\left(\frac{\alpha}{2E}\right)^2\frac{\cos(\chi/2)}{2\sin^3(\chi/2)}.\end{aligned} \tag{2.391}$$

Substituting into equation 2.377, we have

$$\begin{aligned}\sigma(\chi) &= b\left|\frac{db}{d\chi}\right|\frac{1}{\sin\chi} \\ &= \left(\frac{\alpha}{2E}\right)^2\frac{\cos(\chi/2)}{2\sin^3(\chi/2)\,2\sin(\chi/2)\cos(\chi/2)} \\ &= \left(\frac{\alpha}{4E}\right)^2\frac{1}{\sin^4(\chi/2)}.\end{aligned} \tag{2.392}$$

Equation 2.392 is known as the **Rutherford formula**.

Remark 1*:* Note that $\sigma(\chi)$ does not depend on the sign of α (i.e., does not depend on whether the Coulomb potential is attractive or repulsive).

Remark 2*:* Forward scattering is strongly enhanced by the central potential. Back scattering is proportial to $1/E^2$.

Remark 3*:* We can also represent the result given by the Rutherford formula using a **polar diagram** (See Figure 2.16 below), with

$$r = \sigma(\chi) \quad , \quad \varphi = \chi \quad , \quad \left(\frac{\alpha}{4E}\right)^2 = 1. \tag{2.393}$$

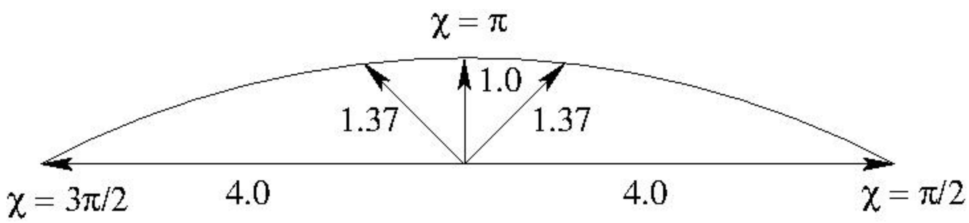

Figure 2.16: Rutherford scattering - $\sigma(\chi)$ polar diagram.

Remark 4*:* The total cross section comes to

$$\begin{aligned} \sigma &= \int d\Omega \sigma(\Omega) \\ &= \infty, \end{aligned} \tag{2.394}$$

due to singular forward scattering. This is a result of the long (infinite) range of the Coulomb potential.

Remark 5*:* Most incident particles go through the Coulomb potential with minimal deflection.

§ 6.5 Scattering by a hard sphere

The **hard sphere** of radius a can be defined as the potential

$$U(r) = \begin{cases} 0, & r \geq a \\ \infty, & r < a \end{cases}. \tag{2.395}$$

Letting φ denote the angle between the incident beam and a line normal to the hard sphere that is in the same plane as the incident beam and the line between the sphere's center and the incident beam, we have

$$b = a \sin\varphi \quad , \quad \chi = \pi - 2\varphi, \tag{2.396}$$

since the incident and scattered beam will make equal angles with the normal, as in Figure 2.17. Then, we have

$$\begin{aligned} & b = -a \sin\left(\frac{\chi}{2} - \frac{\pi}{2}\right) = a \cos\frac{\chi}{2} \\ \Longrightarrow \quad & b\frac{db}{d\chi} = -a^2 \frac{1}{2} \cos\frac{\chi}{2} \sin\frac{\chi}{2} = -\frac{a^2}{4} \sin\chi. \end{aligned} \tag{2.397}$$

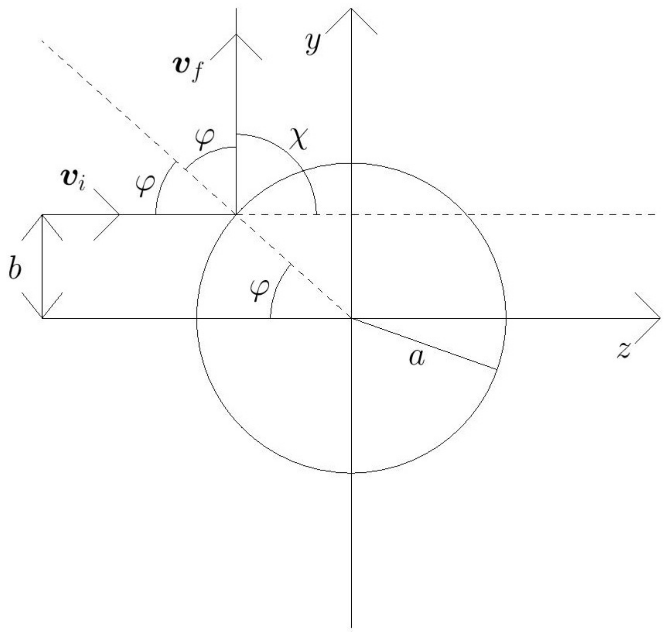

Figure 2.17: Scattering by a hard sphere.

Substituting into equation 2.377 gives the differential cross section as

$$\sigma(\chi) = \frac{a^2}{4}, \tag{2.398}$$

and substituting this result into equation 2.374 gives the total cross section as

$$\sigma = \pi a^2. \tag{2.399}$$

Remark 1*:* Hard-sphere scattering is isotropic.

Remark 2*:* The total cross section σ can be interpreted as the effective scattering area "seen" by the particle that is being scattered. For the hard sphere, σ is equal to the geometric area of the potential. For the Coulomb potential, $\sigma = \infty$ due to the long range of the potential.

§ 7 N-particle systems

§ 7.1 Closed N-particle systems

Consider a **closed system** of N particles with masses m_α, with $\alpha = 1, \ldots, N$, and let

$$q \equiv (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N) \quad , \quad \dot{q} \equiv (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N) \quad , \quad p \equiv (\boldsymbol{p}_1, \ldots, \boldsymbol{p}_N) \tag{2.400}$$

represent the coordinates, velocities, and momenta, respectively, of the particles in the system.

Remark 1: Recall from Ch.2§1.1 that $\boldsymbol{p}_\alpha = m_\alpha \boldsymbol{v}_\alpha$.

Remark 2: A system is **closed** if and only if it has no interaction with the "outside world".

Axiom 1. The Lagrangian of a closed N-particle system is

$$L(q, \dot{q}) = \sum_{\alpha=1}^{N} L_\alpha (\boldsymbol{v}_\alpha) - U(q), \tag{2.401}$$

with L_α the Lagrangian of the α'th free particle (see Ch.2§1.1).

Axiom 2.

$$U(q) \equiv U(r_{12}, \ldots, r_{1N}, r_{23}, \ldots, r_{2N}, \ldots, r_{N-1,N}), \tag{2.402}$$

where

$$r_{\alpha\beta} \equiv |\boldsymbol{x}_\alpha - \boldsymbol{x}_\beta|. \tag{2.403}$$

Axiom 2′ (More specific version of Axiom 2)**.**

$$\begin{aligned} U(q) &\equiv \sum_{\alpha<\beta} U_{\alpha\beta}(r_{\alpha\beta}) \\ &= \frac{1}{2}\sum_{\alpha\neq\beta} U_{\alpha\beta}(r_{\alpha\beta}), \end{aligned} \tag{2.404}$$

where

$$U_{\alpha\beta}(r_{\alpha\beta}) = U_{\beta\alpha}(r_{\beta\alpha}). \tag{2.405}$$

Remark 3: In Axiom 2, U depends only on the $N(N-1)/2$ relative distances of the particles $r_{\alpha\beta}$. Axiom 2′ specifies the form of this dependence.

Remark 4: Axiom 2 gives (using equation 2.400),

$$\begin{aligned} & q \to q + q_0 = (\boldsymbol{x}_1 + \boldsymbol{x}_0, \ldots, \boldsymbol{x}_N + \boldsymbol{x}_0) \\ \Longrightarrow\ & r_{\alpha\beta} = |\boldsymbol{x}_\alpha - \boldsymbol{x}_\beta| \to |\boldsymbol{x}_\alpha + \boldsymbol{x}_0 - \boldsymbol{x}_\beta - \boldsymbol{x}_0| = r_{\alpha\beta} \\ \Longrightarrow\ & U(q + q_0) = U(q), \end{aligned} \tag{2.406}$$

i.e., $U(q)$ is translationally invariant. As a result, by Axiom 1, the Lagrangian is translationally invariant as well.

Remark 5: Axiom 2′ postulates that $U(q)$ is a sum of pair potentials, making direct use of the **principle of superposition**.

Remark 6: Axiom 2′ is a special case of Axiom 2. That is, Axiom 2′ implies Axiom 2, but Axiom 2 does not imply Axiom 2′.

Remark 7: The validity of Axiom 2′ is a matter for experiment to decide, as discussed in Ch.1§1.2.

Example 1: For N particles interacting via gravity, we have

$$U_{\alpha\beta}\left(r_{\alpha\beta}\right) = -G\frac{m_\alpha m_\beta}{r_{\alpha\beta}}. \tag{2.407}$$

Experiments show that Axiom 2′ holds for such potentials.

Example 2: For N charged particles with charges e_α,

$$U_{\alpha\beta}\left(r_{\alpha\beta}\right) = \frac{e_\alpha e_\beta}{r_{\alpha\beta}}. \tag{2.408}$$

Once more, experiments show that Axiom 2′ holds for such potentials.

Consequences of the axioms:

a) **Conservation of momentum**
Under a translation by a in the i_0-direction, the i'th compentent of $\boldsymbol{x}_\alpha$ transforms as

$$x^i_\alpha \to x^i_\alpha = x^i_\alpha + a\delta_{ii_0}. \tag{2.409}$$

From Nöther's theorem (Ch.1§4.4), we have a conserved quantity,

$$\begin{aligned} j\left(q,\dot{q}\right) &= \sum_{\alpha,i} p^i_\alpha \delta_{ii_0} \\ &= \sum_{\alpha} p^{i_0}_\alpha \\ &= p^{i_0} \\ &= \text{constant}, \end{aligned} \tag{2.410}$$

where p^{i_0} is the total momentum in the i_0-direction. But the since the direction of translation i_0 that was chosen was arbitrary, the above equation holds for any direction, so that

$$\begin{aligned} \boldsymbol{p} &= \sum_\alpha \boldsymbol{p}_\alpha \\ &= \text{constant}, \end{aligned} \tag{2.411}$$

i.e., the total momentum is conserved.

b) **Conservation of energy**
From Axiom 1, the N-particle system is autonomous (see Ch.1§4.1). Then,

$$\begin{aligned} H\left(q,\dot{q}\right) &= \sum_{\alpha,i} \dot{q}^i_\alpha p^i_\alpha - L\left(q,\dot{q}\right) \\ &= \sum_{\alpha,i} m_\alpha \dot{q}^i_\alpha \dot{q}^i_\alpha - L\left(q,\dot{q}\right) \\ &= \sum_{\alpha,i} \frac{p^i_\alpha p^i_\alpha}{m_\alpha} - L\left(q,\dot{q}\right) \\ &= E, \end{aligned} \tag{2.412}$$

a constant. Again applying Axiom 1, and substituting each free particle Lagrangian, we have that

$$E = \sum_\alpha \frac{\boldsymbol{p}_\alpha^2}{2m_\alpha} + U\left(q\right), \tag{2.413}$$

the total kinetic plus potential energy, is conserved.

c) **Conservation of angular momentum**
In cylindrical coordinates, the α'th position is given by

$$\left(r_\alpha \cos\varphi_\alpha, r_\alpha \sin\varphi_\alpha, z_\alpha\right), \tag{2.414}$$

and the α'th free-particle Lagrangian is given by

$$L_\alpha = \frac{m_\alpha}{2}\left(\dot{r}_\alpha^2 + r_\alpha^2\dot{\varphi}_\alpha^2 + \dot{z}_\alpha^2\right). \tag{2.415}$$

From equations 2.403 and 2.414,

$$\begin{aligned} r_{\alpha\beta}^2 &= \left(z_\alpha - z_\beta\right)^2 + \left(r_\alpha\cos\varphi_\alpha - r_\beta\cos\varphi_\beta\right)^2 + \left(r_\alpha\sin\varphi_\alpha - r_\beta\sin\varphi_\beta\right)^2 \\ &= \left(z_\alpha - z_\beta\right)^2 + r_\alpha^2 + r_\beta^2 - 2r_\alpha r_\beta\cos\left(\varphi_\alpha - \varphi_\beta\right). \end{aligned} \tag{2.416}$$

Under a rotation about the z-axis by angle φ_0, the position of each particle transforms as,

$$\varphi_\alpha \to \varphi_\alpha + \varphi_0, \tag{2.417}$$

so that

$$\begin{aligned} & \varphi_\alpha - \varphi_\beta \to \varphi_\alpha + \varphi_0 - \varphi_\beta - \varphi_0 = \varphi_\alpha - \varphi_\beta \\ \Longrightarrow\quad & r_{\alpha\beta}^2 \to r_{\alpha\beta}^2, \end{aligned} \tag{2.418}$$

where the second line follows from equation 2.416. From Axiom 2, $U\left(q\right)$ is unchanged, and we can see from equation 2.415 that each L_α is unchanged ($\dot{\varphi}_\alpha$ is unchanged under the rotation since clearly $\dot{\varphi}_0 = 0$). Then, from Axiom 1, the Lagrangian for the N-particle system is unchanged under rotations. Applying Nöther's theorem, the quantity

$$\begin{aligned} j\left(q, \dot{q}\right) &= \sum_\alpha \frac{\partial L}{\partial \dot{\varphi}_\alpha} \\ &= \sum_\alpha m_\alpha r_\alpha^2 \dot{\varphi}_\alpha \\ &= \sum_\alpha \left(\boldsymbol{x}_\alpha \times \boldsymbol{p}_\alpha\right)_z \\ &= \sum_\alpha l_{\alpha,z} \\ &= l_z \end{aligned} \tag{2.419}$$

is conserved. The choice of the z-axis as the axis of rotation was arbitrary - choosing the x-axis or y-axis, respectively, and following identical logic would have led to the conclusion that l_x and l_y are conserved as well. Then, we have that

$$\boldsymbol{L} = \sum_\alpha \boldsymbol{l}_\alpha, \tag{2.420}$$

the total angular momentum, is conserved.

d) **The law of inertia**
From equation 2.411,

$$\begin{aligned} \boldsymbol{p} &= \sum_\alpha \boldsymbol{p}_\alpha \\ &= \sum_\alpha m_\alpha \boldsymbol{v}_\alpha \\ &= \frac{d}{dt}\sum_\alpha m_\alpha \boldsymbol{x}_\alpha \\ &= \text{constant.} \end{aligned} \tag{2.421}$$

Defining the **total mass** M and **center of mass** $\boldsymbol{X}$ as

$$M \equiv \sum_\alpha m_\alpha \quad , \quad \boldsymbol{X} \equiv \frac{1}{M}\sum_\alpha m_\alpha \boldsymbol{x}_\alpha, \tag{2.422}$$

and applying equation 2.421 gives

$$\begin{aligned} & M\frac{d}{dt}\boldsymbol{X} = \text{constant} \\ \Longrightarrow \quad & \boldsymbol{X}(t) = \boldsymbol{X}_0 + \frac{1}{M}\boldsymbol{p}t, \end{aligned} \tag{2.423}$$

which is the **law of inertia**.

Remark 8*:* The center of mass $\boldsymbol{X}$ moves with constant velocity $\boldsymbol{p}/M$ along a straight line.

Remark 9*:* The law of inertia is also called **Newton's first law**, and it holds for the center-of-mass motion, no matter how complicated U may be. It is a corollary of conservation of momentum, and so follows from the axioms indirectly.

e) **The equations of motion**
Starting from the Euler-Lagrange equations,

$$\begin{aligned} \frac{d}{dt}p^i_\alpha &= \frac{\partial L}{\partial x^i_\alpha} \\ &= -\frac{\partial U}{\partial x^i_\alpha} \\ &= -\frac{1}{2}\sum_{\gamma\neq\delta}\frac{\partial U_{\gamma\delta}}{\partial r_{\gamma\delta}}\frac{\partial}{\partial x^i_\alpha}r_{\gamma\delta}, \end{aligned} \tag{2.424}$$

where lines 2 and 3 follow directly from the axioms. From equation 2.403,

$$\begin{aligned} \frac{\partial}{\partial x^i_\alpha}r_{\gamma\delta} &= \frac{\partial}{\partial x^i_\alpha}\sqrt{(\boldsymbol{x}_\gamma - \boldsymbol{x}_\delta)^2} \\ &= \frac{1}{r_{\alpha\beta}}\left(x^i_\gamma - x^i_\delta\right)\left(\delta_{\alpha\gamma} - \delta_{\alpha\delta}\right). \end{aligned} \tag{2.425}$$

Writing each component of the **force** on particle α as

$$\begin{aligned} F^i_\alpha &= \sum_{\beta\neq\alpha} F^i_{\alpha\beta} \\ &= -\sum_{\beta\neq\alpha} \frac{\partial U_{\alpha\beta}}{\partial r_{\alpha\beta}} \frac{1}{r_{\alpha\beta}} \left(x^i_\alpha - x^i_\beta\right), \end{aligned} \tag{2.426}$$

where α is not varied in the above summations, and substituting gives **Newton's second law**,

$$\frac{d}{dt}\boldsymbol{p}_\alpha = \boldsymbol{F}_\alpha. \tag{2.427}$$

Remark 10*:* $\boldsymbol{F}_{\alpha\beta}$ is the force exerted on particle α by particle β.

Remark 11*:* Noting that $\boldsymbol{F}_{\beta\alpha} = -\boldsymbol{F}_{\alpha\beta}$, we have **Newton's third law**.

Remark 12*:* Consider that

$$\begin{aligned} \sum_\alpha \boldsymbol{F}_\alpha &= \sum_\alpha \sum_{\beta\neq\alpha} \boldsymbol{F}_{\alpha\beta} \\ &= \sum_{\alpha,\beta} \left(1-\delta_{\alpha\beta}\right) \boldsymbol{F}_{\alpha\beta} \\ &= -\sum_{\alpha,\beta} \left(1-\delta_{\alpha\beta}\right) \boldsymbol{F}_{\beta\alpha} \\ &= -\sum_{\alpha,\beta} \left(1-\delta_{\alpha\beta}\right) \boldsymbol{F}_{\alpha\beta} \\ &= -\sum_\alpha \boldsymbol{F}_\alpha \\ &= 0, \end{aligned} \tag{2.428}$$

where the last line follows from comparing lines 1 and 5. Considering this and Newton's second law (equation 2.427), we have that the total force $\boldsymbol{F}$ is equal to 0 if and only if the total momentum is conserved,

$$\boldsymbol{F} = 0 \quad \Longleftrightarrow \quad \frac{d}{dt}\boldsymbol{p} = 0. \tag{2.429}$$

§ 7.2 The two-body problem

Consider the closed N-particle system for which $N = 2$. From equation 2.422, **center-of-mass coordinates** for a such a system can be defined as

$$M\boldsymbol{X} \equiv m_1\boldsymbol{x}_1 + m_2\boldsymbol{x}_2 \tag{2.430}$$

and **relative coordinates** as

$$\boldsymbol{x} \equiv \boldsymbol{x}_1 - \boldsymbol{x}_2. \tag{2.431}$$

Similarly, the **center-of-mass velocity** can be defined as

$$\begin{aligned} \boldsymbol{V} &\equiv \frac{d}{dt}\boldsymbol{X} \\ &= \frac{m_1}{M}\boldsymbol{v}_1 + \frac{m_2}{M}\boldsymbol{v}_2, \end{aligned} \tag{2.432}$$

and the **relative velocity** as

$$\begin{aligned} \boldsymbol{v} &\equiv \frac{d}{dt}\boldsymbol{x} \\ &= \boldsymbol{v}_1 - \boldsymbol{v}_2 \end{aligned} \tag{2.433}$$

Solving equations 2.430 and 2.431 for $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ gives

$$\boldsymbol{x}_1 = \boldsymbol{X} + \frac{m_2}{M}\boldsymbol{x} \quad , \quad \boldsymbol{x}_2 = \boldsymbol{X} - \frac{m_1}{M}\boldsymbol{x}. \tag{2.434}$$

From §7.1, the Lagrangian for the system is

$$\begin{aligned} L &= \frac{m_1}{2}\boldsymbol{v}_1^2 + \frac{m_2}{2}\boldsymbol{v}_2^2 - U\left(|\boldsymbol{x}_1 - \boldsymbol{x}_2|\right) \\ &= \frac{m_1}{2}\left(\boldsymbol{V} + \frac{m_2}{M}\boldsymbol{v}\right)^2 + \frac{m_2}{2}\left(\boldsymbol{V} - \frac{m_1}{M}\boldsymbol{v}\right)^2 - U\left(|\boldsymbol{x}|\right) \\ &= \frac{M}{2}\boldsymbol{V}^2 + \frac{m_1 m_2}{2M}\boldsymbol{v}^2 - U\left(|\boldsymbol{x}|\right) \\ &= L_{\mathrm{CM}} + L_r, \end{aligned} \tag{2.435}$$

We have effectively defined

$$L_{\mathrm{CM}} \equiv \frac{M}{2}\boldsymbol{V}^2, \tag{2.436}$$

$$\begin{aligned} L_r &\equiv \frac{m_1 m_2}{2M}\boldsymbol{v}^2 - U\left(|\boldsymbol{x}|\right) \\ &= \frac{m}{2}\boldsymbol{v}^2 - U\left(r\right), \end{aligned} \tag{2.437}$$

where

$$m = \frac{m_1 m_2}{m_1 + m_2} \tag{2.438}$$

is the **reduced mass**, and

$$\begin{aligned} r &= |\boldsymbol{x}| \\ &= |\boldsymbol{x}_1 - \boldsymbol{x}_2|\,. \end{aligned} \tag{2.439}$$

Remark 1*:* From equation 2.435, the center-of-mass motion and relative motion are decoupled.

Remark 2*:* The center-of-mass motion is free-particle motion, see §7.1 (d).

Remark 3*:* In equation 2.437, the problem has been reduced to the motion of a single particle in a central field.

Remark 4*:* We solved such a problem in §§4, 6.

Remark 5*:* This justifies applying the results of §§4, 6 to celestial mechanics.

Remark 6*:* For $N > 2$, the Lagrangian could still be decoupled into a center-of-mass-motion Lagrangian and a relative-motion Lagrangian. However, the problem would still be hard to solve, since U in the relative-motion Lagrangian would still depend on $N-1$ relative distances.

Example 1 (Binary stars)**:** Recall Kepler's third law from §4.4,

$$T^2 = 4\pi^2 \frac{m}{\alpha} a^3 \quad , \quad \alpha = G m_1 m_2. \tag{2.440}$$

Then,

$$\left(m_1 + m_2\right) T^2 = \frac{4\pi^2}{G} a^3. \tag{2.441}$$

Remark 7*:* Measurement of the period T and semi-major axis a yields $m_1 + m_2$.

Remark 8*:* This is the *only* direct way to measure stellar masses.

Remark 9*:* If the center of mass can be determined, then we could still determine the ratio of the masses, and so determine m_1 and m_2 separately using our knowledge of $m_1 + m_2$.

Example 2 (Easy experiment)**:** Consider the measurement of the motion of a light component m_2 relative to a heavy coordinate m_1. This yields an ellipse with the heavy component as a focus. Measuring T and a, gives

$$m_1 + m_2 = \frac{4\pi^2}{G} \frac{a^3}{T^2}. \tag{2.442}$$

Example 3 (Hard experiment)**:** Consider the measurement of both components in Example 2 relative to a fixed background. Observing two ellipses with the center of mass as a focus and with semi-major axes a_1 and a_2, we have

$$m_1 + m_2 = \frac{4\pi^2}{G} \frac{\left(a_1 + a_2\right)^3}{T^2}, \tag{2.443}$$

where, by the definition of the center of mass (equation 2.422),

$$\frac{a_1}{a_2} = \frac{m_2}{m_1}. \tag{2.444}$$

Then,

$$m_1 = \frac{4\pi^2}{G} \frac{a_2 \left(a_1 + a_2\right)^2}{T^2} \quad , \quad m_2 = \frac{4\pi^2}{G} \frac{a_1 \left(a_1 + a_2\right)^2}{T^2}. \tag{2.445}$$

Remark 10*:* Suppose that one of our components, m_2, cannot be observed directly (e.g., because it is too dim, or because it is a black hole, etc.). Then, we can still measure a_1 relative to a background (this is a hard experiment), giving

$$\begin{aligned}
& \frac{a_1}{a} = \frac{m}{m_1} = \frac{m_2}{m_1 + m_2} \\
\implies \quad & (m_1 + m_2) \frac{m_2^3}{(m_1 + m_2)^3} = \frac{4\pi^2}{G} \frac{a_1^3}{T^2} \\
\implies \quad & \frac{m_2^3}{(m_1 + m_2)^2} = \frac{4\pi^2}{G} \frac{a_1^3}{T^2},
\end{aligned} \tag{2.446}$$

so that the combination $m_2^3 / (m_1 + m_2)^2$ can still be measured.

Remark 11*:* If the orbit is at an unknown angle α relative to the line of sight, then only $a \sin \alpha$ can be measured. The methods of astronomy can be used to deal with this and other complications.

§ 8 Problems for Chapter 2

17. Free particles

a) Determine the energy of a free particle according to

   i) Galilean mechanics.

   ii) Einsteinian mechanics.

b) Discuss and draw the momentum and the energy of a free Einsteinian particle. What follows for the role that c plays in the theory?

18. Escape velocity

The escape velocity v_e is the minimum velocity (or rather, speed) an object must have at launch in order to fly infinitely far from the earth. Assume that the earth and your rocket are the only objects in the universe. Determine the flight path which yields the smallest v_e, and determine that minimum value of v_e.

19. Gauge transformations

a) Show that the forces $\boldsymbol{F}^{(1)}$ and $\boldsymbol{F}^{(2)}$ in §1.4 are separately gauge-invariant.

b) Show that a scalar potential $U(\boldsymbol{x}, t)$ can always be "gauged away", i.e., that there always exists a gauge transformation such that the transformed U is identically equal to zero.

c) Show that a vector potential $\boldsymbol{V}(\boldsymbol{x}, t)$ can be gauged away if it is the gradient of a scalar function, i.e., if a function $\chi : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}$ exists such that $\boldsymbol{V}(\boldsymbol{x}, t) = \nabla \chi(\boldsymbol{x}, t)$.

20. Pendulums

Find the Lagrangian for a planar pendulum (mass m, length l) whose suspension point

a) has mass M and is free to move without friction along a horizontal axis.

b) moves uniformly with angular velocity γ along a vertical circle with radius r. Use a gauge transformation to write the Lagrangian as

$$L = \frac{m}{2} l^2 \dot{\varphi}^2 + mrl\gamma^2 \sin(\varphi - \gamma t) + mgl\cos\varphi. \tag{2.447}$$

c) oscillates horizontally according to $x_p(t) = a\cos\gamma t$. Use a gauge transformation to write the Lagrangian as

$$L = \frac{m}{2} l^2 \dot{\varphi}^2 + mal\gamma^2 \cos\gamma t \sin\varphi + mgl\cos\varphi. \tag{2.448}$$

21. Einstein's law of falling bodies

Consider a particle in a homogeneous gravitational potential, and determine its motion according to special relativity.

a) Use the fact that two of the three coordinates are cyclic variables to show that the orbit lies in a plane.
Hint: Take the cyclic coordinates to be x and y and show that $x(t)\,p_y - y(t)\,p_x = \text{constant}$.

b) Determine and discuss the velocity in the xz-plane, i.e., $v_x(t)$ and $v_z(t)$. Show in particular that for $c \to \infty$ (nonrelativistic limit) the Galilean result is recovered, and discuss what happens within Einsteinian mechanics for $t \to \infty$ (ultrarelativistic limit).

c) Determine and discuss the path, i.e., $x(t)$ and $z(t)$, by integrating the result obtained in part (b). Discuss again the nonrelativistic and ultrarelativistic limits.

d) Eliminate time to obtain the orbit, and discuss the nonrelativistic and ultrarelativistic limits.

22. Larmor's theorem

Show that the following statement holds within Galilean mechanics:

A charged particle (charge e, mass m) moves in a static homogeneous magnetic field B and in a scalar potential $U(\boldsymbol{x})$ that is rotationally invariant about the field direction. Consider a coordinate system that rotates about the field direction with an angular velocity $\omega_L = \omega_c/2 = eB/2mc$. In the rotating coordinate system the particle's motion is the same as that of a particle subject to a potential $\tilde{U}(\boldsymbol{x}) = U(\boldsymbol{x}) + m\omega_c^2 r^2/8$ and no magnetic field, with r the distance from the axis of rotation.
Hint: Use cylindrical coordinates.

23. Particle on a moving inclined plane

A point mass m moves under the influence of gravity (g in a negative z-direction) on an inclined plane (inclination angle α) with mass M. The inclined plane can move without friction along the x-axis.

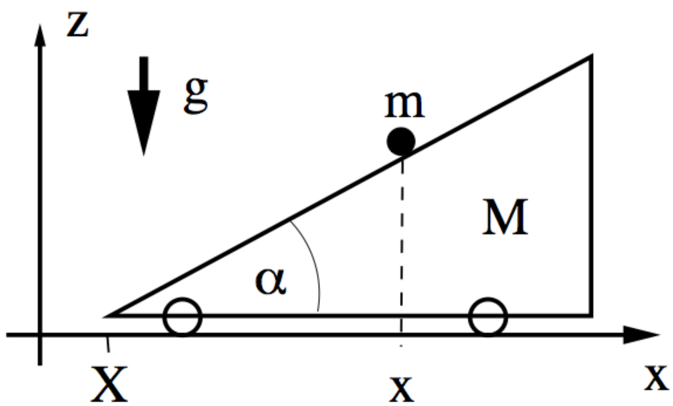

Figure 2.18

a) Solve the equations of motion for the initial conditions

$$\dot{X}(t=0) = \dot{x}(t=0) = 0. \tag{2.449}$$

b) Now let $\dot{y}(t=0) = 0$. What is the orbit of the point mass? Discuss the solution of the equations of motion in the limits $M \to \infty$ and $M \to 0$, respectively.

c) Discuss the energy E_m of the point mass and the energy E_M of the inclined plane as functions of time.

24. Lissajous figures

Consider a two-dimensional oscillator with potential

$$U(x_1, x_2) = \frac{m}{2} \sum_{i=1}^{2} \omega_i^2 x_i^2. \tag{2.450}$$

Let $\omega_2/\omega_1 = \Omega$.

a) Solve the equations of motion.

b) Show that for $\Omega = 1$ the orbit is an ellipse.

c) Show that the result of part (a) can be written, without loss of generality,

$$x_1(t) = \sin(t) \quad , \quad x_2(t) = a_2 \sin(\Omega t + \varphi_2). \tag{2.451}$$

d) Draw the orbits for

i) $\Omega = 1 \quad , \quad a_2 = 1/2 \quad , \quad \varphi_2 = \pi/2,$

ii) $\Omega = 2 \quad , \quad a_2 = 1 \quad , \quad \varphi_2 = \pi/4,$

iii) $\Omega = 2/3 \quad , \quad a_2 = 1/2 \quad , \quad \varphi_2 = \pi/3,$

iv) $\Omega = 3/\pi \quad , \quad a_2 = 1 \quad , \quad \varphi_2 = 0.$

e) What do these results suggest for the nature of the orbit depending on what number set Ω belongs to?

25. Effective potential

Consider the rotationally invariant problem from §4.1. The use of rotational symmetry yields

$$\theta \equiv \pi/2 \quad , \quad \dot{\varphi} = l/mr^2. \tag{2.452}$$

One might get the idea of using these results in the Lagrangian and taking things from there. Show that this procedure yields a result that is different from what is obtained in §4. Which is correct, and why?

26. Particle on a hyperboloid

A point mass m moves under the influence of gravity (g in the negative z-direction) on a hyperboloid of revolution,

$$\frac{r^2}{a^2} - \frac{z^2}{c^2} = 1 \quad , \quad \left(r^2 = x^2 + y^2\right). \tag{2.453}$$

Discuss the possible types of motion depending on the values of the energy E and the parameter

$$\Delta = \left(\frac{l^2/ma^2}{mgc}\right)^{1/2}. \tag{2.454}$$

Give physical interpretations of your results.

27. $1/r^2$ potential

Determine, discuss, and sketch all possible types of motion for a point mass in a potential

$$U(r) = -\alpha/r^2 \quad , \quad (\alpha > 0). \tag{2.455}$$

In the case of bounded motion, find the oscillation period as a function of the energy E and the z-component of the angular momentum l.

28. Relativistic Kepler problem

Consider the motion of a point mass m in an attractive Coulomb potential within the framework of special relativity,

$$L = mc^2 - mc^2\sqrt{1-(\boldsymbol{v}/c)^2} + \alpha/r \quad , \quad (\alpha > 0)\,. \tag{2.456}$$

Let the z-component of the angular momentuntum be given by $l \geq 0$.

a) Use the conservation laws for the energy and the angular momentum to find the radial equation of motion,

$$\dot{r}^2 = c^2 f\,(r; E, l) \tag{2.457}$$

and the equation for the sectorial velocity,

$$r^2\dot{\varphi} = g\,(r; E, l)\,. \tag{2.458}$$

Explicitly determined the functions f and g and show that in the nonrelativistic limit they correctly reduce to the Galilean case.

b) Assume that $l = 0$, and let $\dot{r}\,(t=0) = 0$, $r\,(t=0) = 2a$. Discuss and draw $\dot{r}$ as a function of r, and compare with the Galilean case. Determine the oscillation period $T\,(a)$ in terms of a dimensionless integral that depends only on the parameter $\xi = a/2amc^2$. Discuss your result in the nonrelativistic and ultrarelativistic limits ($\xi \to 0$ and $\xi \to \infty$, respectively). Show that Kepler's third law gets modified within special relativity, and plot $T/T_{\xi=0}$ as a function of ξ.

c) Now assume that $l > 0$. Use $f\,(r; E, l)$ from part (a) to discuss all possible types of motion. Show that there are qualitative changes compared to Galilean mechanics; in particular, that the particle can reach the center provided $\zeta \equiv a^2/l^2c^2 > 1$. Determine the perihelion and the apohelion and show that in the Keplerian result is recovered in the nonrelativistic limit $\zeta \to 0$. Find the condition for the existence of an allowed region and again show that the Keplerian result is recovered for $\zeta \to 0$.
Hint: Write f in the form

$$f\,(r; E, l) = \frac{2x + x^2}{(1+x)^2} - \zeta\frac{(p/r)^2}{(1+x)^2} \tag{2.459}$$

with $p = l^2/m\alpha$ the Keplerian parameter and $x = (E + \alpha/r)/mc^2$ a local energy parameter. Discuss x as a function of r and the two contributions to f as functions of x to obtain a qualitative picture of f as a function of r. Then solve the condition $f\,(r_\pm; E, l) = 0$ to obtain the turning points.

d) Assume that $\zeta < 1$ and determine the equation of orbit. Show that the result can be written in the form

$$p^*/r = 1 + e^*\cos\left(\sqrt{1-\zeta}\varphi\right). \tag{2.460}$$

Determine p^* and e^*.

e) Consider the motion of an electron in the field of a Th nucleus ($\alpha = Ze^2$, $Z = 90$). Draw the orbit on a scale of $1 : 10^{-10}$ for the four cases $l = \hbar, 2\hbar$, $E = \pm 12keV$. (Use $\hbar = 10^{-27}$ erg, $\hbar c/e = 137$.) Also draw the corresponding Galilean orbits.

29. Pendulum

Consider a mathematical pendulum (mass m, length l) with Lagrangian

$$L = \frac{m}{2} l^2 \dot{\varphi}^2 + mgl \cos \varphi. \tag{2.461}$$

a) Show that the system can display bounded or unbounded motion, depending on the value of the energy. Show that the motion is always bounded if $E < U_0$. Determine U_0 and the turning points $\pm\varphi_0$ of the oscillation.

b) For the case of bounded motion, find the oscillation frequency as a function of E. Discuss your result as a function of $\epsilon = E/2U_0 + 1/2$, especially the limiting cases $\epsilon \to 0$ and $\epsilon \to 1$.
Hint: The result can be expressed in terms of the complete elliptic integral of the first kind, $K(k)$, defined as

$$K(k) = \int_0^{\pi/2} dx \frac{1}{\sqrt{1 - k^2 \sin x}}. \tag{2.462}$$

c) Show that $\dot{\varphi}(t)$ can be expressed in closed form in terms of the Jacobi elliptic function cn:

$$\dot{\varphi}(t) = 2\omega_0 \sin(\varphi_0/2) \operatorname{cn}(\omega_0 t, k), \tag{2.463}$$

where $\omega_0 = \sqrt{U_0/ml^2}$.
Hint: The elliptic integral of the first kind $F(x,k)$ is defined as

$$F(x,k) = \int_0^x dz \frac{1}{\sqrt{1 - k^2 \sin z}} \quad , \quad (k \leq 1). \tag{2.464}$$

Its inverse is known as the "amplitude",

$$x = F(y,k) \quad \Longrightarrow \quad y = \operatorname{am}(x,k), \tag{2.465}$$

and cn is defined as

$$\operatorname{cn}(x,k) = \cos \operatorname{am}(x,k). \tag{2.466}$$

Also useful are

$$\operatorname{sn}(x,k) = \sin \operatorname{am}(x,k), \tag{2.467}$$

and

$$\operatorname{dn}(x,k) = \frac{d}{dx} \operatorname{am}(x,k) = \sqrt{1 - k^2 \sin^2 y}. \tag{2.468}$$

d) Define the nonlinearity parameter

$$\alpha = \frac{E}{\omega(E)} \frac{d\omega}{dE} \tag{2.469}$$

as a measure of how "nonlinear" the oscillation is, i.e., how strongly ω depends on E. Determine α to leading order in the energy parameter ϵ for small ϵ. The result shows that for small ϵ the oscillation is almost linear.

e) From the exact solution for $\dot{\varphi}(t)$ found in part (c), find $\varphi(t)$ in terms of a Fourier series. Show that for small ϵ the oscillation is almost harmonic, i.e., only the first term in the Fourier series contributes appreciably.
Hint: Look up the Fourier expansion for $\operatorname{cn}(x,k)$ and integrate term-by-term.

30. Anharmonic oscillator

Consider the anharmonic oscillator from §5.4:

$$L = \frac{m}{2}\dot{q}^2 - \frac{m}{2}\omega_0^2 q^2 + m\sum_{n=3}^{\infty} \frac{\alpha_n}{n} q^n. \tag{2.470}$$

a) Calculate $q(t)$ to $O\left(a^3\right)$, where $a/2$ is the first Fourier coefficient.

b) Show that there are no corrections of $O\left(a^3\right)$ (or any odd order) to the frequency.

31. Perihelion advance of Mercury

Use the results of Problem 28 to determine the perihelion advance of Mercury to leading order in $1/c^2$. Estimate the accuracy of this approximation. Compare the answer with the experimental result of $43''$/century.
Hint: The semi-major axis, period, and eccentricity of Mercury's Galilean orbit are:
$\alpha = 5.79 \times 10^{12}$ cm, $T = 88$ days, and $e = 0.206$, respectively.

32. Scattering by a potential well

Consider a potential (in 3-d)

$$U(r) = U\Theta(a - r) \tag{2.471}$$

with Θ the step function.

a) Consider the case $U < 0$. Show that the maximum scattering angle is given by

$$\chi_{\max} = 2\cos^{-1}(1/n), \tag{2.472}$$

with $n = \sqrt{1 - U/E}$, where E is the energy of the scattered particle. Determine and discuss the differential scattering cross section as a function of χ and E. Sketch $\sigma(\chi)$ for $E \to 0$ and for $E \to \infty$, and sketch a polar diagram for either case.

b) Consider the case $U > 0$. Show that the maximum scattering angle is now given by

$$\chi_{\max} = 2\cos^{-1} n, \tag{2.473}$$

and that $b(\chi)$ is a double-valued function. Sketch $\sigma(\chi)$ for $E \to U$ and for $E \to \infty$, and sketch a polar diagram for either case. What is the qualitative difference between the cases $U < 0$ and $U > 0$?

33. Attraction by a $1/r^2$ potential

Find the total cross section σ for a particle to reach the center of a potential $U(r) = -\alpha/r^2$ $(\alpha > 0)$. Sketch σ as a function of the energy E of the scattered particle.
Hint: Use the results of Problem 27.

34. Scattering by a central field in two dimensions

Consider a point mass moving in the xy-plane in a central potential

$$U(x,y) = U(r) \quad , \quad \left(r^2 = x^2 + y^2\right). \tag{2.474}$$

a) Define the scattering cross section in analogy to the three-dimensional case. Show that the differential cross section can be written as

$$\sigma(\chi) = \sum_{l=1}^{n} \left|db_l/d\chi\right|. \tag{2.475}$$

What is the dimension of σ?

b) Calculate and discuss the differential scattering cross section for the case of a hard disc:

$$U(r) = \begin{cases} \infty & \text{for r} < \text{a} \\ 0 & \text{otherwise} \end{cases}. \tag{2.476}$$

Compare your results with the corresponding ones for scattering by a hard sphere in three dimensions.

c) Now consider a point mass in three dimensions subject to a cylindrically symmetric potential of the form

$$U(x,y,z) = U(r) \quad , \quad \left(r^2 = x^2 + y^2\right). \tag{2.477}$$

Show that the equations of motion can be solved in analogy to the case of a spherically symmetric potential. What is the differential cross section for scattering by a hard cylinder? (Assume that the incident beam is perpendicular to the cylinder axis).

35. Small-angle scattering

a) Show that for weak potentials U, i.e., for small scattering angles χ, the scattering angle to leading order in the small potential can be written as

$$\chi = -\frac{2b}{E}\int_b^\infty dr \frac{dU/dr}{\sqrt{r^2 - b^2}}. \tag{2.478}$$

Hint: Use the same tricks as in §5.5 and assume that $U(r)$ is continuously differentiable, at least for large r. See also the pertinent homework problem in Landau & Lifshitz.

b) Find the differential cross section $\sigma(\chi)$ for small χ for a power-law potential $U(r) = \alpha/r^n$. What is the total cross section?

c) Express $\sigma(\chi \to 0)$ for a screened Coulomb potential,

$$U(r) = \frac{\alpha}{r} e^{-r/a}, \tag{2.479}$$

in terms of two integrals. Show that for $a \gg b$ one recovers the result for a Coulomb potential. Determine $\sigma(\chi \to 0)$ for $a \ll b$. What is the total cross section in this case?

Hint: To leading exponential accuracy (i.e., neglecting power-law prefactors of the exponential), one has, for $\mu \gg 1$,

$$\int_1^\infty dx \frac{e^{-\mu x}}{x^n \sqrt{x^2-1}} \approx e^{-\mu} \int_1^\infty dx \frac{1}{x^n \sqrt{x^2-1}} \quad , \quad (n > 0)\,. \tag{2.480}$$

Note: One can do better by making an ansatz

$$\int_1^\infty dx\, \frac{e^{-\mu x}}{x^n \sqrt{x^2-1}} = e^{-\mu} f\,(\mu)\,, \tag{2.481}$$

and determining $f\,(\mu)$ term-by-term as a power series for large μ, but this is not necessary for our purposes.

Chapter 3

The rigid body

§ 1 Mathematical preliminaries - vector spaces and tensor spaces

§ 1.1 Metric vector spaces

Definition 1: Let $V = \{x, y, \ldots\}$ be a set, and let $+$ be an operation on V such that

a) V is an abelian group under $+$.

Furthermore, define a multiplication of elements of V with the real numbers[1] $\lambda, \mu, \ldots \in \mathbb{R}$ such that $\forall x, y \in V$ and $\forall \lambda, \mu \in \mathbb{R}$,

b1) $\lambda x \in V$,

b2) $\lambda (\mu x) = (\lambda \mu) x$,

b3) $(\lambda + \mu) x = \lambda x + \mu x$, and

b4) $\lambda (x + y) = \lambda x + \lambda y$.

Then, we call V a **vector space** or **linear space** over $\mathbb{R}$.

Remark 1: The elements $x, y, \ldots \in V$ are called **vectors**, and the elements $\lambda, \mu, \ldots \in \mathbb{R}$ are called **scalars**.

Example 1: Let $x = \left(x^1, x^2, \ldots, x^n\right) \in \mathbb{R}^n$. Define $x + y \equiv \left(x^1 + y^1, \ldots, x^n + y^n\right)$, and define $\lambda x = \left(\lambda x^1, \ldots, \lambda x^n\right)$. Then, $\mathbb{R}^n$ is a vector space over $\mathbb{R}$.

Let V be n-dimensional and let $\{e_i | i = 1, \ldots, n\}$ be a basis in V, i.e., a s set of linearly independent vectors that span the space. Then, every $x \in V$ can be written as

$$x = \sum_{i=1}^{n} x^i e_i \quad \text{with } x^1, \ldots, x^n \in \mathbb{R}. \tag{3.1}$$

[1] This definition could equally well apply with an arbitrary set S in place of $\mathbb{R}$. In such a case, V would be called a vector space over S.

Remark 2: The numbers x^i are called the **coordinates** of x in the basis $\{e_i\}$.

Definition 2: Let $g_{ij} = g_{ji}$ be the coordinates of a symmetric $n \times n$ matrix with $\det g \neq 0$. Define a mapping $V \times V \to \mathbb{R}$, denoted by $\cdot$, as

$$x \cdot y = \sum_{i,j=1}^{n} x^i g_{ij} y^j. \tag{3.2}$$

Then, $\cdot$ is called a **scalar product**, g is called a **metric**, and V is called a **metric vector space**.

Remark 3: In $\mathbb{R}^n$, $e_i = \left(\delta_0{}^i, \ldots, \delta_n{}^i\right)$ form a basis. The coordinates of each e_i are $(e_i)^l = \delta_i{}^l$, so that

$$\begin{aligned} e_i \cdot e_j &= \sum_{k,l} (e_i)^k \, g_{kl} \, (e_j)^l \\ &= \sum_{k,l} \delta_i{}^k g_{kl} \delta_j{}^l \\ &= g_{ij}. \end{aligned} \tag{3.3}$$

Remark 4: The function

$$\begin{aligned} \delta_i{}^j &= \delta_{ij} \\ &= \delta^i{}_j \\ &= \delta^{ij} \\ &= \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \end{aligned} \tag{3.4}$$

is called the **Kronecker delta**.

Definition 3: The **inverse metric** is denoted by g^{-1}, and its components are given by

$$\left(g^{-1}\right)_{ij} \equiv g^{ij}. \tag{3.5}$$

We have that

$$\begin{aligned} & gg^{-1} = g^{-1}g = 1 \\ \Longrightarrow \quad & \sum_{k=1}^{n} g_{ik} g^{kj} = \sum_{k=1}^{n} g^{ik} g_{kj} = \delta_i{}^j. \end{aligned} \tag{3.6}$$

Definition 4: An **adjoint basis** $\left\{e^i\right\}$ is given by

$$e^i \equiv \sum_{j=1}^{n} g^{ij} e_j. \tag{3.7}$$

Remark 5*:* The inverse of this relation is

$$\begin{aligned} e_i &= \sum_{j=1}^{n} \delta_i{}^j e_j \\ &= \sum_{j,k=1}^{n} g_{ik} g^{kj} e_j \\ &= \sum_k g_{ik} e^k. \end{aligned} \tag{3.8}$$

Remark 6*:* We have

$$\begin{aligned} x &= \sum_{i=1}^{n} x^i e_i \\ &= \sum_{i=1}^{n} x^i g_{ij} e^j \\ &= \sum_{i=1}^{n} x_j e^j \end{aligned} \tag{3.9}$$

with

$$\begin{aligned} x_j &\equiv \sum_i x^i g_{ij} \\ &= \sum_i g_{ji} x^i. \end{aligned} \tag{3.10}$$

Remark 7*:* The coordinates x^i are called **contravariant coordinates**. They are the coordinates in the basis $\{e_i\}$. Similarly, the coordinates x_i are called the **covariant coordinates**. They are the coordinates in the adjoint basis $\left\{e^i\right\}$.

Remark 8*:* We can define a summation convention

$$\begin{aligned} x \cdot y &= \sum_{ij} x^i g_{ij} x^j \\ &= \sum_i x^i x_i \\ &= \sum_j x_j x^j \\ &\equiv x_i x^i. \end{aligned} \tag{3.11}$$

Remark 9*:* More genearlly, a summation over covariant and contravariant indices is implied when they appear together as in equation 3.11. This is a summation convention. More specifically, it is the Einstein summation convention.

§ 1.2 Coordinate transformations

Consider an $n \times n$ matrix with components[2] $D^i{}_j$, and its inverse with components $\left(D^{-1}\right)^i{}_j$. We have

$$DD^{-1} = D^{-1}D = 1 \quad \Longleftrightarrow \quad D^i{}_j \left(D^{-1}\right)^j{}_k = \left(D^{-1}\right)^i{}_j D^j{}_k = \delta^i{}_k. \tag{3.12}$$

Definition 1: Define a new basis $\{\tilde{e}_i\}$ through the **basis tranformation**

$$\tilde{e}_i = e_j \left(D^{-1}\right)^j{}_i. \tag{3.13}$$

Remark 1*:* The inverse basis transformation is given by D, as expected:

$$\begin{aligned} \tilde{e}_i D^i{}_j &= e_k \left(D^{-1}\right)^k{}_i D^i{}_j \\ &= e_k \delta^k{}_j \\ &= e_j. \end{aligned} \tag{3.14}$$

Proposition 1: In performing a basis transformation or its inverse, the coordinates $\tilde{x}^i$ or x^i, respectively, are given by a **coordinate transformation**,

$$\tilde{x}^i = D^i{}_j x^j \quad , \quad x^i = \left(D^{-1}\right)^i{}_j \tilde{x}^j. \tag{3.15}$$

Proof: Using equations 3.9 and 3.14,

$$\begin{aligned} x &= x^i e_i \\ &= x^i \tilde{e}_j D^j{}_i \\ &= \left(D^j{}_i x^i\right) \tilde{e}_j \\ &\equiv \tilde{x}^j \tilde{e}_j, \end{aligned} \tag{3.16}$$

and similarly for the inverse transformation. □

Proposition 2: In performing a basis transformation or its inverse, the **transformation of the metric** proceeds as

$$\tilde{g} = \left(D^{-1}\right)^{\mathrm{T}} g D^{-1} \quad , \quad g = D^{\mathrm{T}} \tilde{g} D. \tag{3.17}$$

Proof: From equations 3.3 and 3.13,

[2] When dealing with the components of a matrix, the leftmost index indicates the row, and the rightmost index indicates the column. By convention, we denote the components of a matrix M as $M^i{}_j$, i.e., $M^i{}_j$ is the component in the i'th row and j'th column of M. By contrast, $M_j{}^i$ represents the component in the j'th row and i'th column of M - that is, it represents the components of M^{T}, so that $\left(M^{\mathrm{T}}\right)^i{}_j = M_j{}^i$. Index raising and lowering, on the other hand, can be viewed as occuring through interaction with the metric, and thus represents something else entirely.

$$\begin{aligned}
\tilde{g}_{ij} &= \tilde{e}_i \cdot \tilde{e}_j \\
&= e_k \left(D^{-1}\right)^k{}_i \cdot e_l \left(D^{-1}\right)^l{}_j \\
&= \left(D^{-1}\right)^k{}_i \, e_k \cdot e_l \left(D^{-1}\right)^l{}_j \\
&= \left(\left(D^{-1}\right)^{\mathrm{T}}\right)_i{}^k g_{kl} \left(D^{-1}\right)^l{}_j \\
&= \left(\left(D^{-1}\right)^{\mathrm{T}} g D^{-1}\right)_{ij} .
\end{aligned} \tag{3.18}$$

□

§ 1.3 Proper coordinate systems

Lemma 1: Theres exists a coordinate transformation D such that (no summation implied)

$$\tilde{g}_{ij} = \lambda_i \delta_{ij} \quad \text{with } \lambda_i \neq 0 \quad \forall i. \tag{3.19}$$

Proof: The lemma follows directly from the fact that g is a real symmetric matrix with $\det g \neq 0$ (§1.1 Definition 2). □

Remark 1*:* The scalars λ_i are the eigenvalues of g (and of $\tilde{g}$).

Theorem 1: There exists a coordinate transformation D such that

$$\tilde{g} = \begin{array}{cl} \begin{pmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ & & & -1 & & \\ & 0 & & & \ddots & \\ & & & & & -1 \end{pmatrix} & \begin{array}{l} \text{row 1} \\ \vdots \\ \text{row m} \\ \text{row m} + 1 \\ \vdots \\ \text{row n} \end{array} \end{array} , \tag{3.20}$$

i.e., such that $\tilde{g}$ is (m-times-(+1) and $(m-n)$-times-(-1))-diagonal.

Proof: From Lemma 1, we have that $\tilde{g}_{ij} = \lambda_i \delta_{ij}$, with $\lambda_i \neq 0$. Choosing a basis such that

$$\lambda_1, \ldots, \lambda_m > 0 \quad , \quad \lambda_{m+1}, \ldots \lambda_n < 0, \tag{3.21}$$

and defining

$$\left(D^{-1}\right)^i{}_j = \frac{\delta^i{}_j}{|\lambda_i|^{1/2}}, \tag{3.22}$$

we have, from equation 3.17,

$$\begin{aligned}
\tilde{\tilde{g}}_{ij} &= \left(D^{-1}\right)^k{}_i \, \tilde{g}_{kl} \left(D^{-1}\right)^l{}_j \\
&= \delta^k{}_i \frac{1}{|\lambda_i|^{1/2}} \delta_{kl} \lambda_k \delta^l{}_j \frac{1}{|\lambda_j|^{1/2}} \\
&= \delta_{ij} \frac{\lambda_i}{|\lambda|_i} .
\end{aligned} \tag{3.23}$$

□

Definition 1: Coordinate systems in which g has the form of equation 3.20 are called **proper**.

Remark 2: The number m in Theorem 1 is characteristic of the vector space. This is referred to as **Sylvester's rigidity theorem**.

Example 1: In the case that $m = n$, we have

$$g_{ij} = \delta_{ij}. \tag{3.24}$$

In such a case, the vector space is called **Euclidean**, and the proper coordinate systems are called **Cartesian**. For such a system, we have

$$\begin{aligned} x_i &= g_{ij}x^j \\ &= \delta_{ij}x^j \\ &= x^i, \end{aligned} \tag{3.25}$$

i.e., the covariant components and contravariant components are equal. Furthermore, the scalar product for such a system is

$$x \cdot x = x_1^2 + \ldots + x_n^2, \tag{3.26}$$

so that the **Pythagorean theorem** applies.

Example 2: For $m = 1$, $n \geq 2$, we have

$$g = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & -1 & \ddots & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{pmatrix}. \tag{3.27}$$

Such a vector space is called a **Minkowski space**. The proper coordinate systems in such a space are called **inertial frames**. Additionally, we have

$$x_0 = x^0 \quad , \quad x_i = -x^i \quad , \quad (i = 1, \ldots, n-1)\,, \tag{3.28}$$

and

$$\begin{aligned} x \cdot x &= x_i x^i \\ &= x_0^2 - \sum_{i=1}^{n-1} x_i^2. \end{aligned} \tag{3.29}$$

§ 1.4 Proper coordinate transformations

Definition 1: A coordinate transformation is called **proper** if it transforms proper coordinate systems into proper coordinate systems.

Remark 1*:* If D^{-1} is a proper transformation, then $\tilde{g} = \left(D^{-1}\right)^{\mathrm{T}} g D^{-1} = g$, so that

$$g_{ij} = \left(D^{-1}\right)^{k}{}_{i}\, g_{kl} \left(D^{-1}\right)^{l}{}_{j}. \tag{3.30}$$

Example 1: In Euclidean space, **orthogonal transformations**, i.e., those satisfying $\left(D^{-1}\right)^{\mathrm{T}} D^{-1} = 1$, are proper.

Example 2: In Minkowski space, **Lorentz transformations** are proper.

Lemma 1:

a) If D is a proper transformation, then so is D^{-1}.

b) If D_1 and D_2 are proper transformations, then so is $D_1 D_2$.

Proof:

a) D is proper. Then,

$$\begin{aligned} & g = D^{\mathrm{T}} g D \\ \Longrightarrow \quad & \left(D^{\mathrm{T}}\right)^{-1} g D^{-1} = \left(D^{-1}\right)^{\mathrm{T}} D^{\mathrm{T}} g D D^{-1} = g. \end{aligned} \tag{3.31}$$

b)

$$\begin{aligned} g &= D_2^{\mathrm{T}} g D_2 \\ &= D_2^{\mathrm{T}} D_1^{\mathrm{T}} g D_1 D_2 \\ &= \left(D_1 D_2\right)^{\mathrm{T}} g \left(D_1 D_2\right). \end{aligned} \tag{3.32}$$

□

Theorem 1: The set of proper coordinate transformations forms a (non-abelian, in general) group under matrix multiplications.

Proof: From parts (a) and (b) of Lemma 1, respectively, the inverse of each element of the set is in the set, and the group satisfies closure. The identity matrix 1 is clearly a proper transformation, and so is in the set. Associativity is satisfied since matrix multiplication is associative. □

Proposition 1: For proper transformations D, $\det D = \pm 1$.

Proof:

$$\begin{aligned} & \det g = \det D^{\mathrm{T}} \det g \det D = (\det D)^2 \det g \\ \Longrightarrow \quad & (\det D)^2 = 1 \\ \Longrightarrow \quad & \det D = \pm 1. \end{aligned} \tag{3.33}$$

□

§ 1.5 Tensor spaces

Let the system under consideration, or one of its properties, be characterized by k numbers $t_1, \ldots, t_k$, collectively referred to as t in a coordinate system cs. Let it be characterized by by k numbers $\tilde{t}_1, \ldots, \tilde{t}_k$, collectively referred to as $\tilde{t}$, in a coordinate system $\widetilde{\text{cs}}$.

Question. Is there a relation between the numbers t_i and the numbers $\tilde{t}_i$?

Such a question leads to the idea of classifying mathematical objects by their properties under coordinate transformations.

Example 1: Let t be the vector norm squared, given by

$$\begin{aligned} t &= x \cdot x \\ &= x_i x^i. \end{aligned} \tag{3.34}$$

Under a transformation, we have

$$\begin{aligned} \tilde{t} &= \tilde{x}_i \tilde{x}^i \\ &= \tilde{x}^j \tilde{g}_{ji} \tilde{x}^i \\ &= D^i{}_k x^k \tilde{g}_{ji} D^i{}_l x^l \\ &= D^j{}_k \tilde{g}_{ji} D^i{}_l x^k x^l \\ &= g_{kl} x^k x^l \\ &= x_l x^l \\ &= t. \end{aligned} \tag{3.35}$$

We have therefore arrived a property - that t is unchanged under coordinate transformations. Such objects are classified as **scalars**.

Example 2: Let t be a vector, with components $t^i = x^i$. Then,

$$\begin{aligned} \tilde{t}^i &= \tilde{x}^i \\ &= D^i{}_j x^j \\ &= D^i{}_j t^j. \end{aligned} \tag{3.36}$$

We have arrived at another property - that t transforms under coordinate transformations through multiplication by the transformation. Such objects are classified as **vectors**.

Example 3: Let t be an object with two indices, $t^{ij} = x^i y^j$, and n^2 components. Then,

$$\begin{aligned} \tilde{t}^{ij} &= \tilde{x}^i \tilde{y}^j \\ &= D^i{}_k D^j{}_l x^k y^l \\ &= D^i{}_k D^j{}_l t^{kl}. \end{aligned} \tag{3.37}$$

We have arrived at another property - that t transforms under coordinate transformations D through multiplication by the direct product of the transformation with itself, $D \otimes D$. Such objects are called **dyadic tensors**, or **tensors of rank** 2.

Definition 1: Let t be a set of n^N (with n the dimension of the vector space) numbers in a proper coordinate system cs, writeable as $t = t^{i_1,\ldots,i_N}$. Let $\tilde{t}$ be a set of n^N numbers is a proper coordinate system $\widetilde{cs}$, writeable as $\tilde{t} = \tilde{t}^{i_1,\ldots,i_N}$. Let D be a proper coordinate transformation. Then, we call t a **tensor of rank** N if

$$\tilde{t}^{i_1,\cdots,i_N} = D^{i_1}{}_{j_1} \ldots D^{i_N}{}_{j_N} t^{j_1,\ldots,j_N},$$

and we call t a **pseudotensor of rank** N if

$$\tilde{t}^{i_1,\cdots,i_N} = (\det D)\, D^{i_1}{}_{j_1} \ldots D^{i_N}{}_{j_N} t^{j_1,\ldots,j_N},$$

where $\det D = \pm 1$ (see §1.4 Proposition 1).

Remark 1: Note that

$$t^{i_1 \ldots i_{j-1}}{}_{i_j}{}^{i_{j+1}\ldots i_N} = g_{i_j k} t^{i_1 \ldots i_{j-1} k i_{j+1} \ldots i_N}. \tag{3.38}$$

Remark 2: Scalars are tensors of rank 0. Vectors are tensors of rank 1.

Remark 3: Not everything is a tensor. For instance, the first component of a vector x^1, is a single number. So, if it is a tensor, it is a scalar. But $\tilde{x}^1 = D^i{}_j x^j \neq x^1$. So, x^1 is not a scalar, and therefore not a tensor.

Remark 4: The **tensor product**, or **outer product**, given by

$$t = x_{(1)} \otimes x_{(2)} \otimes \ldots \otimes x_{(N)}, \tag{3.39}$$

is defined as

$$t^{i_1 \cdots i_N} = x^{i_1}_{(1)} x^{i_2}_{(1)} \ldots x^{i_N}_{(N)}. \tag{3.40}$$

Then, t is a tensor of rank N.

Lemma 1: Let s and t be (pseudo)tensors of rank N, and let $\lambda, \mu \in \mathbb{R}$. Then, $u = \lambda s + \mu t$, defined by

$$u^{i_1 \ldots i_N} \equiv \lambda s^{i_1 \ldots i_N} + \mu t^{i_1 \ldots i_N} \tag{3.41}$$

is a (pseudo)tensor of rank N.

Proof: Problem 36. □

Proposition 1: The set of all (pseudo)tensors of rank N forms a vector space over $\mathbb{R}$ of dimension n^N.

Proof: Problem 36. □

Lemma 2: Let s and t be tensors of rank N and rank M, respectively. Define $u \equiv s \otimes t$ by

$$u^{i_1 \ldots i_{N+M}} \equiv s^{i_1 \ldots i_N} t^{i_{N+1} \ldots i_{N+M}}. \tag{3.42}$$

Then, u is a tensor of rank $N + M$.

Proof: Problem 36. □

Corollary 1: Let s and t be tensors, and let u and v be pseudotensors. Then, $s \otimes t$ is a tensor, $s \otimes u$ is a pseudotensor, $u \otimes s$ is a pseudotensor, and $u \otimes v$ is a tensor.

Proof: Problem 36. □

Lemma 3: Let $t^{ijk_1 \ldots k_N}$ be a tensor of rank $N+2$. Define the $(1,2)$**-trace** of t, $u = t^{(1,2)}$, by

$$\begin{aligned} u^{k_1 \ldots k_N} &\equiv g_{ij} t^{ijk_1 \ldots k_N} \\ &= t_i{}^{ik_1 \ldots k_N} . \end{aligned} \tag{3.43}$$

Then, u is a tensor of rank N.

Proof:

$$\begin{aligned} \tilde{u}^{k_1 \ldots k_N} &= \tilde{g}_{ij} \tilde{t}^{ijk_1 \ldots k_N} \\ &= \underbrace{\tilde{g}_{ij} D^i{}_m D^j{}_n}_{g_{mn}} D^{k_1}{}_{l_1} \ldots D^{k_N}{}_{l_N} t^{mnl_1 \ldots l_N} \\ &= D^{k_1}{}_{l_1} \ldots D^{k_N}{}_{l_N} g_{mn} t^{mnl_1 \ldots l_N} \\ &= D^{k_1}{}_{l_1} \ldots D^{k_N}{}_{l_N} u^{l_1 \ldots l_N} . \end{aligned} \tag{3.44}$$

□

Remark 5*:* Other traces can be defined analogously to the $(1,2)$-trace.

Remark 6*:* Traces are also called **contractions**.

Lemma 4: The object ε, defined by

$$\varepsilon_{i_1 \ldots i_n} = \begin{cases} +1, & (i_1, \ldots i_n) \text{ is an even permutation of } (1, \ldots, n) \,, \\ -1, & (i_1, \ldots i_n) \text{ is an odd permutation of } (1, \ldots, n) \,, \\ 0, & \text{otherwise,} \end{cases} \tag{3.45}$$

is a pseudotensor of rank n.

Proof: Problem 36. □

Remark 7*:* The object ε is called the **Levi-Civita tensor**, or the **completely antisymmetric tensor of rank** n.

Remark 8*:* Let $n = 3$, and let x and y be Euclidean vectors. Define z by

$$z_i \equiv \varepsilon_{ijk} x^j y^k , \tag{3.46}$$

so that

$$\begin{aligned} z_1 &= x^2 y^3 - x^3 y^2 \\ &= (x \times y)_1 \,, \end{aligned} \tag{3.47}$$

and similarly for cylclic permutations. Then, the cross product of two vectors is not a regular vector, but a pseudovector.

§ 2 Orthogonal transformations

§ 2.1 Three-dimensional Euclidean space

Consider three-dimensional Euclidean space, i.e., $\mathbb{R}^3$. From §1, the Euclidean metric is given by

$$\begin{aligned} g_{ij} &= g^{ij} \\ &= \delta_{ij} \\ &= e_i \cdot e_j, \end{aligned} \tag{3.48}$$

with each e_i a basis vector in a **Cartesian coordinate system**, satisfying

$$\begin{aligned} (e_i)^j &= (e_i)_j \\ &= \delta_{ij}. \end{aligned} \tag{3.49}$$

In such a space, proper coordinate transformations D are orthogonal, obeying

$$\left(D^{-1}\right)^k{}_i \left(D^{-1}\right)_{kj} = \delta_{ij}. \tag{3.50}$$

Remark 1*:* From §1.4, the set of orthogonal transformations forms a group under matrix multiplication, called $O(3)$.

Remark 2*:* The **reflection** (or **inversion**) element $R \in O(3)$ is defined by

$$R_{ij} = -\delta_{ij}. \tag{3.51}$$

Note that this implies that $\det R = -1$ and that $R^{-1} = R$. Then, upon multiplication by R, vectors x transform as

$$\begin{aligned} \tilde{x}^i &= R^i{}_j x^j \\ &= -x^i, \end{aligned} \tag{3.52}$$

and pseudovectors y transform as

$$\begin{aligned} \tilde{y}^i &= -R^i{}_j y^j \\ &= y^i. \end{aligned} \tag{3.53}$$

Remark 3*:* Orthogonal matrices a with components a_{ij} have the property that

$$\sum_{j=1}^{3} a_{ij}^2 = 1. \tag{3.54}$$

This can be seen as follows:

$$\begin{aligned} & aa^{\mathrm{T}} = 1 \\ \Longrightarrow\quad & 1_{ii} = \sum_{j=1}^{3} a_{ij} a_{ji}^{\mathrm{T}} = \sum_{j=1}^{3} a_{ij}^2, \end{aligned} \tag{3.55}$$

where each diagonal component, 1_{ii}, of the identity matrix is 1.

Proposition 1: Let $t^{ij} = -t^{ji}$ be the components of an antisymmetric rank 2 tensor parameterized by three numbers t_1, t_2, and t_3, so that t is writeable as

$$t = \begin{pmatrix} 0 & t_3 & -t_2 \\ -t_3 & 0 & t_1 \\ t_2 & -t_1 & 0 \end{pmatrix}. \tag{3.56}$$

Then,

$$\boldsymbol{t} \equiv (t_1, t_2, t_3) \tag{3.57}$$

is a pseudovector.

Proof: Consider the pseudovector[3] $\boldsymbol{u}$ with components

$$u_i \equiv \varepsilon_{ijk} t^{jk}. \tag{3.58}$$

Writing out each component explicitly,

$$u_1 = t^{23} - t^{32} = 2t^{23} = 2t_1, \tag{3.59}$$

$$u_2 = -t^{13} + t^{31} = -2t^{13} = 2t_2, \tag{3.60}$$

and

$$u_3 = t^{12} - t^{21} = 2t^{12} = 2t_3. \tag{3.61}$$

Then,

$$\boldsymbol{t} = \frac{1}{2}\boldsymbol{u}, \tag{3.62}$$

and therefore $\boldsymbol{t}$, like $\boldsymbol{u}$, is a pseudovector. □

Corollary 1: The set of antisymmetric rank 2 tensors under addition is isomorphic to the set of pseudovectors under addition.

Proof: Let $\varphi : \boldsymbol{t} \mapsto t$ take pseudovectors as written in equation 3.57 to antisymmetric rank 2 tensors as written in equation 3.58,

$$\begin{aligned} \varphi[\boldsymbol{t}] &= \varphi[(t_1, t_2, t_3)]. \\ &= \begin{pmatrix} 0 & t_3 & -t_2 \\ -t_3 & 0 & t_1 \\ t_2 & -t_1 & 0 \end{pmatrix} \\ &= t. \end{aligned} \tag{3.63}$$

[3]From §1.5 Corollary 1, the product of a tensor with a pseudotensor is a pseudotensor. Then, since $\boldsymbol{u}$ is the product of the Levi-Civita tensor and the tensor t, and since the Levi-Civita tensor is a pseudotensor, $\boldsymbol{u}$ is a pseudotensor.

Then,

$$\begin{aligned}
\varphi(\boldsymbol{t}+\boldsymbol{u}) &= \varphi[(t_1,t_2,t_3)+(u_1,u_2,u_3)] \\
&= \varphi[(t_1+u_1,t_2+u_2,t_3+u3)] \\
&= \begin{pmatrix} 0 & t_3+u_3 & -(t_2+u_2) \\ -(t_3+u_3) & 0 & t_1+u_1 \\ t_2+u_2 & -(t_1+u_1) & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & t_3 & -t_2 \\ -t_3 & 0 & t_1 \\ t_2 & -t_1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & u_3 & -u_2 \\ -u_3 & 0 & u_1 \\ u_2 & -u_1 & 0 \end{pmatrix} \\
&= t+u \\
&= \varphi(\boldsymbol{t})+\varphi(\boldsymbol{u}),
\end{aligned} \tag{3.64}$$

so that φ is a homomorphism. If $\boldsymbol{t} \neq \boldsymbol{u}$, then one or more of $t_i \neq u_i$, with $i \in \{1,2,3\}$. Then, one or more of the components of t is unequal to the corresponding component of u, so that $t \neq u$, implying that φ is an injection. For any given antisymmetric tensor t parameterized by t_1, t_2, and t_3, there is a corresponding pseudovector $\boldsymbol{t} = (t_1,t_2,t_3)$, so that $\varphi(\boldsymbol{t}) = t$, so that φ is a surjection. Then, φ is a bijective homomorphism, i.e., an isomorphism, so that

$$(\{t\},+) \cong (\{\boldsymbol{t}\},+). \tag{3.65}$$

□

§ 2.2 Special orthogonal transformations - the group $SO(3)$

Let the sets $O_\pm$ be defined as

$$O_\pm \equiv \{D \in O(3) \,|\, \det D = \pm 1\}. \tag{3.66}$$

Remark 1: $O_+ \cup O_- = O(3)$, and $O_+ \cap O_- = \emptyset$.

Remark 2: $R \in O_-$ and $1 \in O_+$, so that $O_+ \neq \emptyset \neq O_-$, with R the reflection element of $O(3)$.

Lemma 1: Let $S \in O_-$. Then,

$$\exists! T \in O_+ S = RT. \tag{3.67}$$

Proof: Let $S \in O_-$, and let

$$T \equiv RS. \tag{3.68}$$

Then, since $O(3)$ is a group under matrix multiplication and $R, S \in O(3)$, $T \in O(3)$. Furthermore, since $R, S \in O_-$, $\det S = \det R = -1$, so that

$$\begin{aligned}
\det T &= \det R \det S \\
&= 1.
\end{aligned} \tag{3.69}$$

Then, $T \in O_+$. Recalling that $R = R^{-1}$, we have

$$\begin{aligned}
RT &= R^2 S \\
&= S.
\end{aligned} \tag{3.70}$$

Therefore, $\exists T \in O_+ S = RT$. Suppose also that $\exists U \in O_+ S = RU$. Then, we have $S = RT$ and $S = RU$. Setting these equal to each other and again recalling that $R = R^{-1}$,

$$\begin{aligned} & RT = RU \\ \implies \quad & R^2 T = R^2 U \\ \implies \quad & T = U, \end{aligned} \tag{3.71}$$

so that T is unique. We can therefore write

$$\exists! T \in O_+ S = RT. \tag{3.72}$$

□

Proposition 1: $SO\,(3) \equiv O_+$ is a subgroup of $O\,(3)$.

Proof: O_+ was defined as a subset of $O\,(3)$, so $SO\,(3) \subseteq O\,(3)$. Then, because $O\,(3)$ is a group, associativity must be satisfied in $O\,(3)$, so that it is satisfied $SO\,(3)$ as well. Let $D_1, D_2 \in SO\,(3)$. Then, from the definition of O_+, i.e., the definition of $SO\,(3)$,

$$\begin{aligned} & \det D_1 = \det D_2 = 1 \\ \implies \quad & \det\left(D_1 D_2\right) = \det D_1 \det D_2 = 1, \end{aligned} \tag{3.73}$$

so that $D_1 D_2 \in SO\,(3)$, i.e., $SO\,(3)$ satisfies closure. Since $\det 1 = 1$ and $1 \in O\,(3)$, $1 \in SO\,(3)$, so that $SO\,(3)$ contains the identity. Finally, recalling that

$$\det D = 1 \implies \det D^{-1} = 1, \tag{3.74}$$

and that, since $O\,(3)$ is a group,

$$D \in O\,(3) \implies D^{-1} \in O\,(3)\,, \tag{3.75}$$

we have that

$$D \in SO\,(3) \implies D^{-1} \in SO\,(3)\,, \tag{3.76}$$

i.e., existence of the inverse is satisfied. □

Corollary 1:

$$D \in O\,(3) \implies D \in SO\,(3) \vee \tilde{D} \in SO\,(3)\,, \tag{3.77}$$

where $D = R\tilde{D}$.

Proof: Recalling that $O_+ \cup O_- = SO\,(3) \cup O_- = O\,(3)$, we have that

$$D \in O\,(3) \implies D \in SO\,(3) \vee D \in O_-. \tag{3.78}$$

Assume that $D \in O_-$, and let $D = R\tilde{D}$. Then,

$$\begin{aligned} & \det D = -1 \\ \implies \quad & \det R \det \tilde{D} = -1 \\ \implies \quad & -\det \tilde{D} = -1 \\ \implies \quad & \det \tilde{D} = 1. \end{aligned} \tag{3.79}$$

Recalling that $R = R^{-1}$ and that $O(3)$ must satisfy closure, we have that $RD = R^2\tilde{D} = \tilde{D}$, so that $\tilde{D} \in O(3)$. Then, $\tilde{D} \in SO(3)$, so we have

$$D \in O_- \iff \tilde{D} \in SO(3)\,. \tag{3.80}$$

Substituting into equation 3.78,

$$D \in O(3) \implies D \in SO(3) \vee \tilde{D} \in SO(3)\,. \tag{3.81}$$

$\square$

Remark 3: Since we know R explicitly, it suffices to study $SO(3) \leq O(3)$ rather than $O(3)$.

Remark 4: O_- is not a group.

§ 2.3 Rotations about a fixed axis

Proposition 1: The most general element of $SO(3)$ that leaves the 3-axis fixed can be written as

$$D_3(\varphi) = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{3.82}$$

with $\varphi \in \big[0, 2\pi\big)$.

Proof: Recalling that proper transformations D take proper coordinate systems to proper coordinate systems, and that orthogonal transformations are proper in $O(3)$ (and hence in $SO(3)$), we have that

$$\begin{aligned} \tilde{e}_3 &= \left(D^{-1}\right)^i{}_3\, e_i \\ &= D_3{}^i e_i \\ &= D_3{}^1 e_1 + D_3{}^2 e_2 + D_3{}^3 e_3 \\ &\equiv e_3, \end{aligned} \tag{3.83}$$

where line 2 follows from the fact that orthogonal transformations must satisfy $D^{-1} = D^{\mathrm{T}}$, and line 4 follows from the fact that we are considering only transformations that leave the 3-axis fixed. Since each basis vector e_i is linearly independent, the only values for $D_3{}^i$ that yield line 4 from line 3 are

$$D_3{}^1 = D_3{}^2 = 0 \quad , \quad D_3{}^3 = 1. \tag{3.84}$$

Then, we have

$$D = \begin{pmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.85}$$

Again applying the fact that $D^{-1} = D^{\mathrm{T}}$, we have

$$\begin{aligned} DD^{\mathrm{T}} &= \begin{pmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & d & 0 \\ b & e & 0 \\ c & f & 1 \end{pmatrix} \\ &= \begin{pmatrix} a^2+b^2+c^2 & ad+be+cf & c \\ ad+be+cf & d^2+e^2+f^2 & f \\ c & f & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{3.86}$$

Setting the appropriate components equal, we have

$$\begin{aligned} c &= f \\ &= 0, \end{aligned} \tag{3.87}$$

so that setting the other appropriate components equal and substituting the above equation,

$$\begin{aligned} a^2+b^2 &= d^2+e^2 \\ &= 1, \end{aligned} \tag{3.88}$$

and

$$ad + be = 0. \tag{3.89}$$

Recalling that $a, b, d, e \in \mathbb{R}$, we have from equation 3.88 that $\exists \varphi, \psi$ such that

$$a = \cos\varphi \quad , \quad b = \sin\varphi \quad , \quad d = -\sin\psi \quad , \quad e = \cos\psi. \tag{3.90}$$

Substituting into equation 3.89 gives

$$\begin{aligned} & -\cos\varphi\sin\psi + \sin\varphi\cos\psi = 0 \\ \implies & \sin\varphi\cos\psi = \cos\varphi\sin\psi \\ \implies & \tan\varphi = \tan\psi \\ \implies & \varphi = \psi \text{ or } \varphi = \psi \pm \pi. \end{aligned} \tag{3.91}$$

From equations 3.85 and 3.90, we have that

$$\begin{aligned} \det D &= ae - bd \\ &= \cos\varphi\cos\psi + \sin\varphi\sin\psi \\ &= \begin{cases} 1, & \varphi = \psi \\ -1, & \varphi = \psi \pm \pi \end{cases}. \end{aligned} \tag{3.92}$$

But since $D \in SO(3)$, we must have $\det D = 1$, so that $\varphi = \psi$. Substituting into equation 3.90, we have

$$a = \cos\varphi \quad , \quad b = \sin\varphi \quad , \quad d = -\sin\varphi \quad , \quad e = \cos\varphi. \tag{3.93}$$

Substituting equations 3.87 and 3.93 into equation 3.85 gives

$$D = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.94}$$

□

Remark 1: Applying an element of $SO\left(3\right)$ that leaves the 3-axis fixed to a Cartesian coordinate system gives

$$\begin{aligned} \tilde{e}_1 &= D_1{}^i e_i \\ &= \cos\varphi e_1 + \sin\varphi e_2, \end{aligned} \tag{3.95}$$

$$\begin{aligned} \tilde{e}_2 &= D_2{}^i e_i \\ &= -\sin\varphi e_1 + \cos\varphi e_2, \end{aligned} \tag{3.96}$$

$$\tilde{e}_3 = e_3. \tag{3.97}$$

Such elements, which we can denote as $D_3\left(\varphi\right)$, rotate the coordinate system about the 3-axis by an angle φ.

Remark 2: By analogous proofs, the most general elements of $SO\left(3\right)$ that leave the 1-axis or 2-axis fixed can be denoted, respectively, as $D_1\left(\varphi\right)$ and $D_2\left(\varphi\right)$, and are writeable as

$$D_1\left(\varphi\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{pmatrix} \quad , \quad D_2\left(\varphi\right) = \begin{pmatrix} \cos\varphi & 0 & -\sin\varphi \\ 0 & 1 & 0 \\ \sin\varphi & 0 & \cos\varphi \end{pmatrix}. \tag{3.98}$$

Remark 3: For $i = 1, 2, 3$, we have that , $D_i\left(\varphi_1 + \varphi_2\right) = D_i\left(\varphi_1\right) D_i\left(\varphi_2\right)$, i.e., D_i is a homomorphism. We also have that $D_i\left(\varphi = 0\right) = 1$, $D_i^{-1}\left(\varphi\right) = D_i\left(-\varphi\right)$, and that $\left(\left\{D_i\left(\varphi\right)\right\}, \times\right)$ is associative by associativity of matrix multiplication. Then, since $\left\{D_i\left(\varphi\right)\right\} \subseteq SO\left(3\right)$, $\left(\left\{D_i\left(\varphi\right)\right\}, \times\right)$ is a one-dimensional subgroup of $SO\left(3\right)$. Additionally, since for $\varphi \in \mathbb{R}^+/2\pi$ each φ maps to a single $D_i\left(\varphi\right)$ and for every $D_i\left(\varphi\right)$ there exists a corresponding φ, D_i is a bijection, so that

$$\left(\left\{D_i\left(\varphi\right)\right\}, \times\right) \cong \left(\mathbb{R}, +\right)/2\pi. \tag{3.99}$$

§ 2.4 Infinitesimal rotations - the generators of $SO\left(3\right)$

Proposition 1: With $D_i\left(\varphi\right)$ defined as in §2.3, we can write

$$D_i\left(\varphi\right) = e^{I_i\varphi} \quad , \quad \left(I_i\right)_{jk} \equiv \varepsilon_{ijk} \quad , \quad \left(i = 1, 2, 3\right). \tag{3.100}$$

Proof: From equations 3.82 and 3.98, we have

$$\begin{aligned} \frac{d}{d\varphi} D_3\left(\varphi\right) &= \begin{pmatrix} -\sin\varphi & \cos\varphi & 0 \\ -\cos\varphi & -\sin\varphi & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= I_3 D_3\left(\varphi\right), \end{aligned} \tag{3.101}$$

$$\begin{aligned}
\frac{d}{d\varphi}D_2(\varphi) &= \begin{pmatrix} -\sin\varphi & 0 & -\cos\varphi \\ 0 & 0 & 0 \\ \cos\varphi & 0 & -\sin\varphi \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \cos\varphi & 0 & -\sin\varphi \\ 0 & 1 & 0 \\ \sin\varphi & 0 & \cos\varphi \end{pmatrix} \\
&= I_2 D_2(\varphi), \qquad (3.102)
\end{aligned}$$

and

$$\begin{aligned}
\frac{d}{d\varphi}D_1(\varphi) &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\sin\varphi & \cos\varphi \\ & -\cos\varphi & -\sin\varphi \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{pmatrix} \\
&= I_1 D_1(\varphi). \qquad (3.103)
\end{aligned}$$

Solving the differential equations above, we arrive at $D_i(\varphi) = e^{I_i\varphi}$. □

Remark 1*:* For any square matrix M, we can define e^M as

$$e^M \equiv \sum_{n=1}^{\infty} \frac{M^n}{n!}. \qquad (3.104)$$

Remark 2*:* I_1, I_2, and I_3 are called the **generators** of $SO(3)$.

Remark 3*:* Noting that $O(\varphi^2)$ terms vanish as $\varphi \to 0$, we have

$$\lim_{\varphi\to 0} D_i(\varphi) = 1 + I_i\varphi + O(\varphi^2). \qquad (3.105)$$

Theorem 1:

$$\forall D \in SO(3) \quad \exists \varphi_1, \varphi_2, \varphi_3 \quad D = e^{\sum_i I_i\varphi_i}. \qquad (3.106)$$

Proof: The proof can be found in texts on group theory. □

Remark 4*:* In equation 3.106, $D \neq D(\varphi_1) D(\varphi_2) D(\varphi_3)$. Otherwise, equation 3.106 would imply that rotations about different axes commute.

§ 2.5 Euler angles

Proposition 1: Consider the vector $re_3 = (0, 0, r)$ and let x be a vector with $x^2 = r^2$. Then there exist angles φ, θ such that

$$D_3^{-1}(\varphi) D_2^{-1}(\theta) re_3 = x. \qquad (3.107)$$

Proof: In spherical coordinates, we have

$$x_1 = r\sin\theta\cos\varphi \quad , \quad x_2 = r\sin\theta\sin\varphi \quad , \quad x_3 = r\cos\theta, \tag{3.108}$$

with

$$0 \le \theta \le \pi \quad , \quad -\pi \le \varphi < \pi. \tag{3.109}$$

From §2.3, we have

$$\begin{aligned} D_3^{-1}(\varphi)\,D_2^{-1}(\theta)\,re_3 &= \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ r \end{pmatrix} \\ &= \begin{pmatrix} (\cdots) & (\cdots) & \cos\varphi\sin\theta \\ (\cdots) & (\cdots) & \sin\varphi\sin\theta \\ (\cdots) & (\cdots) & \cos\theta \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ r \end{pmatrix} \\ &= \begin{pmatrix} r\sin\theta\cos\varphi \\ r\sin\theta\sin\varphi \\ r\cos\theta \end{pmatrix} \\ &= x. \end{aligned} \tag{3.110}$$

□

Corollary 1: Let x and y be vectors with $x^2 = y^2 = r^2$. Then there exist angles φ_1, φ_2, and θ such that

$$y = D_3^{-1}(\varphi_2)\,D_2^{-1}(\theta)\,D_3^{-1}(\varphi_1)\,x. \tag{3.111}$$

Proof: Recalling that $D_i^{-1}(\varphi) = D_i(-\varphi)$, we have from Proposition 1 that $\exists\varphi_1,\theta_1$ such that

$$\begin{aligned} & D_3^{-1}(-\varphi_1)\,D_2^{-1}(-\theta_1)\,re_3 = x \\ \Longrightarrow\quad & re_3 = D_2^{-1}(\theta_1)\,D_3^{-1}(\varphi_1)\,x, \end{aligned} \tag{3.112}$$

and that $\exists\varphi_2,\theta_2$ such that

$$\begin{aligned} y &= D_3^{-1}(\varphi_2)\,D_2^{-1}(\theta_2)\,re_3 \\ &= D_3^{-1}(\varphi_2)\,D_2^{-1}(\theta_2)\,D_2^{-1}(\theta_1)\,D_3^{-1}(\varphi_1)\,x \\ &= D_3^{-1}(\varphi_2)\,D_2^{-1}(\theta_1+\theta_2)\,D_3^{-1}(\varphi_1)\,x \\ &= D_3^{-1}(\varphi_2)\,D_2^{-1}(\theta)\,D_3^{-1}(\varphi_1)\,x, \end{aligned} \tag{3.113}$$

where line 2 follows from substituting equation 3.112, and where we have effectively defined $\theta \equiv \theta_1 + \theta_2$. □

Remark 1*:* Analogous statements hold for different pairs of axes, e.g., $\exists\varphi,\theta$ such that

$$D_3^{-1}(\varphi)\,D_1^{-1}(\theta)\,re_3 = x. \tag{3.114}$$

Theorem 1: Every element $D^{-1} \in SO(3)$ is writeable as

$$D^{-1} = D_3^{-1}(\psi)\,D_1^{-1}(\theta)\,D_3^{-1}(\varphi), \tag{3.115}$$

with

$$-\pi \le \varphi < \pi \quad , \quad -\pi \le \psi < \pi \quad , \quad 0 \le \theta \le \pi. \tag{3.116}$$

Proof: From §2.1 Remark 3, we have that

$$D^{-1} = \begin{pmatrix} (\cdots) & (\cdots) & x_1 \\ (\cdots) & (\cdots) & x_2 \\ (\cdots) & (\cdots) & x_3 \end{pmatrix} \quad , \quad x_1^2 + x_2^2 + x_3^2 = 1. \tag{3.117}$$

From Proposition 1, $\exists \psi, \theta$ such that

$$\begin{aligned} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= D_3^{-1}(\psi)\, D_1^{-1}(\theta) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ \implies \quad D^{-1} &= D_3^{-1}(\psi)\, D_1^{-1}(\theta) \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ e & f & 1 \end{pmatrix} \\ \implies \quad D_1^{-1}(-\theta)\, D_3^{-1}(-\psi)\, D^{-1} &= \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ e & f & 1 \end{pmatrix} \equiv C. \end{aligned} \tag{3.118}$$

Since $SO(3)$ satisfies closure and $D_1^{-1}(-\theta), D_3^{-1}(-\psi), D^{-1} \in SO(3)$, $C \in SO(3)$. Then, from §2.3 Proposition 1, e and f in equation 3.118 are both 0, implying that C leaves the 3-axis fixed. This implies, by the same proposition, that

$$\exists \varphi \quad C = D_3^{-1}(\varphi), \tag{3.119}$$

so that

$$D^{-1} = D_3^{-1}(\psi)\, D_1^{-1}(\theta)\, D_3^{-1}(\varphi). \tag{3.120}$$

□

Remark 2: The angles φ, θ, and ψ, described graphically in Figure 3.1 below, are called **Euler angles**. They characterize the transformation from $(1, 2, 3)$ to $\left(\tilde{1}, \tilde{2}, \tilde{3}\right)$, or vice versa.

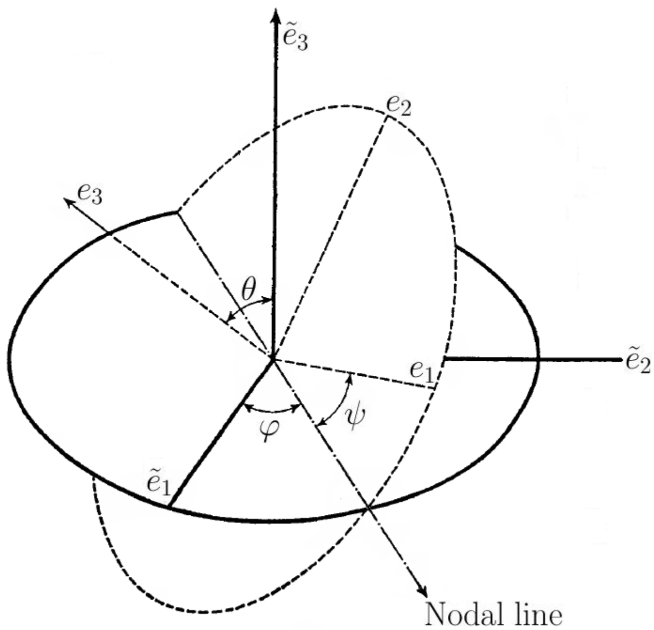

Figure 3.1: Euler angles. Adapted from Landau and Lifschitz Third Edition, Figure 47.

Remark 3*:* From §1.2, we have that

$$e_i = \left(D^{-1}\right)^j{}_i \, \tilde{e}_j, \tag{3.121}$$

so that the coordinate system cs is obtained from the coordinate system $\widetilde{\text{cs}}$ by rotation about the 3-axis by φ if $j = 1$, by rotation about the new 1-axis by θ if $j = 2$, and by rotation about the new 3-axis by ψ if $j = 3$.

§ 3 The rigid body model

§ 3.1 parameterization of a rigid body

Definition 1: A system of $N \geq 3$ point masses, at least three of which are not colinear[4], with fixed mutual distances is called a **rigid body**.

Remark 1*:* Rigid bodies are an obvious idealization of a real objects.

Remark 2*:* The interaction energy (Ch.2 §7.1 Axiom 2′) of a rigid body,

$$\begin{aligned} U &= \sum_{\alpha<\beta} U\left(r_{\alpha\beta}\right) \\ &= \text{constant}, \end{aligned} \tag{3.122}$$

[4]The non-colinearity requirement is not strictly necessary. However, a colinear rigid body is not a very interesting system, and the solution for such a system is fairly trivial. The formalism to be developed for the rigid body becomes useful only for the non-colinear case.

can be eliminated with a gauge transformation.

Remark 3*:* Because the relative distances among each of the point masses is known, the three point masses, x_0, x_1, and x_2, can be characterized by six coordinates - 3 coordinates for x_0 (x_0^1,x_0^2, and x_0^3), 2 coordinates for x_1 (e.g., 2 angles θ and φ), and 1 coordinate for x_2 (an angle). An additional point mass can be characterized by three additional coordinates, as well as three constraints, namely, the distances between the new point mass and each original point mass. This implies that a rigid body has six degrees of freedom[5].

Definition 2: An origin $\boldsymbol{R}$, together with three orthogonal basis vectors $\tilde{e}_i$, $(i = 1, 2, 3)$, which are fixed to a rigid body, is called a **moving coordinate system**, or a **body system** (abbreviated **BS**). An origin $(0, 0, 0)$, together with three basis vectors e_i, $(i = 1, 2, 3)$, is called a **laboratory system**, or an **inertial system** (abbreviated **IS**)[6].

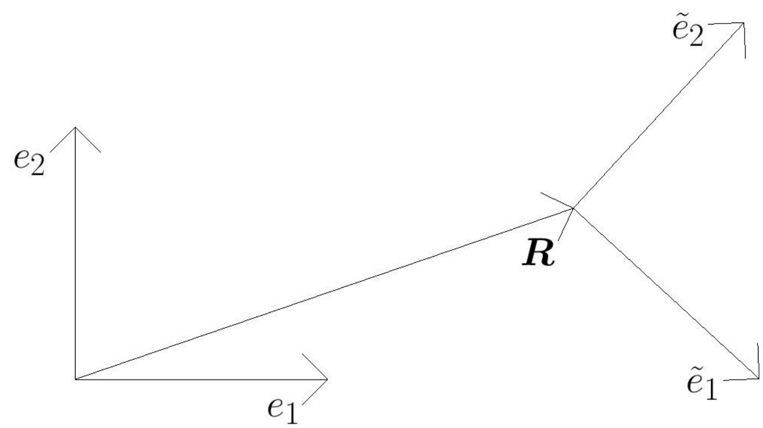

Figure 3.2: An inertial system with basis vectors e_i and a body system with basis vectors $\tilde{e}_i$.

Remark 4*:* A body system is related to an inertial system by a translation by $\boldsymbol{R}$ plus a basis transformation, given by

$$\begin{aligned} \tilde{e}_i &= \left(D^{-1}\right)^j{}_i \, e_j \\ &= D_i{}^j e_j. \end{aligned} \tag{3.123}$$

Remark 5*:* Taking the dot product of the basis vectors of the body system with those of the inertial

[5]Perhaps more intuitively, the point masses that make up a rigid body cannot move relative to each other. Then, a rigid body can move only as a whole. Its movement can occur as translations in the x-, y-, or z-directions, or rotations about the x-, y-, or z-axes, for a total of six degrees of freedom (three translational, three rotational).

[6]We denote body systems and inertial systems with and without tildes, respectively.

system,

$$\begin{aligned}\tilde{e}_i \cdot e_j &= D_i^{\ k} e_k \cdot e_j \\ &= D_i^{\ k} g_{kj} \\ &= D_{ij} \\ &= \left(D^{-1}\right)_{ji}. \end{aligned} \tag{3.124}$$

Remark 6: The position of the α'th point mass in the inertial system is given by

$$\boldsymbol{x}_{(\alpha)} = \boldsymbol{R} + \tilde{x}_{(\alpha)}^{\ \ i} \tilde{e}_i, \tag{3.125}$$

where

$$\boldsymbol{R} = \boldsymbol{R}(t) \quad , \quad \tilde{e}_i = D_i^{\ j}(t)\, e_j = \tilde{e}_i(t), \tag{3.126}$$

and hence,

$$\boldsymbol{x}_{(\alpha)} = \boldsymbol{x}_{(\alpha)}(t) \quad , \quad \tilde{x}_{(\alpha)}^{\ \ i} = \text{constant}. \tag{3.127}$$

Remark 7: The state of a rigid body is completely determined by the position of the body system with respect to the inertial system, so that we need only determine the body system's movement as a whole, rather than the movement of each of its components.

Remark 8: The body system, and hence the rigid body, is completely characterized by the three coordinates of $\boldsymbol{R}$, plus three Euler angles.

Example 1: If the Euler angles, ψ, θ, and φ, characterizing the basis transformation (see §2.5 Corollary 1) in equation 3.123 are constant, then the body system is always related to the inertial system purely by the translation $\boldsymbol{R} = \boldsymbol{R}(t)$.

Example 2: If the translation, $\boldsymbol{R}$, characterizing the relation between the body system and the inertial system is constant, then the body system is related to the inertial system purely by a rotation about the 3-axis. Then, equation 3.126 can be replaced with

$$\tilde{e}_3 = e_3 \quad , \quad D^{-1}(t) = \begin{pmatrix} \cos\psi(t) & \sin\psi(t) & 0 \\ -\sin\psi(t) & \cos\psi(t) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.128}$$

§ 3.2 Angular velocity

Proposition 1: Let

$$\boldsymbol{x}(t) \equiv \boldsymbol{R}(t) + \boldsymbol{r}(t), \tag{3.129}$$

be a point in the rigid body as described in an inertial system, with $\boldsymbol{r}(t)$ the same point as described in the body system (See Figure 3.3). Then,$\boldsymbol{v}(t) \equiv \dot{\boldsymbol{x}}(t)$ is writeable as

$$\boldsymbol{v}(t) = \boldsymbol{V}(t) + \boldsymbol{\Omega}(t) \times \boldsymbol{r}(t), \tag{3.130}$$

where $\boldsymbol{V}(t) \equiv \dot{\boldsymbol{R}}(t)$, and where, recalling §2.1 Corollary 1, $\boldsymbol{\Omega}$ is a pseudovector corresponding to the antisymmetric tensor

$$\dot{D}^{-1} D = \begin{pmatrix} 0 & \Omega_3 & -\Omega_2 \\ -\Omega_3 & 0 & \Omega_1 \\ \Omega_2 & -\Omega_1 & 0 \end{pmatrix}. \tag{3.131}$$

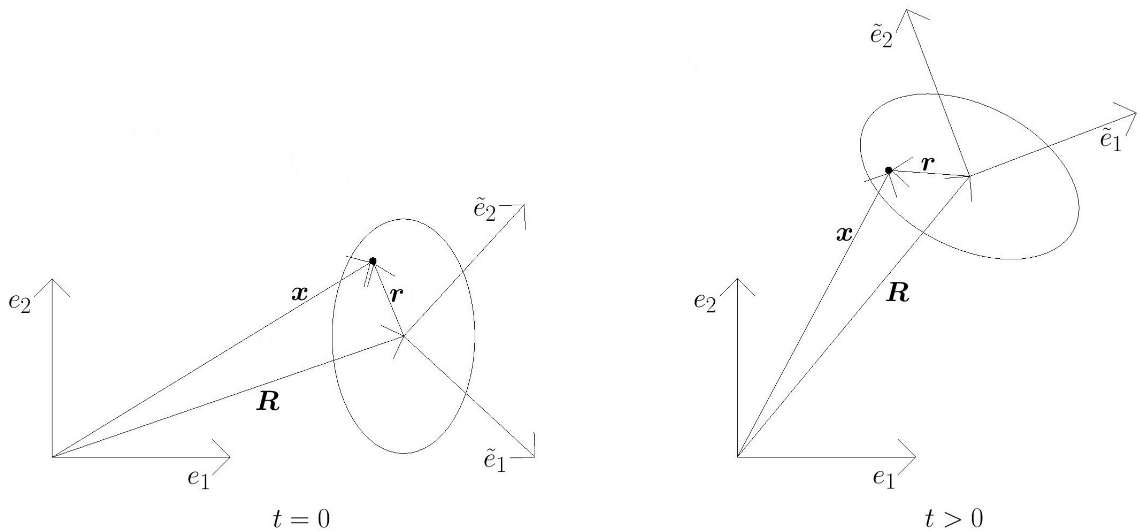

Figure 3.3: A rigid body and a point on it at different times, as described in the inertial system and the body system.

Proof: Taking the time-derivative of equation 3.129 and substituting equation 3.130,

$$\boldsymbol{v} = \boldsymbol{V} + \dot{\boldsymbol{r}}. \tag{3.132}$$

We also have that

$$r_i(t) = \left(D^{-1}(t)\right)_i^{\ j} \tilde{r}_j, \tag{3.133}$$

where $\tilde{r}_j = \tilde{r}_j(t) = \text{constant}$ is a component of $\boldsymbol{r}(t)$ as defined above. Then,

$$\begin{aligned} \dot{r}_i(t) &= \left(\dot{D}^{-1}(t)\right)_i^{\ j} \tilde{r}_j \\ &= \left(\dot{D}^{-1}(t)\right)_i^{\ j} D(t)_j^{\ k} r_k \\ &= -\tau_i^{\ k} r_k, \end{aligned} \tag{3.134}$$

where we have effectively defined τ as a second-rank tensor,

$$\begin{aligned} \tau_i^{\ k} &\equiv -\left(\dot{D}^{-1}(t)\right)_i^{\ j} D(t)_j^{\ k} \\ &= \left(\dot{D}^{-1} D\right)_i^{\ k}. \end{aligned} \tag{3.135}$$

But, we also have that

$$\begin{aligned} & D^{-1}D = 1 \\ \Longrightarrow\quad & \dot{D}^{-1}D + D^{-1}\dot{D} = 0 \\ \Longrightarrow\quad & \tau^{\mathrm{T}} = -\left(\dot{D}^{-1}D\right)^{\mathrm{T}} = -D^{\mathrm{T}}\left(\dot{D}^{-1}\right)^{\mathrm{T}} = -D^{-1}\dot{D}. \end{aligned} \tag{3.136}$$

Substituting line 2 into line 3,

$$\begin{aligned}\tau^{\mathrm{T}} &= \dot{D}^{-1}D \\ &= -\tau.\end{aligned} \tag{3.137}$$

Then, τ is antisymmetric, implying that

$$\exists \Omega_1, \Omega_2, \Omega_3 \quad \tau = \begin{pmatrix} 0 & \Omega_3 & -\Omega_2 \\ -\Omega_3 & 0 & \Omega_1 \\ \Omega_2 & -\Omega_1 & 0 \end{pmatrix}. \tag{3.138}$$

Then, we can write

$$\begin{aligned}\dot{r}_1 &= -{\tau_1}^j r_j \\ &= -\Omega_3 r_2 + \Omega_2 r_3 \\ &= (\boldsymbol{\Omega} \times \boldsymbol{r})_1 ,\end{aligned} \tag{3.139}$$

and similarly for $\dot{r}_2$ and $\dot{r}_3$, so that

$$\dot{\boldsymbol{r}} = \boldsymbol{\Omega} \times \boldsymbol{r}. \tag{3.140}$$

Substituting into equation 3.132, we have

$$\boldsymbol{v} = \boldsymbol{V} + \boldsymbol{\Omega} \times \boldsymbol{r}. \tag{3.141}$$

□

Example 1: Consider a rotation about the 3-axis by an angle $\psi(t)$. We have

$$\begin{aligned}
& D^{-1} = \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
\Longrightarrow\quad & \dot{D}^{-1} = \begin{pmatrix} -\sin\psi & -\cos\psi & 0 \\ \cos\psi & -\sin\psi & 0 \\ 0 & 0 & 0 \end{pmatrix} \dot{\psi} \quad , \quad D = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
\Longrightarrow\quad & \tau = -\dot{D}^{-1}D = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \dot{\psi} \\
\Longrightarrow\quad & \boldsymbol{\Omega}(t) = \left(0, 0, \dot{\psi}(t)\right),
\end{aligned} \tag{3.142}$$

where lines 3 and 4 follow from equations 3.137 and 3.138, respectively.

Remark 1*:* Consider points $\boldsymbol{r}$ satisfying $\boldsymbol{r} = k\boldsymbol{\Omega}$, with k a constant, so that from equation 3.130, $\boldsymbol{v} = \boldsymbol{V}$. These points, and only these points, do not rotate about the body system's 3-axis. For this reason, $\boldsymbol{\Omega}(t)$ is called the **instantaneous axis of rotation**.

Remark 2*:* If we change the origin of the body system as

$$\boldsymbol{R} \to \boldsymbol{R}' \equiv \boldsymbol{R} + \boldsymbol{a}, \tag{3.143}$$

then we have

$$
\begin{aligned}
& \boldsymbol{x} = \boldsymbol{R} + \boldsymbol{r} = \boldsymbol{R}' + \boldsymbol{r}' = \boldsymbol{R} + (\boldsymbol{r}' + \boldsymbol{a}) \\
\Longrightarrow \quad & \boldsymbol{r} = \boldsymbol{r}' + \boldsymbol{a} \\
\Longrightarrow \quad & \boldsymbol{v} = \boldsymbol{V} + \boldsymbol{\Omega} \times (\boldsymbol{r}' + \boldsymbol{a}) = \boldsymbol{V} + \boldsymbol{\Omega} \times \boldsymbol{r}' + \boldsymbol{\Omega} \times \boldsymbol{a} \equiv \boldsymbol{V}' + \boldsymbol{\Omega}' \times \boldsymbol{r}', \\
\Longrightarrow \quad & \boldsymbol{V}' = \boldsymbol{V} + \boldsymbol{\Omega} \times \boldsymbol{a} \quad , \quad \boldsymbol{\Omega}' = \boldsymbol{\Omega}.
\end{aligned} \tag{3.144}
$$

From the last line of the above equation, $\boldsymbol{\Omega}$ is independent of the choice of $\boldsymbol{R}$.

Remark 3: $\tau = -\dot{D}^{-1} D$ is called the **angular velocity tensor**, and[7] $\boldsymbol{\Omega} \cong \tau$ is called the **angular velocity (pseudovector)**.

Remark 4: The coordinates of $\boldsymbol{\Omega}$ in the inertial system can be found as as

$$
\begin{aligned}
\tau_{ij} &= -\left(\dot{D}^{-1} D\right)_{ij} \\
&\cong \Omega_i.
\end{aligned} \tag{3.145}
$$

In the body system, the coordinates can be found as

$$
\begin{aligned}
& \tilde{\tau}_{ij} = D_i^{\ k} D_j^{\ l} \tau_{kl} = \left(D \tau D^{\mathrm{T}}\right)_{ij} \\
\Longrightarrow \quad & \tilde{\tau} = D \tau D^{\mathrm{T}} = -D \dot{D}^{-1} D D^{\mathrm{T}} = -D \dot{D}^{-1} = \begin{pmatrix} 0 & \tilde{\Omega}_3 & -\tilde{\Omega}_2 \\ -\tilde{\Omega}_3 & 0 & \tilde{\Omega}_1 \\ \tilde{\Omega}_2 & -\tilde{\Omega}_1 & 0 \end{pmatrix}
\end{aligned} \tag{3.146}
$$

Remark 5: The convention for writing the components of $\boldsymbol{\Omega}$ in the body system and inertial system is

$$
\boldsymbol{\Omega} \underset{\text{BS}}{\longrightarrow} \left(\tilde{\Omega}_1, \tilde{\Omega}_2, \tilde{\Omega}_3\right) \equiv (p, q, r) \quad , \quad \boldsymbol{\Omega} \underset{\text{IS}}{\longrightarrow} (\Omega_1, \Omega_2, \Omega_3). \tag{3.147}
$$

Remark 6: An alternative convention, used in Landau and Lifschitz, denotes $\boldsymbol{\Omega}$ in the body system as

$$
\boldsymbol{\Omega} \underset{\text{BS,LL}}{\longrightarrow} (\Omega_1, \Omega_2, \Omega_3). \tag{3.148}
$$

Remark 7: Note that $\boldsymbol{\Omega}$, $\boldsymbol{V}$, $\boldsymbol{r}$, etc. are abstract vectors independent of the coordinate system. Only their components depend on the choice of coordinate system. For instance,

$$
\begin{aligned}
\boldsymbol{\Omega}(t) &= \Omega_i(t)\, e_i \\
&= \tilde{\Omega}_i(t)\, \tilde{e}_i(t).
\end{aligned} \tag{3.149}
$$

Proposition 2: The components of $\boldsymbol{\Omega}$ in the body system, (p, q, r), can be written in terms of the Euler angles as

$$
\boxed{p = \dot{\psi} \sin\theta \sin\varphi + \dot{\theta} \cos\varphi \quad , \quad q = \dot{\psi} \sin\theta \cos\varphi - \dot{\theta} \sin\varphi \quad , \quad r = \dot{\psi} \cos\theta + \dot{\varphi}. \qquad (3.150)}
$$

[7] In the second line of equation 3.145, §2.1 Corollary 1 is applied. Often in what follows, this corollary is being applied when a "≅" is seen.

Proof: From §2.5 Theorem 1, we can write

$$D = D_3(\varphi) D_1(\theta) D_3(\psi), \tag{3.151}$$

so that

$$\begin{aligned}\dot{D}^{-1} &= \dot{D}_3^{-1}(\psi) D_1^{-1}(\theta) D_3^{-1}(\varphi) + D_3^{-1}(\psi) \dot{D}_1^{-1}(\theta) D_3^{-1}(\varphi) \\ &\quad + D_3^{-1}(\psi) D_1^{-1}(\theta) \dot{D}_3^{-1}(\varphi).\end{aligned} \tag{3.152}$$

Then, from equation 3.135,

$$\begin{aligned}-\tilde{\tau} &= D\dot{D}^{-1} \\ &= D_3(\varphi) D_1(\theta) \left[D_3(\psi) \dot{D}_3^{-1}(\psi)\right] D_1^{-1}(\theta) D_3^{-1}(\varphi) \\ &\quad + D_3(\varphi) \left[D_1(\theta) \dot{D}_1^{-1}(\theta)\right] D_3^{-1}(\varphi) + D_3(\varphi) \dot{D}_3^{-1}(\varphi).\end{aligned} \tag{3.153}$$

Recalling §2.4 Theorem 1, we have

$$\begin{aligned}D_i(\alpha) \dot{D}_i^{-1}(\alpha) &= e^{\alpha I_i} \frac{d}{dt} e^{-\alpha I_i} \\ &= -\dot{\alpha} I_i.\end{aligned} \tag{3.154}$$

Then,

$$\begin{aligned}D_3(\varphi) \dot{D}_3^{-1}(\varphi) &= -\dot{\varphi} I_3 \\ &= -\dot{\varphi} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &\cong -\dot{\varphi} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.\end{aligned} \tag{3.155}$$

Similarly,

$$\begin{aligned}D_3(\varphi) \left[D_1(\theta) \dot{D}_1^{-1}(\theta)\right] D_3^{-1}(\varphi) &= -\dot{\theta} D_3(\varphi) I_1 D_3^{-1}(\varphi) \\ &= -\dot{\theta} \begin{pmatrix} 0 & 0 & \sin\varphi \\ 0 & 0 & \cos\varphi \\ -\sin\varphi & -\cos\varphi & 0 \end{pmatrix} \\ &\cong -\dot{\theta} \begin{pmatrix} \cos\varphi \\ -\sin\varphi \\ 0 \end{pmatrix},\end{aligned} \tag{3.156}$$

and

$$D_3(\varphi) D_1(\theta) \left[D_3(\psi) \dot{D}_3^{-1}(\psi)\right] D_1^{-1}(\theta) D_3^{-1}(\varphi) \cong -\dot{\psi} \begin{pmatrix} \sin\theta\sin\psi \\ \sin\theta\cos\varphi \\ \cos\theta \end{pmatrix}, \tag{3.157}$$

where the SO(3) generators I_i were substitued in each of equations 3.155, 3.156, and 3.157. Substituting into equation 3.153,

$$\tilde{\tau} \cong -\dot{\psi}\begin{pmatrix}\sin\theta\sin\varphi\\ \sin\theta\cos\varphi\\ \cos\theta\end{pmatrix} + \dot{\theta}\begin{pmatrix}\cos\varphi\\ -\sin\varphi\\ 0\end{pmatrix} + \dot{\varphi}\begin{pmatrix}0\\0\\1\end{pmatrix}. \tag{3.158}$$

Writing equation 3.158 as the product of a matrix and a vector and applying equations 3.138 and 3.147 gives

$$\begin{pmatrix}p\\q\\r\end{pmatrix} = \begin{pmatrix}\sin\theta\sin\varphi & \cos\varphi & 0\\ \sin\theta\cos\varphi & -\sin\varphi & 0\\ \cos\theta & 0 & 1\end{pmatrix}\begin{pmatrix}\dot{\psi}\\ \dot{\theta}\\ \dot{\varphi}\end{pmatrix}. \tag{3.159}$$

□

Remark 8*:* The Landau and Lifschitz convention for Euler angles differs from ours as $\varphi \leftrightarrow \psi$.

Remark 9*:* The relation between the BS components of $\boldsymbol{\Omega}$, (p, q, r), and the time derivatives of the Euler angles, $\left(\dot{\psi}, \dot{\theta}, \dot{\varphi}\right)$, is given by a linear but non-orthogonal transformation.

Remark 10*:* If $p(t)$, $q(t)$, and $r(t)$ are known, then equation 3.150 effectively gives $\psi(t)$, $\theta(t)$, and $\varphi(t)$ as the solution of a system of three coupled first-order nonlinear ODEs.

§ 3.3 The inertia tensor

Let $\boldsymbol{R}$ be the origin of a given body system. Then, the position and velocity of the α'th particle as viewed in the inertial system are

$$\boldsymbol{x}_{(\alpha)} = \boldsymbol{R} + \boldsymbol{r}_{(\alpha)} \quad , \quad \boldsymbol{v}_{(\alpha)} = \boldsymbol{V} + \boldsymbol{\Omega} \times \boldsymbol{r}_{(\alpha)}, \tag{3.160}$$

and the center of mass of the body system as viewed in the inertial system is

$$\boldsymbol{X} = \frac{1}{M}\sum_{\alpha} m_\alpha \boldsymbol{x}_{(\alpha)} \quad , \quad M = \sum_{\alpha} m_\alpha. \tag{3.161}$$

Definition 1: The tensor I which has components[8]

$$\boxed{I_{ij} = \sum_{\alpha} m_\alpha \left[\delta_{ij} {r_{(\alpha)}}^l {r_{(\alpha)}}^l - {r_{(\alpha)}}^i {r_{(\alpha)}}^j\right] \tag{3.162}}$$

in the inertial system, and

$$\boxed{\tilde{I}_{ij} = \sum_{\alpha} m_\alpha \left[\delta_{ij} {\tilde{r}_{(\alpha)}}^l {\tilde{r}_{(\alpha)}}^l - {\tilde{r}_{(\alpha)}}^i {\tilde{r}_{(\alpha)}}^l\right] \tag{3.163}}$$

in the body system is called the **inertia tensor** of the rigid body with respect to $\boldsymbol{R}$.

[8] In much of what follows, the Einstein summation convention becomes cumbersome. Because we are in Euclidean space, upper indices and lower indices are interchangeable, and we will not distinguish between them in any way. Summations are therefore implied over any repeated indices, regardless to their position, unless noted otherwise.

Definition 2: The body system with $\boldsymbol{R} = \boldsymbol{X}$, where $\boldsymbol{X}$ is given in equation 3.161, is called the **center-of-mass system**, abbreviated **CMS**.

Remark 1*:* The inertia tensor I is a symmetric rank-2 tensor.

Remark 2*:* In general, the components of the inertia tensor as viewed in the inertial frame, I_{ij}, are time-dependent. However, the components of the inertia tensor as viewed in the body system, $\tilde{I}_{ij}$, are time-independent.

Theorem 1: If $\boldsymbol{R} = \boldsymbol{X}$, then the Lagrangian of the rigid body is given by

$$\begin{aligned} L &= \frac{M}{2} V_i V^i + \frac{1}{2} I^{ij} \Omega_i \Omega_j \\ &= \frac{M}{2} \tilde{V}_i \tilde{V}^i + \frac{1}{2} \tilde{I}^{ij} \tilde{\Omega}_i \tilde{\Omega}_j. \end{aligned} \tag{3.164}$$

Proof: Recalling §3.1 Remark 2, we have

$$\begin{aligned} L &= T \\ &= \sum_\alpha \frac{m_\alpha}{2} \left(\boldsymbol{v}_{(\alpha)}\right)^2 \\ &= \sum_\alpha \frac{m_\alpha}{2} \left[\boldsymbol{V}^2 + 2\boldsymbol{V} \cdot \left(\boldsymbol{\Omega} \times \boldsymbol{r}_{(\alpha)}\right) + \left(\boldsymbol{\Omega} \times \boldsymbol{r}_{(\alpha)}\right)^2\right], \end{aligned} \tag{3.165}$$

where line 2 follows from equation 3.160. Applying the vector identity

$$(\boldsymbol{a} \times \boldsymbol{b}) \cdot (\boldsymbol{c} \times \boldsymbol{d}) = (\boldsymbol{a} \cdot \boldsymbol{c})(\boldsymbol{b} \cdot \boldsymbol{d}) - (\boldsymbol{a} \cdot \boldsymbol{d})(\boldsymbol{b} \cdot \boldsymbol{c}), \tag{3.166}$$

to the last term in equation 3.165, rearrangin, and applying equation 3.161, we have

$$L = \frac{M}{2} \boldsymbol{V}^2 + \boldsymbol{V} \cdot \boldsymbol{\Omega} \times (\boldsymbol{X} - \boldsymbol{R}) M + \frac{1}{2} \sum_\alpha m_\alpha \left[\boldsymbol{\Omega}^2 \left(\boldsymbol{r}_{(\alpha)}\right)^2 - \left(\boldsymbol{\Omega} \cdot \boldsymbol{r}_{(\alpha)}\right)^2\right]. \tag{3.167}$$

Substituting $\boldsymbol{R} = \boldsymbol{X}$,

$$\begin{aligned} L &= \frac{M}{2} \boldsymbol{V}^2 + \frac{1}{2} \sum_\alpha m_\alpha \left[\delta_{ij} {r_{(\alpha)}}^l {r_{(\alpha)}}^l - {r_{(\alpha)}}^i {r_{(\alpha)}}^j\right] \Omega_i \Omega_j \\ &= \frac{M}{2} V_i V^i + \frac{1}{2} I^{ij} \Omega_i \Omega_j \\ &= \frac{M}{2} \tilde{V}_i \tilde{V}^i + \frac{1}{2} \tilde{I}^{ij} \tilde{\Omega}_i \tilde{\Omega}_j, \end{aligned} \tag{3.168}$$

where the last two lines follow from Definition 1. □

Remark 3*:* The Lagrangian L is the sum of the kinetic energy due to translational motion and the kinetic energy due to rotational motion of the rigid body.

§ 3.4 Classification of rigid bodies

Let $\tilde{I}_{ij}$ be the components of the inertia tensor as described in the CMS (see §3.3). Because I is real-symmetric, there exists a body system, called the **principal axis system** (abbreviated **PAS**), satisfying (no summation implied)[9]

$$\boxed{\tilde{I}_{ij} = \delta_{ij} I_i. \qquad (3.169)}$$

Remark 1*:* Because the components $\tilde{I}_{ij}$ are time-independent, they characterize the rigid body at rest.

Remark 2*:* The eigenvalues of I are called the **principal moments of inertia**. The eigenvectors of I are called the **principal axes of inertia**.

Remark 3*:* Note that each $I_i \geq 0$, implying that I is positive-definite. Additionally, note that

$$\begin{aligned} I_1 + I_2 &= \sum_\alpha m_\alpha \left[2\left(\boldsymbol{r}_{(\alpha)}\right)^2 - \left(\tilde{r}_{(\alpha)}{}^1\right)^2 - \left(\tilde{r}_{(\alpha)}{}^2\right)^2 \right] \\ &= \sum_\alpha m_\alpha \left[\left(\boldsymbol{r}_{(\alpha)}\right)^2 + \left(\tilde{r}_{(\alpha)}{}^3\right)^2 \right] \\ &\geq \sum_\alpha m_\alpha \left[\left(\boldsymbol{r}_{(\alpha)}\right)^2 - \left(\tilde{r}_{(\alpha)}{}^3\right)^2 \right] \\ &= I_3, \end{aligned} \qquad (3.170)$$

so that

$$I_1 + I_2 \geq I_3. \qquad (3.171)$$

Based on the above remarks, we can classify rigid bodies based on their principal moments of inertia, as follows:

a) A rigid body satisfying $I_1 = I_2 = I_3$, such as a cube, is called a **spherical top**. For such rigid bodies, every body system is a principal axis system.

b) A rigid body satisfying $I_2 = I_2 \neq I_3$, such as a pear, is called a **symmetric top**. For such rigid bodies, the third principal axis is fixed, and the other two may be rotated about it.

c) A rigid body satisfying $I_1 \neq I_2 \neq I_3$, such as a (non-square) book, is called an **asymmetric top**.

§ 3.5 Angular momentum of a rigid body

Definition 1: The **orbital angular momentum** $\boldsymbol{M}$ of a rigid body is defined as

$$M_i \equiv I_{ij} \Omega^j. \qquad (3.172)$$

[9]To be consistent with our convention, we should write $\tilde{I}_{ij} = \delta_{ij}\tilde{I}_i$ in place of equation 3.169. However, I_i is used without a tilde in almost all texts, despite referring to body system componets, so we will use this notation as well. Additionally, note that I_i is a one-index shorthand for an entity with two indices. This shorthand is effectively defined in equation 3.169, which says that $\tilde{I}$, as described in the PAS, is diagonal.

Theorem 1: The total angular momentum $\boldsymbol{L}$ of a rigid body for $\boldsymbol{R} = \boldsymbol{X}$ is given by

$$\boldsymbol{L} = \boldsymbol{X} \times \boldsymbol{P} + \boldsymbol{M}, \tag{3.173}$$

where

$$\boldsymbol{P} = M\boldsymbol{V}. \tag{3.174}$$

Proof: The total angular momentum due to all point masses in the rigid body is given by

$$\begin{aligned} \boldsymbol{L} &= \sum_\alpha \boldsymbol{l}_{(\alpha)} \\ &= \sum_\alpha m_\alpha \boldsymbol{x}_{(\alpha)} \times \boldsymbol{v}_{(\alpha)} \\ &= \sum_\alpha m_\alpha \left(\boldsymbol{X} + \boldsymbol{r}_{(\alpha)}\right) \times \left(\boldsymbol{V} + \boldsymbol{\Omega} \times \boldsymbol{r}_{(\alpha)}\right) \\ &= M\boldsymbol{X} \times \boldsymbol{V} + \boldsymbol{X} \times \left(\boldsymbol{\Omega} \times \sum_\alpha m_\alpha \boldsymbol{r}_{(\alpha)}\right) + \sum_\alpha m_\alpha \boldsymbol{r}_{(\alpha)} \times \boldsymbol{V} \\ &\quad + \sum_\alpha m_\alpha \boldsymbol{r}_{(\alpha)} \times \left(\boldsymbol{\Omega} \times \boldsymbol{r}_{(\alpha)}\right), \end{aligned} \tag{3.175}$$

where lines 3 and 4 follow from equations 3.160 and 3.161, respectively. Consider that

$$\begin{aligned} \boldsymbol{\Omega} \times \sum_\alpha m_\alpha \boldsymbol{r}_{(\alpha)} &= \sum_\alpha m_\alpha \left(\boldsymbol{x}_{(\alpha)} - \boldsymbol{R}\right) \\ &= M\left(\boldsymbol{X} - \boldsymbol{R}\right) \\ &= \boldsymbol{0}, \end{aligned} \tag{3.176}$$

where line 3 follows from substituting $\boldsymbol{R} = \boldsymbol{X}$. Substituting equations 3.174 and 3.176 into equation 3.175 gives

$$\begin{aligned} \boldsymbol{L} &= \boldsymbol{X} \times \boldsymbol{P} + \sum_\alpha m_\alpha \boldsymbol{r}_{(\alpha)} \times \left(\boldsymbol{\Omega} \times \boldsymbol{r}_{(\alpha)}\right) \\ &= \boldsymbol{X} \times \boldsymbol{P} + \sum_\alpha m_\alpha \left[\boldsymbol{\Omega}\left(\boldsymbol{r}_{(\alpha)}\right)^2 - \boldsymbol{r}_{(\alpha)}\left(\boldsymbol{r}_{(\alpha)} \cdot \boldsymbol{\Omega}\right)\right]. \end{aligned} \tag{3.177}$$

Then,

$$\begin{aligned} L_i &= \left(\boldsymbol{X} \times \boldsymbol{P}\right)_i + \sum_\alpha \left[\Omega_i {r_{(\alpha)}}^l {r_{(\alpha)}}^l - {r_{(\alpha)}}^i {r_{(\alpha)}}^j \Omega_j\right] \\ &= \left(\boldsymbol{X} \times \boldsymbol{P}\right)_i + \sum_\alpha \left[\delta_{ij} {r_{(\alpha)}}^l {r_{(\alpha)}}^l - {r_{(\alpha)}}^i {r_{(\alpha)}}^j\right] \Omega_j \\ &= \left(\boldsymbol{X} \times \boldsymbol{P}\right)_i + I_{ij}\Omega_j \\ &= \left(\boldsymbol{X} \times \boldsymbol{P}\right)_i + M_i, \end{aligned} \tag{3.178}$$

where equations 3.162 and 3.172 were applied. Then, we can write

$$\boldsymbol{L} = \boldsymbol{X} \times \boldsymbol{P} + \boldsymbol{M}. \tag{3.179}$$

$\square$

Remark 1*:* Note that $\boldsymbol{X} \times \boldsymbol{P}$ is the angular momentum of a point mass M moving along the path of the center of mass, and that $\boldsymbol{M}$ is the angular momentum of the rigid body due to rotation with respect to the center of mass.

Remark 2*:* The relation between $\boldsymbol{M}$ and $\boldsymbol{\Omega}$ in the IS is given by

$$M_i = I_{ij}\Omega_j, \tag{3.180}$$

where the components I_{ij} are time-dependent, in general. The relation between $\boldsymbol{M}$ and $\boldsymbol{\Omega}$ in the principal axis system is given by (no summation implied)

$$\tilde{M}_i = I_i\tilde{\Omega}_i, \tag{3.181}$$

where the components I_i are time-independent, in general.

Remark 3*:* $\boldsymbol{M}$ is parallel to $\boldsymbol{\Omega}$, in general, only if the rigid body is a spherical top, and only in the PAS.

§ 3.6 The continuum model of rigid bodies

Consider a continuous rigid body characterized by a mass density $\rho(\boldsymbol{x})$. The continuum limits of the mass M, center of mass $\boldsymbol{X}$, and inertia tensor I are respectively given by

$$M = \int d\boldsymbol{x}\, \rho(\boldsymbol{x}), \tag{3.182}$$

$$\boldsymbol{X} = \frac{1}{M}\int d\boldsymbol{x}\, \boldsymbol{x}\rho(\boldsymbol{x}), \tag{3.183}$$

and

$$\tilde{I}_{ij} = \int d\boldsymbol{x}\, \rho(\boldsymbol{x}) \left[\delta_{ij}\boldsymbol{x}^2 - x_i x_j\right]. \tag{3.184}$$

Remark 1*:* As in the discrete model, rigid bodies in the continuum model have six degrees of freedom, given, for instance, by the components of $\boldsymbol{R}$ and by ψ, θ, and φ.

Remark 2*:* The relations between body systems and inertial systems discussed for the discrete model still hold in the continuum model.

Remark 3*:* The equations of motion, to be discussed in §3.7, hold for both models.

Example 1: Consider a homogeneous cube, with mass density

$$\rho(\boldsymbol{x}) = \begin{cases} \rho, & \boldsymbol{x} \in \text{cube with sides of length } a \\ 0, & \text{otherwise} \end{cases}. \tag{3.185}$$

From equation 3.184, we have

$$\tilde{I}_{ij} = \rho\int_{-a/2}^{a/2} dx_1 \int_{-a/2}^{a/2} dx_2 \int_{-a/2}^{a/2} dx_3 \left[\delta_{ij}\left(x_1^2 + x_2^2 + x_3^2\right) - x_i x_j\right]. \tag{3.186}$$

For $i \neq j$,

$$\begin{aligned}
\tilde{I}_{i,j\neq i} &= (\cdots)\int_{-a/2}^{a/2} dx_j\, x_j \\
&= 0.
\end{aligned} \tag{3.187}$$

For $i = j = 1$,

$$\begin{aligned}
\tilde{I}_{11} &= \rho\int_{-a/2}^{a/2} dx_1 \int_{-a/2}^{a/2} dx_2 \int_{-a/2}^{a/2} dx_3\ \left[x_2^2 + x_3^2\right] \\
&= 2\rho a^2 \int_{-a/2}^{a/2} dx\, x^2 \\
&= \frac{1}{6}\rho a^5 \\
&= \frac{1}{6} M a^2,
\end{aligned} \tag{3.188}$$

where the last line follows from the fact that $M = \rho a^3$. Identical logic applies for $i = j = 2$ and $i = j = 3$, so combining the above results, we have

$$\tilde{I}_{ij} = \frac{1}{6} M a^2 \delta_{ij}. \tag{3.189}$$

Remark 4*:* The homogeneous cube is a spherical top.

§ 3.7 The Euler equations

We will go about finding the equations of motion for rigid bodies by first finding them in the principal axis system, and then transforming them to find them in the inertial system. Below, we denote the principal axis system and inertial system with and without a tilde, respectively.

Theorem 1: Letting $\boldsymbol{N}$ denote torque, the equations of motion in the principal axis system are given by the **Euler equations**,

$$\boxed{I_i \dot{\tilde{\Omega}}_i - (I_{i+1} - I_{i+2})\,\tilde{\Omega}_{i+1}\tilde{\Omega}_{i+2} = \tilde{N}_i \quad , \quad (i = 1, 2, 3 \text{ and cyclic})\,.} \tag{3.190}$$

Proof: We have from equation 3.172,

$$\begin{aligned}
M_i &= \left(D^{-1}\right)_i^{\ k} \tilde{M}_k \\
&= I_{ij}\Omega_j \\
&= \left(D^{-1}\right)_i^{\ k} \tilde{I}_{kj}\tilde{\Omega}_j,
\end{aligned} \tag{3.191}$$

where, since we are in the PAS,

$$\tilde{I}_{kj} = \delta_{kj} I_j. \tag{3.192}$$

Then,

$$\begin{aligned}
\dot{M}_i &= \left(\dot{D}^{-1}\right)_i^{\ k} \tilde{I}_{kj}\tilde{\Omega}_j + \left(D^{-1}\right)_i^{\ k} \tilde{I}_{kj}\dot{\tilde{\Omega}}_j \\
&= N_i \\
&= \left(D^{-1}\right)_i^{\ j} \tilde{N}_j,
\end{aligned} \tag{3.193}$$

where line 2 follows from conservation of angular momentum. Comparing the RHS of line 1 with line 3 and recalling §3.2 Remark 4, we have

$$\begin{aligned}
\begin{pmatrix}\tilde{N}_1\\ \tilde{N}_2\\ \tilde{N}_3\end{pmatrix} &= \tilde{I}_{ij}\dot{\tilde{\Omega}}^j + D_i^{\ j}\left(\dot{D}^{-1}\right)_j^{\ k}\tilde{I}_{kl}\tilde{\Omega}^l \\
&= \tilde{I}_{ij}\dot{\tilde{\Omega}}^j - \tilde{\tau}_i^{\ k}\tilde{I}_{kl}\tilde{\Omega}^l \\
&= \begin{pmatrix} I_1 & 0 & 0\\ 0 & I_2 & 0\\ 0 & 0 & I_3\end{pmatrix}\begin{pmatrix}\dot{\tilde{\Omega}}_1\\ \dot{\tilde{\Omega}}_2\\ \dot{\tilde{\Omega}}_3\end{pmatrix} \\
&\quad - \begin{pmatrix} 0 & \tilde{\Omega}_3 & -\tilde{\Omega}_2\\ -\tilde{\Omega}_3 & 0 & \tilde{\Omega}_1\\ \tilde{\Omega}_2 & -\tilde{\Omega}_1 & 0\end{pmatrix}\begin{pmatrix} I_1 & 0 & 0\\ 0 & I_2 & 0\\ 0 & 0 & I_3\end{pmatrix}\begin{pmatrix}\tilde{\Omega}_1\\ \tilde{\Omega}_2\\ \tilde{\Omega}_3\end{pmatrix}.
\end{aligned} \tag{3.194}$$

Comparing the first and last lines, we have equation 3.190. □

Remark 1*:* The Euler equations describe the motion of the instantaneous axis of rotation in the principal axis system (which, recall, is a special case of a body system).

Remark 2*:* The coordinates of $\mathbf{\Omega}$ in the body system can be found by solving equation 3.190. Then, the Euler angles that describe the motion in the inertial system can be found by solving equation 3.150 in §3.2.

Remark 3*:* The Euler equations are nonlinear, even for free motion. Furthermore, equation 3.150 in §3.2 adds additional nonlinearities.

Remark 4*:* See Landau and Lifschitz for how equation 3.190 follows directly from the Euler-Lagrange equations.

§ 4 Simple examples of rigid-body motion

§ 4.1 The physical pendulum

Definition 1: A **physical pendulum** is a rigid body that rotates about a fixed axis $\boldsymbol{a}$ under the influence of gravity.

Let $\boldsymbol{a} = (0, 0, 1)$, and choose $\boldsymbol{R}$ to point from the origin of the IS to $\boldsymbol{a}$. Let $\boldsymbol{X} \neq \boldsymbol{R}$, and let

$$\boldsymbol{b} \equiv |\boldsymbol{X} - \boldsymbol{R}|\,. \tag{3.195}$$

Remark 1*:* In the above setup, we are not working in a CMS.

Letting φ denote the angle between $\boldsymbol{b}$ and a vertical line through $\boldsymbol{a}$, and recalling equations 3.126 and 3.128 from §3.1, the transformation from the BS to the IS proceeds as

$$\tilde{e}_i = D_i^{\ j}(t)\, e_j = \tilde{e}_i(t)\,, \tag{3.196}$$

with

$$D^{-1} = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\Longrightarrow \quad \dot{D}^{-1} = \begin{pmatrix} -\sin\varphi & -\cos\varphi & 0 \\ \cos\varphi & -\sin\varphi & 0 \\ 0 & 0 & 0 \end{pmatrix} \dot{\varphi} \quad , \quad D = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.197}$$

From equations 3.137 and 3.138 in §3.2,

$$-\tau = \dot{D}^{-1} D = \dot{\varphi} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\Longrightarrow \quad \Omega_i = -\delta_{i3}\dot{\varphi}. \tag{3.198}$$

The kinetic energy (see §3.3) and potential energy of the system are writeable as

$$T = \frac{1}{2} I_{33} \dot{\varphi}^2 \quad , \quad U = -Mgb\cos\varphi. \tag{3.199}$$

We also have that (see problem 38)

$$\begin{aligned} I_{33} &= I_3 + Mb^2 \\ &\equiv \theta. \end{aligned} \tag{3.200}$$

Substituting equation 3.200 into equation 3.199, we can write our Lagrangian as

$$L = \frac{1}{2}\theta\dot{\varphi}^2 + Mgb\cos\varphi. \tag{3.201}$$

Remark 2*:* The equation for a **mathematical pendulum** is given by

$$L = \frac{1}{2} m l^2 \dot{\varphi}^2 + mgl\cos\varphi. \tag{3.202}$$

Comparing with equation 3.201, it is clear that the physical pendulum is equivalent to the mathematical pendulum, but with **effective length** and **effective mass**, respectively, given by

$$l = \frac{\theta}{Mb} \quad , \quad m = \frac{M^2 b^2}{\theta}. \tag{3.203}$$

§ 4.2 Cylinder on an inclined plane

Consider an unconstrained, homogeneous cylinder of radius a. We can choose our body system with

$$\begin{aligned} \boldsymbol{R} &= \boldsymbol{X} \\ &= (x, y). \end{aligned} \tag{3.204}$$

Remark 1*:* The BS chosen in equation 3.204 is a CMS.

Let $I = I_{33}$ be the moment of inertia with respect to the cylinder axis, and let φ denote the angle between a vertical line through $\boldsymbol{X}$ and a fixed point on the circumfrance on the cylinder. Then, we can write the Lagrangian as (see §3.3 for the kinetic energy term)

$$L = \frac{M}{2}\left(\dot{x}^2 + \dot{y}^2\right) + \frac{1}{2} I \dot{\varphi}^2 - Mgy. \tag{3.205}$$

Our system is not unconstrained - instead, it is on an inclined plane and is not allowed to "slip". Then, in addtion to equation 3.205, we have two constraints, respectively given by (with α the angle between the horizontal and the inclined plane)

$$\begin{aligned} & y = x \tan\alpha \quad , \quad a\varphi = \sqrt{x^2 + y^2}. \\ \Longrightarrow \quad & x = Ry \quad , \quad \varphi = \frac{y}{a}\sqrt{1 + R^2} \quad , \quad R = \cot\alpha. \end{aligned} \tag{3.206}$$

Substituting into equation 3.205 gives

$$L = \frac{M}{2}\left(1 + R^2\right)\dot{y}^2 + \frac{1}{2}\frac{I}{a^2}\dot{y}^2 - Mgy. \tag{3.207}$$

Letting

$$\mu \equiv \left(M + \frac{I}{a^2}\right)\left(1 + R^2\right) \quad , \quad g_{\text{eff}} = \frac{gM}{\mu}, \tag{3.208}$$

equation 3.207 becomes

$$\begin{aligned} L &= \frac{\mu}{2}\dot{y}^2 - Mgy \\ &= \frac{\mu}{2}\dot{y}^2 - \mu g_{\text{eff}} y. \end{aligned} \tag{3.209}$$

Remark 2*:* From equation 3.209, the problem is equivalent to that of a free-falling body with mass μ and effective gravitational acceleration g_{eff}, both of which are given by equation 3.208.

Remark 3*:* Applying the Euler-Lagrange equations to equation 3.209 gives

$$\begin{aligned} & \mu\ddot{y} = Mg \\ \Longrightarrow \quad & \ddot{y} = \frac{gM}{\mu} = g_{\text{eff}}, \end{aligned} \tag{3.210}$$

as expected from Remark 2.

Remark 4*:* The same solution can be arrived at through a more elementary means as follows: If the inertia tensor is taken with respect to the axis of roation, the torque is given by

$$Mga\sin\alpha = \theta\ddot{\varphi}, \tag{3.211}$$

with (see Problem 38)

$$\begin{aligned} \theta &\equiv I_{33} \\ &= I + Ma^2. \end{aligned} \tag{3.212}$$

Solving equation 3.211 for $\ddot{\varphi}$, we have

$$\ddot{\varphi} = \frac{Mga}{\theta} \sin\alpha. \tag{3.213}$$

From geometrical considerations (which are effectively equivalent to applying the constraints given in equation 3.206), we have

$$\begin{aligned} \ddot{y} &= \frac{a}{\sqrt{1+R^2}}\ddot{\varphi} \\ &= \frac{Ma^2}{\theta} g\sin^2\alpha \\ &= \frac{Ma^2}{I+Ma^2} g\sin^2\alpha \\ &= \frac{gM}{\mu} \\ &= g_{\text{eff}}, \end{aligned} \tag{3.214}$$

so that equation 3.210 is reproduced.

§ 4.3 The force-free symmetric top

Let the $\tilde{3}$-axis be the **figure axis**, so that $I_1 = I_2 \neq I_3$. Then, from the Euler equations given in §3.7, and recalling the notation introduced in §3.2 Remark 5, we have

$$I_1\dot{p} - (I_1 - I_3)\, qr = 0, \tag{3.215}$$

$$I_1\dot{q} - (I_3 - I_1)\, rp = 0, \tag{3.216}$$

and

$$I_3\dot{r} = 0. \tag{3.217}$$

From equation 3.217, we have that $\dot{r} = 0$, so we can write

$$r\,(t) \equiv r_0. \tag{3.218}$$

Substituting into equations 3.215 and 3.216 and rearranging gives

$$\dot{p} = \frac{I_1 - I_3}{I_1} r_0 q \quad , \quad \dot{q} = -\frac{I_1 - I_3}{I_1} r_0 p. \tag{3.219}$$

Substituting

$$\omega \equiv \left(1 - \frac{I_3}{I_1}\right) r_0 \tag{3.220}$$

gives

$$\begin{aligned} & \dot{p} = \omega q \quad \dot{q} = -\omega p \\ \Longrightarrow \quad & \ddot{p} = -\omega^2 p. \end{aligned} \tag{3.221}$$

Solving the differential equation gives

$$p(t) = p_0 \cos \omega t + \dot{p}_0 \sin \omega t. \tag{3.222}$$

Choosing $t = 0$ such that $\dot{p}_0 \equiv \dot{p}(t = 0) = 0$ eliminates the second term in the above equation. Then, substituting into the first line of equation 3.221, we can solve for $q(t)$. The solution for the force-free symmetric top is then

$$p(t) = p_0 \cos \omega t \quad , \quad q(t) = -p_0 \sin \omega t \quad , \quad r(t) = r_0. \tag{3.223}$$

Remark 1*:* In the BS, where the $\tilde{3}$-axis, i.e., the figure axis, is fixed, the instantaneous axis of rotation $\mathbf{\Omega}$ with components $\left(\tilde{\Omega}_1, \tilde{\Omega}_2, \tilde{\Omega}_3\right)$ rotates about the figure axis with a frequency of $\omega = (1 - I_3/I_1)\,\tilde{\Omega}_3$, where $\tilde{\Omega}_3$ is constant. This motion is called the **precession**.

Remark 2*:* The earth can be approximated as a symmetric top, with $2\pi/\omega \approx 10$ months.

Remark 3*:* From equation 3.181 in §3.5, we have

$$\tilde{M}_1 = I_1 p_0 \cos \omega t \quad , \quad \tilde{M}_2 = -I_1 p_0 \sin \omega t \quad , \quad \tilde{M}_3 = I_3 r_0, \tag{3.224}$$

so that the orbital angular momentum $\boldsymbol{M}$ in the BS is always coplanar to the figure axis and $\mathbf{\Omega}$ (i.e., is always in the plane created by the figure axis and $\mathbf{\Omega}$).

Remark 4*:* In the IS, $\boldsymbol{M}$ is fixed, while $\mathbf{\Omega}$ and the figure axis rotate about $\boldsymbol{M}$ with frequency ω.

Remark 5*:* The heavy symmetric top (i.e., the symmetric top subject to gravity) displays a precession, as is seen here, as well as additional phenomena (see problems 40 and 42).

§ 4.4 The force-free asymmetric top

Let $I_1 \neq I_2 \neq I_3$. Equation 3.181 from §3.5 and the Euler equations from §3.7 give

$$\begin{aligned} \dot{\tilde{M}}_1 &= \dot{I}_1 \dot{\tilde{\Omega}}_1 \\ &= \frac{I_2 - I_3}{I_2 I_3} \tilde{M}_2 \tilde{M}_3 \\ &= \left(\frac{1}{I_3} - \frac{1}{I_2}\right) \tilde{M}_2 \tilde{M}_3 \\ &= \alpha_1 \tilde{M}_2 \tilde{M}_3, \end{aligned} \tag{3.225}$$

$$\begin{aligned} \dot{\tilde{M}}_2 &= \left(\frac{1}{I_1} - \frac{1}{I_3}\right) \tilde{M}_3 \tilde{M}_1 \\ &= \alpha_2 \tilde{M}_3 \tilde{M}_1, \end{aligned} \tag{3.226}$$

and

$$\begin{aligned} \dot{\tilde{M}}_3 &= \left(\frac{1}{I_2} - \frac{1}{I_1}\right) \tilde{M}_1 \tilde{M}_2 \\ &= \alpha_3 \tilde{M}_1 \tilde{M}_2, \end{aligned} \tag{3.227}$$

where

$$\alpha_1 \equiv 1/I_3 - 1/I_2 \quad , \quad \alpha_2 \equiv 1/I_1 - 1/I_3 \quad , \quad \alpha_3 \equiv 1/I_2 - 1/I_1. \tag{3.228}$$

Lemma 1: The force-free asymmetric top has the property that

$$\frac{d}{dt}\left(\tilde{M}_1^2 + \tilde{M}_2^2 + \tilde{M}_3^2\right) = 0. \tag{3.229}$$

Proof: From the product rule and from equations 3.225 through 3.228,

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\left(\tilde{M}_1^2 + \tilde{M}_2^2 + \tilde{M}_3^2\right) &= \tilde{M}_1\dot{\tilde{M}}_1 + \tilde{M}_2\dot{\tilde{M}}_2 + \tilde{M}_3\dot{\tilde{M}}_3 \\ &= \underbrace{(\tilde{\alpha}_1 + \tilde{\alpha}_2 + \tilde{\alpha}_3)}_{=0}\tilde{M}_1\tilde{M}_2\tilde{M}_3 \\ &= 0. \end{aligned} \tag{3.230}$$

□

Remark 1: Lemma 1 follows from conservation of angular momentum, since $\boldsymbol{M}^2$ is a scalar, and hence must be the same in both the BS and the IS.

Remark 2: Note that $\boldsymbol{M}$ itself is constant in the IS, but that its direction changes in the BS.

For notational convenience, let

$$(x_1, x_2, x_3) \equiv \left(\tilde{M}_1, \tilde{M}_2, \tilde{M}_3\right). \tag{3.231}$$

Then, from equations 3.225, 3.226, and 3.227, we have a system of nonlinear ODEs,

$$\dot{x}_1 = \alpha_1 x_2 x_3 \quad , \quad \dot{x}_2 = \alpha_2 x_3 x_1 \quad , \quad \dot{x}_3 = \alpha_3 x_1 x_2, \tag{3.232}$$

with each α_i given by equation 3.228. From Lemma 1, with x_i's in place of $\tilde{M}_i$'s, we have

$$x_1^2 + x_2^2 + x_3^2 = \text{constant}, \tag{3.233}$$

so that equation 3.232 describes motion on a 2-sphere.

Remark 3: From equation 3.232, the asymmetric top is described by a system of nonlinear ODEs. In the case of the symmetric top, by contrast, one of the α_i's is 0, so that one of the x_i's is constant. As a result, the system of ODEs becomes linear for the symmetric top (see §4.3).

§ 4.4.1 Fixed points

Definition 1: A point $\boldsymbol{x} = (x_1, x_2, x_3)$ is called a **fixed point** if $\dot{\boldsymbol{x}} = 0$.

Proposition 1: The fixed points for which equation 3.232 is satisfied are

$$\boldsymbol{x}_{1\pm} = (\pm 1, 0, 0) \quad , \quad \boldsymbol{x}_{2\pm} = (0, \pm 1, 0) \quad , \quad \boldsymbol{x}_{3\pm} = (0, 0, \pm 1), \tag{3.234}$$

where the points $\boldsymbol{x}$ are redefined so as to be normalized.

Proof: For a fixed point, we have that $\dot{\boldsymbol{x}} = 0$, so that $\dot{x}_1 = \dot{x}_2 = \dot{x}_3 = 0$. But from equation 3.228 and the fact that $I_1 \neq I_2 \neq I_3$, none of α_1, α_2, or α_3 are 0, so from equation 3.232, we have that

$$\dot{x}_i = 0 \implies x_j = 0 \vee x_k = 0 \quad , \quad (i \neq j \neq k)\,, \tag{3.235}$$

where $i, j, k \in \{1, 2, 3\}$. Equation 3.235 must be satisfied for each possible permuation of values for i, j, and k, which is only possible if

$$x_k \neq 0 \implies x_i = x_j = 0 \quad , \quad (i \neq j \neq k)\,. \tag{3.236}$$

But the points $\boldsymbol{x}$ are normalized (i.e., confined to a unit 2-sphere), so we must also have that

$$\boldsymbol{x}^2 = 1. \tag{3.237}$$

The only points satisfying equations 3.236 and 3.237 are those given in the proposition. □

Remark 1*:* Note that equation 3.232 is invariant under any transformation of the form

$$(x_i, x_{j\neq i}) \to (-x_i, -x_{j\neq i})\,, \tag{3.238}$$

where $i, j \in \{1, 2, 3\}$. Then, $\boldsymbol{x}_{i+}$ and $\boldsymbol{x}_{i-}$, as written in equation 3.234, are equivalent, so that it suffices to study

$$\boldsymbol{x}_1 = (1, 0, 0) \quad , \quad \boldsymbol{x}_2 = (0, 1, 0) \quad , \quad \boldsymbol{x}_3 = (0, 0, 1) \tag{3.239}$$

in place of equation 3.234.

§ 4.4.2 Linearization near fixed points

Question. Does a system prepared near, but not at, a fixed point have any interesting properties?

Consider the following point displaced from the fixed point $\boldsymbol{x}_1$

$$\begin{aligned} \boldsymbol{x} &= \left(\sqrt{1 - \varepsilon_2^2 - \varepsilon_3^2}, \varepsilon_2, \varepsilon_3\right) \\ &= (1, \varepsilon_2, \varepsilon_3) + O\left(\varepsilon^2\right) \\ &= \boldsymbol{x}_1 + (0, \varepsilon_2, \varepsilon_3) + O\left(\varepsilon^2\right), \end{aligned} \tag{3.240}$$

where $\varepsilon_2 \backsimeq \varepsilon_3 \simeq O(\varepsilon)$. Then, applying equation 3.232, we have

$$\dot{x}_1 = O\left(\varepsilon^2\right) \quad , \quad \dot{x}_2 = \alpha_2 \varepsilon_3 + O\left(\varepsilon^2\right) \quad , \quad \dot{x}_3 = \alpha_3 \varepsilon_2 + O\left(\varepsilon^2\right), \tag{3.241}$$

or equivalently, from equation 3.240,

$$\dot{x}_1 = O\left(\varepsilon^2\right) \quad , \quad \dot{x}_2 = \alpha_2 x_3 + O\left(\varepsilon^2\right) \quad , \quad \dot{x}_3 = \alpha_3 x_2 + O\left(\varepsilon^2\right). \tag{3.242}$$

Then, in the limit $\varepsilon \to 0$, i.e., in the vicinity of $\boldsymbol{x}_1$, we have

$$\dot{x}_1 = 0 \quad , \quad \dot{x}_2 = \alpha_2 x_3 \quad , \quad \dot{x}_3 = \alpha_3 x_2, \tag{3.243}$$

which is a system of linear ODEs for x_1, x_2, and x_3. Identical logic can be applied to find equations analogous to equation 3.243 in the vicinity of the fixed points $\boldsymbol{x}_2$ and $\boldsymbol{x}_3$. Then, in the vicinity of the fixed point $\boldsymbol{x}_i$ we have

$$\dot{x}_i = 0 \quad , \quad \dot{x}_{i+1} = \alpha_{i+1} x_{i+2} \quad , \quad \dot{x}_{i+2} = \alpha_{i+2} x_{i+1}, \tag{3.244}$$

so that equivalently, the $\dot{x}_{i+1}$ and $\dot{x}_{i+2}$ equations can be written as

$$\begin{pmatrix} \dot{x}_{i+1} \\ \dot{x}_{i+2} \end{pmatrix} = \begin{pmatrix} 0 & \alpha_{i+1} \\ \alpha_{i+2} & 0 \end{pmatrix} \begin{pmatrix} x_{i+1} \\ x_{i+2} \end{pmatrix} + O\left(\varepsilon^2\right). \tag{3.245}$$

where the indices $i \in \{1, 2, 3\}$ cycle as $i + 3 \equiv i$ in equations 3.244 and 3.245.

§ 4.4.3 Solution to the linearized equations

Lemma 1: Let M be a 2×2 matrix with eigenvalues λ_1 and $\lambda_2 \neq \lambda_1$, and with eigenvectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$, so that

$$M\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i \quad , \quad (i = 1, 2). \tag{3.246}$$

Furthermore, let a system of linear ODEs be writeable as

$$\dot{\boldsymbol{x}} = M\boldsymbol{x}. \tag{3.247}$$

Then, the general solution to equation 3.247 is given by

$$\boldsymbol{x}(t) = c_1 \boldsymbol{v}_1 e^{\lambda_1 t} + c_2 \boldsymbol{v}_2 e^{\lambda_2 t}, \tag{3.248}$$

where c_1 and c_2 are constants.

Proof: Taking the time-derivative of equation 3.248 gives

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= c_1 \lambda_1 \boldsymbol{v}_1 e^{\lambda_1 t} + c_2 \lambda_2 \boldsymbol{v}_2 e^{\lambda_2 t} \\ &= c_1 M \boldsymbol{v}_1 e^{\lambda_1 t} + c_2 M \boldsymbol{v}_2 e^{\lambda_2 t} \\ &= M\left(c_1 \boldsymbol{v}_1 e^{\lambda_1 t} + c_2 \boldsymbol{v}_2 e^{\lambda_2 t}\right) \\ &= M\boldsymbol{x}(t), \end{aligned} \tag{3.249}$$

where line 2 follows from equation 3.246. Comparing the first and last lines, we have equation 3.247. Then, equation 3.248 is indeed a solution to equation 3.247. Furthermore, the solution must be unique, since the system is linear. Then, equation 3.248 is the general solution to equation 3.247. □

The matrix associated with the fixed points is given in equation 3.245 from §4.4.2. Its eigenvalues are

$$\lambda^i_{\pm} = \pm\sqrt{\alpha_{i+1}\alpha_{i+2}}, \tag{3.250}$$

and its corresponding eigenvectors are

$$\boldsymbol{v}_{\pm} = \begin{pmatrix} \pm\sqrt{\alpha_{i+1}\alpha_{i+2}} \\ 1 \end{pmatrix}. \tag{3.251}$$

Then, Lemma 2 can be applied to equation 3.245 to find the general solution for $x_{i+1}(t)$ and $x_{i+2}(t)$ in the vicinity of the fixed point $\boldsymbol{x}_i$.

Definition 1: A fixed point is called **stable** if the eigenvalues associated with it are imaginary, and **unstable** if they are real.

Remark 1*:* A system in the vicinity of a stable fixed point shows oscillatory motion in the vicinity of that fixed point. A system in the vicinity of an unstable fixed point shows run-away behavior of the time-derivatives of the orbital angular momentum about that fixed point, so that the system does not stay in the linear regime.

Theorem 1: Of the three fixed points of a force-free asymmetric top, two are stable and one is unstable.

Proof: From equation 3.228,

$$\alpha_1 + \alpha_2 + \alpha_3 = 0. \tag{3.252}$$

Since $I_1 \neq I_2 \neq I_3$, none of α_1, α_2, or α_3 are 0 (see equation 3.228), so we must have that two of the α_i's are positive and the other is negative, or two are negative and the other is positive. Let

$$\alpha_1 > 0 \quad , \quad \alpha_2 > 0 \quad , \quad \alpha_3 < 0. \tag{3.253}$$

Then,

$$\alpha_1\alpha_2 > 0, \tag{3.254}$$

so that $\boldsymbol{x}_3$ is unstable, and

$$\alpha_1\alpha_3 < 0 \quad , \quad \alpha_2\alpha_3 < 0, \tag{3.255}$$

so that $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are stable. Identical logic gives one unstable fixed point and two stable fixed points for other choices of which α_i's are positive or negative. □

Remark 2*:* A graphical representation of a numerical solution for $\alpha_1 = 1$, $\alpha_2 = -2$, and $\alpha_3 = 1$, corresponding to $I_1 = 1$, $I_2 = 1/2$, and $I_3 = 1/3$, can be found in Bender and Orszag Figure 4.31. This is given as Figure 3.4 below.

Remark 3*:* The instability described in Theorem 1 is sometimes called the **Euler wobble**. It can lead to catastrophic instability of high-speed aircraft when the engine is hot.

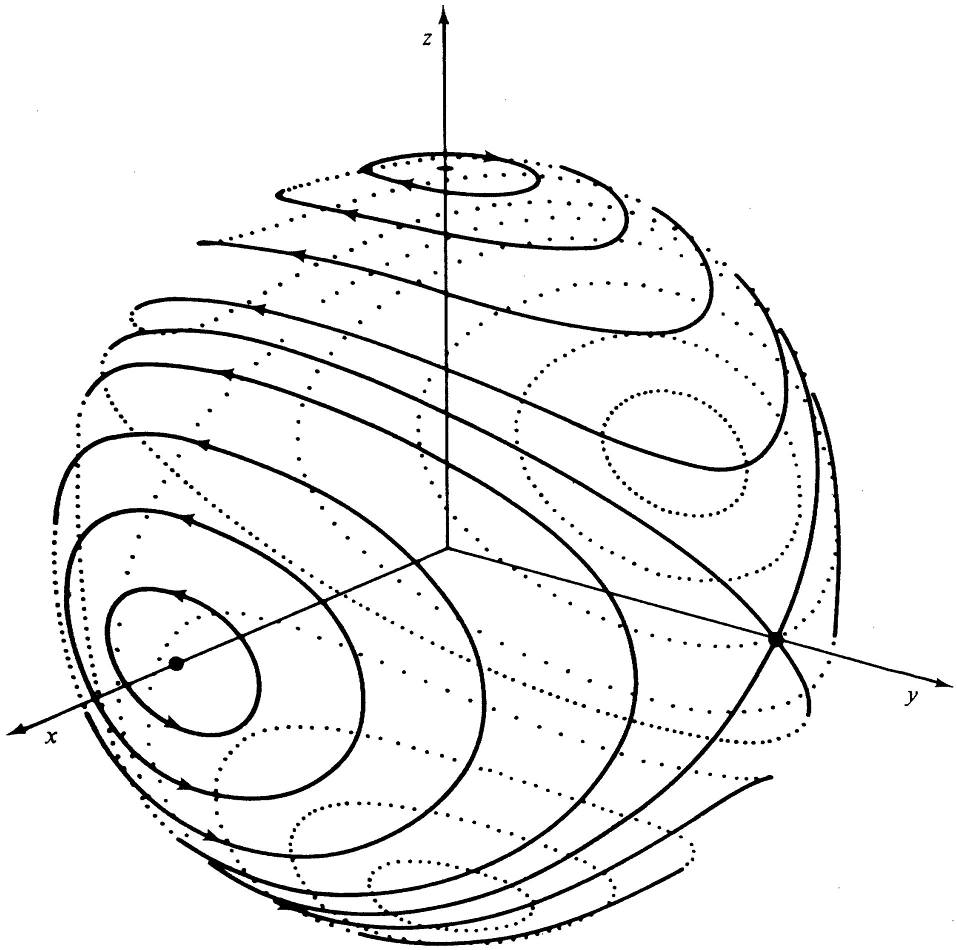

Figure 3.4: Graphical representation for the case $\alpha_1 = 1$, $\alpha_2 = -2$, and $\alpha_3 = 1$. Taken from Bender and Orszag, Figure 4.31.

§ 5 Problems for Chapter 3

35. Lorentz transformation

Consider a 2-d Minkowski space with metric

$$g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.256}$$

Prove the following statements:

a) $D(\varphi) = \begin{pmatrix} \cosh\varphi & \sinh\varphi \\ \sinh\varphi & \cosh\varphi \end{pmatrix}$ is a proper coordinate transformation.

b) The set $\{D(\varphi)\}$ forms a group under matrix multiplication, and the mapping $\varphi \to D(\varphi)$ defines an isomorphism between this group and the real numbers under addition.

c) There exists a fixed matrix J (the *generator* of the group) such that any $D(\varphi)$ can be written $D(\varphi) = e^{J\varphi}$. Determine J explicitly.

36. Properties of tensors

a) For every proper coordinate system in n-dimensional space one defines

$$\varepsilon^{i_1 i_2 \dots i_n} = \begin{cases} +1, & (i_1 \dots i_n) \text{ is an even permutation of } (1 \dots n) \\ -1, & (i_1 \dots i_n) \text{ is an odd permutation of } (1 \dots n) \\ 0, & \text{otherwise} \end{cases} . \tag{3.257}$$

Show that ε is a pseudotensor of rank n.

b) Show that every second-rank tensor t^{ij} can be uniquely written as

$$t^{ij} = t_0^{ij} + t_+^{ij} + t_-^{ij}, \tag{3.258}$$

where t_0, t_+, and t_- have the following properties: $t_0^{ij} = t g^{ij}$ with t a scalar, $t_-^{ij} = -t_-^{ji}$ (i.e., t_- is antisymmetric), and $t_+^{ij} = t_+^{ji}$ with $t_+^{ij} g_{ji} = 0$ (i.e., t_+ is symmetric and traceless).

c) Show that an object t with coordinates t^{ij} is a second-rank tensor if and only if for every (pseudo)vector x the object defined by $y^i = t^{ij} x_j$ is a (pseudo)vector.

d) Prove §1.5 Lemma 1 and Proposition 1.

e) Prove §1.5 Lemma 2 and Corollary 1.

37. Small-angle scattering

Approximate the orbit of the earth as a Keplarian ellipse with $e = 0.0167$ and $T = 365.24$ days. The plane of the orbit is called the *ecliptic*. The *equator* of the earth defines a plane that is tilted against the ecliptic by an angle $\gamma = 23.44^\mathrm{o}$. The two points common to both the equator plane and the orbit are the called *spring equinox* (SE) and *fall equinox* (FE). 90^o away are the *winter solstice* (WS) and *summer solstice* (SS). The angle $\tilde{\omega}$ between the SE and the perihelion (P) is called the *longitude of perihelion*, because of the precession of the earth's axis, $\tilde{\omega}$, is time-dependent ($\Delta\tilde{\omega} \approx 50''$/year). In geocentric coordinates, one analogously defines the equator, the ecliptic, etc., on the *celestial sphere*. The *nodal line* is defined by the points SE and FE.

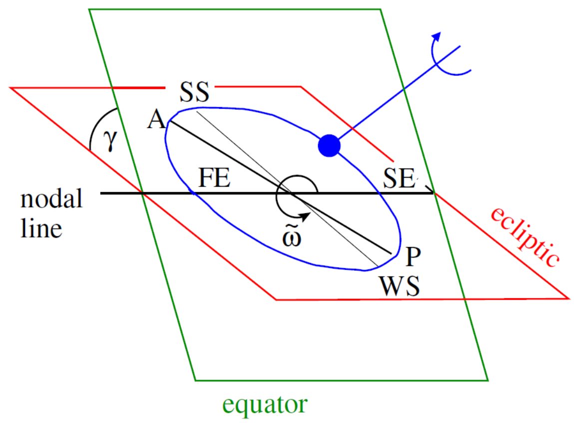

Figure 3.5

a) In 2011, th earth will be in the SE on March 20, 23:21 universal time (UT, see below) (or 7:21pm EDT), and $\tilde{\omega} \approx 283.1^o$. Calculate the times when the earth will be at SS, A, FE, WS, and P with an accuracy of 1 h. Show that within that accuracy you don't have to worry about the precession. Which is shorter in 2011, winter or summer, and why?
Hint: First show that for small e, the time evolution of the azimuthal angle is given by $\varphi(t) = \omega(t) + 2e\sin(\omega t) + O\left(e^2\right)$, with $\omega = 2\pi/T$, and use this expression.

Consider the position of the sun in the geocentric system, where the sun moves on the celestial sphere along the ecliptic. Define the angle ψ between the SE and the projection of the sun's position onto the equator parallel to the earth's axis (i.e., perpendicular to the equator). We distinguish between

i) ψ_s for the apparent or physically observed sun,

ii) ψ_{us} for a fictitious *uniform sun* that travels uniformly ($\varphi(t) = \omega t$), along the ecliptic, and

iii) ψ_{ms} for a fictitious *mean sun* that travels uniformly ($\varphi(t) = \omega t$), along the equator.

Universal time (UT) is defined as follows: 12:00 UT is the time of transit of the mean sun over the meridian at Greenwich (i.e., the time the position of the mean sun as seen from Greenwich is due south). The time of difference between the transit of the apparent and 12:00 UT is called the *equation of time* ET; the corresponding difference between the transit of the uniform sun and 12:00 UT is called $\mathrm{ET_u}$.

b) Calculate and plot ET adn $\mathrm{ET_u}$ as functions of time. Also plot the ET that would result if γ were zero. What are the physical effects reflected by these three quantities? Determine the time of transit for the apparent sun in Eugene, OR on the day that this assignment is due.
Hint: Transform from the equator system to the ecliptic system by means of rotation about the nodal line.

c) The sun's *declination* is defined as the angle δ between the position of the apparent sun and the equator. Provide a parametric plot of ET on the horizontal axis versus δ on the vertical axis with time as the parameter. The resulting curve is called the *analemma*. Give an interpretation of its meaning.

Note: For a cool related project initiated a few years ago by Prof. John Nicols, Dept. of History, and conducted by him with various other faculty and a group of students, see
http://solarium.uoregon.edu/,
and especially
http://solarium.uoregon.edu/analemma.html.

38. Steiner's theorem

Let I_{ij} be the inertia tensor with respect to the center of mass $\boldsymbol{X}$. Let Θ_{ij} be the inertia tensor with respect to a point $\boldsymbol{R} \neq \boldsymbol{X}$. Show that

$$\Theta_{ij} = I_{ij} + M\left(\boldsymbol{a}^2\delta_{ij} - a_i a_j\right), \tag{3.259}$$

where M is the total mass and $\boldsymbol{a} = \boldsymbol{R} - \boldsymbol{X}$.

39. Moments of inertia

Calculate the principal moments of inertia for the following systems of point masses.

a) N particles on a straight line (masses and distances are in general pairwise different). Express your result in terms of the masses and the distances between nearest neighbors. What is the result for $N = 2$?

b) 3 particles, 2 of which have equal masses, that form an equilateral triangle.

40. Heavy symmetric top

Note: This problem will be continued in Problem 42.
Consider a heavy symmetric top, i.e., a symmetric top in a homogeneous gravitational field, whose lowest point is fixed.

a) Starting with the Lagrangian in the principal axes system, derive the Lagrangian in the inertial system.
Hint: You can find the result in Landau & Lifshitz.

b) Find the three conserved quantities that arise from the fact that two of the three Euler angles are cyclic variables, and that the Lagrangian is time-independent.

c) Use the three conservation laws found in part (b) to show that the top's energy can be written as

$$E = \frac{1}{2}J_1\dot{\theta}^2 + U_{\text{eff}}(\theta), \tag{3.260}$$

where J_1 is constant. Determine the effective potential U_{eff}.

d) Express the three time-dependent Euler angles in terms of three one-dimensional integrals.

41. More moments of inertia

Calculate the principal moments of inertia for the following homogeneous rigid bodies:

a) Hollow sphere: $\rho(r) \neq 0$ for $R_1 \leq r \leq R_2$.

b) Hollow cylinder: $\rho(r, z) \neq 0$ for $-h/2 \leq z \leq h/2$, $R_1 \leq r \leq R_2$.

c) Cone: $\rho(r, z) \neq 0$ for $0 \leq z \leq h$, $r \leq z \tan \alpha$.

42. Heavy symmetric top (continued)

Consider the heavy symmetric top of Problem 40.

a) Use the implicit solution obtained in Problem 40 part (d) to give a qualitative discussion of the top's motion. Make sure that you distinguish among all of the different cases that lead to qualitatively different types of motion.

b) Suppose that you start the top with the figure axis in a vertical position. How fast must the top spin in order for this type of motion to be stable?

43. Rotating physical pendulum

A physical pendulum can swing about a horizontal axis which is rotated about a vertical axis with a fixed angular velocity Ω.

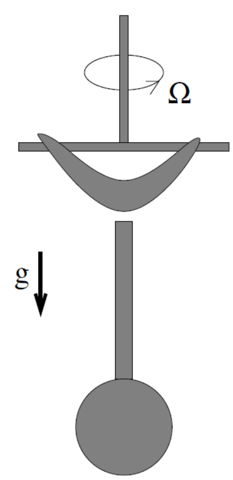

Figure 3.6

a) Determine the Lagrangian.

b) Discuss the equilibrium position as a function of Ω.

c) Discuss the frequency of the small oscillations of the pendulum close to its equilibrium position.